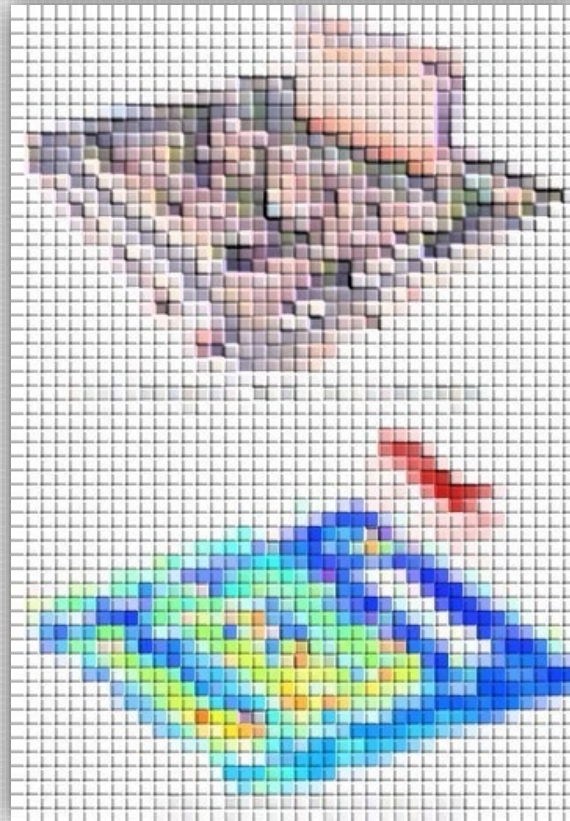


Big Data Research



Big Data Research



Dr. Rama Venkat
Dean, College of Engineering
Phone: (702) 895-1094
Email: Rama.Venkat@unlv.edu

For more than a decade, UNLV researchers have been conducting a world-class effort in various aspects of big data research. This program has been funded by federal and state agencies, as well as many industrial partners. Our researchers have addressed questions related to many fields, including big data, relative to national security and health issues.



Dr. Mohamed Trabia
Associate Dean, College of
Engineering
Phone: (702) 895-0957
Email: Mohamed.Trabia@unlv.edu

We would like to introduce you to some of our researchers. Please feel free to contact us if we can help with future collaboration.

Big Data

Research Areas of Expertise

- Unstructured data analysis
- Cloud computing
- Deep learning
- Document layout analysis
- Human-computer interaction
- Information retrieval
- UAS and GIS integration
- Medical image analysis
- Data mining
- Physiological data sensing applications
- Social network mining using NLP and ML
- Image processing and computer vision
- Machine learning for object detection, semantic segmentation and temporal action localization

Big Data Research

Why UNLV?

- UNLV is a leader among the state's public entities dedicated to advancing big data research in the region and beyond.
- UNLV is located centrally in the west and houses the high-performance computing center, also known as The National Supercomputing Institute.
- Additional data center space is located at the National Supercomputing Institute, which houses the Intel Supercomputer "Cherry Creek" (26,000 cores) cluster. The site provides Federal and State researchers with state-of-the-art, highly-secure cloud computing.



Big Data Research

Why UNLV?

- UNLV's outstanding achievements in big data research, its success in forging public/private partnerships, and its excellent academic programs place the university at the forefront of the field.
- UNLV has been awarded funding from various agencies in the Big Data area.



Faculty Involved in Big Data Research

Dr. Shaikh Arifuzzaman

Assistant Professor, Ubiquitous Data

Dr. Mingon Kang

Assistant Professor, Department of Computer Science

Dr. Jeehee Lee

Assistant Professor, Department of Civil and Environmental Engineering and Construction

