Author Correction: Subtle adversarial image manipulations influence both human and machine perception

Correction to: *Nature Communications* https://doi.org/10.1038/s41467-023-40499-0, published online 15 August 2023

https://doi.org/10.1038/s41467-024-44826-x

Published online: 16 January 2024

Check for updates

Vijay Veerabadran, Josh Goldman, Shreya Shankar, Brian Cheung, Nicolas Papernot, Alexey Kurakin, Ian Goodfellow, Jonathon Shlens, Jascha Sohl-Dickstein, Michael C. Mozer & Gamaleldin F. Elsayed

The original version of this Article included incorrect acknowledgements to the sources of Fig. 1c. The following sentence in the original caption to Fig. 1 "Illustration images in this Figure were obtained with permission from[33,35]." has been changed to "Illustration images in panel (c) were obtained with permission from Papernot et al.[70], Nguyen et al.[71], and Athalye et al.[72], left to right, respectively."

The following missing references have also been added to the reference list:

[70] Papernot, N. et al. The limitations of deep learning in adversarial settings. In *2016 IEEE European symposium on security and privacy (EuroS&P)*, 372–387 (IEEE, 2016).

[71] Nguyen, A., Yosinski, J. & Clune, J. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 427–436 (2015).

[72] Athalye, A., Engstrom, L., Ilyas, A. & Kwok, K. Synthesizing robust adversarial examples. In *International conference on machine learning*, 284–293 (PMLR, 2018).

This has been corrected in the PDF and HTML versions of the Article.