Sandipan Choudhuri

[+1] 480-300-9385 | s.choudhuri@asu.edu | s-choudhuri.github.io | Linkedin | Github | Google Scholar

EDUCATION

- **Doctor of Philosophy** — *Aug 2017-Present • USA*
  Arizona State University • *Computer Science (Machine Learning)* — *CGPA: 4/4*
- **Masters** — *Aug 2013-July 2015 • India*
  Jadavpur University • *Computer Science & Engineering (Computer Vision)* — *CGPA: 8.44/10*
- **Bachelors** — *Aug 2009-July 2013 • India*
  West Bengal University of Technology • *Computer Science & Engineering* — *CGPA: 8.6/10*

PROFESSIONAL EXPERIENCE

- **Rovicare** - *Machine Learning Intern* — *May-Aug 2022 • USA*
  - Automated the patient intake process by designing a cloud-based machine learning framework for extracting proprietary data from scanned records. Enhanced in-house efficiency by streamlining customer ticketing and contract systems.

- **Interbiz** - *Machine Learning Intern* — *May-Aug 2021 • USA*
  - Engineered a document parser for automated processing of scanned health/auto insurance claim forms, improving the turnaround time by a factor of 18. Devised a HIPAA-compliant cloud-based application for secure health-record access.

- **NetXT Lab, Arizona State University** - *Research Assistant* — *Aug 2019-Present • USA*
  - Designed unsupervised adaptation frameworks to alleviate annotation overhead, achieving state-of-the-art accuracies. Developed a time-series pipeline to estimate future opioid incidents and optimal sites for medication-assisted treatment facilities. Leveraged language models to assess social media data's effectiveness in monitoring the opioid crisis.

- **CMATER Lab, Jadavpur University** - *Junior Research Fellow* — *Aug 2015-Mar 2017 • India*
  - Developed a production-level breast cancer diagnostic system, overseeing requirements, data sourcing, and model architecture. Designed frameworks and authored publications on scene labeling, contour detection, and image segmentation.

TECHNICAL SKILLS & RELEVANT COURSEWORK

Language: Python, Java, C++, C, SQL, JavaScript, HTML + CSS
Technology: Docker, Apache Spark, Tableau, Matlab, Git, ClickUp, Scikit-learn, OpenCV, Statsmodels, Spark MLlib, Seaborn, Flask, SpaCy, D3.js
Platform: Google Cloud, Microsoft Azure
Library: Pytorch, Keras, Pandas, NumPy,
Course: Machine Learning, Data Mining, Deep Learning, Statistical Learning, Computer Vision, Natural Language Processing, Data Visualization

RESEARCH PUBLICATIONS

- *Summary* of works leveraging knowledge from a known dataset (*source*) to *mitigate annotation costs on an unlabelled dataset* (*target*), assuming the *source* has a broader category set than the *target* (*Partial Domain Adaptation*)
  - Utilizing source prototypes and negative ensemble learning for target supervision — *(In Review) IEEE ICMLA, 2023*
  - Leveraging objectives beyond first-order moments for category-level distribution alignment — *(Accepted) IEEE ACSSC, 2023*
  - Aligning class distributions with complement entropy objective and adaptative target label refinement — *AIA Journal, 2023*
  - Coupling adversarial learning with selective voting for distribution alignment — *AdvML@KDD, JCCE Journal, 2022*
  - Domain-invariant feature alignment using variational inference — *IEEE ACSSC, 2022*
  - Adaptation using selective representation learning for class-weight computation — *IEEE ACSSC, 2020*

- *Summary* of works identifying object positions (*localization*) and assigning semantic labels to every pixel (*scene labeling*)
  - Pixel-level image segmentation and scene labeling using multi-scale super-pixel contextuality — *Springer RMLDA, 2019*
  - A survey of object localization techniques on natural scenes — *IJPRAI Journal, 2018*

- *Summary* of works detecting boundaries between objects at individual pixel resolution (*contour detection*) in images
  - Leveraging neighborhood attributes and particle swarm optimization for object boundary detection — *Springer FICTA, 2017*
  - Identifying contours by detecting image discontinuities using combined feature-channel information — *IEEE ICACCI, 2016*

