AI & Future Society (AFS): Technology, Ethics, and Governance in an Interdisciplinary Perspective
DOI: https://doi.org/10.63802/afs.V2.I1.215

PANORAMA SCHOLARLY GROUP PSG

AI & Future Society (AFS)



When Robots Become More Human: Emotional Disruption and the Boundaries of Social Acceptance

Yang Hu[a,] *

[a]*Institute of Malaysia and International Studies (IKMAS), Universiti Kebangsaan Malaysia (The National University of Malaysia), Malaysia*

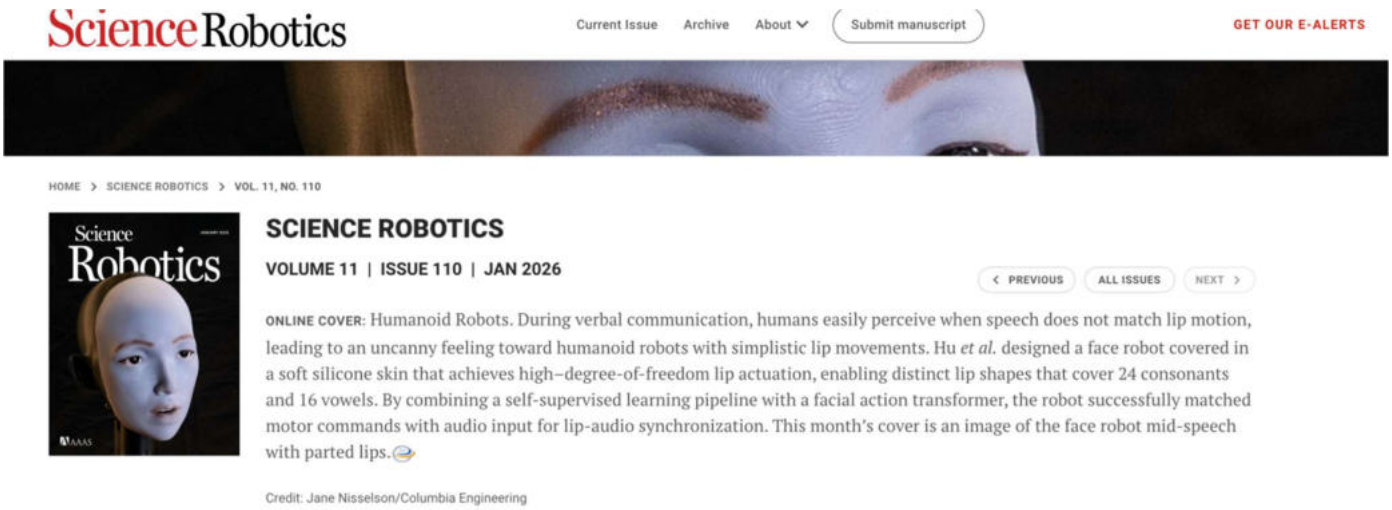
ARTICLE INFO

Keywords:
Humanoid Robots
Lip Motion Generation
Human–Robot Interaction
Uncanny Valley
Ethical and Social Implications

