

Yogesh Kulkarni

(213) 551-4049 | Los Angeles, CA, USA | ykulkarn@usc.edu | [Github](#) | [Linkedin](#) | [Google Scholar](#) | [Portfolio](#)

EDUCATION

University of Southern California

Master of Science in Computer Science

Aug 2022-May 2024

Los Angeles, USA

University of Pune

Bachelor of Engineering in Computer Engineering — Honors in Data Science

Aug 2018-Jul 2022

Pune, India

PUBLICATIONS

1. **Y. Kulkarni**, S. Hussain, K. Ramamritham and N. Somu, "EnsembleNTLDetect: An intelligent framework for electricity theft detection in smart grid," in *21st IEEE International Conference on Data Mining Workshops*, Auckland, NZ, December 2021, pp. 527-536. [\[Paper\]](#) [\[Code\]](#)
2. **Y. Kulkarni**, K. Bhambani, "Kryptonite: An adversarial attack using regional focus," in *19th International Conference on Applied Cryptography and Network Security (Workshops)*, Kamakura, Japan, 2021, pp. 463-481. [\[Paper\]](#)
3. **Y. Kulkarni** and A. Gorkar, "Intensive image malware analysis and least significant bit matching steganalysis," in *Proceedings of IEEE International Conference on Big Data (Big Data)*, Atlanta, USA, 2020, pp. 2309-2317. [\[Paper\]](#)
4. A. Gorkar, P. Patil, **Y. Kulkarni**, S. Joshi, "Covi-Care: Secure Covid-19 Vitals Diagnosis and Disease Identification Application," to appear in *Springer Nature - Transactions on Computational Science & Computational Intelligence*. [\[Paper\]](#)

TECHNICAL SKILLS

Languages: Python, Java, C/C++, SQL, JavaScript

Libraries/Frameworks: PyTorch, TensorFlow, Pandas, Scikit-learn, SpaCy, Flask, NumPy, NLTK, Android

Analytical Tools & Databases: Linux, MySQL, MongoDB, Docker, Hadoop, Spark, MLFlow, Kubernetes, GCP

Techniques: Data Mining & Analysis, Machine Learning, Natural Language Processing, Computer Vision, Statistics

RESEARCH EXPERIENCE

Research Assistant, DEVCOM US Army Research Laboratory Los Angeles, USA | Jan 2023-Present

- Leading efforts in 3D Point Cloud Style Transfer leveraging advanced 3D GAN architectures, notably Pix2Pix and CycleGAN.
- Conducting in-depth experiments with cutting-edge 3D Segmentation models, including Point-Transformer, KPConv, and SparseUNet.
- Working on a diffusion approach to enhance existing GAN approaches.

Research Intern, Nokia Bell Labs

New Providence, USA | June 2023-Aug 2023

- Implemented automatic model parallelism and partitioning for GPT-3 and LLaMA foundational models, increased model training throughput by 15% across heterogeneous clusters.
- Designed and executed communication and compute efficient inter-node pipeline parallelism approach for training LLMs on heterogeneous cluster GPUs.

NLP Intern, Episource LLC

Mumbai, India | Jan 2022-Mar 2022

- Implemented and deployed a solution for Abbreviation - Disambiguation of real-time Clinical Texts using PyTorch.
- Prepared an annotated dataset (from AWS S3) and re-engineered an ACL research paper, fine-tuned Bio_ClinicalBERT and PubmedBERT to achieve accuracy of 99 % and 98%.
- Mapped 200+ clinical drugs to their strength, dosage, form with a custom Python script and boosted company's existing NER model by 15%.

Research Intern, RBCDSAI (IIT Madras)

Chennai, India | Jul 2021-Oct 2021

- Proposed an End-to-End framework for detecting Electricity Theft in Industrial Smart Grids.
- Applied Enhanced Dynamic Time Warping for imputation, Stacked Auto-Encoder for dimensionality reduction & Conditional GAN's for robustness attaining an impressive accuracy of 99% & Matthews Correlation Coefficient of 0.98.

Research Intern, IIT Jodhpur

Jodhpur, India | Apr 2021-Jun 2021

- Devised a Spatial-Temporal Learning Framework for Abnormal Event Prediction utilizing Hierarchical Attention Networks. Built bidirectional-recurrent network to capture correlations changing over time in the data.
- Incorporated 15+ spatial & temporal features to generate a rich set of feature representations, achieved an average F1 score of 0.81 (SoTA: 0.86) on NYC dataset.

