An Immersive Virtual Try-On System with Augmented Reality

Rana Shiba
Artificial Intelligence, College of Information Technology, Misr University for Science and Technology (MUST), Cairo, Egypt
89597@must.edu.eg

Ahmed Fadel
Computer Science, College of Information Technology, Misr University for Science and Technology (MUST), Cairo, Egypt.
94118@must.edu.eg

Ahmed Awad
Computer Science, College of Information Technology, Misr University for Science and Technology (MUST), Cairo, Egypt
94091@must.edu.eg

Kyrillos Emad
Computer Science, College of Information Technology, Misr University for Science and Technology (MUST), Cairo, Egypt
94289@must.edu.eg

Rania El Gohary
Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
Rania.elgohary@cis.asu.edu.eg

***Abstract*—Virtual try-on systems using augmented reality (AR) have garnered significant attention in recent years for their potential to revolutionize the fashion industry. Augmented reality is a technology that superimposes digital information, such as images, videos, or 3D models, onto the real-world environment, enhancing or augmenting the user's perception of reality. Unlike virtual reality (VR), which creates entirely immersive digital environments, AR overlays virtual content onto the physical world, allowing users to interact with both the real and virtual elements simultaneously. AR combines computer vision, image processing, and graphics rendering techniques to overlay digital content onto the real-world environment in real-time. By leveraging these techniques, AR systems enable users to virtually try on products in real-time, enhancing the shopping experience and reducing the need for physical try-ons. Despite obstacles such as precise mapping of the 3D model onto the user's body and high development techniques to accurately represent the product throughout the real-time feed, AR virtual try-on systems offer numerous benefits, including time savings, improved customer satisfaction, and enhanced online sales conversions. By implementing an AR solution that not only aims to fit 3D products onto the user but also enhances previously implemented studies by applying it to the real-time environment, our study focuses on achieving high accuracy results through extensive deep learning techniques and aims to reduce customer challenges for an improved shopping experience. Virtual try-on systems with AR hold promises beyond the fashion industry, with potential applications in cosmetics, eyewear, and home design.**

Keywords— Virtual Try-On, Augmented Reality, MediaPipe, Three.js, Fashion.

Abstractions:

VTO	***Virtual Try-On***
AR	***Augmented Reality***
ERD	***Entity-Relationship Diagram***
TAM	***Technology Acceptance Model***

I. INTRODUCTION

The integration of augmented reality (AR) technologies into virtual try-on systems marks a transformative era in consumer interaction within the fashion industry. By harnessing sophisticated algorithms for pose tracking and 3D rendering, these systems offer an immersive and interactive shopping experience that closely mimics the in-store experience. This paper introduces a cutting-edge approach to the development of a real-time virtual try-on system - for glasses and watches - that seamlessly affiliates company products directly within the VTO web browser or integrates the VTO service into an e-commerce platform, enhancing user engagement and bridging the gap between digital conveniences and the tactile appeal of physical/online shopping.

In AR, scholars agree on the following three main characteristics: (1) a combination of real and virtual worlds, (2) real-time interactivity, and (3) computer-generated content [1] that interacts with physical surroundings. In contrast, in a VR experience, the existing surroundings play a minor role as a completely virtual environment is created. Thus, AR can be positioned between reality and VR (see Figure 1)[2]

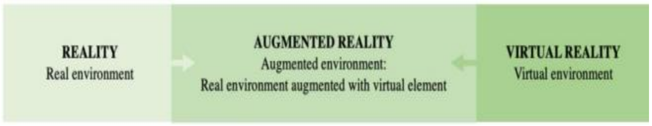

Fig.1. Comparison of Reality, Augmented Reality, and Virtual Reality (MICHELLE ET AL. [3]).

Virtual try-on systems have gained significant traction by enabling consumers to visualize how items would look and fit on their bodies without the need for a physical try-on. Our study utilizes AR technology to overlay virtual items onto live video feeds captured from the user's device, providing an effective try-on experience that allows users to customize styles and colors[4].