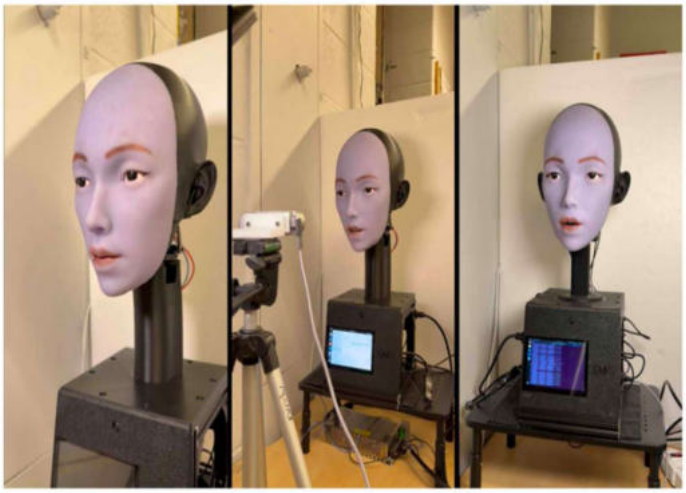
ABSTRACT

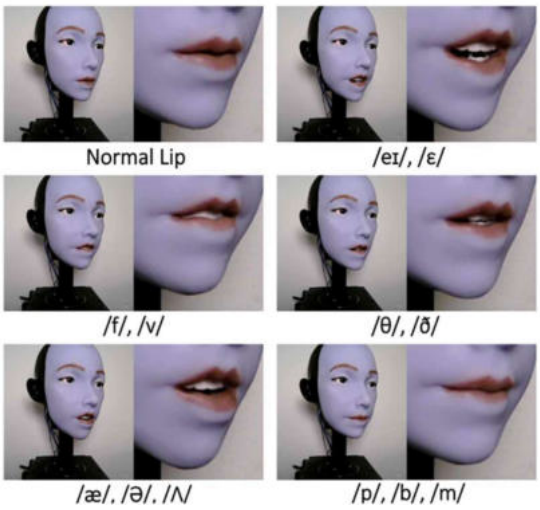
This commentary uses the humanoid facial robot lip-motion generation study by Hu et al., published in Science Robotics, as a starting point to explore the broader cognitive, social, and ethical implications behind its technical achievements. By integrating a soft silicone facial structure, a ten-degree-of-freedom lip actuation system, and a self-supervised learning framework that combines a variational autoencoder with a facial action transformer, the study enables robots to learn natural lip trajectories directly from speech audio. This approach achieves notable improvements in visual naturalness, multilingual generalization, and the mitigation of the "uncanny valley" effect. Yet this seemingly subtle enhancement in lip–speech synchrony effectively pushes human–robot interaction across a critical threshold: robots begin to be perceived not merely as sound-producing tools but as socially present quasi-others. As artificial expressions grow increasingly lifelike, the boundaries by which humans judge emotional authenticity, trust, and intimacy are quietly being rewritten, bringing with them emerging risks of emotional manipulation, over-trust, and even the substitution of interpersonal relationships. At the same time, the study's limited training data and cultural contexts remind us that technical realism does not automatically translate into universal applicability, and that its implications extend well beyond algorithmic performance. The article therefore argues that, within the widening temporal gap between rapid technological advancement and comparatively slower social norm formation, designers and policymakers share a dual responsibility: to establish transparent and comprehensible interaction boundaries, strengthen informed-consent mechanisms, and enhance public literacy in interpreting artificial social signals. In this sense, the work stands not only as an engineering milestone but also as a societal mirror—it demonstrates how machines are learning to "speak like us" while compelling us to reconsider how we define authenticity, agency, and trust in an increasingly synthetic social world.

* Corresponding author. e-mail: p144622@siswa.ukm.edu.my



In the January 2026 issue of Science Robotics, Hu et al. present a humanoid facial robot capable of generating highly realistic lip movements synchronized with speech, emphasizing that "lip motion holds outsized importance in human communication, capturing nearly half of our visual attention during conversation." To address the long-standing mismatch between robotic speech audio and visual articulation, the authors integrate a soft silicone facial structure with a 10-degree-of-freedom lip actuation mechanism and a self-supervised learning pipeline that combines a variational autoencoder with a facial action transformer. This approach enables the robot to infer natural lip trajectories directly from audio without relying on manually predefined rules. Experimental results demonstrate improved visual coherence, multilingual generalization, and reduced uncanny perception compared to traditional heuristic methods. Beyond its technical innovation, this research marks a critical shift in human–robot interaction—from machines that merely produce sound to systems that replicate subtle social and perceptual cues embedded in human communication. However, the training data and evaluation contexts remain limited, raising questions about cross-cultural generalizability. Writing a commentary on this work is therefore meaningful not only for assessing its engineering contributions but also for exploring its broader cognitive, social, and ethical implications, particularly as increasingly lifelike robotic expressions begin to reshape how humans perceive authenticity, agency, and emotional connection in artificial agents.

