Smart and Secure Augmented Reality for Assisted Living

Eduardo Machado

A thesis submitted in partial fulfilment of the

requirements of De Montfort University

for the degree of Doctor of Philosophy

De Montfort University

December 2022

Copyright

Declaration

I declare that the work described in this thesis was originally carried out by me during the period of registration for the degree of Doctor of Philosophy at De Montfort University, U.K., from July 2016 to December 2022. It is submitted for the degree of Doctor of Philosophy at De Montfort University. Apart from the degree that this thesis is currently applying for, no other academic degree or award was applied for by me based on this work.

Acknowledgements

At the end of this important but also a long journey of my life, I would like to thank all the people who in some way contributed to the completion of this doctoral thesis, to whom I convey my sincere thanks.

First of all, a special thanks to my supervisor Feng Chen and co-supervisor Laleh Kasraian for your enthusiasm, motivation, dedication, guidance, patience, and all the support given during these years. It was a great pleasure working with you as a PhD student.

To ISOIN and Daniel Saldana for providing the conditions for conducting this thesis. A special thanks to Fernando Salvago, Jose Sosa, Alejandro Garcia and Fernando Guerrero for their unceasing availability, help and support. I would also like to thank all the professors and colleagues of the ACROSSING project.

A special thanks to Professor Liming Chen for believing in me and opening the world of scientific research to me.

On a more personal level, a deep and sincere thanks to my parents and brother for their unconditional motivation and support. Thank you for the love, patience and understanding during this journey. Finally, a very special thanks to Alejandra Acedo, she was the reason why I joined this adventure, without your love none of this would be possible.

Abstract

Augmented reality (AR) is one of the biggest technology trends which enables people to see the real-life surrounding environment with a layer of virtual information overlaid on it. Assistive devices use this match of information to help people better understand the environment and consequently be more efficient. Specially, AR has been extremely useful in the area of Ambient Assisted Living (AAL). AR-based AAL solutions are designed to support people in maintaining their autonomy and compensate for slight physical and mental restrictions by instructing them on everyday tasks.

The discovery of visual attention for assistive aims is a big challenge since in dynamic cluttered environments objects are constantly overlapped and partial object occlusion is also frequent. Current solutions use egocentric object recognition techniques. However, the lack of accuracy affects the system's ability to predict users' needs and consequently provide them with the proper support. Another issue is the manner that sensitive data is treated. This highly private information is crucial for improving the quality of healthcare services. However, current blockchain approaches are used only as a permission management system, while the data is still stored locally. As a result, there is a potential risk of security breaches. Privacy risk in the blockchain domain is also a concern. As major investigation tackles privacy issues based on off-chain approaches, there is a lack of effective solutions for providing on-chain data privacy. Finally, the Blockchain size has been shown to be a limiting factor even for chains that store simple transactional data, much less the massive blocks that would be required for storing medical imaging studies.

To tackle the aforementioned major issues, this research proposes a framework to provide a smarter and more secure AR-based solution for AAL. Firstly, a combination of head-worn eye-trackers cameras with egocentric video is designed to improve the accuracy of visual attention object recognition in free-living settings. A heuristic function is designed to generate a probability estimation of visual attention over objects within an egocentric video. Secondly, a novel methodology for the storage of large sensitive AR-based AAL data is introduced in a decentralized fashion. By leveraging the power of the IPFS (InterPlanetary File System) protocol to tackle the lack of storage issue in the Blockchain. Meanwhile, a blockchain solution on the Secret Network blockchain is developed to tackle the existent lack of privacy on smart contracts, which provides data privacy at both transactional and computational levels. In addition, is included a new off-chain solution encapsulates a governing body for permission management purposes to solve the problem of the lost or eventual theft of private keys.

Based on the research findings, that visual attention-object detection approach is applicable to cluttered environments which presents a transcend performance compared to the current methods. This study also produced an egocentric indoor dataset annotated with human fixation during natural exploration in a cluttered environment. Comparing to previous works, this dataset is more realistic because it was recorded in real settings with variations in terms of objects overlapping regions and object sizes. With respect to the novel decentralized storage methodology, results indicate that sensitive data can be stored and queried efficiently using the Secret Network blockchain. The proposed approach achieves both computational and transactional privacy with significantly less cost. Additionally, this approach mitigates the risk of permanent loss of access to the patient on-chain data records.

The proposed framework can be applied as an assistive technology in a wide range of sectors that requires AR-based solution with high-precision visual-attention object detection, efficient data access, high-integrity data storage and full data privacy and security.

List of Tables

List of Figures

List of Abbreviations

AAL	Ambient Assisted Living
AES	Advanced Encryption Standard
AI	Artificial Intelligence
AUI	Adaptive User Interface
ASD	Autism Spectrum Disorders
BPS	Brain Power System
CBOD	Centre Biases Object Detection
CID	Content Identifier
CLT	Cognitive Load Theory
CNN	Convolutional Neural Network
DApp	Decentralized Application
DLA	Daily Living Activities
ECDH	Elliptic-Curve Diffie-Hellman
EEG	Electro Encephalo Graphy
ECDSA	Elliptic Curve Digital Signature Algorithm
EHR	Electronic Health Records
EMR	Electronic Medical Record
FBOD	Fixation-based Object Detection
FDM	Fixation Density Maps
FHE	Fully Homomorphic Encryption
fMRI	functional Magnetic Resonance Imaging
ICT	Information and Communication Technology
HCI	Human-Computer Interaction
HE	Homomorphic Encryption
HR	Heart Rate
NN	neural network
NIZK	Non-Interactive Zero-Knowledge
PoS	Proof-of-Stake
RAM	Random-Access Memory
RA	Reentrancy Attack
RoI	Regions of Interest
SGX	Software Guard Extension
SSD	Single Shot Detector
SSS	Shamir Secret Sharing
TEE	Trust Execution Environment
Zk-SNARK	Zero-knowledge Succinct Non-interactive ARGument

Research Activities Completed

- **Conference Paper:** Machado, E.; Singh, D.; Cruciani, F.; Chen, L.; Hanke, S.; Salvago, F.; Kropf, J.; Holzinger, A. *A Conceptual framework for Adaptive User Interfaces for older adults*. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece, 19–23 March 2018; pp. 782–787.

- **Talk:** *Machine Learning Is Everywhere* at Cyber Security Department, De Montfort University August 2019.

- **Conference Paper:** Machado, E.; Carrillo, I.; Saldana, D.; Chen, F.; Chen, L. An *Assistive Augmented Reality-Based Smartglasses Solution forIndividuals with Autism Spectrum Disorder.* In Proceedings of the 2019 IEEE Intl Conf on Pervasive Intelligence and Computing, Fukuoka, Japan, 5–8 August 2019; pp. 245–249.

- **Conference Paper:** Machado, E.; Carrillo, D.; Chen, *Visual Attention-Based Object Detection in Cluttered Environments* The 16th IEEE Int'l Conf. on Ubiquitous Intelligence and Computing, Athens, Greece, 12-15 August 2018; pp 133–139

- **Conference Paper**: iappas, N.; Terius-Padron, J.G.; Machado, E.; Loghmani, M.R.; García-Betances, R.I.; Vincze, M.; Carrillo, I.; Cabrera-Umpierrez,M.F. *Best practices on personalization and adaptive interaction techniques in the scope of Smart Homes and Active AssistedLiving*. In Proceedings of the Workshop on Advanced Technologies for Smarter Assisted Living Solutions 2019, Leicester, UK, 19–23 August 2020.

- **Talk**: *Blockchain Future without privacy.* At Tecnologías Informáticas, Universidad Sevilla 2022.

- **Journal Paper:** Machado, E; Chen, F; Kasraian, L; *An Efficient Blockchain-based Data Storage Solution with Enhanced Security and Privacy* On Future Generation Computer Systems December 2022, vol 141 (Submited)

Table of Contents

Chapter 1

Introduction

1.1 Research Motivation

In recent years, augmented reality (AR) has been emerging as an assistive technology for children and adults with autism spectrum disorder (ASD). Evidence claims that AR technology is a key factor to improve their communication skills, emotional skills and daily living activities (DLA) skills. Despite the number of potential benefits, current interventions are mostly based on in-person occupational therapies that are both costly and difficult to access. These interventions explore the use of the Discrete Trial Teaching (DTT) approach [1] to address some of the deficits that characterize ASD such as executive function problems and (specific) learning difficulties [2]. This approach relies on an orderly and intensive instructional task that involves four components: (a) presentation of a discriminative stimulus (SD), (b) occurrence or approximation of the targeted response, (c) delivery of a reinforcing consequence, and (d) a specified intertrial interval [1]. Depending on the children's ASD degree, the task complexity can be increased or decreased. Nevertheless, this approach is very task repetitive and usually presents unappealing content, resulting in low levels of user engagement and frustration during task performance. Consequently, they spend most of their time off task and face difficulties in skill acquisition [3]. Moreover, evidence-based research shows that despite DTT being highly effective in single units of behaviour (e.g. practising a tennis serve over and over or colour recognition) it is not effective in sequential behaviour activities (e.g. making the bed, meal preparation etc.). Sequential behaviour activities require the display of several responses in a sequential chain. For example, washing dishes demand a set of actions beginning with location, retrieving, and preparing materials, followed by performing the act of washing the dish. Children with ASD typically have deficits in a wide array of skills that require the display of a chain of behaviours to learn a specific living skill. In detail, sequential behaviour activities involve two main components: identification of actions needed to perform a complex task and prompting instructional procedures during task performance.

Here, AR assistive technologies can be extremely useful in prompting instructional procedures as children with ASD have shown to be very receptive and engaged to AR technology. The benefits of AR technology among children with ASD cohort led to considerable growth in the development of AR assistive technologies on a variety of platforms. In particular, smartphones and tablet devices gained special interest due to their advantages in terms of portability, capability and ubiquity. These unique features empower the user to be assisted anytime and anywhere compared to traditional person-to-person interventions that are very limited in frequency and duration. Nevertheless, despite the promising results, evidence suggests that the head-down posture inherent to smartphone or tablet usage may decrease the user's awareness of their social and physical environment [3]. This is especially concerning due to the fact that autism people already suffer from issues in social communication and interaction. Moreover, AR-based smartphone increases the risk of injuries through distraction and the development of postural and grip strain [3]. In this regard, AR-based smartglasses present key advantages when compared to smartphones or tablet platforms. The usage of AR-based smartglasses

is found to be less distracting and less demanding in terms of cognitive workload. By looking through smartglasses, users can continue to look heads-up at the environment around them and remain hands-free because smartglasses are head-worn [4]. These advantages gain special importance in the context of assistive applications, allowing the user's hands to be unoccupied as well as the user's visual attention continuously remain on task. Yet, most of the existent AR applications based on smartglasses are not designed to be used in multiple situations, but mainly to assist the user in a specific scenario or activity (i.e. cooking training social communication, washing dishes, etc.). In that sense, there is a clear need for a portable technology readily accessible to self-instruct a person in a wide range of activities. This would alleviate the need to force children to use different assistive technologies for different activities.

In that sense, the goal of this work is to develop a smart and secure augmented reality framework for applications based on sequential behaviour activities. By smart, it means that system must be provided with sufficient intelligence to perceive the user's mental state and understand his context. Only by meeting these requirements is possible to prompt personalized assistance at a given right time. Currently, the investigation has shifted towards improving or creating new models of user context for situation-ware interaction. Here, the ability to recognize and predict objects of interest assumes to be critical for constructing an accurate model. In this regard, researchers have been using raw data from egocentric cameras and head-worn eye trackers to train artificial intelligence models in order to predict objects of interest. The use of egocentric cameras is becoming very popular as it overcomes the major limitation of common approaches which is the need for sensor-equipped environments (i.e. RFID attached to objects, or fixed cameras distributed over a place). Current approaches vary in the granularity of predictions, either focusing on predicting saliency regions in an image that are likely to attract user attention or discovering specific object instances based of on the prediction of user visual attention upon those objects. Nevertheless, despite the progress that has been made, it still remains a challenge to apply those models in daily living environments. In most cases, the existent solutions perform well because they are trained using data exclusively collected for task-specific and in a controlled environment. The main challenge is to extract visual attention in dynamic cluttered environments, where objects are constantly overlapped, and partial object occlusion is also frequent.

By secure, it refers to the capability of ensuring access and exchange of electronic health records (EHRs) between multiple entities of a system in a secure manner. Because patients move between medical facilities due to life events, EHRs are often maintained in a private local database, which results in data being dispersed across numerous medical institutions. Moreover, patients lose simple access to their medical records after diagnosis even if they own them.

To tackle these issues, researchers leverage cloud platforms to manage EHRs [5]. The use of cloud technology has, to some extent, enhanced the share of EHRs. On the other hand, there are potential security issues with this cloud-based approach. Given that cloud providers are part of centralized domains, the difficulty of cloud-based data sharing lies in the degree of trust that users place in cloud service providers. Such a lack of trust is resultant form constants lack of transparency, and the security data breaches [6]. Accidents involving the leakage of medical records is an example of what have been happening in recent year [7][8].

Blockchain, which is widely leveraged in cryptocurrency systems [9] is a promising technology that has been used recently in a variety of sectors, with the financial sector being the most prominent. Blockchain is a digital ledger that stores all activity performed

with the network. It creates blocks that store collections of the transactions among the participants using a distributed peer-to-peer network.

Each of those blocks includes a collection of signed transactions, and using a consensus algorithm, the network itself verifies the validity of each block. Each network node that is actively participating stores a copy of the blockchain state. The blockchain is an immutable ledger thanks to its cryptographic algorithms that prevent the data stored to be tampered. Blockchain technology and eHealth systems have been successfully merged in several countries. For instance, Estonia[10] adopted blockchain technology as an additional layer of security over their patient's medical records. Several approaches, i.e. [8][10] aim to improve the operability and confidentiality of eHealth systems by utilizing cutting-edge technology like smart contracts and privacy protection modules. Based on the literature review, using blockchain to act as a decentralized storage platform offers many security advantages when compared to centralized local/cloud storage servers, still, there are certain obstacles to be solved.

Firstly, the capacity for blockchain data management systems to ensure data privacy at both transactional and computational privacy [11–13]. Most of the existent solutions rely on off-chain modules that make use of encryption schemas to encrypt the data before storing it in the blockchain. However, these types of techniques result in high computational complexity and latency which makes them hard to bring to real-world application scenarios. Other solutions have solved the problem of transactional privacy whereas providing computational privacy is still an open research challenge.

Secondly, blockchain provides independent storage of data over multiple nodes of the network in the form of a distributed ledger. However, due to the very limited storage space available in each node of the network, storing all kinds of existent data types would be extremely expensive and consequently unfeasible for healthcare organizations. As result, current approaches use blockchain for authentication purposes and data indexing whereas the data is stored locally or in cloud serves, making it prone to be undermined by malicious actors. Finally, existent literature does not pay relevant attention to access control methodologies for preventing the cases of lost or theft of private keys.

This research presents a framework that aims to increase the intelligence, security and data interoperability of AR-based assistive technologies. It includes the investigation of models capable of extracting objects of interest through the user's visual attention in real-world settings. In addition, this research also introduces a decentralized data storage system that protects user data privacy and promotes secure data interoperability among multiple entities with low cost and high computation efficiency.

1.2 Problem Statement

Build an smart AR-based assistive solution that can achieve high performance under complex environments is a challenge that requires investigation in a multidisciplinary field of study. In order to design and conceiving such a system, this study focus on addressing five main challenges: (i) the design of a device that offers a natural and easy user interface, (ii) the development of a computational framework that offers high scalability and high performance in real-time, (iii) discovery of biological features to assess ongoing user cognitive state (iv) the creation of algorithms to perceive both user intention and his world surroundings, (v) a data sharing solution that ensures the interoperability, immutability and privacy of the data. These challenges are going to be exposed in detail in the following paragraphs:

1.2.1 Design of a Device Based on Natural Interaction with the User

Independently on the aim of any assistive technology, its adoption is always dependent on its capability of providing a natural interaction experience to the user. In this regard, the design of a natural interaction interface is focused on enhancing system usability by making it more self-explanatory, easy to use, entertaining and non-intrusive [14]. Only by satisfying these premises is possible to activate a positive interaction between the user and product interface at the time of performing a specific task.

Immersive technologies like AR introduced a new interaction paradigm. Its natural user interface is proven to reduce the user's cognitive load over a specific task (by overlaying digital content onto the real world) and consequently preventing failures. However, most of the existing AR-based solutions are built for tablets and smartphones. And despite its advantages in terms of portability, capability, and ubiquity, the inherent head-down posture decreases the user's awareness of the surrounding environment. This highly increases the chances of accidents or errors. Moreover, AR-based smartphone increases the risk of injuries through distraction and the development of postural and grip strain [3]. In this respect, there is a clear need to solve issues related to head-down posture and distraction without compromising the existing advantages.

1.2.2 Understanding the Human Cognition

In order to better assist the user, the understanding of his environment surroundings becomes very relevant. However, the sole interpretation of the user environment is not enough to effectively achieve good performance. It is also requires information over the ongoing user cognitive state. Only by have these two sources of knowledge, is possible for a system to have a correct understanding of the user needs. In that sense, the research community has been investigating on new forms of measuring human cognitive load. This is highly valuable resource for assistive applications.

However, due to its nature, conventional forms of assessing human cognitive load like subjective measures or dual-task measures are not feasible to be integrated into real-time assistive systems. In that sense, the study of human physiological measures as a way of estimating different levels of cognitive effort is taking the lead in this research area. Specifically, eye function measures like pupil dilation or fixations have been highly studied due to their advantages of accuracy and unobtrusiveness to the user. Nevertheless and despite the promising results in extracting the user cognitive state, most of the studies were conducted in controlled settings. Only a few studies attempted to replicate these results in free-living conditions because these biological responses are highly affected by external factors. The problem is finding a type of user cognitive assessment technique that can be applied in free-living settings, perform in real-time and still achieving high accuracy.

1.2.3 The problem of Object Recognition-Cluttered Environments

The ability to recognize an object of interest assumes to be critical for the construction of an accurate model of situation-aware interaction. In this regard, egocentric cameras

and head-worn eye tracker are becoming very popular as they overcome the major limitation of common approaches, the need for a sensor-equipped environment.

Classic approaches use egocentric cameras and start from the assumption that user visual attention is allocated at the centre of the visual field. Then, this pre-defined spatial pixels region of the egocentric images is fed as features input to object recognition algorithms in order to predict the object of interest. However, this approach arise issues related to the variability of the human natural gaze exploration. It is shown that human attention can be allocated in peripherical regions distanced from the spatial centre image. This loss of information can have a significant impact on the accuracy of the model. In an attempt to overcome this issue, researchers have been focused on the use of eye-tracking cameras to combine fixation data with object detection algorithms. Despite that progress has been made in controlled experiment environments, it remains a challenge to apply egocentric visual attention object recognition algorithms in normal daily living environments. In most cases, the existent solutions perform well because they are trained using data exclusively collected for task-specific and controlled environments. The main issue is to extract visual attention in dynamic cluttered environments, where objects are constantly overlapped, and partial object occlusion is also frequent.

1.2.4 The problem of Preserving the Data Security and Privacy.

Independently of the use case of assistive technology, might it be industrial or healthcare etc., the manner in which sensitive data is treated is highly relevant. Assistive devices produce EHRs that need to preserved from a data security and privacy perspective. Data security refers to ensuring that sensitive information is protected against eavesdroppers or attackers. On other hand, data privacy means that data belonging to a specific user or entity is processed privately or a certain authorization needs to be conceded to access the data. The lack of security or privacy can potentially leave valuable information exposed, which can negatively generate a lack of trust and a drop in the reputation of an organization. In that sense, the problem lies in building a robust solution that maintains sensitive information in a highly controlled and protected environment while ensuring its privacy.

1.2.5 Design of a Scalable Framework to Handle a Variety of Problems

The recent number of AR assistive solutions based on smart glasses has been growing at a considerable pace. However, most of the existent solutions are not designed to be task free, but mainly to assist the user in a specific scenario or activity (i.e. cooking, training in social communication, washing dishes, etc.). In that sense, there is a clear need for portable technology readily accessible to assist a person in a wide range of activities. This would alleviate the need for different assistive technologies for different purposes. This study is focused on building a framework that fit the requirements of a such “one-size-fits-all” assistive solution. This is quite challenging because it demands the integration of a multitude of tools. Those include raw sensor data gathering, data integration protocols, AI-based knowledge discovery, data security and privacy, and both task and user action modelling. Moreover, the performance of the framework must be reliable on low computational resources and feasible to operate in real-time.

1.3 Research Hypothesis

Above is exposed the main challenges to building an AR assistive solution to apply in a variety of scenarios and not only to a specific task. The presented hypotheses that have been raised in this work are going to provide a roadmap towards the solution for the aforementioned research issues.

Considering the first problem, raises the question if an AR technology-based smartphone or table is an ideal assistive solution considering the requirements of a natural user interface. In an attempt to solve this question, this work hypothesizes that AR-based smartglasses offer a better natural interaction considering their very nature, are less intrusive, and are easy to use. Along with this, by integrating eye tracker cameras as well as egocentric cameras as part of the smartglasses, it will be possible to not only understand the user's surrounding environment but also his intentions in real-time. In that sense, this study hypothesises that information from eye trackers such as pupil dilation, fixation periods, and gaze points are suitable for extracting measures of the ongoing user cognitive state. This assumption will be explored to solve the second-mentioned problem.

The recognition of objects of interest is considered an important topic in the field of computer vision with high applicability in assistive applications. However, few investigations have been made to tackle the issue of the low accuracy of egocentric video-based activity recognition algorithms when applied in cluttered environments. In this thesis, it is hypothesized that data from eye movements in combination with ego-motion patterns and visual features observed through the camera can provide an extra layer of information and significantly improve the accuracy of classification.

To solve the problem of lack of privacy and security of AR-based assistive solutions, the author of this thesis hypothesises that blockchain technology alone improves but does not entirely solve the problems of data security and privacy. This is supported by the fact that public blockchains like Ethereum do not offer privacy features by default. Along with that, concerns related to the low storage capacity of such networks limit their applicability. In that sense, the hypothesis raised is that blockchain must be combined with decentralized storage protocols along with off-chain solutions to ensure the high security and privacy standards of assistive technologies.

Finally, to address the challenge of scalability this investigation hypothesises that the architecture foundations of this framework must rely on a modular approach, similar to microservices approaches. This, combined with an event-driven protocol for the communication between the core components of the framework, will foster performance, responsiveness, and the extensibility of the functionalities in the application domain.

1.4 Research Objectives and Questions

The main aim of this study is to develop an AR framework for smart and secure applications in the area of assistive living. To achieve the presented goal, this study formulated a set of research questions that were converted into the research objectives of this investigation. In this respect, the research questions and the related research objectives are the followings:

a) **RQ:** What are the requirements of smart assistive technology to provide a secure, private and easy-to-use interface?

RO: To investigate different interaction modalities for providing natural and easy-to-use assistive technology tailored to the needs of individuals with ASD. In addition, investigate decentralized protocols for security and privacy purposes as well as machine learning algorithms to improve the intelligence whole system.

b) **RQ:** Are the existent frameworks for assistive AR technology flexible enough to cover a wide range of different scenarios and applications?
RO: Development of an AR-based assistive framework designed to fit the requirements of real-time applications on multiple use case scenarios.

c) **RQ:** Can the existent eye-functions measures provide an assessment of the user's visual attention in real time? And can those be used for the discovery object of interest?
RO: To investigate biological function metrics that coupled with artificial intelligence algorithms can enable the discovery of objects of interest.

d) **RQ:** Is it possible to build a smart decentralized AR framework capable of providing high data security and privacy?
RO: Development of a decentralized storage system that encapsulates cryptographic mechanisms to enable the management of EHRs in a secure and private fashion.
RO: Investigate a solution to enable the user to still have access to his EHRs in case of lost or stolen private keys.

e) **RQ:** Can a decentralized storage system be more affordable and efficient than a traditional cloud or local-based system?
RO: Development of a smart contract that achieves low transactional cost and data while preserving privacy features at both the transactional and computational levels.

1.5 Research Methodology

Research methodology refers to a comprehensive set of organizing principles around which empirical data is gathered and analysed [15]. In this context, a variety of methodologies can be applied to solve a research problem, often to fully understand the nature of the problem it is a common practice to combine different methodologies. Typically, the two different methodologies for conducting research are quantitative and qualitative. This research work adopted a quantitative methodology coupled with a qualitative methodology. The methodology design chosen aims to ensure that this particular research is reliable by testing the hypotheses and not extraneous variables.

1.5.1 Qualitative Research Methodology

A variety of methodologies and methods are employed by researchers in science. A particular methodology is defined as qualitative research. This scientific methodology

inserts into the realm of observation, in contrast to the quantitative methodology which is typically rooted in numbers and statistics [16]. Qualitative methodology is based on interpretivism, it exploits non-numerical data to understand the meaning of concepts, definitions and features of things [17]. The analyses of the data gathered will enable the elaboration of a theoretical framework giving the possibility to explain the finding results in a coherent manner instead of directly accepting or refusing an a-priori-defined hypothesis.

In this study, the qualitative research methodology will be used to understand which is a suitable user interface for AR-based assistive technologies that fits the requirements of people with different cognitive degrees. In addition, qualitative methods will be applied to find evidence of which biological metric can be used to extract users' visual attention, considering the requirements of a real-time application and the complexity of daily living environments.

1.5.2 Quantitative Research Methodology

Quantitative research methodology can be defined as a measurement of quantity or amount. It is concerned with quantifiable data, objective measures and analysis variables in order to obtain results [16]. Specifically, it is rooted in the collection of numerical data and the application of statistical techniques to support or refute alternative knowledge claims [18]. In that sense, the major difference between qualitative and quantitative research methodology is concerned with how the data are collected and analyzed. In [19] is remarked that quantitative research requires the reduction of phenomena to numerical values in order to carry out the statistical analysis.

By using quantitative research methods, this study aimed to evaluate the effectiveness of the proposed framework by calculating its operational cost, the computational response time of the different modules involved and the accuracy of the object detection algorithm.

1.6 Success Criteria

In this section, it is presented the success criteria for this project:

a) The User Interface was natural and easy to use.

b) The decentralized storage solution is cheaper than traditional cloud-based solutions.

c) The assistive AR-based smart glasses framework meets the requirements for operating in real-time.

d) The model to extract user visual attention and object of interest is effective in cluttered environments.

e) The developed smart contract can offer both transactional and computational privacy.

f) The user can still have access to their data in the event of loss or stolen private key.

1.7 Research Ethics

In this study, a set of ethical considerations were taken to the design and practices of this research. Firstly, all the participants involved in the investigation were free to choose to participate without any pressure or coercion. The participants were also informed that they were free to quit and leave the study at any given moment. Anonymity was a guarantee to the participants by not collecting any personally identifying information. Finally, this research did not engage in any type of misconduct such as falsification, data analysis manipulation or any form of plagiarism.

1.8 Research Contributions

The findings of this study make several contributions to the current research literature on AR-based smartglasses for assistive living. The majority of previous research in AR for assistive living has only studied solutions to solve a specific use case. The researcher believes that this is the first study to create a solution seeking to promote the scalability of such frameworks by introducing a new architecture, novel approaches for data security and privacy as well as methodologies to increase the overall intelligence of such systems. In that sense, the outcome of this work provides research contributions to the existent body of knowledge :

a) **Smart and secure real-time assistive framework.** A new solution for real-time assistive applications on AR-based smartglasses. It enables multiple-task modelling to improve user skills in DLA by delivering assistive cues and instructions based on user behaviour while preserving the security and privacy of user data.

b) **Visual attention-based object detection**. A novel heuristic function that is able to generate a probability estimation of visual attention over objects within an egocentric video.

c) **Egocentric video dataset.** An egocentric indoor video dataset annotated with human fixation during natural exploration in a cluttered environment. This dataset targets general multiple-task conditions. It becomes a more realistic dataset as it was recorded in real settings with variations in terms of objects overlapping regions and object sizes.

d) **Decentralized EHRs storage system.** An approach to attribute access control and decentralized storage was proposed using three novel technologies, smart contracts, IPSF and SSS. This approach offers secure storage with a fast encryption schema to promote data privacy.

e) **Privacy-preserving smart contracts for the management of EHRs.** A novel smart contract that guarantees on-chain data privacy mitigating the inherent risk that off-chain privacy solutions take by having to manage the custodian of user cryptographic keys.

f) **Governing Body for accessing control.** A new approach based on an off-chain governing body for permission authentication entities, providing systems with a secure approach to deal with events of loss or stolen of private keys.

1.9 Structure of Work

The description of the research and developments presented in this thesis is organised as illustrated in Figure 0-1. The thesis comprises the following chapters and references:

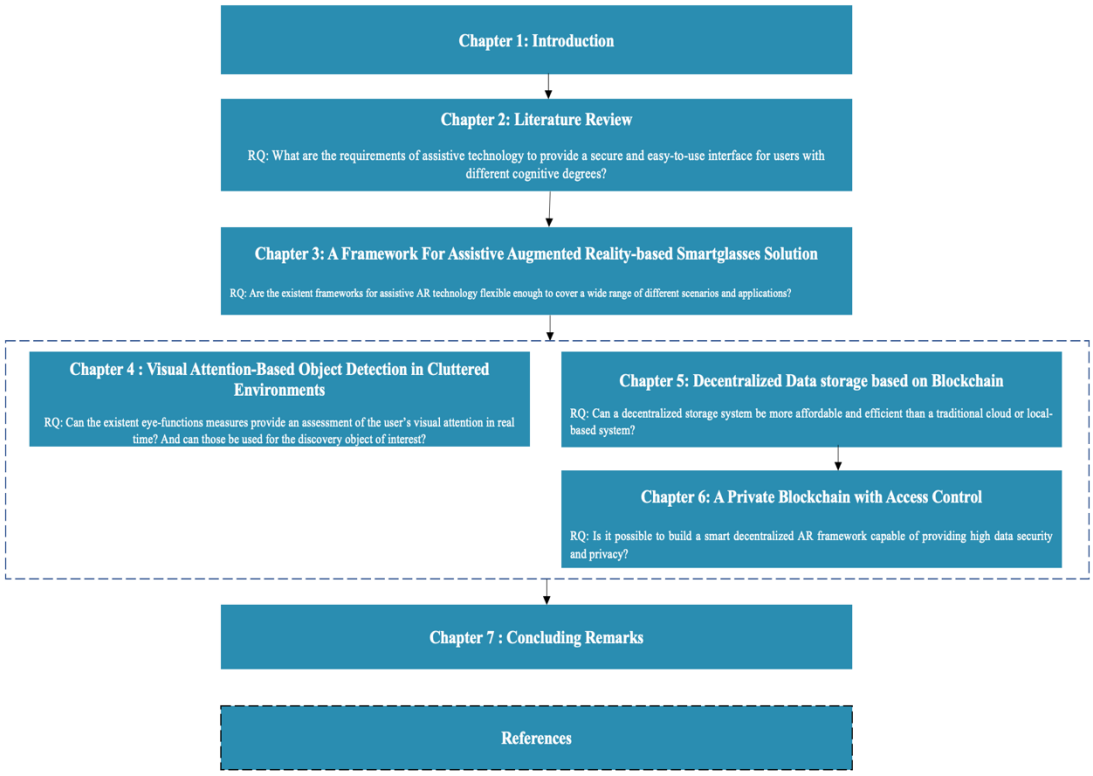

Figure 0-1-Schematic representation of the structure of this thesis and Research Questions (RQ) answered in each chapter.

Chapter 1 corresponds to this introduction.

Chapter 2 provides the foundations required to understand the state of the art of frameworks for assistive living technologies focused on individuals with ASD. This is accompanied by a literature review of the existent models for object detection, along with

eye function metrics used for visual attention. Within this chapter, a literature review on decentralized approaches for data security and privacy is also documented.

Chapter 3 describes the development of a framework for AR assistive living applications. The performance of the framework is also provided within the course of this chapter.

Chapter 4 applies a visual attention-based object detection model. An experiment is conducted, and the following are evaluated. A dataset is also presented.

Chapter 5 introduces a decentralized storage system aiming to improve the security and privacy of the EHR shared by different entities. Within this chapter, is possible to observe the results from this methodology in terms of cost and overall performance.

Chapter 6 evaluates the new approach for enabling both transactional and computational privacy within the network. In addition, it introduced an off-chain solution designed to mitigate the issues related to the loss or stolen of a user's private key.

Chapter 7 summarizes the main conclusions of the author's work in the preceding chapters and describes possible avenues for further research.

Chapter 2

Literature Review

At the present time, assistive living technologies are considered a critical component for the therapy of individuals with ASDs and cognitive issues in general. In this context, Augmented Reality (AR) has become very popular due to its benefits in terms of engagement and its easy-to-use and natural user interface. This chapter will explore the realm of AR-based assistive living applications as described in the introductory chapter. It will discuss the state of the art of existing frameworks for building AR applications. This includes the analysis of methodologies to improve their scalability, intelligence, security and privacy. In this respect, a comprehensive explanation of the current models for object detection and visual attention is given. Finally, it reviewed the existent literature on decentralized solutions for the protection and privacy of EHR data.

2.1 Augmented Reality for Assisted Living

A series of neurodevelopmental disabilities known as autism spectrum disorders (ASDs) affect children's ability to communicate and connect with others in social situations. ASDs is also characterized by the children's limited and repetitive patterns of behaviour or interest. Acquiring daily living skills can be also hard for children and adults with ASD. The complexity of the disorder as well as the heavy raising burden on families turned therapists into major players in diminishing the symptoms provoked by ASD. Recent advances in information and communication (ICTs) technologies have enabled therapists to be provided with better tools to assist individuals during interventions. Multimedia, collaborative interactive environments, virtual reality, avatars, and robotics are a few examples of these assistive technologies.

Not only for therapeutic interventions, assistive technologies can also be found vital for family members of people with special needs by reducing their daily workload [20]. In overall, assistive technologies show to improve both families' and patients' quality of life as well as to improve therapeutic programs in terms of cost-effectiveness.

Recent research has leveraged technological advancements to develop AR-based solutions for real-world social interactions and living skills with the aid of sensing, inference, and delivery of in situ social cues via multimodal feedback. In [2] was developed an AR-based video-modelling storybook as a training tool to improve children's social skills. Augmented virtual visual hints were used to indicate and amplify the nonverbal social cues in videos. In this way, AR attracted the attention of the children to nonverbal social cues and helped them to understand the facial expressions and emotions of others. In the same direction, in [4] is proposed an AR game book for the tablet to promote children with ASD to recognize and acquire emotions by engaging their attention and motivation. The AR application enabled the creation of more attractive and interactive interfaces that could be manipulated by hand, without using conventional peripherals such as the keyboard or the mouse. The authors in [9] developed a mobile AR application to teach science vocabulary words to college students with ASD. The results indicated that all the students involved in the experiment acquired definition and labelling knowledge for new science vocabulary terms. Another system has been proposed to improve the lack of imagination in children with ASD [4]. The system is an AR-based computer application that elicits pretend-to-play. The computer works as a "mirror" that

augments different types of animations over the toys that the kid is playing with. In that sense, it allows the user to interact with the system without wearing or holding the display equipment and manipulate physical toys with both hands. The results showed a significant improvement in pretend play in terms of frequency, duration and relevance using the AR system in comparison to a non-computer-assisted situation. Nevertheless, despite the positive results, this system lack in terms of portability, the user is very limited in the area where he can play. In this regard, the implementation of such a system in an AR-based SmartGlasses platform would benefit from its advantages in terms of portability and accessibility. Following the premise, in [9] was developed Brain Power System (BPS) to enhance social and cognitive skills in children and adults with ASD. BPS is a combination of hardware and software add-ons that can be integrated into Smartglasses devices. It includes sensors to extract real-time data such as user's movement and physiology as well as video and audio. Thanks to a variety of AR-based smartglasses game applications, individuals with ASD are coached to understand emotions and increase eye contact with persons. In detail, these applications detect human faces over the user's central field of view and overlaid them with AR cartoons to attract the user's attention. The user is then rewarded with "points" if continuously looks at the cartoon, meanwhile, the cartoon gradually fades revealing the human face. However, the author starts from the assumption that the user's visual attention is allocated to the centre region of the smart glasses' camera image frame. This approach would fail in situations where the cartoon is located in the centre region, but the user's visual attention is distanced from it. For instance, the centre of the smartglasses visual field can match a human face but the child is focusing his attention on an object nearby. Other approaches include the use of video-based instructions, commonly known as video modelling or video prompting to teach daily living skills to individuals with ASD [1], [10], [12]. Here, the learner watches a video clip of a model completing a target task in its entirety at the onset of each instructional session. A recent study examined existent literature to analyse the effectiveness of using video-based instruction procedures while teaching DLA (e.g. cooking, cleaning, using a screwdriver, sorting recycling items), and the results show evidence that VM positively influences skill acquisition [6]. Nevertheless, as reported in [2] normal video-based instructions are too long and too difficult for children with ASD to handle. This argument is supported by the lack of visual-guidelines feedback and the absence of an interactive mechanism to help perceive and feel engaged with the videos. As result, they feel bored and helpless, tending to stop watching the video. Those issues prevent video-based instructions from providing effective aid.