Providing consumers with an accurate virtual try-on experience solves many significant challenges. Traditional online shopping lacks the ability for buyers to physically try on items, leading to uncertainty about fit, style, and overall satisfaction. This can result in high return rates, adding extra effort and time, for both customers and brands, and reducing trust and confidence in the purchasing process. Moreover, long queues and wait times in physical stores pose significant challenges, leading to customer dissatisfaction and potential loss of future business. Limited accessibility is also a major concern, particularly for people with disabilities, who may find it difficult to navigate crowded stores or reach items on shelves. This also applies to individuals with hearing or visual impairments, for whom physical shopping can be frustrating. While physical stores remain important, online shopping with virtual try-on offers a valuable alternative, providing greater accessibility, convenience, and visualization, particularly for those with mobility challenges.

The rise of e-commerce has significantly presented persistent challenges that necessitate solutions. One such solution is the introduction of Virtual Try-On Systems with Augmented Reality (AR), which promises to revolutionize online shopping. Based on the research insights, the aim of the study is to develop a 3D and AR fashion platform offering an interactive, and engaging shopping experience that will facilitate social sharing of virtual try-on experiences and integrate brand products seamlessly within the virtual try-on system. AR overlays virtual images onto real-time video feeds, allowing customers to visualize products from multiple angles and perspectives, enhancing understanding and engagement.

Businesses adopting AR technology gain a competitive advantage by reducing returns, increasing satisfaction, and providing standardized sizing experiences. AR-powered virtual try-ons fosters trust and engagement in e-commerce.

II. Related Works

A detailed examination of the survey paper by Michelle et al. [1] was undertaken, aiming to extract pivotal insights that will inform and enhance the development of our study's concept. The objective of the survey was to explore distinctions between two generational groups, namely millennials and Generation Z (see Figure 2 for the gender distribution), concerning their experiences with augmented reality (AR) through virtual try-ons (VTO) in the context of online retail.

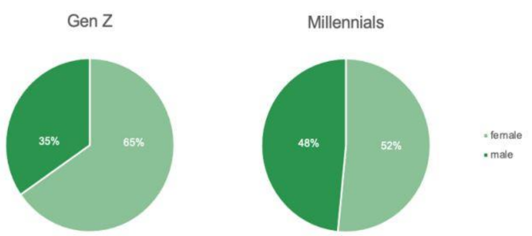

Fig.1. Gender Distribution Among Millennials and Gen Z (MICHELLE ET AL. [3]).

A questionnaire was conducted to assess prior AR experiences among millennials and Gen Z, revealing significant insights. Utilizing a quantitative online survey and experiment, the study incorporates variables from the Technology Acceptance Model (TAM) and AR-specific measures, as their workflow is represented in Figures 3 and 4.

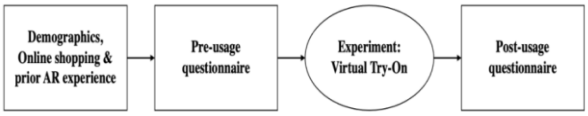

Fig.2. Structure of Online Survey (MICHELLE ET AL. [3]).

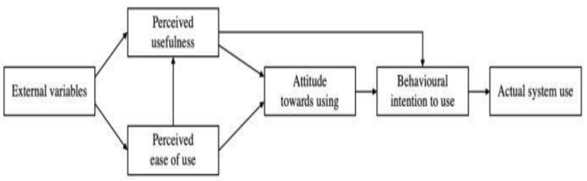

Fig.3. Technology Acceptance Mode (MICHELLE ET AL. [3]).

According to Figure 5, 80.3% of participants have had some form of AR experience. Notably, a larger percentage of Gen Z (25.7%) are unaware of AR compared to millennials (13.4%), while 40.2% of millennials have tried AR, indicating a greater exposure among this demographic. Furthermore, Figure 6's bar chart shows that, on a 7-point Likert scale, both generations scored above the neutral midpoint of 4 across various items. Gen Z generally scored higher than millennials, except in areas such as post-usage hedonic motivation, spatial presence, and the need for touch, where millennials outscored their younger counterparts.

