EnvironAI: Extending AI Research into the Whole Environment

Jingyi Duan [1*], Song Tong [1,2*], Hongyi Shi [1], Honghong Bai [1,3] Xuefeng Liang [4†] Kaiping Peng [1,2†]
1. Department of Psychology, Tsinghua University, Beijing, China
2. Environmental Psychology AI Lab, Tsinghua University, Beijing, China
3. Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing, China
4. School of Artificial Intelligence, Xidian University, Shaanxi, China
⋆ equal technical contribution
† xliang@xidian.edu.cn, pengkp@tsinghua.edu.cn

1. Introduction

Embodied AI has made significant progress in simulating human-like behavior and improving performance [22]. However, it's crucial to recognize that AI is not operating in isolation. As AI continues to develop rapidly, it also reshapes its environment, for instance, human experience [16]. It is a critical historical point for not only AI but also humans and the whole ecosystem. Beyond the functional improvement of AI, it is equally vital to recognize how humans, or the whole environment (including other AIs, humans, the ecosystem, the physical world, etc.) may develop with AI. In short, this perspective, referred to as Environment with AI (EnvironAI), the focus shifts towards not only noticing how the environment is changing with AI but also broadening the research scope to the whole interrelated dynamic environment [9] where AI develops. Instead of solely prioritizing the functionality or the abilities of specific AI agents, EnvironAI aims to optimize the interactive system as a whole.

To pursue the harmonious coexistence between AI and humans, it first requires acknowledging the interdependence in between [5]. Former studies in embodied AI have already brought perception and action together to improve AI functionality. To go deeper into the interrelated system, however, more light should also be shed on the dynamic and reciprocal relationship between AI and the environment. Thus, EnvironAI emphasizes the influence of AI on its environment. Environment refers not only to the natural environment but also to cultural and social environments. Specifically, environment refers to any region (of different meaning systems such as language system or the physical world) where effect from and to the agent (the "me") can happen. Current research in embodied AI emphasizes the perception-action loop (e.g., [8, 14]), involving perceiving the world, modeling it, taking actions, and constantly refining the model. Expanding from this idea, EnvironAI highlights the importance of mutual perception and influence between the environment and AI. For example, human interaction in the physical and social world constitutes a great part of the AI's environment [2], while technology such as AI also constitutes a great part of the human environment, thus affecting human behavior and psychological state [18], and also the cultural domains [16].

EnvironAI further broadens the scope of AI research to encompass the entire environment where AI is embedded, incorporating the concept initially proposed by Ron Chrisley [3] regarding the interaction between artificial intelligence, humans, and other environmental elements. While current AI research primarily aims to enhance AI as an independent agency, EnvironAI deliberately highlights the environment as a whole and its interrelationships. Interdisciplinary discussions across computer science, social science, art, and other fields are encouraged to foster multifaceted insights and applicable suggestions. Ultimately, EnvironAI aims at the optimization of this whole interactive system. Aligning with evolving trends across industries, AI is hoped to coordinate and develop together with humans [7]. EnvironAI aims explicitly for a harmonious coexistence between AI and humans.

In the following part, a case study is dedicated to exemplifying a possible application of EnvironAI. By combining human experience in the physical environment and AI's big-data processing capabilities, we reached an understanding of the "big picture" metaphor considering both individuality and universality. This is a new dimension of the way of thinking only because of this interaction in the whole system. This case was published in [19].

2. Illustrating EnvironAI: A Case Study

Connecting elements in the environment. The studying of the "big picture" proverb presents us with the difficulty to bridge individual perceptual experience with a cross-cultural high-level metaphor.

In tackling these problems, photographs and AI processing are chosen. Photographs create a connection between a photographer and the environment, representing the very personal experience (including emotion, and perception) of the photographer [1, 13, 17]. Then, a dataset consisting of 46,906 tourism photographs was rated for their visual breadth using neural networks. In short, this approach first combined humans with the physical world, then introduced AI's macro comprehension ability into the model to reach a more comprehensive understanding of a theme.