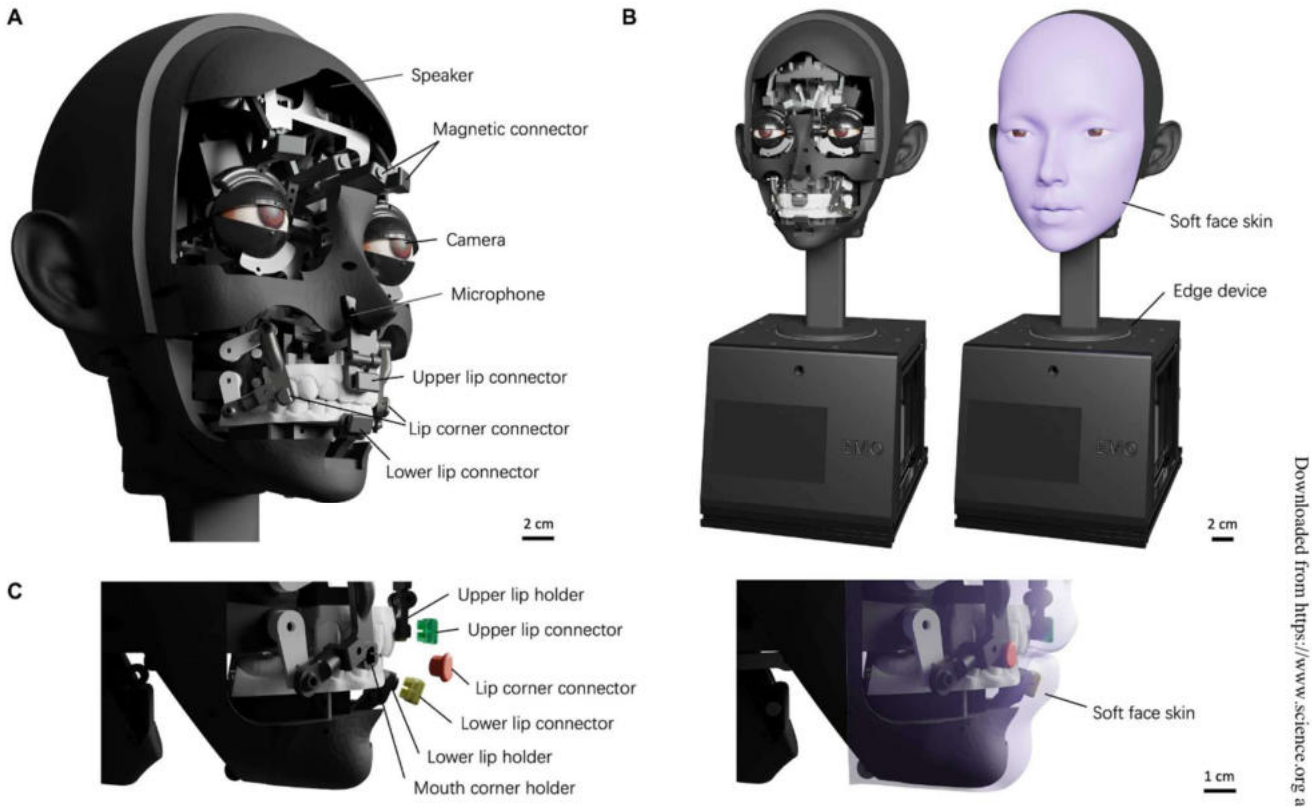




Movie 1. Overview of learning realistic lip motions for humanoid face robots.
The song sung by the robot in movie S5 was generated using the Suno platform.