Dr. Fatma Nasoz

Associate Professor, Department of Computer Science

Dr. Emma Regentova

Professor, Department of Electrical and Computer Engineering

Dr. Haroon Sahotra

Professor, Department of Civil and Environmental Engineering and Construction

Dr. Bryar Shareef

Assistant Professor, Department of Computer Science

Dr. Kazem Taghva

Professor and Chair, Department of Computer Science

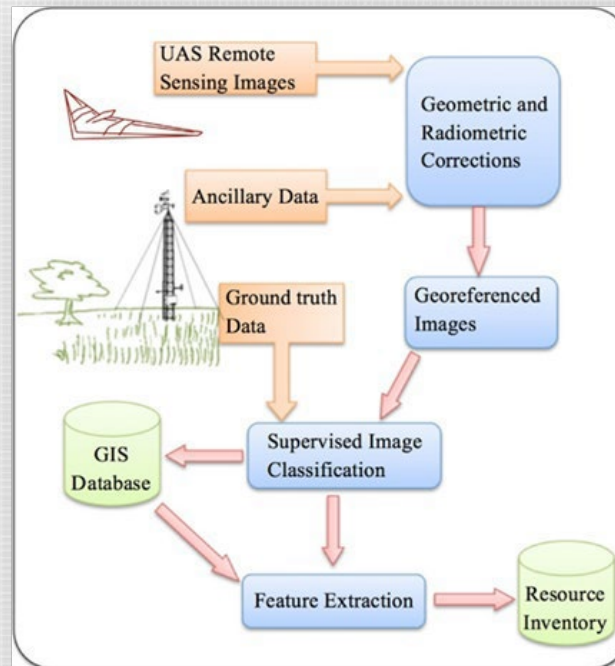
Big Data Research

Additional Resources

www.NevadaCompareCare.Net

Big Data

Research Highlights



Dr. Shaikh Arifuzzaman

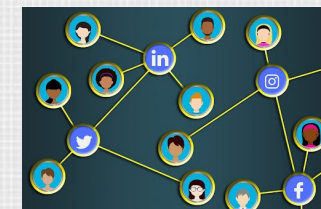
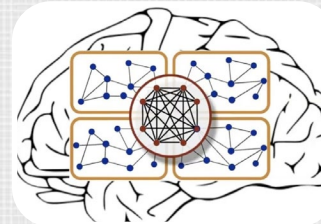
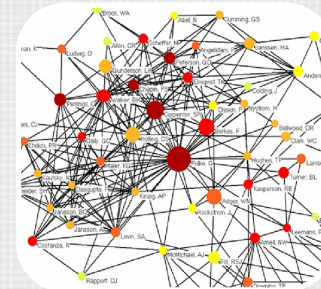
Assistant Professor, Department of Computer Science

Director, Data-intensive Scalable Computing Lab (DiSC Lab)

Phone: (702) 774-3406

Email: shaikh.arifuzzaman@unlv.edu

- Expertise (Big Data + AI/ML + HPC)
 - Parallel algorithms and high-performance computing for data-intensive domains
 - AI/ML methods and applications:
 - Deep learning
 - Graph convolutional network
 - Scaling ML methods
 - Social network mining using NLP and ML
 - Computing on biological/health datasets
 - Building big data applications and systems
 - Graph data analytics, modeling, and visualization



Dr. Shaikh Arifuzzaman

Assistant Professor, Department of Computer Science

Relevant Publications

- Sattar N, Buluc A, Ibrahim K, Arifuzzaman S. "Exploring temporal community evolution: algorithmic approaches and parallel optimization for dynamic community detection." *Applied Sciences*, Springer. 2023; 8(1):64.
- Faysal M, Bremer M, Arifuzzaman S, Popovici D, Shalf J, Chan C. "Fast Community Detection in Graphs with Infomap Method using Accelerated Sparse Accumulation." In Prof. of 37th IEEE International Parallel & Distributed Processing Symposium Workshops (IPDPSW 2023), IEEE; 2023; USA.
- Faysal M, Bremer M, Chan C, Shalf J, Arifuzzaman S. "Fast Parallel Index Construction for Efficient K-truss-based Local Community Detection in Large Graphs." 52nd International Conference on Parallel Processing (ICPP 2023), ACM; 2023; USA.
- Sattar NS., Arifuzzaman S. "Scalable distributed Louvain algorithm for community detection in large graphs." *Journal of Supercomputing (JSC)*. 2022 January; 78:10275.
- Sattar NS, Arifuzzaman S. COVID-19 Vaccination Awareness and Aftermath: Public Sentiment Analysis on Twitter Data and Vaccinated Population Prediction in the USA. *Applied Sciences*. 2021; 11(13):6128.
- Sattar N, Arifuzzaman S. "Data Parallel Large Sparse Deep Neural Network on GPU." 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW 2020). 2020 May 20.
- Sattar NS., Arifuzzaman S. "Community Detection using Semisupervised Learning with Graph Convolutional Network on GPUs." 2020 IEEE International Conference on BigData (BigData 2020). 2020 December.
- Arifuzzaman S, Khan M, Marathe M. "Fast Parallel Algorithms for Counting and Listing Triangles in Big Graphs." *ACM Transactions on Knowledge Discovery from Data (TKDD)*. 2020 January 01; 14(1).
- Sattar NS, Arifuzzaman S, Zibran MF and Sakib MM, "Detecting Web Spam in Webgraphs with Predictive Model Analysis," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 4299-4308.

Dr. Mingon Kang

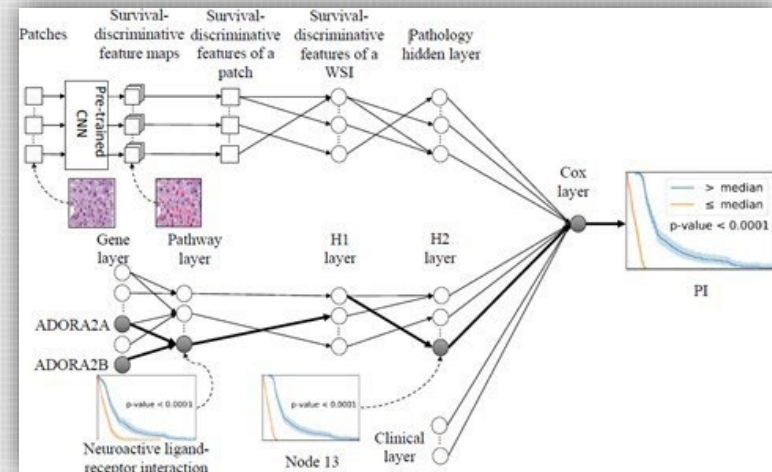
Assistant Professor, Department of Computer Science