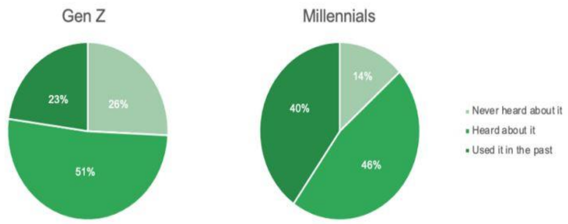

Fig.4. Previous AR Experience Among Millennials And Gen Z (MICHELLE ET AL. [3]).

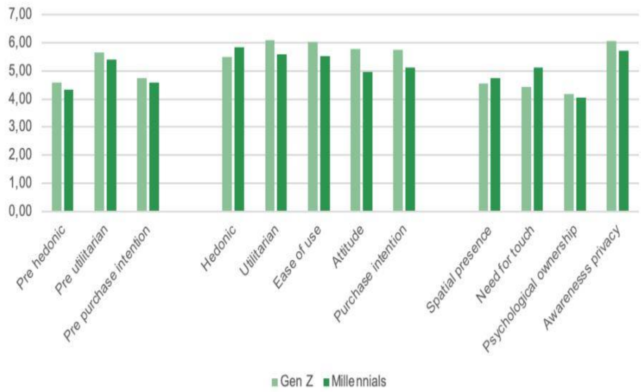

Fig.5. Overview Of Mean Scores (MICHELLE ET AL. [3]).

The bar chart in Figure 6 and results in Table 1, illustrate the mean scores of both generations which confirms that VTOs significantly increased hedonic and utilitarian values and purchase intentions for both generations. Gen Z's higher tech-savviness and familiarity with technology contributed to their elevated higher mean scores than millennials across all items, except for post-usage hedonic motivation, spatial presence, and need for touch, where millennials score higher. Data privacy awareness remained equally important for both generations.

TABE 1. Overview Of Mean Scores Comparing Generations (Michelle et al. [3]).

Measure	Gen Z Mean (SD)	Millennials Mean (SD)	Difference in means d	Total sample Mean (SD)
Pre-usage variables				
Prehedonic	4,58 (0,97)	4,33 (1,12)	***0,25***	4,46 (1,05)
Preutilitarian	5,66 (1,05)	5,42 (0,96)	***0,25***	5,54 (1,01)
Prepurchaseintention	4,74 (1,16)	4,57 (1,31)	***0,17***	4,66 (1,23)
Post-usage variables				
Hedonic	5,5 (0,93)	5,83 (0,81)	***-0,33***	5,66 (0,89)
Utilitarian	6,08 (0,9)	5,59 (1,05)	***0,49***	5,84 (1,01)
Easeofuse	6,02 (0,64)	5,53 (1,09)	***0,49***	5,78 (0,92)
Attitude	5,77 (0,71)	4,96 (1,09)	***0,81***	5,37 (1)
Purchaseintention	5,75 (0,93)	5,11 (0,89)	***0,64***	5,44 (0,96)
AR-specific variables				
Spatial presence	4,55 (1,24)	4,75 (1,17)	***-0,2***	4,65 (1,2)
Need for touch	4,42 (1,54)	5,11 (1,03)	***-0,69***	4,77 (1,35)
Psychological ownership	4,17 (1,3)	4,04 (1,61)	***0,13***	4,11 (1,46)
Awareness for privacy concerns	6,07 (1,14)	5,7 (1,44)	***0,37***	5,89 (1,31)

Practical implications include the need for tailored strategies when targeting each generation. For millennials, emphasizing the playful and joyful aspects of AR is crucial, while for Gen Z, focusing on utilitarian values and enhancing product information is recommended. The survey contributes to the fields of AR and generational marketing by highlighting differences between millennials and Gen Z.

