BURAK UZKENT, Ph.D.

Santa Clara, CA | +1-650-861-8068 | uzkent.burak@gmail.com | uzkent.github.io

PROFESSIONAL SUMMARY

Principal Member of Staff with 10+ years of experience developing and deploying large-scale machine learning systems. Specialized in generative AI, computer vision, multimodal and video-language modeling, and efficient transformer architectures. Published 40+ papers in top-tier venues (CVPR, ICCV, ICLR, NeurIPS, AAAI, EMNLP) with extensive experience in both academic research and industrial applications. Proven track record of leading research initiatives at AMD, Amazon, Samsung, and Stanford University.

PROFESSIONAL EXPERIENCE

Principal Member of Staff

April 2026 - Present

AMD | Santa Clara, CA

- Work on applications of Generative AI on AMD hardware
- Develop and evaluate ML systems optimized for AMD accelerators and platforms

Machine Learning Scientist

April 2022 - March 2026

Amazon Prime Video | Sunnyvale, CA

- Led development of Video LLMs for advanced video understanding and content moderation in long-form videos
- Designed and implemented multimodal foundation models for video summarization pipeline
- Developed transformer-based NLP models for subtitle analysis enabling automated content moderation at scale
- Published papers (CVPR, WACV) and filed multiple patents; papers under review at EMNLP and ECCV

Senior Research Scientist

November 2020 - April 2022

Samsung Research America | Mountain View, CA

- Optimized vision transformer architectures achieving significant model compression while maintaining accuracy
- Developed efficient multi-modal transformers for on-device applications with reduced inference latency
- Led team of 3 researchers in developing novel computer vision and multimodal ML models
- Published 4 papers in top-tier conferences (CVPR, ICLR, AAAI) and filed 6 patents

Postdoctoral Fellow

July 2018 - October 2020

Stanford University, Department of Computer Science | Stanford, CA

- Published 15+ papers in top-tier conferences (ICCV, CVPR, ICLR, AAAI, IJCAI, KDD)
- Developed novel self-supervised and weakly supervised learning approaches for remote sensing
- Created efficient deep learning models using reinforcement learning for adaptive computation
- Built ML models for sustainability applications including poverty mapping and farmland delineation

Computer Vision Engineer

June 2017 - July 2018

Planet Labs | San Francisco, CA

- Built large-scale object detection dataset with 100K+ annotated satellite images
- Improved small object detection accuracy in low-resolution aerial imagery using convolutional detectors
- Conducted research to tackle unique challenges of satellite image object detection

Computer Vision Engineer

August 2016 - June 2017

Autel Robotics | San Ramon, CA

- Designed long-term target following system for next-generation drones
- Implemented online learning method for real-time single object tracking on embedded platforms
- Deployed and optimized tracking algorithms on low-end embedded systems

Computer Vision Algorithm Engineer Intern

November 2015 - May 2016

Futurewei Technologies (Huawei R&D) | Bridgewater, NJ

- Designed subspace learning method for stranger detection in family photo albums using deep CNNs
- Developed probabilistic graph-based approach for semantic role assignment in family photos

EDUCATION

Rochester Institute of Technology

August 2011 - May 2016

Ph.D. in Imaging Science, Chester F. Carlson Center for Imaging Science

Thesis: Aerial visual vehicle detection and tracking using an adaptive, multi-modal sensor

Advisor: Matthew J. Hoffman, Ph.D.

University of Bridgeport

August 2009 - May 2011

M.S. in Electrical Engineering

Thesis: Environmental non-speech sound classification with a new set of time-domain features

Advisor: Buket D. Barkana, Ph.D.

Eskisehir Osmangazi University

September 2004 - May 2009

B.S. in Electrical and Electronics Engineering

Thesis: Autonomous parallel parking of non-holonomic vehicles

Advisor: Osman Parlaktuna, Ph.D.

RESEARCH EXPERIENCE

Graduate Research Assistant

April 2012 - July 2016

RIT, Chester F. Carlson Center for Imaging Science | Rochester, NY

- Conducted research on aerial vehicle detection and tracking using adaptive, multi-modal sensors
- Developed computer vision and ML methods for vehicle detection, association, and tracking in aerial video
- Addressed challenges of medium-to-high altitude tracking through efficient use of hyperspectral data

Graduate Research Assistant
Rochester Institute of Technology
Advisor: Elizabeth Cherry, Ph.D.

