Sensors and Materials, Vol. 38, No. 4 (2026) 2233–2244
MYU Tokyo

S & M 4433

High-precision Software-defined Sensing for Smart Retail and Interactive 3D Displays

Lisa C. F. Ying* and Wei-Hsi Hung

Department of Management Information Systems, National Chengchi University,
Ph.D. Program in Communication, Shih Hsin University
No. 64, ZhiNan Road, Section 2, Wenshan District, Taipei City 11605, Taiwan

(Received January 29, 2026; accepted April 6, 2026)

Keywords: SDS, AI-image sensor, naked-eye 3D, augmented reality, seamless interaction

This study was carried out to validate a software-defined sensing (SDS) architecture that combines AI-vision-enabled image sensors with a deep learning model to replace traditional hardware sensors, such as infrared and ultrasonic sensors, for complex human–computer intent recognition. The architecture was implemented in Picbot, an AI-driven, naked-eye augmented reality interactive platform for smart retail that integrates visual AI with high-performance displays. The Picbot model identifies 1293 2D facial key points, significantly exceeding the 106 points provided by Google MediaPipe, an open-source framework widely used for real-time perception tasks such as facial mesh and skeletal tracking. In operational reliability testing, the architecture achieved a 94.56% success rate in recognizing and triggering composite user intent in controlled indoor environments. Deployment in dynamic outdoor environments achieved a 50.41% social share rate, demonstrating the SDS architecture's practicality for brand engagement and social dissemination. These results highlight SDS architectures as flexible, high-precision AIoT sensing solutions, marking a transition from hardware-constrained sensing to intelligent, algorithm-defined perception. Despite these results, the transition to algorithm-defined sensing has challenges of system stability in high-traffic environments with fluctuating ambient light. Such challenges can be addressed by employing a hybrid AI tracking strategy and real-time edge processing in complex interactive settings.

1. Introduction

Human–computer interaction (HCI) has enabled intuitive and seamless natural user interfaces (NUIs). NUIs require sensing systems that interpret human behavioral intent, posture, and facial expressions without reliance on traditional physical inputs such as buttons or touchscreens.(1) The transformation from HCI to NUI redefines conventional imaging components as smart sensors capable of real-time, high-fidelity perception, establishing a technical foundation for immersive and personalized interactive experiences. Advances in AI have further enhanced image generation and interaction, leading to the development and integration of new HCI and NUI frameworks.(2)

*Corresponding author: e-mail: 108356504@nccu.edu.tw
https://doi.org/10.18494/SAM6266

ISSN 0914-4935
https://myukk.org/

Conventional sensor design requires specialized hardware, such as infrared or time-of-flight sensors, to enable depth perception or behavioral intent. In contrast, software-defined sensing (SDS) leverages computational methods to extract these dimensions from standard visual data through AI-driven spatial mapping and skeletal estimation. This approach reduces hardware complexity while increasing data utility, aligning with contemporary sensor research on soft or virtual sensors that derive high-value information through signal processing rather than direct measurement. The integration of SDS into naked-eye 3D display systems advances sensor technology. Embedding AI-enabled perception within the display interaction enables high-reliability sensing, which is essential for commercial deployment. Performance reliability is defined by hardware accuracy and the robustness of the algorithm–hardware integration.

An example of such integration is Picbot, a naked-eye augmented reality (AR) interactive platform developed by Speed 3D Inc. in Taiwan. Picbot combines AI-driven visual sensing with advanced display technologies to optimize interaction processes and deliver immersive experiences in commercial environments.(3) Powered by Qualcomm Edge AI, Picbot functions as an innovative retail system and interactive vending machine, employing 3D avatars and visual AI to support flexible applications. Despite these innovations, related research has mainly focused on algorithmic optimization or on optical display materials,(4) and studies on integrating AI image sensors with naked-eye display components and their impact on user experience and operational efficiency remain scarce.

