Thesis

MIXED REALITY APPLICATION FOR DISPLAYING ARTWORKS

Nikolas Constantinides

UNIVERSITY OF CYPRUS

DEPARTMENT OF COMPUTER SCIENCE

May 2024

UNIVERSITY OF CYPRUS DEPARTMENT OF COMPUTER SCIENCE

Mixed Reality Application for Displaying Artworks

Nikolas Constantinides

Supervisor

Dr. Vasos Vassiliou

The thesis was submitted in partial fulfilment of the requirements for the degree of Computer Science of the Department of Computer Science of the University of Cyprus.

May 2024

Acknowledgments

I am deeply grateful for the guidance and support I received during the development of this thesis.

I would like to express my deep gratitude to Dr. Belk for his expertise and encouragement in developing and completing this project.

Special thanks to Dr. Argyris Constantinides, whose guidance was very important in the direction of my work. His unique approach to problem solving and willingness to share his extensive knowledge were invaluable. His guidance not only influenced this dissertation, but also laid a solid foundation for my future endeavors.

I must also express my sincere gratitude to Dr. Vassiliou for his valuable help and support. His optimistic attitude and willingness to help was a source of motivation for the trip. His contribution was crucial to the realization of this project.

This journey would not have been possible without the support of my family.

This work is not only a reflection of my efforts, but also a testament to the valuable contributions of everyone who took part in this journey.

Abstract

The rapid advancements in technology have led to the emergence of many new opportunities for displaying artworks. One such technology is Mixed Reality (MR), which bridges the physical and digital realms, allowing users to manipulate and interact with virtual objects as if they were part of the real environment. In this context, the integration of Artificial Intelligence (AI) has introduced a new dimension for improving MR experiences, enabling users to interact with virtual environments using natural language.

The need for a system that could display artworks digitally was increased due to the coronavirus pandemic. During this time, protective measures against the virus made it impossible to display artworks in physical spaces. For example, some governments restricted gatherings to only a few people, and large gatherings were prohibited.

This thesis targets a system that runs on a MR environment, and can be deployed on Microsoft HoloLens (1st gen) devices. Also, a web application and a backend system were designed and implemented for demonstration purposes. The MR application allows users to view an artwork that is digitized. More specifically, the user has the option to view an artwork that is in the system, as well as all the associated media of a specific artwork. Associated media are either images, videos, or text and they could be multiple. Whereas a goal is viewing digitized artworks, another vital functionality is recognizing an artwork that is in the user's place using AI and showing all the associated information. The main goal was to develop the client-side part, but a server-side part was also developed for demonstration purposes.

Furthermore, the latest advancements of AI has opened up new possibilities for the Mixed Reality landscape, particularly Large Language Models (LLMs), which have yet to be fully explored. As a result, one of the goals successfully attained was to investigate the integration of LLMs into the MR environment. For doing so, a user study was conducted, which evaluated the user experience and perceived trustworthiness of integration of an LLM within the MR application. Results of the evaluation revealed the positive aspect of utilizing LLMs within MR applications in terms of user experience and perceived trustworthiness when individuals interact with virtual art exhibitions.

Content

Chapter 1

Introduction

1.1 Thesis overview

Given that Mixed Reality is a cutting-edge technology that improves significantly the user's experience, it could be used for displaying digitally artworks digitally. Also, the evolution of AI has brought the opportunity for integrating LLMs into the Mixed Reality environment. Because the effect of the LLMs to the user in the Mixed Reality is currently unexplored, a lot of effort was put for conducting a user evaluation study to investigate the integration of LLMs in the MR application.

After the coronavirus pandemic, the need for a system for displaying digital artworks was discovered. This has resulted in the development of this project, a system that aims to display digital artworks in a mixed reality environment. The program that was developed targets Mixed Reality technology that runs on the HoloLens 1 device.

More specifically, on the system was developed a Login System, a Gallery System that displays 2D artworks with the capability of viewing associated media that are either texts, or videos, or images. Moreover, an artwork recognition system was created that could match an image of an artwork with an artwork that is on the server and display to the user all the associated media. An extended gallery view system was developed for giving a more realistic and better experience to the user which illustrates the artworks in a different way. Also, a website for adding or deleting artworks and related associated media was designed and some endpoints for letting the client to get the list of the stored artworks with their related associated media and logging in.

Additionally, the integration of LLMs in the Mixed Reality environment was explored, under the HCI umbrella. For being able to measure the affection of the LLMs to the user, two versions of virtual galleries were constructed, one that supports voice-questions in the context of the artworks and returns an answer that is generated through an LLM and one that is not using any intelligent system. A total of 39 individuals took part in a user evaluation study which compared the two systems and evaluated the user experience and perceived trustworthiness when interacting with a virtual art exhibition.

1.2 Problem Statement

I recognized the limitations imposed by physical galleries and museums, including spatial and physical limitations, the static form of engagement, and the limited scope of interactive information within artworks. I find that these traditional methods often fail to capture the depth and context of the art, which diminishes the experience and understanding of the viewer.

In addition, I recognized a gap in the use of new technologies and art. Although digital platforms offer new possibilities for organizing art exhibitions, they rarely use the full potential of surrounding technologies such as Mixed Reality (MR). This gap represents a missed opportunity to increase art appreciation and education, and to make art accessible outside of traditional settings.

Additionally, the effect that LLMs are causing to the user in the MR were currently underexplored. This is the reason that an aim of this thesis was to explore how the users are experiencing this inventive technology in the MR environment.

My goal is to overcome the physical and interactive limitations of traditional art presentation methods, providing viewers with a more engaging and informative experience. This approach not only increases access and appreciation of art, but also addresses the broader challenge of effectively using new technologies to enhance cultural and educational experiences of art.

1.3 Mixed Reality

Mixed reality (MR) on the HoloLens device represents a fusion of the real and virtual world, delivering an immersive experience that fundamentally changes the way we interact with technology.

This innovative technology bridges the physical and digital realms, allowing users to manipulate and interact with virtual objects as if they were part of the real environment. One of the most appealing aspects of the MR environment on HoloLens is the user interaction. Users can interact with virtual elements naturally using gestures, voice commands, and eye tracking. This interaction paradigm creates a more organic and immersive experience because users can interact with holograms and digital objects as easily as they would with physical objects in the environment.

MR has been extensively used in various domains, including Education and Training [7, 2, 4], Architecture and Real Estate [5], Design and Visualization [17], Manufacturing and Engineering [13], Retail [9, 10, 37], Gaming and Entertainment [16], Military and Defense, Urban Planning and Development [18, 38, 39, 40], Interactive Exhibitions and Presentations, Law Enforcement and Forensic Analysis [20].

1.4 Motivation

The motivational part of my thesis focuses on the innovative integration of mixed reality (MR) technology into artistic performance, especially for the HoloLens device. This effort stems from the recognition that traditional art presentation methods are limited in their ability to fully engage and immerse viewers in the experience. Also, during a pandemic, artworks could be viewed digitally, so they would be more accessible. Utilizing the capabilities of MR, my thesis offers a transformative approach that goes beyond mere visual representation and offers an interactive three-dimensional experience that can greatly enhance the appreciation, understanding of works of art and also gives the opportunity to people to view artworks without being physically to the place that they are located.

This new application of MR technology in the presentation of art aims at several main goals: first, it aims to break the physical and spatial limits of traditional art galleries or museums, enabling a more dynamic and individualized interaction with art. Second, it seeks to enrich the educational perspective of art appreciation by providing users with contextual information, artist knowledge, and historical background in an engaging, interactive format. Third, a main goal is to make artworks more accessible. Overall, my research is motivated not only to redefine the way art is experienced, but also to push the boundaries of the integration of Artificial Intelligence in Mixed Reality.

1.5 Scope of thesis

The main goal of the thesis is to develop a HoloLens-based MR interface specifically designed for viewing digitized works of art. More specifically, we aim to improve the user experience regarding art perception and education through immersive MR technology, aiming to bridge the gap between traditional art viewing methods and modern technological advances. A backend system and a website were created only for demonstration purposes but most of the effort was put to the client-side. Also, a task for this thesis was to explore how the LLMs affect the user experience in the Mixed Reality environment and it was found that it has a positive effect to the user experience.

Chapter 2

Background Theory

2.1 Introduction

This chapter will discuss the existing artwork displaying methods and will explain in depth what displaying artworks in mixed/augmented/virtual reality is and what displaying methods are already available. Furthermore, it will talk about Mixed Reality (MR), what it is and how it can be. Also, a comparison between MR and Augmented Reality (AR) will take place since it is common to be confused with Augmented Reality (AR).

2.2 Artwork Displaying Methods

2.2.1 Explanation of MR in Art:

The concept of Mixed Reality (MR) is revolutionizing the way we experience and interact with art, offering a blend of real and digital worlds to create immersive environments. This innovative approach allows for a new dimension of artwork display, where digital elements are seamlessly integrated with the physical world, enhancing the viewer's perception and interaction with art.

A ground-breaking technological advancement in the art world, mixed reality combines digital content with real-world. Through this merging, digital items are not only exhibited but also interactively merged with the physical world, providing enhanced content about viewing artworks. An example of this would be the addition of digital images, animations, films, or contextual information to a traditional painting or sculpture. This allows the viewer to understand and enjoy the artwork on a deeper level. More people may now access and fully understand the work due to the active interaction between the actual and digital worlds, which adds layers of meaning and engagement.

Methods and Implementation:

It requires specialized hardware and software to implement mixed reality in art exhibitions. In this integration, devices like the Microsoft™ HoloLens, an MR headset, are necessary. This device can display digital content into the physical world using sensors, optics, and processor, resulting in a complete mixed reality experience. This technology, when used in combination with art, enables viewers to interact with works of art in a variety of ways.

For instance, when a spectator looks at an artwork, the MR device can display animated contexts that are relevant to the piece, the artist's notes, or even interactive parts that interact with the spectator. This approach not only enhances the experience of the viewer but also creates additional opportunities for interaction and storytelling inside the framework of the artwork.

Furthermore, MR uses go beyond simple visual enhancement. They can provide both text and audible feedback, allowing for a more complete sensory experience. An atmosphere for learning that is more interactive is created by this multisensory interaction with art using MR technologies. With the use of technology that blurs the lines between the material and the ethereal, viewers explore the depths of the artwork as active participants rather than as passive viewers.

Also, if the user cannot view an asset in a museum or an exhibition on the physical world, the possibility for viewing it digitally anywhere is possible.

To sum up, Mixed Reality is refashioning the way that art is exhibited and experienced. Artists may produce exhibits that are more dynamic, educational, and interactive by utilizing MR technology. This opens new possibilities for art creation and interpretation in addition to changing the experience for the observer. We may anticipate even more creative methods of showcasing art as technology develops further, increasing accessibility and making art more interesting for all.

2.2.2 Explanation of AR in Art:

Augmented reality, or AR, is a revolutionary method for showcasing and appreciating art by placing digital content on the physical world. With the help of this innovative technology, art watching may become even more engaging and interactive as real artworks are enhanced with virtual layers that provide depth, context, and interaction that is not achievable with conventional displays [6].

Fundamentally, Augmented Reality improves our view of art by giving real works of art a digital component, producing a composite image that smoothly combines the two. As a result, spectators are able to enjoy artworks that are enhanced digitally with things like interactive stories, animations, or thorough historical and contextual details. The purpose of these digital layers is to further improve the physical artwork, giving the viewer a more engaging and immersive experience. They are not just static overlays. When a spectator views a sculpture, for instance, through their augmented reality device, the technology can make the sculpture come to life by illuminating its making, highlighting its historical significance, or even animating parts of the narrative it represents.

Methods and Implementation:

The widespread use of augmented reality in art has been made possible by mobile technology. Many augmented reality art displays may be accessed via smartphone applications that use the camera on the device to recognize and add digital elements to the actual artwork. To display the augmented content, users only need to aim their device at the artwork; the program will identify specific markers or identifiers. By allowing everyone with a smartphone or tablet to interact directly with art, this method empowers access to artworks.

Numerous digital advancements can be obtained using these mobile applications. When a device is pointed at a painting, for example, it may provide information on the artist's biography, the historical setting in which the artwork was made, or a full description of the picture's artistic style. Thematic or story-related animated tales or visual effects that are displayed in more sophisticated apps can further augment the viewer's comprehension and appreciation of the artwork by providing a cinematic experience [1].

Additionally, AR art exhibits can promote a more interactive learning environment. Certain applications provide two-way interaction between users and digital content, hence enhancing the viewer's engagement with the artwork. These interactions can be achieved through touch or voice commands. A passive watching experience can become an active exploration with this level of participation, bringing up art to a wider population and increasing its attraction, particularly for educational or new audiences.

To summarize, the frontiers of art presentation and appreciation are being redefined by Augmented Reality. Augmented Reality (AR) provides a more immersive, interactive, and enriched art experience by fusing digital content with the real environment. This enhances viewers' appreciation and understanding of art. AR technology promises to open up new channels for artistic expression and participation as it develops and becomes more ingrained in our daily lives, making art more approachable and interesting for all.

2.2.3 Explanation of VR in Art:

A revolutionary change in the ways that art can be viewed, experienced, and interacted with by virtual reality (VR). VR technology provides an incomparable platform for exhibiting art by constructing completely immersive virtual worlds that let people tour virtual art galleries or interact with artwork in novel and profound ways [14]. With virtual reality technology, the user is cut off from the actual world and immersed in a virtual environment. This immersive experience is particularly powerful when used to art, since it creates countless opportunities for production, presentation, and engagement. With virtual reality (VR), curators and artists may create virtual galleries that exceed the limitations of physical space, providing a platform for exhibition and expression. VR can also be used to digitally remake already-existing artworks, presenting them in novel ways or incorporating interactive features that deepen their significance and engage viewers.

The viewer's relationship with art can be significantly changed by this degree of immersion and participation, making the experience more memorable, compelling, and intimate [12].

Methods and Implementation:

Most of the time, a VR headset like the Oculus Rift or HTC Vive is required to access these virtual art experiences. With the help of advanced tracking and display technology, these devices are able to effectively express the existence in virtual environments. Users can move from room to room in virtual galleries as if they were in a real museum once they have a VR headset on. Users can interact with artworks in ways that are not possible in real life, so the interaction goes beyond simple observation. For example, one may handle a sculpture to observe it from every angle or "enter" a painting to explore its surroundings from within. Anyone could even experience the enhanced artwork from the perspective of the artist.

Furthermore, viewing VR art exhibits has no limitations to static illustrations. Interactive narratives are one of them; in these, the viewer's actions affect the artwork, leading to a dynamic artwork that changes with each interaction. This could entail changing the surroundings around the artwork, discovering its hidden layers, or even helping to create it in a collaborative virtual setting.

Virtual reality technology has significant art-related implications. By enabling anybody, wherever, to visit virtual galleries, it not only democratizes access to art but also broadens the definition of artistic expression. With the physical constraints of traditional media long gone, artists may now create large-scale, interactive works that connect with viewers in whole new ways [15].

Finally, by providing immersive, interactive experiences that enhance viewers' involvement and comprehension of art, virtual reality is revolutionizing the art world. VR technology promises to further push the frontiers of artistic expression and enjoyment as it develops and becomes more widely available, making the art world even more vibrant and inclusive.