Director, DataX Lab (<http://www.dataxlab.org>)

Phone: (702) 895-4884

Email: mingon.kang@unlv.edu

- Expertise
 - Machine learning, big data analytics, data science
 - Deep learning:
 - Interpretable deep learning
 - Evidential deep learning for trustworthy prediction
 - Integrative deep learning
 - Integrative analysis of multiple types of data, e.g., images and structured data
 - Medical image analysis



Dr. Mingon Kang

Assistant Professor, Department of Computer Science

Relevant Publications

- A. Yaganapu, M. Kang, "Multi-layered Self-attention Mechanism for Weakly Supervised Semantic Segmentation", *Computer Vision and Image Understanding* (IF: 4.5), 2023
- S. Han, M. Park, S. Kosaraju, J. Lee, J. H. Lee, T. Oh, M. Kang, "Evidential deep learning for trustworthy prediction of enzyme commission number", *Briefings in Bioinformatics* (IF: 9.5), 2023
- S. Kosaraju, J. Park, H. Lee, J. W. Yang and M. Kang, "Deep learning-based framework for slide-based histopathological image analysis", *Scientific Reports* (IF: 4.996), 2022
- J. Jo, S. Jung, J. Park, Y. Kim, and M. Kang, "Hi-LASSO: High-performance python and apache spark packages for feature selection with high-dimensional data", *PLOS ONE* (IF: 3.752), 2022
- M. Kang, E. Ko, T. Mersha, "A Roadmap for Multi-Omics Data Integration using Deep Learning", *Briefings in Bioinformatics*, 2022
- H. Lee, M. Kang, Y. Li, D. Seo, D. Kim, "Epidemic Vulnerability Index for Effective Vaccine Distribution against Pandemic", *ISBRA*, 2021
- J. H. Oh, W. Choi, E. Ko, M. Kang, A. Tannenbaum, and J. O. Deasy, "PathCNN: Interpretable convolutional neural networks for survival prediction and pathway analysis applied to glioblastoma", ppi443-i450, *Bioinformatics*, 2021
- S. Kim, S. Yang, K. Lim, E. Ko, H. Jang, M. Kang, P. Suh, and J. Joo, "Prediction of Alzheimer's disease-specific phospholipase c gamma-1 SNV by deep learning-based approach for high-throughput screening", *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 2021
- J. Hao, S. Kosaraju, N. Tsaku, D. H. Song, and M. Kang, "PAGE-Net: Interpretable and Integrative Deep Learning for Survival Analysis Using Histopathological Images and Genomic Data", *Pacific Symposium on Biocomputing (PSB)*, 2019
- T. Mallavarapu, J. Hao, Y. Kim, J.H. Oh, M. Kang, "Pathway-based Deep Clustering for Molecular Subtyping of Cancer," *Methods*, 2019

