Govind Mittal

NYU Center for Cybersecurity, New York, NY 11201
mittal@nyu.edu mittalgovind mittalgovind (917)-292-6704

EDUCATION

New York University (NYU) *2019 - 2024 (Expected: Aug '24)*
Doctor of Philosophy (Ph.D.) Candidate in Computer Science, *A.B.D.* GPA : 3.8 / 4.0
Dissertation: Human-AI Multimodal Approaches to Media Authentication.
Advisors: Prof. Nasir Memon and Prof. Chinmay Hegde
Committee: Prof. Brendan Dolan-Gavitt and Prof. David Fouhey

Birla Institute of Technology and Science (BITS) Pilani, India *2014 - 2019*
Dual Majors: Bachelor of Engineering (Honours) in Computer Science, GPA: 8.0/10.0
Master of Science (Honours) in Mathematics

ACADEMIC PUBLICATIONS

***** - Equal Contribution ; - Created New Dataset;

1. PITCH: AI-assisted Tagging of Deepfake Audio Calls using Challenge-Response.
   G Mittal, A Jakobbson, KO Marshall, C Hegde, N Memon.
   Under Review at 32nd ACM International Conference on Multimedia (**MM** 2024)
2. GOTCHA: Real-Time Video Deepfake Detection via Challenge-Response.
   G Mittal, C Hegde, N Memon.
   Ninth IEEE European Symposium on Security and Privacy (**EuroS&P** 2024).
3. Circumventing Concept-Erasure Methods for Text-to-Image Generative models.
   M Pham, KO Marshall, N Cohen, **G Mittal**, C Hegde.
   Twelfth International Conference on Learning Representations (**ICLR** 2024).
   NeurIPS 2023 Workshop on Diffusion Models.
4. Identity-preserving aging of face images via Latent Diffusion Models.
   S Banerjee*, **G Mittal***, A Joshi, C Hegde, N Memon.
   IEEE Transactions on Biometrics, Behavior, and Identity Science (**TBIOM** 2024)
   IEEE International Joint Conference on Biometrics (**IJCB** 2023). **Best Reviewed Paper.**
5. FiFTy: Large-Scale File Fragment Type Identification Using Convolutional Neural Networks.
   G Mittal, P Korus, N Memon.
   IEEE Transactions on Information Forensics and Security (**TIFS** 2020) Vol. 16.
6. Spoken Language Identification Using ConvNets.
   Sarthak, S Shukla, **G Mittal**.
   European Conference on Ambient Intelligence (**AmI** 2019). **Most Cited Paper across Proceedings.**

WORK EXPERIENCE

Research Intern Jun - Aug 2020
Vidrovr: *Company helps businesses create insights from multimedia using multi-modal AI* *New York, NY*

- Automated extraction of daily events from news transcripts for incorporation into the company's knowledge graph.
- Augmented company's codebase with an end-to-end natural language processing pipeline for extracting structured events (like dependency graphs) using raw news transcripts, *adding a new source modality for deriving insights.*

Graduate Research Assistant Sep 2019 - Aug 2024
NYU Tandon School of Engineering *New York, NY*

- Mentored three junior researchers in machine learning, cybersecurity, and high-performance computing (HPC).
- Presented technical research to a technically diverse research group using meaningful visualizations.

NOTABLE PROJECTS

AI-assisted Deepfake Audio Calls Detection Spring - Fall 2023
Research Work

- Automated human speech data collection using interactive sessions and generated 1.6 million deepfake speech samples.

- Trained 20 machine-based detectors using Wav2Vec to score samples with 85% accuracy on a challenging subset.
- Developed an AI assistant for pre-screening deepfake phone scams and aided under-confident human predictions, improving their detection accuracy from 72% to 83%, using AI warnings.

Challenge-Response for Detecting Real-time Video Deepfakes Spring - Fall 2022
Research Work

- Invented GOTCHA, a CAPTCHA-like verification system that guards against video deepfakes in online live interactions.
- Boosted humans and AI deepfake detection capabilities to an area-under-curve (AUC) of 88% and 82% against four SoTA deepfake generation pipelines across all considered challenges (baseline was 58%).
- Generated a dataset of 56,000 deepfake videos performing various challenges, like face occlusions and distortions.

Continual Learning for Language Models Spring 2021
DS-GA 1012: Natural Language Understanding *Center for Data Science, NYU*

- Quantified catastrophic forgetting in the pretrained BERT language model, which worsened when trained on new tasks.
- Increased resilience of language models by 30%, preventing forgetting of older tasks after the introduction of newer ones.

Trusted Forensic Analysis with Uncertainty Estimation using Bayesian Learning Fall 2020 - Spring 2021
Research Work in Progress

- Created a modular, readable framework for enhancing vanilla neural networks with uncertainty quantification capabilities.
- Accelerated ML pipelines by 6x through data pipeline optimizations and curating computational graphs for JIT compilers.
- Maximized code execution efficiency for High-Performance Computing (HPC) clusters to scale hyper-parameter tuning and ablations to 1000+ models.

Large-Scale File Fragment Type Identification Spring 2019
Center for Cyber Security, NYU Abu Dhabi *Abu Dhabi, UAE*

- Built a 10x faster SoTA filetype inference system with a 10% increase in accuracy using embedding layers and CNNs.
- Curated the largest and most diverse corpus of 75 commonly used file types, ushering in 10 research papers as of 2024 to compete for the state-of-the-art, and influencing 20 others.

TECHNICAL SKILLS

Pythonic Frameworks:	PyTorch, TensorFlow. PyCharm IDE, NumPy, SciPy, Matplotlib, Jupyter, PyTorch Lightning, Pandas, Sklearn, Weights & Biases, MLflow.
ML Paradigms:	Unsupervised Learning, Supervised Learning, Semi-supervised Learning, Probablistic Machine Learning, Interpretable Machine Learning, Ensemble Learning, Transfer Learning.
ML Concepts:	Generative Videos, Hyperparameter Optimization, Model Explainability, NLP, CV, Classification, Feature Engineering, Regression,, Cross-Validation, Bias-Variance Tradeoff, Anomaly Detection, Adversarial Attacks, GAN, Diffusion Models, Transformers.
Cyber Security Concepts:	Threat Modeling, Authentication, Indicators of Compromise, Reverse Engineering, Forensics, Capture-the-Flag, Human Studies.
Softwares:	Linux, Unix, Shell Scripting, Docker, Singularity, HPC, Git.

AWARDS AND FELLOWSHIPS

Funded with $100,000 by **Google's Cyber NYC** Program in 2023 - 2024.

Recipient of **Dean's Ph.D. Fellowship** at NYU Tandon School of Engineering in 2019 - 2020.

Awarded **Young Scientist Fellowship (KVPY)** of Indian Institute of Science (IISc), and funded by the Department of Science and Technology, Government of India, from 2014 to 2019 (Scored **above 99.8 percentile across applicants**.).

VOLUNTEER AND SERVICE

Served as Reviewer for IEEE Transactions on Information Forensics and Security 2021-2024, IEEE Access 2023-2024, ACM International Conference on Multimedia (MM) 2024.

Co-organized Applied Research Competition, Cyber Security Awareness Week (CSAW 2021) at NYU Tandon.