2.2.4 Explanation of Web Artwork Display (2D and 3D)

The creation, presentation, and experience of art have all undergone significant change in the digital age, with web artwork displays being a key component of this development. Nowadays, there are many different ways that artistic creations can be exhibited online, from conventional two-dimensional screens to more immersive three-dimensional web technology. This change offers viewers new methods to interact with art while also expanding the artist's potential audience [21].

A vast range of formats, from digital paintings and photos to interactive constructions, can be exhibited on websites using art. Although 2D screen displays have been widespread, enabling users to navigate image galleries or watch movies, the emergence of 3D web technologies has started to alter the online art exhibition environment. With the use of these cutting-edge platforms, viewers can be engaged with artwork in a virtual three-dimensional environment for a more immersive experience. By offering a sense of depth and space, this technology enhances the viewing experience by enabling visitors to examine artworks from various perspectives and to move through virtual art installations as though they were real.

2.3 Mixed Reality

Through the use of digital and physical elements, mixed reality (MR) technology creates a new world in which virtual and actual items coexist and interact in real time. It offers a more dynamic and integrated experience than either technology alone by fusing the immersive aspects of virtual reality (VR) with the interactive components of augmented reality (AR)

Definition and Difference of AR:

MR vs AR: MR attaches virtual objects to the actual world and allows more complicated interactions, while AR overlays digital content over the real environment and provides simple interactions. Digital content usually acts as an overlay over the real world in augmented reality (AR), but in mixed reality (MR), digital and real-life elements can interact. In MR, the artwork may come to life and interact with the real objects in the room, but in AR, you might see a virtual painting hanging on the right wall [22].

How MR is Used:

2.3.1 MR in art and museum

A major advancement in our ability to interact with and understand art is being brought by Mixed Reality (MR), which is transforming the way that art and exhibitions are experienced in galleries and museums. Through the merging of the real and virtual world, MR presents a unique chance for museums to design incredibly interactive and engaging displays. Imagine exploring a museum where sculptures and paintings come alive in front of you because of MR headsets. In order to give voice to the silent stories that are contained in the artworks, figures can come out from their frames and tell the story of their historical context or the artist's intention. Sculptures may be brought to life through animation, providing an animated representation of their place in history or process of creation [32]. Beyond simple observation, guests with MR technology can access extra layers on the actual artwork, enhancing the viewing experience. This content could include artist biographies, extensive comments, and historical details. This interactive method increases art accessibility and engagement for more people while also augmenting the educational value of museum visits.

A deeper awareness and comprehension of cultural legacy and artistic expression are promoted by MR, which turns viewers from passive spectators into active participants in the art.

2.3.2 MR in design and visualization

Mixed Reality (MR) has proven to be a very useful tool in the fields of art and design, especially for creation and visualization. With the help of this cutting-edge technology, designers and artists can now imagine, develop, and produce their creations in the same space in which they will be used or exhibited. Before any actual work is done, creatives can see firsthand how their installations or art will fit into and change a particular area thanks to mixed reality (MR), which allows digital material and the real world to come together.

This ability has revolutionary effects on the finished product as well as the process of creation. An artist preparing a large-scale artwork, for example, can utilize MR to overlay their design onto the selected wall, giving a realistic preview of how the piece will blend in with its surroundings. This helps with the decision-making process not only from an aesthetic standpoint but also in spotting possible problems with scale, color contrasts, or interactions with neighboring elements (such doors or windows) that could affect how visible or useful the design is.

Similarly, MR can be used by designers working on public projects to test out various locations, setups, and scales. Designers are able to predict more accurately how their work will interact with natural lighting, pedestrian flow, and other architectural elements by seeing their works in real-time within the intended environment. This kind of vision can result in designs that improve public areas by being more interesting, approachable, and harmonious.

Furthermore, no matter where team members are located, MR technology makes it easier for them to collaborate by enabling them to see and work with the same design in a physical setting. Through stimulating a more unified design approach and streamlining the feedback process, this shared experience can ultimately result in a final product that is more refined and coherent.

MR can be an effective tool for stakeholder interaction and client presentations in addition to helping with the design stage. Clients and other stakeholders may simplify the approval process, improve decision-making, and ensure that expectations are in line with the creative vision by seeing a realistic picture of how an artwork or installation will look in its intended location.

In conclusion, a major advancement in the way artists and designers approach their trade is represented by the usage of Mixed Reality in design and visualization. In addition to improving the creative and decision-making processes, mixed reality (MR) lowers the possibility of expensive mistakes by allowing the visualization of artworks and installations within their intended environments before physical creation, guaranteeing that the finished pieces are both aesthetically pleasing and flawlessly integrated with their surroundings.

2.3.3 MR in Education and Training

Mixed reality (MR), which provides a full and immersive learning environment, is revolutionizing the field of education and training across multiple sectors. This technology creates a hybrid environment where interesting and interactive learning opportunities are realized by fusing the digital and physical worlds [31]. In domains like medical [30], engineering and military applications where practical experience and situational awareness are critical, MR's exceptional capacity to replicate real-life scenarios and environments makes it a priceless tool.

MR is transforming healthcare professionals' training in the medical field. It provides a very realistic practice environment without the hazards associated with working with live patients by enabling medical professionals and students to perform virtual surgeries or diagnostic procedures. With the use of MR, students can develop confidence and enhance their skills by simulating medical procedures, understanding the spatial relationships between various organs and tissues, and visualizing complicated anatomical structures in three dimensions. With instantaneous feedback and the ability to repeat steps until mastery, this degree of realism and involvement improves the learning process.

The use of MR technology has significantly improved engineering training and education. With the use of mixed reality (MR), engineers of all skill levels may create, view, and work with three-dimensional (3D) models of buildings, machines, or systems in a realistic setting. An enhanced comprehension of design principles and operational mechanics is made possible by this practical method of instruction and problem-solving. Moreover, MR simulations may simulate real-world engineering problems, giving students the opportunity to troubleshoot and find solutions in a virtual setting, encouraging creativity and critical thinking.

When it comes to military applications, MR is essential for preparing soldiers for dangerous and complex scenarios that would be hard or impossible to safely reenact in real life. Through the use of realistic combat scenarios, MR simulations can give trainees the chance to hone their tactical, strategic, and decision-making abilities in a safe setting. In order to ensure that people are ready for a variety of circumstances, these virtual battlefields and operational scenarios provide a dynamic training experience that can adjust to the evolving nature of military operations.

Given that MR offers an unequaled level of immersion and engagement, it has a significant impact on training and education. MR offers learners a secure place in which to practice, experiment, and refine their abilities by simulating real-life scenarios and environments. In addition to improving education, this technology helps people get ready for the real world and gives them the knowledge, abilities, and self-assurance they need to succeed in their chosen industries [5].

2.3.4 MR in Architecture and Real Estate

Mixed Reality (MR) is a rapidly developing technology that is revolutionizing the real estate and architectural industries by providing clients, designers, and architects with never-before-seen possibilities to view and engage with building designs in real time. Mixed reality (MR) enables the production of realistic, three-dimensional virtual models of architectural projects by fusing digital content with the real world. This allows stakeholders to virtually walk through their prospective houses, offices, or public structures before any real construction starts.

The vibrant, dynamic experience offered by this cutting-edge method of architectural visualization exceeds the experience offered by traditional 2D drawings or even 3D computer models. With the use of MR devices, clients can look into every facet of a proposed design, including the arrangement of rooms and the way light and shadow change throughout the day. They are able to virtually walk regarding the structure, experiencing its volume, space, and overall functionality. Customer happiness and engagement are greatly increased when clients are able to make intelligent choices and request modifications or additions on the spot thanks to this direct and simple understanding of the design.

MR technology is a potent tool for creativity and accuracy for designers and architects. A very flexible and effective design process is achievable by the ability to visualize changes to the design immediately and make them in real time. By assisting in the early detection and resolution of possible problems during the design phase, this real-time visualization helps to minimize expensive changes during construction. Moreover, MR makes it easier for clients, designers, engineers, and architects to communicate and work together, ensuring that everyone is aware of the objectives and constraints of the project.

MR has the power to revolutionize the marketing and sales of real estate. A virtual tour of a property, especially one that is still under construction or in the planning stages, can be provided by real estate professionals to interested purchasers. In the current global market, when purchasers might not always be able to visit a property in person, this feature is very valuable. MR can give potential customers a more accurate impression of the area and its possibilities, enabling them to make more assured selections.

Moreover, MR applications may simulate different furniture configurations, interior design motifs, and even restorations, enabling clients to see a property's entire potential. In addition, for improving the purchasing experience, this degree of customization and interaction creates new opportunities for creativity and personalization in the marketing and design of real estate.

In conclusion, the way that buildings are designed, built, and sold is changing as a result of the incorporation of Mixed Reality into these fields. MR technology is opening the door for a new era of architectural innovation and customer service by providing a more dynamic, engaging, and efficient approach that helps architects, designers, and real estate professionals meet the changing demands and expectations of their clients.

2.3.5 MR in Manufacturing and Engineering

Assembly Line Setup: It requires accuracy and careful planning to set up or reconfigure assembly lines, which is a complex process. By allowing engineers to visualize the production line in a mixed reality setting, MR technology transforms this process. Without requiring any actual restructuring, they are able to virtually organize workstations, conveyors, and machines to optimize the layout for both safety and efficiency. By identifying any bottlenecks or safety issues, this virtual configuration makes it possible to make improvements prior to implementing changes in the actual environment. Moreover, MR can help with assembly line worker training by offering a realistic, interactive setting where they can pick up new machinery or processes without interfering with current operations [13].

Maintenance Operations: The efficient running of factories depends on the maintenance of machinery and equipment. By directly placing digital data, such schematics, operational details, or step-by-step repair manuals, on top of the actual machinery, MR helps maintenance staff. Machine downtime is decreased thanks to this hands-free access to vital information, which enables quicker diagnosis and repairs. By being able to see inside machinery, understand how systems work together, and follow instructions in real time, technicians can do maintenance jobs more quickly and with less chance of error.

Additionally, MR makes remote support possible, allowing professionals to guide staff members through difficult repairs or troubleshooting assignments from any location in the real world. This feature not only makes better use of senior engineers' experience, but it also speeds up problem solving, increasing overall operational effectiveness.

2.3.6 MR in Retail

Mixed Reality (MR), which offers a combination of ease, personalization, and participating that traditional and online retail experiences find difficult to match, will certainly change the dynamic world of retail and the way customers purchase. MR allows shoppers to browse products in a highly immersive virtual world and virtually try on clothing, accessories, and makeup. This is accomplished by integrating digital content into the actual shopping experience. This creative strategy gives sellers new ways to interact with their target market while also improving customer happiness [9, 10, 37].

Virtual Try-Ons: One of the most interesting applications of MR in retail is the virtual try-on feature. Customers no longer need to physically try on clothing, spectacles, or accessories because MR technology allows them to see how they look on them. This technology improves the efficiency and enjoyment of shopping by not only saving time but also lessening the discomfort of fitting rooms. Additionally, it gives consumers the chance to try out products or fashions that they might not have otherwise considered, which could boost sales.

Improved Product Interaction: MR gives consumers the opportunity to engage with a variety of products in a virtual setting, not just apparel. Customers can picture, for instance, how a certain paint color would appear on their walls or how furniture would fit and look in their own houses. This level of interaction lowers the possibility of returns, improves consumer happiness, and helps make more accurate purchases.

Educational Content and Demonstrations: Retailers have the ability to take advantage of Machine Learning (MR) to offer clients classes, product information, and demos. Customers may obtain information on products' manufacturing processes, origin, and usage instructions, for example. A retail business that offers this extra content can stand out from the competition and attract customers who want to learn more about the things they buy.

Remote Shopping Experiences: Moreover, MR creates the opportunity for online purchasing, allowing clients to browse and engage with products just as if they were in-person. When compared to typical internet shopping, this offers a level of accessibility and interaction that is especially encouraging to customers who are unable to visit a business in the real world.

In summary, Mixed Reality has the potential to revolutionize the retail industry by providing a more personalized, efficient, and enjoyable fusion of purchasing experiences. Virtual try-ons, improved product interaction, personalized shopping experiences, instructional content, and remote shopping capabilities are just a few of the ways that MR technology helps sellers stand out from the competition and keep their customers in an increasingly competitive marketplace.

2.3.7 MR in Gaming and Entertainment

By combining virtual and real-world gaming settings, mixed reality (MR) is revolutionizing the gaming industry and providing gamers with a never-before-seen level of interaction. With the help of this state-of-the-art technology, players may engage with digital information in their real-world surroundings, fusing real-world and virtual components to create experiences that are more realistic and captivating than ever [16].

Immersive Gaming Experiences: MR allows players to enter a game environment that exceeds typical screen gaming. Mixed Reality (MR) incorporates digital items into the physical world so that they can interact with both the player and the real world. This results in a distinct gaming experience that blurs the lines between the real and virtual worlds and offers a higher degree of immersion. Through physical navigation and interaction with both virtual and real-world items, players can experience the game environment as though it were an extension of their real-world surroundings.

Linking the Real and Virtual Worlds: Mixed reality games are made to detect and react to the player's physical surroundings. As a result, public areas like parks and living rooms can be transformed into game environments where digital elements react to natural obstacles and features. For instance, during a game, a sofa in the living room may change into a virtual barricade, giving players a real-world and virtual challenge to overcome. Because real-world components are included in the game experience, every play session is distinct and extremely customized.

Future of Gaming: MR gaming has the potential to become a popular form of entertainment as MR devices become more widely available and reasonably priced. It offers experiences that are not only entertaining and engaging but also potentially instructive and physically demanding.

In summary, mixed reality could revolutionize the gaming industry by enabling immersive experiences that smoothly merge the virtual and physical worlds. MR gives gamers a new level of involvement by improving gameplay, and interaction. This makes gaming more dynamic, intuitive, and exciting. As this technology grows, it will open up new avenues for game creation and gameplay, bringing in a new phase of interactive entertainment.

2.3.8 MR in Military and Defense

Mixed Reality (MR) technology is becoming a game-changer in the military and defense industry. It is going to revolutionize mission planning capabilities and change the way training takes place. Military troops can operate in extremely realistic, interactive environments that simulate the intricacies and unpredictability of real-world operations thanks to mixed reality (MR), which blends digital information with the real world. With the demands of modern combat, where it is more and more important to integrate digital information with physical action, this technology is proving to be useful in preparing soldiers for those needs.

Realistic Training Environments: Mixed reality (MR) technology generates training simulations that combine digital and real-world military operations [36]. These models are capable of simulating a variety of situations [33], including medical and transportation challenges [35], urban combat, and counterterrorism operations. Through controlled environment interaction with both virtual and actual elements, military personnel can obtain practical experience with the tactics [34], techniques, and procedures they will encounter in the field. In addition to being more interesting, this type of training enables a degree of complexity and realism that is not achievable with conventional training techniques. Because scenarios can be readily altered to highlight particular skills or consider new risks, MR is a flexible and useful training aid.

Safety and Efficiency: The ability to perform detailed instruction and planning without the risk and expense associated with real exercises is one of the main advantages of mixed reality (MR) in military applications. This not only guarantees worker safety during training but also makes it possible for more frequent and varied simulations, which raise preparation levels.