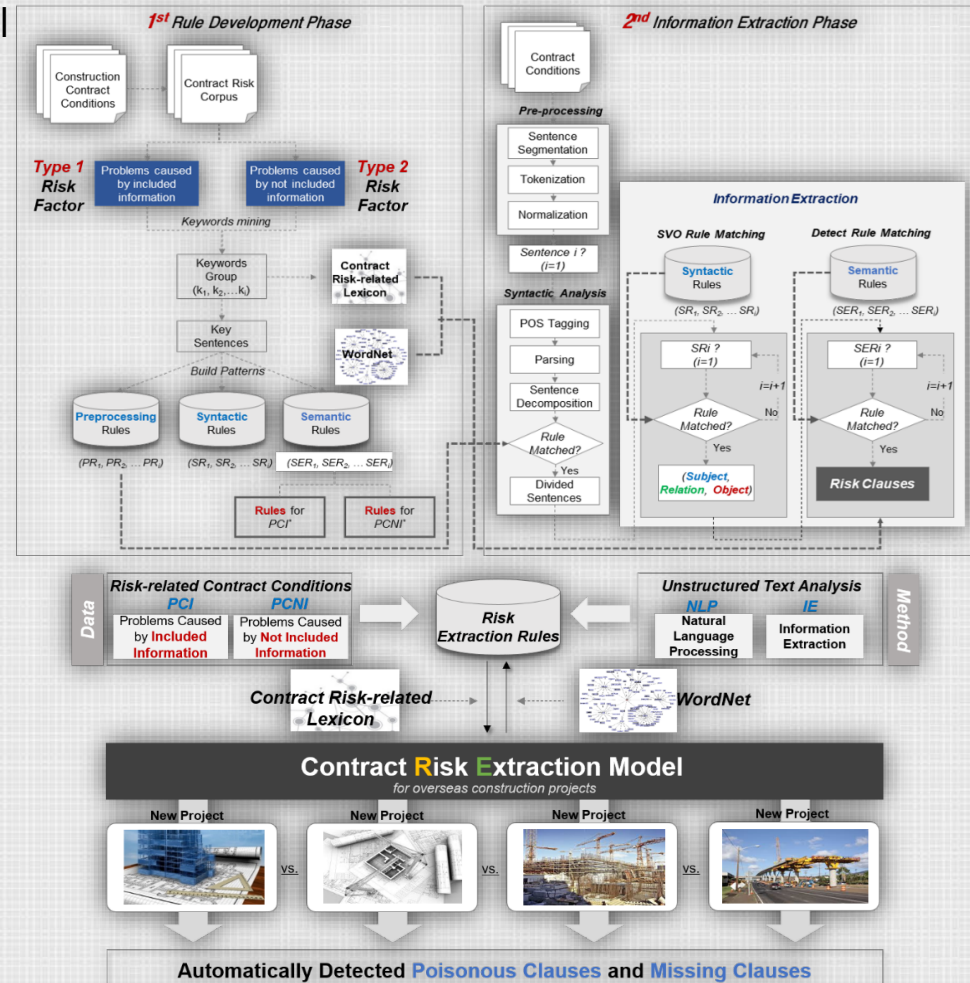
Dr. Jeehee Lee

Assistant Professor, Department of Civil and Environmental Engineering and Construction

Email: jeehee.lee@unlv.edu

Expertise

- Unstructured data analysis: natural language process (NLP), text mining, information retrieval, information extraction
- Physiological data sensing applications in smart building and smart city
- Interpretable machine learning



Dr. Jeehee Lee

Assistant Professor, Department of Civil and Environmental Engineering and Construction

Relevant Publications

- Ko, T., Lee, J., & David Jeong, H. (2024). "Project Requirements Prioritization through NLP-Driven Classification and Adjusted Work Items Analysis." *Journal of Construction Engineering and Management*, 150(3), 04023171.
- Chang, S., Oh, H. J., Lee, J., & Perkins, J. (2023). "Proof-of-Concept Study for Model-Based Construction Safety Diagnosis and Management Driven by Prevention through Design." *Journal of Management in Engineering*, 39(6), 04023044.
- Shrestha, R., Ko, T., and Lee, J. (2023). "Uncertainties Prevailing in Construction Bid Documents and Their Impact on Project Pricing through the Analysis of Prebid Requests for Information." *Journal of Management in Engineering*, 39(6), 04023040.
- Ko, T., David Jeong, H., and Lee, J. (2023). "Natural Language Processing–Driven Similar Project Determination Using Project Scope Statements." *Journal of Management in Engineering*, 39(3), 04023005.
- Oh, J., Wong, W., Castro-Lacouture, D., Lee, J., and Koo, C. (2023). "Indoor environmental quality improvement in green building: Occupant perception and behavioral impact." *Journal of Building Engineering*, 69, 106314.
- Koo, C., Si, K., Li, W., and Lee, J. (2022). "Integrated approach to evaluating the impact of feed-in tariffs on the life cycle economic performance of photovoltaic systems in China: A case study of educational facilities." *Energy*, 254, 124302.
- Lee, J., Ham, Y., and Yi, J., (2021). "Construction Disputes and Associated Contractual Knowledge Discovery Using Unstructured Text-Heavy Data: Legal Cases in the United Kingdom". *Sustainability (MDPI)*, 13(16), 9403.
- Lee, J. and Ham, Y. (2021). "Physiological sensing-driven personal thermal comfort modeling in consideration of occupant activity variations." *Building Research & Information (Taylor & Francis)*.
- Lee, J., Ham, Y., Yi, J., and Son, J. (2020). "Effective risk positioning through automated identification of missing contract conditions from the contractor's perspective: based on the FIDIC contract cases".
- Hwang, S., Lee, J., Yi, J., and Kim, M. (2019). "Korean Public Rental Housing for Residential Stability of the Younger Population: Analysis of Policy Impacts using System Dynamics". *Journal of Asian Architecture and Building Engineering (Taylor & Francis)*. 18(3).

Dr. Fatma Nasoz

Associate Professor, Department of Computer Science

Phone: (702) 895-0097

Email: fatma.nasoz@unlv.edu

- Expertise
 - Machine learning and deep learning
 - Data analysis and visualization
 - Human-computer interaction
 - Cloud computing

