

# Hassan Ali

PhD Candidate, UNSW, Sydney

 hassan.ali@unsw.edu.au    Google Scholar    Sydney, Australia  
 0000-0002-1701-0390    hassanalikhatim    hassanalikhatim.github.io

## How do I see myself?

I am a self-motivated machine learning engineer and researcher. As a researcher, I strive to enable real-world deployment of machine learning models that people can trust. As an engineer, I want to use my machine learning skills to assist people in their daily routine tasks.

## Education

Sep 2023 – Ongoing	<b>University of New South Wales (UNSW), Sydney, Australia</b> <i>PhD in Computer Science and Engineering</i> <ul style="list-style-type: none"><li>Research focuses on Trustworthy Machine Learning</li></ul>
Sep 2017 – Aug 2019	<b>National University of Sciences and Technology (NUST), Islamabad, Pakistan</b> <i>Master of Science in Electrical Engineering (CGPA: 4.0/4.0)</i> <ul style="list-style-type: none"><li><u>Thesis Title</u>: “Analyzing the Security Vulnerabilities of Deep Neural Networks: Attacks and Defenses”</li></ul>
Sep 2013 – Aug 2017	<b>University of Engineering and Technology (UET), Lahore, Pakistan</b> <i>Bachelor of Science in Electrical Engineering (CGPA: 3.645/4.0)</i>

## Work Experience

Feb 2024 - Present	<b>University of New South Wales (UNSW)</b> <i>Casual Research Assistant</i> <ul style="list-style-type: none"><li>Large Language Models</li></ul>
Sep 2021 - Sep 2023	<b>Information Technology University (ITU)</b> <i>Research Assistant</i> <ul style="list-style-type: none"><li>Human-centric Robust ML-driven IoT Smart Services</li></ul>
Jan 2021 - Nov 2021	<b>Information Technology University (ITU)</b> <i>Research Assistant</i> <ul style="list-style-type: none"><li>Mitigating Anti-social Behavior through Beneficial AI</li></ul>

## Tools and skillset

- Python, PyTorch, TensorFlow (last 5 years)
- Java, C, MATLAB, Verilog, VHDL, HTML

## Publications

---

- |      |  |
|------|--|
| 2023 | <ol style="list-style-type: none"><li>1. <b>Ali, H.</b>, Butt, M. A., Filali, F., Al-Fuqaha, A. &amp; Qadir, J. Consistent Valid Physically-Realizable Adversarial Attack Against Crowd-Flow Prediction Models. <i>IEEE Transactions on Intelligent Transportation Systems</i>, 1–16. doi:10.1109/TITS.2023.3343971 (2023).</li><li>2. <b>Ali, H.</b>, Khan, M. S., AlGhadhban, A., Alazmi, M., Alzamil, A., Al-utaibi, K. &amp; Qadir, J. Con-detect: Detecting adversarially perturbed natural language inputs to deep classifiers through holistic analysis. <i>Computers &amp; Security</i> <b>125</b>, 103367 (2023).</li><li>3. Butt, M. A., Qayyum, A., <b>Ali, H.</b>, Al-Fuqaha, A. &amp; Qadir, J. Towards secure private and trustworthy human-centric embedded machine learning: An emotion-aware facial recognition case study. <i>Computers &amp; Security</i> <b>125</b>, 103058 (2023).</li><li>4. Qayyum, A., Butt, M. A., <b>Ali, H.</b>, Usman, M., Halabi, O., Al-Fuqaha, A., Abbasi, Q. H., Imran, M. A. &amp; Qadir, J. Secure and Trustworthy Artificial Intelligence-Extended Reality (AI-XR) for Metaverses. <i>ACM Comput. Surv.</i> (2023).</li></ol> |
| 2022 | <ol style="list-style-type: none"><li>5. <b>Ali, H.</b>, Khan, M. S., Al-Fuqaha, A. &amp; Qadir, J. Tamp-X: Attacking explainable natural language classifiers through tampered activations. <i>Computers &amp; Security</i> <b>120</b>, 102791 (2022).</li></ol>  |
| 2021 | <ol style="list-style-type: none"><li>6. <b>Ali, H.</b>, Khan, M. S., AlGhadhban, A., Alazmi, M., Alzamil, A., Al-Utaibi, K. &amp; Qadir, J. All your fake detector are belong to us: evaluating adversarial robustness of fake-news detectors under black-box settings. <i>IEEE Access</i> <b>9</b>, 81678–81692 (2021).</li><li>7. Petrick, N., Akbar, S., Cha, K. H., Nofech-Mozes, S., Sahiner, B., Gavrielides, M. A., Kalpathy-Cramer, J., Drukker, K., Martel, A. L. &amp; BreastPathQ Challenge Group, f. t. SPIE-AAPM-NCI BreastPathQ Challenge: an image analysis challenge for quantitative tumor cellularity assessment in breast cancer histology images following neoadjuvant treatment. <i>Journal of Medical Imaging</i> <b>8</b>, 034501–034501 (2021).</li></ol>   |
| 2020 | <ol style="list-style-type: none"><li>8. Khalid, F., <b>Ali, H.</b>, Hanif, M. A., Rehman, S., Ahmed, R. &amp; Shafique, M. <i>FaDec: A Fast Decision-based Attack for Adversarial Machine Learning in 2020 International Joint Conference on Neural Networks (IJCNN)</i> (2020), 1–8.</li></ol>   |
| 2019 | <ol style="list-style-type: none"><li>9. <b>Ali, H.</b>, Khalid, F., Tariq, H. A., Hanif, M. A., Ahmed, R. &amp; Rehman, S. SSCNets: Robustifying DNNs using Secure Selective Convolutional Filters. <i>IEEE Design &amp; Test</i> <b>37</b>, 58–65 (2019).</li><li>10. Khalid, F., <b>Ali, H.</b>, Tariq, H., Hanif, M. A., Rehman, S., Ahmed, R. &amp; Shafique, M. <i>QuSecNets: Quantization-based defense mechanism for securing deep neural network against adversarial attacks in 2019 IEEE 25th International Symposium on On-Line Testing and Robust System Design (IOLTS)</i> (2019), 182–187.</li></ol>   |