This study was conducted to explore how to implement SDS in hardware. Traditional software-defined sensor networks in wireless sensor networks emphasize resource optimization and energy efficiency.(5) In contrast, SDS, as a foundation for NUIs, focuses on perceptual content and intent interpretation. AI-based image recognition algorithms replace or augment hardware sensors, enabling the real-time tracking of complex human intentions such as skeletal movement and facial recognition, and triggering automated virtual augmentations.(6) Such reconceptualization enhances the role of AI at the perceptual level and introduces a novel analytical framework for sensor research. In addition, methodologies for analyzing system stability under variable conditions, such as fluctuating ambient light and dense user traffic, were identified as essential for transitioning AI-vision sensors.

2. Literature Review

2.1 AI-enabled sensing and naked-eye 3D displays in HCI

HCI has evolved significantly to provide immersive, multisensory experiences that bridge the gap between digital content and physical reality. Current technological paradigms enable seamless user interaction through NUIs, which necessitate sensing systems capable of interpreting and responding to complex human behaviors in real time. On the basis of the theoretical foundations of HCI and NUIs, it is emphasized that technological implementations must align with users' natural operational habits to minimize cognitive load and maximize intuitive use.(7) In the context of naked-eye 3D technology, input sensing requires the sophisticated coordination of multimodal technologies, including visual tracking, gesture recognition, and eye-tracking, to ensure spatial accuracy and responsiveness.[8] Furthermore, iterative design processes involving sketching and prototyping are essential for refining system usability and enhancing user satisfaction.[9]

AI has considerably transformed image generation and interaction, rendering them more dynamic and context-aware. Multimodal image generation methods now facilitate real-time naked-eye visual manipulation via lenticular lens arrays, meeting the stringent technical requirements of high-performance Mini LED display systems.[10] At the hardware level, the integration of Mini LED technology with transparent conductive film materials provides a stable, high-resolution image for autostereoscopic effects. Previous research on the photoelectric properties of transparent conductive zinc oxide films and the sensing stability of aluminum-coated flexible substrates has confirmed the feasibility of integrating these display components with advanced sensors.[11,12] The application of value co-creation theory within user-generated content (UGC) enhances brand interaction and user participation, thereby validating the commercial potential of integrated sensing-display systems.[13]

AI-based naked-eye 3D technology has been rapidly adopted across the retail and outdoor media industries, where personalized content generation significantly enhances advertising effectiveness and user engagement.[14] The technology enables remote 3D consultations, preoperative simulations, and clinical training, and improves doctor–patient communication and diagnostic precision by providing high-fidelity spatial visualizations. Such platforms represent a convergence of sensing technology, AI algorithms, and multimedia presentation, echoing recent breakthroughs in optoelectronic sensing and signal processing.[15]

In these advancements, AI image sensors have played a vital role, transforming a passive camera lens into an active, software-defined sensor by integrating computer vision and deep learning. By complex skeletal tracking, facial feature analysis, and gesture recognition without wearable devices, this technology enables somatosensory interaction. In addition, the innovation of naked-eye 3D displays, combined with AR and virtual reality, is driving a mainstream trend toward integrated, multisensory immersion. These cross-domain applications, ranging from smart retail scheduling to immersive medical diagnostics, demonstrate that integrating software-defined sensing and advanced display technology contributes to the development of the next-generation intelligent HCI.

2.2 HCI

HCI is the design of interactions between people and computer systems that enhance the user experience. The seamlessness of NUIs requires intuitive interaction so that users do not need additional learning or physical effort. In conventional user interfaces, users face the Gulf of Execution and the Gulf of Evaluation. The design philosophy of SDS is to solve the problems caused by the Gulf of Execution and Evaluation. With immediate, accurate feedback, the AI image sensor proactively identifies user intent and automatically initiates the interactive process. This software-led automation eliminates the friction that traditional interfaces may produce in input, navigation, or initiation links. It is a key technical means to achieve high usability and seamless interaction. Presence is a key indicator for measuring the quality of an immersive experience. Naked-eye display technology, particularly the integration of a Mini LED array with lenticular lens optics, produces multiview images that convey a strong sense of visual depth and enhance the perception of presence.