Measurement. Grounded in lexical analysis of the "big picture" and the "Broaden and build" theory [6], we propose the hypothesis of understanding the link between visual breadth and affective experience in terms of both physical breadth and contextual breadth. A dataset consisting of 29,216 images from the most famous 20 globally renowned tourist attractions was collected for evaluating the physical and contextual breadth using Environ AI. The destination ratings of tourists regarding attractions are sourced from TripAdvisor. [4] For physical breadth (images that present a vast and wide scene) evaluation, we trained a modified Alexnet model [20] using 46,906 travel photos, which achieved a testing accuracy of 93.67%. For contextual breadth (images that present a diverse range of visual elements) assessment, we trained an Alexnet model to extract feature vectors into 102 scenes using the SUN attribute model [15]. Attribute kurtosis was then calculated based on the distribution of recognized scenes in every spot, referring to visual diversity in a certain spot.

Results. Logistic regression analysis was employed to verify the impact of physical wide and contextual breadth on general experience ratings in specific destinations. The logistic regression model was highly significant, X^2 (186) = 52.735, p $<$ .001, accounting for 32.3% (Nagelkerke X^2) of the total variance in destination ratings. The physical breadth exhibited a significant positive main effect on ratings (OR = 2.56, 95% CI [4.569, 18.650]). Lower attribute kurtosis was associated with a higher likelihood of high destination ratings (OR = 0.526, 95% CI [0.013, 0.888]), indicating the influence of contextual breadth on the overall tourism experience. An interaction effect was observed between view type and attribute kurtosis on ratings (OR = 2.669, 95% CI [1.342, 5.311]), suggesting the intertwined relationship between physical breadth and contextual breadth in promoting positive affective experience. This effect reveals a possible phenomenal base of the "big picture" metaphor.

For its possible application, we may also use the framework of EnvironAI. First, it promotes the understanding of emotion and vision for both humans and AI. Second, this new knowledge may help facilitate education, environment design, and AI programming, thus changing the environment. Third, it highlights what new affordances both AI and human is providing for each other, paving new avenues for future research across various disciplines and domains, thus highlighting the interconnectedness.

This case study serves as a preliminary illustration of the potential of the EnvironAI approach, showing how human perception and "AI perception" can be combined to understand the "big picture" metaphor. In this case, humans employed technology (photography) to document their personalized and emotive experiences in the environment [17], while AI exhibited a universal and high-level understanding derived from large-scale data [12]. Combining these two, we reached a model balancing the intuitive personal feelings and the cross-cultural universality of the metaphor. On the other hand, EnvironAI can help people get information from others around the world, across time and space to filter out attractions that suit them within a limited time.

3. Discussion

Compare to the current research trends in Embodied AI, EnvironAI underscores the reciprocal relationship between technology, including AI, humans, and other possible elements, each shaping the environment of the other and hence their understanding. It is always a dynamic system of co-construction [11]. Integrating explanation and prediction in computational social science complements the perspective of EnvironAI by emphasizing the need to consider both causal effects and predictive outcomes [10]. This integration encourages collaboration with researchers from humanistic sciences and social psychology. The fusion of AI processing and human understanding offers a new perspective of the world, a comprehensive view that considers both individual elements ("trees") and the whole system ("forest"). This new understanding is not exclusive to AI or humans but emerges from the interaction within the environment itself.

While promising, it is crucial to approach the EnvironAI with a balanced view, acknowledging the challenges and ethical considerations that might arise in its application, especially as AI capabilities continue to grow [21]. For instance, the integration of AI with human experience might lead to issues around privacy, data security, and user consent. These considerations should be included in future research and development of EnvironAI.

Future studies should aim at the optimization of the whole dynamic system, and seek better modes of collaboration between different intelligent agents. The new understanding produced by the interaction of AI and humans may be a solid foundation for harmonious coexistence in the future. Expanding our idea, intelligence may be considered in the sense of the whole environment instead of any single agent. This will be a focus of our future research.