Dr. Fatma Nasoz

Associate Professor, Department of Computer Science

Relevant Publications

- Z Fitzhugh, MR Schiller Ph D, F Nasoz (2021). "Predicting Variant Pathogenicity with Machine Learning." *UNLV Undergraduate Research Symposium Posters*. 46.
- Wu, Q., Nasoz, F., Jung, J., Bhattarai, B., Han, M.V., Greenes, R.A., and Saag, K.G. (2021). "Machine learning approaches for the prediction of bone mineral density by using genomic and phenotypic data of 5130 older men." *Scientific Reports*. 11(1):4482, doi: 10.1038/s41598-021-83828-3
- Wu, Q., Nasoz F., Jung, J., Bhattarai, B, Han, MV. (2020). "Machine learning approaches for fracture risk assessment: a comparative analysis of genomic and phenotypic data in 5,130 older men." *Calcified Tissue International*. 107 (4): 353–361, doi: 10.1007/s00223-020-00734-y
- Singh, S. and Nasoz, F. (2020). "Facial expression recognition with convolutional neural networks." *10th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, pp. 0324- 0328, doi: 10.1109/CCWC47524.2020.9031283.
- Shrestha, N. and Nasoz, F. (2019). "Deep Learning Sentiment Analysis of Amazon.com Reviews and Ratings." *International Journal on Soft Computing, Artificial Intelligence, and Applications*, Vol. 8(1): 1-15.
- Nasoz F. and Shrestha C. (2017) "A Web-Based User Interface for Machine Learning Analysis". In: Yamamoto S. (eds) *Human Interface and the Management of Information: Supporting Learning, Decision-Making and Collaboration*. HIMI 2017. *Lecture Notes in Computer Science*, vol. 10274. Springer, Cham. doi: 10.1007/978-3-319-58524-6_35

Dr. Emma E. Regentova

Professor, Department of Electrical and Computer Engineering

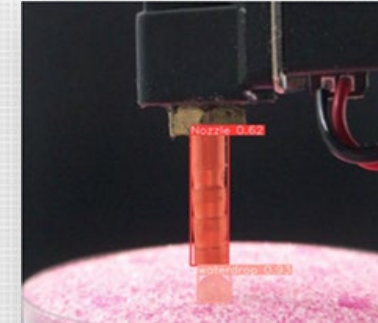
Phone: (702) 895-3187

Email: emma.regentova@unlv.edu

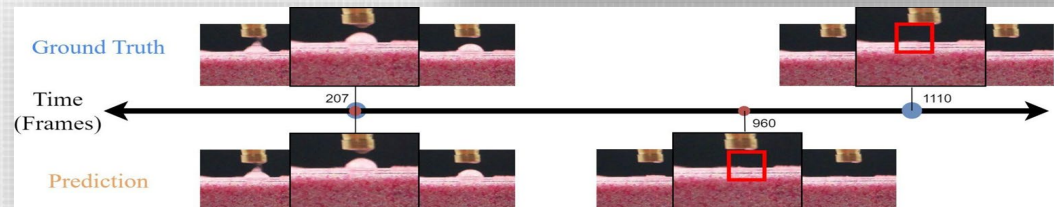
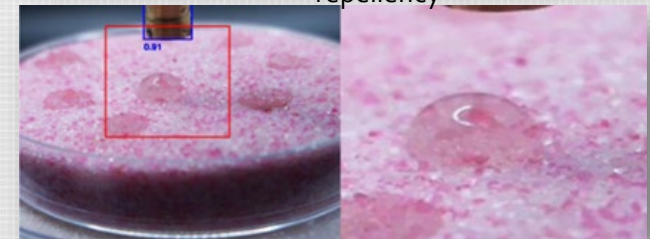
- Expertise
 - Analysis of images (medical, photon- and neutron CAT), video, multi- and hyperspectral data for applications
 - Machine learning for object detection, semantic segmentation and temporal action localization
 - DSP accelerator circuits on FPGA



Methane gas leaks detection on the drone platform



Estimation of water drop penetration time and evaluation of soil water repellency



Dr. Emma Regentova

Professor, Department of Electrical and Computer Engineering

Relevant Publications

- D.Wang, E. Regentova, V. Muthukumar, M. Berli, F. Harris, "A Machine Learning Framework to Measure Water Drop Penetration Time (WDPT) for Soil Water Repellency Analysis," *Machine Learning with Applications*, Elsevier, April 2024.
- Chato, L.; Regentova, E., "Survey of Transfer Learning Approaches in the Machine Learning of Digital Health Sensing Data." *J. Pers. Med.* 2023, 13, 1703.
- S. Nasr-Esfahani, V. Muthukumar, E.E. Regentova, K. Taghva; M. B. Trabia, "Olive pit detection using Hyperspectral imagery", *IEEE Access*, Volume: 10, 30 May 2022, pp. 58525 – 58536.
- S. Nasr-Esfahani, V. Muthukumar, E. Regentova, K. Taghva, M. Trabia, "Food Recognition Improvement by using Hyper-Spectral Imagery", *International Journal of Advanced Computer Research (IJACR)*, Vol 11(53) ISSN (Print): 2249-7277 ISSN (Online): 2277-7970.
- Yang Jiao, Barbara St. Pierre Schneider, Emma Regentova, Mei Yang, "DeepQuantify: deep learning and quantification system of white blood cells in light microscopy images of injured skeletal muscles", *Journal of Medical Imaging*, 6(2), 024006 (2019).
- Ali Pour Yazdanpanah, Jessica Hartman, Emma Regentova, Alexander Barzilov, "Sparse-View Photon-Neutron Computed Tomography: Object Reconstruction and Material Discrimination," *Applied Radiation and Isotopes*, *Applied Radiation and Isotopes*, Volume 132, Pages 1-232 (February 2018).
- Ali P. Yazdanpanah, E.E. Regentova, "Compressed Sensing MRI Based on Shearlet Sparsity and Nonlocal Total Variation: SS-NLTV," *Journal of Medical Imaging*, 201, 4(2), 026003 (Apr - Jun 2017) 2017.