This work pursues to overcome the issues related to the video-based instructions approach by taking advantage of the capabilities of AR-based smart glasses technology. To accomplish this goal, it is necessary to develop a framework for assistive applications that offers real-time interactive AR visual content to help children with ASD toward task completion.

There is a wide range of works in a variety of fields that proposes theoretical frameworks for the adoption of AR ranging from education, healthcare, industry, gaming, tourism etc. In [21] was introduced a framework that can provide both optical and video see-through-based augmentations and it features a robust optical tracking algorithm. The framework runs on a single computing platform and facilitates the deployment of AR-based applications on specialized headsets for image-guided and surgical intervention simulation. Another widely used software framework designed specifically for creating projection-based AR applications in the surgery rooms was introduced by [22]. The core of the multi-layer architecture is based on a communication module implemented using Google Protocol Buffers for exchanging messages among peers through a transport layer.

In [23], a framework that makes use of AR targeting to enhance tourism experiences. The framework combines AR with an electroencephalogram to measure the visitor's emotions and consequently optimize the AR content displayed during a tour. The authors in [24], developed a framework for motivational AR applications to improve students' motivation in learning settings. Following the same path, in [25] is described a framework for AR to enhance the knowledge, attention and practical skills of students. Especially, this framework provides students with the possibility to interact with virtual objects that describes the basic principle of a DC motor and generator. In the industry field, the authors in [26] developed a framework for maintenance systems. The framework offers different information types and ways users may engage with AR to increase productivity in the maintenance of sophisticated machinery. Essentially, it offers a generic implementation of AR in maintenance operations including diagnostics, repair, and analysis from a data management standpoint.

Despite most of the current literature presenting noble results in terms of engagement, performance and benefit to the achievement of a task, still, significant research gaps have been found in the literature:

a) The contextualization of the information in these frameworks must be defined. It should be based on information gathered from the users, the situation, the surroundings, and their relationships with one another.

b) Previous research in AR has barely paid attention to the challenges that arise in terms of security and privacy. The existing AR frameworks should leverage the standard security best practices such as on-device and network encryption.

c) The need for effective storage, processing, and sharing of the enormous real-time data gathered from various health monitoring systems.

d) Addressing the complexity of data processing and transfer while satisfying security and privacy requirements.

e) Enabling remote access to EHRs by authorized parties in order to attain widespread, affordable healthcare and tailored treatment.

Two major challenges arise from the gaps identified above. Firstly, developing AR systems with sufficient intelligence for understanding human behaviour which critical to provide effective assistance. Given the complexity, unpredictable nature, and diversity of each user, AR-based assistive interventions must be intelligent enough to react to untrained human feedback. Only by following this premise, is possible for AR systems continuously and rapidly adapt to the user’s unique needs by generating personalized assistance via virtual agents. In this regard, this work is focused on human visual attention over objects of interest. This provides AR systems with capabilities to not only understand the user environment but to record the interaction in real-time between the user and objects under his attention. Understanding this interaction enables AR systems to adjust their assistance more naturally and effectively. However, capturing human visual attention upon objects can be hard. In literature, common approaches include object detection models in combination with visual sight estimations to discover objects of interest [27][28][29]. Nevertheless, the data used to train those models, clear and well

disposed of, is not similar to the data found in free-living settings, where objects are repetitively overlapped, and it is also frequent to occur partial object occlusions.

The second challenge is related to the capabilities of assistive frameworks in ensuring the right security and privacy standards. The lack of privacy prevents the general population from adopting AR assistive technologies because of fear of privacy and confidentiality breaches. The author in [20] also argues that ASD individuals and parents are reluctant to use assistive technologies because of fear of privacy and confidentiality breaches. It affects the perceived trustworthiness of the technology and creates a psychological barrier to the risk of embracing it.

Typically, in AR frameworks the data flows and is stored in cloud environments that are prone to be attacked by malicious actors. Current research is investigating Blockchain as a promising technology for mitigating privacy and security breaches [30][31][32]. However, current literature still pays poor attention to addressing blockchain privacy issues at the computational level. It is hard to address this issue and existing off-chain solutions rise other issues regarding the degree of decentralization of the whole system. Access management and related issues are also open topics of research. Finally, storing data on the blockchain is very expensive which makes hard the adoption of technology when compared to traditional cloud-based approaches.

2.2 Eye Movements as an Indicator of User's Cognitive Load and Visual Attention

Cognitive load (or other similar terms such as mental effort or mental workload) plays an important role in the field of human-computer interaction (HCI). Researchers are analysing users' interaction with machines and technological devices based on the cognitive load associated with it. In the field of human-web interactive systems, the information can be displayed in different formats such as audio, video, text or multiple forms of multimedia, requiring different amounts of mental resources to process the information which sometimes can exceed the limit of the user‘s capacity. Poor interface designing can potentiate the overload of mental resources leading to a decline in task performance in task performance, frustration, and an increment in the chances of errors and mistakes. A lot of research has been done on cognitive load in the realms of education and instructional design in order to predict the cognitive burden implied by various learning materials or interface designs. It is possible to determine whether material or design will be more effective by contrasting the cognitive load indexes.

The idea of cognitive load is not new. Over the last 40 years, researchers have been studying a similar psychological term called "mental load" or "mental workload". The mental workload was used to refer to the quantity of cognitive work required for a person or group of individuals to do a specific task over time. However, this general definition of the concept has evolved over time becoming more complex and detailed. Currently, it does not exist in literature as a general or commonly accepted definition of mental workload. Some authors extended the definition by making reference to the limitations of human information processing [33][34], defining mental workload as "the difference between the capacities of the information processing system that are required for task performance to satisfy performance expectations and the capacity available at any given time" [35]. On the other hand, in [36] suggested that mental workload is "not an inherent property, but rather it emerges from the interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours and perceptions of the operator". This definition not only makes emphasises the limited processing

capacity of a person but also highlights his individual characteristics such as skills, behaviours and perceptions as well as the context of the task like its difficulty and the environment in which the user is performing it. In sum, the mental workload can be seen as a multifaceted and multidimensional nature as the execution of a specific task and its performance is influenced by the following factors [37]:

a) Exogenous factors- are related to factors inherent to the tasks, such as the task difficulty, environment complexity and uncertainty.

b) Endogenous factors- are concerned with the individual characteristics that affect the task performance such as the processing resources, skills, background and expertise.

c) External factors- context and situations where the task is performed.

Mental workload has been relevant to multiple distinct fields such as medicine and health care, automation and manufacturing system, critical environments such as military and nuclear power plants, and HCI and web-based environments [38]. Generally, the main goal of the study of mental workload in such a variety of fields is the improvement of the design of interfaces and systems by analysing and evaluating the mental workload of a user interacting with a system. As result, the user's mental burden is reduced leading to an increment in terms of task performance.

The development of Cognitive load Theory (CLT) by [39] was clearly influenced by research findings in the context of mental workload. The first work to assess cognitive load during the process of instruction was investigated with metal workload-based measures [40]. This work was inspired by [41] which used subjective measures to perceive items' difficulty in tests of intellectual performance capacity. Both mental workload and cognitive load are similar in terms of trying to assess the demands inherent to a task accomplishment, however, cognitive load differentiates itself by not taking into consideration an individual’s physiological aspects like his beliefs, goals and efforts on his load perceptions. This view is discussed by [42] [43] as one of the limitations of cognitive load theory.

CLT was originally developed based on the model of working memory where short-term memory is responsible for recall and information processing that can be stored in long-term memory in form of schemas[44]. The theoretical foundation of CLT is that working memory resources are limited. The human brain has several cognitive resources to process and reason over the information coming from different modalities of the human sensory system (auditory, visual, verbal, tactile etc.). Each modality has a limited number of resources dealing with it. This means that if during the execution of a task is required more resources than those available at a given time can occur an excess of cognitive load. Resulting in errors or decreasement in task performance.

The existing literature is mainly focused on the CLT assessment to increase the student’s learning capability in the context of education. Despite that, this project considered that research findings in the assessment of cognitive load can be very useful for other fields of study namely the HCI to help older people to overcome the difficulties in web interaction. Though that web interaction could be seen as an activity much more casual than formal education, it is still very important to ensure that web interfaces are designed with less extraneous cognitive load. This could foster the learning abilities of older people in browsing a new website for instance in the use of a navigation section,

buttons, and formularies. Even if the website is familiar to the user, users must spend the effort to filter the information that is relevant to their goal. In that sense, CLT can be an important factor for designers to understand and accommodate the limitations of the user.

The notion of cognitive load refers to the load that the execution of a task carries out on the user's cognitive system. According to [39], cognitive load can be distinguished into three classes:

a) Intrinsic cognitive load- refers to "the effort of absorbing new information. It is directly related to the complexity or nature of material content or task. It is defined as the number of elements that have to be processed and the relationships among them. An element means "anything that needs to be or has to be learned such as concept and procedure" [39].

b) Extraneous cognitive Load- refers to" the mental resources that are devoted to elements that are not related to learning or the schema acquisition. In other words, is mental resources required to process the manner in which the information is presented but doesn't help users learn or understand the content" [39].

c) Germane cognitive load- refers to the cognitive resources bounded for the construction and storage of schemata into long-term memory [39].

In the respect to the model of CLT still, there is much controversy around the community. Some authors disagree about the concepts and relationship between the three classes of CLT. In [45] the author defends that learning can occur only by ICL, while GCL is a way to enhance learning for instance with the conscious definition of a learning strategy. The same idea is supported by presenting evidence that GCL is responsible for setting strategies to enhance learning instead of schemata acquisition. Another view that goes a step further is shared by [46] arguing that ICL and GCL are concepts purely based on theoretical constructs as there wasn't discovered any empirical evidence of their existence. Therefore, the author refers to these concepts as redundant. An explanation for the contrast of ideas of CLT models is provided by [47] the author claims that such controversy is justified by the post-hoc problem. Due to the lack of accurate metrics for each load, cognitive load is measured in overall. This means that there is always the possibility to justify variabilities in the overall cognitive load by a cause that corroborates initial assumptions. For instance, if the overall of cognitive load maintains constant baseline but is possible to observe a decline in task performance, this effect can be attributed to the high extraneous load. On the other hand, in the case of the overall is kept constant but the task performance is increased, this may mean that GCL was successful in terms of schemata acquisition. In this regard, the present study will be focused on the measurements of the cognitive load as overall due to the lack of reliable measurement methods for each load. In the next section, it will be given an explanation of the existing measurement methods as well as their advantages and limitations.

2.2.1 Measurement Methods

Over the last decades has been applied different methods for the evaluation of cognitive load. The present section will be made a review of the existent methods and in which context they may be applied. There are three methods used to assess cognitive load:

2.2.1.1 Subjective measures

It refers to the measurement of cognitive load by forcing the user to analyse their own cognitive process and the level of mental effort necessary to complete the task in hand. Subjective measures are very attractive in the area of HCI because it is strongly believed that the user is the "right" person to give the most truthful judgment about the amount of mental effort dispensed during an activity. Moreover, subjective measures can be easily administered and applied in distinct fields. Accordingly to [47], subjective measurements can be sorted into two scales: multi-dimension scale and unidimensional scale. The unidimensional scale refers to the measurement of cognitive load as overall. It is very effective when one single item difficulty is rated. Multidimensional scale refers to the measurement of various dimensions of cognitive load. For example, the multidimensional scale provided by NASA Task Load Index to assess cognitive load includes six dimensions such as performance, mental effort, frustration level, task demand, physical demand and temporal demand. Later, the authors [48] [49] [50] adapted NASA's multidimensional scale to assess each of the three cognitive loads in the context of CLT. Another method for assessing cognitive load is provided by [51] requesting subjects to rate their experience with the learning content, complexity of material and concentration during learning. Those dimensions are respectively associated with the measurement of ICL, ECL and GCL. Nevertheless, in this case, the author concluded that it was not possible to differentiate between ICL and GCL.

Subjective measures can be also classified according to other variables such as evaluation style and immediacy. Evaluation styles are concerned with the fact that the assessment of cognitive load can be obtained by absolute rating or relative rating depending on if the rating of a current experiment is compared to others or not. In terms of immediacy, it is related to the temporal space in that the subjective rating is given. In this case, the rating can be realised after the end of the task, after a specific set of tasks or after completing an experiment.

2.2.1.2 Dual-task measures

The present approach is based on the assumption that human working memory resources are limited and can be shared when multiple tasks are executed in parallel. In that sense, if a primary task is more demanding than a secondary task, more processing resources will be allocated to the primary task. This variation of processing resources will result in a decrease in performance in the tasks with fewer employed resources and vice-versa. To summarize, the consumption of cognitive resources during the execution of a primary task is measured by the resulting performance in the secondary task. Usually, the performance of a secondary task is assessed by the response to an auditory stimulus, tapping or visual stimulus. For example, in [52] the dual-task method was used to assess in which format the information should be presented to the user to reduce the cognitive load. They presented as the primary task an interface with information about a historic city in different formats (audio-visual or visual only) and, as a secondary task, it was presented a letter rendered in the same interface (above the information content) that when it changed its colour the user was forced to tap the button space in the keyboard. The reaction time was used to assess the ECL inherent to the different formats. The results showed higher reaction times for visual-only multimedia than audio-visual multimedia. Meaning that audio-visual multimedia has less ECL than visual-only multimedia. However, this type of procedure is not well accepted in the field of HCI because the implementation of a secondary task can be very intrusive leading to an increase in the

user's demand for ECL and, consequently, to a decrease in performance in the primary task. In an attempt to solve this problem, in [53] was developed a technique called *embedded secondary-task* that aims to minimize the intrusiveness of secondary tasks. Instead of forcing the user to slip his attention between the primary task and secondary task, where the second task does not bring substantial benefit for his goal accomplishment (i.e. the inclusion of the letter "a" in the example above). The secondary task is built into the system as part of the current task and at the same time is also assessing his cognitive load. For example, measuring the reaction time to see a notification of "few fuels" (i.e. secondary task) when a subject is driving a car (i.e. primary task). With this method, the user is expecting a secondary task during the execution of the primary task, increasing its acceptance.

2.2.1.2.1 Physiological measures

This method has been researched in an effort to link changes in cognitive load with human physiological responses. The key benefit of physiological measurements is the lack of necessity to submit users to their feelings about a task experience because it can be very subjective. Another reason is the higher rate measurements and high friability that this approach offers compared to secondary-task and subjective measures. On the opposite side, the physiological measure has been subject to criticism due to the high obtrusiveness that measurement devices represent to the user. But currently, with technological advances, those devices became more minimized and less intrusive increasing user acceptance and suitability for many fields of study. In general, the state of the art in the physiological assessment of cognitive load has found several human functions that can be linked to cognitive load. From the perspective of [54], physiological measures can be sorted into three categories: eye function, brain function, and cardiac muscle function. Each category is different from the point of view of intrusiveness and reliability:

a) Muscle function- There is evidence that correlates cognitive load with variations in muscle tension. The underlying nature of this assumption regards the static muscle tension during cognitive stress and the resulting changes in blood flow in the muscle tissues. In [55] was found significant results appointing to an increase in cognitive load followed by a growth of tension in neck and back muscles. Tension in the arm and shoulder girdle muscles was also linked to an increase in cognitive load [56]. Usually, the experiments are based on the comparison of variations in muscle tension and different levels of cognitive load. The muscle tension is represented by values of signal amplitude of EMG [55]. In its turn, the cognitive load is evaluated by subjective, secondary-task or quantitative methods. However, in the view of [54], there still exists criticism about this approach because there are many contradictory results and the equipment required is very intrusive to be applied in HCI systems.

b) Cardiac Measures- Empirical evidence has linked heart rate, blood pressure and respiration as a predictor of cognitive load. However, the existent literature has been emphasizing cardiac measures due to their lack of intrusiveness. Cardiac measures are highly used in the field of aviation to evaluate the mental workload of a pilot during a flight. Related findings include Heart Rate (HR) and Heart Rate Variability (HRV) parameters as indexes to assess the cognitive load. The

assumption is that HR and HRV decrease along with an increase in cognitive load. Despite the less intrusiveness that this method represents to the user, HR and HRV alone as index parameters for cognitive load assessment have serious reliability problems because they are influenced by several psychological processes [57]. Moreover, this technique is not very accurate because it may only detect variations in cognitive load when the task demand is very high. Also, the period in which the measurements are reliable is between 40 seconds and 5 minutes, placing a big limitation on the applications that require instantaneous evaluations. In another study, to overcome this issue, the author [58] used breath rating along with HRV and HR to assess cognitive load achieving significant results.

c) Brain function measures- The most attractive neuroimaging technique for the assessment of cognitive load is called electroencephalography (EEG). Currently, it is considered the least invasive and the most cost-effective way to measure brain activity compared to the other techniques in the state of the art such as functional Magnetic Resonance Imaging (fMRI) or positron emission tomography (PET) [59]. The EEG aims to record the brain electrical activity generated by a neuronal activity that is passing through the scalp. It generates a complex waveform by using signal processing techniques to process the data. Currently, researchers attempt to extract an index of cognitive load from the spectrum analysis of the bands Alpha, Beta, Gamma, Theta and Delta[60]. However, this technique presents some limitations, according to [54] the recent results are often inaccurate and imprecise, with significant variability and low reliability. Those limitations may be justified by various factors that affect EEG like fatigue, anxiety, mild hypoxia, expertise, body movements etc. Moreover, the specific equipment and required technical experience to operate increases the unsuitability of this technique to assess cognitive load in modern HCI environments.

d) Eye function measures- In the past years, eye movements have sparkle the interest of the research community to investigate cognitive processes. Neuronal activity in the cortical location provokes small nervous responses in the eye movements and small dilation of the pupil. Those reactions have been studied with the assumption that they represent an index of cognitive load. The primary benefit of this method is that it has been demonstrated to be very unobtrusive and well accepted by the users regarding other existent techniques. Eye movements can be categorized into voluntary and involuntary movements. Voluntary movements consist of saccades and fixations. In its turn, pupil dilations and blinks represent involuntary movements. Fixations refer to the state in when the visual gaze remains in a single position during a fixed period of time. It has been linked to cognitive load assuming that long periods of fixations are related to an increase in attention that is influenced by the cognitive load [61]. On the other hand, fixations have some limitations, the fixations rate and durations can be highly influenced by factors such as drowsiness and fatigue [60]. Saccades refer to the shift of fixations between two locations. Findings appoint that saccade size and speed can be used as an index of Cognitive Load. However, saccade speed can be also affected by other factors such as habituation and fatigue [62][63].

The combination of multiple features of eye movements has proved to achieve higher accuracy for cognitive load assessment. In [63], the author found strong evidence that longer fixation and shorter saccades are related to higher cognitive load. In terms of involuntary movements, blinks have shown the most controversy while some authors claim that blink rate and blink latency are related to mental effort, others suggest that blink latency does not seem correlated to cognitive load [64][61]. In its turn, pupil dilation measurements have been considered the most sensitive method for cognitive assessment. While blinks have been mainly investigated in the field of visual-related tasks, pupil dilations have proved their accuracy in distinct fields involving perception, memory, problem-solving and reading [50]. In [65] the author found that during language processing there are variations in pupil dilations that reflects changes in cognitive load. Another study presented attractive results in the assessment of cognitive load in the field of visual searching tasks [66]. Despite the promising results, this method has a crucial limitation, the pupil is very sensitive to the environment bright and screen luminosity compromising its reliability in real scenarios of HCI. In order to overcome this limitation in [67] was developed and implemented a system of calibration that almost removed the geometry-based distortion allowing to increase in the accuracy of pupil dilation even when the conditions of the display bright are not constants.

2.2.2 Advantages and Limitations Cognitive Assessment Techniques

After a detailed review of existent techniques for the assessment of cognitive load, the present section aims to identify the advantages and disadvantages of each technique taking into account their applicability in the field of AUI. Those results can be consulted in Table 2-1.

Table 2-1-Advantages and disadvantages of cognitive load assessment techniques.

Technique	Advantages	Disadvantages
Subjective measures	+ Easy to manage and analyse. + Allow multi-dimensional and multi-factorial measures of cognitive load.	- Low reliability in long task. - User's cognitive limitations might influence it accuracy. - Unviable for instantaneously adaptive systems.
Primary – task measures	+ Very accurate for task performance. + Accurate for assessment of long periods of cognitive load. + Suitable for instantaneously AUI systems with only one task.	- Cognitive Load measurement severely influenced by individual characteristics. - Unable to measure cognitive load in dual-task scenarios.
Secondary-task measures	+ High accurate measurement of cognitive load for short periods. + High accuracy in distinct fields of study (i.e designing, text writing, information searching)	- Highly invasive for adaptive interface systems since it affects how the primary task is performed.
Physiological measurements	+ High sensitive for measurements of cognitive load for continuously periods. + Suitable for instantaneous AUI system	- In general, the required equipment is very expensive. -Equipment requires trained experts for the complex data analysis. - Highly intrusive for the user.

As demonstrated in table 2, physiological measurements are recognized as one of the most accurate measures to assess cognitive load. However, most of its techniques such as cardiac function, muscle function and brain function require the attachment of very invasive equipment to the user. As consequence, unless the aim of those techniques is evaluating systems or applications, they are useless for real scenarios because the user performance might be significantly reduced due to the invasiveness of the equipment.
In the opposite direction, eye function measures are not invasive and still reach the accuracy levels of the main physiological measurement techniques. Observing the user from an eye tracker provides a non-intrusive means which is crucial for the success of an adaptive system in real-time. In that sense, eye function measures have sparked the interest of this research. In literature is possible to find many existing works that correlate user cognitive abilities with a different set of eye-tracking measures such as fixation, gaze patterns, blinks, saccades and pupil dilation. However, this research is particularly focused on the fixations metric as it has been considered the most accurate eye function measurement of cognitive load.

2.3 Fixations as Metric of Cognitive load and Attention.

A fixation corresponds to a biological event where a gaze is maintained in a single location. The eye movement can be characterized by fixations and saccade, with the significant exception of smooth pursuit, which is mediated by a separate neurological substrate that appears to have evolved for hunting prey [68]. The term "fixation" refers to the point in time and location of focus or the act of fixating [68]. Fixation is the period of time between any two saccades when the eyes are essentially motionless. Perceptions have a tendency to vanish quickly in the absence of retinal jitter, a laboratory condition known as retinal stability. Fixational eye movement, which the nervous system does to preserve sight, continually activates neurons in the early visual regions of the brain to react to fleeting inputs [68][69].

Fixations have been studied as a useful indicator of cognitive load assessment, however, no consensus has been reached so far [70][71][72]. Commonly, such studies manipulate the mental cognitive load by adding a concurrent secondary task or changing the difficulty level of the task at hand [73]. By this, fixation-related information is found to be a reflective pattern of cognitive load by some studies, well others found opposite results. In [73] the author argues that this divergent pattern of results comes with the fact that those studies have assessed different types of mental cognitive load. Essentially, the studies that found that fixation duration decreased with mental cognitive load, were in fact, studying the effect of perception load on fixation duration. More specifically, in those studies, individuals were subject to experiments where their cognitive load was used to suppress irrelevant items during a task (i.e. choosing the right road based on its properties and complexity). In this type of experiment, the perceptual load is high, and subjects are forced to process a few relevant items within a single fixation which in turn decreases the fixation duration and increases the fixation frequency.

Fixations have been also intensively explored over the last decade for the understanding of the nature of human attention. In this respect, the eye tracker has thus far proven to be a key technology enabling determining the position of a user's visual attention allocation at any given moment. The attention information has been represented by long sequences of eye positions, saccades, and fixations. However, fixations have been the primary measure of visual attention in psychological studies of reading

comprehension [74], memory [75] and visual perception [76]. Fixations also have been used to analyse users' visual attention on a wide range of applications such as assessment of online learning, typing, multimedia content-aware image resizing and so on [77].
The fact that fixations-related parameters can be used to infer different levels of attention and cognitive load, places it as a very useful physiological measure to be employed in assistive technologies. Those metrics can serve for instance to locate places or objects where is the user's attention upon or to infer situations where the user is under high cognitive load.

Current literature applies artificial intelligence techniques for discovering user objects/regions of interest based on human visual attention in real-time. Typically, those techniques involve the training of deep learning models with fixations maps [21][67][68]. In its turn, cognitive load is usually assessed in offline settings, as its analysis needs statistical functions that are unfeasible for real-time settings. In this regard, pupil dilation is becoming more important as it shows to be very sensitive to mental activity. However, the current investigation is still undertaken in laboratory settings due to the noise of free-living settings (i.e. luminosity conditions) that severely affects the metrics used in this type of study such as the pupillary diameter mean and pupillary diameter standard deviations [80][81][82]. In that sense, human fixations assume to be the most efficient physiological measure of mental activity under free-living settings. Moreover, and from eye fixations, it's possible to extract not only the index of cognitive load but also insights over visual attention.

In this regard, fields such as automotive, education, healthcare, aviation etc. where assistive systems are becoming more and more prevalent, would benefit from this extra layer of intelligence because it would open a new range of possibilities for personalized assistance in real-time.

2.4 The Born of Deep Learning to the Beauty of Object Detection.

The first step toward deep learning was made in 1943 when Walter Pitts, a young mathematician, and Warren McCulloch, a neurophysiologist, published a paper on the behaviour of a neuron. They demonstrated a basic neural network with an electrical circuit and noted that neural connections became stronger with each use. A decade later Frank Rosenblatt began his work in the perceptron [83]. The Perceptron was a prototype of a type of neural network where binary neural units are connected via adjustable weights. That was constructed in accordance with the human brain's biological principles and showed an ability to learn. Essentially, it is a single-layer network based on an algorithm that performs the classification of inputs into two categories. A prediction is made by the neural network, such as "vehicle or bike" or "lion or zebra," and if it is incorrect, it is adjusted so that the following prediction will be more accurate. It becomes more accurate over thousands or millions of iterations. In this particular case, Rosenblatt's perceptron was able to classify different images of shapes or letters.

Over time, the understanding of biology turns out to be less literal, but the name has remained. At its core, remains a few fundamentals that can be found in a large of networks these days:

a) The layers that are frequently used to describe the alternation of linear and nonlinear processing units.

b) The use of the backpropagation technique that changes every one of the network's parameters at once.

However, despite promising and rapid advances, research in neural networks started to fall in interest from 1995 to 2005. Two main reasons can be pointed out as explanations for this lack of progress. Firstly, the enormous computational cost associated with the network training process. Computers back then just weren't fast enough or able to store enough data. Compared to these days, computers were millions of times weaker in terms of computational resources, making it almost impossible to exhibit intelligence. Secondly, the available datasets were very small. Till the 2010s the promise of Machine learning was limited by the prohibitive expense of getting labelled training data. It resulted in an academic focus on testing different algorithms on a relatively small number of canonical datasets [84].

Strong statistical techniques including kernel methods, decision trees, and graphical models have proven to be experimentally better given the lack of data and processing power. This comes from the fact that, unlike neural networks, they did not demand high computational power, and the existent datasets were sufficient to achieve expectable results.

However, over the last few years was possible to observe a turnover in neural networks' popularity. It started to slow down due to the appearance of cheap data storage and computation in a particular form of GPUs. These were the requirements needed for the success of Deep Learning. Along with this, computation power also rases in hands with the information explosion, another advantage of Deep Learning as it performs better with a large amount of data.

At a certain point, the random-access memory (RAM) was outpaced by the data available which lead to statistical models' need for becoming more memory efficient (generally achieved by adding nonlinearities) while simultaneously spending more time on optimizing these parameters. As result, the deep learning performance determined the transition between statistical methods (typically linear models and kernel methods) to deep learning methods, such as convolutional neural networks [85], long short-term memory [86] and Q-learning [87].

2.4.1 Convolution neural networks on object detection

The rapid progress of Deep Learning sparked the interest of the research community in a variety of fields including, robotics, healthcare, finances, automobile, etc. The performance of state-of-the-art conventional machine learning techniques, used till 2010 (linear regression, Naïve Bayes, Random Forest, etc), were outperformed by Deep Learning. Its ability to learn over the massive amount of data and to extract higher-level features directly from the raw data turned Deep learning into a more reliable, adaptive and accurate model.

There are plenty of advancements in deep learning that boosted some hot research topics. For instance, in computer vision, several deep learning methods have been employed on object detection for improving traffic sign classifiers in autonomous vehicles [88].

Convolutional Neural Networks (CNN) were firstly introduced by Fukushima in 1998 [89] and have been extensively applied in a wide range of different fields including activity recognition [54][91], sentence classification [92], face recognition [93], object detection and localization [94][95], image characterization [96][97], etc.

The idea behind CNN was to mimic the way visual system structures work and in particular, the models of it proposed in [98]. In fact, recent findings demonstrate that modern CNNs and the human brain share many key structural and functional substrates [99]. CNN take images and from them, they learn the patterns, the building blocks that make them up. For instance, as exposed in Figure 2-1 at the shallow levels of a network it might learn things like line segments that are at different angles and then at subsequent layers those get built into shapes like faces or elements of cars depending on the images that it is trained the network on.

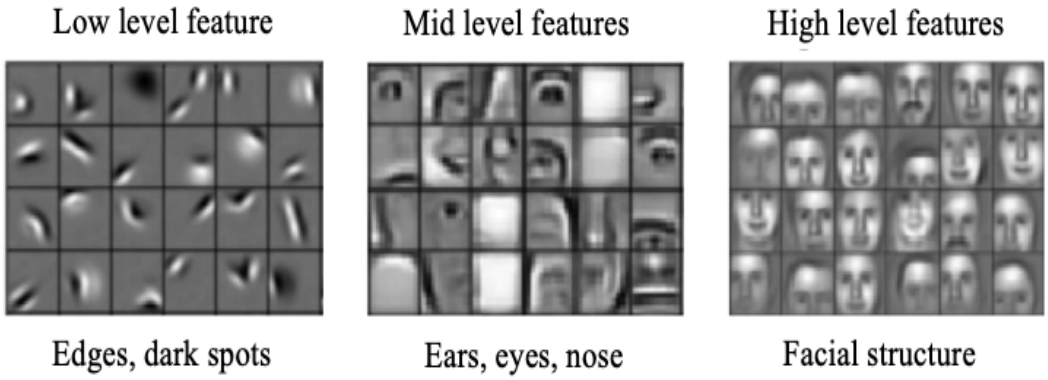

Figure 2-1- A hierarchy of features [100] .

The major challenging part that occurs during the training process of a CNN, is the overfitting problem. This phenomenon takes palace because of the huge number of parameters that have to be learned. To this end, techniques such as stochastic pooling, dropout, and data augmentation have been proposed [101,102].

CNNs are also frequently subjected to pretraining, that is, to a process that initializes the network with pretrained parameters instead of randomly set ones. Pretraining can accelerate the learning process as well as enhance the generalization capability of the model.

Before CNN, feature extraction methods for image recognition were very labour process. Nevertheless, CNN came to enable a more scalable approach to both images and object recognition tasks, leveraging principles from linear algebra, specifically matrix multiplication, to identify patterns within an image.

2.4.2 Application of CNN for Object detection.

As mentioned in the previous subchapter, the great performance exhibited by CNN found its application in the field of object detection. Prior to deep learning's rise in popularity in computer vision, object detection algorithms were mainly based on hand-crafted machine learning features such as shift invariant feature transform [103], histogram of oriented gradients (HOG) [104] etc. However, little progress has been made until the appearance of CNN. In 2012, the resurgence of “AlexNet” sparked the interest of researchers from a variety of fields within computer vision like image classification, localization, object detection, segmentation etc. and from that several object detection models have emerged in coming years such as Fast R-CNN, YOLO, SSD and MobileNet.

The ultimate purpose of object detection is to detect multiple objects of interest within an image and understand how they are placed within the image. Object detection falls into two groups from the standpoint of applications: generic object detection and

detection applications. The former focuses on exploring the techniques for detecting various types of objects within a single framework to mimic human vision and cognition, while the latter refers to the detection under specific application scenarios, such as vehicle detection, pose detection, animal detection, etc. This research is focused on the topic of "detection applications" aiming to detect objects derived from human visual attention.

2.4.3 Object Detector Algorithms

In this section is going to be given a detailed description of the architecture of popular object detection algorithms such as Fast R-CNN, YOLO, Single Shot, and MobiletNet.

2.4.3.1.1 Fast R-CNN

Fast R-CNN is known as an enhanced version of the R-CNN object detector model. It was proposed by [105] to overcome some R-CNN limitations such as a slow training process and low inference speed. Fast R-CNN provides a new architectural paradigm that increases detection accuracy while simultaneously speeding up training and testing process.

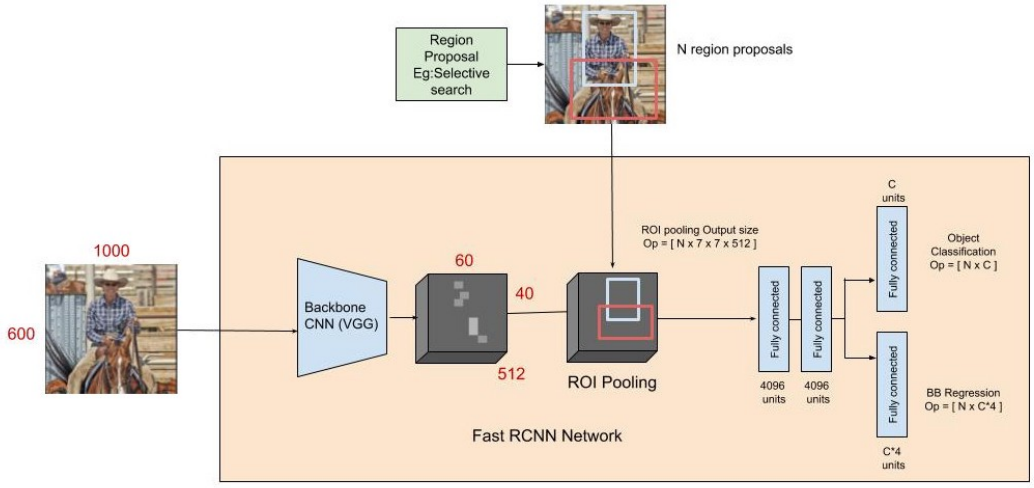

Figure 2-2-Fast R-CNN architecture [105].

As can be observed, instead of applying convolution layers on each object proposal, Fast R-CNN, image is fed into convolutional layers only once. After image features are extracted using CNN, in the regions of interest (RoIs) feature extraction is applied on the convolutional feature map. Each RoI is polled into a fixed-size vector from the feature map. Then, vectors are fed into a sequence of fully connected layers and produce two branches of outputs per RoI. One is used to output a class of the object by applying SoftMax probabilities. Finally, another branch is used to define the bounding box position of one of the classes.

2.4.3.1.2 YOLO

YOLO is the abbreviation of “You Only Look Once”. As can be perceived by the name, the authors changed paradigms of previous object detection algorithms which relied on proposal detection plus verification [106]. YOLO introduced a new philosophy: it applies a single neural network to the full image. The iamge is divided into regions by this network, which concurrently preficts bounding boxes and probabilities for each region.

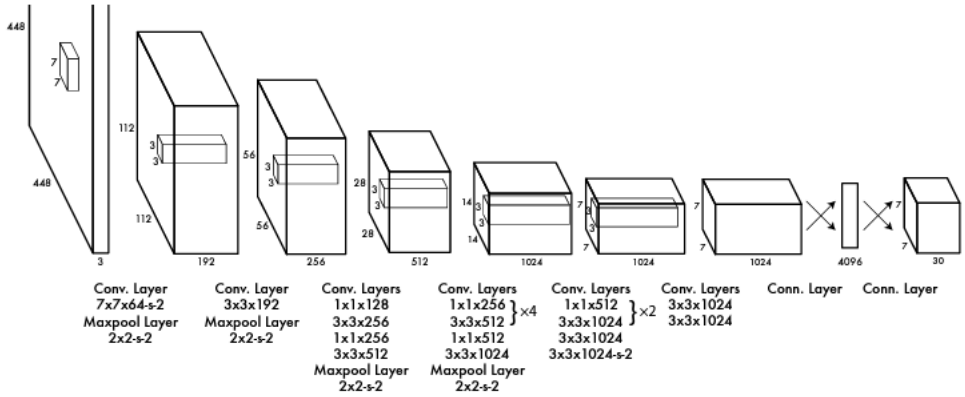

Figure 2-3- YOLO architecture [106].

The three essential components of the YOLO model are the head, neck, and backbone. The backbone of the network is composed of convolutional layers aiming to detect and process the key features of an image. It is first trained on a classification dataset, such as ImageNet, and it is commonly trained at a lower resolution than the final detection model. It is designed in that manner because detection demands finer details than classification. The neck makes predictions on probabilities and bounding box coordinates by making use of features from the convolution layers in the backbone with fully connected layers. The head is the network's last output layer and can be switched out for other layers with the same input shape to facilitate transfer learning.

With the introduction of its v2 and v3 editions [107][108], that also make a higher detection accuracy while maintaining a very fast detection speed, YOLO has seen its model undergo a series of upgrades since its beginnings.