This survey study confirms our hypothesis of the potential of AR to enhance the environmental sustainability of online retail. VTOs, by facilitating the evaluation of product fit, can reduce returns, leading to decreased shipping, lower emissions, and reduced customer travel to stores. Concluding the survey analysis, if AR is implemented appropriately, it can be an effective and environmentally sustainable marketing tool.

In preparation for our study, a comprehensive search was conducted to identify relevant related work papers within the domain of virtual try-on systems and augmented reality (AR) technologies. Each selected paper provides valuable insights into different aspects of virtual try-on, including pose-tracking algorithms, image processing techniques, 3D rendering methodologies, and user experience considerations.

The studies have explored different approaches to virtual footwear try-ons in AR. Marelli, D et al. [5] proposes an eyewear virtual try-on experience based on advanced deep learning and computer vision techniques for a realistic fit estimation. Anand Piyush et al. [6] presents an AI-based virtual try-on web application for glasses, utilizing AR and web technologies for a real-time, immersive shopping experience. An et al. [7] introduces a real-time AR system for virtual shoe try-on on smartphones, focusing on footwear with novel methods for occlusion generation and stabilization. Chou et al. [8] addresses the challenge of realistically rendering shoes on users in various poses with a novel dataset and conditional image completion techniques. Greci et al. [9] develops a mechanical haptic device for simulating the four fundamental dimensions of a shoe.

Moreover, as per research, it has been concluded that the Virtual Try-on System has not been implemented in Egyptian Local Companies: only implemented by international companies based in Egypt. This gives us more motive to implement our study and revolutionize the Egyptian fashion industry. By bringing this technology to local companies, we can enhance the customer experience, provide a more innovative and interactive shopping experience, and ultimately increase sales and brand loyalty. Overall, implementing the Virtual Try-on System in Egyptian local companies will not only benefit the fashion industry but also contribute to the country's economic development and technological advancement[10].

As illustrated in Table 2, a comparative study meticulously outlines the distinctions between the features of various papers and our study, emphasizing the diversity in approach methods and their subsequent impacts. This comparison is crucial for understanding the uniqueness and contributions of our study within the broader research landscape. By examining the methods employed and the effects achieved, Table 1 provides a comprehensive overview of how our study differentiates itself from existing literature, highlighting the innovative aspects of our approach and the significant advancements it offers. This analysis not only showcases the novelty of our study but also situates it within the context of ongoing research, demonstrating its relevance and potential to contribute to the field.

TABLE 2. COMPARATIVE STUDY.

Paper's Author	**Features**		
	Techniques	***Key Findings***	***Limitations***
Marelli, D et. al. [5]	Face Detection, 3D Reconstruction from a Single Image, Face Size Estimation, Fitting Parameter Estimation, using Vue.js, Dropzone.js, and babylon.js.	Emphasizes markerless techniques and web-based solution for an immersive try-on experience, using 3D models for a detailed view.	Challenges with occlusions, reliance on static images, rendering performance issues with complex scenes or high-resolution textures.
Anand Piyush et al. [6]	Facial recognition, AR technology, HTML5, WebGL, WebRTC for real-time interaction and 3D model rendering.	Real-time AR implementation with 3D visual models for a detailed and interactive shopping experience via web technologies.	Technical complexity, reliance on specific hardware, potential tracking errors, limited to footwear customization.
An et. al. [7]	Multi-Branch Network for pose estimation and segmentation, Occlusion Generation, 3D Model Stabilization, 6-DoF Pose Annotation.	Highlights the importance of real-time AR integration, accurate pose estimation, and platform-specific optimizations for mobile applications.	Computational demands of real-time processing, accuracy in pose estimation and occlusion generation.
Chou et. al. [8]	Data Collection and Pre-processing, Detection and Masking, Conditional Image Completion, Loss Function Optimization.	Focuses on accurate rendering of shoes in various poses using a novel dataset, emphasizing the role of deep learning in virtual try-on systems.	Reliance on specific dataset, complexity of generating realistic images from a single view, flexibility, and scalability concerns.
Greci et. Al. [9]	Integration of engineering principles, material science, computational modeling for the development of the FG device, including material characterization and FEA.	Accurate simulation of shoe dimensions and user-friendly design, offering versatility and robust mechanical design for a realistic shoe fitting experience.	Complex design increasing manufacturing complexity, potential high cost, regular maintenance and upkeep required.
Our Proposed System	Real-time video processing with MediaPipe Pose for pose tracking, Three.js for 3D model rendering for immersive web-based AR experiences.	The project the feasibility of a web-based, real-time AR VTO system that offers an immersive and interactive try-on experience across demonstrates	Optimization for Various Real-world Scenarios, Seamless Performance Across Web Platforms