Junior ML Engineer, Omdena

Remote | Oct 2020-Jan 2021

- Worked with Kenya Red Cross Society to classify Sentinel1 and Sentinel2 satellite imagery into different cropland types for 16 counties of Kenya which were severely affected by desert locust attacks.

- Devised a custom version of U – Net Convolutional Neural Network for image segmentation achieving an impressive 81% accuracy on the LandCoverNet dataset.
- Devised a custom LSTM network for predicting future NDVI (Normalized difference vegetation index) values and to prepare the counties for upcoming desert locust attacks on croplands having an RMSE score of 0.024.

Research Intern, DRDO HQ

New Delhi, India | Jul 2020-Oct 2020

- Investigated Hex dump, EXIF data of images for identifying embedded payloads with sophisticated string-matching algorithms in Python.
- Developed a novel, robust and scalable framework for malware analysis of images.
- Constructed a Stacked Ensemble classifier using XGBoost, Catboost & Feedforward Neural Net for detecting LSB Matching Steganography both for color & grayscale images, with an AUC of 0.98 & 0.87 respectively.

PROJECTS

Constrictive Region Focused Adversarial Attack | *NumPy, OpenCV, Pandas, TensorFlow, PyTorch*

- Proposed a novel adversarial attack using Regional Focus for Medical Deep Learning Systems.
- Executed explicit extraction of Region of Interest (RoI) from images to add imperceptible adversarial perturbations with resultant 100% mis-classification rate by Deep Neural Nets on real-world datasets beating SoTA attacks in terms of max plummet in accuracy, smallest perturbation size and high efficiency.
- The attack holds up well against adversarial defence techniques such as Adversarial Training, Pixel Deflection & Defensive Distillation.

Detecting Gravitational Wave Signals from Binary Black Hole Collisions [\[Code\]](#) | *NumPy, Pandas, PyTorch, SciPy*

- Transformed simulated gravitational wave measurements into 2D images using Continuous Wavelet Transform for efficient cloud-based training.
- Achieved a notable AUC-ROC score of 0.86 by fine-tuning a Vision Transformer to predict the presence of GW signals.

Toxic Pharmaceutical Pills Detection | *NumPy, Pandas, PyTorch*

- Utilized Python's Beautiful Soup library to parse and annotate the pills dataset and fine-tuned Vision Transformer (ViT) on large-scale pills image database to identify name of pill which achieved precision of 0.50 and recall of 0.45.
- Performed Interpretability analysis of ViT using Grad-CAM.

Covi-Care: Covid-19 Detection Application | *Dart, Flask, Java, Librosa, Python, Scikit-learn, TensorFlow*

- Envisioned first-of-its kind mobile application for comprehensive Covid-19 vitals diagnosis using Raspberry Pi 3b. Created a rule-based learning algorithm using fuzzy logic to generate susceptibility score to Covid-19 based on vitals collected by Embedded device.
- Implemented a Cough Sound Classifier based on spectrogram derived feature space to identify potential Covid-19 patients with an accuracy of 94% and deployed it on mobile device.
- Developed a CT-scan based MobileNet model to detect Covid-19 with AuC of 0.78 and a U-net model to predict infection mask in a given CT-scan with dice coeff. of 0.84.
- Built & deployed a machine learning-based Mental Health evaluator to evaluate person's mental health and detect if professional help is needed with AuC = 0.85.

OSIC Pulmonary Fibrosis Progression [\[Code\]](#) | *NumPy, OpenCV, Pandas, TensorFlow*

- Built an ensemble of EfficientNetB5, ElasticNet, Quantile Regression & Lasso Regression to predict severity of decline of lung function based on CT scans of patient lungs.
- Optimized model hyper-parameters with GridSearchCV & implemented a weighed-weight average ensemble of 4 ML models to achieve a Laplace Log-Likelihood of -6.92 for prognosis of this lung disease.

AWARDS / GRANTS

Google Research India Graduate Symposium	Jan 2022
Conference Travel Grant for ICDM conference in Auckland, NZ	Dec 2021
Conference Travel Grant for IEEE Big Data Conference in Atlanta, USA	Dec 2020

PRESENTATIONS / SEMINARS

1. EnsembleNTLDetect: An Intelligent Framework for detecting NTLs in Smart Grids [\[Slides\]](#) [\[Recording\]](#)
2. Kryptonite: An Adversarial Attack using Regional Focus [\[Slides\]](#) [\[Recording\]](#)
3. Pruned Generative Networks for Robustifying Deep Neural Nets [\[Slides\]](#) [\[Recording\]](#)
4. Intensive Image Malware Analysis and Least Significant Bit Matching Steganalysis [\[Slides\]](#) [\[Recording\]](#)