2.4.3.1.3 Single Shot Detector (SSD)

SSD (Single Shot Detector) is one of the most advanced object detection algorithms, it introduced high detection accuracy with real-time speed [109]. The main difference between them is how the bounding boxes are generated. SSD proposes a much simple approach, for each pixel, it is generated multiple point anchor boxes centred at this pixel. In the end, those multiple anchors are represented by different shapes and ratios.

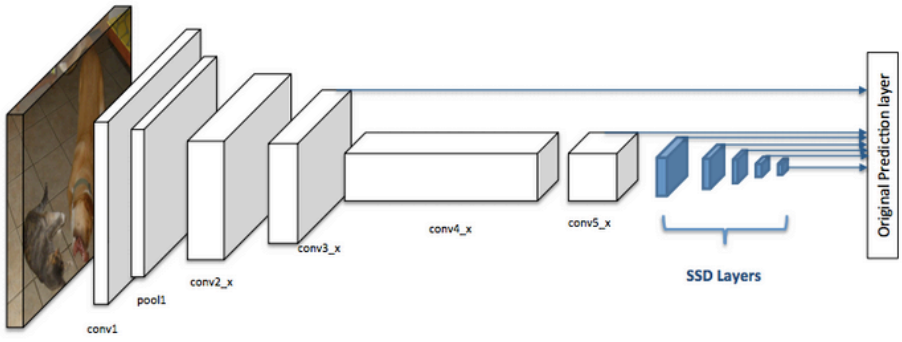

Figure 2-4- Architecture of a convolutional neural network with an SSD detector [109].

In terms of structure, SSD is composed of two main components, a backbone model and an SSD head. Backbone model is a pre-trained image classification network as a feature extractor for instance a ResNet network [110]. Then anchor boxes are generated and the output of each anchor is represented as the bounding boxes and classes of objects in the spatial location of the final activation layers. This is followed by a multiscale feature to reduce the width and height of the feature map. This process is repeated multiple times and enables earlier layers bearing smaller receptive fields to represent smaller-sized objects and predictions from later layers to help in dealing with bigger-sized objects.

2.4.3.1.4 MobileNet

MobileNet is a lightweight convolutional network that is much smaller and faster in size than mainstreaming popular object detector models. To give a comparison in regards to size, MobileNet is hundreds of times lighter than a full VGG16 network [111]. This huge difference gives MobileNet the requirements needed for deploying a model to run on a mobile app for example. The lightness of the network is due to the number of parameters or weights and biases contained in the model. There is a direct correlation between the number of parameters that a model has and the space in memory it will use. However, to be much faster and lighter there is a trade-off involved, and it concerns the model accuracy. MobileNet is not as accurate as most of the popular models.

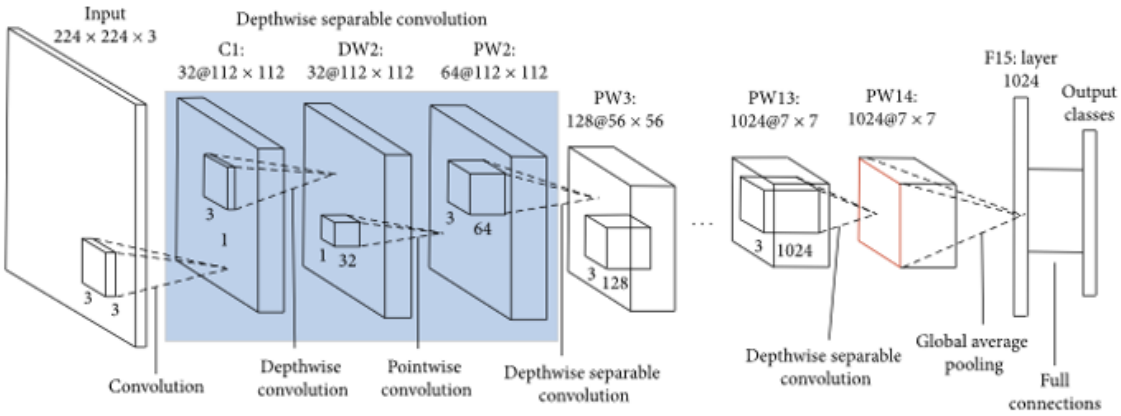

Figure 2-5- MobileNet Architecture [111].

In terms of its structure, MobileNet is designed on depth-wise separable convolutions except for the first layer which is a full convolution. In Figure 2-5, the MobileNet architecture is described. The distribution of the layers is succeeded by a batchnorm [112] and ReLU nonlinearity with the exception of the final fully connected layer which has no nonlinearity and feeds into a softmax layer for classification. An important difference between the regular convolutions layer is that mobileNet uses a convolutional layer with a Depthwise and 1 x 1 Pointwise layer with a batch norm and ReLU after each of the convolutional layers. Before the fully connected layer, there is an average pooling aimed to reduce the spatial resolution to 1.

2.5 Visual Attention-Based Object Detection in Cluttered Environments

Over the last decade, eye movements have been intensively explored for the understanding of the nature of human attention. In this regard, the eye tracker emerged as an important enabling technology to extract where a user's visual attention allocation is at any given point in time. The attention information has been represented by long sequences of eye positions, saccades and fixations. Among them, fixations have been the core metric for psychological experiments on visual attention in the context of reading comprehension [74], memory [75] and visual perception [76]. Fixations also have been used to analyse users' visual attention on a wide range of applications such as assessment of online learning, typing, multimedia content-aware image resizing and so on [77].

More recently, the investigation has shifted towards improving or creating new models of user context for situation-ware interaction. Here, the ability to recognize objects of interest assumes to be critical for constructing an accurate model. In this regard, egocentric cameras and head-worn eye trackers are becoming very popular as they overcome the major limitation of common approaches which usually require a sensor-equipped environment.

More advanced approaches make use of Fixation Density Maps (FDM) generated from eye-tracking experiments to train neural networks to predict objects' location of the user's visual attention in an image [113][114]. The result is a saliency map of regions that are then extrapolated to discover objects of interest.

Although existing research proves to be promising, there are also limitations and open issues. Due to the nature of this approach, it is designed to be task-specific, because the predicted objects do not necessarily correspond to the ones the user is focused on. Instead, they show potential objects of interest due to their shape contours, size and location. Here the variability of the individual natural gaze exploration is not considered because it depends on the user's motivation and the specific task that he is performing [77]. Moreover, in daily living contexts, i.e. driving, where a user's environment is continuously changing, the nature of this approach turns out to be unviable for this type of setting.

In an attempt to overcome this obstacle, other approaches start by using egocentric cameras based on the assumption that user visual attention is allocated at the centre of the visual field [114]. This assumption is supported by the fact that in the centre of the human retina there is a much higher resolution compared to its more peripheral regions [115], so the human gaze tends to be focused on the centre region.

Studies have used a pre-defined spatial pixels region of egocentric images used as feature input to object recognition algorithms. However, results from psychological experiments have shown that humans can focus their attention on peripheral regions

distanced from the gaze centre. An example is a car driver who is fixated on the road but is still able to monitor road signs or potential pedestrians trying to cross the road. Evidence show that this redirection of visual attention is preceded by prior fixations in peripheral regions [77]. This effect is only perceived by analysing eye movements on eye trackers. Meaning that the fixed centre region approach would fail to analyse important information outside of the centre of the visual field.

An alternative approach that seeks to reduce the low accuracy of fixed region models, is to use head-worn eye-tracking equipment and egocentric cameras for gaze-based object detection [115][116][117]. Here, object detection algorithms are used in combination with fixation data to discover the object of interest. This approach was applied to the creation of digital episodic memories to support people cognitively impaired people, which has where demonstrated good results [118].

Despite that progress has been made in controlled experiment environments, it remains a challenge to apply egocentric visual attention object recognition when moving these solutions to real-life environments. In most cases, the existent solutions perform well because they are trained using data exclusively collected for task-specific and in a controlled environment. The main issue is to extract visual attention in dynamic cluttered environments, where objects are constantly overlapped and partial object occlusion is also frequent.

2.5.1 Visual Attention Detection Based on Saliency Map

Recently in the field of computer vision, researchers use human fixations as input to train neural network algorithms to create saliency maps of an image. The result is a combination of the probability of saliency with a spatial colour variation of pixels, which represents human visual attention over a specific region. Saliency prediction models can be divided into two categories: bottom-up and top-down [119]. The bottom-up category is focused on building saliency maps based on intrinsic features of the image like its colour intensity, object shape, size and location. In addition to bottom-up models, the top-down category takes also into consideration external cues that influence human attention like facial presence, attractive objects and motion. In [120] is presented a bottom-up model that combines global contrast with a probability of saliency. A different approach is proposed by [121] that combines local and global contract image features to predict saliency in cluttered scenes. In an attempt to increase the accuracy of the saliency map, in [119] is included top-down cues such as face presence and movement combined with global contrast features.

Other works exploited cognitive features and scales for a top-down visual search model by using multivariate Gaussian distributions [122]. More recently, [123] combined several bottom-up and top-down saliency models using several combination strategies. Saliency maps reveal to be very useful to extend this work on the understanding of human visual attention and what object features influence it the most. However, in the context of this work that envisions the detection of objects under visual attention in runtime, saliency maps would not perform well because it is designed to predict fixations in specific regions of an image which not necessarily correspond to the user attention in a particular temporal space.

2.5.2 Egocentric-based Object Detection

The recent proliferation of wearable cameras prompted an interest in exploring visual analysis in a wide range of research fields such as activity recognition, handled object recognition, and object prediction of joint visual attention. Independently of the application domain, there is a common goal among these research topics which is producing a model capable of inferring the object of interest. Here, the discovery of the user's visual attention assumes to be a critical step for the model's accuracy. In the absence of visual data, common approaches define a specific fixed region in the first-person video, for which it is extracted features to describe objects being viewed. In [124], the centre of the image frame is assumed to be the region under the user's visual attention. Then, a Euclidean function is applied to measure the distance from the object detected to the centre frame in order to estimate how likely an object is being focused. In the same direction, [125] computes the distance between the centroid of an object and the centre of the image as input features for saliency detection on egocentric video.

In the context of this work, the fixed-size region's approach would not work well because due to the variability in the size of objects and the different angle perspectives, only a limited part of object features is often described in the centre region. Moreover, findings suggest that despite human visual attention being mostly focused on the spatial centre region of the visual field, attention can be also observed in peripheral regions [126].

2.5.3 Fixation-Based Object Detection

Recent technological advances have made head-worn eye trackers more affordable and widely accessible. This effect sparked the interest in the development of an interactive system that incorporates user visual data as an indicator of user visual attention for the object of interest detection. A comprehensive list of different application domains for head-worn eye trackers is provided by [127]. However, one of the issues regarding the eye tracker cameras is the lack of accuracy of the gaze data points resulted from inefficient calibrations. Most of the approaches interpolate data gaze with output from state of art object detection models, for summarizing of object of interest. As result, the accuracy of gaze data becomes very important for the performance of those models. Few studies have tackled this challenge. In [118], it is cropped a rectangle region from the image centred on the fixation point and then extracted SIFT features from that region for object detection. The rectangle region is composed based on a fixed number of fixations that occurred around and after the first fixation. A different approach is proposed by where fixations in fixation density maps that capture the spatial density of fixation over the image. Further, a Gaze Polling layer combines visual feature and fixation density maps for the prediction of object categories [113].

However, prior solutions are designed exclusively on physically manipulated objects, that are specific to certain tasks. None of them, fully exploit common cluttered environments that can be found in free-living settings.

Considering a scenario where fixations point disposed on the centre of an object that is overlapping a second object (i.e., keyboard and laptop or book and pencil) both approaches would consider a region that is concurrent to both objects. As result, the input features provided to the object detection model would classify both objects instead of the target one.

2.6 History of Blockchain Technology

Blockchain technology has recently sparked the interest of not only the research community but also regular citizens. This interest is justified by its grounds conceptualization of shifting the natural course of the financial system, a centralized digital current and a central baking system to a decentralized currency concept named Bitcoin. In its essence, what Satoshi Nakamoto proposed in 2008, was a distributed timestamp server on a peer-to-peer basis, that works as a generator of computational proof of the chronological orders of transactions [128]. Here, exists the concept of a "digital coin", which is defined as a chain of digital signatures. Each transition occurred in the network is signed by the owner of the transaction as well as the public key of the next owner.

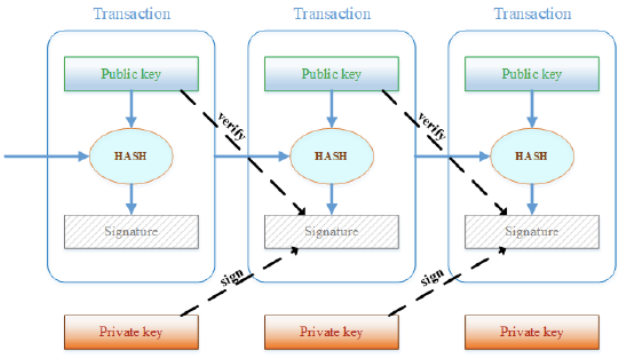

Figure 2-6- Transaction structure in a Bitcoin Blockchain.

As can be observed in Figure 2-6, both private and public keys have different goals. The public key aims to verify the transactions. In its turn, the signing process of the transactions is performed by the private key. Both keys are stored in the wallet, which can be a device, physical medium, program or service. The transaction records are held on blocks which cannot be modified as they are linked between each other as chain of blocks, secured by cryptographic techniques.

2.6.1 Digital signature

Digital signatures are a key component for constructing new blocks in blockchains, as they are responsible for the authentication of transactions. A digital signature's purpose is to demonstrate to each node in the network, that a user is authorized to submit a transition, meaning that his funds are safe from being spent by other users. This transaction is then verified by all the network nodes. For instance, if user "X" decides to send 1 Ethereum to user "Y", "X" must sign a transaction of 1 Ethereum with his private key and send it to nodes on the network. If all the conditions presented in the transaction are met, the miners that know their public key are ready to validate the signature. A validator or miner can now finalize the block containing that transaction after validity has been established. Most popular blockchains such as Ethereum and bitcoin, apply the same digital signature schema. It is known as Elliptic Curve Digital Signature Algorithm (ECDSA) and is the elliptic curve analogue of the Digital Signature Algorithm (DSA) [129]. Both Blockchains work with the elliptic curve secp256k1 which is recommended by the Stands for Efficient Cryptography Group.

The ECDSA algorithm relies on the elliptic curve Secp256k1 to generate signatures that are hard to forge and easy to verify. The emergence of Secp256k1 and its rise in terms of popularity was due to its designed criteria. Essentially it has an internal structure that was designed to increase the speed of calculations [130].

2.6.2 Blockchain Blocks

The blockchain is composed of blocks that contain data about the state of the blockchain. They can be seen as data structures that include a block header, a block identifier and transactions.

2.6.2.1.1 Block header

The blockhead is composed of information like the hash of the previous block, merkel rout, transactions, timestamp, version and a nonce number. The Merkle tree is responsible to store the hash of the data blocks of transactions validated by the blockchain network. It ensures that data sent through a peer-to-peer network is secure. In its turn, the timestamp registers the exact time when the mining process of the block is finalized. The timestamp of the current block needs to be greater than the timestamp of the previous block. To prevent Time Warp Attacks , typically orchestrated by malicious miners that undermine incorrect timestamps on the blocks they mine to decrease the level of mining difficulty, the blocks are generated at every 10 minutes. Finally, using the algorithm of proof of work (described in subsection 2.6.3.1.1 the hash of each header is generated.

2.6.2.1.2 Block Identifier

The block identifier is a data structure that store the hash of the block, the hash of the previous block, the version, the registers of the valid transactions and a nonce number.

The immutability of blockchain is achieved by the process represented in Figure 2-7. The hash of a block is generated by passing through a Secure Hash Algorithm (SHA-256) the current block metadata along with the hash of the previous block. By this, the blocks become linked to each other as a chain of blocks. Meaning that, if a bit of information has tampered in one block it will break the chain next blocks [131].

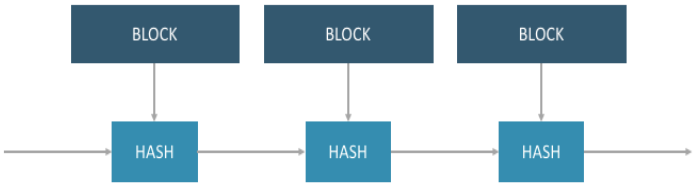

Figure 2-7- The blockchain scheme.

2.6.3 Consensus

One of the most valuable features of blockchain technology is its capacity to validate and verify transactions without a central authority. This is autonomous network behaviour that ensures that every node within the network participates to reach a consensus on the state of the network. Essentially, the network has a consensus algorithm to ensure that every new block broadcasted to the network, is a legitimate block. In the realm of blockchain technology, there are two popular consensus algorithms such as Proof-of-work (PoW) and Proof-of-stake (PoS). The process of how these algorithms operate will be discussed in the next subchapters.

2.6.3.1.1 Proof-of-work

This consensus algorithm involves solving a complex computational exercise performed by what are called validators, miners or nodes.

Miners compete against each other to solve this computational exercise, as rewards are granted to the ones who solve them first. Despite the verification of the correct solution being easy, the challenge of solving this exercise demands high computational power. When a miner accomplishes to find the correct solution, the result is forwarded to the other miners for verification. After this verification succeeds, the new block is approved for the blockchain.

PoW is performed by adding a nonce number to the block. Essentially, this value is obtained by applying a cryptographic hash function similar to Hashcash [132] and based on SHA-256 [133]. The PoW is considered valid when the prefix of the hash function output is composed of a predefined number of zeros bits. The effort of "guessing" the correct nonce is directly correlated to a number of zeros encountered in the prefix of the hash. An example of the PoW mechanism is illustrated in

Figure 2-8. The miner tries to guess a nonce value that in combination with the metadata of the block passes through a cryptographic schema and matches the number of four zeros that is part of the prefix of the hash function. The rewards are only distributed to the miner when the given conditions are met and verified by the other miners.

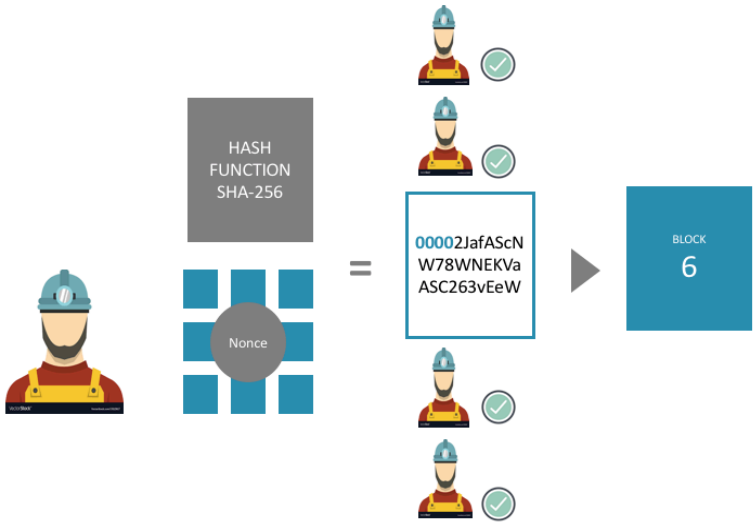

Figure 2-8-Proof-of-work schema.

By this, the PoW is able to mitigate the double spend attack where malicious actor can potentially try to spend the same digital currency unit twice. One of the downsides of the PoW is the amount of hardware resources needed to solve the computational exercise. Despise being a good mechanism to secure the network, it demands high energy consumption that rises issues in terms of sustainability.

2.6.3.1.2 Proof-of-stake

The PoS algorithm emerged as an alternative to the popular PoW. Both algorithms are aimed to reach the consensus of integrating new blocks into the network, however, the PoS operates slightly differently. Its design was conceived to overcome the major limitation of PoW such as the higher energy consumption and need for high computational power. Instead of using forcing nodes to have powerful hardware resources to compete in solving a complex computational exercise, it uses a pseudo-random election process to select the node to be the validator of the next block. This selection takes into account a set of parameters that includes the staking age, randomisation and the number of the network digital coins staked by the node. Here, the node's wealth is what significantly affects the probability of a node becoming the next validator. This design significantly reduces the network congestion as well as environmental sustainability concerns surrounding the PoW.

In PoS the block creators are called validators, whereas in PoW are called miners. Another difference between both consensus algorithms is that in PoS validators are rewarded by the network transaction fees, instead of receiving block rewards for the mining process. In order to not favour the network's wealthiest nodes of becoming validators, there were conceived two methods for the selection process:

a) **Randomized block selection:**

The selection process to find the new validator takes into consideration a combination of the candidate lowest hash value and the highest stakes. As this information is public and can be consulted by the other nodes, it is possible to predict which node is going to be the next forger [134].

b) **The coin age selection:**

The validator is selected accordingly to the period of time that the node's digital coin has been staked for in the network. The equation of the coin age selection can be consulted below:

$$Coin\ age = Number\ of\ days\ stacked \times Number\ of\ coins\ staked$$

In order to prevent the increasing dominance of the wealthiest nodes within the network, the Coin age parameter is set to zero and the node is forced to spend a certain period of time without being able to forge another block [135].

Another security measure adopted by the PoS to prevent the dominance of the network from a group of participants, in this case, the 51% attack, is that it would require them to have 51% of the staked digital coin which would be extremely expensive. Another

measure that PoS against malicious activities is that in case a fraudulent transaction being detected, the node loses a percentage of the digital coins staked as well as its right to participate as a validator.

A node that is selected to forge the next block will verify that the transactions in the block are legitimate, signs the block, and adds it to the blockchain. The node is then rewarded with transaction fees derived from the transactions. In case of the node decides to stop being a forger, it will be in a locking period for the safety of the network, primarily to verify that they do not added any fraudulent blocks to the network. After the locking period is finished, the node receives both stakes and earned rewards [136].

The main advantages of the Pos algorithm are energy efficiency and security. As there is no need for users to be provided with powerful computational resources, it incentivises more users to participate in the network as it is more affordable. Moreover, since mining pools are no longer required for the mining process of the blocks, together with the randomization process of the PoS makes the network more decentralized and consequently more secure.

1.1.2. Smart Contract

Smart contracts can be compared as physical contracts that are established in the everyday life, however, smart contracts are digital. In detail, a smart contract is a form a computer program which automatically execute transactions without an outside entity to intermediate it. It is a piece of coding that exist within a distributed and decentralized blockchain protocol. This piece of code is responsible for the execution of transactions that are irreversible and can be trackable across the network. These smartcontracts can be used to register a house, transfer funds or trace products.

The smart contracts operate on a blockchain that supports the Ethereum Virtual Machine, which is s software based on Ethereum blockchain that executes and computes the state of the protocol chain after a new block being integrated to the network.

1.1.3. Taxonomy of blockchain systems

Blockchain protocols can be differentiated based on the type of permission that are conceived for the network participants. In Table 2-2 can be consulted the three different type and their specifications.

Table 2-2- Classification of the different types of blockchain technology.

Blockchain Type	Description	Resources
Public blockchain	A public blockchain means that any participant can be a part of the blockchain network as a validator or a miner. Users can join the network having permissions for reading and writing. The truth essence of a public blockchain is that is permissionless.	[137–139]
Private Blockchain	A private blockchain is the opposite of a public blockchain. Meaning that only a restricted set of users are invited to participate in the network activity. Here the control of the network is considered centralized as only restrict entities have the authority over the mining process and the consensus algorithm.	[140–142]
Consortium blockchain	This type of blockchain is an extension of the private blockchain, it attempts to remove the sole autonomy of the private blockchain. Despite being a permission network, it offers a decentralized structure. Usually, is a group of companies or institutes that collaborate, share data and maintain the rules and records within the network.	[143,144]

1.1.4. Motivations for Blockchain-based EHR Systems

Towards the end of 2013, blockchain developers were about to make an important decision. That would mean or either decider to continue the legacy Bitcoin by continuing to develop on top of it or embrace a new whole idea and conceive a new blockchain. Building on top of Bitcoin, meant to work constantly trying to overcome and walkaround over the constraints of the network. The limited set of transaction types, data types, and size of data storage would probably reduce the number and type of applications that could run directly on bitcoin blockchain. That meant, that a significant amount of work was expected to be built on off-chain solutions renouncing the main advantages of a pure decentralized blockchain. These limitations motivated the launch of Ethereum blockchain as a way overcome the need for supporting semantically richer applications. With Ethereum emerged the notion of smart contract, a representation of autonomous programs that lead to a new paradigm of decentralized applications. Those were characterized by having open source code, work without third-party intervention, need for cryptographic token for access as well as a consensus method for generating tokens. Those features made Ethereum blockchain popular and its success resulted on the adoption of the blockchain by a variety of fields ranging from digital assets, Artificial Intelligence, Internet of Things and HealthCare etc.

Assistive technologies would also benefit from the adoption of blockchain technology in as a mean to foster the exchange of Electronic Health Records (EHR) between multiple parties in a more secure and decentralized manner. In [145] is proposed a public blockchain framework for health-related assistive technologies that supports real-time patient monitoring and medical interventions while maintaining EHRs in a cloud based storage. In the same direction, the author in [146] presented a blockchain solution that not only facilitates EHR sharing between different healthcare institutions as well as researchers. Additionally, the proposed system uses an encryption schema to share data in a secure manner among involved parties.

Another framework proposes a model to offer patients the possibility to share and control their personal health data in a secure and private fashion by leveraging the benefits of blockchain technology [147]. This framework provides achieves higher security standards as it makes use of Multi-party Computing and Indicator-Centric Schema. Alternatively, MedChain in [148] introduces an efficient data-sharing framework for health data sharing thanks to a integration of different technology such blockchain, digest

chain, and structured P2P network mechanisms. MedChain aims to overcome scalability and latency issues of a blockchain network in scenarios where devices are sharing sensitive data among multiple entities.

Most of the presented studies are focused on solving two main issues:

a) Lack of privacy, where solutions are developed to protect the authorization of private access to personal data of individual patients that are often compromised by malicious actors within the system.

b) Lack of security, meaning that in order keep personal data safe solutions are proposed to secure the channels of interactions among the system, professionals, patients, and assistive devices.

Nevertheless, most of the contributions presented in these studies offer solutions based on external databases for storing the user's raw data while the index and location of the data are preserved in the blockchain. Or instead of a decentralized solution, it is offered a more centralized-oriented approach where the data is kept on a private blockchain, and the index of the data is kept on a consortium blockchain. However, with this type of solution, there are some drawbacks. Forcing the data to be stored at a single entry increases the risk of a single point of failure while using two types of blockchain might significantly affect the computational cost of the whole system. Also, encryption techniques used to protect access to personal information, which is a common practice, may result in high computational complexity and latency.

2.7 Privacy Protection in Blockchain Systems

In the last few years, blockchain technology has sparked the interest not only of the research community but also at an enterprise level to integrate it into business applications. Blockchain technology promised to bring a decentralized trust where applications would benefit from being robust, unstoppable, censorship-resistant, and transparent. This would radically transform the way industry and services developed their models of data storage, tracking, and payment systems. In other words, these features would minimize the role of the middleman, increase the layer of security and ultimately achieve economic growth. However, most blockchain protocols and as consequence smart contracts have a critical issue that put at risk its adoption- all the data stored is public.

With this lack of privacy, no single enterprise would integrate blockchain into their business knowing a priori that their most sensitive or valuable data is exposed to the public domain. In that sense, it is clear that privacy plays a major role in the adoption of blockchain technology. Currently, the investigation on the privacy side of the blockchain is focused on identity privacy preservation and transaction privacy preservation.

2.7.1 Identity Privacy Preservation

Identity privacy preservation refers to the ability to preserve the anonymity between senders and recipients of a transaction. In a blockchain ecosystem it is possible to establish a link between entities of transactions which means that by analysing the available public data it's possible to infer some privacy information. In case of the identity

of a user being liked to its digital address the address of a transaction, it may cause the exposure of all his transactions. In this respect, there are three major solutions to preserve identity privacy such as mixing services, centralized mixing services and decentralized mixing services. Mixing services were first introduced by [149] that build a methodology that enabled systems to hide the communication and its content between participants. The concept is exposed in Figure 2-9 and denotes that one entity intents to deliver a message to another participant at the address R. For the process to succeed, participant A needs to encrypt the M with the receiver's PK, appending the address R. The result is encrypted with the intermediary's public key KI. In its turn, the intermediary decrypts the message using his private key and encrypts its again using the public key of participant R. Finally, the intermediary delivers the encrypted message to R who then decrypts it with his private key. Once multiple inputs and outputs are processed by the intermediary, this method will hide the correspondences between each message's origin and destination. The order of arrival is hidden by outputting the uniformly sized items in random patterns. Moreover, to reduce the likelihood of the single intermediary suffering an attack, multiple intermediaries might be linked together by generating a mixed cascade [150]. In the blockchain ecosystem, the intermediate is often a crypto mechanism that random exchanges a user's coins with other users' coins. Consequently, ownership of coins is hidden from outside observers.

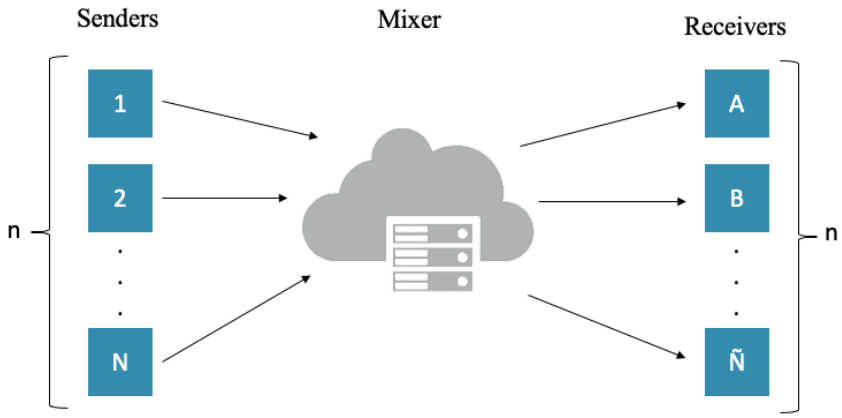

Figure 2-9-Basic Architecture of mixing services.

These types of services have been utilized among blockchain users with an intent to enable privacy when executing transactions and preserving their identity anonym. The research has been mostly investigated in two main research areas such as centralized mixing and decentralized mixing.

2.7.1.1 Centralized Mixing Services

Mixcoin was one first centrilized mixer services to achieve a strong anonymous payment in Bitcoin and related cryptocurrencies by providing cryptographic accountability, randomized mixing fees, and an adaptation of mix networks to Bitcoin. Moreover, Mixcoin uses a signature-based accountability mechanism to detect fraud or stealing [151]. Another project that seeks to provide identity anonymity is Dash [152]. At its core, Dash uses a coin-mixing service named PrivateSend that is designed to

provide privacy to the user by removing from the blockchain all sort of confidential information that can be linked to the user. The service is composed of multiple nodes that controls the mixing mechanism to only accept specific denominations. Nevertheless, depending on this service load, it is restricted to a certain number of online users.
TumbleBit was one of the first projects to deliver a service aiming to simultaneously achieve full unlinkability and prevent coin theft [153]. The methodology used behind this project is based on a central service, however, through a zero-knowledge proof algorithm and two-party computation, it achieves the user identity's privacy and transaction fairness. Nonetheless, the author in [154] pointed out that TumbleBit presents several issues such as the inability to perform multiple payments in one single transition and high execution time that makes this service unfeasible for real-time applications.

In the end, centralized mixer services face relevant issues that compromise their adoptions. They are still vulnerable to Denial-of-Service(DoS) attacks and users can incur in high mixing fees.

2.7.1.2 Decentralized Mixing Services

One of the reasons for the existence of decentralized mixing services was to mitigate the vulnerability of a DoS attack as mentioned above. Such services, as claimed in [150], allow peers who are not trusted by each other to broadcast their messages concurrently and anonymously without the requirement for a third-party anonymity proxy.

In this regard, CoinJoin proposes a special way to perform transactions in order to preserve the anonymity of users [155]. The core idea behind this project is to make a joint payment between users. Let's say that a number of users desire to execute a payment, the mechanism mixes the link between the inputs and outputs of their transactions leading to a significant reduction of the likeability of being tracked by other peers.

A user needs to find another user that wants to make a payment and the resulting joint transaction mixes the link between inputs and outputs which reduces significantly the likelihood to trace the direction of data among other peers. However, this service has a limit number of users to reduce the likelihood of a DoS attack, this lowers the internal unlinkability between users in a joint transaction. To address this issue CoinShuffle provides an extra layer of privacy by excluding trusted third parties from the process of mixing transactions . This project is referenced as a total decentralized coin-mixing protocol offering a security solution that mitigates theft. Moreover, it integrates Dissent, an enhanced protocol that enables privacy at the accountable level [156].

2.7.2 Transaction Privacy

Transaction privacy is a needed requirement in the blockchain ecosystem in order to ensure that only the legitimated participants can have access to their transaction information (i.e. amount or transaction patterns) and this data is kept obfuscated in the public blockchain. In the last couple of years, cryptographic privacy-preserving tools have been integrated to blockchain protocols such as Non-Interactive Zero-Knowledge (NIZK) Proof and Homomorphic Encryption (HE).

2.7.2.1 Non-Interactive Zero-Knowledge

Zero-Knowledge (ZK) proof is a cryptographic technology that provides a solution to put users in control of their data. This mathematical method proposed in the early 1980s by the author [157] was meant to ensure data sharing without leaking any personal information. The core idea behind this method is to enable one party to prove to a verifier that some assertion is true without revealing that information to the verifier.

Another variant of ZKP is Non-Interactive Zero-Knowledge. This variant allows one to reach computational ZK with no need for certifier or verifier to be included in the process, in case of both to sharing the same reference string [158].

In the blockchain ecosystem, it means that all account information is encrypted and stored in the blockchain. In case of a participant do a transfer to other user, he can effortlessly demonstrate or prove that the transfer can be executed successfully with no need to disclose his account balance. Zero-knowledge Succinct Non-interactive ARgument of Knowledge (zk-SNARK) is an alternative to Zero-Knowledge proof and is employed by Zcash protocol to verify transaction is a private fashion[159]. Moreover, in Zcash the coin value is added in the commitment scheme and zero-knowledge proof so that the value is arbitrary and publicly verifiable. It also employs the new zk-SNARK verifier to enforce an original coin mixing contract, which uses a simplified version of Zerocash, an academic protocol whose implementation is used to build Zcash. The contract enables a participant to store discrete amounts by adding a "serial number" as a commitment into a Merkle tree, which is maintained by the contract. The cost of being one of the most preeminent solutions for privacy comes along with the high computational power needed to generate the transaction proofs.

Transactional privacy is also a major concern among emerging blockchain-based smart contract technologies. Despise being able to generate decentralized applications, which is an enormous advantage, all transactions including the amounts and respective addresses are publicly available on the blockchain. This means that everyone can easily consult the flow of money generated by an institution or individual. This issue was addressed by the author in[160], the first to simultaneously offer transactional privacy and programmability privacy in a decentralized blockchain network. Hank project is inspired by the fundaments of the Zerocash project and the smart contract technology. Essentially, it gives users the privacy needed to share information in an encrypted form with the smart contract. Its foundations rely on the NIZK proofs to ensure the correct execution of the contract and the transactions. In this way, Hank gives the guarantee to the users that the transactions are private, however, the outcome of the smart contract operations is public.

2.7.2.2 Homomorphic Encryption

Homomorphic Encryption (HE) is a type of encryption that enables to compute mathematical operations on encrypted with no need to decrypt the data.

In detail, HE enables users to perform inference and analysis on encrypted data and keep the results encrypted as well. Different from other related technologies like Secure Multiparty Computation it doesn't demand interaction between the data and model owners. The Confidential Transaction (CT), conceptualized by Gregory Maxwell was the first to implement a homomorphic comment scheme at blockchain and achieve transactional privacy. In detail, CT uses commutative prosperity which allows changing the order of binary operations without changing its result output. By this is possible to

obscure the transaction amount between two parties by shielding it from the broad network so that only the tracking parties would know how much had been transferred. It also allows to share of a scanning key used to establish the shared secret used by the rewindable range proofs. This method is fully compatible with watching wallets; In case of an auditory the user can easily generate and share a "viewing key" with auditors or any other entity to have access to the user's transactions.

However, this is considered a Partial Homomorphic Encryption scheme because it is limited to a unique calculous operation (i.e. addition or multiplication) over the encrypted data. This is considered a limitation for the requirements of current Dapps that demands complex types of calculous operations. Here is where Fully Homomorphic Encryption (FHE) can be a "game changer" in Dapps as it enables to perform an infinite number of different sorts of evaluation operations on the encrypted data. Afterwards, several FHE constructions emerged [161][162][163], unfortunately, they are only test implementations and are not for effective use. As discussed in [164] amount of work and time a computer takes to processing FHE schemes is at the moment too long to make them feasible to be applied in current Dapps.