May 2014 – June 2014
Rochester, NY

- Performed research on 3-D MRI cardiac segmentation for understanding ventricular fibrillation mechanics

Graduate Research Assistant
University of Bridgeport, Electrical Engineering Department
Advisor: Buket D. Barkana, Ph.D.

August 2009 – May 2011
Bridgeport, CT

- Conducted research on environmental sound classification for abnormal event detection
- Proposed novel pitch range-based features to improve classification accuracy

TEACHING EXPERIENCE

Teaching Assistant – Data for Development — Sept 2019 – Dec 2019
Stanford University — Instructors: David Lobell, Marshall Burke, Stefano Ermon

Graduate Teaching Assistant – Pattern Recognition — Jan 2015 – May 2015
Rochester Institute of Technology — Instructor: John Kerekes, Ph.D.

Graduate Teaching Assistant – Programming for Imaging Science — Aug 2011 – May 2012
Rochester Institute of Technology — Instructor: Jeff Pelz, Ph.D.

Graduate Teaching Assistant – Digital Image/Audio/Speech Signal Processing — Aug 2010 – May 2011
University of Bridgeport — Instructor: Buket D. Barkana, Ph.D.

Graduate Teaching Assistant – Non-linear Control Systems — Jan 2009 – May 2009
Eskisehir Osmangazi University — Instructor: Osman Parlaktuna, Ph.D.

TECHNICAL SKILLS

Machine Learning: PyTorch, TensorFlow, Caffe, Transformers, Multi-modal Learning, Computer Vision, LLMs

Programming: Python, C/C++, MATLAB, Shell Scripting, Git, HTML

Frameworks & Tools: OpenCV, Linux, Docker, AWS, IDL/ENVI, LaTeX

Research Areas: Video Understanding, Multi-modal ML, Self-Supervised Learning, Object Detection & Tracking

REFEREED JOURNAL PUBLICATIONS

[1] E. Velterop, **B. Uzkent**, J. Suckale, "Safe Shelter: A Case for Prioritizing Housing Quality in Climate Adaptation Policy by Remotely Sensing Roof Tarps in the San Francisco Bay Area", *Earth's Future* 10, no. 8 (2022)

[2] **B. Uzkent**, A. Rangnekar, M.J. Hoffman, "Tracking in Aerial Hyperspectral Videos Using Deep Kernelized Correlation Filters", *IEEE Transactions on Geoscience and Remote Sensing*, 57(1): 449-461, August 2018. [Code]

[3] **B. Uzkent**, M.J. Hoffman, A. Vodacek, "Integrating Hyperspectral Likelihoods in a Multi-dimensional Assignment Algorithm for Aerial Vehicle Tracking", *IEEE Journal of Selected Topics in Remote Sensing and Observation*, 9(9): 4325-4333, May 2016.

[4] **B. Uzkent**, M.J. Hoffman, A. Vodacek, B. Chen, "Feature Matching with an Adaptive Optical Sensor in a Ground Target Tracking System", *IEEE Sensors Journal*, 15(1): 510-519, January 2015.

[5] **B. Uzkent**, B.D. Barkana, H. Cevikalp, "Non-speech environmental sound classification using SVMS with a new set of features", *International Journal of Innovative Computing, Information and Control*, 8(5): 3511-3524, May 2012.

[6] B.D. Barkana, **B. Uzkent**, I. Saricicek, "Normal and abnormal non-speech audio event detection using MFCC and PR-based feature sets", *Advanced Materials Research*, Volume 601, pp: 200–208, December 2012.

[7] B.D. Barkana, **B. Uzkent**, I. Saricicek, "Environmental noise classifier using a new set of feature parameters based on pitch range", *Applied Acoustics*, 72(11): 841–848, November 2011.

[8] **B. Uzkent**, B.D. Barkana, J. Yang, "Automatic environmental noise source classification model using fuzzy logic", *Expert Systems with Applications*, 38(7): 8751–8755, July 2011.

REFEREED CONFERENCE PUBLICATIONS

[1] J. Yi, **B. Uzkent**, O. Ignat, Z. Li, A. Garg, X. Yu, L. Liu, "Augment the Pairs: Semantics-Preserving Image-Caption Pair Augmentation for Grounding-Based Vision and Language Models", *IEEE Winter Conference on Applications of Computer Vision*, **WACV-24**, January 2024.

[2] K. Doshi, A. Garg, **B. Uzkent**, X. Wang, "A Multimodal Benchmark and Improved Architecture for Zero Shot Learning", *IEEE Winter Conference on Applications of Computer Vision*, **WACV-24**, January 2024.