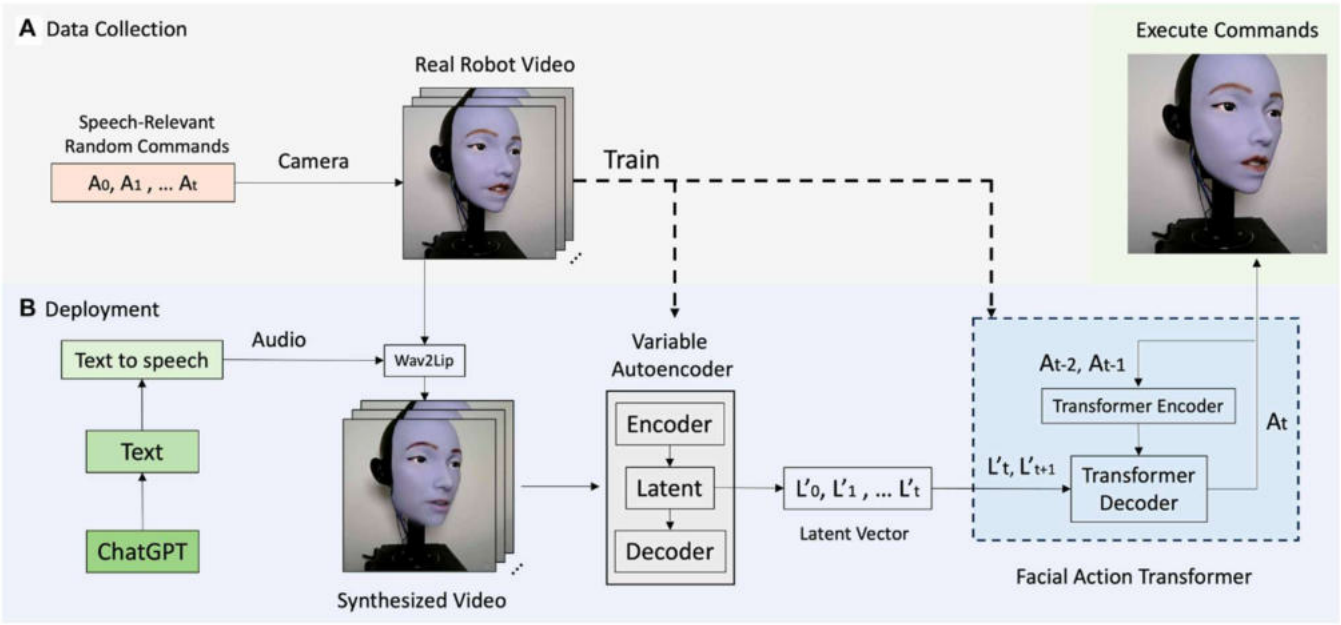
As this research has circulated globally, its influence has rapidly extended beyond the fields of robotics and artificial intelligence, prompting interdisciplinary discussions in multiple countries, including China. The attention it has garnered stems not only from improvements in engineering performance but also from its direct engagement with the long-standing perceptual threshold known as the "uncanny valley" in human–robot interaction. When speech and lip movements achieve a high degree of synchrony, robots cease to be perceived as mere sound-producing devices and begin to be interpreted as socially present, interactive agents. However, increased realism simultaneously blurs the emotional and cognitive boundaries between humans and artificial entities: outward displays of affect may invite genuine emotional attribution, thereby reshaping patterns of trust and attachment in human–machine relationships. The resulting ethical tensions are difficult to ignore—risks of emotional manipulation, over-trust

among vulnerable populations, and the substitution of digital identities for real interpersonal connections may all intensify concurrently. This work, therefore, represents more than a mechanical or algorithmic advance; it signals an emerging societal transition in which humans move from interacting with tools to coexisting with quasi-social agents whose norms and boundaries remain to be newly defined.



Downloaded from https://www.science.org a

Looking ahead, the central question is no longer whether machines can reproduce human expressive signals, but how societies will regulate, interpret, and coexist with these capabilities. Advances in realism often outpace the evolution of social norms, creating a temporal gap in which public expectations, institutional regulations, and ethical safeguards remain underdeveloped. Designers and policymakers therefore bear a dual responsibility: to advance interactive intelligence while simultaneously establishing clear and transparent boundaries that preserve human autonomy and informed consent. Equally important is enhancing public literacy regarding artificial social signals, so that emotional responsiveness is not conflated with emotional authenticity. In this light, the future of human–AI interaction may depend less on the sophistication of algorithms than on society's collective capacity to determine which forms of machine expressiveness should be encouraged, constrained, or deliberately limited. This study thus serves as both a milestone and a mirror—marking technical progress while reflecting the societal questions it inevitably raises. From this perspective, the challenge ahead is not merely technological but fundamentally philosophical—an ongoing negotiation over the meanings of presence, agency, and trust within an increasingly synthetic social environment.



Author Contribution

Yang Hu is the sole author of this commentary. She independently conceptualized the theme, conducted the critical analysis of the source study, and composed the full manuscript. All views and interpretations reflect her own academic perspective.

Compliance with Ethical Standards

Not Required

Funding

None

Competing Interest declaration

The author declare that, to their knowledge, they have no financial, personal, or professional relationships that could be construed as a potential competing interest regarding the work described in this manuscript.

Acknowledgments

None

References

1. Hu, Yuhang, Jiong Lin, Judah Allen Goldfeder, Philippe M. Wyder, Yifeng Cao, Steven Tian, Yunzhe Wang, Jingran Wang, Mengmeng Wang, Jie Zeng, Cameron Mehlman, Yingke Wang, Delin Zeng, Boyuan Chen, and Hod Lipson. 2026. "Learning Realistic Lip Motions for Humanoid Face Robots." S c i e n c e R o b o t i c s.