2.7.3 Computational Privacy

While transactional privacy has long been a point of interest for a variety of privacy-focused protocols, there remains the need for methodologies to protect privacy at the computational level. Here, the investigation has been using on Trust Execution Environment (TEE) to solve computational privacy issues [165–167]. TEE offers a fully isolated environment for executing applications in it. This ensures that other unauthorized entities from tampering with data or accessing the state of the application running in it. The Intel Software Guard eXtensions (SGX) is a representative technology that enables software to be allocated inside a TEE [168].

In this regard, Secret Network is the first and only blockchain that provides a platform to develop Dapps with computational privacy [169]. It uses TEEs and other cryptographic schemas to isolate and encrypted the data, preventing the nodes of the network to have access to it. Rather than relying on off-chain solutions to enable data privacy, Secret Network relies on a decentralized network of secure processors. To ensure the security of the network, the node operators are forced to be equipped with specific hardware as well as a specific version of the TEEs. As result, Secret Network is able to guarantee that any actor including the nodes can access the raw data that is being decrypted and processed in the network.

2.8 The rise of Blockchain in Healthcare

In this section, it is discussed relevant research efforts in the field of medical data sharing, via local or cloud service and decentralized protocols. There are several solutions that adopted blockchain technology for EHRs management systems. Since this work focuses on both the security and privacy aspects, specific works with different types of approaches and architectures on security or/and privacy issues are selected for review.

In terms of security, the MedBlock project developed by [170], which proposes a secure system based on blockchain to share electronic records among authorized users. In detail, blockchain is used as an index database to store summaries of patient medical records and hash values of electronic medical records (EMRs). The EMRs are however

sorted in hospitals' databases. Patients can query the blockchain for summaries of their past medical records stored on the chain and use their private keys to encrypt the medical information. On the other hand, hospitals can upload encrypted summaries by signing the data with their private keys. In this operational model, it uses a Certificated Authority (CA) that serves as a system administrator and an authority management agency. It aims to prevent the blockchain network from malicious nodes and hold the private keys of the patients.

Another solution, named MedRec, that utilizes blockchain for EHRs decentralized record management is proposed by [171]. MedRec blockchain is based on the Ethereum blockchain aiming to build smart contracts for the representation of existent patient medical records. In detail, the patient's EHRs are encrypted and stored in form of references in the Ethereum ledger. These references are arranged to provide a navigable breadcrumb trail for These references are arranged to provide a navigable breadcrumb trail for health information. Additionally, medical records are linked to data retrieval instructions or data pointers for use on external databases, as well as viewing permissions. In the same direction, the author in [172] proposes a blockchain privacy-preserving cloud storage of EHRs. A ciphertext policy attribute encryption scheme ensures the confidentiality of the patient data and protects the access policy. A cloud server is used to store both ciphertext and symmetric key cypher text while a consortium blockchain is employed for authentication purposes and access policies.

A different approach proposed by [173] employs the use of smart contracts and an access control mechanism to effectively trace the data generated by users as well as revoke access to violated rules and permissions on data. This access control mechanism resides on access policies stored in the blocks, where the patient or health organizations have different permissions based on the type of service required. In [174] the author implemented a layer one blockchain named Medchain, also targeting the sector of healthcare. Patients' records were uniformized in JSON files to be distributed among the existing network blocks. A framework for decentralized storage of EHRs was seen in [175]. By using the AES256 cryptographic schema, patients' records are encrypted and stored in a decentralized manner by IPFS protocol. The resulting hash from the IPFS is then uploaded into the blockchain.

The noted projects could have successfully built solutions to improve the security of patient health records. One of the main reasons is due to the use of blockchain technology, specifically, the implementation of smart contracts to control the permission level of access to patient data. However, despite the efforts to improve the security of EHR, there are some major issues which were not covered.

One of them is the fact that patients' EHRs are stored in a centralized manner where big entities are responsible to process and store the existent medical records on their infrastructure layer. This type of approach is shared among the works mentioned above and significantly arises the risk of sensitive data being exposed. The assumption that it is worth storing medical data in a centralized manner, goes against the whole purpose of adopting a decentralized blockchain.

Regardless of the authors claiming that the security of the system is achieved by using blockchain for permission level control, off-chain queries can still be performed by malicious actors within these centralized providers. Another issue relies on works where their solutions are based on a consortium blockchain [172,176–178]. The centralized nature of this network makes consortium blockchains vulnerable to corrupted nodes.

In terms of the problem of data privacy, the analysed works have enhanced their privacy features by using off-chain solutions based on cryptographic schemas to encrypt the data. Nevertheless, it increases the complexity of such a system in terms of

computational efficiency and raises concerns over the custodian of the cryptographic keys.

In an attempt to solve the lack of privacy on blockchain, some alternatives have been presented. A decentralized anonymous payment system that explicitly encapsulates the functionality and security assurances of a full-fledged decentralized electronic currency with strong anonymity guarantees is introduced via the blockchain protocol Zcash [179]. In detail, Zcash uses zero-knowledge Succinct Non-interactive ARguments of Knowledge (zk-SNARKs) to provide privacy protection, by hiding user identities, transaction amounts, and account balances from public view [180]. Another project named Monero, a fork of Bytecoin, is a secure, private and untraceable currency built on Cryptonote protocol. Monero makes use of ring signatures, ring confidential transactions (RCT) and stealth addresses to obfuscate transactions at the protocol level [181]. Despite both protocols presenting a solution that brings privacy to the blockchain, they only introduce privacy at a transactional level. Following the use case of our work, privacy only at the transactional level is not enough to guarantee that sensitive data stored in the blockchain is not publicly exposed. Finally, they do not contemplate any solution to the possibility of an event related to the loss or stealing of a private key by a malicious actor. As result, numerous data records from different patients would be potentially leaked or irreversibly lost.

2.9 Summary

In this chapter, a literature review on concepts, theories and state-of-the-art technologies with relevance to this research has been presented. The author of this research started by showing how the fundamental proprieties of AR technology can benefit individuals in everyday life activities. Although the scope of this work is individuals with ASD, it is described the underlying value of this technology across the assistive application domain. This literature review also highlights the existent limitations that might prevent the proliferation of AR-based assistive applications. Essentially this literature covers two main issues, the lack of intelligence of such systems and low data security and privacy practices. Here this review focused on the use of artificial intelligence methods and decentralized network protocols.

Whilst this literature review's main concerns are technologies such as Deep Learning and Blockchain, some challenges in those areas have been highlighted. Firstly, it is described Deep learning technology as a mechanism for extracting meaning out of raw data and its application in the field of computer vision. As this review is more oriented to the field of object detection, the most popular algorithms such as Fast R-CNN, YOLO, SSD and MobileNet are described in detail. This serves as an introduction to the understanding of how object detection algorithms are used for perceiving the user's visual attention and object of the interest. Here, again, is given a detailed review of the state-of-the-art algorithms, which can be sorted into egocentric-based object detection and fixation-based object detection. To complement the understanding of the reader it is explained how eye function metrics such as fixations have been studied as an indicator of user visual attention and how it is coupled to the field of object detection.

As mentioned previously, a significant part of this literature review pays attention to blockchain technology as a means of improving the underlying data security of assistive technology. By this, it is given an introduction to the technology behind it and relevant works that have been done in this area. It is also highlighted the importance of enabling privacy to blockchain-based applications. Here, the author of this work gives an extensive

review the core topics of research in the matter of privacy such as identity privacy, transaction privacy and computational privacy.

Chapter 3

A Framework for Assistive Augmented Reality-based Smartglasses Solution

In this section, a framework for an assistive AR-based smartglasses system is presented. This solution aims to coach individuals with ASD in a wide range of DLA that may improve and enhance their self-sufficiency. Specifically, it solves the issues related to in-person occupational interventions: the constant need for professional supervision during an intervention, the lack of user engagement and the lower effectiveness of common approaches on sequential behaviour activities. The system is built over a framework that integrates different modules that handle the data from the input till the security and privacy of the data storage. Specifically, with this framework, it was deployed head-worn eye tracker cameras and world camera for a better understanding of the user's intentions as well as his context. The head-worn eye tracker aims to explore the user's visual attention and objects of interest discovery. On the other hand, the world camera enables the system to perceive the user environment. The framework also integrates a blockchain module responsible for the security and privacy of patients' EHRs storage. Other modules are responsible for the processing of the environmental sensor input data and integrations of output assistive devices.

It is worth mentioning that despite this framework targeting DLA, it can be also applied to model tasks related to any other activity segment. In addition, it also integrates a web platform that focuses on two critical components. Firstly, enabling the therapist or caregiver to easily model a task suiting the user's needs. This component assumes to be very important because the targeting group of study is characterized by a wide variety of deficits that demands personalized interventions. On the other hand, the web platform also includes a component aimed to deliver statistical data on users' visual attention behaviour, the object of interest as well as feedback on task performance. Here, the goal is to provide an assessment tool allowing therapists to have a better clue of the effectiveness of the interventions.

In this work the contribution to ASD is focused on three major aspects:

a) Multiple instances of assistive AR technologies based on smartglasses.

b) Visual attention analysis tool.

c) Secure and privacy of user's data.

As such, a framework is proposed based on a modular infrastructure approach similar to micro services approaches, in the next subchapter it is depicted how this infrastructure behaves during all the processes, starting from the user data gathering to the generation of AR information on the smartglasses.

3.1 Framework Structure.

The framework design follows a modular approach seizing its extensibility to other service infrastructures. Structure of system modules and their interaction is outlined in Figure 3-1.

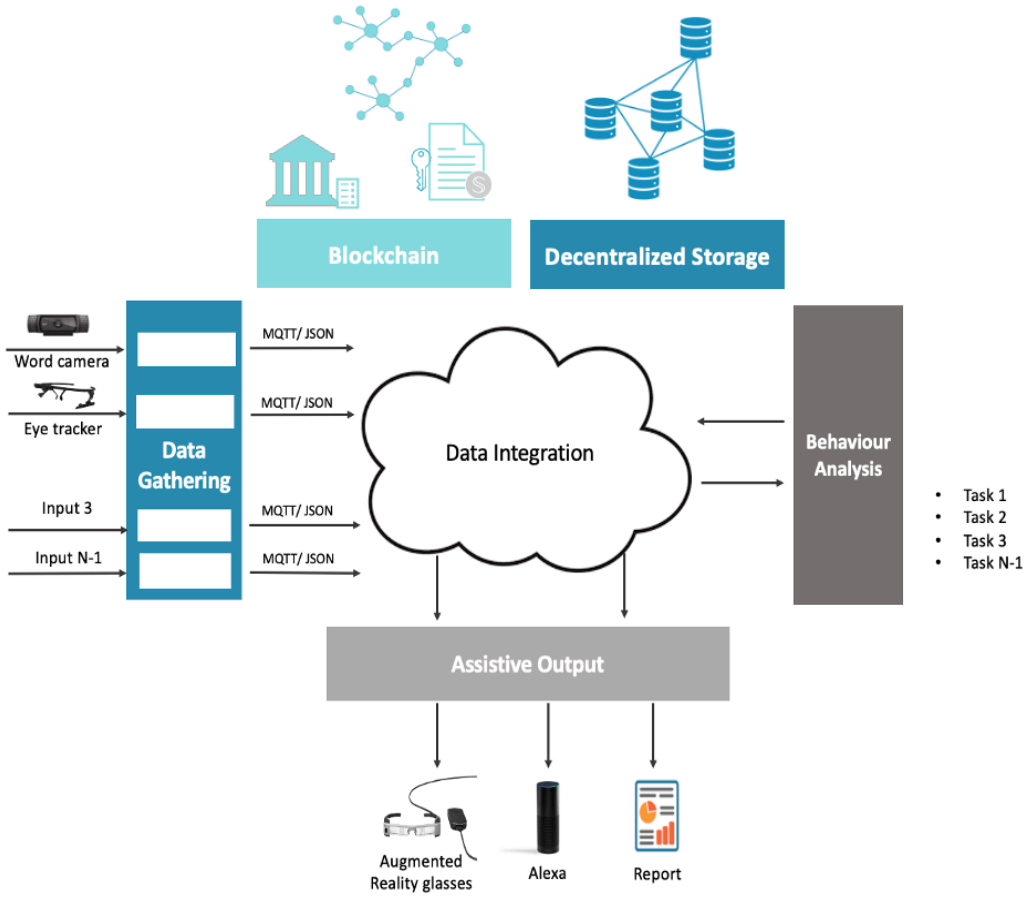

Figure 3-1- System architecture scheme exposing the different modules involved. The arrows represent the flow of information between connected entities.

In the next subsections, the main functions of each module and how they operate will be described in detail.

3.1.1 Data Gathering Module

The Data Gathering module enables a group of devices (i.e. IOT sensors, eye trackers, cell phones, etc.) or APIs to send their data to be managed by the Data Integration module using their own native protocols. In detail, this brings a standard interface to all kinds of devices at the context information management level. Specifically, this module processes the input data (i.e. images, raw temperature data, audio, etc.) with the aim of ensuring that it is converted to the MQTT protocol and then transmitted to the Data Integration module. By this, it is enabled the interoperability between this system and devices that use incompatible protocols.

In the context of this assistive AR-based smartglasses system, the Data Gathering module processes and converts to MQTT protocol all the events from the Pupil Labs binocular camera as well as world camera device. Specifically, it collects data about user's fixations and gaze direction in real-time. This data is pre-processed by the manufacturer [10] At the same time, it also collects image data over the user's context. Here, a Deep learning state of art algorithm is applied for visual attention analysis. Specifically, a pre-trained MobileNet model (trained on ImageNet dataset) is employed for object detection aiming to record all the object part of the user's environment.

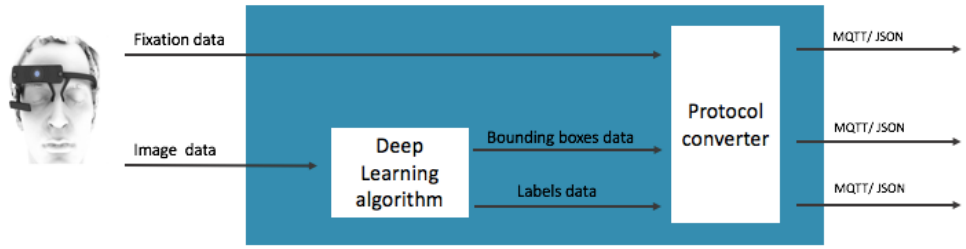

Figure 3-2- An example of how data from input devices flows on the Data Gathering module. Specifically, the fixation data from the eye-tracker flows directly to the protocol converter sub-module as it is already pre-processed. In its turn, the output image data from egocentric cameras need to be pre-processed before reaching the Protocol converter module.

The output are bounding boxes coordinates as well as labels of the object detected. In Figure 3-2, can be seen that all the data is then forwarded to the sub-module Protocol Converted to convert the data to MQTT protocol. Finally, the data is then transmitted to Data Integration module. It is worth to mention that the major advantage that this module offers, is its ability to support very heterogeneous data inputs, meaning that different devices with different protocols can be easily integrated with the system.

3.1.2 Data Integration Module

The Data Integration module is designed over a publish/subscribe architecture that uses a client/server model. This protocol overcomes the challenges of connecting the physical world of sensors, actuators and phones with software processing technologies. Compared with traditional protocols like HTTP, MQTT protocol presents key advantages such as faster response and throughput and low bandwidth [11]. In addition, its modular structure allows data to be discovered by, and delivered to, any number of subscriber clients. Clients can be data systems, applications or devices. In this application scenario, the Data Gathering module works as a publisher to MQTT. In its turn, the AR-based smartglasses works as subscribers. In the case of the Behaviour Analysis module and Blockchain module, and Decentralized Storage module, they are both publisher and subscriber, due to the bidirectional flow of information between them.

3.1.3 Behaviour Analysis Module

The main goal of the system is to deliver AR visual information to assist the user during task performance. In this regard, the system's capabilities to extract the user's context plays an important role to provide effective assistance. It is also considered crucial to deliver mechanisms for therapists or caregivers to model tasks to be executed by the

user. Regarding the extraction of the user's context, this work focus on the following goals:

a) Object detection over the user's surrounding environment, by analysing the user's egocentric images from the world camera.

b) Objects of interest discovery by exploiting a sliding window-based time series approach in conjunction with a Heuristic probabilistic function to analyse user's fixations around a potential object of interest in an egocentric video [182].

In the respect to task modelling, a web platform was developed for this purpose. Here, therapists are provided with tools for modelling a task by defining a set of steps and related actions. For each step within the task, actions are defined and then transmitted to the Data Integration module to be displayed as procedural instructions on the AR-based smartglasses [13]. Meanwhile, the behaviour analysis module analyses the user's context information to verify whether an action was performed or not. When all the actions are executed the user moves to the next step. The web platform also enables therapists to assess user's task performance by delivering a detailed task report with data over objects upon visual attention (i.e. number and period of fixations on objects, number of recalls on objects out of scope of the task, etc.), and statistical data related to execution times over steps and actions (i.e. average, maximums, minimums, etc.). Finally, the behaviour analysis module is built over Node.js, a robust and mature programming language for deploying the backend and front-end platforms. The backend is built a platform in Node.js due to its event-driven nature as well as great advantages in terms of performance, scalability and high interoperability between the front end and backend [14]. Regarding the frontend, it was chosen Angular framework as it is mobile and desktop-ready, meaning one framework for multiple platforms. Moreover, Angular is built on top of JavaScript language which improves the efficiency and interoperability between the backend and front end.

3.1.4 Assistive Output Module

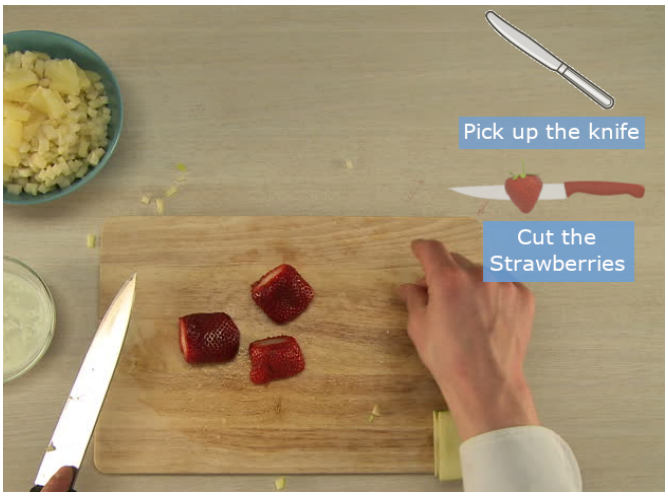

Figure 3-3-AR-based smartglasses user interface during task performance.

The Assistive Output module represents the output devices connected to this system. All the output devices work as subscribing clients of the Data Integration module, where they receive information through MQTT topics that they are subscribed to. In the context of this application scenario, it is defined as two external devices such as AR-based smartglasses and Alexa devices. The AR-based smartglasses device has an Android application aiming for delivering AR procedure instructions to the user during task performance. As can be observed in Figure 3-3, it incorporated the use of pictographic cued instructions because has shown benefits in terms of acquisition, generalization and maintenance of DLA skills [16]. Alexa device is also incorporated into the system, aiming to deliver audio notifications to caregivers in the case of the user struggling to complete a step or action.

3.1.5 Blockchain Module

The Blockchain module is a core component of the framework. It aims to protect access to the users' data from malicious and unauthorized agents. Essentially, it leverages a smart contract that works as a permission management database for authentication purposes. In addition, it also stores users' metadata and a hash that points to the data stored in the Decentralized Storage Module. In detail, access to the applications is managed by the Blockchain Module. If the access is granted, the data generated and processed by the Behaviour Module is then forwarded to the Decentralized Storage Module in order to be stored. In its turn, it returns a hash that is stored in the Blockchain Module along with other users' metadata. Here, data privacy is achieved thanks to the specificity of the Secret Network protocol that has the particularity of providing data privacy by default at both transactional and computational levels. Moreover, due to the nature of the blockchain, a peer-to-peer network, it preserves the immutability of records and ledger entries, making this electronic database highly secure. Finally, the governing body is introduced in order to solve the issue of stolen or loss of the private key. This centralized authority can be seen as a system admin that has the power of recovering users' access.

3.1.6 Decentralized Storage Module

The decentralized Storage Module is responsible to store user data in a decentralized fashion. It offers multiple benefits, like offering cost-efficient storage and securing sensitive data by not storing it with one centralized provider. This module relies on IPFS protocol that utilizes a peer-to-peer network of operators that keeps the data secure during storage and transfers. By this, it exponentially decreases the possibility of data breaches and provides data privacy as the data is distributed into multiple encryptions that are spread throughout several nodes.

In the context of this framework, the Decentralized Storage Module receives data that is processed by the Behaviour Module and stores it in the IPFS protocol by dividing it into multiple chunks as "blocks". The result is a unique cryptographic hash that identifies each file, which is computably derivate from the original content file. The hash is then forwarded to the blockchain module to store it along with other metadata belonging to the user.

3.2 An Example of Use Scenario.

As a generic case study to evaluate the proposed solution, it targets the self-management skills of children with ASD in the context of DLA. Specifically, it is shown how both therapist and user can benefit from this system in a meal preparation interaction. In particular, the focus relies on providing support in three different scenarios. One regards to situations where the therapist needs to model a task to be executed by the user. Another situation, regards the use of AR-based smartglasses to provide assistance to the user in DLA. Finally, it presents a detailed statistical report enabling therapists to assess user performance.

3.2.1 Task Modelling

In this subsection, it is demonstrated the system capabilities to provide therapists the tools needed for modelling a task in a high level of abstraction. In detail, it is chosen a fruit salad preparation task which is a very common exercise in occupational interventions. It was designed an intuitive interface for easily defining specific sequences of steps as well as actions to be performed by the user in each step. Each action involves an interaction between the user and a specific object (i.e. knife, bread, toaster, etc.). Here, the therapist can define the following actions:

a) Fixation action- This type of action is used by therapists to ensure that the user locates a specific object or is focusing his attention on it. Firstly, the therapist defines the object to be located, then when the system detect fixation on the target object, the action is considered completed. In the context of this scenario, therapists might define fixation actions to ensure that the user has located all the ingredients needed to perform the task.

b) General action- The general action includes generic actions to be performed by the user like cut, grab, wash, open, close, etc. To complement those actions, the system presents a list of objects to be selected (i.e. action: grab, object: knife).

c) Interaction action- The interaction action aims to detect the interaction between two objects. The system perceives an interaction between objects when the bounding boxes of one object overlap the other one. For example, placing fruit into the cup.

d) Time-out action- This action aims to alert therapists that the user is facing difficulties over a specific step of the task. In this action, a period of time is defined to perform a step. In case of the user exceeds the defined time-out period, an audio notification is reproduced in Alexa.

Moreover, for each action, there is also an input field aiming to specify a guideline as well as an image to be displayed in the augmented reality glasses. If the action is performed on time, a green check icon is displayed next to the corresponding guideline and image, otherwise, a red check icon is displayed.

3.2.2 AR-Based Smartglasses

The AR-based smartglasses aim to assist children with ASD in living activities. When children use smartglasses they can still see the world around them. This is fundamentally better than other AR-based platforms such as mobile and tablets because with those, the children are constantly looking down and locked in another world. The AR-based smartglasses keep the children heads-up, hands-free, and engaged with their surrounding environment. As outlined in Figure 3-3, the smartglasses display both pictograph and procedural instructions previously defined by the therapist. Meanwhile, the system automatically monitors user performance. When the children accomplish a step, a green check icon is exposed over the pictograph. This feature is designed to promote the children's engagement as well as help them on recalling the next action to be performed. The location of visual cues and procedural instructions were strategically disposed on the right side of the AR interface to not disturb the user during task execution. This is supported by the fact that human visual attention is mostly focused on the centre region of the visual field. Finally, AR-based smartglasses can be used anytime and anywhere which is considered a major advantage when compared to occupational therapies that provide interventions with low frequency.

3.2.3 Accessibility and Data Sharing

Blockchain technology was employed as an alternative to centralize systems to promote transparency and trust with participants of the platform being able to have relatively unfettered access to their medical records securely and privately. Also, record sharing like statistical reports or medical records from a patient to the other party such as a doctor from another hospital is performed in a secure and transparent fashion. After a successful authentication to the platform, a user can either decide to grant or revoke access to his medical records by adding the public key of the other party to an access list. This access list is stored on the blockchain along with the user metadata. By this, every party on the access list can have access to the user’s medical records. The user, in his turn, can also revoke access by removing the public key of the external party. To increase the transparency of the whole process, all the records related to authentications and access to medical records are stored in the blockchain. The user can verify those records and have total control of this account.

In an event of lost or stolen of the user's private key, he can contact the governing body and claim a new public and private key to continue to have access to the platform and his medical records.

3.3 Statistical Report

The measurement of user behaviour in various environmental contexts is an ongoing process that plays an important role for therapists. Assessment can be said to drive the design and modification of the task model. Moreover, analysing the user performance on regular basis enables the therapist to have a clear assessment of their interventions as well as the progress made by the user. In this regard, this system collects data on every response, all timings and even measures of fixations metrics related to a set of objects. In addition, it collects information regarding the number of visual recalls that occurred on each of the objects detected. This data is extremely useful to detect for instance irrelevant

objects (objects outside of the scope of the task) that are capturing the user's attention. Due to the extreme sensitivity of the data that is collected, the access records to the data itself are saved into the blockchain. To preserve data privacy a combination of cryptographic schemas is also employed.

3.4 Discussion

This section presents the performance evaluation of this framework. The goal is to measure the execution time of each of the modules and discover if the framework can perform in runtime. For this test, it was used a laptop with the following specifications - 2.2GHz dual-core Intel Core i7 (Turbo Boost up to 3.2GHz) with 4MB shared L3 cache. Regarding the experiment, it was evaluated two different scenarios. First, it is tested the whole process that involves the storage of data in the blockchain. Specifically, input images and pupillometric data regarding objects under user visual attention and data of user performance on a specific task.

Here, it is measured the execution time of the different modules involved, starting from the input image on the data gathering module, passing by the data Integration module, following the behaviour analyses module and finalizing with the storage of the data on the Blockchain module. The other scenario concerns the goal of retrieving information from the Assistive Output module. In this case, the assistive cues are redirected to the AR-based smartglasses. This test is very similar to the first scenario, however, the Blockchain module execution time is excluded because the flow of information does not pass through it.

Based on the results shown in Table *3-1*, it can be perceived that there is a huge difference in the total execution time when comparing both scenarios. In the case of the Data storage scenario, the execution time result is 0.719 s, very different from the Assistive cues scenario with only 0,10 s. This difference is due to the Blockchain module that presents the slowest execution time, 0.62s. This time is justified due to the nature of the chosen blockchain protocol that due to its privacy features proves to be very costly in terms of execution time. The Behavior Analysis module and Data Integration module take quite similar times in both scenarios. In the case of the Behavior Analysis, it takes 0.067 s for data storage and 0.071 s for assistive cues. The Data Integration module presents times of 0.032 s for data storage and 0.029 s for assistive cues.

Those results show that this framework meets the timing requirements for building real-assistive applications in a secure and private manner. It is worth mentioning that despise framework has achieved good results, the margin to be improved is considerable in case of the system runs over better hardware specifications. Also, the framework was tested against a hard demanding condition, usually, assistive solutions do not require much of an amount of data to be processed in terms of size and quantity.

Table 3-1- Execution time of the different framework modules.

	Behaviour Analysis module (s)	Data Integration module (s)	Blockchain module (s)	Total (s)
Data storage	0.067	0.032	0.62	0,719
Assistive cues	0.071	0.029	-	0,10

3.5 Summary

In this chapter is presented a framework for developing assistive AR-based smartglasses systems targeting children and adults with ASD on improving their DLA skills. Specifically, it introduced a real-time assistive framework to deliver AR visual cues instruction based on the user behaviour during task performance. It is also presented a web platform that enables caregivers and therapists to model DLA tasks by defining specific sets of steps and actions. The conducted literature review identified some limitations of existing solutions such as the inability of the current AR-based framework to perceive both user intentions and context as well as the low level of data security and privacy.

The first limitation was addressed by integrating a module that is able to discover objects of interest by combining object detection algorithms, egocentric images and user-daze data. In order to improve the level of security and privacy, it is offered a decentralized storage system based on blockchain technology with a combination of cryptographic schemas to achieve data privacy.

To validate this work, the different modules that constitute this framework were tested in terms of their execution times. The results show that the proposed framework can meet the requirements for building real-assistive applications in a secure and private manner.

Chapter 4

Visual Attention-Based Object Detection in Cluttered Environments

The importance of detecting user objects of interest in real-time is critical to provide accurate cues about the user's intentions. This kind of information can have a major impact on assistive technologies for providing the right assistance. However, current methods for visual attention extraction and object detection suffer from low performance when moving to an ongoing condition. The inherent complexity of cluttered environments is considered the major barrier to achieving good performance. This is justified by the dynamic of these environments where objects are constantly overlapped, and partial object occlusions also frequently occur. In this section, it is in presented a novel methodology to discover objects of interest based on the user's visual attention. This methodology utilizes a combination of sensors that includes head-worn eye tracker cameras and egocentric video. The data of these sensors is then passed through a heuristic function that is able to generate probability estimation of visual attention over objects within an egocentric video.

4.1 The Proposed Methodology

In this work, it is proposed to use information from the image, fixation location and gaze location for object detection in runtime. The features extracted from each domain along with the proposed heuristic function are described below. A schema diagram of the proposed methodology is shown in Figure 4-1.

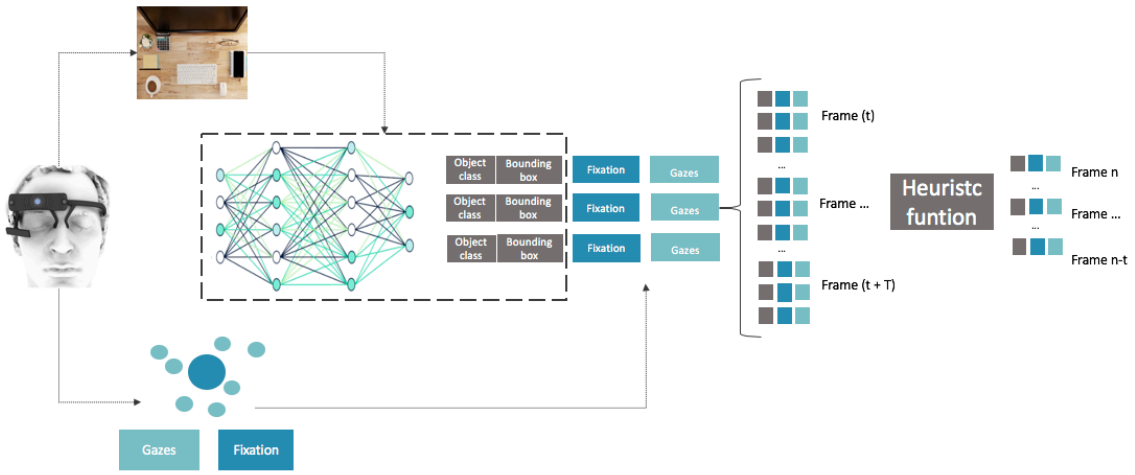

Figure 4-1-Detailed scheme exposing the different modules involved in the object of interest detection task.

4.1.1 Fixation Guided Object Recognition

The major advantage of this object detection system compared to approaches based only on egocentric cameras is that in addition to images it is also obtained useful information about the user's visual gaze provided by the head-worn eye-tracker. Common practices to discover an object upon user visual attention needs to perform image analysis to locate where the object of interest is. Thus, for instance, when an image is captured in a highly cluttered environment it becomes quite hard to obtain good performance on detection tasks. Unlike such a system, it is possible to take the advantage of having fixation data to infer the location of an object of interest and attenuate the noise of objects with close spatial distance.

A typical gaze-based object detection system extracts only a region of the image centered on a fixation point, then neural network models use this region for feature extraction and object detection. Nevertheless, this approach is not designed to tackle the challenges inherent to cluttered environments. Firstly, because objects have different sizes and that can also vary with different camera perspectives, meaning that the region extracted might be too small to capture sufficient features of an object. On the other hand, in case of the size of the fixation region is too large it is likely to capture features related to other objects as well. Finally, for instance, in a situation where the fixation point takes place on two overlapped objects or between borders of two objects, it would not be capable to distinguish which object has the user's attention upon.

Therefore, a completed image is used for feature extraction in order to not lose any relevant information. In detail, in computing the total image from the egocentric video with the resolution of 1280x720 pixels at 30 FPS. Regarding the technique, it used convolutional neural networks due to their robustness and high representation power. A pre-trained MobileNet model, trained on ImageNet dataset, is employed for this purpose (the architecture of this model can be consulted in chapter 2.4.3.1.4. The basis of this choice was its efficient trade-off between latency and accuracy.

As described in Figure 4-1 once a fixation is obtained, it is matched to the frame image most closely in time. Then, for each of the resulting bounding boxes of the image, the corresponding gazes' points that occurred during the fixation duration period are counted. The decision of representing gaze points instead of just fixation points in a frame, it is due to the nature of cluttered environments. A fixation point is located at the center of the gazes' region, meaning that in cluttered situations these gazes can match multiple bounding boxes and not only one like if it is considered to use only a fixation point. By this, it is possible to achieve a better representation of user visual attention during a fixation event.

4.1.2 Time Series Sliding Window Approach

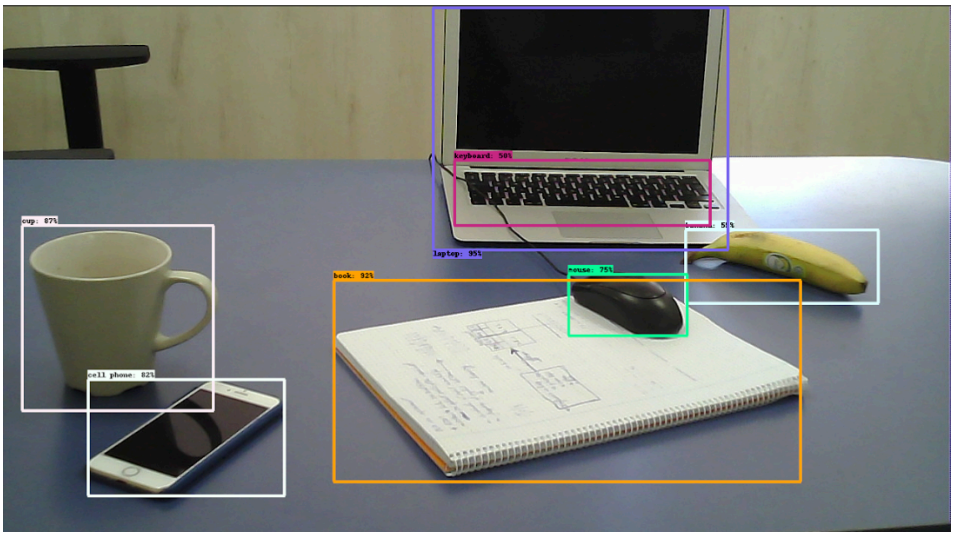

Figure 4-2-Cluttered environments as characterized by the variability of

As can be observed in Figure 4-2, in cluttered environments the boundaries of bounding boxes between objects are often overlapped or very close to each other. In that sense, the task of detecting objects of interest become very hard because if a fixation occurs over an overlapping region, it is needed to decide in which object does it belongs to. Also, minimal dispersion of user's gaze direction due to the subconscious visual attention effect, can provoke a fixation in a region out of the user's visual attention. As result, in a cluttered environment it is possible to observe a significant number of false positive errors.

For those reasons, it is critical to analyse not only each fixation point but all their gaze points in a time series sliding window before and after a fixation event. By this, a detailed overall picture of the user's visual action can be obtained and determine in which object was the user visual attention upon.

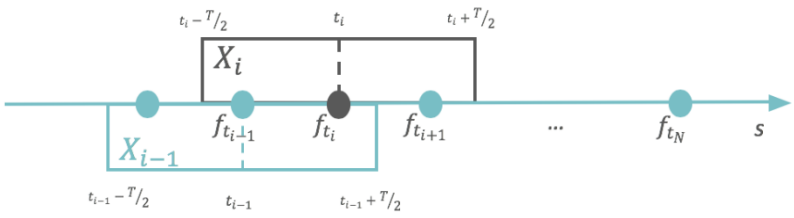

Figure 4-3-Example of the time series sliding window approach.

To solve the aforementioned problem, a sliding window time series approach and a Heuristic probabilistic function are used. Giving an array of N frames, it is defined a vector $f =< f(t_0), f(t_1)..., f(t_N) >$ where $f(t_i)$ corresponds to the frame captured in timestamp t_i. Then, as can be observed in the Figure 4-3, for each $f(t_i)$, it is defined a sequence of $f(t_i)$ as a vector $X_i =< \cdots, f_k, .. >$, $t_k = [t_i - T2, t_i + T2)$ during the period of T. In which i is the number of each frame and T corresponds to 500

milliseconds. Then, it is computed the total gazes occurred in sequence as $R = \sum_{k=0}(X_k)$.

4.1.3 Probabilistic Heuristic Function

At this time, for each element f of the vector X it is computed the percentage of the interception area between element f and the following elements of X as and $P(A \cap I) = \{I \in A \wedge I \in B\}$ for each, A corresponds the target object area, I to the interception area between both objects and B to the area of the secondary object.