III. METHODOLOGIES

Our research delves into the methodology of a real-time Virtual Try-On (VTO) system enhanced by Augmented Reality (AR) technologies. This system integrates sophisticated pose tracking with advanced 3D rendering techniques to leverage a seamless and interactive virtual fitting experience. Here, we detail the comprehensive approach used to realize this integration effectively[11].

The initial stage in our VTO system involves capturing video frames through the user's camera. This is the foundational step where pose tracking plays a crucial role by accurately detecting and tracking the user's body landmarks in real time. Utilizing the MediaPipe Pose framework, the system analyzes the incoming video feed to identify 33 key body landmarks [12] for wrist landmarks extraction, and Face Mesh Landmarks [13] for face landmarks detection. These landmarks include critical anatomical points such as joints and extremities, essential for the precise alignment of virtual products with the user’s actual physique.

In the illustrated workflow (Figure 10), we demonstrate how the video stream is processed. The captured frames are fed into deep learning algorithms, where MediaPipe Pose calculates the positions of these landmarks. This pose landmark model, as detailed in Figure 7 and Figure 8, is integral to our methodology as it provides a detailed map of the body positions necessary for the next steps in virtual overlay. Figure 8 illustrates the pose landmark model used in MediaPipe Pose for wrist landmarks extraction (including the wrist and elbow landmarks needed for watches VTO) which tracks 33 body landmark locations, representing the precise location of the following body parts:

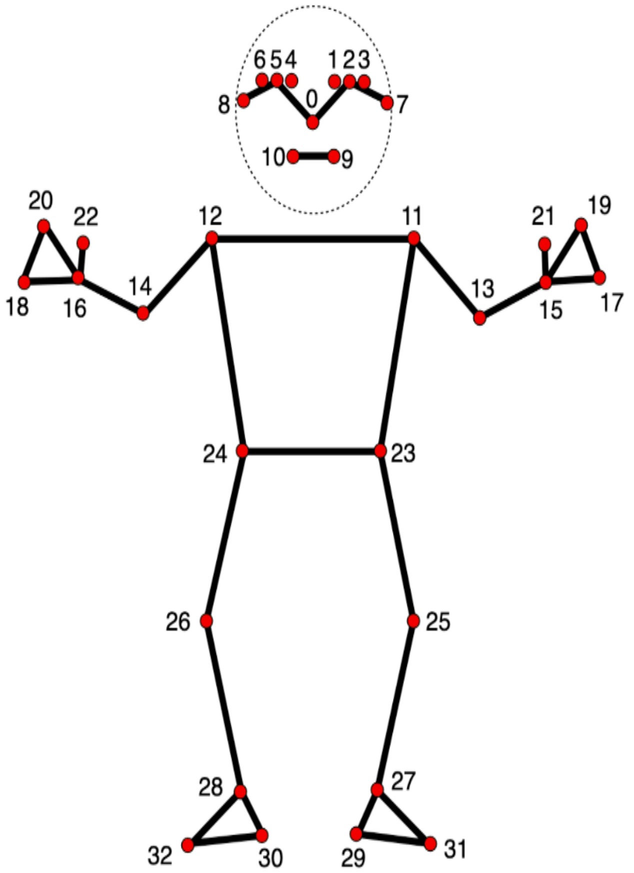

Fig. 6. Pose Landmarks Index. [9]

Whereas, Figure 8 illustrates the face mesh landmarks provided by MediaPipe that are utilized for the glasses rendering onto the user’s face.