Dr. Haroon Sahotra

Professor, Department of Civil and Environmental Engineering

Director, GIS and Remote Sensing Core Lab

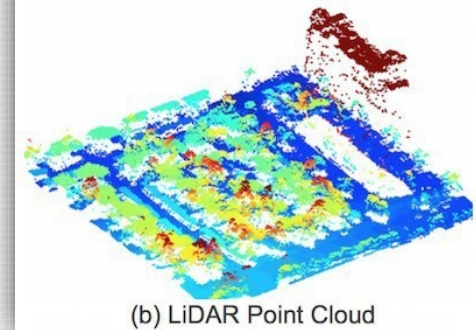
Phone: (702) 774-1463

Email: haroon.stephen@unlv.edu

- Expertise
 - Sensors aboard UAS for remote sensing: testing and applications
 - UAS and GIS integration
 - Applications of UAS to resource mapping, monitoring, and management
 - Data fusion and data mining of multi-scale sources: UAS, aerial, and spaceborne sensors



(a) Google Earth 3D Buildings



(b) LiDAR Point Cloud

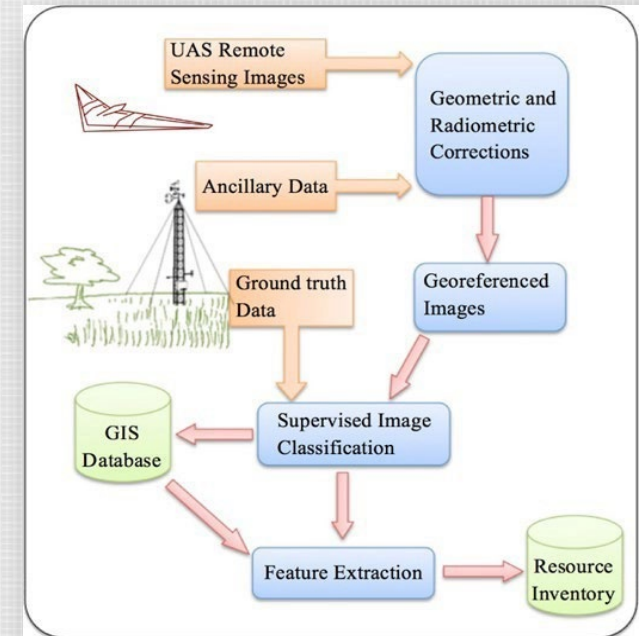
Example of 3D view of buildings and land cover

Dr. Haroon Sahotra

Professor, Department of Civil and Environmental Engineering and Construction

Relevant Publications

- Adjovu, G. E., Ahmad, S. & Stephen, H., (2023). "A Machine Learning Approach for the Estimation of Total Dissolved Solids Concentration in Lake Mead Using Electrical Conductivity and Temperature." *Water*, 15(13), 2439.
- Shrestha, B., Ahmad, S., & Stephen, H. (August 2021). "Fusion of Sentinel-1 and Sentinel-2 data in mapping the impervious surfaces at city scale." *Environmental Monitoring and Assessment*, 193(9), 1-21.
- Shrestha, B., Stephen, H., & Ahmad, S. (August 2021). "Impervious Surfaces Mapping at City Scale by Fusion of Radar and Optical Data through a Random Forest Classifier." *Remote Sensing*, 13(15), 3040.
- Abedin, S. J. H., & Stephen, H. (2019). "GIS Framework for Spatiotemporal Mapping of Urban Flooding." *Geosciences*, 9(2), 77.
- Stephen, H. (2018). "Trend Analysis of Las Vegas Land Cover and Temperature Using Remote Sensing." *Land*, 7(4), 135.
- Lachniet, M. S., Lawson, D. E., Stephen, H., Sloat, A. R., & Patterson, W. P. (2016). "Isoscapes of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ reveal climatic forcings on Alaska and Yukon precipitation." *Water Resources Research*, 52(8), 6575-6586.
- Black, A., & Stephen, H. (2014). "Relating temperature trends to the normalized difference vegetation index in Las Vegas." *GIScience & Remote Sensing*, 51(4), 468-482. doi: 10.1080/15481603.2014.940695.