After this step, it is initiated a process that aims to assign the gazes points of overlapped areas to objects by the following rules:

a) In the case of the $P(A \cap I) \geq T$ and $P(B \cap I) \geq T$: it is considered that all the gazes points in the interception area belong to the object with the smallest area.
b) In the case of the $P(A \cap I) \geq T$ and $P(A \cap I) < T$: it is determined that all the gazes points in the interception belong to Object A.
c) In case of the $P(A \cap I) < T$ and $P(B \cap I) < T$: the gazes' points of the interception area are assigned to both objects.

It is defined T as threshold $T = 70$. Since within a sequence X it is possible to find duplicated object classes a new vector O is created, with unique object class and their corresponding gazes within a sequence X as $O =< o_{(i_0)}, o_{(i_1),\dots,} o_{(i_N)} >, o \in X$. Then, it is calculated the percentage of gazes bellowing each object class o during a sequence x x as $i \in O. P(o_i) = \frac{o_i}{R}, O \in X$ with O and R representing respectively the total number of gazes during a sequence. Finally, for each element of the vector o it is compared the percentage of gazes with a predefined threshold $G = 50$. The result of this comparison is the following:

a) In case $P(o_i) \geq G$: it is defined that in this particular frame of the sequence, the user visual attention was focused on the object $P(o_i)$.
b) In case $P(o_i) < G$: it is concluded that there is not enough confidence to attribute user visual attention to a specific object, so it is defined as *Null.*

It is also worth mentioning that in case of more than one element in the vector O shows a percentage greater than G, it is defined as object class upon user visual attentional *Null*.

4.2 Experiments and Results

To evaluate the effectiveness of this approach and given the lack of an appropriate dataset, it was designed a new visual attention study to collect visual data over a typical cluttered environment. The experiment aimed to evaluate the effectiveness of a new visual attention approach for object detection in cluttered environments. To collect visual data, a new study was designed with six different objects arranged in groups of two, intentionally overlapped in different regions and magnitudes of space. A group of eight participants wore head-mounted cameras and eye-tracking devices to capture first-person videos and collect gaze data as a group. The collected data set is the first to use point-of-gaze sources in video vision tasks targeting cluttered environments. The experiment

aimed to explore the impact of subconscious visual attention and obtain ground truth labels of the time intervals when a participant looked at a specific object. The developed methodology was compared to two baseline methods for visual attention object detection, Central Biases Object Detection and Fixation Based Object Detection. The results showed that the developed approach outperformed the baseline methods and demonstrated its effectiveness in cluttered environments with great variability of overlapping conditions.

4.2.1 Data Collection

In contrast to previous approaches that strategically well-spaced-out objects [121], the goal in this work was to simulate a cluttered environment characterized by overlapped objects of different sizes and disposed in different positions. In that sense, it was set as an experiment composed of six different objects and arranged into groups of two objects. For each group, objects were intentionally disposed in a way that they got overlapped in different regions and different magnitudes of space. The aim was to test the developed approach in a scenario with great variability of overlapping conditions. The experiment settings are shown in Figure 4-2. Each participant was equipped with a head-worn camera and eye trackers to record first-person videos and gaze data collectively. To the best of this research knowledge, this data set is the first to use point-of-gaze sources in video vision tasks targeting cluttered environments.

During each recording, eight participants were asked to focus their attention upon various objects such as a banana, book, laptop, mouse, mobile and cup in the same manner as they do in their lives, instead of forcing them to fixate on the centre of the objects. In addition, it was asked to them to pay attention to the contours of each object. In this way, it is possible to explore the impact of the subconscious visual attention effect as well as make sure that some of the gaze points reach the overlapped regions.

Here, it was used the Pupil Lab eye trackers to record HD resolution first-person video at 30 FPS with points-gaze-data at 120Hz. Parameters of fixation detection were left at their defaults where fixation duration was between 100-200 milliseconds. Eye trackers were calibrated before each recording session and validated by the eye tracker accuracy. As soon as the participants started to pay attention to an object, they pressed a key. The same procedure was repeated when they finished the contour of object visualization. This aimed to obtain ground truth labels of the time intervals where a participant looked at a specific object. In detail, each frame that a fixation occurred was annotated with a label based on the object that the user was looking at that specific time interval.

4.2.2 Experimental results

In this section, it is presented the evaluation of the conducted experiments. Here, was calculated the precision based on the results obtained from each methodology and the ground truth labels. Firstly, it is shown a comparison of the developed method against some baseline methods for visual attention object detection. Here, the following two methods for baseline were implemented:

a) **Central Biases Object Detection:**

The authors [124] introduced the Central biases object detection methodology which uses a central fixed region in an image to extract object features for object detection

algorithms. In the experiment, a central region was defined for each captured frame and used the Convolutional Neural networks (CNN) object detection algorithm to find an object of interest.

b) **Fixation Based Object Detection:**

To provide evidence of the effectiveness of the developed methodology it was implemented a simplified version of it that is also used by the authors in [117][122]. This methodology combines fixations with CNN object detection algorithms to determine the object of interest. In the experiment, fixations were manually configured as a unique parameter for object detection, excluding the time series gaze analysis along with the heuristic probabilistic function.

4.2.3 Evaluation of the Research Method by Participant and Objects

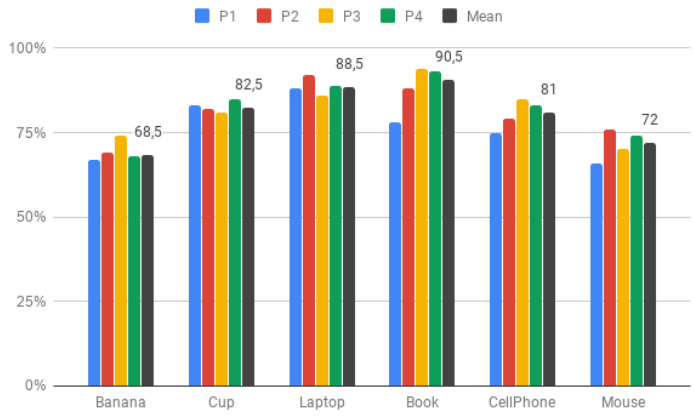

Figure 4-4-Results of real-time simulation for each object class. Although precision drops significantly in both classes banana and mouse, it remains at an acceptable level.

Figure 4-4 shows the results of the developed methodology in terms of accuracy regarding each object class and participants. It can be observed that the developed method in overall performed better on the objects class laptop and book with 88,5% and 90,5% respectively. This significant improvement can be explained by the fact that those objects had major bonding box areas over the rest of the objects. So, the majority of user visual gaze tends to be situated outside of overlapped regions. In the opposite direction, can be perceived a decrease in the average accuracy among the eight participants on the object's banana and mouse with 68,5% and 72% respectively. One of the reasons that can justify these poor results is the lack of precision of the eye trackers. Small deviations of user gazes can have a negative impact when dealing with reduced bounding box areas. Even so, this method shows to be robust against this type of adversity.

4.2.4 Evaluation of key parameters on the performance

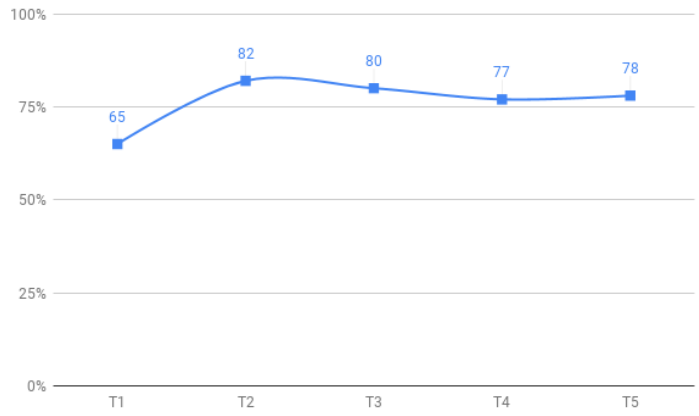

Figure 4-5-Results of the influence of the varying the values of threshold T in the performance.

In addition to the previous results, this section shows the influence of different thresholds on the overall accuracy. It was tested with different parameters the T threshold of the heuristic function that is responsible to decide within a time window the object of interest based on its percentage of user gazes. Figure 4-5 shows the average accuracy of this method by changing $T1 = 40, T2 = 50, T3 = 60, T4 = 70, T5 = 80$. It can be observed that $T1$ threshold achieves the higher accuracy with 84%. On the other, when decreasing this optimal parameter, the accuracy drops drastically as can be seen by $T1$.

4.2.5 Evaluation of Methods for Object Detection Under Visual Attention

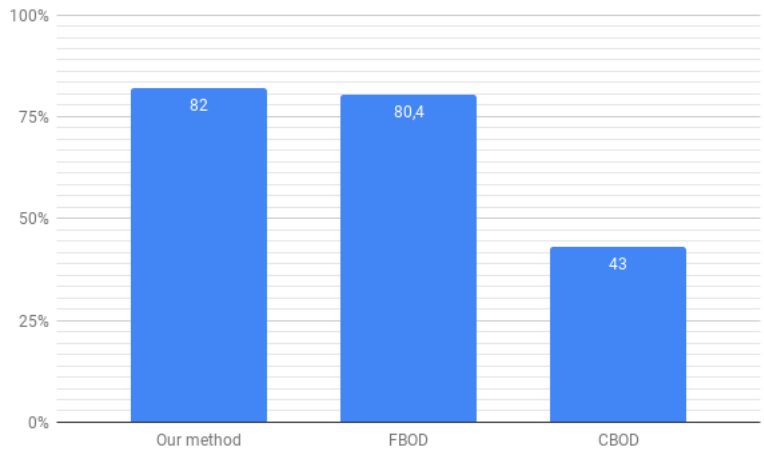

Figure 4-6-Comparison of overall accuracy between baseline methods.

In this subsection, it is shown the results of this experiment in real-time processing. All the detection methods were compared in the same experiment setting conditions. The average accuracy of each method was computed for all the participants. Based on the results in Figure 4-6, it was concluded that the developed method outperforms both centre biases object detection (CBOD) and fixation-based object detection (FBOD) methodologies. The developed methodology achieved a mean average of 82%, outperforming by 2% the FBOD method and by 39% the CBOD method. These results indicate that in cluttered environment the developed methodology performs significantly better than CBOD and slightly better than FBOD.

4.3 Discussion

This study points out an important but overlooked issue of visual attention-based object detection in mobile settings. Since most of the existent approaches are applied in offline conditions, it started by addressing the challenge of performing it in real-time settings. In this regard, it was demonstrated that the salience object detection approach is not suitable for this type of setting, instead, it was presented an approach based on head-worn eye tracker cameras and egocentric video. The change from offline to real-time settings introduces one of the major challenges in this field of study, which is the lack of performance in detecting objects of interest in cluttered environments. This study was the first to address this challenge by exploiting visual gaze data analysis with CNN object detection algorithms. It provides the first mobile data set of visual attention in cluttered environment settings. In addition, it was also conducted an in-depth evaluation of the developed methodology against the existing widely used methodologies in mobile settings.

It is encouraging to see that the developed methodology can perform well in cluttered environments. It significantly outperforms the centre biases-based object detection methodology. As was expected, this algorithm also slightly outperforms fixation-based object detection methodology thanks to the developed time series sliding window approach and heuristic probabilistic function. Experimental results also reveal that it is very robust against extremely overlapping regions and small objects, showing a small difference in its accuracy compared to objects with a reduced overlapping area and large size.

Given the technological advance of head-worn eye tracking and emerging interest in mobile computing, it is believed that this work can open numerous opportunities for assistive technology studies as well as follow-up visual behaviour research.

Assuming that the goal of this work is to study the detection of the object upon the user's visual attention in a cluttered environment, this experiment was performed with the participants in controlled settings. In future work, will be evaluated the developed approach on a novel data set covering free-living settings.

4.4 Summary

This work introduced a new method for discovering objects upon user visual attention. It tackles one of the major challenges in this field which is the poor performance of existent methods in both real-time settings and cluttered environments.

To address this problem, it was presented a novel method that combines object detection algorithms with egocentric video and user gaze analysis. This research method uses a heuristic probabilistic function as well as time series windows approach to analyse user gazes' data around objects in egocentric images.

The results show that in cluttered environments this research method outperforms commonly used methods for the detection of objects of interest in real-time.

To evaluate this method was conducted an experiment and constructed an indoor dataset with data annotated of eight participants during visual attention tasks in a cluttered scenario. It is considered that this dataset fulfils all requirements needed to evaluate this method. In the next chapter will be introduced a methodology to increase the security and privacy of user data storage in a decentralized fashion.

Chapter 5

Decentralized Data storage based on Blockchain

The healthcare industry heavily relies on electronic health records (EHRs) from assistive technologies for disease prevention and treatment. However, storing this sensitive data in the cloud or local healthcare infrastructures has made it vulnerable to cyberattacks, raising concerns over privacy and security. To address these challenges, a decentralized cloud system based on blockchain technology has been proposed in this section. This system not only minimizes the impact of malicious attacks but also mitigates some of the major challenges of storing information on public blockchains, such as expensive costs and lack of privacy. By using blockchain, assistive technology data can be stored and shared in a secure and transparent manner, allowing for better collaboration between patients, healthcare providers, and technology developers. This can lead to more accessible and reliable assistive technologies that significantly improve the quality of life for people with disabilities while maintaining their privacy and control over their information.

5.1 Proposed system-Decentralized data storage and Public Key Governance IPFS

In the proposed system, it was envisioned a solution that brings the possibility to store health data records in a decentralised and secure manner. Inevitably, blockchain technology appeared as a primary solution due to its distributed decentralized network that enables tamper-resilience and practical immutability for stored data. However, the blockchain is not designed for the storage of large files.

Ethereum's Yellow Paper states that it costs 20 000 gas to store one 256-bit word [183]. If taken as an example an insurance policy document that weighs 1 Mb, would mean at the time of writing a cost of 747.52 Dollars.

This comes from the fact that all the files are appended to the blockchain network, which needs to be executed and stored at every mining or verifying node. Since it is too expensive to store non-transaction data, such as medical images, contracts, and personal information in the actual blockchain ledger, some form of off-chain or sideDB storage is required. Here, a hash or signature for the off-chain item is generated and stored in the blockchain ledger.

In addition, it integrates a second off-chain solution to provide an extra layer of privacy by encrypting the EHR to be stored in the smart contract.

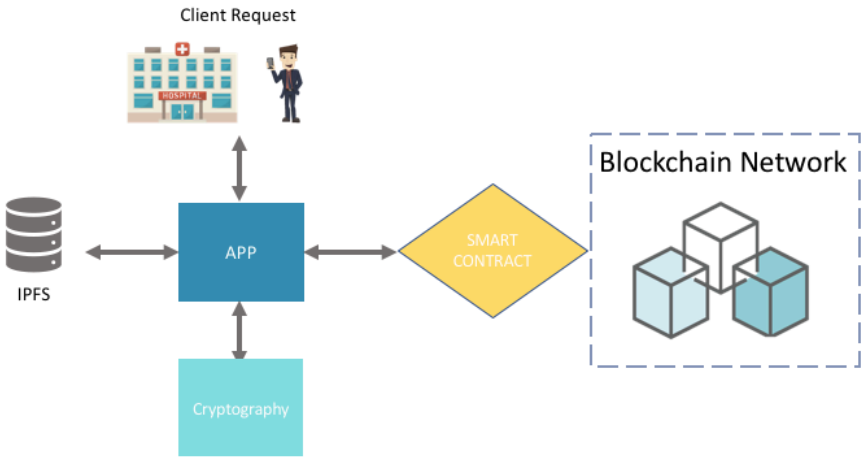

Figure 5-1- Detailed scheme exposing the different modules involved in the decentralized storage system.

This section, as outlined in Figure 5-1, it is presented details describing the different concepts and modules underlying the Dapp for the storage of eHealth records. Firstly, it introduced the smart contract based on the Ethereum blockchain. Thereafter, will be described the integration of the off-chain solution with the smart contract.

5.1.1 Smart contract

The central part of this system architecture is given by the smart contract module. Essentially, it was developed a smart contract based on the Ethereum blockchain that operates in two different modes. Firstly, as a permission management authentication system, where the smart contract maintains a list of authorized entities for access rights to upload and query existent EHRs. Ultimately, it works as a database, ensuring that the patient's EHR is stored in a secure fashion thanks to the Ethereum blockchain nature.

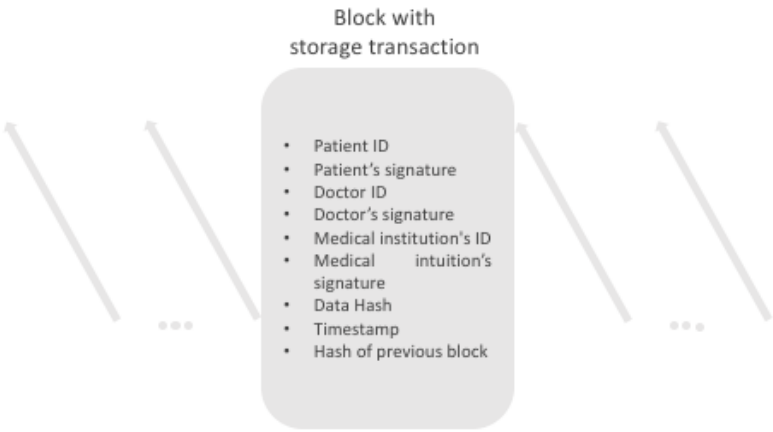

Figure 5-2-Mainchain storage.

As mentioned above, the smart contract supports types of functions to be executed such as transactions and queries. Transactions are created by healthcare professionals from healthcare providers. This type of transaction is mainly designed to update the patient's EHRs. It contains the following characteristics (see Figure 5-2): the patient ID (public key, unique for each transaction), the patient's digital signature for proving that

the transaction has been accepted by the patient, the doctor's ID, the doctor's signature, the medical institution's ID, the medical institution's signature, the timestamp, the data type, the data hash and the hash of the previous block.

Let, H represent the healthcare provider where P is the total patients in the network and R is the total of patient health records. After a patient being consulted within a medical institution or healthcare facility, the system authenticates his identity. For authentication purposes, the smart contract has a mapping between the healthcare provider h and authorized the patients based on their public ID $h, h \in \{1, \dots, P\}$. Then the healthcare provider that generated a new EHR is responsible for creating a new transaction in order to store the EHR in the blockchain. Here, the healthcare provider broadcast the transaction to all the other nodes. If the transaction passes the requirements, which include validating all the intervenient signatures, as well as authenticating the sending node, the new block is appended to the blockchain.

Also, each patient $p, p \in \{1, \dots, P\}$ can query his health records r, r $\in \{1, \dots, P\}$ performed by a specific health care provider $h, h \in \{1..R\}$.

In its turn, a Healthcare provider h, can query and update the records of each patient p.

5.1.2 Off-chain data storage

IPFS protocol was the solution off-chosen for this project. Specifically, IPFS is peer-to-peer hypermedia protocol designed to connect all nodes (computing devices) with the same system of files. In addition to that, IPFS protocol includes valuable features such as moving the identification of files from location-based addressing to content-based addressing.

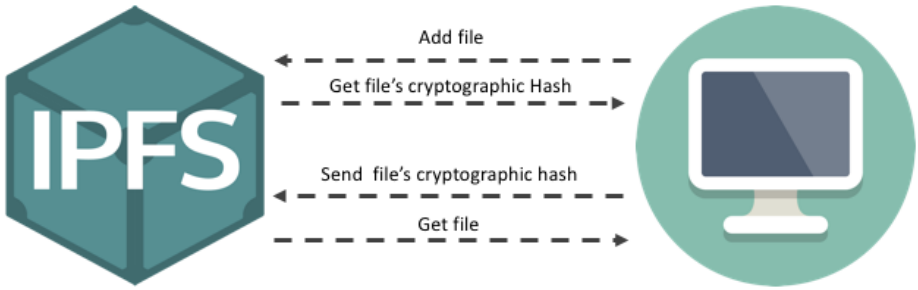

Figure 5-3- IPFS off-chain solution. The arrows represent the flow of information between IPFS protocol and the developed system.

Consequently, each file is identified by a unique cryptographic hash, which is computably derivate from the original content file.

As can be observed in Figure 5-3, the system executes an add function to add a file to the IPFS protocol that returns the corresponding cryptographic hash. To obtain the original file, a cat function is executed passing by parameter the file's cryptographic hash.

Another advantage of using this protocol is the fact that using hashes for addresses can eliminate the duplication issue. In an event of the same file being published multiple times on IPFS, it will only be created once making the network very efficient.

In detail, IPFS make use of acyclic graphs (DAG), basically, it stores files in IPFS Objects of 256kb. In case of a file exceeds the size of an IPFS Object, it is split into blocks. Meaning that these different parts of the files can be stored in multiple nodes. The

mapping of all the blocks of the file is perpetuated by an empty IPFS Object that links them. By this is possible to reconstruct the original file from the different blocks.

5.1.3 Off-chain encryption model

This research includes the development of a smart contract based on the Ethereum blockchain for the security and transparency of data storage. However, public blockchains like Ethereum have a critical privacy issue since all the data stored is public. In that sense, the storage of the IPFS hashes itself is not secure because it represents a unique fingerprint belonging to specific data. In case of some bad actor gets access to the hash, automatically he gets access to the original file stored in the IPFS. For this reason, in this work, data privacy is achieved by applying an encryption scheme to the hashes of uploaded data on IPFS. Thereafter and as it is shown in Figure 5-4, these hashes are encrypted by using the Shamir secret sharing (SSS)[184] scheme, which divides the hash into n number of encrypted shares. The encrypted shares are stored in the smart contract.

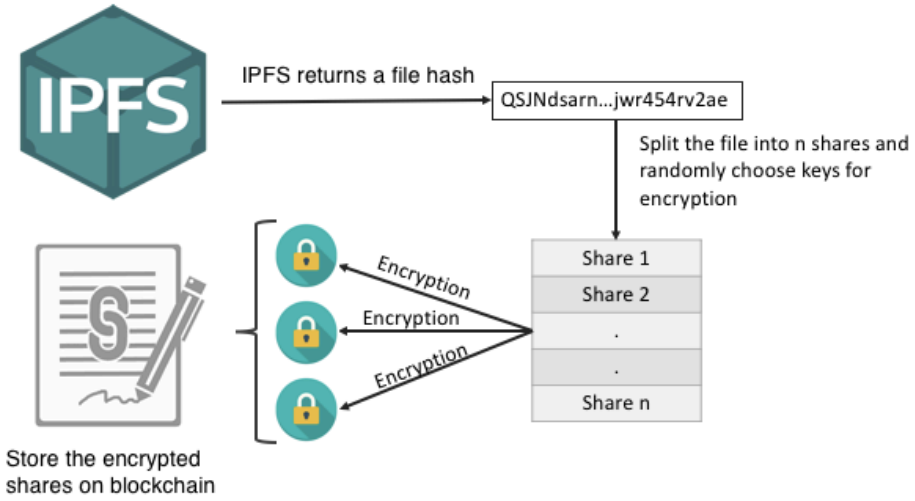

Figure 5-4- Shamir secret sharing algorithm integrated into this system.

Consider a secret, S, which is to be divided into n pieces so that each piece of data can be represented as $S_1, \ldots S_n$.

1. Combination of k or greater number of S_i makes it easy to compute original file

2. In another case, if there is k-1 or a smaller number of S_i then it would be difficult to reconstruct the actual data, therefore S becomes undetermined. It is known to be $an\ (k, n)$ threshold scheme, where $k = n$ means that all the involved entities need to reconstruct the actual secret S [184]. In the context of this research, split size and number of shares are taken as 5 and 3, respectively. When shares are generated using SSS, these shares are encrypted and stored in a smart contract. Regarding the keys, it has chosen to distribute 3 keys, where one is stored in the smart contract and the other two are distributed between the health care provider and the patient. Since the algorithm requires the majority of keys for decryption, in this way both patient and health care providers can have easy and secure access to the EHR as they do not need keys from each other to obtain the majority of distributed keys.

In order to add, remove or consult patient files, the requester must also be part of a list of entities with granted access for that purpose. This list is stored in the smart contract and granted entities are approved by the owner of the smart contract.

5.2 System model

In this subsection, it is first delineated the workflow of the whole system which is represented by a process model in Figure 5-5. The system model has the following entities:

Healthcare institutions: It might be a health organization that owns patient health data records to be stored in the blockchain. This entity can also query patient data records.

Patient: The patient can have access to his health data records. This means that he is not allowed to write data on the blockchain, instead, he only has privileges to query his data records from the blockchain.

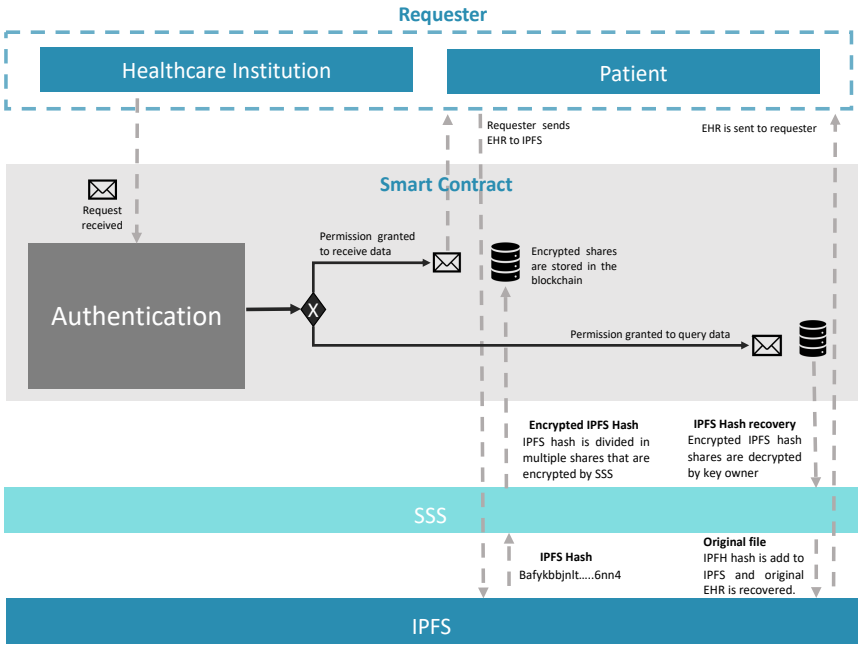

Figure 5-5- System architecture scheme exposing the different modules involved. The arrows represent the flow of information between connected entities.

As can be observed, the process is initiated by a request from a *Healthcare Institution* or a *Patient*. The request might be either related to the storage or access to patient medical records in the blockchain.

Here, the authentication block can grant or deny access to different entities based on the permission policies of the smart contract. In case of the permission is granted, there are two possible scenarios:

c) **Permission granted to stored medical records:**

In Figure 5-6, is exposed how the system behaves in case of permission granted to store medical records.

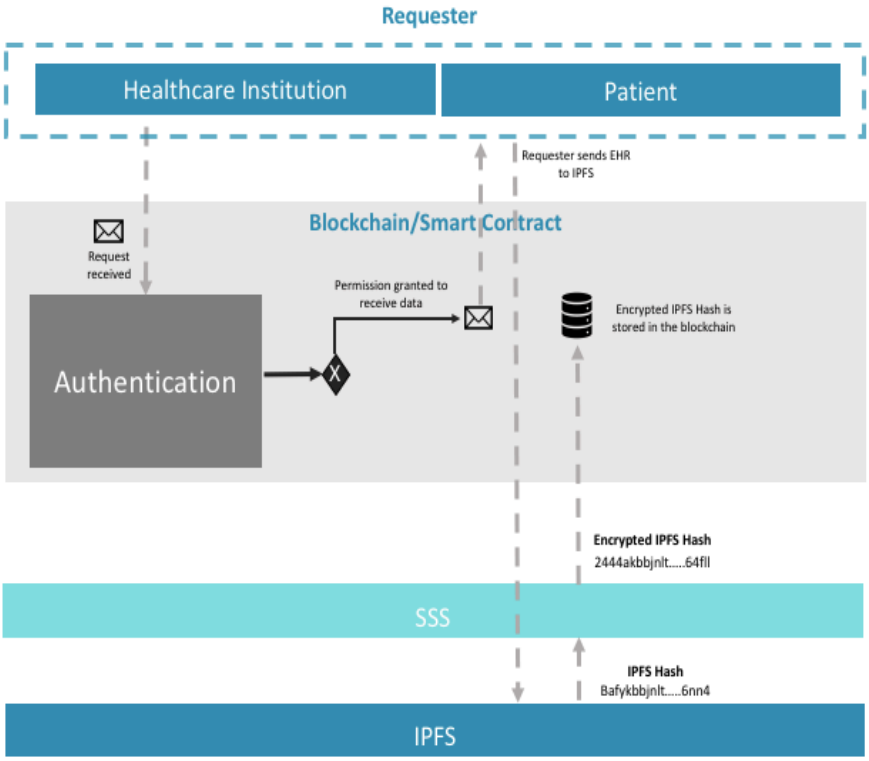

Figure 5-6- Data upload request by Requester.

Here, the requester initializes IPFS, to connect with the network. The requester is given an entity ID, public key and the addresses corresponding to the peers that he is connected on the IPFS network.

After the requester adding the content to the IPFS network, the protocol stores it into multiple IPFS Objects. In its turn, the client receives from the IFPS protocol a Content Identifier (CID) that is based on the content's cryptographic hash. Meaning that any difference in the content will produce a different CID and the same content added to two different IPFS nodes using the same settings will produce the same CID.

Thereafter the IFPS hash or CID is forwarded to the SSS module in order to be encrypted. Here, the CID is converted into multiple shares that are encrypted by the SSS algorithm.

The next step is again to interact with the Ethereum smart contract to store the shares in the network.

This interaction begins with the connection to the Ethereum blockchain, which is performed using the web3.py library based on Python programing language. Here there are there three basic ways to connect to Ethereum nodes. Via HTTP, WebSockets and OPC. For the simplicity of processes, HTTP was the chosen protocol, despise the via IPC being more secure as they rely on the local file system.

For testing purposes, it is used in a local Ethereum blockchain environment (Ganache). After a successful deployment of the smart contract to the local Ethereum

blockchain environment, the client is now ready to create transactions and queries to the blockchain.

In the context of the use case exposed in Figure 5-6, a client creates a transaction to store in the blockchain the multiple shares of the CID encrypted by the SSS algorithm. For the transaction to be authorized, it must be to be signed by the client using the private key. This will require gas since a transaction makes changes to the state. This alters data on the blockchain, which requires consensus through mining before the transaction can be committed to a block. Then, a smart contract function establishes a link between the files with a specific Patient for query purposes.

d) **Permission granted to access medical records:**

In Figure 5-7, it is depicted the whole process related to situations where permission to query medical records is granted.

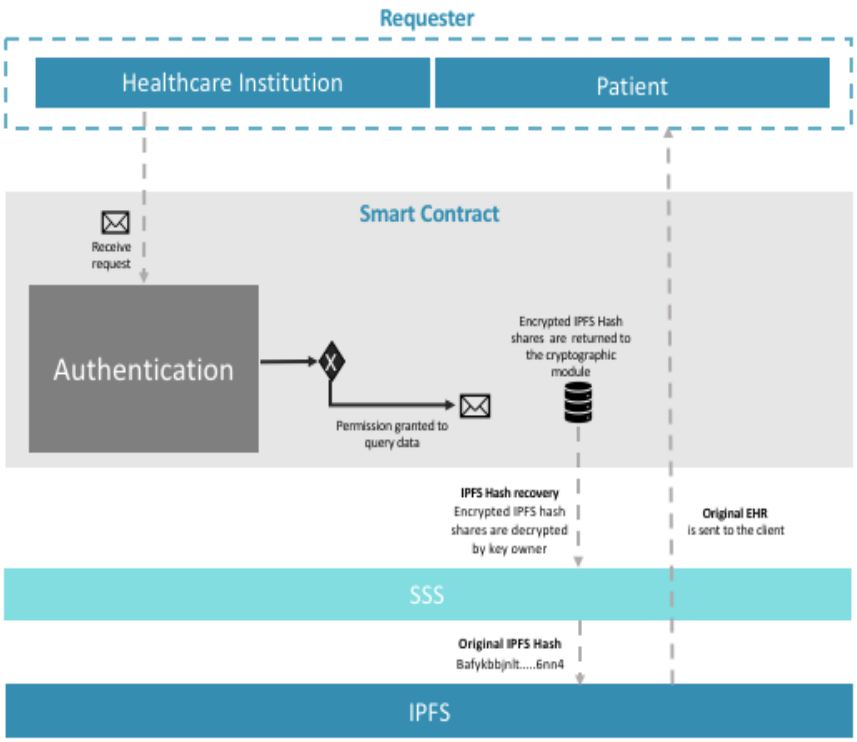

Figure 5-7- Data download request by Requester.

In this situation, the client executes a call function to the smart contract requesting a specific EHR stored in the smart contract. In its turn, the smart contract returns the encrypted shares of the corresponding IPFS hash.

Thereafter, on the SSS block, the encrypted shares are decrypted by the client key returning a CID in JSON file format. The client now connects to the IPFS network and is ready to read the content by passing to it the content's cryptographic hash. Finally, the IPFS network returns the original EHR to the client.

5.3 System Implementation details

In this section, can be observed the implementation details. The proposed system is conceived under a private network of Ethereum Blockchain. Ethereum is a decentralized, open source blockchain with smart contract functionality that makes efficient use of solidity. An object-oriented programming language for writing smart contracts. The implementation of the private network is due to the fact that the nodes are not connected to the main network. This decision is based on the context of testing purposes, because is reserved or isolated rather than protected or secure such as Ethereum public network.

5.3.1 Simulation setup

The implementation setups are based on the following specifications: 2.2GHz dual-core Intel Core i7 (Turbo Boost up to 3.2GHz) with 4MB shared L3 cache. Solidity is the chosen programming language to write the smart contract. Regarding the interactive forms, it is used Python programming language. The main tools used to conceive this system are exposed below:

a) **Visual Studio Code**

Visual Studio is a lightweight, freeware source-code editor designed by Microsoft for multiple operative systems. It allows developing, packing, and testing smart contracts.

b) **Ganache**

Ethereum Ganache is a private Ethereum blockchain. Specifically, it is a local in-memory blockchain designed for testing and executing commands. It simulates the features of a real Ethereum blockchain, making available to developers several accounts funded with test Ether.

5.4 Discussion

In this subsection, it is discussed the results of this system. It introduced a decentralized storage smart contract based on the Ethereum network that integrates IPFS protocol in combination with an off-chain solution for data privacy. For evaluation purposes, it was compared the results with some baseline methods for data storage and encryption algorithms. In addition, it was shown the average transaction gas cost of smart contract functions for data storage.

In that sense, tests were performed to study the system behaviour in a situation where files with different sizes are uploaded. In detail, the computational time that takes to upload different file sizes was tested and compared with conventional cloud storage systems. As can be observed in Figure 5-8, this system performance is constant, 430ms on average, independently of the file size. In contrast, conventional cloud storage systems show an exponential increase in computational time with the increase of file size, reaching a maximum of 2370ms with a 2048MB file size. These results prove that this system above all of the advantages of decentralization still takes less time and energy to store data than conventional storage systems.

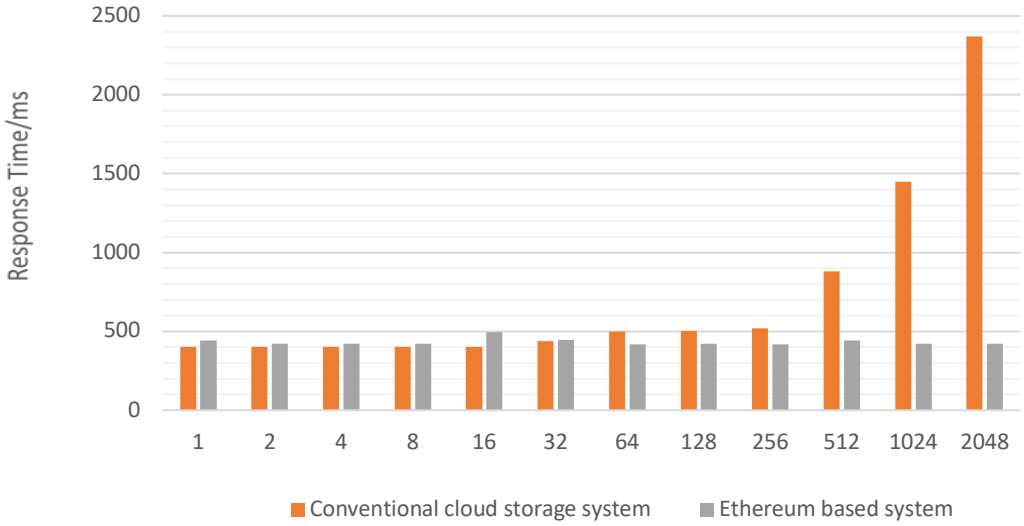

Figure 5-8- Average Response Time vs File Size.