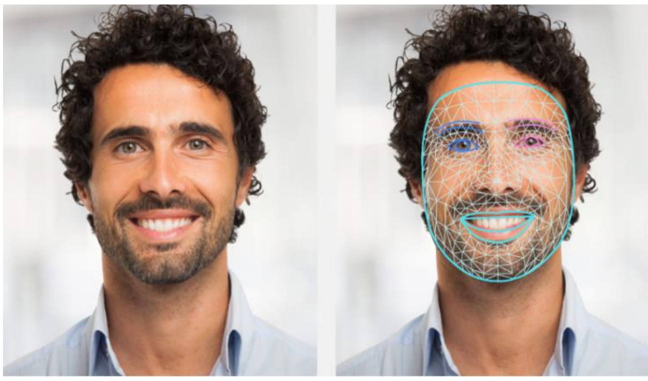

Fig.7. Face Mesh Landmarks. [10]

These key points represent critical anatomical points on the user's body, such as joints or extremities, and serve as the basis for aligning virtual products with the user's physique. Next, the landmarks of the detected locations are marked on the user as illustrated in Figure 9, where the landmarks needed are detected to help indicate the rendering positions of the 3D model.

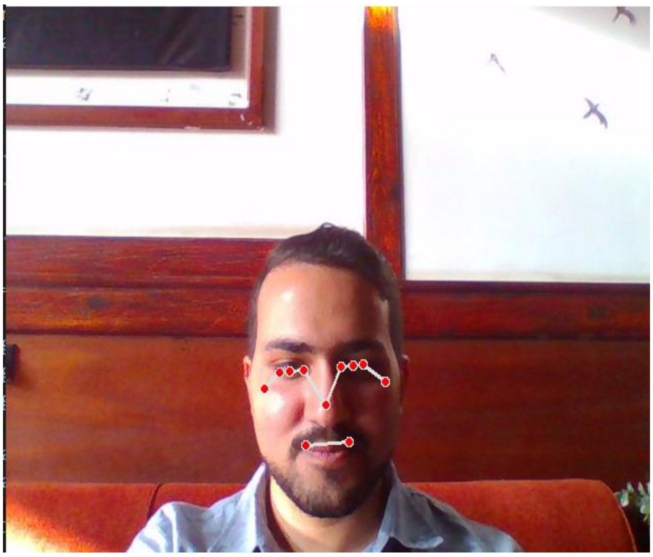

Fig.9. Head Pose Estimation Sample Result.

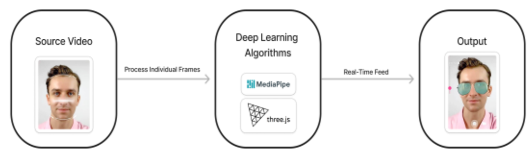

Fig.10. VTO Workflow.

Following the detection of the landmarks, orientations, and rotations, the next critical phase involves the rendering of 3D models corresponding to virtual products, glasses or watches. This is executed using Three.js [14], a powerful library for 3D graphics on the web, which enables the rendering of high-quality 3D models that closely mimic the appearance of real-world objects. The 3D models are dynamically aligned and scaled based on the detected landmarks to ensure that they fit naturally over the users. This dynamic alignment is crucial for maintaining the realism of the virtual try-on experience, ensuring that the virtual objects integrate seamlessly with the user's movements and poses.