2.3 Value co-creation theory

Value co-creation theory emphasizes experience-centric value creation, in which customers and enterprises co-create value through the experience of products or services. In this study, UGC is regarded as the specific mechanism for co-creating value. Users' motivation is shared to provide a visually impactful (high presence) naked-eye interactive experience. The high sharing rate is related to social capitalization. When the quality of the experience is high, users experience show-off motivation, converting their personal enjoyment of the experience into social capital and disseminating it on social networks. This mechanism transforms the interactive results into quantifiable social dissemination value, increasing brand influence. UGC is closely correlated with intellectual property (IP). IP acts as a multiplier, increasing sales and users' sharing rates. Therefore, the mechanism is considered a high-performance IP-to-social-value conversion mechanism.

In this study, we analyzed how AI-enabled sensing technology empowers users to take on the role of content co-creator. By leveraging personalized AR photos or videos created by users, the marketing communication model is established from a one-way brand display to a two-way value dialogue and co-creation between the brand and users. Such a UGC mechanism transforms users' experience into social values, creating additional, measurable commercial value for companies.

3. Methods

3.1 Picbot

Picbot, developed by Speed 3D Inc. in Taiwan, is an AI-driven, naked-eye, augmented-reality interactive system (Appendices I and II). It integrates advanced sensing, display, and edge computing technologies (Fig. 1). The system functions as both an AR photo booth and a smart retail platform, making it a practical case for analyzing SDS in commercial deployments.

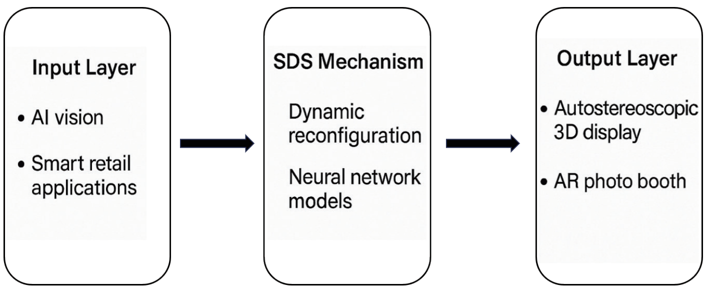

Fig. 1. Process of Picbot (created in this study).

The Picbot system begins with the input layer, which includes an AI vision module and CMOS imaging sensors. These components capture user posture, facial expressions, and gestures in real time. User interaction is initiated in various commercial environments, such as retail spaces and exhibition venues, where the system collects raw visual data for further processing. The second layer is the SDS mechanism. In this stage, the system virtualizes multiple sensing functions by dynamically reconfiguring the imaging hardware's roles using deep learning models. These models enable the system to perform skeletal tracking, facial landmark detection, and proximity analysis. Edge AI processing ensures real-time inference, allowing the system to respond adaptively to user behavior without latency. The final layer is the output layer, where the processed data are transformed into immersive visual experiences. The naked-eye 3D display generates autostereoscopic AR content without requiring wearable devices. Users receive personalized AR avatars and interactive content that can be shared instantly using QR codes or directly on social media platforms. These outputs support both innovative retail applications and interactive photo booth experiences, enhancing user engagement and commercial effectiveness.

Picbot represents an advancement in sensor technology by shifting the interaction trigger mechanism from traditional hardware-based sensors, such as infrared or pressure sensors, to sophisticated software algorithms that interpret human intent. The device integrates a 4K wide-angle lens with a high-performance edge-computing unit, enabling the system to operate as an active smart sensor. Through this software-led approach, the system autonomously identifies composite intent and recognizes when a user enters an interactive zone, pauses, and performs a specific gesture to trigger content automatically. To balance computational precision with development efficiency, the system utilizes a hybrid AI tracking strategy. For body skeleton tracking, the system integrates standard tools to ensure sensitive full-body interaction. However, for high-precision facial localization, Picbot employs a proprietary deep learning model that significantly outperforms existing solutions on the market. Picbots were deployed across three locations to evaluate different performance metrics under varying environmental conditions: Kaohsiung Pili Ceremony (indoor), the Yilan Traditional Arts Park (outdoor), and the Leofoo Village Theme Park (outdoor). Table 1 describes the Picbot SDS, which interprets a complex

Table 1
Logic table for Picbot AR photo booth.