[3] **B. Uzkent**, A. Garg, W. Zhu, K. Doshi, J. Yi, X. Wang, M. Omar, "Dynamic Inference with Grounding Based Vision and Language Models", *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, **CVPR-23**, pp. 2624-2633, 2023.

[4] S. Gao, **B. Uzkent**, Y. Shen, H. Huang, H. Jin, "Learning to Jointly Share and Prune Weights for Grounding Based Vision and Language Models", *International Conference on Learning Representations*, **ICLR-23**, May 2023.

[5] M. Yin, **B. Uzkent**, Y. Shen, H. Jin, "GOHSP: A Unified Framework of Graph and Optimization-based Heterogeneous Structured Pruning for Vision Transformer", *AAAI Conference on Artificial Intelligence*, **AAAI-23**, February 2023.

[6] Q. Lu, Y.C. Shu, **B. Uzkent**, T. Hua, Y. Shen, H. Jin, "Lite-MDETR: A Lightweight Multi-Modal Detector", *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, **CVPR-22**, pp. 12206-12215, 2022.

[7] S. Chakraborty, **B. Uzkent**, K. Ayush, K. Tanmay, E. Sheehan, S. Ermon, "Efficient Conditional Pre-training for Transfer Learning", *CVPR Workshop*, **CVPRW-22**, pp. 4241-4250, 2022.

[8] **B. Uzkent**, K. Ayush, C. Meng, M. Burke, D. Lobell, S. Ermon, "Geography-Aware Self-Supervised Learning", *IEEE/CVF International Conference on Computer Vision*, **ICCV-21**, pp. 10181-10190, 2021. **[Code]**

[9] K. Ayush, A. Sinha, J. Song, **B. Uzkent**, H. Jin, S. Ermon, "Negative Data Augmentation", *International Conference on Learning Representations*, **ICLR-21**, May 2021. **[Code]**

[10] **B. Uzkent**, K. Ayush, M. Burke, D. Lobell, S. Ermon, "Efficient Poverty Mapping from High Resolution Remote Sensing Images", *AAAI Conference on Artificial Intelligence*, **AAAI-21**, February 2021.

[11] J. Lee, D. Grosz, **B. Uzkent**, S. Zheng, M. Burke, D. Lobell, S. Ermon, "Predicting Livelihood Indicators from Community-Generated Street-Level Imagery", *AAAI Conference on Artificial Intelligence*, **AAAI-21**, February 2021. **[Code]**

[12] **B. Uzkent**, K. Ayush, M. Burke, D. Lobell, S. Ermon, "Generating Interpretable Poverty Maps Using Object Detectors in Satellite Images", *International Joint Conference on Artificial Intelligence*, **IJCAI-20**, pp. 4410-4416, August 2020.

[13] **B. Uzkent**, S. Ermon, "Learning Where and When to Zoom Using Deep Reinforcement Learning", *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, **CVPR-20**, pp. 12345-12354, June 2020. **(Oral Presentation)** **[Code]**

[14] H.L. Aung, **B. Uzkent**, M. Burke, D. Lobell, S. Ermon, "Farmland Parcel Delineation using Spatio-temporal Convolutional Networks", *CVPR Workshop on Agriculture-Vision*, **CVPRW-20**, pp. 76-77, June 2020. **[Code]**

[15] **B. Uzkent**, C. Yeh, S. Ermon, "Efficient Object Detection in Large Images Using Deep Reinforcement Learning", *IEEE Winter Conference on Applications of Computer Vision*, **WACV-20**, pp. 1824-1833, March 2020. **[Code]**

[16] V. Sarukkai, A. Jain, **B. Uzkent**, S. Ermon, "Cloud Removal from Satellite Images using Spatiotemporal Generative Networks", *IEEE Winter Conference on Applications of Computer Vision*, **WACV-20**, pp. 1796-1805, March 2020. **[Code]**

[17] **B. Uzkent**, E. Sheehan, C. Meng, D. Lobell, M. Burke, S. Ermon, "Learning to Interpret Satellite Images using Wikipedia", *International Joint Conference on Artificial Intelligence*, **IJCAI-19**, pp. 3620-3626, July 2019. **[Code]**

[18] E. Sheehan, C. Meng, M. Tan, **B. Uzkent**, N. Jean, D. Lobell, M. Burke, S. Ermon, "Predicting Economic Development using Geolocated Wikipedia Articles", *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, **KDD-19**, pp. 2698-2706, July 2019. **[Code]**

[19] **B. Uzkent**, Y. Seo, "EnKCF: Ensemble of Kernelized Correlation Filters for High-Speed Object Tracking", *IEEE Winter Conference on Applications of Computer Vision*, **WACV-18**, pp. 1133-1141, March 2018. **[Code]**

[20] **B. Uzkent**, A. Rangnekar, M.J. Hoffman, A. Vodacek, "Aerial Vehicle Tracking by Adaptive Fusion of Likelihood Maps", *IEEE Workshop on Perception Beyond the Visible Spectrum, CVPR Workshop*, **CVPRW-17**, pp. 39-48, July 2017.