In summary, by offering realistic training environments, strengthening mission planning, and dispersing crucial information in the field, mixed reality is enhancing the capabilities of military and defense operations. This technology improves operational effectiveness and safety while also preparing military personnel for the intricacies of modern combat. As MR technology develops, it will play an even more crucial role in military planning, training, and operations, highlighting its significance in the defense industry.

2.3.9 MR in Art and Design

Mixed reality (MR) offers artists the unique capability of extending their digital works into the actual environment, enhancing them with enhanced artworks that are exclusively visible on MR devices. This has the power to change regular spaces into dynamic, interactive art galleries where the boundaries between reality and virtual world are merged. These digital improvements can be exhibited in public areas, museums, and galleries, providing new perspectives on art and culture.

Interactive Exhibitions and Presentations: Exhibitions and presentations are much improved thanks to Mixed Reality (MR) technology; viewers get to enjoy an unrivaled level of immersion and interaction as these events benefit from the incredible extra dimension that MR offers, enabling organizers to present ideas, knowledge and products in a dynamic, interactive and much more effective way than with conventional display techniques.

If you are working in an industry or with products, services, or concepts that are inherently complex or abstract, MR makes for a powerful tool. For example, engineering companies can demonstrate exactly how their machinery works on the inside in three dimensions, while a biotech firm can show you exactly what is happening at a microscopic level inside your cells. Also, assets of a museum could be viewed digitally. By creating environments where these ideas can be visualized in three dimensions on an immersive scale, it allows for the conveyance of intricate detail in ways that are much more digestible for an audience and vastly more engaging.

2.3.10 MR in Law Enforcement and Forensic Analysis

In the field of law enforcement and forensic analysis, the utilization of advanced technologies has dramatically expanded the capabilities of investigators and forensic scientists. While the use of Mixed Reality (MR) for crime scene reconstruction and analysis may be the most innovative of these technologies. This groundbreaking application allows for a more interactive and immersive means of investigating and understanding the complexities of a crime scene and redefines the way forensic analysis is conducted [20].

Chapter 3

Tools, Technologies and Architecture

3.1 Introduction

This chapter will present the various technologies and tools used to create the Mixed Reality interface. In addition, an explanation of the system architecture will be provided.

3.2 Technologies and Tools Used

For the design and implementation of the system, the following tools and devices were used:

1. **HoloLens**, developed by Microsoft to integrate Mixed Reality (MR) into our daily lives through a computer embodied in a wearable headset. It allows users to experience MR immediately upon wearing it, featuring semitransparent holographic lenses that produce three-dimensional holograms. These holograms merge with the user's environment and adhere to surfaces and objects, created by directing light from the top of the lenses into the user's eyes. It maps the space around the user, recognizing surfaces to facilitate the placement of holograms. It includes various sensors for spatial mapping and a 2-megapixel camera for taking photos and videos.

   The HoloLens is a revolutionary device that not only enhances everyday tasks and experiences but also gives us entirely new possibilities. It has a several applications such as remote assistance through video calls with screen sharing for complex tasks, 3D computer-aided design for creating and visualizing new products in real size within the physical environment, the gamification of routine tasks to increase engagement and productivity, immersive mixed reality gaming, and holographic entertainment, allowing users to explore distant locations virtually without the need for travel. It is illustrated in **Figures 3.1, 3.2**.

Figure 3.1 – HoloLens see-through lenses

Figure 3.2 – HoloLens Sensors and Camera

2. ***Microsoft's HoloLens Clicker*** which is a peripheral device used only when connected with the HoloLens to replace hand gestures. It allows an extra way of interacting and controlling the holograms projected by the device. The basic functionalities that someone can use are: scroll, zoom, click, click-and-hold. It is shown on **Figure 3.3** and **Figure 3.4**.

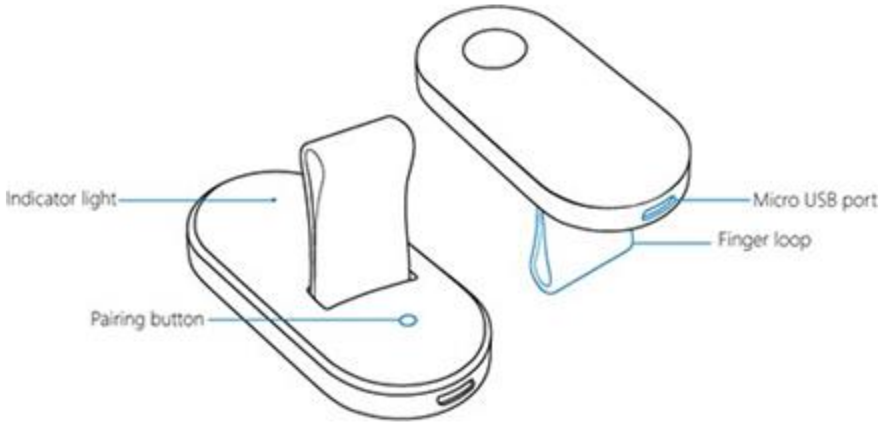

Figure 3.3 – Clicker

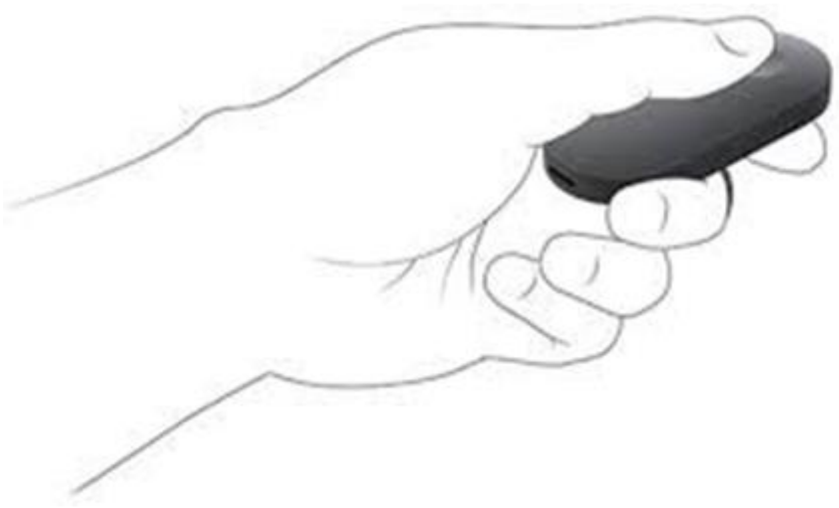

Figure 3.4 – How the clicker is used

3. ***Unity 2019.2*** is a game engine that developers are using to create interactive 2D and 3D applications. It is scripted in C#. The application of mixed reality was built on this game engine.
4. ***Unity HoloToolkit*** is a collection of scripts (opensource) and components that are used by developers to create applications on HoloLens through Unity. It contains a collection of objects designed for achieving interactive Mixed Reality interfaces and scenes. Many components from this toolkit were used from the application and too much time was saved.
5. ***C#*** is a programming language developed by Microsoft and is the language used in Unity and HoloLens. It can easily be easily used for contacting with APIs. The mixed reality application was written on this language.
6. ***Microsoft Visual Studio*** is an IDE (integrated development environment), and it is widely used for easier programming. It can be used for many languages and C# is built in. The main IDE that was used for the mixed reality application was visual studio.
7. ***Django*** is a high-level web framework that was written in Python, that helps developers to achieve rapid development. It was developed in 2005. It makes it much easier for developers to build complex, database-driven websites. It was used for the back-end part of the thesis.

8. ***SQLite*** is the default database that the Django framework uses, and it is suitable for development and testing purposes. The entire database is stored in a single file. This was the database that was used for the backend part of the thesis.

3.3 System Architecture

The system's architecture is the 3-tier model which is a client-server design. It consists of a presentation tier (the interface), logic/application tier (the server) and data tier (the database). Many developers are using this architecture because these modules could be developed simultaneously but in this thesis case it was selected because any tier could be scaled or updated without having an impact on another tier [8]. The architecture is illustrated in **Figure 3.5**.

More specifically the three tiers main tasks are:

1. **Presentation tier:** This is the front-end part of the system. It regularly sends a request to the logic tier and gets some data as a response. Those data when it is necessary are presented in a specific way to the user. For instance, when the bytes of some images of artworks are downloaded, they are displayed as images to the user.
2. **Application layer:** All the functionalities that the user can access are implemented in there. It retrieves data from the database when it is necessary, and it may convert it to a Json format for creating a structure that the client can handle. The response that this layer is returning is always in Json format.
3. **Data tier:** The main functionality that this tier has is storing and retrieving data. These data could be accessed through API calls that are done from the application layer. More specifically when the time comes for some data to be sent to the presentation layer, a request is sent to the application layer and the application layer executes an API call to the data tier for retrieving that information. After, the application returns them converted to Json format.

As was mentioned before this architecture was chosen as the development could be speeded up as each part could be developed concurrently and any part could be updated without having impact on the other. Another major benefit is that the availability could be expanded because each part could be checked and debugged individually.

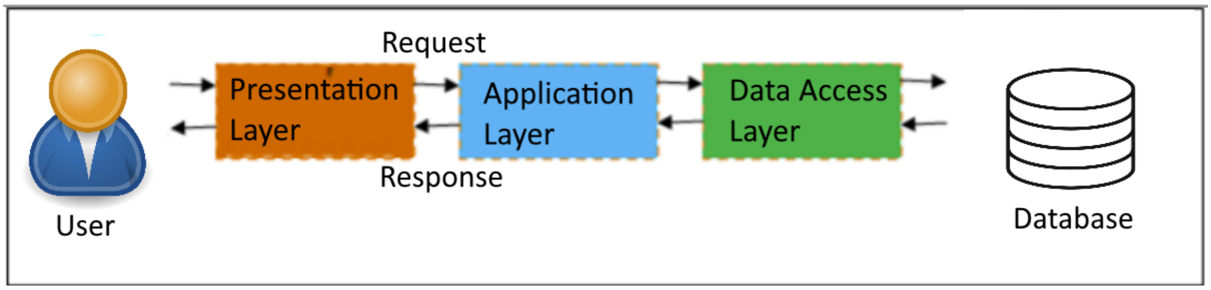

Figure 3.5 - Three tier Architecture model

Figure 3.6 – Three tier scheme with the user and the HoloLens device included.

Chapter 4

Design and implementation

4.1 Introduction

This chapter contains information about the system that was developed. Furthermore, frontend and backend system will be explained. To be more specific the Menu, Gallery, Extended Gallery View and the Photo Recognition System will be discussed with the architecture of the database and the website that is used for managing the artworks.

4.2 Database Architecture

The software that the database is running is SQL Lite and it is the default database that Django uses. This database is suitable for local development because it is lightweight and easy to set up and there is no need for doing complex configurations. Also, many developers are using it for prototyping and rapid development as they can focus more on their application logic rather than database administration. Furthermore, this database is often used on embedded systems where a small self-convenient database is needed. There are many programmers that are using it on mobile applications as both android and iOS are supporting this type of database and for testing purposes as they can run unit testing without having the overhead of setting up a server that the database will run on [19].

Now the architecture of the database will be explained. Before that two explanations about the name of some fields and tables will be provided:

1. The characters before the '_' character indicate the name of the database in the context of the table name that is displayed on the scheme. For instance, the name that is displayed on **Figure 4.1** is diplomatikibackend_associatedmusicmedia indicates that the table name is associatedmusicmedia and the database name is diplomatikibackend.
2. The fields that are ending with _id are indicating that they are storing a foreign key. On all purposes, the field artwork_id (that is shown on **Figures 4.1, 4.2, 4.3**) is a foreign key of the table diplomatikibackend_photo. The field on the photo table is storing a foreign key of the user table (**Figure 4.4)**.

The fields that are ending with *id are foreign keys. More specifically user*_id is the primary key of the Auth user table and the fields named artwork_id are the primary key of the diplomatikibackend_photo table.

Every artwork (**Figure 4.2**) has a 1: N relationship with any associated media, either its type is music, video, or photo (**Figures 4.3, 4.4, 4.5**) [3].

Let's explain what the use of each table and what data is are stored in there:

1. **user** table, which is illustrated in **Figure 4.1**, stores all the registered data of the user. Django had this table predefault.
   a. **id** is the primary key of the table
   b. **password** is the user's password
   c. **last_login** is the specific date and time that the last login was done.
   d. **is_superuser** indicates if the account has administrator rights.
   e. **username** is the username that is used for login
   f. **email** is the user's email.

g. **last_name** is the user’s last name.

h. **is_staff** indicates if the account has staff rights.

i. **is_active** indicates if the user can login.

j. **date_joined** indicates the date and time that the user registered.

k. **first_name** indicates the user’s name.

auth_user	
id	integer
password	varchar(128)
last_login	datetime
is_superuser	bool
username	varchar(150)
last_name	varchar(150)
email	varchar(254)
is_staff	bool
is_active	bool
date_joined	datetime
first_name	varchar(150)

Figure 4.1 - The structure of the user table.

2. **photo** table which is illustrated in **Figure 4.2**, stores all the artworks that were uploaded.

    a. **id** is the primary key.

    b. **title** is the title of the artwork.

    c. **image** is a pointer that indicates a relative path, where the file is stored on the server.

    d. **uploaded_at** indicates the date that the artworks were uploaded.

    e. **user_id** is a foreign key of the user table. The id is the uploader’s id.

diplomatikibackend_photo	
id	integer
title	varchar(255)
image	varchar(100)
uploaded_at	datetime
user_id	integer

Figure 4.2 - The structure of the photo table.

3. **associatedmusicmedia,** which is illustrated in **Figure 4.3** is a table that stores the URL of the music associated media.

    a. **id** is the primary key of that record.

b. **artwork_id** is a foreign key of a record that is stored in the photo table. It is the artwork that this music associated media is stored.

c. **url** is the music's url.

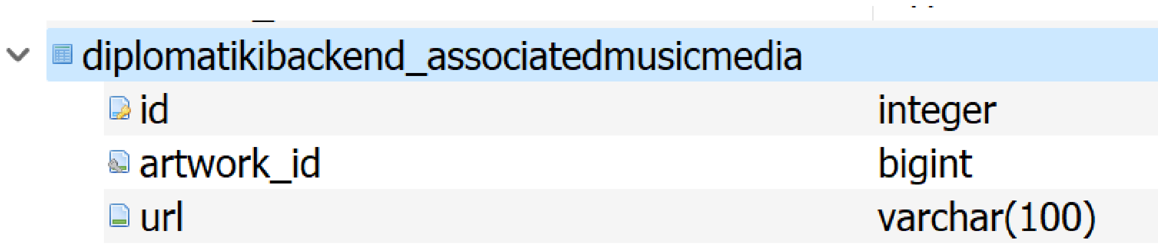

Figure 4.3 - The structure of the associatedmusicmedia table.

4. **associatedphotomedia** which is illustrated in **Figure 4.4**, is a table that holds the relative path on the server of a photo associated media.

   a. **id** is the primary key of that record.

   b. **artwork_id** is a foreign key of a record that is stored in the photo table. It is the artwork that this associated media is stored.

   c. **photografia** is a pointer that indicates a relative path of that artwork on the hard drive.