Our AR integration technique involves rendering an invisible base 3D model, which acts as a foundational layer over which the actual product models are overlaid. This base model, depicted in Figure 11, is crucial as it helps achieve depth occlusion—blocking out elements behind it, thereby enhancing the realism of the overlay. This effect is essential for creating the illusion that the virtual, 3-dimensional products are physically being worn, as shown in Figure 12. The virtual objects are parented to this base model, meaning they inherit transformations applied to the base. This hierarchical relationship is critical when pose estimation is used to calculate transformations of the base model; the overlaid 3D models will accurately follow the user's movements.

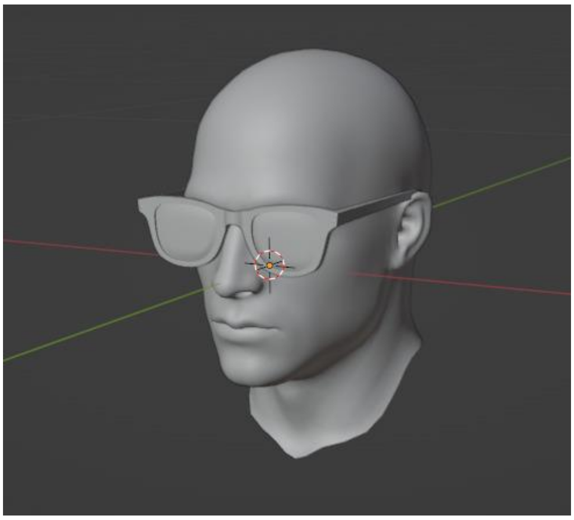

Fig.11. Base 3D Mesh Model.

Fig.12. Result Of Depth Occlusion.

IV. RESULTS

The culmination of these methodologies is depicted in Figure 13, showcasing the application of glasses in a VTO scenario, and Figure 14, showcasing the application of watches in a VTO scenario.

Fig.13. Glasses Virtual Try-On Implementation.

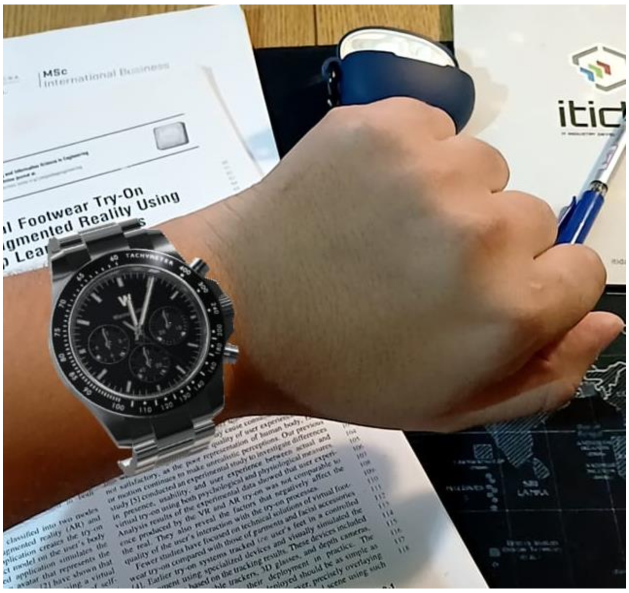

Fig.14. Glasses Virtual Try-On Implementation.

By integrating sophisticated algorithms for pose tracking and 3D rendering within an AR framework, we enable users to try on various products in real-time. This system not only allows for the exploration of different styles and fits but also provides an engaging and interactive shopping experience directly from the user's device when integrated into an e-commerce platform or utilized on our website's product affiliation service.

V. CONCLUSION

In this paper, we presented an augmented reality system for 3D products virtual try-on. After a critical analysis of existing virtual try-on applications that aimed to identify their limitations, we introduced the workflow of an immersive virtual try-on system and described the main modules used to provide the try-on experience to address the previous study's limitations. Our study is developed to detect the user's landmarks in the live stream video, the virtual try-on is performed in the 3D space using a 3D model of the user's face or wrist, rendered through depth occlusion of the invisible 3D base model, allowing the user to interact with the system giving the possibility of observing the results from different points of view. Second, the fitting takes into account the actual size of the frames and the actual size of the user, providing a realistic size fitting. The results of the glasses and watches try-on showed a realistic representation of the 3D product, providing optimal results that ensure high customer satisfaction during online shopping.

