NILUTHPOL CHOWDHURY MITHUN

Advanced Computer Scientist, SRI International
https://niluthpol.github.io/
Phone:+1-951-907-2396 ⋄ Email: niluthpol.mithun@sri.com

RESEARCH INTERESTS

· Multimedia · Computer Vision · Machine Learning

EDUCATION

University of California, Riverside (UC Riverside) June 2019
Ph.D. in Electrical & Computer Engineering *Riverside, CA, USA*

Bangladesh University of Engineering & Technology (BUET) May 2014
M.Sc. in Electrical Engineering *Dhaka, Bangladesh*

Bangladesh University of Engineering & Technology (BUET) Feb. 2011
B.Sc. in Electrical Engineering *Dhaka, Bangladesh*

EXPERIENCE

Center for Vision Technologies (CVT), SRI International, Princeton, NJ
Advanced Computer Scientist I (June 2019 - Present)

- Vision-based navigation in GPS challenged environments • Matching Geo-referenced Data across Modalities
- Multimodal Data Analysis with Limited Supervision • Deep Reinforcement Learning for Navigation

Video Computing Group, UC Riverside Advisor: Dr. Amit K. Roy-Chowdhury
Graduate Student Researcher (Sept. 2014 - June 2019)

- Cross-Modal Vision-Language Retrieval · Learning with Limited Supervision
- Weakly Supervised Learning • Image Collection Summarization

Center for Vision Technologies, SRI International, Princeton Mentor: Dr. Han-Pang Chiu
Research Intern (Jun. 2018 - Sept. 2018)

- Cross-modal visual localization by matching query RGB images to Geo-referenced point cloud

Smart Campus Lab, Bosch Research Center, Pittsburgh Mentor: Dr. Sirajum Munir
Research Intern (Jun. 2017 - Sept. 2017)

- Real-time Object Detection from RGBD Images • Video to Text Retrieval

Digital Signal Processing Group, BUET Advisor: Dr. S. M. Mahbubur Rahman
Student Researcher (Jan. 2011 - Aug. 2014)

- Detection, Classification and Tracking of Vehicles from Video • Retinal Image Analysis

Advanced R&D Group [Now Solution Lab], Samsung R&D Institute Bangladesh, Dhaka
Sr. Software Engineer (Jan. 2013 - Aug. 2014), Software Engineer (Mar. 2011 - Jan. 2013)

- Virtual Platform based Optimization of SoC • Algorithm optimization for Samsung Reconfigurable Processor

Design Solution Lab, DMC R&D Center, Samsung Electronics, Suwon, Korea
Software Engineer [Expat] (Apr. 2012 - Jul. 2012) and (Aug. 2011 - Nov. 2011)

- Post-silicon Verification of Display Processor of DRIMe IV SoC. • Virtual HW IP for Pre-Silicon Verification

SELECTED PUBLICATIONS (Please see Google Scholar for full list.)

1. M. Sizintsev, **N. C. Mithun**, H. P. Chiu, S. Samarasekera, R. Kumar, Text-based Localization of Moments in a Video Corpus, *IEEE Trans. Visualization and Computer Graphics (**TVCG**)*, 2021.

2. S. Paul, **N. C. Mithun**, A. K. Roy-Chowdhury, Text-based Localization of Moments in a Video Corpus, *IEEE Trans. Image Processing (**TIP**)*, 2021.

3. Z. Seymour, K. Thopalli, **N. C. Mithun**, H. P. Chiu, S. Samarasekera, R. Kumar, "MaAST: Map Attention with Semantic Transformers for Efficient Visual Navigation", in *Int. Conf. Robotics Automation (**ICRA**)*, 2021.

4. **N. C. Mithun**, K. Sikka, H. P. Chiu, S. Samarasekera, R. Kumar, "RGB2LIDAR: Towards Solving Large-Scale Cross-Modal Visual Localization", in *ACM Multimedia Conference (**ACM MM**)*, 2020.[**Best Paper Candidate**]

5. **N. C. Mithun**, S. Paul, A. K. Roy-Chowdhury, "Weakly Supervised Video Moment Retrieval from Text Queries", in *IEEE Conf. on Computer Vision & Pattern Recognition (**CVPR**)*, 2019.