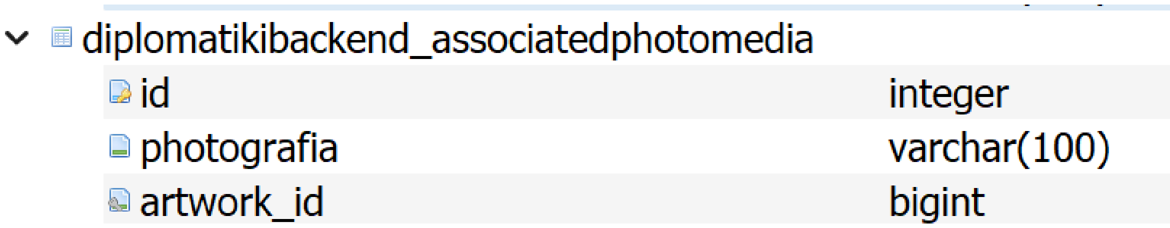

Figure 4.4 - The structure of associatedphotomedia table.

5. **associatedvideomedia** which is illustrated in **Figure 4.5**, is a table that holds the url of an associated media of a video.

   a. **id** is the primary key of that record.

   b. **artwork_id** is a foreign key of a record that is stored in the photo table. It is the artwork that this associated media is stored.

   c. **url** is the video's url.

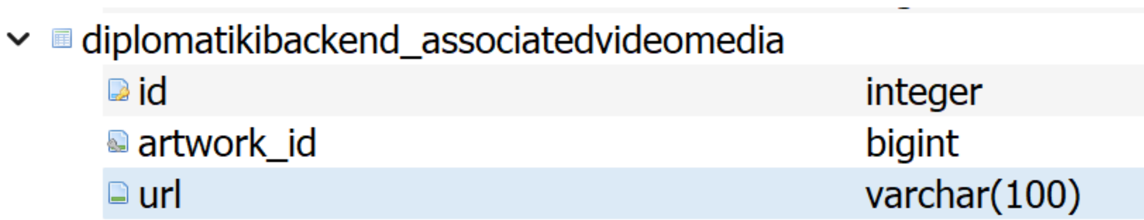

Figure 4.5 - The structure of the associatedvideomedia table.

4.3 User Interface and Implementation

A 3-step approach was followed for designing this interface. Furthermore, a low fidelity prototype was designed on a piece of paper as shown in **Figure 4.6.1**, a medium fidelity prototype as shown in **Figures 4.6.2 – 4.6.5** and a high fidelity as shown in **Figures 4.7 – 4.18**

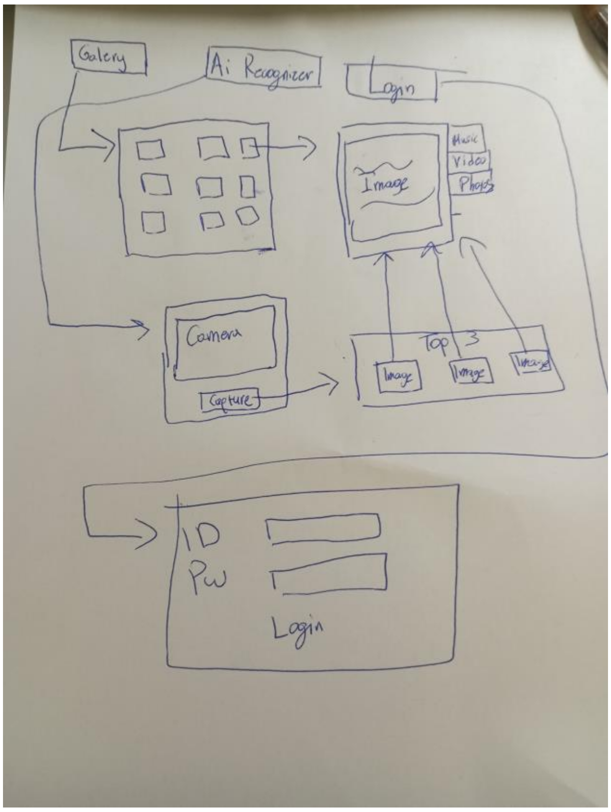

Figure 4.6.1 - Low fidelity prototype

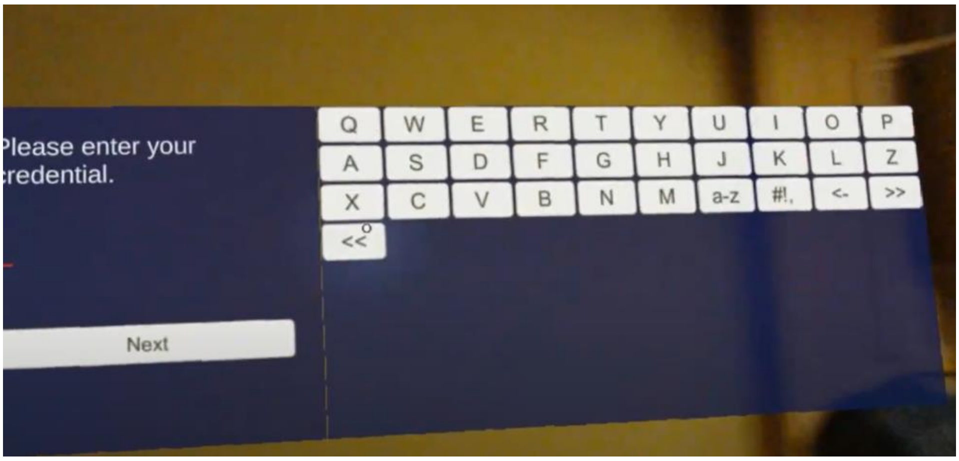

Figure 4.6.2 – Medium Fidelity Login

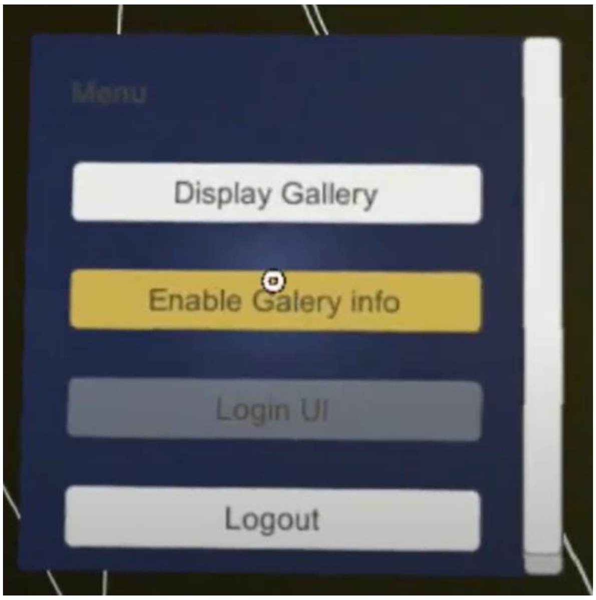

Figure 4.6.3 – Medium Fidelity Menu

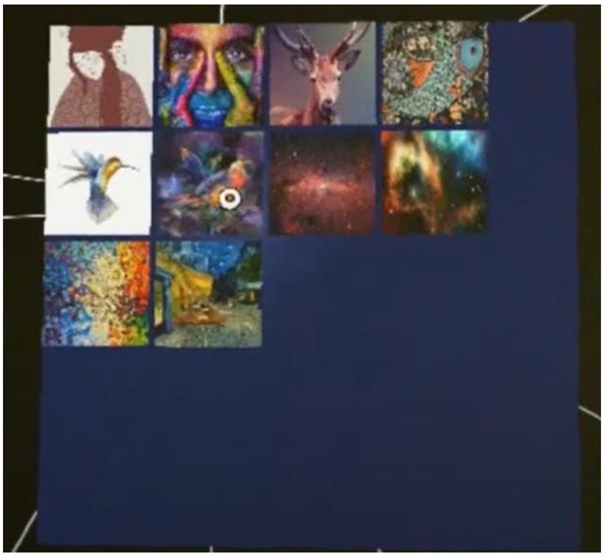

Figure 4.6.4 – Medium Fidelity Gallery

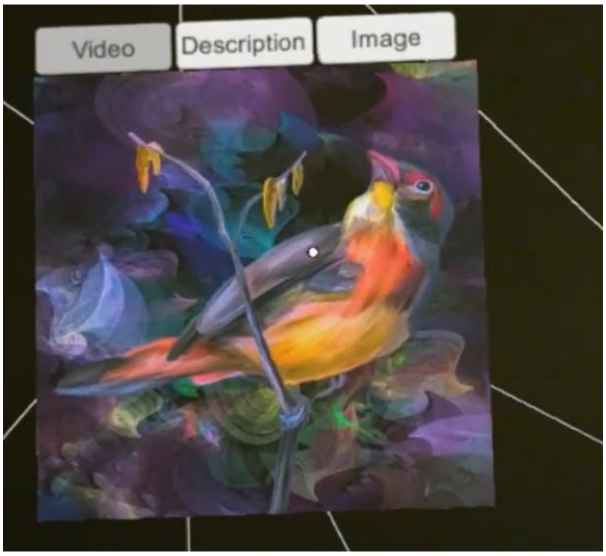

Figure 4.6.5 – An image is selected in the medium fidelity gallery.

For being able to do the first thoughts and starting the design of the project I had to explore the HoloLens's documentation on Microsoft's website and read many articles in forums. After finishing these tasks, I installed the project with the correct settings that are displayed on **Figure 4.6.6** to be able to deploy it on the HoloLens. I also installed HoloLens's toolkit for using its components.

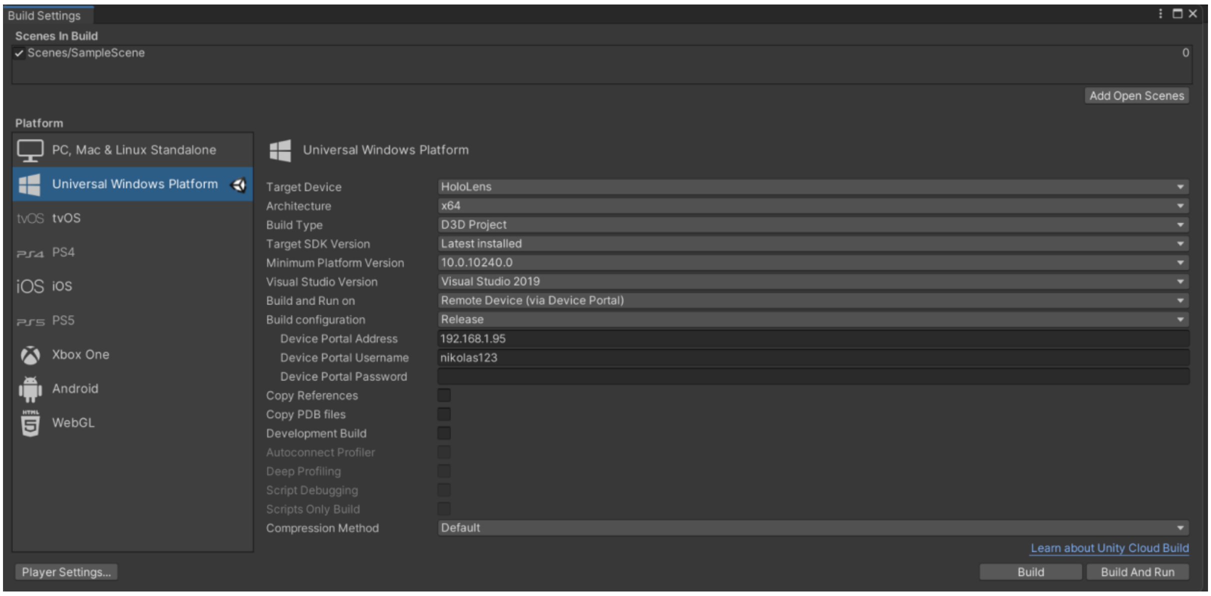

Figure 4.6.6 – Unity Build Settings for deploying on HoloLens

4.3.1 Main Menu

The menu that displays the available functionalities that the user can use

Menu as shown in **Figure 4.7** is displayed after logging in successfully. The user can also access the menu by clicking anywhere four times subsequently (in a short time).

If the user is not logged-in, is not allowed to click on the buttons for Gallery, Get Info, and Logout.

We should note here that every entity in unity is a GameObject and various components such scripts, images could be added in to it.

The home button on every canvas disables the activated canvas and displays this menu.

By clicking any image, the user can access the corresponding functionality.

The main menu is a canvas that has a GameObject as a child that corresponds to each functionality. Each gameobject has an image component and a text component for helping the user to be able to create a mental model for the system easier. When a functionality is selected the Menu's canvas GameObject gets disabled.

Figure 4.7. The menu that displays the available functionalities that the user can use

4.3.2 Gallery

Gallery is a canvas that can dynamically display many photos to the user. Furthermore, the user when selects an image, the capability to view associated media of a photo, or select the extended view to view them is available. Each image corresponds to a page and the user should go to the corresponding page to view the image. The extended view functionality will be explained to the N section. The Back and Next buttons are used for changing pages. Users can see the current page number at the bottom right. When a user selects (double taps) on an image, that image expands and becomes full size. Then, the user has the option to see the associated media. They could be video, text, images, and voice.

Regarding the download phase, Gallery's photos are requested by the server via a GET request, and are then displayed in the grid, as shown in **Figure 4.8**. More specifically, a Json file is downloaded to the client that contains all the names of the images and each image corresponds to a link in the server. After that the client sends GET requests for each image's corresponding link to download its image and the associated media of it. Associated media's type is either image, voice, textual and video and there is not a number limit. The maximum number of artworks that are downloaded is the maximum number that a page can display based on the width, height and spacing of each photo (a function in source code exists that calculates the maximum number.), so each time only the current page's images are downloaded. This is done for memory and time efficiency as the HoloLens's device is not as strong as a mean laptop. When an image is downloaded, a corresponding GameObject is created, and a GridElement (it handles the gallery's image) with an image is added on it and they are managed through a GridMananger object.

About the display phase, the photos automatically have a default width, height and spacing. The default configuration (width, height and spacing) can be changed by clicking the Configure button, which will open the configuration screen that is illustrated on **Figure 4.9**. and based on them the number of photos that are displayed per page is calculated. After that, the maximum number per page is recalculated and images are re-downloaded and re-displayed.

When an image is selected (double tapped), all the images' GameObjects are disabled except from the selected and the default size is changed to the Gallery's canvas's dimensions. So, the only image that is displayed is the selected one and it fits to the full size of the canvas.
By double tapping the photos again all the images's GameObjects are restored, and the default size is restored to normal.

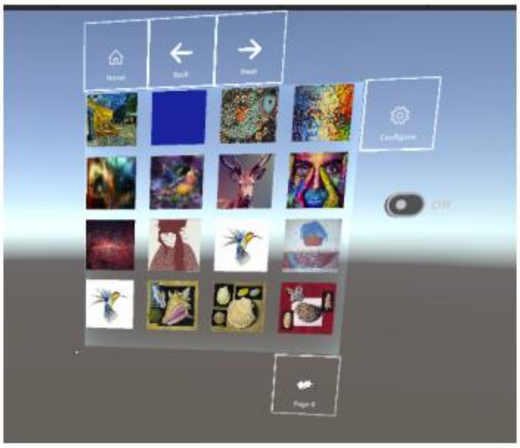

Figure 4.8. The Gallery displays all the artworks that were uploaded in the server

4.3.3 Gallery's Image Size Adjuster

Users can adjust the size of the images. For being displayed the configure button must be pressed. The Submit button should be clicked to save the setting, as shown in **Figure 4.9**.
After adjusting those 3D sliders, an event is triggered immediately that changes the default spacing, width and height of a photo and deletes all the displayed images's gameobjects and re-downloads all the photos and reconstructs the images's gameobjects. As mentioned at section 4.3.3 an image's gameobject contains a GridElement component.