Phase	Triggering sequence (sensory input)	Decision mechanism	Output result (virtual augmentation)
Start interaction	1. Facial detection 2. Presence in interaction zone > 2 s (proximity/dwell time)	Software state machine	System enters standby animation; displays prompts (e.g., "Please prepare")
Trigger capture	1. Specific gesture detected (e.g., peace sign) 2. Face alignment to camera 3. Skeletal stability > 1 s	Dedicated software logic module	Countdown initiated; capture of photo/video
Exit/end session	1. Exit detection (leaving zone > 3 s) 2. "Exit" gesture detected	Software state machine	Ends current session and displays QR code for content download

interaction intent from a sequence of independent sensory events using software logic, and Table 2 provides essential hardware configuration details for Picbot.

To ensure the privacy of the data collected by Picbot, the Picbot architecture uses a local-inference-only model. The Edge AI processing unit performs all high-density facial landmarking and gesture recognition on-site. Raw visual data from the CMOS sensors are processed in volatile memory and immediately discarded after the intent is interpreted. No raw biometric data or video streams are transmitted to the cloud or stored in backend logs. Only the final user-generated AR content is uploaded to a secure cloud server, where it is accessible exclusively via a unique, transient QR code provided to the user. Furthermore, the system complies with data protection standards by requiring explicit user initiation for any social platform sharing or the collection of contact information.

3.2 Data collection

The quantitative data for this study were obtained from Picbots' backend operational logs during public deployments. All interactions between visitors and Picbot were voluntary and occurred naturally in the following public events or venue visits, with no prior instructions given to users. In the Kaohsiung Pili Ceremony (indoor), audiences were general attendees who engaged with Picbot spontaneously as part of the exhibition experience. Because the environment was relatively controlled and the interference was minimal, this deployment was used to evaluate conversion efficiency. The Yilan Traditional Arts Park and the Leofoo Village Theme Park (outdoor) are open cultural and recreational venues with heavy, unpredictable visitor traffic.

Picbot was installed at these locations as part of the public attractions, and visitors initiated interactions freely. The data collected from these outdoor deployments were used to evaluate the efficiency of social dissemination, since users generated and shared AR content on their personal social platforms without researcher prompting.

The amount of data collected from Picbot was substantial. Across different locations, the backend operational logs recorded more than 1000000 interactions. Specifically, during the Kaohsiung Pili Ceremony (indoor), approximately 3500 visitors engaged with Picbot, resulting in 1745 logged interactions and 1650 content downloads, corresponding to a conversion efficiency of 94.56%. At the Yilan Traditional Arts Park (outdoor), 300369 interactions were recorded, including 151,418 shares on social platforms, resulting in a social share rate of 50.41%. At the Leofoo Village Theme Park (outdoor), visitor engagement was very high, with a

Table 2
Hardware specifications of Picbot.

Component	Specification
Sensor/camera	4K wide-angle lens (model or key parameters)
Processing chip	Edge computing chip (specific model or performance level of Qualcomm Edge AI)
Display type	Mini LED naked-eye 3D display
Optical element	Lenticular lens (focal length, matching relationship between pixel pitch and lens pitch)
Display performance	Brightness, contrast ratio, and resolution
System dimensions	Reference Picbot 2.0 ($170 \times 200 \times 220$ cm^3, width × length × height)

satisfaction rate of 97.75%. These large-scale operational logs provided a robust dataset for evaluating both conversion efficiency and social dissemination, without reliance on questionnaire surveys or recruited participants.

In addition to quantitative logs, qualitative assessments were conducted through semistructured interviews with five of Picbots' partners and clients, who were responsible for system operation, technology development, product and service management, marketing, and manufacturing. The interviews were conducted mainly on AI algorithm optimization, challenges in NUI design, and feedback regarding Picbots' effectiveness in commercial applications. Complementary on-site observations were also conducted to record user behavior, including the comprehension of the interaction flow, the frequency of interaction failures, and emotional responses. These qualitative data were used to assess the seamlessness of human–computer interaction. As a result, the foloowing two metrics were used in the quantitative analysis:

1. Conversion rate: the proportion of users who successfully generated AR content and either provided contact information or completed the designated task;
2. Sharing rate: the proportion of users who successfully generated AR content and shared it on personal social platforms (e.g., Instagram and Facebook) using the device or a Quick Response (QR) code.