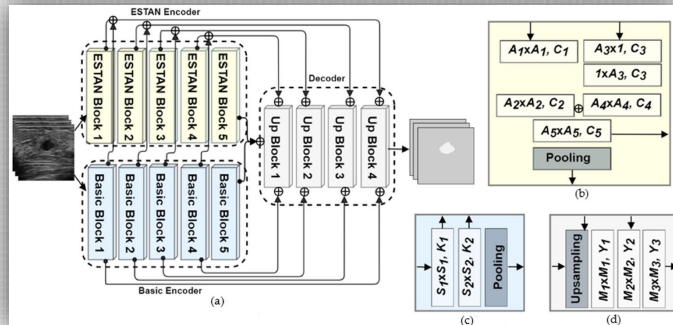
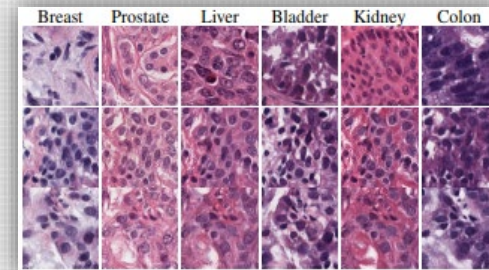
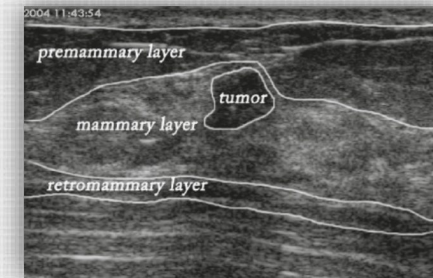
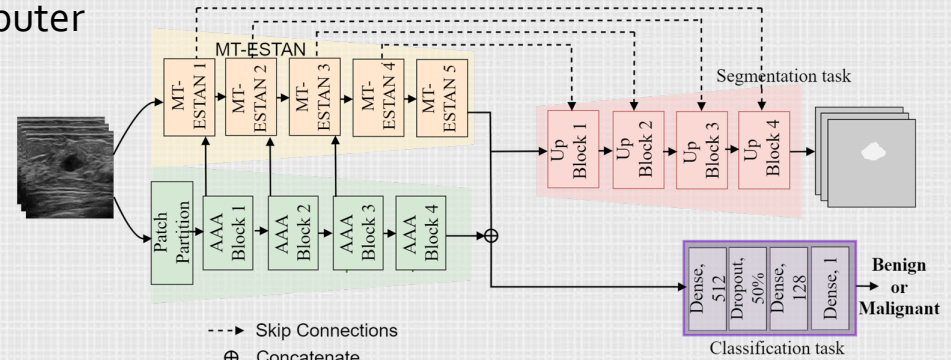
Dr. Bryar Shareef

Assistant Professor, Department of Computer Science

Phone: (702) 774-3420

Email: bryar.shareef@unlv.edu

- Expertise
 - Machine learning
 - Deep learning
 - Medical image analysis
 - Image processing and computer vision



Dr. Bryar Shareef

Assistant Professor, Department of Computer Science

Relevant Publications

- Bryar Shareef, Min Xian, Aleksandar Vakanski, Haotian Wang, "Breast Ultrasound Tumor Classification using a Hybrid Multitask CNN-Transformer Network," 2023, *MICCAI* 2023.
- B. Shareef, M. Xian, A. Vakanski, J. Ding, C. Ning, H.D. Cheng, "A benchmark for breast ultrasound image classification," *Ultrasound in Medicine & Biology*, 2023.
- H. Wang, M. Xian, A. Vakanski, and B. Shareef, "SIAN: style-guided instance-adaptive normalization for multi-organ histopathology image synthesis," *IEEE International Symposium on Biomedical Imaging*, 2023, pp. 1-5.
- B. Shareef, A. Vakanski, P. E. Freer, and M. Xian, "Estan: Enhanced small tumor-aware network for breast ultrasound image segmentation," *Healthcare*, vol. 10, no. 11, pp. 2262, 2022.
- Y. Zhang, M. Xian, H.-D. Cheng, B. Shareef, J. Ding, F. Xu, K. Huang, B. Zhang, C. Ning, and Y. Wang, "BUSIS: A Benchmark for Breast Ultrasound Image Segmentation," *Healthcare*, vol. 10, no. 4, pp. 729, 2022-04-14, 2022.
- B. Shareef, M. Xian, and A. Vakanski, "STAN: Small tumor-aware network for breast ultrasound image segmentation," *IEEE International Symposium on Biomedical Imaging*, Iowa City, Iowa, USA, 2020, pp. 1-5.
- R. E. Hiromoto, M. Haney, A. Vakanski, and B. Shareef, "Toward a Secure IoT Architecture," *Advanced Control Techniques in Complex Engineering Systems: Theory and Applications: Dedicated to Professor Vsevolod M. Kuntsevich*, pp. 297-323, 2019.