The displayed value of a dimension attribute is normalized to a max and a min value that is a constant number in the source code and that is why it is ranging from 0 to 1.

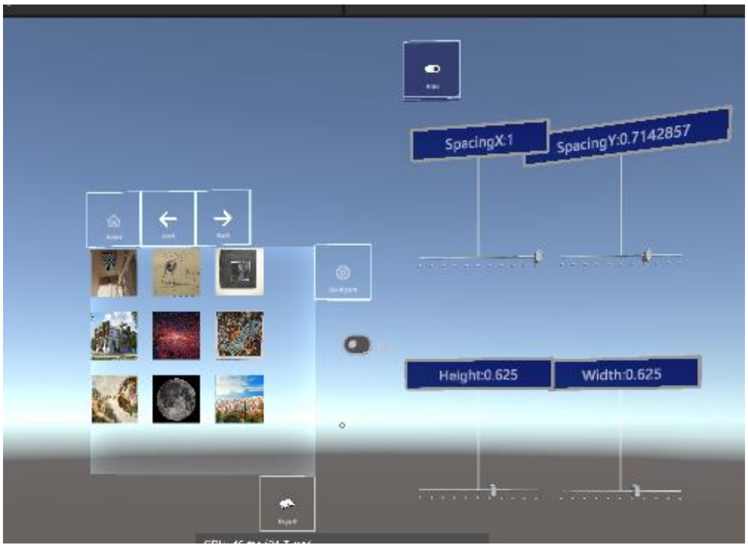

Figure 4.9. The configuration canvas through which the user can adjust the width, height, and spacing of the artworks in the Gallery.

4.3.4 Recognize artwork with AI and display information

In the screen that is displayed at **Figure 4.10**, the application allows the user to take a photo of an artwork that exists in the real world. Once the Capture button is clicked the three images that have the highest probability to be the captured one are presented.

When the user clicks the Submit button, the captured image is sent to the server via a POST request, and a machine learning algorithm extracts the main features of the captured photo and each stored photo and after that calculates the mean square distance difference. After that, the name of the three matched artwork with the least distance are returned to the client application. Then the corresponding images that have the highest probability are downloaded to the client and displayed in a new Canvas named Top3Canvas. Each image has a click event that when triggered it displays in the Gallery the clicked image.

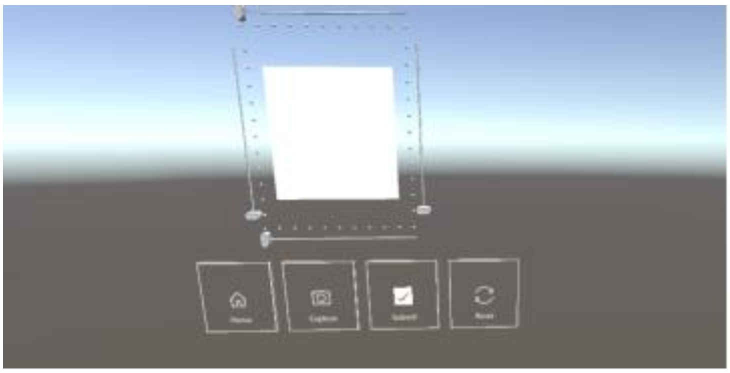
Figure 4.10 - The canvas that is used for image capturing and cropping

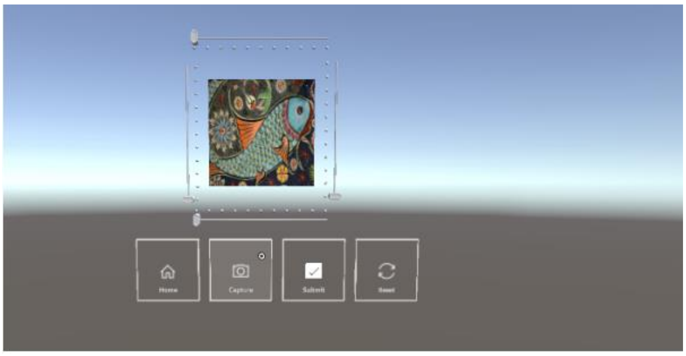

Figure 4.11 - The sample photo of a captured artwork

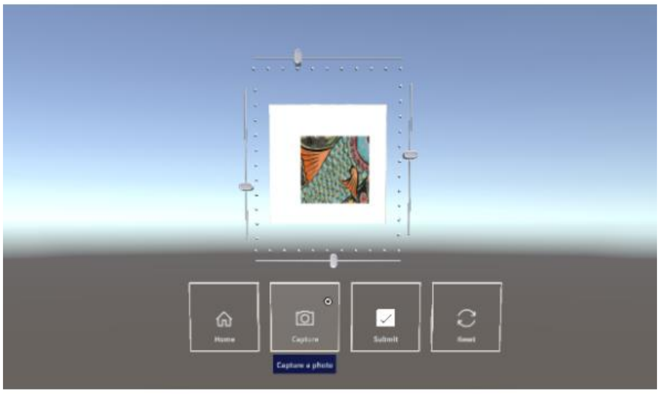

Figure 4.12 - The image can be cropped by adjusting the sliders accordingly

4.3.5 Extended Gallery View

The extended Gallery View allows the user to have a more realistic experience. The user can view 4 associated media on each page which are either image or text or video or music, a sample is illustrated on **Figure 4.13, 4.14 and 4.14.2**. The main image is always displayed at the center.

The user is able to change page each time and the images are always sorted by type. By changing page, the user views the four associated media are replaced with the next four of the same type. The associated media could be placed anywhere in the place while the tap gesture is made. Also, the associated media could still remain after the page changes if it is double tapped. This state is named “selected” and it could be deselected by a re-double tapping action. After tapping, a small green tick is displayed upwards right from the associated media that indicates that the image is selected, and it disappears after it is deselected. For further understanding the **Figures** that are numbered from **4.15 to 4.18.**

The deselected images only disappear after the page is changed.

From the technical view, the main GameObject parent of the associated media game objects is the central image. The name is “NodeForAssociatedMedia” indicates a component that is attached to the associated media gameobject that is displayed regardless of its type. The component named “DoubleTapForExtendedView” manages the trigger-event and also saves the selected or deselected state. The “VideoMananger” component manages the video to be displayed correctly.

When an associated media is selected and the page is changed, the selected gameobject is cloned with the Instantiate() method. Instantiate() method clones a gameobject with its components. The only difference between the cloned object and a regular one is that a flag exists that indicates that this is a cloned object, which is needed for destroying purposes after it is de-selected, and the page is changed.

When a video is played, it is done throught the VideoPlayer component that can display a video to a texture. The texture that the video is shown is named render texture and it is placed on a canvas component (of the associated media’s gameobject). The VideoPlayer component that is developed from Unity implements the Play, Reset and Pause methods.

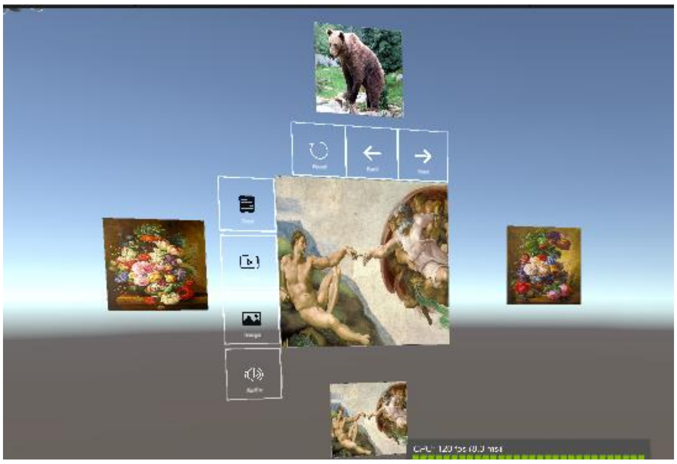

Figure 4.13 The image associated media

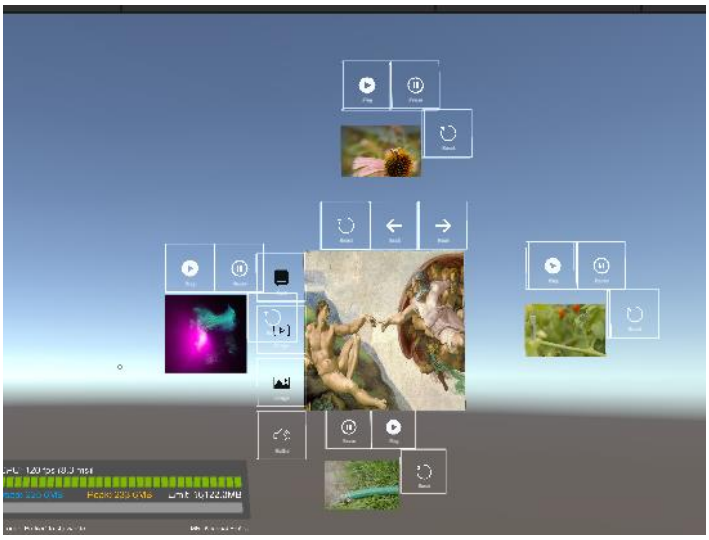

Figure 4.14 - The video associated media. A reset button exists at the right and upwards the play and pause.

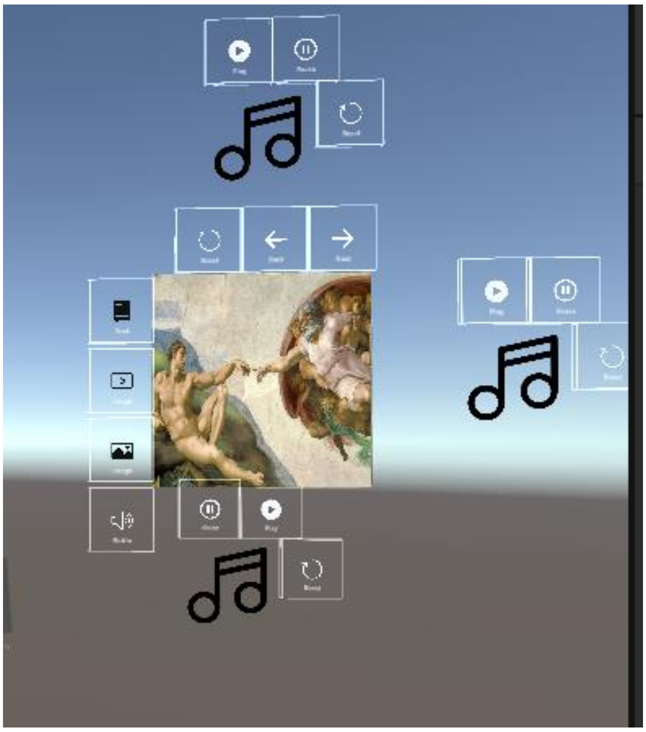

Figure 4.14.2 - The music associated media. The play and pause buttons are upwards and the reset at the right. Respectively they start the music to play, pause it and reset the second that is being played to zero.

Figure 4.15 - A color effect that affects the selected image and remains only for two seconds when the image is selected. It triggers much more the user’s caution to the photo to realize that the image is selected.

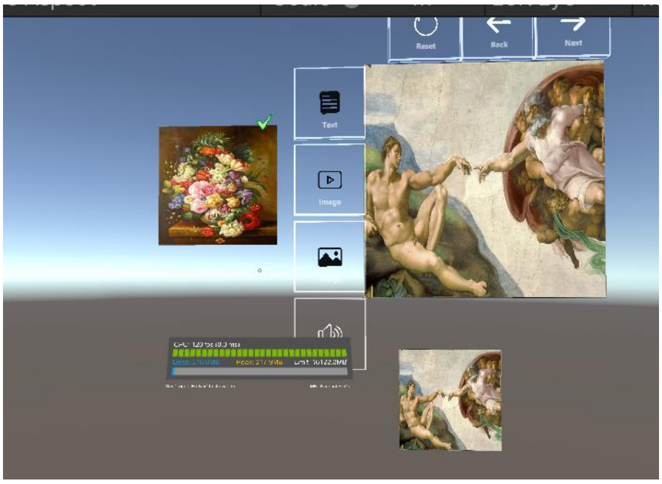

Figure 4.16 - An image after is selected is having a green tick. It will remain after the page changes at the same place.

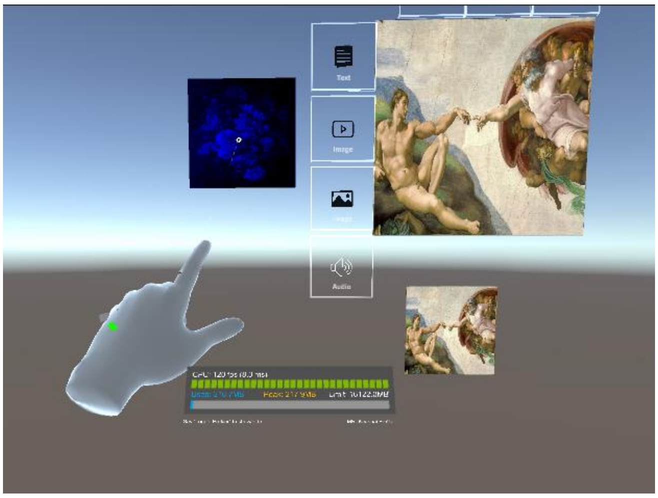

Figure 4.17 The blue color effect that affects the de-selected image remains only for two seconds and indicates the de-selected state. It triggers much more the user’s caution to the photo to realize that the image is de-selected.

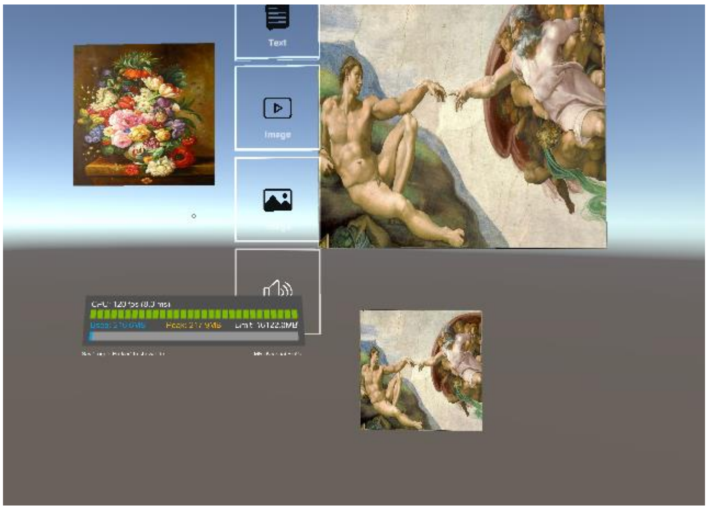

Figure 4.18 The green tick disappears after the image is deselected (re-double tapped).