4. Results and Discussion

4.1 Architecture and SDS

The SDS architecture of Picbot comprises three layers that facilitate sensing, processing, and presentation, as detailed in Table 3. The input layer (integrated sensing) replaces traditional hardware sensors with AI-vision-enabled image sensors to capture high-density behavioral data. These data are fed into the SDS architecture (processing), where deep learning models interpret complex human–computer intent in real time. Finally, the output layer (presentation) utilizes a combination of mini LED arrays and lenticular lenses to deliver an immersive AR experience. This layered approach ensures that the software-led logic can dynamically adjust system triggers without requiring hardware reconfigurations.

The Picbot model identifies 1293 2D facial key points, exceeding the 106 points provided by Google MediaPipe (Fig. 2). It also demonstrates a significantly higher point density than the benchmark. Google MediaPipe is an open-source framework developed by Google for building

Table 3
Three-layer SDS architecture of the Picbot system.

Layer	Function	Components/processes	System objective
Input layer	Integrated sensing	CMOS image sensors, high-density 2D/3D facial landmarks	Capture raw environmental and behavioral data
SDS mechanism	Data processing	Deep learning models, intent recognition algorithms	Translate visual landmarks into actionable digital commands
Output layer	Presentation	Lenticular lens, mini LED array, and AR overlay	Render immersive, depth-enhanced interactive content

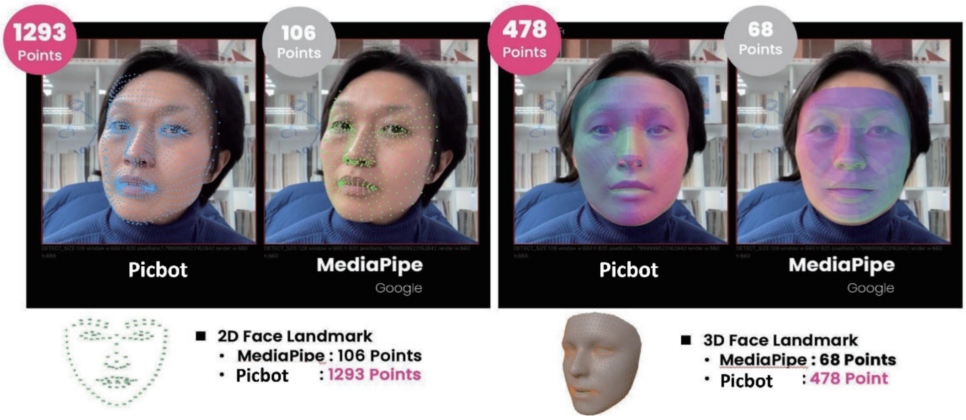

Fig. 2. (Color online) Picbots 2D facial key points compared with those in Google MediaPipe.

cross-platform, multimodal machine learning pipelines. It is widely used for real-time perception tasks such as facial landmark detection, facial mesh reconstruction, and skeletal tracking. MediaPipe provides lightweight, efficient models that run on mobile and embedded devices, making it a standard tool for vision-based applications.

In 3D facial mesh applications, the Picbot model provides 478 points, compared with 68 in standard tools such as Google MediaPipe. Its high-density landmark detection is critical to sensor development, as it allows for hyper-realistic virtual-real fusion. Such precision enables the system to align user movements with virtual characters by restoring exact lighting and camera parameters, thereby significantly increasing immersion. The operational flow of the SDS architecture is detailed in Table 3, which categorizes the system into three distinct functional phases.

4.2 HCI and user experience validation

The results of on-site observations and the interviews showed that the Picbot system satisfied the requirements for NUIs, particularly in terms of seamlessness and presence. High-reliability AI algorithms reduce the user's cognitive load by minimizing the Gulf of Execution, while immediate visual feedback from the naked-eye display narrows the Gulf of Evaluation. Users characterized the experience as intuitive, requiring no prior learning to operate.