- Other works
  - Optimal cost network design for bounded delay data transfer from PMU to control center — *IEEE GLOBECOM, 2021*
  - Structural dependency-aware service chain mapping for network function virtualization — *IEEE DRCN, 2020*
  - User satisfaction-driven bandwidth allocation for image transmission in a crowded environment — *MMTC Journal, 2018*
  - Identification of at-risk groups for opioid addiction through web data analysis — *epiDAMIK@KDD, 2018*

AWARDS & SERVICES

- Doctoral Fellowship Award for "strong academic work and research progress" • *Arizona State University* — *2021-2023*
- Web chair • *INFOCOM workshop on Network Science for Quantum Communication Networks (NETSCIQCOM)* — *2022, 2023*
- Ph.D. Conference Travel Award • *Arizona State University* — *2020, 2022, 2023*
- Research award reviewer for the *Graduate Grants Program* • *Arizona State University* — *2020-2023*
- Reviewer for *IEEE journals* and *conferences* — *2018-2023*
- Co-authored and secured *NSF* and *AHCCCS* grants for combating drug trafficking and Arizona's opioid crisis — *2019, 2021*
- Top 0.5% among 0.25M candidates • *Indian Graduate Aptitude Test in Engineering (GATE) for Computer Science* — *2013*

Other Projects

- **Domain Adaptation in Unconstrained Label-Space**

  Proposed a framework for unsupervised target classification under universal adaptation → Conditional distribution alignment achieved by coupling source prototypes and reciprocal points, with an iterative category-importance estimator for *intra* and *inter* class distance optimization → Bagged highest accuracies across benchmark models on *Office-31, Office-home, VISDA2017*, and *ImageNet-Caltech* datasets.

- **DNN Visualization Using Knowledge Distillation**

  Designed a visualization framework to interpret the functioning of a deep-neural network → Utilized variational autoencoder to generate neighbors around data instances and trained a linear network to probe and replicate the complex network's behavior on the generated samples using knowledge distillation → Examined platform's efficacy using *ResNet* and *VGG* models on *MNIST* and *SVHN* datasets.

- **Time-Series Forecasting on Opioid Data**

  Formulated a pipeline for estimating the geospatial distribution of future opioid occurrences, leveraging time-series analysis on Arizona Emergency Medical Services (EMS) and demographic data → Tasks involved time-series decomposition, stationarity and seasonality testing, estimating series forecastability through running *Granger causality tests*, and developing forecasting models using *ARIMA, SARIMA, RNN* and *uni/bi-directional LSTM.*

- **Colonoscopic Image Segmentation**

  Proposed a method to bypass the necessity for constant monitoring during optical colonoscopy → Polyp regions estimated from informative frames (informativeness determined with supervision) using an iterative entropy-based clustering policy and a *deep-siamese network* trained on contrastive loss → A 6% improvement in *mean Intersection Over Union (mIOU)* witnessed over the benchmark *UNet*, on a real-world *Mayo-Clinic* dataset.

- **Natural Scene Labelling**

  Designed a multi-scale solution to capturing context information for scene labeling → map super-pixel groups at different neighborhood scales to object labels using deep classifiers. A consensus-labeling strategy employed on the output probabilities through multiple voting routines → Outperformed state-of-the-art super-pixel-based methods on *Stanford B.* dataset.

- **Contour Detection on Natural Scenes**

  Developed a game-theoretic, multi-scale contour detection approach, treating texture and color-based feature extraction as players in a two-player game → A mixed strategy Nash Equilibrium is searched across various extraction scales, aiming to minimize spurious edges and highlight salient contours → Secured the second position (average precision of 0.67 with a peak recall of 0.91), on the *BSDS-500* dataset, next only to the state-of-the-art *hierarchical image segmentation model.*

- **Activity Recognition using Myo Gesture Control Armband**

  Acquired gesture data over two continuous days from four students wearing sensor bands, with eating durations logged → Utilized classical machine learning and deep learning methodologies to discern between eating and non-eating behaviors with a 94.76% accuracy, leveraging features from the *Inertial Measurement Unit (IMU)* and *Electromyography (EMG)* data.

Graduate Teaching

Arizona State University

- CSE 110 - Introduction to programming   *2017*
- CSE 205 - Object Oriented Programming and Data Structures   *2018*
- CSE 551 - Foundation of Algorithms   *2017-2020*
- CSE 572 - Data Mining   *2019, 2023*