4.4 Endpoints

In this section the endpoints that the mixed reality interface is using will be presented. The program that sends requests to the server is named Postman. The client is response is on the “Body” tab. All these endpoints are running on the Django framework.

4.4.1 ReturnPhotoAsResponse

Returns the relative path of all the records of the photo database converted to a Json structure for helping the client to handle this information. An example could be viewed at **Figure 4.19**.

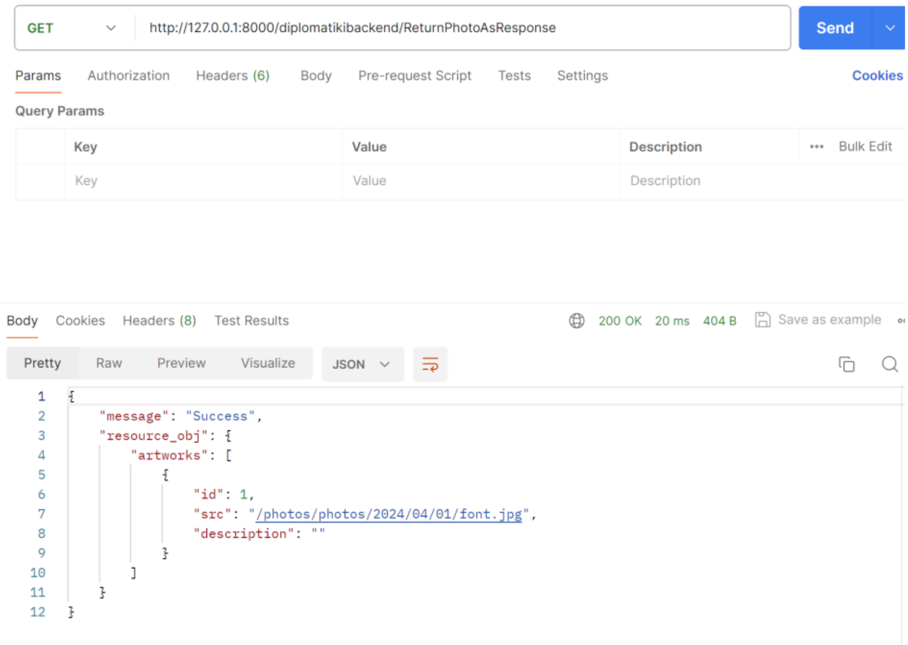

Figure 4.19 - The response of the ReturnPhotoAsResponse endpoint.

4.4.2 ReturnAssociatedMediaAsResponse

Returns the relative path of all the associated media (images, music, video) of a specific artwork (this artwork exists in the photo database) converted to a Json structure for helping the client to handle this information. The artwork is specified by the id parameter. An example could be viewed in **Figure 4.20**.

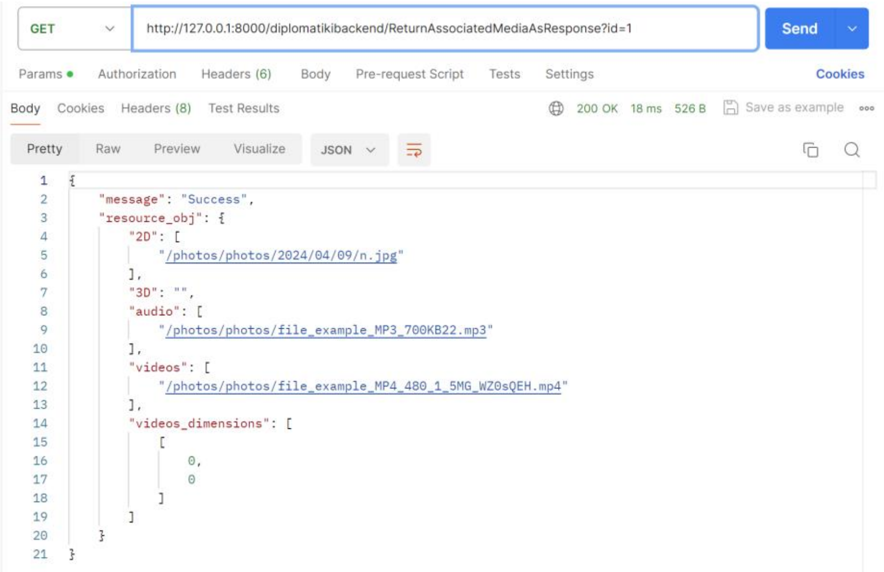

Figure 4.20 - The endpoint gets as parameter the id=1 which means the record in the photo database with id equals to one and returns all its associated media. The video_dimensions field is unused for now, so it is always returned as 0,0.

4.4.3 token

This endpoint could be used for authorized access for specific actions. For now, there are not any actions that require authorization. For achieving authorization, the client should write in the request's body that will be sent, the username and password like in **Figure 4.21**. The mixed reality interface has a label named e-mail instead of username because the authorization was done by communicating with another online service. This functionality was implemented from Django. An example could be viewed in **Figure 4.21**.

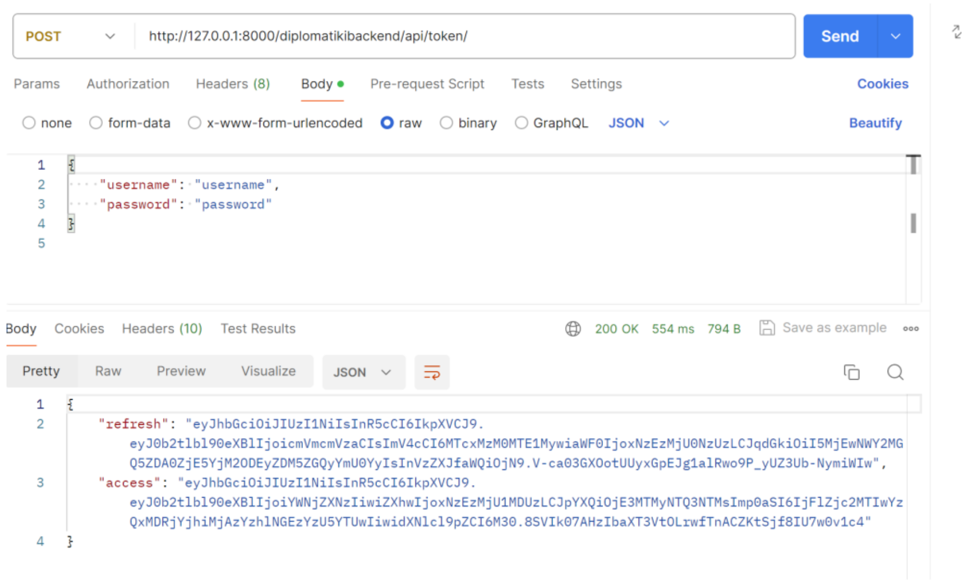

Figure 4.21 - A request is sent with a username and password in a Json format, and it is returned back a refresh and an access token.

4.5 Web Application

This section is going to illustrate a website that could be used for managing the artworks for the system. As was mentioned before, the main focus on this thesis was not creating a website but an interface in Mixed Reality, so the website was developed only for demonstration purposes.

4.5.1 Menu

The menu is always displayed on the top right. There are two different types of menus that are displayed to the visitor, one is only for logged-in users and the other is for non-logged-in users. The **figures 4.22 and 4.23** are displaying them.

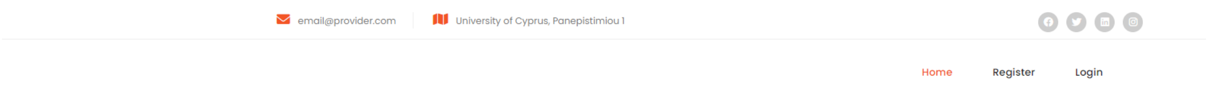

Figure 4.22 - The selections that a visitor could make is getting redirected to Home, Register and Login page.

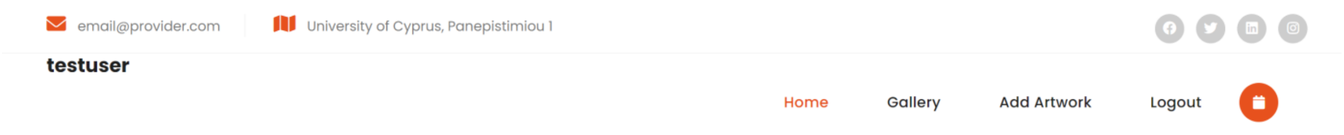

Figure 4.23 - The selections that a logged-in user could make are getting redirected to Home, Gallery, Add Artwork and Logout respectively. The name of the logged in user is always displayed on the top left.

4.5.2 Home

Home is a page that shows a small description about the website and some server messages. These messages may be "You have been successfully Logged in" or "You have been successfully logged out". Home page is illustrated on **Figure 4.24**.

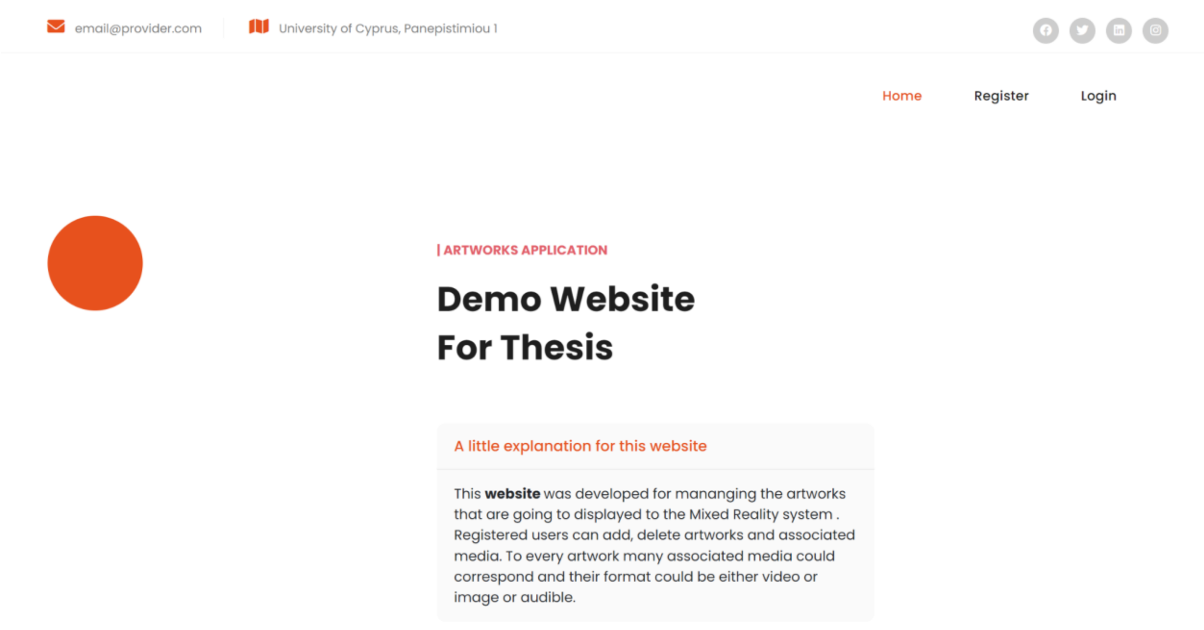

Figure 4.24 - The home page

4.5.3 Register

The register page could be visited only from not logged-in users and there are some restrictions about the password. The registered data is saved in the user table (which is displayed in Figure 4.1).

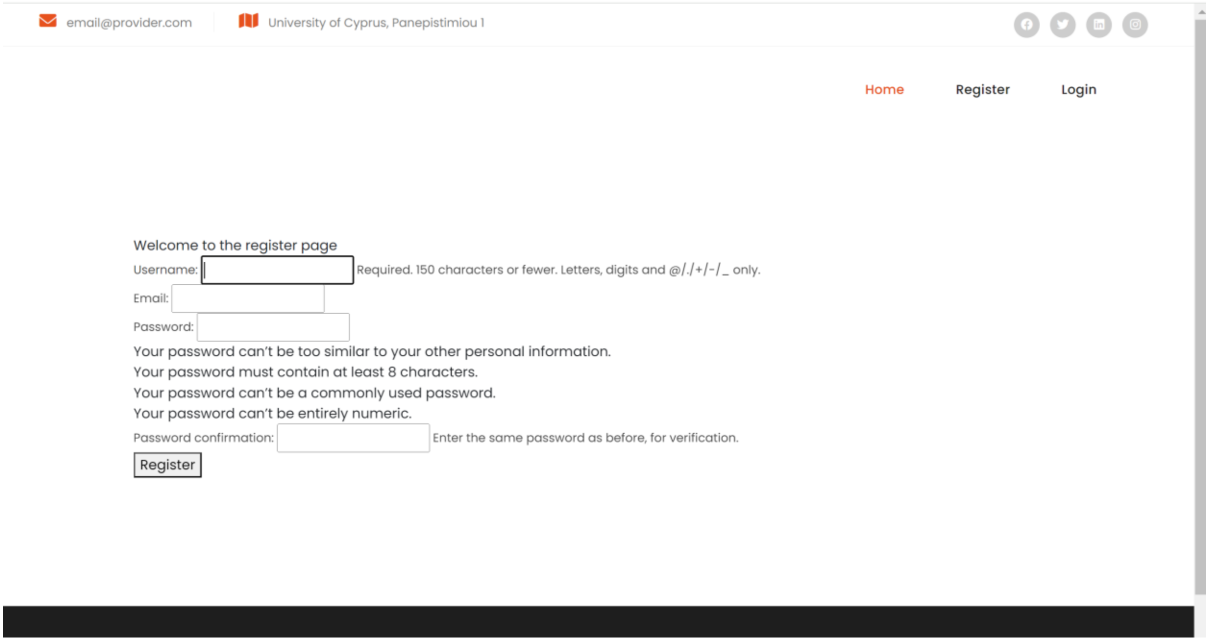

Figure 4.25 - The register page

4.5.4 Login

The login page needs to enter a username and password as is illustrated at **Figure 4.26**. After pressing Login a request is sent to the server and if the credentials are valid the user gets logged in and if they are wrong a failure message is displayed.

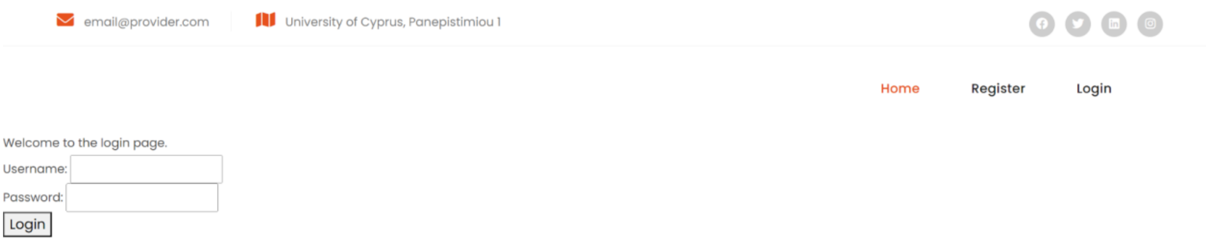

Figure 4.26 - The Login page

4.5.5 Add Artwork

This page allows the user to upload a new artwork which will be stored on the database. It requires inputting a title and an image and to be logged in. After clicking upload a message will be displayed and will inform the user if the upload was done successfully or not. A sample could be viewed at **Figure 4.27**.