The integration of Mini LED technology with inverse perspective algorithms led to high levels of presence. This integration ensures high clarity and accurate depth-of-field effects, enabling users to perceive a seamless fusion between virtual objects and real-world space. As noted by multiple interviewees, this transition from passive seeing to active participating represents the effective integration of smart displays and human–computer interaction. Another interviewee emphasized that future displays are not just about seeing, but about participating and interacting. Picbot subverts traditional display applications while also opening a new chapter in the fusion of smart displays and human–computer interaction.

From an industry perspective, interviewees highlighted Picbot's potential to transform commercial applications. As one interviewee stated, "This innovative solution will help sectors such as amusement parks and retail more easily apply high-end display hardware and naked-eye visual technology, bringing far-reaching impact to the industry, while providing new sensory experiences for consumers. We hope this cross-domain innovation and collaboration result can expand across Taiwan, and even internationally, to jointly create new business opportunities and drive industrial development."

The qualitative results show that Picbot meets technical requirements for seamless NUIs and demonstrates significant commercial and industrial value, reinforcing its role as a next-generation SDS architecture. To validate the monetization and dissemination value of the system's mechanism, more than 1000000 interaction records were analyzed. These data were automatically captured from the Picbot backend management system, which logged real-time telemetry from the SDS state machine across various commercial settings. Each record encompasses a complete interaction lifecycle, from the input layer detecting a user's presence, through the SDS mechanism processing gesture triggers, to the output layer generating a unique QR code for content retrieval. Using the automated logs over the multimonth deployment period at the Kaohsiung, Yilan, and Leofoo Village sites, we evaluated the performance of the SDS system, as shown in Table 4.

5. Conclusion and Recommendation

We developed and validated an SDS architecture integrating AI and vision sensor technologies within the Picbot interactive system. SDS fundamentally advances human–computer interaction by shifting from passive hardware triggers to the seamless, closed-loop, algorithm-led recognition of composite user intent, achieved through dedicated software logic or a state machine.

The Picbot architecture demonstrated superior perceptual performance, utilizing a proprietary deep learning model to capture 1293 2D facial key points (significantly outperforming industry benchmarks such as Google MediaPipes' 106 points) to facilitate hyper-realistic virtual-real fusion and enhance immersion in the Mini LED naked-eye 3D display. Empirical validation across diverse deployments confirmed the system's reliability and adaptability. It achieved a conversion efficiency of 94.56% in controlled indoor environments and a social share rate of 50.41% with a high user satisfaction score of 97.75% in dynamic outdoor settings.

Table 4
Interaction effectiveness was observed in different locations.

Location	Metric	Data collected
Kaohsiung Pili Ceremony (indoor)	Interaction rate	50% (1745 interactions/3500 foot traffic)
	Content download rate	94.56% (1650 downloads/1745 interactions)
Yilan Traditional Arts Park (outdoor)	Interaction count	300369
	Social share rate	50.41% (151418 shares/300369 interactions)
Leofoo Village Theme Park (outdoor)	Satisfaction	97.75%

Through the conceptualization and validation of SDS, a flexible, high-precision AIoT sensing solution is achieved, overcoming the challenges posed by traditional hardware. However, it is necessary to detail the optical foundation of the display, such as the Lenticular Lens and Reverse Perspective Computation principles for new analytical frameworks, and to address the technical challenges of NUIs in complex, multiuser environments. With the conventional camera lens as an intelligent sensor for real-time semantic content interpretation, the SDS architecture can be used for next-generation NUIs, intelligent human–computer interaction systems, and AI-generated content.

The conceptualization and validation of SDS enable a flexible, high-precision AIoT sensing solution that overcomes the challenges of traditional hardware. However, maintaining perceptual accuracy in unconstrained, multiuser environments and refining the optical-computational integration of lenticular displays remain challenges for further research. The results of this study serve as a basis for related research. By replacing passive hardware triggers with a closed-loop system, a software-led state machine can interpret composite intent. A deep learning model can also be integrated into the SDS architecture to enable high-density landmark detection, minimizing the impact of environmental interference in next-generation intelligent human–computer interaction.