References

[1] Albert Bandura. Social cognitive theory of mass communication. *Media psychology*, 3(3):265–299, 2001. 2

[2] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems*, 32, 2019. 1

[3] Ron Chrisley. Embodied artificial intelligence. *Artificial intelligence*, 149(1):131–150, 2003. 1

[4] Alton YK Chua and Snehasish Banerjee. Reliability of reviews on the internet: The case of tripadvisor. In *World Congress on Engineering & Computer Science: International Conference on Internet and Multimedia Technologies*, pages 453–457. York, 2013. 2

[5] David De Cremer and Garry Kasparov. Ai should augment human intelligence, not replace it. *Harvard Business Review*, 18, 2021. 1

[6] Barbara L Fredrickson. The broaden–and–build theory of positive emotions. *Philosophical Transactions of The Royal Society of London, Series B: Biological Sciences*, 359(1449):1367–1377, 2004. 2

[7] Kotaro Funakoshi, Hideaki Shimazaki, Takatsune Kumada, and Hiroshi Tsujino. Personal partner agents for cooperative intelligence. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 570–571. IEEE, 2019. 1

[8] David Hall, Feras Dayoub, John Skinner, Haoyang Zhang, Dimity Miller, Peter Corke, Gustavo Carneiro, Anelia Angelova, and Niko Sünderhauf. Probabilistic object detection: Definition and evaluation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1031–1040, 2020. 1

[9] Martin Heidegger and Friedrich-Wilhelm von Herrmann. *Sein und zeit*, volume 2. M. Niemeyer Tübingen, 1977. 1

[10] Jake M Hofman, Duncan J Watts, Susan Athey, Filiz Garip, Thomas L Griffiths, Jon Kleinberg, Helen Margetts, Sendhil Mullainathan, Matthew J Salganik, Simine Vazire, et al. Integrating explanation and prediction in computational social science. *Nature*, 595(7866):181–188, 2021. 2

[11] Brady D Lund and Ting Wang. Chatting about chatgpt: how may ai and gpt impact academia and libraries? *Library Hi Tech News*, 2023. 2

[12] Daniel E O'Leary. Artificial intelligence and big data. *IEEE intelligent systems*, 28(2):96–99, 2013. 2

[13] Steve Pan, Jinsoo Lee, and Henry Tsai. Travel photos: Motivations, image dimensions, and affective qualities of places. *Tourism Management*, 40:59–69, 2014. 2

[14] Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. *arXiv preprint arXiv:2304.03442*, 2023. 1

[15] Genevieve Patterson, Chen Xu, Hang Su, and James Hays. The sun attribute database: Beyond categories for deeper scene understanding. *International Journal of Computer Vision*, 108(1):59–81, 2014. Publisher: Springer. 2

[16] Tess Posner, Li Fei-Fei, et al. Ai will change the world, so it's time to change ai. *Nature*, 588(7837):S118–S118, 2020. 1

[17] Peter Robinson. Emediating the tourist gaze: memory, emotion and choreography of the digital photograph. *Information Technology & Tourism*, 14:177–196, 2014. 2

[18] Marte H Schia, Bo C Trollsas, Hakon Fyhn, and Ola Lædre. The introduction of ai in the construction industry and its impact on human behavior. 27th Annual Conference of the International Group for Lean Construction . . . , 2019. 1

[19] Song Tong, Jingyi Duan, Xuefeng Liang, Takatsune Kumada, Kaiping Peng, and Ryoichi Nakashima. Inferring affective experience from the big picture metaphor: A two-dimensional visual breadth model. *IEEE Conf. Comput. Vis. Pattern Recog. Worksh.*, 2023. 1

[20] Song Tong, Xuefeng Liang, Takatsune Kumada, Peng Zhang, and Kaiping Peng. Detecting the attention scopes from travel photos. In *Proceedings of the 2022 International Conference Information Technology and Biomedical Engineering*, pages 213–217. IEEE, 2022. 2

[21] Eva AM Van Dis, Johan Bollen, Willem Zuidema, Robert van Rooij, and Claudi L Bockting. Chatgpt: five priorities for research. *Nature*, 614(7947):224–226, 2023. 2

[22] Anthony Zador, Sean Escola, Blake Richards, Bence Ölveczky, Yoshua Bengio, Kwabena Boahen, Matthew Botvinick, Dmitri Chklovskii, Anne Churchland, Claudia Clopath, et al. Catalyzing next-generation artificial intelligence through neuroai. *Nature Communications*, 14(1):1597, 2023. 1