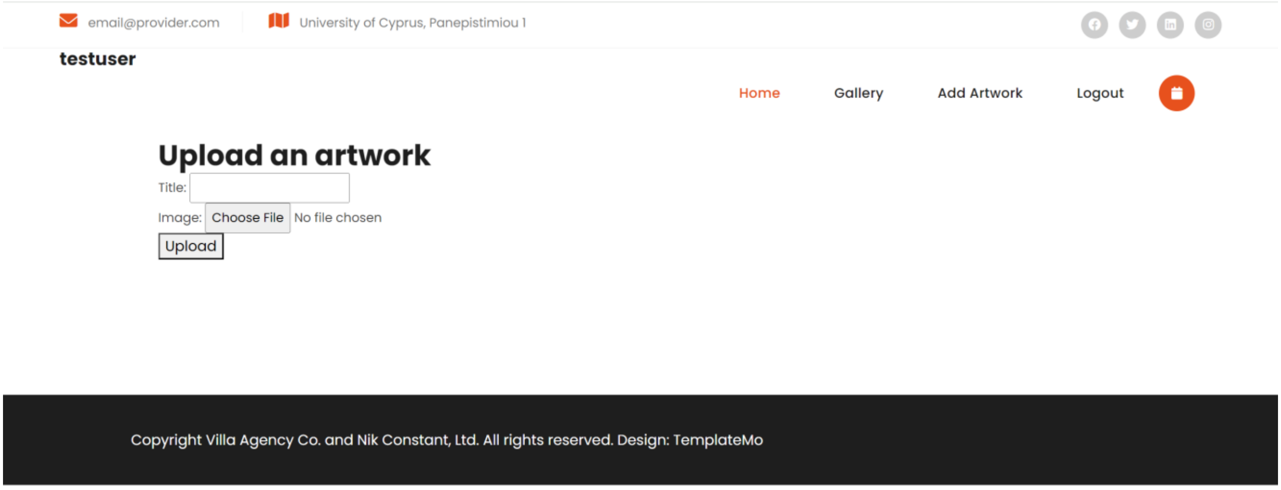

Figure 4.27 - The Add Artwork page

4.5.6 Gallery

Gallery displays all the records of the photo table, in other words the artworks that were uploaded through the “Add Artwork” page. In the table that they are displayed, the user has the capability to add and view associated media of these artworks. The user should be logged in to view this page. A sample could be viewed at **Figure 4.28**.

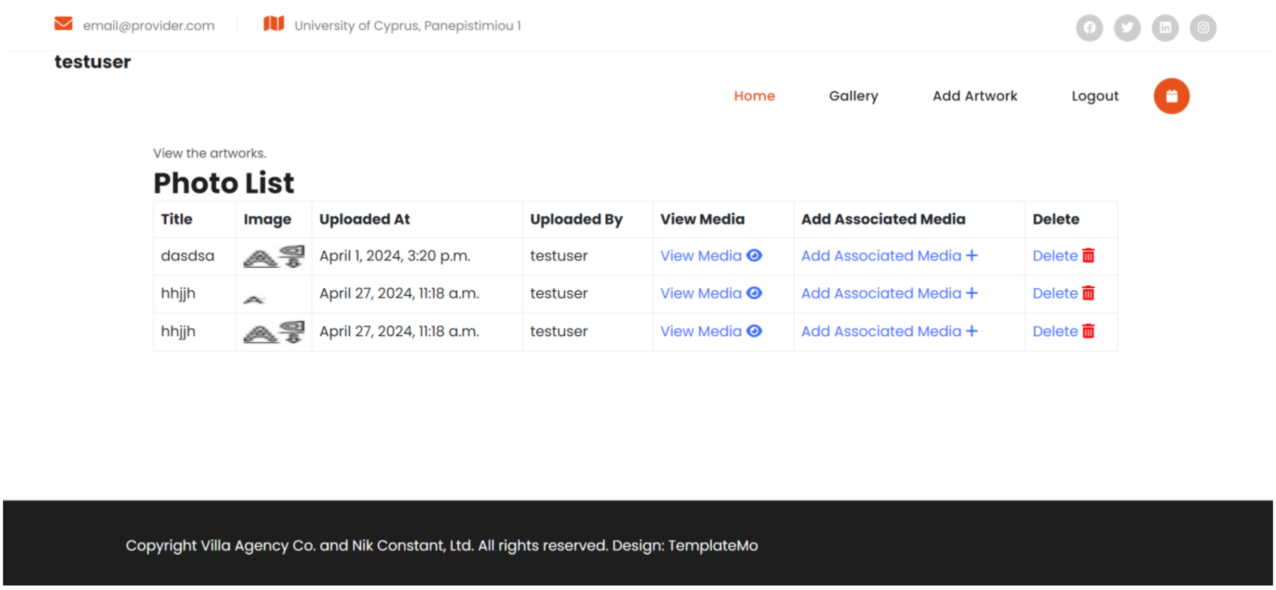

Figure 4.28 - The Gallery page

4.5.7 View Media Page

This page takes as a parameter an artwork and displays its associated media. This is done automatically from the table on the Gallery page. When the view button is clicked a page is opened which illustrates the associated media.

The user has the capability to delete any associated media wants. After that, a message will be displayed that will inform the user about the success of execution of that action.

A sample can be viewed in **Figure 4.29**.

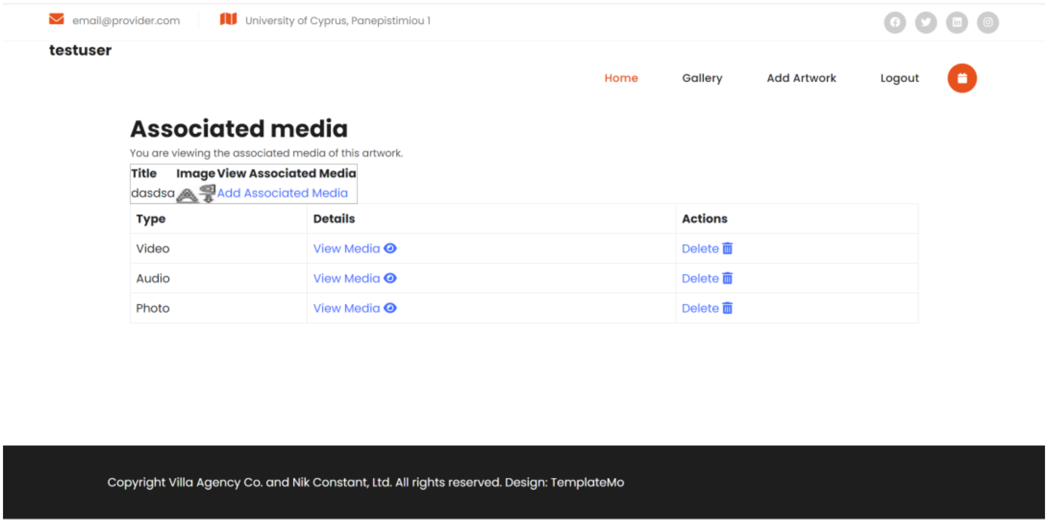

Figure 4.29 - The View Media Page

4.5.8 Add Associated Media

This page allows the user to upload any type of associated media wishes. A sample could be viewed at **Figure 4.30**.

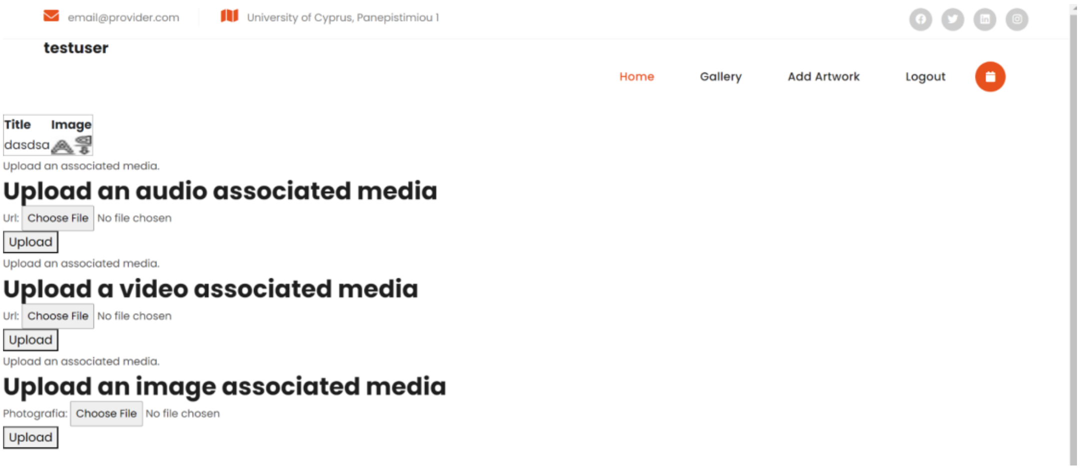

Figure 4.30 - Add Associated Media

Chapter 5

Evaluation

5.1 Introduction

This section contains excerpts derived from a paper I authored, which will appear in the Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (doi: https://doi.org/10.1145/3631700.3664874).

Mixed Reality (MR) technologies have revolutionized how we interact with digital content by creating immersive experiences that merge the real and virtual worlds. In recent years, there has been growing interest in using Artificial Intelligence (AI) to enhance user experiences and reliability in MR. However, the use of Large Language Models (LLMs) in MR applications related to Cultural Heritage (CH) remains relatively unexplored. This study investigates the integration of LLMs in MR environments, specifically focusing on virtual art exhibitions. We developed a HoloLens MR application that allows users to explore artworks while interacting with an LLM using voice commands. To assess user experience and perceived reliability when engaging with an LLM-based virtual art guide, we conducted a study where participants were randomly assigned to either the LLM-based version or a control group using traditional interaction methods. The LLM-based version enables users to ask questions about the displayed artwork, such as details about the artist, the artwork's background, and historical significance. This study discusses the technical aspects of integrating LLMs into MR applications and evaluates how this approach enhances the exploration of virtual art exhibitions by examining user experience and perceived reliability. Initial evaluations suggest that integrating LLMs into MR applications has positive implications. This research contributes to advancing MR technologies for the creation of interactive personalized art experiences in the future.

This research is motivated by the growing interest in using AI technologies to enhance user experience and reliability in cultural settings. Several studies have investigated the use of AI-assisted technologies in the Cultural Heritage (CH) field. For instance, [23] proposed a tailored conversational agent to assist remote museum visitors, employing natural language processing techniques to handle interactions and a hybrid recommender system to increase engagement with the museum's collection. In [24], the authors suggested an interactive museum system allowing voice-based interaction for querying the museum's knowledge base. Another study introduced an AI-based chatbot prototype for museum use, facilitating conversational interaction via text and speech inputs for visitors to inquire about artworks, with responses provided in various formats (text, audio, images, videos) to enrich the visitor experience [25]. A recent project explored GPT-4 as a digital storytelling tool, examining its potential as both a museum guide and a recommender system [26]. However, the integration of Large Language Models (LLMs) into Mixed Reality (MR) applications within the CH domain remains relatively limited. By integrating LLMs into the MR environment, our goal is to develop a dynamic and interactive platform where users can explore artworks, ask questions, and receive insightful and personalized responses in nearly real-time.

5.2 CulturAI

The CulturAI client application is an MR application designed for user interaction with digital artworks. It is compatible with Microsoft HoloLens (1st gen) devices and developed using Unity (v2019.4.0), Mixed Reality Toolkit, and C#. This application presents digital artworks along with their descriptions in written form, depicted in **Figure 5.1**. Furthermore, it enables interaction with the LLM through the following features:

1. Voice Question Recording: Each artwork includes a record button that activates the HoloLens microphone, allowing users to record voice questions.
2. LLM Interaction: After capturing the voice recording, the client sends an HTTP request containing the recorded question as a Waveform Audio (WAV) file to the server for processing and generating the answer.

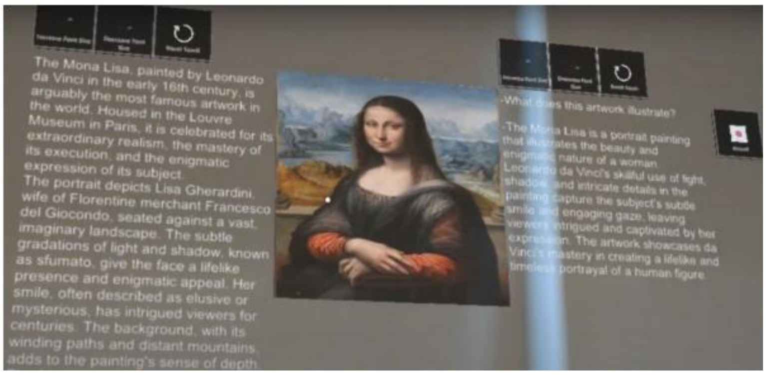

Figure 5.1 –What does this artwork illustrate was asked and an answer was returned.

The server is a web application developed using Python 3.10 and the Django Framework, incorporating the OpenAI Python library for GPT-3.5 integration, and Chroma DB for storing artwork description embeddings. Its primary functions are as follows:

1. Hosting Digital Artworks and Descriptions:
   - The server stores digital artworks and their descriptions. These artworks are retrieved and displayed in the client application, along with their associated descriptions provided in JSON format.
2. Processing Voice Recordings and Generating Answers:
   - Upon receiving an artwork identifier and a WAV file (containing the user's question), the server converts the WAV file content into bytes.
   - The server then sends these bytes to OpenAI for speech-to-text conversion to detect the user's question from the audio.
   - The resulting text question is converted into an embedding, and the server calculates the mean square distance with each sentence of the related artwork description stored in Chroma DB.
   - Next, the server communicates with OpenAI's LLM, passing the three descriptions with the lowest mean square distances and the text question as input parameters, to request an answer.
   - Finally, the server sends back the detected question along with the corresponding answer to the client application.

5.3 Study

A study was conducted with 39 participants, comprising 17.95% female and 82.05% male, aged between 20 and 25 years old. All participants were undergraduate or postgraduate students, and participation was voluntary with the option to leave at any time. The study adhered to the University's human research protocol, ensuring privacy, confidentiality, and anonymity. Participants provided informed consent by signing a consent form before taking part.

The study employed a between-subject design, randomly assigning participants into two groups. The experimental group (20 participants) used the LLM-based version of the application, while the control group (19 participants) used a conventional method to interact with displayed artworks. The selection of artworks included ten of the most famous paintings of all time, with five unique artworks assigned to each group.

The study proceeded through several steps:

i) Participants were informed by the researcher that their data would be used anonymously for research purposes, and they signed a consent form.

ii) Participants were introduced to the study's objectives and familiarized themselves with the HoloLens device.

iii) Participants engaged with the virtual art exhibition using the CulturAI application. To enhance the study's realism, participants were asked to respond to five multiple-choice questions about the artworks. Example questions included identifying the creator of an artwork, the museum where it is housed, and the event that inspired its creation. In the control group, participants read the artwork descriptions and selected answers from multiple-choice options. In contrast, participants in the experimental group asked questions verbally and received answers from the LLM, displayed alongside the artwork's description. They then considered the LLM's response when selecting their answer from the multiple-choice options.

iv) Following the interaction, participants completed a questionnaire to provide feedback on their experience. This feedback aimed to evaluate overall user experience and perceived reliability, particularly for the experimental group. Each session lasted approximately 15 minutes.