Secondly, the Ethereum-based smart contract was evaluated in terms of cost. To run the smart contract, the gas price is adjusted to 3 Gwei. 1Gwei is equivalent 10^9Gwei, which is approximately 10^{-9} ethers. The cost measured for the smart contract can be consulted in Table 5-1. Here, is shown the gas consumption values for multiple functions executed in smart contract. Ethereum smart contracts are executed on EVM and once the byte code is generated it is sent within a transaction and then it exists on the blockchain. This is done once when the smart contract is deployed and the amount of transaction gas consumed is 1891116 where the actual cost belongs to $6,47 USD. In situations where a file is uploaded to the network, addFile function has an inherent transactional cost of $0,54 USD. Similarly, to remove a specific file from the network, removeFile function is invoked prevailing a transactional cost of $0,089 USD. The last function of the smart contract is queryFile. Once a requester queries a file from the network, the cost associated with the query function is $0,27 USD.

*Table 5-1-*Smart *contract cost test (gas price = 3 Gwei, 1 ether = 3247 USD).*

Functions	Transaction Gas (GWei)	Transaction Cost (Eth)	Actual Cost (USD)
Contract creation	1891116	0.00189112	6,47
Add file	158083	0.00015808	0,54
Remove file	25930	0.00002593	0,089
Query file	78337	0.00007834	0,27

Finally, the performance of encryption scheme integrated in this work, was tested by measuring the computational time. SSS was the chosen scheme for encryption of file's Hashes derived from IPFS protocol. Its performance was tested against two well know encryption schemes such as advanced encryption standard (AES) 128 and AES 256 [185]. The results can be consulted in Figure 5-9. SSS and AES128 schemes present the best results at encryption level with both 7ms computational time. On the opposite direction, AES256 performed slightly slower with 9ms. Regarding the decryption response time, again SSS shows the best performance with a 2ms computational time where as AES128 and AES256 recorded a 3ms and 4ms respectively. As can be observed, AES256

performances in both scenarios slower than AES128, this is due to the fact that it has a larger the combination of key which reflects in an increase of computational time response. AES128 offers 2^{128} combinations of keys, while AES256 offer 2^{256} keys.

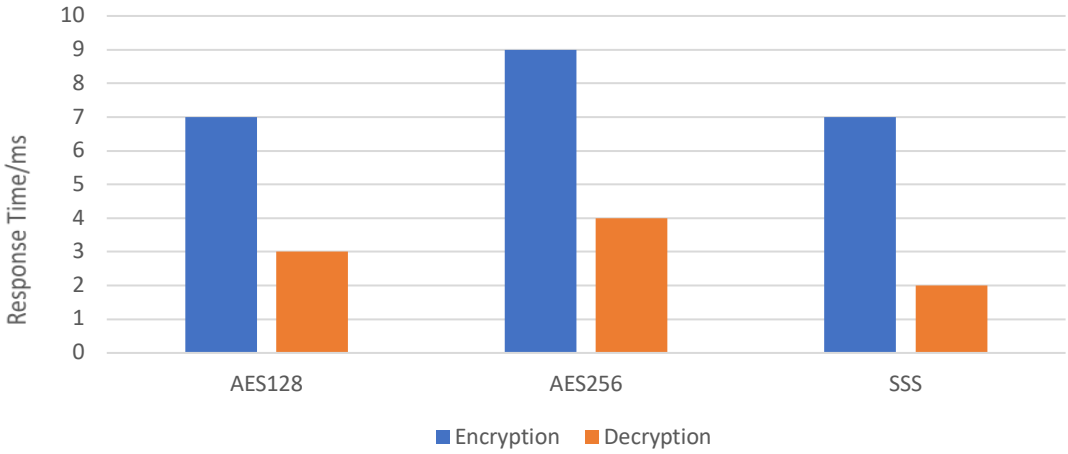

Figure 5-9-Comparative execution times of encryption/decryption algorithms.

5.5 Summary

In this section, it was proposed a new decentralized storage system for EHRs on assistive applications. This system was conceived to overcome one of the major limitations of the current blockchain-based applications such as such as the expensive cost of storing large files and the lack of privacy inherent to the public aspect of blockchains.

To overcome these limitations the proposed system presents a smart contract based on the Ethereum blockchain with a combination of two off-chain solutions. The first off-chain solution is designed to provide a decentralized storage alternative that tackles the issue of the limited storage capacity of the Ethereum blockchain. The further off-chain solution provides an extra layer of privacy by applying a cryptographic schema to the data.

Results show that in the above a significant improvement in terms of security, the proposed decentralized storage system also achieves less computational time to upload files when compared to conventional storage systems. In addition, the SSS cryptographic schema used for data privacy also outperforms in terms of response time popular schemas used in this type of solution such as AES 128 and AES 256. The following chapter introduces a new methodology to achieve on-chain programable privacy.

Chapter 6

A Private Blockchain with Access Control

Currently, the true essence of a decentralized blockchain relies on a public blockchain where no central authority is in control of the nodes of the network. However, those blockchain networks are public by default, meaning that all the data stored in the smart contract is exposed to the public domain. This is considered a critical issue because privacy in the health sector is a major requirement. To solve this issue, some off-chain solutions applied cryptography schemes to encrypt sensitive data and then store it on the blockchain. However, this type of approach is risky because, in case of the loss of control of cryptographic keys, the data is lost indefinitely [186].

Another key concern is related to the adoption of a smart contract as a permission management database for authentication purposes. In the last chapter, it was described a system that has leveraged the use of smart contracts as an authentication mechanism for access to patients' records. Nevertheless, this approach reveals to be very complex in terms of design decisions concerning the governance of public and private key. Here, the identification of an entity is linked to his corresponding public key, which in case of loss or theft of the private key results in a total loss of control of a patient's EHR. This might lead to malicious access from an unauthorized entity to a patient's sensitive data.

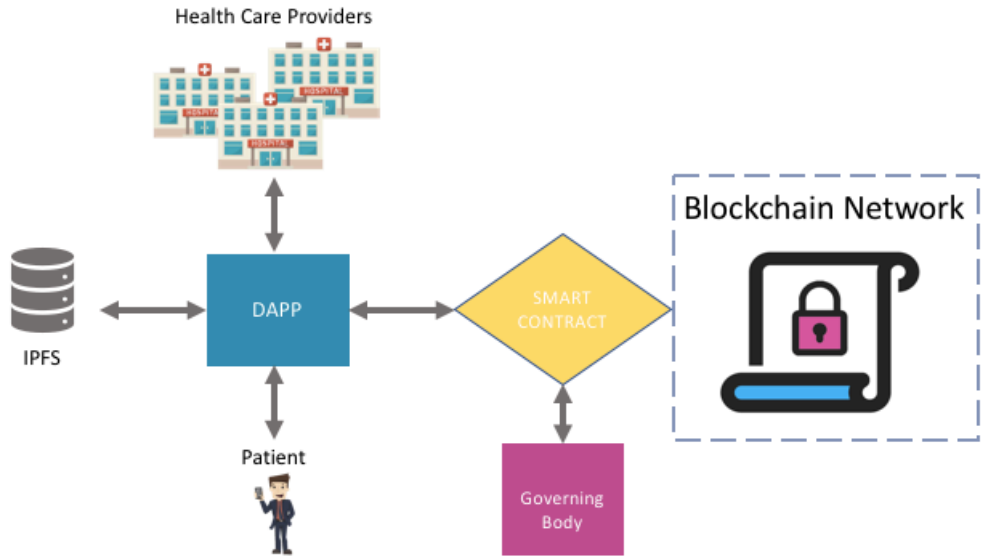

Figure 6-1-System Architecture with the integration of a governing body module.

In this section, it is presented an optimized approach for decentralized storage systems focusing on the eHealth sector. It was developed a new smart contract to tackle the lack of computational privacy on smart contracts. In addition, a governing body is introduced for permission authentication entities. In the following sections, it is described in detail how each module operates and its advantages.

6.1 Governing Body – Public Key Governance

In order to handle the problems related to the authentication of an entity or even the lost or eventual theft of private keys, an off-chain solution is conceived to be responsible for managing the permission of entities to have access to sensitive data storage.

Despite this governing body being considered centralized, it is crucial to increase the security of the whole system. In that sense, in this work, the management of the governing body is attributed to the National Health Institutes. This premise comes from the fact that in the past these organizations were already responsible for the storing and authentication of the digital identities of their citizens.

In detail, the governing body is responsible for a mapping between the healthcare entities and patients, by linking them with their corresponding public keys. Let, H represent the set of health care providers, and P be the set of patients in network ($|H| < |P|$). Each hospital h, can be part of a patient p list of granted access entities $h \in H$.

In this way, it is possible to ensure that in case of loss or theft of the private key, the patient can still have access to his health records. In this specific case, the patient should notify the governing body claiming that his private key is compromised. Then, the governing body would generate a new wallet, associating both private and public keys with the patient. As such, the previous public and private key will no longer be recognized as legitimate digital identity preventing any access attempt from malicious actors.

Despite being focused on eHealth, the proposed approach can be applied to a wide range of sectors that needs to secure their authentication system to access sensitive data.

6.2 Secret Network Protocol

Secret Network is a layer one blockchain solution built with Cosmos. Cosmos network is an ecosystem with a set of tools that allows developers to build custom, secure, scalable and interoperable blockchain applications [187]. Secret Network uses Proof-of-Stake (PoS) as a consensus algorithm. In detail, PoS has been achieved thanks to the use of Tendermint's Byzantine fault-tolerant consensus algorithms. Tendermint enables an application can be replicated on many machines, ensuring that unfaulty machines or nodes can see the same transactions and compute the same state and validators agree on a set of transactions to append to the blockchain[188]. One special characteristic of Tendermint is that it is Byzantine Fault Tolerant, meaning it can only tolerate up to a 1/3 of failures, which can include arbitrary behaviour (Malicious attack or hacking i.e.) which is considered a fundamental problem in the distributed system[189].

Due to the fact that Secret Network is a layer one blockchain protocol, it is considered chain-agnostic, meaning that it is designed for communication of blockchain amongst each other and blockchain with the outside world using the Cosmos InterBlockchain Communication protocol [190][191].

Moreover, Secret Network offers a unique platform to build censorship-resistant applications that keep data encrypted. It is accomplished by leveraging TEEs to enable secure, private computation over encrypted data. Essentially, TEEs separate the kernel into several partitions and guarantees a strong isolation between them. In that sense, the information that flows between those partitions is isolated and controlled. The data with one partition can not be read or modified by other partitions, the shared resources can not be used to leak information into other partition and their communications cannot occur unless explicitly permitted. Unlike dedicated hardware coprocessors, TEE can ensure that data is stored, processed, and protected in a trusted environment that cannot be tampered

with [192]. By this, it is possible to implement programmable privacy, which is defined as arbitrarily complex data privacy controls within an application[191]. Finally, SRCT is a native coin of the network. The token is used for staking, governance, transactional and computational fees.

6.2.1 Secret Contracts

Secret Network enables the creation of a Secret Contract designed to perform privacy-preserving computation on a public blockchain. Its elementary foundations can be seen as similar to Ethereum Smart contracts, however, Secret Contracts enable the encryption of data where input, output and state data are encrypted. These features are achieved thanks to a decentralized network of validators, who run Secret Contracts execution inside TEEs.

Secret Contracts are Rust-based smart contracts that compile to WebAssembly. Secret Contracts, which are based on Go-CosmWasm. CosmWasm is smart contract framework that enables integration with Cosmos SDK and Cosmos Ecosystem to easily add smart contracts Cosmos blockchain [193]. Cosmos wasm, has two majors feature that combined are considered as a major advantage when compared to existent frameworks. Firstly, its architecture is designed to avoid all attack vectors present in Ethereum. Reentrancy Attack which is a procedure where its execution can be interrupted in the middle, initiated over (re-entered) and both runs can be completed without any errors in execution. Cosmwasm avoids this completely by preventing a contract from calling another one directly. Another attack vector comprises denial of service, where a malicious actor can upload a contract that ran an infinite loop to halt the chain or write tons of data to fill up the disk. CosmoWasm provides a tight sandbox with no default access to the OS, tight resource limits can be provided to smart contracts and mitigate eventual DDoS attack to the blockchain.

6.2.2 Validators

The Secret Network validators are responsible for proposing new blocks to the blockchain, and confirming blocks proposed by other validators. A validator is a full node that can also propose and sign blocks. Validators perform all the requested computations in each block via the compute module, which means all computations also occur as part of the consensus process [194]. Validators run Intel SGX chips, and have gone through remote attestation, a process by which Intel SGX chips are verified. It will have also successfully completed a network registration process. As part of registration the validators are provisioned with the secret keys they need to participate in private computations. Validators run the Secret Network code, and execute WASM code within a TEE. They are responsible for achieving consensus on computation results, and proposing and/or validating new blocks in the Secret Network's blockchain. Validators also participate in governance.

6.2.3 Encryption

The Secret Network uses both symmetric and asymmetric encryption protocols. Asymmetric cryptography is used for achieving consensus and sharing secrets between nodes and users, whereas symmetric cryptography is used for input/output encryption with users of Secret Contracts, as well as internal contract state encryption [195].

Specifically, the process of encryption that enables the privacy features of the network starts before the genesis of a new chain, with a creation of a bootstrap node. This node goes through three core processes. Firstly, it creates a remote attestation proof that the TEE or node's enclave is genuine. Secondly, a random 256-bit seed known as a consensus seed is generated inside the enclave. This consensus seed is a critical component of the network since its encryption schemas as all the other derived keys and consequently, functionalities of the protocol are built upon the distribution of this original consensus keys through the network validators. The derived keys from the consensus seed are done deterministically by using HKDF-SHA256 [196] in combination with contextual-relevant data. These derived keys are used for enabling the registering of a new node, I/O encryption and state encryption. Finally, the process is concluded by publishing to the Secrete Network genesis.json the remote attestation proof that the enclave is genuine, the public key for the consensus of I/O exchange and public key consensus seed exchange. Theses asymmetric keys generation is performed by a Curve25519 [197] while Ellipic-curve Diffie-Hellman(ECDH) [198] is used for deriving symmetric encryption keys which are used to encrypt data with AES-128-SIV[199].

One of the main objectives of this cryptographic process is to enable that the original consensus seed can be broadcasted to new validators nodes without being compromised by malicious validators. With the consensus seed already sealed in the new validator node disk, it is now able to execute all necessary key derivations to get network-wide secrets in order to participate in block execution and validation. By this, the Secrete Network protocol can ensure the immutability of data and privacy not only at a transactional level but also at a computational level.

6.2.4 Secret Contracts

Secret Contracts are code executing over encrypted data. This means that the despise the Secret Contracts being public, the data they operate is private. This feature enables users to have confidence that contracts will perform as functioned, while simultaneously ensuring user data cannot be viewed by any counterparty. Secret Contracts are developed in Rust programming language further compiled to WASM binaries. Contracts are stored on the Secret blockchain, where their code is publicly available. They execute inside the trusted part of the Secret Network.

6.2.5 Transactions

In Figure 6-2 is possible to observe how the Secret Network protocol behaves in case of a transaction.

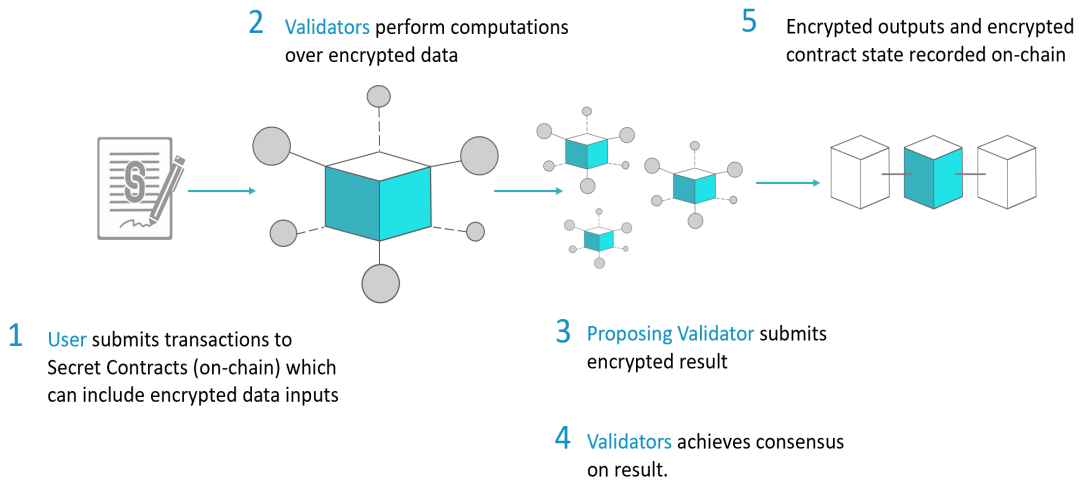

Figure 6-2- Transaction process on Secret Network protocol.

When a user executes a transaction to Secret Contracts the body of the message contains the user's public key and signature (standard Tendermint transaction) enabling validators to also derive a shared key using ECDH. The user encrypts input data with this shared key, using an AES-256-GCM authenticated encryption scheme. The next step implies validator nodes which receive the encrypted data from users and perform computations by executing the Secret Contract. Here, validators decrypt the inputs and execute the contract function inside the node's TEE. The contract state is updated, and the transaction output is encrypted for the transaction sender. Subsequently, the validator proposes a block containing the encrypted outputs and updated encrypted state. If more than two-thirds of participating validators achieve consensus, the proposed block (encrypted output and state) is committed in the secret Network.

6.3 System Model

Similarly to the Section 5.2, this section delineates the workflow of the whole system. The process model includes changes. Specifically, the introduction of the secret network in the replacement of Ethereum Network and the integration of a governing body module which resulted on modifications from the system architecture. In terms of entities, the system model preserved the same structure. In that sense, Health Care institution, Patients and Governing body are entities of the system model.

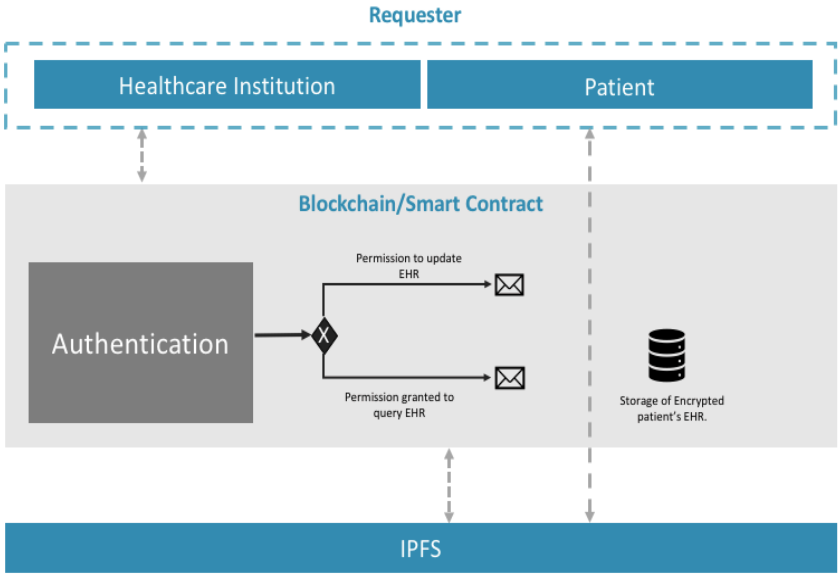

Figure 6-3- System architecture scheme exposing the different modules involved. The arrows represent the flow of information between connected entities.

Due to the privacy features of the Secret Network protocol, the inclusion of the SSS module is no longer required as data encryption is ensured by the Secret Network protocol. In that sense, this optimization resulted in an exclusion of SSS module from the system architecture as can be observed in Figure 6-3. Moreover, an Authentication block that includes a Governing body entity was added for managing the permission of entities for accessing the data.

Regarding the process, it is initialized by a request from the entity's *Healthcare Institution* or *Patient.* This request as exposed in Figure 6-3 goes through an authentication process to verify that the entity in question has legitimated access to specific entity data. Subsequently, if granted the request that can be either related to the storage or query of an EHR is then processed. In the next subchapters, the authentication process as well as both types of requests will be described in detail.

6.3.1 Authentication-public key governance:

For a better understanding of the authentication process and the role of the Governing body, it is presented in Figure 6-4 a BPMN process model that illustrates the whole process of the authentication.

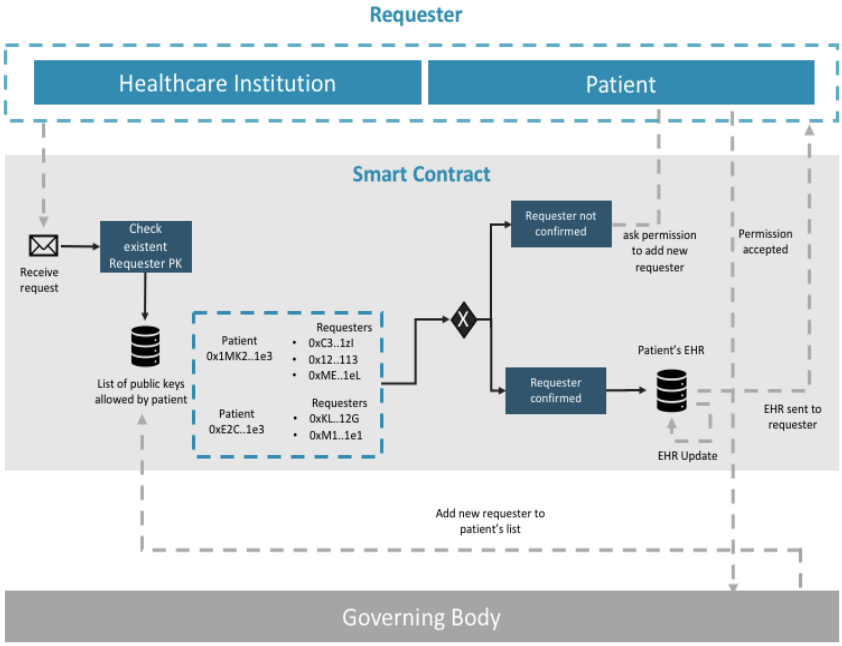

Figure 6-4-A BPMN process model of smart contracts as permission management database.

The smart contract operates as a permission management database maintaining a mapping between the public keys of patients and the public keys of authorized requesters. When a request is performed by a healthcare party, the smart contract starts a verification stage in order to conclude that this party is a licensed and legitimate party and not a malicious actor. In detail, the smart contract searches for a match between the healthcare public key and the patient public key. In case of a match is successfully found, the healthcare party has now granted access to the storage data corresponding to the requested patient public key.

On the other hand, in case of a negative response, the smart contract immediately notifies the corresponding patient for adding access permission to the healthcare agent. A positive response by the patient will result in a request by the smart contract to the governing body for adding permission. Here, the governing body works as an off-chain entity to increase the security level of the whole system. In detail, the governing body is the only authority that can modify the data in the smart contract. It is accomplished by forcing the smart contract to only accept modifications by the owner of the smart contract.

In that sense, the governing body has the responsibility to add or remove new entities' permissions of the smart contract. Once the permission is accepted by the patient, the governing body ensures that the entity is stored within the smart contract as one or multiple mappings between the healthcare agent to the patient's public key.

The introduction of the governing body serves as a potential solution in an event of a loss or steal of the private key. Moreover, in this research was considered that the developed design minimizes the use of private keys. The owner of the smart contract (governing body) only needs to use their private key to add/remove new entities to the smart contract. Meaning that most of the time, the private key will be saved in cold wallet settings, reducing the risk of malicious actors stealing the private key and taking control over the smart contract.

6.3.2 Permission granted for the storage of EHR

In Figure 6-5 is exposed the process of storage of medical records in case of the request being granted by the authentication module.

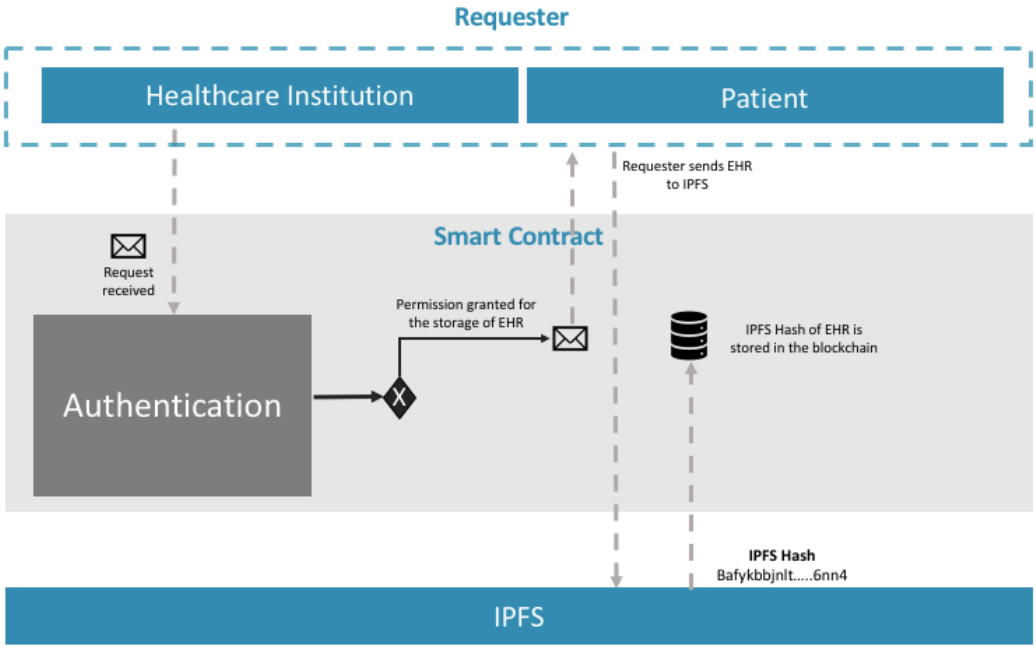

Figure 6-5- Data storage request by Requester.

After the request being granted by the authentication module, the requester sends the EHR to the IPFS module. Subsequently, an entity ID, public key and the addresses corresponding to the peers responsible for the connection to the IPFS network are attributed to the requester. Once the connection to the IPFS network is established, the requester adds the EHR to IPFS which will return a hash that serves as the unique identifier for the content of that file (CID).

The next step is to execute a transaction to Secret Network in order to store the CID corresponding to the EHR. In detail, the smart contract creates handles messages for storing data of a specific entity on the chain. This interaction works in the following way: the client or requester sends a message that is received by the contract as an encrypted Base64-enconded version of the JSON stringify'd version of the original message (i.e., Javascript object) defined in the client code. Depending on the client's request, the message can be type handle or query. In this case, the message is a type handle meaning that a transaction is executed and the CID is stored into the blockchain network. This transaction requires a payment from the requester in order to succeed.

6.3.3 Permission granted to access to medical records:

The process of query EHR from the Secret Network blockchain is exposed in Figure 6-6.

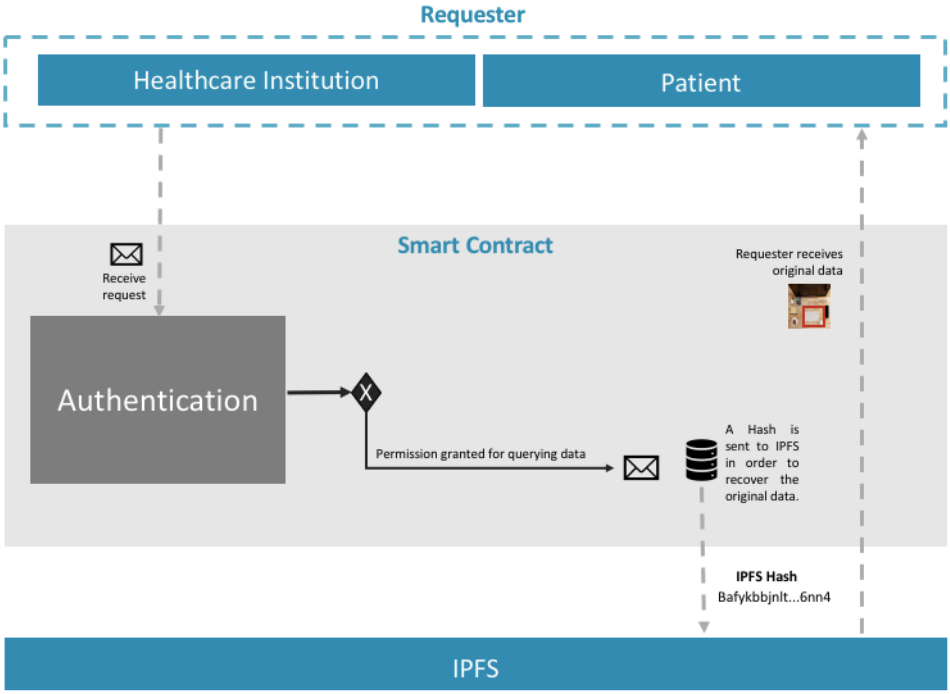

Figure 6-6- Data query request.

Similarly, to the process of storage of an EHR, an entity must have its request granted from the authentication module. In case of its permission being granted, the entity has now access to the data storage in the blockchain. The query process is passed by executing a query message to the Secret Network for the hash of a specific EHR. Again, this message is received by the contract as an encrypted Base64-encoded version of the JSON stringify'd version of the original message defined in the client code. In contrast to the handled message described in subsection 6.3.2, query messages in Secret Network do not incur in any cost. The client receives now a response containing a decrypted CID that is sent to IPFS in order to recover the original EHR to the client.

6.4 System Implementation Details

In this section, can be observed the implementation details. The proposed system is conceived under the Secret Network blockchain. Secret Network is a decentralized, open source blockchain with smart contract functionality that makes efficient use of Rust. Rust is intended to be a language for highly concurrent and highly safe systems, and programming in the large, that is, creating and maintaining boundaries that preserve large-system integrity. The implementation of the private network is due to the fact that the nodes are not connected to the main network. This decision is based on the context of testing purposes, because is reserved or isolated rather than protected or secure such as the Secret Network public network.

Regarding the client side, it was utilized Node js for easily building fast and scalable network applications. Node js uses an event-driven, non-blocking I/O model that makes

it lightweight and efficient, perfect for data-intensive real-time applications that run across distributed devices.

a) Simulation setup

The implementation setups are based on the following specifications: 2.2GHz dual-core Intel Core i7 (Turbo Boost up to 3.2GHz) with 4MB shared L3 cache. Rust is the chosen programming language to write the smart contract. Regarding the interactive forms, it is used Rust programming language. The main tools used to conceive this system are exposed below:

b) Visual Studio Code

Visual studio is a lightweight, freeware source-code editor designed by Microsoft for multiple operative systems. It allows for developing, packing and testing smart contracts.

c) Holodeck-2

Holodeck-2 is a private Secret Network blockchain. Specifically, it is a local in-memory blockchain designed for testing and executing commands. It simulates the features of a real Secret Network blockchain, making available to developers several accounts funded with SCRT tokens for Cosmos test net.

6.5 Discussion

In this section, it is discussed the results from this system. It is presented a decentralized storage system that leverages blockchain technology to increase security, privacy and performance of EHR storage systems. This system differentiates from the overall systems in three major aspects. Firstly, it is introduced a novel smart contract based on Secret Network that provides data privacy by default. In opposition to existents privacy blockchain projects, that offer privacy by integrating off-chain solutions based on cryptographic schemas, here it is presented an on-chain smart contract solution that provides privacy at both transactional and computational level. By this, it is possible to overcome the major issue of off-chain solutions, the loss of the decryption key that would mean the loss of data forever. Instead, users can trust a system that provides data privacy at all the stages of computation, from the input till the output information. Lately, it offers an off-chain solution that works as a governing body for handling the authentication of different entities and solving the problem of the lost or eventual theft of private keys.

In order to evaluate the performance of this system, it was tested the computational average time of different file sizes against two other systems. Those include a conventional cloud system and a decentralized storage system based on the Ethereum blockchain. In section 0, it was exposed the implemented off-chain solution that aims to encrypt the data to then be stored on the Ethereum network. This approach aims to add privacy layers to blockchain which is considered a common practice in the literature. The graph chart in Figure 6-7, demonstrates that as the file size increases the average response time of conventional could storage systems increases exponentially. On the other hand, both blockchain solutions present a constant average time, independent of the file size. In this case, this system shows an increase of 20ms approximately compared to the

Ethereum-based solution. This increment in time is justified by the computational cost of TEEs and both symmetric and asymmetric encryption protocols used in the network. However, this gap in time response is not significant considering the great advantages of this system in terms of privacy and security.

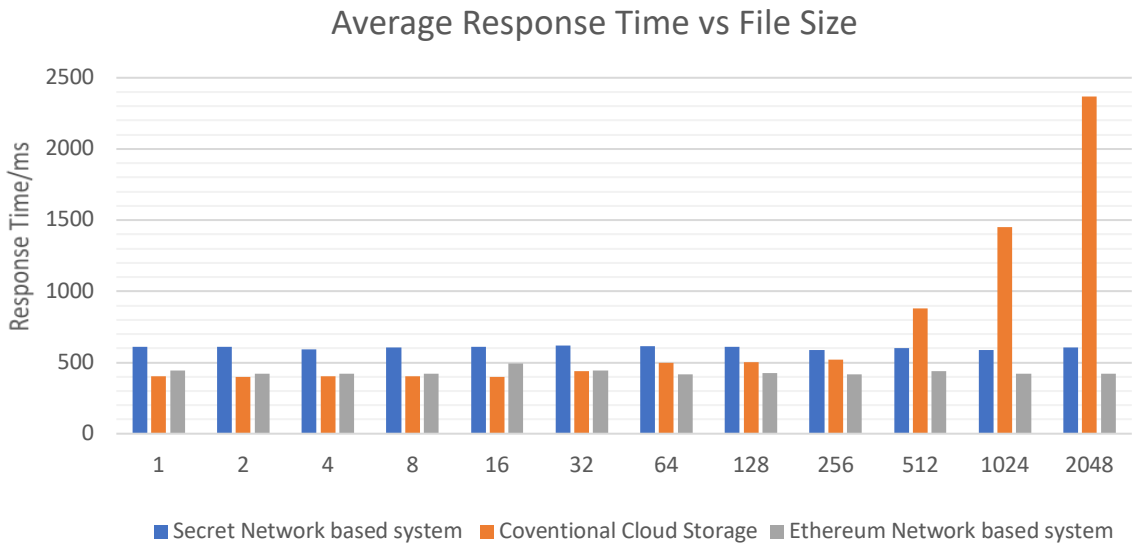

Figure 6-7- Average time versus file size for convectional cloud storage, developed system and Ethereum based system

In Table 6-1 the functions of both smart contracts with different values of transaction gas along with actual cost are shown. Regarding the developed system, the total transaction gas for uploading the smart contract, executing transactions and call functions are 2 000 000, 500 000 and 80 000 (USCRT), respectively. In its turn, the Ethereum Network based system shows a total 380 000 000 transaction gas, 3 003 600 overall execution gas, and 946 300 call function gas. Functions' transaction gas of both systems is noticed to remain unchanged after multiple transactions. Regarding the cost, in the developed system the variation between each function are not significative showing for uploading, executing and calling a cost of 0.20, 0.17 and 0.02 respectively. Nevertheless, currently the cost functions of the Ethereum based system are extremely expensive considering the nature of EHR. For uploading the smart contract to the network, is 119,86 $, execution function is 9,46 $, and finally call function shows a cost of 2.98 $. The high cost on the Ethereum Network-based System is due to the finite block space and an ever-increasing volume of on-chain activity on the Ethereum blockchain. In that sense, Secret Network-based system shows to be a very scalable option, firstly because it presents a faster response than conventional systems. Secondly, it is worth mentioning that any activity that deals with sensitive data prize security and privacy. In that sense, this system is far behind all the conventional storages system and also overall blockchain solutions that do not offer privacy at the computational level. Finally, the system's operating cost is not relevant considering the advantages in terms secure, privacy and computational power.

Table 6-1- Smart contract cost test.

Functions	Transaction Gas (GWei)	Transaction Cost (Eth)	Actual Cost (USD)	Transactions Gas (USCRT)	Transaction Cost (SCRT)	Actual Cost (USD)
Upload	380 000 000	0.038	119,86	2 000 000	2	0.20
Exec	3 003 600	0.003036	9,46	500 000	0.5	0.17
Call	946 300	0.0009463	2,98	80 000	0.08	0.02

6.6 Summary

In this chapter is presented an alternative to smart contracts that operate as a permission management database for EHRs. This alternative aims to tackle the existent lack of computational privacy on the on-chain public networks as well as the risk associated with the lost or eventual theft of private keys.

The first limitation is addressed by introducing a smart contract based on the Secret Network that offers privacy by detail. Meaning that computational privacy and transactional privacy are achieved by leveraging TEEs to enable secure, private computation over encrypted data. The last limitation is tackled by proposing the introduction of a governing body for handling the authentication of different entities and solving the problem of the lost or eventual theft of private keys.