Acknowledgments

We would like to convey our sincerest appreciation to Prof. Dr. Rania Al-Gohary for her invaluable guidance and unwavering support during the entire research and development journey. Moreover, we would like to extend our sincere gratitude to Eng. Shaimaa Bahaa and Eng. Fares El-Tomy for their invaluable guidance, support, and expertise throughout the duration of this study. Their encouragement and mentorship have been a source of inspiration, motivating us to strive for excellence and pursue innovative solutions. We are truly grateful for everyone's contributions and honored to have had the opportunity to collaborate with such esteemed professionals. Thank you, Prof. Dr. Rania Al-Gohary, Eng. Shaimaa Bahaa, and Eng. Fares El-Tomy, for your immense dedication to keep us going forward.

References

[1] R. T. Azuma, "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments/MIT press*, 1997.

[2] C. Flavián, S. Ibáñez-Sánchez, and C. Orús, "The impact of virtual, augmented and mixed reality technologies on the customer experience," *J Bus Res*, vol. 100, pp. 547–560, 2019.

[3] M. Schwarz, "Augmented reality in online retail: Generational differences between millennials and generation Z using virtual try-on's," 2022.

[4] G. Xie, H. Attar, A. Alrosan, S. M. F. Abdelaliem, A. A. S. Alabdullah, and M. Deif, "Enhanced diagnosing patients suspected of sarcoidosis using a hybrid support vector regression model with bald eagle and chimp optimizers," *PeerJ Comput Sci*, vol. 10, p. e2455, 2024.

[5] D. Marelli, S. Bianco, and G. Ciocca, "Designing an AI-based virtual try-on web application," *Sensors*, vol. 22, no. 10, p. 3832, 2022.

[6] P. Anand, R. Gupta, N. K. Tiwari, and S. Gupta, "Glass Virtual Try-On," *Int J Res Appl Sci Eng Technol*, vol. 10, no. 5, pp. 718–722, 2022.

[7] S. An *et al.*, "ARShoe: Real-time augmented reality shoe try-on system on smartphones," in *Proceedings of the 29th ACM International Conference on Multimedia*, 2021, pp. 1111–1119.

[8] C.-T. Chou, C.-H. Lee, K. Zhang, H.-C. Lee, and W. H. Hsu, "Pivtons: Pose invariant virtual try-on shoe with conditional image completion," in *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part VI 14*, Springer, 2019, pp. 654–668.

[9] L. Greci, M. Sacco, N. Cau, and F. Buonanno, "FootGlove: a haptic device supporting the customer in the choice of the best fitting shoes," in *Haptics: Perception, Devices, Mobility, and Communication: International Conference, EuroHaptics 2012, Tampere, Finland, June 13-15, 2012. Proceedings, Part I*, Springer, 2012, pp. 148–159.

[10] M. A. Deif, H. Attar, M. A. Hafez, W. Alomoush, and H. Al-Faiz, "Automatic Sarcoidosis Stage Classification Based on Gray Level Co-occurrence Matrix Features," *Appl. Math*, vol. 19, no. 1, pp. 197–208, 2025.

[11] M. A. Deif, H. Attar, A. Alrosan, A. A. A. Solyman, and S. M. F. Abdelaliem, "Design and development of an intelligent neck and head support system based on eye blink recognition for cervical dystonia," *Discover Applied Sciences*, vol. 6, no. 11, p. 602, 2024.

[12] "Mediapipe Pose Landmarker." https://developers.google.com/mediapipe/solutions/vision/pose_landmarker.

[13] "Mediapipe Face Landmarker." https://developers.google.com/mediapipe/solutions/vision/face_landmarker.

[14] "Threejs Creating a scene." https://threejs.org/docs/index.html#manual/en/introduction/Creating-a-scene.