5.4 Results

To evaluate the effectiveness of the CulturAI solution, we utilized established methods for assessing user experiences within Mixed Reality (MR) applications in the Cultural Heritage (CH) domain [27]. Accordingly, we designed a questionnaire aimed at measuring various aspects including flow (the state of deep engagement in a task), presence (the cognitive experience in a media-mediated environment), natural interaction, user experience as per the User Experience Questionnaire (UEQ) [28], and the perceived trustworthiness of individuals interacting with an LLM-based agent.

The questionnaire employed a Likert scale ranging from 1 to 7, where 1 indicated Strongly Disagree and 7 indicated Strongly Agree. In terms of the flow dimension, which assessed the balance between challenge and skill, as well as the absorption in the task, participants from the experimental group generally scored higher than those from the control group in most questions. Regarding the presence and natural interaction dimensions, participants showed positive sentiments towards the inclusion of a smart agent in MR environments.

Given our experimental setup with one categorical independent variable (experimental versus control groups) and an ordinal dependent variable with 7 values (Likert scale from 1 to 7), we conducted Mann-Whitney U tests to compare the means of the two groups for each question. However, the analysis of results indicated no significant differences observed between the experimental and control groups. Table 1 summarizes the outcomes related to the flow, presence, and natural interaction dimensions.

To assess user experience, we utilized the short version of the User Experience Questionnaire (UEQ) [28], which measures pragmatic quality, hedonic quality, and overall quality across several dimensions: obstructive/supportive, complicated/easy, inefficient/efficient, confusing/clear, boring/exciting, not interesting/interesting, conventional/inventive, and usual/leading edge.

In the control group, the overall quality score was 1.606 (pragmatic quality: 1.600; hedonic quality: 1.613), whereas in the experimental group, the overall quality score was 2.007 (pragmatic quality: 1.961; hedonic quality: 2.053). The results of the UEQ assessment are depicted in **Figure 5.2**.

In assessing the perceived trustworthiness of CulturAI, we utilized state-of-the-art methods for evaluating trust in AI-based systems [29]. Specifically, for the experimental group, we included an additional section in the questionnaire to investigate perceived reliability, perceived technical competence, perceived understandability, and faith in the system. This section of the questionnaire also employed a Likert scale ranging from 1 to 7, where 1 represented Strongly Disagree and 7 represented Strongly Agree.

The outcomes indicated that the proposed LLM-based approach demonstrated high scores in terms of perceived reliability, perceived technical competence, perceived understandability, and faith. Table 2 summarizes the findings related to the dimension of perceived trustworthiness.

Dimension	Question	Control Group (Mean ± SD)	Experimental Group (Mean ± SD)
Flow	I felt I could meet the requirements of the situation.	6.30 ± 0.84	**6.47 ± 0.75**
	I felt I was in control over the situation.	6.05 ± 1.02	**6.37 ± 0.67**
	I knew exactly what I had to do, and I acted accordingly.	6.40 ± 1.02	**6.53 ± 0.60**
	This task was not too difficult.	**6.40 ± 1.11**	6.32 ± 1.34
	I could effortlessly perform well.	6.10 ± 0.77	**6.26 ± 0.96**
	It was boring for me.	1.95 ± 1.60	**2.05 ± 1.57**
	I found the task interesting.	**6.65 ± 0.57**	6.26 ± 0.85
	I forgot about the progress of time all along.	**5.30 ± 1.93**	4.47 ± 2.21
	Time passed faster than I thought it did.	**5.80 ± 1.47**	5.05 ± 2.01
	I forgot about my close environment.	4.15 ± 1.53	**4.47 ± 2.06**
Presence	When I had a question I had immediately an answer.	5.55 ± 1.20	**6.00 ± 1.21**
	The phrases that were provided were realistic.	6.30 ± 0.78	**6.42 ± 0.82**
	The dialogue with the smart agent made me feel that I was having a realistic conversation.	N/A	4.84 ± 1.57
Natural Interaction	The integrated composition of interacting with smart agent mixed-reality content helped understanding.	N/A	6.00 ± 1.05
	Chatbot functionality is considered necessary in mixed-reality experiences.	5.55 ± 1.32	**6.11 ± 0.79**
	The Q & A experience with chatbots was convenient for understanding artifacts.	N/A	6.37 ± 0.83

Table 1 - Summary of the evaluation results regarding the dimensions of flow, presence, and natural interaction.

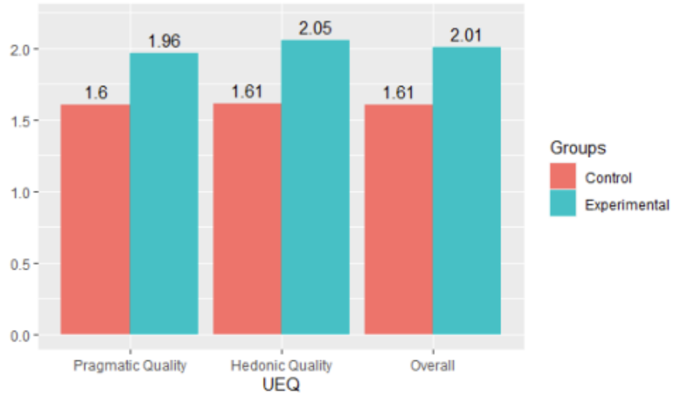

Figure 5.2 - Summary of the evaluation results regarding the dimension of perceived trustworthiness.

Dimension	Question	Experimental Group (Mean ± SD)
Perceived Reliability	The system always provides the advice I require to make my decision.	6.10 ± 1.15
	The system performs reliably.	6.32 ± 1.11
Perceived Technical Competence	The advice the system produces is as good as that which a highly competent person could produce.	5.74 ± 1.19
	The system correctly uses the information I enter.	6.68 ± 0.75
Perceived Understandability	I understand how the system will assist me with decisions I have to make.	5.95 ± 1.22
	It is easy to follow what the system does.	6.37 ± 1.01
Faith	I believe advice from the system even when I don't know for certain that it is correct.	5.95 ± 1.08
	When the system gives unusual advice I am confident that the advice is correct.	5.26 ± 1.48

Table 2 - Summary of the evaluation results regarding the dimension of perceived trustworthiness.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In conclusion, this thesis presents the development of a mixed reality (MR) application designed to display digital artworks, alongside a comprehensive website for managing these artworks. The integration of large language models (LLMs) in MR was explored in the context of enhancing user interaction, providing a richer and more intuitive experience compared to a non-LLM-based system. Together, these contributions advance the field of digital art presentation, offering innovative solutions that bridge technology and the world of art in new and impactful ways.

6.2 Limitations

Despite the promising results and innovative features of the mixed reality (MR) application for displaying digital artworks, several limitations need to be acknowledged.

Limitations:

1. The HoloLens 1 device is having some hardware limitations. I found that when something while I was downloading something in the application (for instance photos), the Frames per Second (FPS) were dropped immediately.
2. The application needs a connection to the internet.
3. The operating system (OS) had many restrictions. For instance it was not possible to write on the disk from the application that I developed.

6.3 Future Work

The study presented the development and initial assessment of CulturAI, a HoloLens Mixed Reality (MR) application designed for virtual art exhibitions. CulturAI integrates a Large Language Model (LLM), enabling users to interact with virtual art exhibitions using voice commands to create personalized and immersive experiences.

Using a between-subject study design, we assessed user experience and perceived reliability of this approach. Analysis of the findings showed the benefits of incorporating LLMs into MR applications, leading to enhanced user experiences of trust when engaging with virtual art exhibitions. However, the study's limitations include the brief duration of user interactions. Longitudinal studies in the future could offer insights into the sustained effectiveness of CulturAI. Additionally, future efforts will explore leveraging LLMs to develop adaptive narratives and dynamic features that respond to user interactions, interests, and preferences.

Regarding the Mixed Reality application, the integration of a 3D gallery is a future task. More specifically, a 3D scanner could be used for scanning objects in to the real world and after the exhibition of it should be possible through the virtual gallery.

Βιβλιογραφία

[1] Agarwal, Anmol, et al. "Mobile application development with augmented reality." *Int. J. Comput. Sci* 2 (2014): 20-25.

[2] Birchfield, David, et al. "Mixed-reality learning in the art museum context." *Proceedings of the 16th ACM international Conference on Multimedia*. 2008.

[3] Wikipedia. “One-to-many (data model)”

[4] Hughes, Charles E., et al. "Mixed reality in education, entertainment, and training." *IEEE computer graphics and applications* 25.6 (2005): 24-30.APA

[5] Ganapathy, Apoorva. "Virtual Reality and Augmented Reality Driven Real Estate World to Buy Properties." *Asian Journal of Humanity, Art and Literature* 3.2 (2016): 137-146.

[6] Geroimenko, Vladimir. "Augmented reality technology and art: The analysis and visualization of evolving conceptual models." *2012 16th International Conference on Information Visualisation*. IEEE, 2012.

[7] Hughes, Charles E., et al. "Mixed reality in education, entertainment, and training." *IEEE computer graphics and applications* 25.6 (2005): 24-30.

[8] IBM, “Three-tier architecture”

[9] Jain, Shubham, et al. "Omnichannel retail customer experience with mixed-reality shopping assistant systems." *Advances in Visual Computing: 16th International Symposium, ISVC 2021, Virtual Event, October 4-6, 2021, Proceedings, Part I*. Springer International Publishing, 2021.

[10] Jiang, Yirui, et al. "Enhancing the Customer Experience by Mixed Reality in the Retail Industry." *2021 5th International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 2021.

[11] Jiawei, Wang, and Nur Azlina Mohamed Mokmin. "Virtual reality technology in art education with visual communication design in higher education: a systematic literature review." *Education and Information Technologies* 28.11 (2023): 15125-15143.

[12] Jiawei, Wang, and Nur Azlina Mohamed Mokmin. "Virtual reality technology in art education with visual communication design in higher education: a systematic literature review." *Education and Information Technologies* 28.11 (2023): 15125-15143.

[13] Juraschek, Max, et al. "Mixed reality in learning factories." *Procedia Manufacturing* 23 (2018): 153-158.

[14] Liu, Qiang. "The virtual reality technology in art design." *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*. IEEE, 2012.

[15] Papasarantou, Chrissa, Rene Alimisi, and Dimitris Alimisis. "Virtual Galleries as Learning Scaffolds for Promoting Problem-Based Learning." *Education Sciences* 13.12 (2023): 1168.

[16] Resolution Games, "Spatial Computing & Mixed Reality Games"

[17] Dunston, Phillip S., and Xiangyu Wang. "Mixed reality-based visualization interfaces for architecture, engineering, and construction industry." *Journal of construction engineering and management* 131.12 (2005): 1301-1309.

[18] Sankowska, Patrycja-Jadwiga. "Mixed realities. Application of geospatial augmented reality in urban planning. Case study: finding places." *INCREaSE 2019: Proceedings of the 2nd International Congress on Engineering and Sustainability in the XXI Century*. Springer International Publishing, 2020.

[19] DevOpsSchool. "What is SQLite and use cases of SQLite?"

[20] Trushchenkov, Igor, et al. "Using virtual reality systems for crime scene reconstruction." *Creativity in Intelligent Technologies and Data Science: 4th International Conference, CIT&DS 2021, Volgograd, Russia, September 20–23, 2021, Proceedings 4*. Springer International Publishing, 2021.

[21] Gulhan, Doga, Szonya Durant, and Johannes M. Zanker. "Aesthetic judgments of 3D arts in virtual reality and online settings." *Virtual Reality* 27.2 (2023): 573-589.

[22] Tremosa. "Beyond AR vs. VR: What is the Difference between AR vs. MR vs. VR vs. XR?"

[23] Trichopoulos, Georgios. "Large language models for cultural heritage." *Proceedings of the 2nd International Conference of the ACM Greek SIGCHI Chapter*. 2023.

[24] Yi, Ji Hyun, and Hae Sun Kim. "User experience research, experience design, and evaluation methods for museum mixed reality experience." *Journal on Computing and Cultural Heritage (JOCCH)* 14.4 (2021): 1-28.

[25] Pisoni, Galena, et al. "Human-centered artificial intelligence for designing accessible cultural heritage." *Applied Sciences* 11.2 (2021): 870.

[26] Sylaiou, Styliani, et al. "From physical to virtual art exhibitions and beyond: Survey and some issues for consideration for the metaverse." *Journal of Cultural Heritage* 66 (2024): 86-98.

[27] Yi, Ji Hyun, and Hae Sun Kim. "User experience research, experience design, and evaluation methods for museum mixed reality experience." *Journal on Computing and Cultural Heritage (JOCCH)* 14.4 (2021): 1-28.

[28] Schrepp, Martin, Andreas Hinderks, and Jörg Thomaschewski. "Applying the user experience questionnaire (UEQ) in different evaluation scenarios." Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I 3. Springer International Publishing, 2014.

[29] Valadi, Shokoufeh, and Ing David Broneske. "Analysis of Current Usability and User Experience Questionnaires and Creating an Optimized Usability Questionnaire." *Magdeburg: Universität Magdeburg* (2020).

[30] Gerup, Jaris, Camilla B. Soerensen, and Peter Dieckmann. "Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review." *International journal of medical education* 11 (2020): 1.

[31] Tang, Yuk-Ming, et al. "Evaluating the effectiveness of learning design with mixed reality (MR) in higher education." *Virtual Reality* 24.4 (2020): 797-807.

[32] Bäck, Regina, et al. "Mixed reality in art education." *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019.

[33] Rathnayake, WGRM PS. "Usage of mixed reality for military simulations." *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*. IEEE, 2018.

[34] Boyce, Michael W., et al. "Enhancing military training using extended reality: A study of military tactics comprehension." *Frontiers in Virtual Reality* 3 (2022): 754627.

[35] Stone, Robert J., et al. "A 'mixed reality'simulator concept for future Medical Emergency Response Team training." *BMJ Military Health* 163.4 (2017): 280-287.

[36] Fan, Lei, et al. "Multi-vehicle cooperative military training simulation system based on augmented reality." *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 2019.

[37] Gil-López, Cristina, et al. "Is mixed reality technology an effective tool for retail? A vividness and interaction perspective." *Frontiers in Virtual Reality* 4 (2023): 1067932.

[38] Wolf, Mario, Heinrich Söbke, and Florian Wehking. "Mixed reality media-enabled public participation in urban planning: A literature review." *Augmented reality and virtual reality: Changing realities in a dynamic world* (2020): 125-138.

[39] Sareika, Markus, and Dieter Schmalstieg. "Urban sketcher: Mixed reality on site for urban planning and architecture." *2007 6th IEEE and ACM international symposium on mixed and augmented reality*. IEEE, 2007.

[40] Kieferl, Joachim, and U. Wossner. "Mixed realities: improving the planning process by using augmented and virtual reality." *Proceedings of the 4th Conference on new technologies in landscape architecture*. 2003.