For evolution purposes, the computational efficiency of the developed solution was tested against two other systems. Those include a conventional cloud system and a decentralized storage system based on the Ethereum blockchain. Moreover, it was also conducted an exercise to test the operating cost of such a system, this time only against an Ethereum-based system

Results show that the proposed solution has high computational efficiency when compared to a conventional cloud-based system and a slight decrease in efficiency from the Ethereum-based solution. However, considering the great advantages in terms of computational and transactional privacy, this slight efficiency difference becomes irrelevant. On the other hand, it was possible to observe that the operating cost of the proposed solution is much lower than the Ethereum-based system. To conclude this work offers a very efficient and cheaper alternative to the current state of art decentralized solutions for the management EHRs.

Chapter 7

Conclusion and Future Work

This chapter summarises the work within this thesis and discusses the key outcomes of the conducted research. Section 7.1 contains a summary of the work presented in this thesis. The limitations are listed in Section 7.3, conclusions and research contributions for the entire thesis are listed in Sections 7.2 and 7.4 respectively, recommendations for further work are presented in Section

7.1 Research summary

In this dissertation, it was developed a framework for assistive applications focused on meeting the requirements for building AR real-time assistive applications. Existent AR-assistive frameworks often lakes in terms of security and privacy of data sharing. As result, the lack of encryption and associated best practices for encryption schemes and key management have allowed attackers to gain access to millions of data records. Moreover, AR technology by itself is not enough to provide effective support during task performance. Environmental sensors are needed to feed systems with data to understand the user's environment. Here this data in combination with artificial intelligence will enable these systems to react accordingly to the user's needs. The next paragraphs summarize the work within this thesis and discuss the key outcomes of the conducted research to address the gaps mentioned above.

The opening chapter, Chapter 1, provided a brief introduction to the subject of AR-assistive applications frameworks, establishing the context for a new approach to improve the intelligence and usability of such systems as well as new methodologies to increase their security and privacy. The problem statements and objectives of the thesis were outlined in sections 1.2 and section 1.4 respectively. These objectives were fully explained in the consequent chapters. The chapter concluded with the research methodology in section 1.5 and the thesis structure in section 1.9.

In Chapter 2 the reader has been provided with an extended theoretical and conceptual introduction to the subjects concerning frameworks of AR-assistive applications that were relevant to the context of the research. Apart from the introduction to state of the art of frameworks of AR-assistive applications, the conducted literature review emphasised the problems that are still unsolved on this topic. In particular, (i) questionable accuracy of current methods for object detection based on human visual attention. (ii) low security and privacy standards in terms of data storage and data-sharing of EHR. (iii) Appropriate solution for authentication procedures under decentralized network protocols.

For each of the problems mentioned, the conducted literature review revealed the benefits and drawbacks of the considered techniques used to mitigate those.

It was considered important by the author of this thesis to highlight the impact of practical systems on the research work. Hence, Chapter 3 presented a framework for AR-based smartglasses applications. This framework was designed to follow a modular approach seizing improve its scalability as well as handle and process different data inputs while ensuring the security and privacy of the data storage. To test the framework, it was

developed a smart application for helping users during procedural tasks. The application provided easy access and tools to create tasks and evaluate the user's performance during tasks. To increase the level of system intelligence, the application leveraged technology to detect objects under the user's visual attention. This smart object detection algorithm as well as the security and privacy features were the research outcomes from Chapter 4, Chapter 5 and Chapter 6 respectively. Essentially these chapters covered the identified research gaps of current AR applications, such as low security and privacy standards as well as lack of intelligence. This framework was tested against ongoing conditions and the results show that its performance fits the requirements of real-time applications.

Chapter 4 opened Part 4 of this thesis with a proposed methodology aimed at increasing the accuracy of object detection models based on human visual attention. The methods examined in chapter 2.2 were scrutinized under the scope of their effectiveness to be used in real-time environments. In particular, the method proposed by [124] was examined in depth. The method was particularly focused on achieving visual attention by defining a specific fixed region in first-person video, assuming that human attention is fixated at the center spatial coordinates of an image frame. Then, features of the centre region were extracted for object detection. However, this technique would not perform well due to the variability in the size of objects and the different angles of perspective, only a limited part of object features is often described in the centre region. So, it is hard to distinguish the right object out of multiple objects at the center of the frame. Moreover, despite most human visual attention being focused on the spatial center region of the visual field, attention can be also observed in peripheral regions meaning. To address the challenge of mimicking human visual attention and discovering objects under it, in this chapter was proposed a methodology that combined both hardware and software solutions. In terms of hardware, it integrated an eye tracker device to address the spatial variability of human visual attention. Nevertheless, eye gaze coordinates on an image frame are not enough to decipher which object is the attention upon, as several objects can be overlapped. In this regard, it was proposed a method that involves a time series sliding window approach and a probabilistic heuristic function to analyse user gazes data around objects in egocentric images. This methodology has shown that cluttered environments outperform commonly used methods for the detection of objects in real-time and is claimed by the author of this thesis as the first major contribution to AR-based smartglasses systems.

In Chapter 5 a decentralized storage system was proposed seeking to store health data records in a decentralized and secure manner. The proposed solution is composed of a smart contract based on the Ethereum network that integrates IPFS protocol in combination with an off-chain cryptographic schema to preserve data privacy. The solution was evaluated by comparing its performance with some baseline methods for data storage and encryption algorithms. These results prove that this system above all of the advantages of decentralization still takes less time and energy to store data than conventional storage systems.

Security and privacy are always considered important requirements of any data-sharing application. Despise the good performance of the system proposed in chapter 5, in terms of privacy its results were not ideal. The lack of computational privacy is still an open issue in the majority of decentralized solutions.

In that sense, Chapter 6 it was introduced a smart contract based on the Secret Network blockchain that offers both transactional and computational privacy. Moreover, the

presented solution abolishes the need for off-chain solutions for data encryption which have the risk of loss of control of cryptographic keys. Finally, a governing body was introduced for permission authentication entities. The system was tested against a conventional cloud system and a decentralized storage system based on the Ethereum blockchain. The results show that this solution is a very scalable option because it presented a faster response than conventional systems. Moreover, it offers a much higher level of security and privacy when compared with other blockchain-based solutions. Finally, its operating cost is not relevant considering the advantages in terms secure, privacy and computational power.

7.2 Conclusions

The conclusions reached in the development of this dissertation are the following:

a) The existent frameworks for assistive living application-based AR are mostly designed to address a specific use case. Those frameworks lack in terms of providing a scalable architecture to serve a wide range of applications. To address this shortfall, it was proposed a real-time assistive framework to deliver AR visual cues instructions based on the user behaviour during task performance. The framework followed a modular approach enabling the extensibility of new functionalities, sensors and new forms interaction between user and the system.

b) Understanding the user's context is a critical step to extending the support capabilities of any assistive technology. The researcher focused his efforts on extracting meaning from the user's visual attention and his object's surroundings. Existent models for discovering objects of user's interest tend to have their accuracy at low levels when shifting to cluttered environments. The proposed methodology contradicts those results thanks to a combination of sensors like an eye-worn eye-tracker with a world camera, and an algorithm based on a series of probabilistic heuristic functions that outperform the current literature models.

c) To mitigate potential cyber-attacks, often linked to the lack of data privacy and security standards, the researcher has extended the framework by developing a decentralized storage system. In addition to the increase in security and privacy, this system was designed to reduce the high storage costs typically found on public blockchains such as Ethereum and traditional cloud-based systems. This approach was able to reduce the storage cost by integrating IPFS protocol into the network along with encryption schemes to ensure the privacy of the data. In such approach, files are stored in the IPFS protocol where their hashes are encrypted and then stored on the Ethereum blockchain.

d) The integration of a decentralized storage system into the framework raised two major issues. These are (i) how to tackle the problem of lost or stolen of the private key and (ii) how to solve the issue of transactional and computational

data privacy in public blockchains like Ethereum. The first challenge was tackled by developing an off-chain solution that works as a governing body which is responsible for legitimate or denying the access and interaction of specific users with the smart contract. In Chapter 5, encryption schemes were applied to bring privacy to the data stored in the blockchain, however, transactional records were still public and sensitive data could also be exposed during the computations performed by the nodes of the network. The development of a new smart contract based on Secret Network, which uses TEEs and symmetric and asymmetric encryption protocols for enabling both transactional and computational privacy to the network.

7.3 Limitations of the Research

This study attempts to contribute to the development of AR-based smartglasses assistive living applications. The study developed a framework with tools to extract user contextual information while preserving its data privacy and security in a decentralized fashion. However, as with any other research, there were a few limitations, detailed as follows:

a) This study designed a methodology for discovering objects upon user visual attention in cluttered environments. However, the range of the experiment was limited to six individuals and the design of the experiment was only tested in one format. Thus, the results could have limited generalisability, and so findings may vary if the methodology is applied to other object distribution settings.

b) The study developed a smart contract based on the Secret Network blockchain that insures both transactional and computational privacy. Despise the key to achieving computational privacy being the use of TEEs has been found to have vulnerabilities that could expose sensitive data.

c) The methodology created to manage the access of different entities to the blockchain is based on an off-chain solution that is ruled by a governing body. This governing body can be seen as a centralised authority, which makes it against the nature of a decentralized system.

Despite the limitations encountered by this study, the research has succeeded to provide a new form of building secure and private AR-based smartglasses assistive applications. It also succeeded, to some extent, in highlighting new findings in the field of object detection. Additionally, the study has also succeeded in demonstrating a low-cost solution for ensuring data privacy and security in a decentralized form.

7.4 Further research

Aside from the interesting research findings, several related areas were encountered, which would benefit from additional investigation. The following are proposals for future research:

a) Conduction additional research on building a new model capable of extracting the user cognitive load out of the eye function measure such as the pupil dilation, saccades and fixations. This model would be useful for a close understanding of the user's phycological state which results in a more accurate form of assistance in real-time.

b) As mentioned previously, the TEEs still have some vulnerabilities which can compromise the computational privacy feature of the blockchain. Further research could usefully be applied to the fully homomorphic encryption area. This would mitigate all the existent vulnerabilities and provide full transactional and computational privacy to the network.

c) Future research could be pointing to investigating new forms of accessibility to the blockchain. Mitigating the issue of lost or stolen private keys without the need for an off-chain centralised governing body authority would extend the decentralization nature of the whole network.

d) Building technologies for assistive living is a process that involves a series of studies to evaluate if the applications are robust to solve real-world problems and meet end-user requirements. In that sense, future research could be related to studies on evaluating the developed application on its user interface and task performance. This would include different users targeting ASD people as well as people with cognitive decline.

References

[1] M.W. Steege, F.C. Mace, L. Perry, H. Longenecker, Applied behavior analysis: Beyond discrete trial teaching, Psychol Sch. 44 (2007) 91–99.

[2] C.S. Rayner, Video-modelling to improve task completion in a child with autism, Dev Neurorehabil. 13 (2010) 225–230. https://doi.org/10.3109/17518421003801489.

[3] R. Liu, J.P. Salisbury, A. Vahabzadeh, N.T. Sahin, Feasibility of an Autism-Focused Augmented Reality Smartglasses System for Social Communication and Behavioral Coaching, Front Pediatr. 5 (2017) 1–8. https://doi.org/10.3389/fped.2017.00145.

[4] J.-J. Yang, J.-Q. Li, Y. Niu, A hybrid solution for privacy preserving medical data sharing in the cloud environment, Future Generation Computer Systems. 43–44 (2015) 74–86. https://doi.org/https://doi.org/10.1016/j.future.2014.06.004.

[5] K. Ren, C. Wang, Q. Wang, Security Challenges for the Public Cloud, IEEE Internet Comput. 16 (2012) 69–73. https://doi.org/10.1109/MIC.2012.14.

[6] Y. Liu, G. Liu, C. Cheng, Z. Xia, J. Shen, A Privacy-Preserving Health Data Aggregation Scheme, KSII Transactions on Internet and Information Systems. 10 (2016) 3852–3863. https://doi.org/10.3837/tiis.2016.08.023.

[7] G. Liu, Z. Yan, W. Feng, X. Jing, Y. Chen, M. Atiquzzaman, SeDID: An SGX-enabled decentralized intrusion detection framework for network trust evaluation, Information Fusion. 70 (2021) 100–114. https://doi.org/https://doi.org/10.1016/j.inffus.2021.01.003.

[8] F. Casino, T.K. Dasaklis, C. Patsakis, A systematic literature review of blockchain-based applications: Current status, classification and open issues, Telematics and Informatics. 36 (2019) 55–81. https://doi.org/https://doi.org/10.1016/j.tele.2018.11.006.

[9] C. Sullivan, E. Burger, E-residency and blockchain, Computer Law & Security Review. 33 (2017) 470–481. https://doi.org/https://doi.org/10.1016/j.clsr.2017.03.016.

[10] C. Esposito, M. Ficco, B.B. Gupta, Blockchain-based authentication and authorization for smart city applications, Inf Process Manag. 58 (2021) 102468. https://doi.org/https://doi.org/10.1016/j.ipm.2020.102468.

[11] E. Ben-Sasson, A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, M. Virza, Zerocash: Decentralized Anonymous Payments from Bitcoin (extended version), 2014.

[12] G. Zyskind Oz, N. Alex ', S.' Pentland, Enigma: Decentralized Computation Platform with Guaranteed Privacy, n.d.

[13] An Implementation of Ekiden on the Oasis Network Oasis Protocol Project, n.d.

[14] M.H. Calp, The importance of human- - computer interaction in the development process of software projects, 05 (2015) 48–54.

[15] J. Singer, M. Storey, D. Damian, Selecting Empirical Methods for Software Engineering Research, (2002).

[16] S. Goundar, Chapter 3 – Research Methodology and Research Method, (2019).

[17] O. Hazzan, Y. Dubinsky, L. Eidelman, V. Sakhnini, M. Teif, Qualitative Research in Computer Science Education, (2004) 0–4.

[18] O.D. Apuke, Arabian Journal of Business and Management Review (Kuwait Chapter), (2017). https://doi.org/10.12816/0040336.

[19] I. Physiological, O. Carlo, G. Gelo, D. Braakmann, G. Benetka, Quantitative and Qualitative Research : Beyond the Debate, (2009). https://doi.org/10.1007/s12124-009-9107-x.
[20] F. Shic, M. Goodwin, Introduction to Technologies in the Daily Lives of Individuals with Autism., J Autism Dev Disord. 45 (2015) 3773–3776. https://doi.org/10.1007/s10803-015-2640-1.
[21] F. Cutolo, B. Fida, Software Framework for Customized Augmented Reality Headsets in Medicine, 8 (2020).
[22] M.S. Pérez, M.C. Trinidad, D. Karatzas, A.C. Calaf, P.P.V. Alcocer, Development of general-purpose projection-based augmented reality systems, in: 2016.
[23] D.D. Han, J. Weber, Blowing your mind : a conceptual framework of augmented reality and virtual reality enhanced cultural visitor experiences using EEG experience measures Marcel Bastiaansen Ondrej Mitas and Xander Lub, 14 (2020).
[24] J. Bacca, F. Universitaria, K. Lorenz, S. Baldiris, R. Fabregat, Framework for designing motivational augmented reality applications in vocational education and training, 35 (2019) 102–117.
[25] S. Gargrish, A. Mantri, D.P. Kaur, Augmented Reality-Based Learning Environment to Enhance Teaching-Learning Experience in Geometry Education, Procedia Comput Sci. 172 (2020) 1039–1046. https://doi.org/https://doi.org/10.1016/j.procs.2020.05.152.
[26] I. Fernández del Amo, J.A. Erkoyuncu, R. Roy, S. Wilding, Augmented Reality in Maintenance: An information-centred design framework, Procedia Manuf. 19 (2018) 148–155. https://doi.org/https://doi.org/10.1016/j.promfg.2018.01.021.
[27] X. Wang, Z. Cai, D. Gao, N. Vasconcelos, Towards universal object detection by domain attention, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: pp. 7289–7298.
[28] S. Chen, X. Tan, B. Wang, X. Hu, Reverse attention for salient object detection, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018: pp. 234–250.
[29] Z. Zou, Z. Shi, Y. Guo, J. Ye, Object detection in 20 years: A survey, ArXiv Preprint ArXiv:1905.05055. (2019).
[30] M. Andoni, V. Robu, D. Flynn, S. Abram, D. Geach, D. Jenkins, P. McCallum, A. Peacock, Blockchain technology in the energy sector: A systematic review of challenges and opportunities, Renewable and Sustainable Energy Reviews. 100 (2019) 143–174.
[31] M. Kouhizadeh, S. Saberi, J. Sarkis, Blockchain technology and the sustainable supply chain: Theoretically exploring adoption barriers, Int J Prod Econ. 231 (2021) 107831.
[32] C.C. Agbo, Q.H. Mahmoud, J.M. Eklund, Blockchain technology in healthcare: a systematic review, in: Healthcare, MDPI, 2019: p. 56.
[33] D.L. Eggemeier, T. F., Wilson, G. F., Kramer, A. F., and Damos, Workload assessment in multi-task environments, 1991.
[34] E. Gopher, D. and Donchin, Workload: An examination of the concept., Handbook of Perception and Human Performance. 2, (n.d.) 41/1–41/49.
[35] E. Gopher, D. and Donchin, Workload: An examination of the concept., Handbook of Perception and Human Performance. 2, (n.d.) 41/1–41/49.

[36] L.E. Hart, S. G. and Staveland, Development of nasa-tlx (task load index): results of empirical and theoretical research, in: P.A. and M. Hancock (Ed.), Advances in Psychology, Human Ment, 1988: pp. 139–183.
[37] M.A. Tsang, P. S. and Vidulich, Mental workload and situation awareness., in: John Wiley & Sons (Ed.), Handbook of Human Factors and Ergonomics, 2006: pp. 243–268.
[38] Formalising Human Mental Workload as a Defeasible Computational Concept Formalising Human Mental Workload as a Defeasible Computational Concept, (2014).
[39] J. Sweller, Cognitive Load During Problem Solving : Effects on Learning, 285 (1988) 257–285.
[40] F.G.W.C.P.J.G. Van Merriënboer, Instructional control of cognitive load in the training of complex cognitive tasks, in: Educ Psychol Rev, 1994: p. pp 351–371 volume 6 issue 4.
[41] O. Bratfisch, Perceived Item-Difficulty in Three Tests of Intellectual Performance Capacity, (n.d.).
[42] R. Moreno, Does the modality principle hold for different media? A test of the method-affects-learning, J Comput Assist Learn. 22 (2006) 149–158.
[43] M. Bannert, Managing cognitive load, recent trends in cognitive load theory., Learn Instr. 12, (n.d.) 139–146.
[44] A.D. Hitch,G.J.,and Baddeley, Verbal reasoning and working memory., Exp.Psychol. 28 (1976) 603–621.
[45] N. Debue, What does germane load mean ? An empirical contribution to the cognitive load theory, 5 (2014) 1–12. https://doi.org/10.3389/fpsyg.2014.01099.
[46] K.E. Deleeuw, R.E. Mayer, A Comparison of Three Measures of Cognitive Load : Evidence for Separable Measures of Intrinsic , Extraneous , and Germane Load, 100 (2008) 223–234. https://doi.org/10.1037/0022-0663.100.1.223.
[47] N. Debue, What does germane load mean ? An empirical contribution to the cognitive load theory, 5 (2014) 1–12. https://doi.org/10.3389/fpsyg.2014.01099.
[48] P. Gerjets, K. Scheiter, R. Catrambone, P. Gerjets, K. Scheiter, R. Catrambone, D. Instructional, Designing Instructional Examples to Reduce Intrinsic Cognitive Load : Molar versus Modular Presentation of Solution Procedures To cite this version :, (2007).
[49] K. Scheiter, P. Gerjets, R. Catrambone, Making the abstract concrete : Visualizing mathematical solution procedures, 22 (2006) 9–25. https://doi.org/10.1016/j.chb.2005.01.009.
[50] U.P.O. Défense, The use of Tholos software for combining measures of mental workload : Toward theoretical, 40 (2008) 988–1000. https://doi.org/10.3758/BRM.40.4.988.
[51] G. Cierniak, K. Scheiter, P. Gerjets, Computers in Human Behavior Explaining the split-attention effect : Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load ?, 25 (2009) 315–324. https://doi.org/10.1016/j.chb.2008.12.020.
[52] R. Brünken, S. Steinbacher, J.L. Plass, D. Leutner, Assessment of Cognitive Load in Multimedia Learning Using Dual-Task Methodology, 49 (2002) 109–119. https://doi.org/10.1027//1618-3169.49.2.109.
[53] C.A. Shingledecker, Behavioral and subjective workload metrics for operational environments, Wright-Patterson AFB, Ohio 45433., 1983. https://doi.org/Technical report ADP002983,.
[54] E. O'Donnell F.T., R.D., Workload assessment methodology, 1986.

[55] I. Engineering, R. Labs, Cervicobrachial muscle response to cognitive load in a dual-task scenario, (2004) 625–645. https://doi.org/10.1080/00140130310001629766.
[56] I. Grabarek, P. Bartuzi, W. Choromański, The influence of mental load on muscle tension, 0139 (2017). https://doi.org/10.1080/00140139.2013.798429.
[57] R.T.O.T. Report, (L ' évaluation de l ' aptitude opérationnelle de l ' opérateur humain), 2004.
[58] D. Mcduff, S. Gontarek, R. Picard, Remote Measurement of Cognitive Stress via Heart Rate Variability, (n.d.) 3–6.
[59] Formalising Human Mental Workload as a Defeasible Computational Concept Formalising Human Mental Workload as a Defeasible Computational Concept, (2014).
[60] R.T.O.T. Report, (L ' évaluation de l ' aptitude opérationnelle de l ' opérateur humain), 2004.
[61] S. Chen, U. Sydney, Eye activity as a measure of human mental effort in HCI Eye activity as a measure of human mental effort in HCI, (2011). https://doi.org/10.1145/1943403.1943454.
[62] M. Pivec, J. Pripfl, C. Trummer, Adaptivno u no okolje na osnovi spremljanja pogleda in koncept oblikovanja u nih gradiv Eye-Tracking Adaptable e-Learning and Content Authoring Support, (2005) 1–6.
[63] R. Dewhurst, Eye Tracking : A Comprehensive Guide To Methods And Measures, (2011).
[64] V. Manuel, G. Barrios, C. Gütl, A.M. Preis, K. Andrews, F. Mödritscher, C. Trummer, AdELE : A Framework for Adaptive E-Learning through Eye Tracking, (n.d.) 1–8.
[65] T.Q. Journal, E. Psychology, J. Hy, Pupil Dilation as a Measure of Processing Load in Simultaneous Interpretation and Other Language Tasks, (2017). https://doi.org/10.1080/14640749508401407.
[66] P. Studies, Inferring User Cognitive Abilities from Eye-Tracking Data, (2015).
[67] M. Pomplun, Pupil dilation as an indicator of cognitive workload in human-computer interaction, (2017).
[68] M. Rucci, M. Poletti, Control and Functions of Fixational Eye Movements, Annu Rev Vis Sci. 1 (2015) 499–518. https://doi.org/10.1146/annurev-vision-082114-035742.
[69] M. Rucci, P. V McGraw, R.J. Krauzlis, Fixational eye movements and perception, Vision Res. 118 (2016) 1–4. https://doi.org/https://doi.org/10.1016/j.visres.2015.12.001.
[70] R. Mallick, D. Slayback, J. Touryan, A. Ries, B. Lance, The use of eye metrics to index cognitive workload in video games, 2016. https://doi.org/10.1109/ETVIS.2016.7851168.
[71] L.L. Di Stasi, A. Antolí, J.J. Cañas, Evaluating mental workload while interacting with computer-generated artificial environments, Entertain Comput. 4 (2013) 63–69. https://doi.org/https://doi.org/10.1016/j.entcom.2011.03.005.
[72] M.A. Recarte, L.M. Nunes, Effects of verbal and spatial-imagery tasks on eye fixations while driving., J Exp Psychol Appl. 6 (2000) 31.
[73] J.-C. Liu, K.-A. Li, S.-L. Yeh, S.-Y. Chien, Assessing Perceptual Load and Cognitive Load by Fixation-Related Information of Eye Movements, Sensors . 22 (2022). https://doi.org/10.3390/s22031187.

[74] J. Snell, S. Mathôt, J. Mirault, J. Grainger, Parallel graded attention in reading: A pupillometric study, Sci Rep. 8 (2018) 3743. https://doi.org/10.1038/s41598-018-22138-7.
[75] D. Amso, G. Scerif, The attentive brain: insights from developmental cognitive neuroscience, Nat Rev Neurosci. 16 (2015) 606.
[76] D. Rahnev, B. Maniscalco, T. Graves, E. Huang, F.P. de Lange, H. Lau, Attention induces conservative subjective biases in visual perception, Nat Neurosci. 14 (2011) 1513.
[77] A. Manuscript, T. Structures, Visual Attention and Applications in Multimedia Technologies Patrick, 6 (2009) 247–253. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial.
[78] M.A. Pedziwiatr, M. Kümmerer, T.S.A. Wallis, M. Bethge, C. Teufel, Meaning maps and saliency models based on deep convolutional neural networks are insensitive to image meaning when predicting human fixations, Cognition. 206 (2021) 104465.
[79] W. Wang, J. Shen, X. Dong, A. Borji, Salient object detection driven by fixation prediction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: pp. 1711–1720.
[80] B. Pfleging, D.K. Fekety, A. Schmidt, A.L. Kun, A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions, CHI Conference on Human Factors in Computing Systems. (2016) 5776–5788. https://doi.org/10.1145/2858036.2858117.
[81] L.R. Chapman, B. Hallowell, A Novel Pupillometric Method for Indexing Word Difficulty in Individuals With and Without Aphasia, Journal of Speech Language and Hearing Research. 58 (2015) 1508. https://doi.org/10.1044/2015_JSLHR-L-14-0287.
[82] S.P. Marshall, The Index of Cognitive Activity: measuring cognitive workload, Proceedings of the IEEE 7th Conference on Human Factors and Power Plants. (2002) 5–9. https://doi.org/10.1109/HFPP.2002.1042860.
[83] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain., Psychol Rev. 65 (1958) 386–408. https://doi.org/10.1037/h0042519.
[84] V. Kaul, S. Enslin, S.A. Gross, History of artificial intelligence in medicine, Gastrointest Endosc. 92 (2020) 807–812.
[85] Y. Lecun, L. Bottou, Y. Bengio, P. Ha, Gradient-Based Learning Applied to Document Recognition, (1998) 1–46.
[86] R. Cascade-correlation, N.S. Chunking, Long Short-Term Memory, 9 (1997) 1–32.
[87] C.J.C.H. Watkins, Q-Learning, 292 (1992) 279–292.
[88] J. Janai, A. Geiger, Computer Vision for Autonomous Vehicles, 2020. https://doi.org/10.1561/0600000079.Joel.
[89] K. Fukushima, S. Miyake, Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, in: Competition and Cooperation in Neural Nets, Springer, 1982: pp. 267–285.
[90] K. Xia, J. Huang, H. Wang, LSTM-CNN Architecture for Human Activity Recognition, IEEE Access. 8 (2020) 56855–56866. https://doi.org/10.1109/ACCESS.2020.2982225.
[91] S. Ha, Multi-Modal Convolutional Neural Networks for Activity Recognition, (2015) 3017–3022. https://doi.org/10.1109/SMC.2015.525.

[92] T. Chen, R. Xu, Y. He, X. Wang, Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN, Expert Syst Appl. 72 (2017) 221–230. https://doi.org/https://doi.org/10.1016/j.eswa.2016.10.065.
[93] G. Guo, N. Zhang, A survey on deep learning based face recognition, Computer Vision and Image Understanding. 189 (2019) 102805. https://doi.org/https://doi.org/10.1016/j.cviu.2019.102805.
[94] Z.-Q. Zhao, P. Zheng, S.-T. Xu, X. Wu, Object Detection With Deep Learning: A Review, IEEE Trans Neural Netw Learn Syst. 30 (2019) 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865.
[95] R. Padilla, S.L. Netto, E.A.B. Da Silva, A survey on performance metrics for object-detection algorithms, in: 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), IEEE, 2020: pp. 237–242.
[96] Y. Sun, B. Xue, M. Zhang, G.G. Yen, J. Lv, Automatically designing CNN architectures using the genetic algorithm for image classification, IEEE Trans Cybern. 50 (2020) 3840–3854.
[97] J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang, W. Xu, Cnn-rnn: A unified framework for multi-label image classification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: pp. 2285–2294.
[98] D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J Physiol. 160 (1962) 106.
[99] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature. 521 (2015) 436–444. https://doi.org/10.1038/nature14539.
[100] D. Mishkin, N. Sergievskiy, J. Matas, Systematic evaluation of CNN advances on the ImageNet, (2016). https://doi.org/10.1016/j.cviu.2017.05.007.
[101] X. He, F. Wang, W. Li, D. Sheng, Deep learning for efficient stochastic analysis with spatial variability, Acta Geotech. 17 (2022) 1031–1051.
[102] A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis, Deep learning for computer vision: A brief review, Comput Intell Neurosci. 2018 (2018).
[103] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int J Comput Vis. 60 (2004) 91–110.
[104] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Ieee, 2005: pp. 886–893.
[105] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE International Conference on Computer Vision, 2015: pp. 1440–1448.
[106] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: pp. 779–788.
[107] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: pp. 7263–7271.
[108] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, ArXiv Preprint ArXiv:1804.02767. (2018).
[109] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: Single shot multibox detector, in: European Conference on Computer Vision, Springer, 2016: pp. 21–37.
[110] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: pp. 770–778.

[111] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, ArXiv Preprint ArXiv:1409.1556. (2014).
[112] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, PMLR, 2015: pp. 448–456.
[113] H. Sattar, A. Bulling, M. Fritz, M. Planck, Predicting the Category and Attributes of Visual Search Targets Using Deep Gaze Pooling, (n.d.).
[114] G. Bertasius, H.S. Park, S.X. Yu, J. Shi, First Person Action-Object Detection with EgoNet, (2016).
[115] J. Steil, M.X. Huang, A. Bulling, Fixation detection for head-mounted eye tracking based on visual similarity of gaze targets, Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications - ETRA '18. (2018) 1–9. https://doi.org/10.1145/3204493.3204538.
[116] T. Toyama, T. Kieninger, F. Shafait, A. Dengel, Gaze guided object recognition using a head-mounted eye tracker, Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '12. (2012) 91. https://doi.org/10.1145/2168556.2168570.
[117] J. Klingner, Measuring cognitive load during visual tasks by combining pupillometry and eye tracking, Perspective. (2010) 130.
[118] M. Barz, D. Sonntag, Gaze-guided object classification using deep neural networks for attention-based computing, Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16. (2016) 253–256. https://doi.org/10.1145/2968219.2971389.
[119] A. Rogalska, P. Napieralski, The visual attention saliency map for movie retrospection, Open Physics. 16 (2018) 188–192. https://doi.org/10.1515/phys-2018-0027.
[120] G. Yildirim, S. Süsstrunk, Fasa: Fast, accurate, and size-aware salient object detection, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9005 (2015) 514–528. https://doi.org/10.1007/978-3-319-16811-1_34.
[121] J. Li, M.D. Levine, X. An, H. He, Saliency Detection Based on Frequency and Spatial Domain Analyses, in: BMVC, 2011.
[122] A. Oliva, A. Torralba, M.S. Castelhano, J.M. Henderson, Top-down control of visual attention in object detection, in: Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429), 2003: pp. I–253. https://doi.org/10.1109/ICIP.2003.1246946.
[123] J. Wang, A. Borji, C.-. J. Kuo, L. Itti, Learning a Combined Model of Visual Saliency for Fixation Prediction, IEEE Transactions on Image Processing. 25 (2016) 1566–1579. https://doi.org/10.1109/TIP.2016.2522380.
[124] Y. Jae, L. Kristen, C. V May, Y.J. Lee, Predicting Important Objects for Egocentric Video Summarization, (2015).
[125] G. Bertasius, H.S. Park, Exploiting Egocentric Object Prior for 3D Saliency Detection, (n.d.).
[126] B.J. Tamber-Rosenau, R. Marois, Central attention is serial, but midlevel and peripheral attention are parallel-A hypothesis, Atten Percept Psychophys. 78 (2016) 1874–1888. https://doi.org/10.3758/s13414-016-1171-y.
[127] A.T. Duchowski, A breadth-first survey of eye-tracking applications, Behavior Research Methods, Instruments, & Computers. 34 (2002) 455–470. https://doi.org/10.3758/BF03195475.
[128] S. Nakamoto, Bitcoin : A Peer-to-Peer Electronic Cash System, (n.d.) 1–9.

[129] D. Johnson, A. Menezes, S. Vanstone, The Elliptic Curve Digital Signature Algorithm (ECDSA), (2001) 36–63.
[130] W. Bi, X. Jia, M. Zheng, A Secure Multiple Elliptic Curves Digital Signature Algorithm for Blockchain, (n.d.).
[131] D. Vuji, D. Jagodi, S. Ran, Blockchain Technology , Bitcoin , and Ethereum : A Brief Overview, (2018) 21–23.
[132] A. Back, Hashcash - A Denial of Service Counter-Measure, (2002) 1–10.
[133] N.T. Courtois, M. Grajek, R. Naik, Optimizing SHA256 in Bitcoin Mining BT - Cryptography and Security Systems, in: Z. Kotulski, B. Księżopolski, K. Mazur (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2014: pp. 131–144.
[134] K. Suresh, An overview of randomization techniques: An unbiased assessment of outcome in clinical research, J Hum Reprod Sci. 4 (2011) 8–11. https://doi.org/10.4103/0974-1208.82352.
[135] D. Mingxiao, M. Xiaofeng, Z. Zhe, W. Xiangwei, C. Qijun, A review on consensus algorithm of blockchain, in: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017: pp. 2567–2572. https://doi.org/10.1109/SMC.2017.8123011.
[136] S. King, S. Nadal, PPCoin : Peer-to-Peer Crypto-Currency with Proof-of-Stake, (2012).
[137] D. Khan, L.T. Jung, M.A. Hashmani, Systematic literature review of challenges in blockchain scalability, Applied Sciences (Switzerland). 11 (2021). https://doi.org/10.3390/app11209372.
[138] M.N.M. Bhutta, A.A. Khwaja, A. Nadeem, H.F. Ahmad, M.K. Khan, M.A. Hanif, H. Song, M. Alshamari, Y. Cao, A Survey on Blockchain Technology: Evolution, Architecture and Security, IEEE Access. 9 (2021) 61048–61073. https://doi.org/10.1109/ACCESS.2021.3072849.
[139] R. Yang, R. Wakefield, S. Lyu, S. Jayasuriya, F. Han, X. Yi, X. Yang, G. Amarasinghe, S. Chen, Public and private blockchain in construction business process and information integration, Autom Constr. 118 (2020). https://doi.org/10.1016/j.autcon.2020.103276.
[140] B. Cao, X. Wang, W. Zhang, H. Song, Z. Lv, A Many-Objective Optimization Model of Industrial Internet of Things Based on Private Blockchain, IEEE Netw. 34 (2020) 78–83. https://doi.org/10.1109/MNET.011.1900536.
[141] D. Huang, X. Ma, S. Zhang, Performance Analysis of the Raft Consensus Algorithm for Private Blockchains, IEEE Trans Syst Man Cybern Syst. 50 (2020) 172–181. https://doi.org/10.1109/TSMC.2019.2895471.
[142] S. Banerjee, B. Bera, A.K. Das, S. Chattopadhyay, M.K. Khan, J.J.P.C. Rodrigues, Private blockchain-envisioned multi-authority CP-ABE-based user access control scheme in IIoT, Comput Commun. 169 (2021) 99–113. https://doi.org/https://doi.org/10.1016/j.comcom.2021.01.023.
[143] Y. Wang, A. Zhang, P. Zhang, H. Wang, Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain, Ieee Access. 7 (2019) 136704–136719.
[144] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, E. ben Hamida, Consortium blockchains: Overview, applications and challenges, International Journal On Advances in Telecommunications. 11 (2018) 51–64.
[145] K.N. Griggs, O. Ossipova, C.P. Kohlios, A.N. Baccarini, E.A. Howson, T. Hayajneh, Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring, J Med Syst. 42 (2018) 130. https://doi.org/10.1007/s10916-018-0982-x.