[21] **B. Uzkent**, M.J. Hoffman, A. Vodacek, "Real time Target Detection and Tracking in Aerial Video using Hyperspectral Features", *IEEE Workshop on Moving Cameras Meet Video Surveillance, CVPR Workshop*, **CVPRW-16**, pp. 36-44, June 2016.

[22] **B. Uzkent**, M.J. Hoffman, A. Vodacek, "Spectral Validation of Measurements in a Vehicle Tracking DDDAS", *International Conference on Computational Science*, Volume 51, pp. 2493-2502, June 2015.

[23] **B. Uzkent**, M.J. Hoffman, A. Vodacek, "Background Image Understanding and Adaptive Imaging for Vehicle Tracking", *SPIE Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications XII*, pp. 94600-94607, April 2015.

[24] **B. Uzkent**, M.J. Hoffman, A. Vodacek, "Efficient Integration of Spectral Features for Vehicle Tracking utilizing an Adaptive Sensor", *SPIE Video Surveillance and Transportation Imaging Applications*, pp. 940707-940717, February 2015.

[25] **B. Uzkent**, M.J. Hoffman, E. Cherry, N. Cahill, "3-D MRI Cardiac Segmentation using Graph Cuts", *IEEE Western NY Image Processing Workshop*, pp. 47-51, November 2014.

[26] **B. Uzkent**, M.J. Hoffman, A. Vodacek, J.P. Kerekes, B. Chen, "Feature matching and adaptive prediction models in an object tracking DDDAS", *Procedia Computer Science*, Volume 18, pp. 1939-1948, 2013.

[27] **B. Uzkent**, B.D. Barkana, "Pitch range-based feature extraction for audio surveillance systems", *IEEE International Conference on Information Technology: New Generations (ITNG)*, pp. 476-480, April 2011.

[28] B.D. Barkana, I. Saricicek, **B. Uzkent**, "Performances of the ANN, SVM, K-means clustering methods recognizing different environmental sounds", *24th European Conference on Operational Research*, Lisbon, Portugal, July 2010.

[29] **B. Uzkent**, O. Parlaktuna, "Autonomous parallel parking of non-holonomic vehicles", *13th National Conference in Middle East Technical University*, Ankara, Turkey, 2009.

PAPERS UNDER REVIEW

[1] T. Poppi, B. Uzkent, A. Garg, L. Porto, G. Kessler, Y. Yang, M. Cornia, L. Baraldi, R. Cucchiara, F. Schiffers, "CounterVid: Counterfactual Video Generation for Mitigating Action and Temporal Hallucinations in Video-Language Models", EMNLP-26.

[2] G. Sun, A. Singhal, B. Uzkent, M. Shah, C. Chen, G. Kessler, "From Frames to Clips: Efficient Key Clip Selection for Long-form Video Understanding", ECCV-26.

[3] A. Blume, B. Uzkent, G. Kessler, "Learning to Rank Caption Chains for Video-Text Alignment", ECCV-26.

[4] R. Jain, K. Doshi, B. Uzkent, G. Kessler, "Narrative Aligned Long Form Video Question Answering", CVPR Workshop-26 (Best Paper Candidate).

PROFESSIONAL SERVICE

Peer Reviewer: IEEE TGRS, IEEE TIFS, NeurIPS, WACV, ICCV, BMVC, IEEE TIP, ICML, ICLR, CVPR, IEEE Sensors Journal, Nature Machine Intelligence, IEEE Access

HONORS & AWARDS

* Amazon Invention Award (June 2024)
* RIT Graduate Scholarship Award (September 2011 - May 2016)
* University of Bridgeport Dean's Scholarship Award (August 2009 - May 2011)
* University of Bridgeport Outstanding Student Award (May 2011)
* Fulbright Opportunity Grant (August 2009)
* Erasmus Exchange Student (September 2007 - January 2008)

LANGUAGES

English - Advanced/Fluent
Turkish - Native