References

1 B. Buxton: Sketching User Experiences: Getting the Design Right and the Right Design (Morgan Kaufmann, San Francisco, 2007) pp. 39–55. https://doi.org/10.1016/B978-012374037-3/50047-X
2 K. Hinckley: The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications (CRC Press, Boca Raton, 2012) p. 151–168.
3 H. Lynam, S. Dascalu, and E. Folmer: Int. J. Jum. Comput. Interact. **41** (2025) 10190. https://doi.org/10.1080/10447318.2024.2431757
4 H. Urey, K. V. Chellappan, E. Erden, and P. Surman: Proc. IEEE **99** (2011) 540. https://doi.org/10.1109/JPROC.2010.2098351
5 J. A. Puente Fernández, L. J. García Villalba, and T.-H. Kim: Entropy **20** (2018) 225. https://doi.org/10.3390/e20040225
6 J. Chen and X. Ran: Proc. IEEE **107** (2019) 1655. https://doi.org/10.1109/JPROC.2019.2921977
7 Y. Jopu, Q. Xie, N. Zhang, and J. Lv: Sci. Rep. **15** (2025) 13732. https://doi.org/10.1038/s41598-025-98891-3
8 T. Wang, C. Yang, J. Chen, Y. Zhao, and J. Zong: Sci. Rep. **17** (2024) 24381. https://doi.org/10.1038/s41598-024-75172-z
9 R. K. Dewi and A. T. Sitorus: J. Comput. Sci. Appl. **3** (2025) https://doi.org/10.61978/digitus.v3i1.1117.
10 Z. Xu, W. Zhao, J. Zong, and X. Li: Electronic **14** (2025) 744. https://doi.org/10.3390/electronics14040744
11 C.-H. Yu, S. T. Wicaksono, S.-T. Wang, and T.-H. Chen: Sens. Mater. **37** (2025) 1825. https://doi.org/10.18494/SAM5316
12 X. Zhang, J. Chai, Y. Zhan, D. Cui, X. Wang, and L. Gao: Micromachies **16** (2025) 330. https://doi.org/10.3390/mi16030330
13 C. Wnag, X. Xhao, and J. Hong: Behav. Sci. **14** (2024) 1177. https://doi.org/10.3390/bs14121177
14 W.-H. Hsiao, P.-H. Tsai, and C.-J. Chen: J. Enterp. Inf. Manag. 2025. https://doi.org/10.1108/JEIM-05-2024-0246
15 M. Obst, J. Arensmeyer, H. Bonsmann, A. Kolbinger, J. Kigenyi, F. Oneka, B. Owere, J. Schmidt, P. Feodorovici, and J. Wynands: JMIR Form Res. (2025) e693000. https://doi.org/10.2196/69300

Appendix I

Picbot awards and recognitions.

Year	Award	Category	Institution/organizer
2020	XR Golden Awards	B2C Best XR Marketing	Yahoo/TV (Implied Media/ Platform)
2022	Future Commerce Awards	Experience Innovation	N/A (Future Commerce Awards)
2023	Asia XR Golden Awards	B2C Best XR Interaction	Asia XR Golden Awards
2024	Digital Singularity Awards	Sports Marketing Bronze	Digital Singularity Awards
2024	Taoyuan Sports Tech Start-up Accelerator	Golden Awards	N/A (Part of the Accelerator Program)
2025	Qualcomm Innovate in Taiwan Challenge	Finalist Team	Qualcomm
2025	Smart Display Application Awards	Smart Edutainment	SDA (Smart Display Application Awards)
2025	Epson Innovation Competition	High-level recognition	Epson

Appendix II

(Color online) Picbot product family and functional evolution.

Year	Product (achievement)	Specification (width × length × height, cm)	Photo
2019	Picbot 1.0 (World's first AR smart picture robot)	85 × 62 × 200	
2023	3D Picbot (3D + AR interactive hub with motion detection and AI face recognition)	97 × 71 × 200	

2024	Picbot 2.0 (AR, video, and audio virtual interaction with stars and celebrities)	170 × 200 × 220	
2025	Picbot AI Hub (AI vending machine with virtual make-up try-on functions)	170 × 200 × 220