[146] L. Chen, W.-K. Lee, C.-C. Chang, K.-K.R. Choo, N. Zhang, Blockchain based searchable encryption for electronic health record sharing, Future Generation Computer Systems. 95 (2019) 420–429. https://doi.org/https://doi.org/10.1016/j.future.2019.01.018.
[147] X. Yue, H. Wang, D. Jin, M. Li, W. Jiang, Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control, J Med Syst. 40 (2016) 218. https://doi.org/10.1007/s10916-016-0574-6.
[148] B. Shen, J. Guo, Y. Yang, MedChain: Efficient Healthcare Data Sharing via Blockchain, Applied Sciences . 9 (2019). https://doi.org/10.3390/app9061207.
[149] D.L. Chaum, Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms, Commun. ACM. 24 (1981) 84–90. https://doi.org/10.1145/358549.358563.
[150] Q. Feng, D. He, S. Zeadally, M.K. Khan, N. Kumar, A survey on privacy protection in blockchain system, Journal of Network and Computer Applications. 126 (2019) 45–58.
[151] J. Bonneau, A. Narayanan, A.K. Miller, J. Clark, J.A. Kroll, E.W. Felten, Mixcoin: Anonymity for Bitcoin with Accountable Mixes, IACR Cryptol. EPrint Arch. 2014 (2014) 77.
[152] Dash is digital cash, (n.d.). https://www.dash.org.
[153] E. Heilman, L. AlShenibr, F. Baldimtsi, A. Scafuro, S. Goldberg, TumbleBit: An Untrusted Bitcoin-Compatible Anonymous Payment Hub, 2017. https://doi.org/10.14722/ndss.2017.23086.
[154] D. Stebila, L. Kuppusamy, J. Rangasamy, C. Boyd, J. Gonzalez Nieto, Stronger difficulty notions for client puzzles and denial-of-service-resistant protocols, in: Cryptographers' Track at the RSA Conference, Springer, 2011: pp. 284–301.
[155] A.A. Maksutov, M.S. Alexeev, N.O. Fedorova, D.A. Andreev, Detection of blockchain transactions used in blockchain mixer of coin join type, in: 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), IEEE, 2019: pp. 274–277.
[156] H. Corrigan-Gibbs, B. Ford, Dissent: accountable anonymous group messaging, in: Proceedings of the 17th ACM Conference on Computer and Communications Security, 2010: pp. 340–350.
[157] U. Feige, A. Fiat, A. Shamir, Zero-knowledge proofs of identity, Journal of Cryptology. 1 (1988) 77–94.
[158] M. Blum, P. Feldman, S. Micali, Non-interactive zero-knowledge and its applications, in: Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali, 2019: pp. 329–349.
[159] D. Hopwood, S. Bowe, T. Hornby, N. Wilcox, Zcash protocol specification, GitHub: San Francisco, CA, USA. (2016) 1.
[160] A. Kosba, A. Miller, E. Shi, Z. Wen, C. Papamanthou, Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts, in: 2016 IEEE Symposium on Security and Privacy (SP), 2016: pp. 839–858. https://doi.org/10.1109/SP.2016.55.
[161] Y. Wang, Octonion algebra and noise-free fully homomorphic encryption (FHE) schemes, ArXiv EPrint Archive Cornell University Library. (2016).
[162] J. Fan, F. Vercauteren, Somewhat practical fully homomorphic encryption, Cryptology EPrint Archive. (2012).
[163] A. Kipnis, E. Hibshoosh, Efficient methods for practical fully homomorphic symmetric-key encrypton, randomization and verification, Cryptology EPrint Archive. (2012).

[164] I. Editor, A. El-Yahyaoui, Fully Homomorphic Encryption: State of Art and Comparison, (2016). https://doi.org/10.6084/M9.FIGSHARE.3362338.
[165] S. Singh, S. Rathore, O. Alfarraj, A. Tolba, B. Yoon, A framework for privacy-preservation of IoT healthcare data using Federated Learning and blockchain technology, Future Generation Computer Systems. 129 (2022) 380–388.
[166] Z. Sisi, A. Souri, Blockchain technology for energy-aware mobile crowd sensing approaches in Internet of Things, Transactions on Emerging Telecommunications Technologies. (2021) e4217.
[167] D. Lee, D. Kohlbrenner, S. Shinde, K. Asanović, D. Song, Keystone: An open framework for architecting trusted execution environments, in: Proceedings of the Fifteenth European Conference on Computer Systems, 2020: pp. 1–16.
[168] B.C. Xing, M. Shanahan, R. Leslie-Hurd, Intel® Software Guard Extensions (Intel® SGX) Software Support for Dynamic Memory Allocation inside an Enclave, in: Proceedings of the Hardware and Architectural Support for Security and Privacy 2016, Association for Computing Machinery, New York, NY, USA, 2016. https://doi.org/10.1145/2948618.2954330.
[169] G. Zyskind, O. Nathan, A. Pentland, Enigma: Decentralized computation platform with guaranteed privacy, ArXiv Preprint ArXiv:1506.03471. (2015).
[170] K. Fan, S. Wang, Y. Ren, H. Li, Y. Yang, MedBlock: Efficient and Secure Medical Data Sharing Via Blockchain, J Med Syst. 42 (2018) 136. https://doi.org/10.1007/s10916-018-0993-7.
[171] A. Azaria, A. Ekblaw, T. Vieira, A. Lippman, MedRec: Using Blockchain for Medical Data Access and Permission Management, in: 2016 2nd International Conference on Open and Big Data (OBD), 2016: pp. 25–30. https://doi.org/10.1109/OBD.2016.11.
[172] S. Niu, M. Song, L. Fang, F. Yu, S. Han, C. Wang, Keyword search over encrypted cloud data based on blockchain in smart medical applications, Comput Commun. 192 (2022) 33–47. https://doi.org/https://doi.org/10.1016/j.comcom.2022.05.018.
[173] C. Service, P. Via, MeDShare : Trust-less Medical Data Sharing Among, (2017). https://doi.org/10.1109/ACCESS.2017.2730843.
[174] R. Johari, V. Kumar, K. Gupta, D.P. Vidyarthi, BLOSOM: BLOckchain technology for Security Of Medical records, ICT Express. 8 (2022) 56–60. https://doi.org/https://doi.org/10.1016/j.icte.2021.06.002.
[175] A. al Mamun, Md.U. Faruk Jahangir, S. Azam, M.S. Kaiser, A. Karim, A Combined Framework of InterPlanetary File System and Blockchain to Securely Manage Electronic Medical Records, in: M.S. Kaiser, A. Bandyopadhyay, M. Mahmud, K. Ray (Eds.), Proceedings of International Conference on Trends in Computational and Cognitive Engineering, Springer Singapore, Singapore, 2021: pp. 501–511.
[176] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, M. Abid, HealthBlock: A secure blockchain-based healthcare data management system, Computer Networks. 200 (2021) 108500.
[177] H.S.A. Fang, T.H. Tan, Y.F.C. Tan, C.J.M. Tan, Blockchain personal health records: systematic review, J Med Internet Res. 23 (2021) e25094.
[178] S. Shi, D. He, L. Li, N. Kumar, M.K. Khan, K.-K.R. Choo, Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey, Comput Secur. 97 (2020) 101966.

[179] G. Kappos, H. Yousaf, M. Maller, S. Meiklejohn, An empirical analysis of anonymity in zcash, in: 27th USENIX Security Symposium (USENIX Security 18), 2018: pp. 463–477.
[180] E. Ben-sasson, A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, M. Virza, Zerocash : Decentralized Anonymous Payments from Bitcoin, (2014) 459–474. https://doi.org/10.1109/SP.2014.36.
[181] N. van Saberhagen, CryptoNote v 2.0, in: 2013.
[182] E. Machado, I. Carrillo, L. Chen, Visual Attention-Based Object Detection in Cluttered Environments, (n.d.).
[183] B.V. Buterin, A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM, (2009) 1–36.
[184] A. Shamir, H o w to Share a Secret, (1979) 612–613.
[185] J. Daemen, Note on naming Rijndael Note on naming Rijndael, 2003.
[186] A. Morgan, R. Pass, On the Security Loss of Unique Signatures, in: A. Beimel, S. Dziembowski (Eds.), Theory of Cryptography, Springer International Publishing, Cham, 2018: pp. 507–536.
[187] J. Kwon, E. Buchman, Cosmos whitepaper, A Netw. Distrib. Ledgers. (2019).
[188] Tendermint, (n.d.).
[189] https://v1.cosmos.network/intro, (n.d.).
[190] E.B. Jae Kwon, Cosmos, n.d.
[191] C. Woetzel, Secret Network : A Privacy-Preserving Secret Contract & Decentralized Application Platform, (2020) 1–16.
[192] M. Sabt, M. Achemlal, A. Bouabdallah, M. Sabt, M. Achemlal, A. Bouabdallah, T. Execution, M. Sabt, M. Achemlal, A. Bouabdallah, Trusted Execution Environment : What It is , and What It is Not To cite this version : HAL Id : hal-01246364 Trusted Execution Environment : What It Is , and What It Is Not, (2015).
[193] R. Belchior, A. Vasconcelos, S. Guerreiro, A Survey on Blockchain Interoperability : Past , Present , and Future Trends, 54 (2021).
[194] L.M. Bach, B. Mihaljevic, M. Zagar, Comparative analysis of blockchain consensus algorithms, in: 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Ieee, 2018: pp. 1545–1550.
[195] C. Cachin, Asymmetric distributed trust, in: International Conference on Distributed Computing and Networking 2021, 2021: p. 3.
[196] H. Krawczyk, Cryptographic Extraction and Key Derivation: The HKDF Scheme, 2010.
[197] D.J. Bernstein, Curve25519: new Diffie-Hellman speed records, n.d.
[198] Dr.S. Vasundhara, Elliptic curve Cryptography and Diffie- Hellman Key exchange, IOSR Journal of Mathematics. 13 (2017) 56–61. https://doi.org/10.9790/5728-1301015661.
[199] P. Rogaway, T. Shrimpton, The SIV Mode of Operation for Deterministic Authenticated-Encryption (Key Wrap) and Misuse-Resistant Nonce-Based Authenticated-Encryption, n.d.

Appendix

A.1 Smart Contract Implementation on Secret Network

```
use cosmwasm_std::{
    to_binary, Binary, Env, Extern, HandleResponse, InitResponse,
Querier, StdError,
    StdResult, Storage,HumanAddr,
Api,log,CanonicalAddr,HandleResult,InitResult,QueryResult,Uint128
};

use secret_toolkit::utils::{pad_handle_result, pad_query_result};
use std::convert::TryFrom;
use crate::msg::{HandleMsg, InitMsg, QueryMsg, HandleAnswer,
QueryAnswer};
use crate::state::{remove, load, may_load, save, State,
metadata,Teste,CONFIG_KEY2,CONFIG_KEY };

pub const BLOCK_SIZE: usize = 256;

pub fn init<S: Storage, A: Api, Q: Querier>(
    deps: &mut Extern<S, A, Q>,
    env: Env,
    msg: InitMsg,
) -> StdResult<InitResponse> {
    let max_size = match valid_max_size(msg.max_size) {
        Some(v) => v,
        None => return Err(StdError::generic_err("Invalid max_size.
Must be in the range of 1..65535."))
    };

    let sender_address =
deps.api.canonical_address(&env.message.sender)?;


    let config = State {
        max_size,
        metadata_count: 0_u64,
        contract_info: env.contract.address,
    };

    save(&mut deps.storage, CONFIG_KEY, &config)?;
    Ok(InitResponse::default())
}

// limit the max message size to values in 1..65535
fn valid_max_size(val: i32) -> Option<u16> {
    if val < 1 {
        None
    } else {
        u16::try_from(val).ok()
    }
}

pub fn handle<S: Storage, A: Api, Q: Querier>(
    deps: &mut Extern<S, A, Q>,
    env: Env,
    msg: HandleMsg,
```

```
) -> HandleResult {

    let response = match msg {
        HandleMsg::Record { metadata } => try_record(deps, env,
metadata),
        HandleMsg::Read { } => try_read(deps, env),
    };

    pad_handle_result(response, BLOCK_SIZE)
}

fn try_record<S: Storage, A: Api, Q: Querier>(
    deps: &mut Extern<S, A, Q>,
    env: Env,
    metadata: String,
) -> StdResult<HandleResponse> {
    let mut status: String;

    let metadata = metadata.as_bytes();
    let sender_address =
deps.api.canonical_address(&env.message.sender)?;
    // retrieve the config state from storage
    let mut config: State = load(&mut deps.storage, CONFIG_KEY)?;

    let  result3: Option<Teste> = may_load(&mut deps.storage,
&sender_address.as_slice().to_vec()).ok().unwrap();

    let sender_address =
deps.api.canonical_address(&env.message.sender)?;

    if metadata.len() > config.max_size.into() {
        // if metadata content is too long, set status message and do
nothing else
        status = String::from("Message is too long. metadata not
recorded.");

    } else {

        match result3{

            Some (mut value2) => {

                let stored_metadata = metadata {
                    content: metadata.to_vec(),
                    timestamp: env.block.time
                };

                value2.metadata_vec_struct.push(stored_metadata);
                //result3.metadata_vec_struct.push(stored_metadata);
                save(&mut deps.storage,
&sender_address.as_slice().to_vec(), &value2)?;

            }
            None => {

                let stored_metadata = metadata {
                    content: metadata.to_vec(),
                    timestamp: env.block.time
                };
```

```
                let test = Teste{
                    metadata_vec_struct: vec![stored_metadata]
                };

                save(&mut deps.storage,
&sender_address.as_slice().to_vec(), &test)?;

            }

        };

        // increment the metadata_count
        config.metadata_count += 1;

        save(&mut deps.storage, CONFIG_KEY, &config)?;

        // set the status message
        status = String::from("metadata recorded!");
    }
    //status = String::from("Message is too long. metadata not
recorded.");
    // Return a HandleResponse with the appropriate status message
included in the data field
    Ok(HandleResponse {
        messages: vec![],
        log: vec![],
        data: Some(to_binary(&HandleAnswer::Record {
            status,
        })?),
    })
}

fn try_read<S: Storage, A: Api, Q: Querier>( deps: &mut Extern<S, A,
Q>,env: Env)
-> StdResult<HandleResponse>  {

    let  sender_address =
deps.api.canonical_address(&env.message.sender)?;
    let result2: Option<Teste> = may_load(&mut deps.storage,
&sender_address.as_slice().to_vec()).ok().unwrap();
    let mut status: String;
    let mut timestamp: Vec<u64> = vec![0];
    let mut user: Option<String>;
    let mut metadata : Vec<Option<String>>= vec![None];
    let mut tempo : u64;

    status = String::from("metadata found.");

    match result2{

        Some(h) => {

        for entry in (h.metadata_vec_struct).into_iter(){

            let iu= entry.content;
            tempo = entry.timestamp;

            user = String::from_utf8(iu).ok();
            //timestamp = Some(entry.timestamp);
            metadata.push(user);
```

```
            timestamp.push(tempo);

            //status = String::from("metadata ai edu found2.");
            //rest = out_response(deps,metadata,timestamp);

            //status = String::from("metadata ai edu found merda.");

        }

    }
            None => {

                status = String::from("metadata not found.");

        }

    };

    Ok(HandleResponse {messages: vec![],log: vec![],
        data: Some(to_binary(&HandleAnswer::Read {
            status,
            metadata,
            timestamp,
        })?),
    })

}
// pub fn out_response<S: Storage, A: Api, Q: Querier>(_deps: &mut
Extern<S, A, Q>,
//     metadata: Option<String>, timestamp: Option<u64>) ->
Vec<Option<String>>{

//     let status = String::from("metadata ai ai ai found.");
//     let mut final_vec : Vec<Option<String>>;
//     //let ass : Option<String>= None;

//     final_vec = vec![None];

//     final_vec.push(metadata);

//     return final_vec;

//     // Ok(HandleResponse {messages: vec![],log: vec![],
//     //     data: Some(to_binary(&HandleAnswer::Read {
//     //         status,
//     //         metadata,
//     //         timestamp,
//     //     })?),
//     // })
// }

pub fn query<S: Storage, A: Api, Q: Querier>(
    deps: &Extern<S, A, Q>,
    msg: QueryMsg
) -> StdResult<Binary> {
    match msg {
        QueryMsg::Stats { } => query_stats(deps)
    }
}
```

```
fn query_stats<S: Storage, A: Api, Q: Querier>(deps: &Extern<S, A, Q>)
-> StdResult<Binary> {

    // retrieve the config state from storage
    let config: State = load(&deps.storage, CONFIG_KEY)?;


    to_binary(&QueryAnswer::Stats{ metadata_count:
config.metadata_count,contract_info_address: config.contract_info})
}
```

A.2 State of the Smart Contract Implementation on Secret Network

```
use std::{any::type_name, collections::HashSet};
use serde::{de::DeserializeOwned, Deserialize, Serialize};

use cosmwasm_std::{ Storage, ReadonlyStorage, StdResult,
StdError,Binary, CosmosMsg, HumanAddr, Querier, Uint128};


use secret_toolkit::serialization::{Bincode2, Serde};

use crate::msg::ContractInfo;

pub static CONFIG_KEY: &[u8] = b"config";

pub static CONFIG_KEY2: &[u8] =
b"secret1anf7y546qxmeae7a6d5m6l9ceraptnxh3jd0py";



#[derive(Serialize, Deserialize)]
pub struct State {
    pub max_size: u16,
    pub metadata_count: u64,
    pub contract_info: HumanAddr,

}

#[derive(Serialize, Deserialize, Clone, Debug, PartialEq)]
pub struct metadata {
    pub content: Vec<u8>,
    pub timestamp: u64,
}

#[derive(Serialize, Deserialize, Clone, Debug, PartialEq)]
pub struct Teste{
     pub metadata_vec_struct: Vec<metadata>,
}

pub fn remove<S: Storage>(storage: &mut S, key: &[u8]) {
    storage.remove(key);
}

pub fn save<T: Serialize, S: Storage>(storage: &mut S, key: &[u8],
value: &T) -> StdResult<()> {
    storage.set(key, &Bincode2::serialize(value)?);
```

```
    Ok(())
}

pub fn load<T: DeserializeOwned, S: ReadonlyStorage>(storage: &S, key:
&[u8]) -> StdResult<T> {
    Bincode2::deserialize(
        &storage
            .get(key)
            .ok_or_else(|| StdError::not_found(type_name::<T>()))?,
    )
}

pub fn may_load<T: DeserializeOwned, S: ReadonlyStorage>(storage: &S,
key: &[u8]) -> StdResult<Option<T>> {
    match storage.get(key) {
        Some(value) => Bincode2::deserialize(&value).map(Some),
        None => Ok(None),
    }
}
```

A.3 Implementation of Ethereum Smart Contract

```
pragma solidity ^0.4.6 <0.7.0;

contract Mapping
{
    string public imagem;
    uint256 public peopleCount;
    string public objecto;
    string public metadata;
    uint public leng;
    string public imga;
    string public obbj;
    uint public i;
    bytes32[5] public frase;
    bytes[] bytesArray;


    struct Entity{

        uint index;
        string metadata;
        string  imagem;
        string  objecto;
    }



    mapping(address => Entity[]) public values;

    function addValuesArray(address a, string img, string obj) public{

        incrementCount();
        values[a].push(Entity(peopleCount,img,obj, metadata));

    }
```

```
    function incrementCount() internal{

        peopleCount += 1;

    }



    function getLenght(address ad) public returns(uint){

        leng = values[ad].length;
        return leng;

    }

    function getValue(address ad, uint l) public returns(bytes) {

        //imagem = values[ad].img;
        //objecto = values[a].obj;
        uint x=0;
        bytes32[5]  frases;
        leng = values[ad].length;

        for ( i=0; i<leng; i++) {

           imga = values[ad][i].imagem;
           obbj = values[ad][i].objecto;

           bytes memory b3 = bytes(imga);
           bytesArray[x] = "b3";
           x++;
           //b[] = string_tobytes(imga,obbj,x);

         }
         return (bytesArray[0]);
        //Entity memory e = values[ad];
    }

    function getValues(address ad, uint l) public returns(string,
string, string) {


        imga = values[ad][i].imagem;
        obbj = values[ad][i].objecto;
        metadata = values[ad][i].metadata;

        return(imga, obbj, metadata);

        //Entity memory e = values[ad];
    }



     function string_tobytes(string s, string a, uint x) constant
returns (bytes){


        uint d=0;
        x--;
        bytes memory b3 = bytes(s);
```

```
        bytes memory b4 = bytes(a);

        bytesArray[x] = b3;

        bytesArray2[x] = b4;

        x++;

        return (bytesArray);
    }

}
```

A.4 Implementation of Visual Attention-base Object Detection

```
import numpy as np
import cv2
import tensorflow as tf

from object_detection.utils import label_map_util
from object_detection.utils import visualization_utils as vis_util


# Path to frozen detection graph.
# This is the actual model that is used for the object detection.
PATH_TO_CKPT = 'model/frozen_inference_graph.pb'

# List of the strings that is used to add correct label for each box.
PATH_TO_LABELS = 'model/mscoco_label_map.pbtxt'
NUM_CLASSES = 90

# Loading label map
label_map = label_map_util.load_labelmap(PATH_TO_LABELS)
categories = label_map_util.convert_label_map_to_categories(
                            label_map,
                            max_num_classes=NUM_CLASSES,
                            use_display_name=True)
category_index = label_map_util.create_category_index(categories)


def detect_objects(image_np, sess, detection_graph,
                   norm_pos, height, width):
    """ Method to detect objects using tensorflow model.

    It works using the frozen graph given and return
    the processed image in addition to the array of
    objects detected in the image.

    This method is called by the worker.

    Args:
        - image_np (Image): Image needed to process.
        - sess (session): Tensorflow model session.
        - detection_graph (graph): Tensorflow graph of the model.
        - norm_pos (list): List of the fixation position (x,y).
        - height (int): Height of the input frame.
        - width (int): Width of the input frame.

    Returns:
```

```
        - image_np (Image): Image processed.
        - objects_detected (list): List of objects detected data.
        - object_detected (string): Name of the Object where fixation
points.

    """

    # Expand dimensions since the model expects images
    # to have shape: [1, None, None, 3]
    image_np_expanded = np.expand_dims(image_np, axis=0)
    image_tensor =
detection_graph.get_tensor_by_name('image_tensor:0')

    # Each box represents a part of the image
    # where a particular object was detected.
    boxes = detection_graph.get_tensor_by_name('detection_boxes:0')

    # Each score represent how level of confidence for each of the
objects.
    # Score is shown on the result image, together with the class
label.
    scores = detection_graph.get_tensor_by_name('detection_scores:0')
    classes =
detection_graph.get_tensor_by_name('detection_classes:0')
    num_detections =
detection_graph.get_tensor_by_name('num_detections:0')

    # Actual detection
    (boxes, scores, classes, num_detections) = sess.run(
        [boxes, scores, classes, num_detections],
        feed_dict={image_tensor: image_np_expanded})

    # Convert results into arrays
    boxes_m = np.squeeze(boxes)
    classes_m = np.squeeze(classes).astype(np.int32)
    scores_m = np.squeeze(scores)

    # Visualization of the results of a detection
    vis_util.visualize_boxes_and_labels_on_image_array(
        image_np,
        np.squeeze(boxes),
        np.squeeze(classes).astype(np.int32),
        np.squeeze(scores),
        category_index,
        use_normalized_coordinates=True,
        line_thickness=4)

    # Build the list of dictionaries with objects detected data
    objects_detected = create_objects_detected_list(boxes_m,
classes_m,
                                                    scores_m)

    # Return object detected by the fixation
    object_detected = detect_fixation(boxes_m, classes_m,
                                      scores_m, norm_pos,
                                      height, width)

    return image_np, objects_detected, object_detected
```

```
def detect_fixation(boxes_m, classes_m, scores_m, norm_pos, height,
width):
    """ Return what object the fixation is pointing

    If there is no object detected in the fixation, it returns
    None as a value.
    Args:
        - boxes_m (list):
        - classes_m (list):
        - scores_m (list):
        - norm_pos (list):
        - height (int):
        - width (int):

    Returns:
        - detected (string):

    """

    # Create variables needed
    detect = []
    ymin, xmin, ymax, xmax = [], [], [], []
    index = None
    object_detected = None
    posx, posy = 0, 0

    # Convert normalized position into real position
    if norm_pos is not None:
        posx = norm_pos[0]*width
        posy = norm_pos[1]*height

    # Loop over object detected with a max of 5
    for i in range(min(5, boxes_m.shape[0])):
        # If the score of this object detected is up 0.5
        if scores_m[i] > 0.5:
            # If the class is inside our categories
            if classes_m[i] in category_index.keys():
                # Extract this object data
                object_class = category_index[classes_m[i]]['name']
                ymin_norm, xmin_norm = boxes_m[i][0], boxes_m[i][1]
                ymax_norm, xmax_norm = boxes_m[i][2], boxes_m[i][3]

                # Append the data to auxiliar lists
                detect.append(object_class)
                ymin.append(int(ymin_norm*height))
                xmin.append(int(xmin_norm*width))
                ymax.append(int(ymax_norm*height))
                xmax.append(int(xmax_norm*width))

    # If any of the object fit with the fixation pos
    for i in range(len(detect)):
        if((ymin[i] < posy < ymax[i]) and (xmin[i] < posx < xmax[i])):
            index = i
            break

    if detect is not None and index is not None:
        object_detected = detect[index]

    # Return the object detected pointing by the fixation or None
    return object_detected
```

```
def create_objects_detected_list(boxes_m, classes_m, scores_m):
    """ Creates an object detected list.

    It creates a list with the most important objects
    detected in the image.

    Args:
        - boxes_m (list): List of boxes data.
        - classes_m (list): List of object detected classes.
        - scores_m (list): List of object detected scores.

    Returns:
        - objects_detected (list): List with the most important
features of
                                   each object detected.

    """

    objects_detected = []

    # Loop over object detected with a max of 5
    for i in range(min(5, boxes_m.shape[0])):
        # If the score of this object detected is up 0.5
        if scores_m[i] > 0.5:
            # If the class is inside our categories
            if classes_m[i] in category_index.keys():
                object_class = category_index[classes_m[i]]['name']
                ymin_norm, xmin_norm = boxes_m[i][0], boxes_m[i][1]
                ymax_norm, xmax_norm = boxes_m[i][2], boxes_m[i][3]
                object_score = scores_m[i]

                # Append relevant data into list
                objects_detected.append({"class": object_class,
                                         "score": object_score,
                                         "ymin": ymin_norm,
                                         "xmin": xmin_norm,
                                         "ymax": ymax_norm,
                                         "xmax": xmax_norm})
    return objects_detected


def worker(input_q, output_q):
    """ Object detection main Engine.

    It uses Tensorflow frozen model Mobilenet
    to detect objects in the frame.
    Also it detects what object the fixation is pointing.

    Args:
        input_q (queue): Input Queue of the process. It is composed:
            - frame_timestamp (Timestamp): Timestamp of the image.
            - frame (Image): Image to detect objects.
            - fixation_id (int): Id of the pupil fixation.
                                 If the fixation is pointing
                                 to the same time this ID is gonna
                                 be the same in both fixations.
            - fixation_timestamp (timestamp): Timestamp of the
fixation.
```

```
            - fixation_duration (time): Fixation duration time in
milliseconds.
            - norm_pos (list): Position of the current fixation.
        output_q (queue): Output Queue of the process. It is composed:
            - frame_timestamp (Timestamp): Timestamp of the image.
            - image (Image): Image processed. Bounding boxes drawed.
            - objects_detected (list): List with Objects detected
info.
            - detected (string): Name of the Object where fixation
points-
            - fixation_id (int): Id of the pupil fixation.
                                 If the fixation is pointing
                                 to the same time this ID is gonna
                                 be the same in both fixations.
            - fixation_timestamp (Timestamp): Timestamp of the
fixation.
            - fixation_duration (time): Fixation duration time in
milliseconds.
            - norm_pos (list): Position of the current fixation.

    """

    # Load a (frozen) Tensorflow model into memory.
    detection_graph = tf.Graph()
    with detection_graph.as_default():
        od_graph_def = tf.GraphDef()
        with tf.gfile.GFile(PATH_TO_CKPT, 'rb') as fid:
            serialized_graph = fid.read()
            od_graph_def.ParseFromString(serialized_graph)
            tf.import_graph_def(od_graph_def, name='')
        sess = tf.Session(graph=detection_graph)

    # Loop of the process
    while True:
        # Get data from queue
        input_message = input_q.get()

        # Split list of data
        if len(input_message) == 6:
            (frame_timestamp, frame, fixation_id,
             fixation_timestamp, fixation_duration, norm_pos) =
input_message

            height, width, depth = frame.shape

            frame_rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)

            # Pass image to detection engine
            image, objects_detected, detected = detect_objects(
                                                frame_rgb, sess,
                                                detection_graph,
norm_pos,
                                                height, width)

            # Return data into the queue
            output_q.put([frame_timestamp, image, objects_detected,
                          detected, fixation_id, fixation_timestamp,
                          fixation_duration, norm_pos])

    # When finish, close the session
```

```
    sess.close()
```

A.4 Integration of IPFS protocol with SSS and Connection to Ethereum Smart Contract

```
def get_images(out_q):
    #logging.info("Thread %s: starting", name)


    onlyfiles = [f for f in listdir(mypath) if isfile(join(mypath,
f))]

    for i in range(len(onlyfiles)):

            time.sleep(0.033333)
            ol= out_q.put(onlyfiles[i])
            #print("queue", ol)

    #logging.info("Thread %s: finishing", name)


def consumer(in_q):
    while True:
        # Get some data
        data = in_q.get()
        ipfs(data)


        # Process the data

def ipfs(data):

    ####### IPSF PROTOCOL FOR IMAGES OR TXT #######
    #####https://pypi.org/project/ipfshttpclient/#usage

    client = ipfshttpclient.connect()  # Connects to:
/dns/localhost/tcp/5001/http
    #res = client.add("../Video2/image1.jpg")

    res = client.add("../Video2/"+data)

    ####OBTEN THE HAHS FROM IPFS PROTOCOL#####

    x= client.cat(res['Hash'])

    hashx = res['Hash']
    print("SendImage->hash", hashx)
    secretshares(hashx)

    #####CHECK THE FILE IN IPFPS

###http://127.0.0.1:5001/ipfs/bafybeigkbbjnltbd4ewfj7elajsbnjwinyk6tii
lczkqsibf3o7dcr6nn4/#/explore

    #return hashx

def secretshares(x):
```

```
    ####ENCRYPTION OF THE HASH FROM IPFS, SPECIFICLY THE USE O SECRET
SHARES#################

    shares = PlaintextToHexSecretSharer.split_secret(x, 2, 3)
    print ("SendImage->Secret Shares:", shares)

    ####RECOVER THE SECRET VY USING THE MAJORATY OF SHARES###########

    sh=PlaintextToHexSecretSharer.recover_secret(shares[0:2])
    print ("SendImage->Recovered Secrets:", sh)

    tests(shares)

if __name__ == "__main__":
    format = "%(asctime)s: %(message)s"
    logging.basicConfig(format=format, level=logging.INFO,
                        datefmt="%H:%M:%S")

    logging.info("Main    : before creating thread")
    #x = threading.Thread(target=thread_function, args=(1,))
    logging.info("Main    : before running thread")

    q = Queue()
    t1 = Thread(target = get_images, args =(q, ))
    t2 = Thread(target = consumer, args =(q, ))
    t1.start()
    t2.start()


    #x.start()
    logging.info("Main    : wait for the thread to finish")
    # x.join()
    logging.info("Main    : all done")
```

A.5 Implementation of Data Integrating Module for Assistive framework

```
public class MqttClient {

    public MqttAndroidClient client;

    private static final String TAG = "PahoMqttClient";
    private static String mqttUrl;
    private static String clientId;

    private Context mContext;

    public MqttClient(Context context){
        mContext = context;
        mqttUrl = MqttConnectionConstants.getUrl();
        clientId = MqttConnectionConstants.getClientId();
        String m_androidId =
Settings.Secure.getString(context.getContentResolver(),
Settings.Secure.ANDROID_ID);

        if ( client == null ) {
            client = new MqttAndroidClient(context, mqttUrl,
m_androidId + "_" + clientId);
        }
```

```
        client.setCallback(new MqttCallbackExtended() {
            @Override
            public void connectComplete(boolean b, String s) {
            }

            @Override
            public void connectionLost(Throwable throwable) {
            }

            @Override
            public void messageArrived(String topic, MqttMessage
mqttMessage) throws Exception {
            }

            @Override
            public void deliveryComplete(IMqttDeliveryToken
iMqttDeliveryToken) {
            }
        });
//        connect();
    }

    public void setCallback(MqttCallbackExtended callback) {
        client.setCallback(callback);
    }

    public void connect(){

        try {
            client.connect(getMqttConnectionOption(), null, new
IMqttActionListener() {
                @Override
                public void onSuccess(IMqttToken asyncActionToken) {

client.setBufferOpts(getDisconnectedBufferOptions());
                }

                @Override
                public void onFailure(IMqttToken asyncActionToken,
Throwable exception) {
                    Log.w("Mqtt", "Failed to connect to: " + mqttUrl +
exception.toString());
                }
            });


        } catch (MqttException ex){
            ex.printStackTrace();
        }
    }

    public IMqttToken connectWithResult(){
        try {
            IMqttToken token =
client.connect(getMqttConnectionOption());
            return token;
        } catch (MqttException ex){
            ex.printStackTrace();
        }
        return null;
```

```
    }

    public void publishMessage(String msg, String topic) throws
MqttException, UnsupportedEncodingException {
        byte[] encodedPayload = msg.getBytes(StandardCharsets.UTF_8);
        MqttMessage message = new MqttMessage(encodedPayload);
        message.setId(320);

message.setRetained(MqttConnectionConstants.getRetainMessage());
        message.setQos(MqttConnectionConstants.getQoS());
        client.publish(topic, message);
    }

    public void subscribe(final String topic) {
        try {
            client.subscribe(topic, MqttConnectionConstants.getQoS(),
null, new IMqttActionListener() {
                @Override
                public void onSuccess(IMqttToken asyncActionToken) {
                    Log.w("Mqtt","Subscribed!");
                    Log.d(TAG, "Subscribe Successfully " + topic);
                }

                @Override
                public void onFailure(IMqttToken asyncActionToken,
Throwable exception) {
                    Log.w("Mqtt", "Subscribed fail!");
                }
            });

        } catch (MqttException ex) {
            System.err.println("Exceptionst subscribing");
            ex.printStackTrace();
        }
    }

    @NonNull
    private static MqttConnectOptions getMqttConnectionOption() {
        MqttConnectOptions mqttConnectOptions = new
MqttConnectOptions();

mqttConnectOptions.setCleanSession(MqttConnectionConstants.getCleanSes
sion());

mqttConnectOptions.setAutomaticReconnect(MqttConnectionConstants.getAu
tomaticReconnect());
        // mqttConnectOptions.setWill(Constants.PUBLISH_TOPIC,
MqttConnectionConstants.getWill().getBytes(), 1, true);
        //
mqttConnectOptions.setUserName(MqttConnectionConstants.getUsername());
        //
mqttConnectOptions.setPassword(MqttConnectionConstants.getPassword());
        return mqttConnectOptions;
    }

    @NonNull
    public DisconnectedBufferOptions getDisconnectedBufferOptions() {
        DisconnectedBufferOptions disconnectedBufferOptions = new
DisconnectedBufferOptions();
        disconnectedBufferOptions.setBufferEnabled(true);
```

```
        disconnectedBufferOptions.setBufferSize(100);
        disconnectedBufferOptions.setPersistBuffer(false);
        disconnectedBufferOptions.setDeleteOldestMessages(false);
        return disconnectedBufferOptions;
    }

    public void disconnect(){
        try {
            client.disconnect();
        } catch (MqttException e) {
            Log.d(TAG, "Error disconnecting " + e.toString());
        }
    }
}
```

A.6 Task Modeling Design on Assistive Framework

```
public class TaskModel {
      /**
       * task ID
       */
      private String id;

      /**
       * task title
       */
      private String title;

      /**
       * task description
       */
      private String description;

      /**
       * actions
       */
      private List<Step> steps;

      /**
       * Constructor
       */
      public TaskModel() {
            steps = new ArrayList<Step>();
      }

      /**
       * @return the id
       */
      public String getId() {
            return id;
      }

      /**
       * @param id
       */
      public void setId(String id) {
            this.id = id;
      }
```

```
	/**
	 * @return the title
	 */
	public String getTitle() {
		return title;
	}

	/**
	 * @param title
	 */
	public void setTitle(String title) {
		this.title = title;
	}

	/**
	 * @return the description
	 */
	public String getDescription() {
		return description;
	}

	/**
	 * @param description
	 */
	public void setDescription(String description) {
		this.description = description;
	}

	/**
	 * @return the steps
	 */
	public List<Step> getSteps() {
		return steps;
	}

	/**
	 * @param steps the actions to set
	 */
	public void setActions(List<Step> steps) {
		this.steps = steps;
	}


	@Override
	public String toString() {
		return "Step [id=" + id +  " title=" + title +  "
description=" + description +  " steps=" + steps.size() + "]";
	}

}
```

A.7 Action Modeling on Assistive Framework

```
public class ActionModel {
	/**
	 * description
	 */
	private String description;
```

```
	/**
	 * type
	 */
	private String img;

	/**
	 * Constructor
	 */
	public ActionModel() {
	}

	/**
	 * @return description
	 */
	public String getDescription() {
		return description;
	}

	/**
	 * set description
	 */
	public void setDescription(String description) {
		this.description = description;
	}

	/**
	 * @return img
	 */
	public String getImg() {
		return img;
	}

	/**
	 * set img
	 */
	public void setImg(String description) {
		this.img = img;
	}

	@Override
	public String toString() {
		return "Action [description=" + description +  " img=" +
(img != null) + "]";
	}

}
```