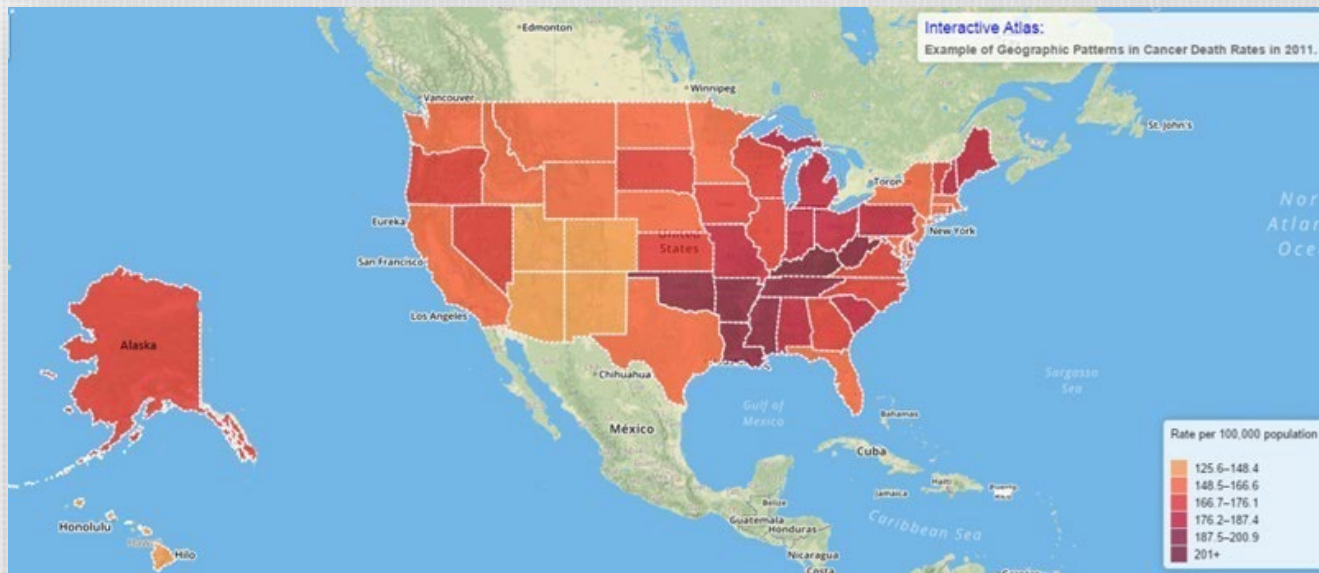
Dr. Kazem Taghva

Chair and Professor, Department of Computer

Science Phone: (702) 895-0873

Email: kazem.taghva@unlv.edu

- Expertise
 - Databases
 - Machine learning
 - Information retrieval



Dr. Kazem Taghva

Chair and Professor, Department of Computer Engineering

Relevant Publications

- P Puranik, K Taghva, K Ghaharian. "Descriptive Analysis of Gambling Data for Data Mining of Behavioral Patterns". *International Conference on Interactive Collaborative Robotics*, 40-51, 2023, Springer.
- P Puranik, K Taghva, K Ghaharian. "Check for updates Descriptive Analysis of Gambling Data for Data Mining of Behavioral Patterns". *Proceedings of the Second International Conference on Innovations in Computing Research (ICR'23)*, 2023, Springer.
- K Ghaharian, P Puranik, B Abarbanel, K Taghva, SW Kraus, A Singh, A Feldman, B Bernhard. "Payments transaction data from online casino players and online sports bettors". *Data in Brief* 48, Elsevier, 2023.
- P Nahed, MEZN Kamar, JRF Cacho, G Lee, J Cummings, K Taghva. "A Recommendation Model for Predicting Alzheimer's Drugs' Mechanism of Action". *Intelligent Sustainable Systems: Selected Papers of WorldS4 2022, Volume 1*, pp 63-73, Springer, 2023.
- S Choi, P Puranik, B Dahal, K Taghva. "How to generate data for acronym detection and expansion". *Advances in Computational Intelligence* 2 (2), 1-8, 2022.
- J Cummings, G Lee, K Zhong, J Fonseca, K Taghva. "Alzheimer's disease drug development pipeline: 2021" *Alzheimer's & Dementia: Translational Research & Clinical Interventions* vol. 7 issue 1, e12179, 2021.
- Pouyan Nahed; Mina E. Zahed Nojoo Kamar; Jorge Ramón Fonseca Cacho, Garam Lee; Jeffrey Cummings; Kazem Taghva. "Clinical Text Classification of Alzheimer's Drugs' Mechanism of Action using BERT.", *Proceedings of Sixth International Congress on Information and Communication Technology* pp 513-521, 2021.
- Cacho, Jorge Ramón Fonseca, Ben Cisneros, and Kazem Taghva. "Building a Wikipedia N-GRAM Corpus." *Proceedings of SAI Intelligent Systems Conference*. Springer, Cham, 2020.
- Cacho, Jorge Ramón Fonseca, and Kazem Taghva. "Aligning ground truth text with OCR degraded text." *Intelligent Computing-Proceedings of the Computing Conference*. Springer, Cham, 2019.
- Fonseca Cacho J.R., Taghva K. "Reproducible Research in Document Analysis and Recognition." *Advances in Intelligent Systems and Computing*, vol. 738, pp. 389-395, 2018, Springer.