6. **N. C. Mithun**, R. Panda, A. K. Roy-Chowdhury, "Construction of Diverse Image Datasets from Web Collections with Limited Labeling", *IEEE Trans. Circuits and Systems for Video Technology (**TCSVT**)*, 2019.

7. **N. C. Mithun**, J. B. Li, Florian Metze, A. K. Roy-Chowdhury, "Joint Embedding with Multimodal Cues for Video-Text Retrieval", in *International Journal Multimedia Information Retrieval (**IJMIR**)*, 2019.

8. **N. C. Mithun**, R. Panda, E. Papalexakis, A. K. Roy-Chowdhury, "Webly Supervised Joint Embedding for Cross-Modal Image-Text Retrieval", in *ACM Multimedia Conference (**ACM MM**)*, 2018.

9. **N. C. Mithun**, J. B. Li, F. Metze, A. K. Roy-Chowdhury, "Learning Joint Embedding with Multimodal Cues for Cross-Modal Video-Text Retrieval", in *ACM Int. Conf. Multimedia Retrieval (**ICMR**)*, 2018. [**Best Paper**]

10. **N. C. Mithun**, S. Munir, K. Guo, C. Shelton, "ODDS: Real-Time Object Detection using Depth Sensors on Embedded GPUs", in *ACM/IEEE Conf. on Information Processing in Sensor Networks (**IPSN**)*, 2018.

11. **N. C. Mithun**, C. Simons, R. Casey, S. Hilligardt, A. K. Roy-Chowdhury, "Learning Long-Term Invariant features for Vision-based Localization", in *IEEE Winter Conf. on Computer Vision (**WACV**)*, 2018.

12. R. Panda, **N. C. Mithun**, A. K. Roy-Chowdhury, "Diversity Aware Multi-Video Summarization", *IEEE Trans. Image Processing (**TIP**)*, 2017.

13. **N. C. Mithun**, R. Panda, and A. K. Roy-Chowdhury, "Generating Diverse Image Datasets with Limited Labeling", in *ACM Multimedia Conference (**ACM MM**)*, 2016.

14. **N. C. Mithun**, N. U. Rashid, S. M. M. Rahman, "Detection and Classification of vehicles from video using multiple time-spatial images", in *IEEE Trans. Intelligent Transportation Systems (**TITS**)*, 2012.

Preprints

1. J. Tian, **N. C. Mithun**, Z. Seymour, H. Chiu, Z. Kira, Striking the Right Balance: Recall Loss for Semantic Segmentation", arXiv preprint arXiv:2106.14917 , 2021.

2. M. Irshad, **N. C. Mithun**, Z. Seymour, H. Chiu, S. Samarasekera, R. Kumar, SASRA: Semantically-aware Spatio-temporal Reasoning Agent for Vision-and-Language Navigation in Continuous Environments, arXiv preprint arXiv: 2108.11945, 2021.

Patent Applications

1. **N. C. Mithun**, S. Munir, C. Shelton, "Real-Time Object Detection using Depth Sensors", Int. Patent Pub. No: WO 2019/162241 A1, Pub. Date: August 29, 2019

2. H. P. Chiu, Z. Seymour, K. Sikka, S. Samarasekera, R. Kumar, **N. C. Mithun**, "Semantically-Aware Image-based Visual Localization", US Patent Pub. No: US 2020/0357143 A1, Pub. Date: Nov. 12, 2020

3. H. P. Chiu, Z. Seymour, **N. C. Mithun**, S. Samarasekera, R. Kumar, Y. Yao, "Physics-Guided Deep Multimodal Embeddings for Task-Specific Data Explorations", Int. Patent Pub. No: WO 2021/183256 A1, Sept. 16, 2021

SELECTED TALKS

· **Joint Embedding with Multimodal Cues for Cross-Modal Video-Text Retrieval**, *Best Paper Presentation* in ACM International Conf. on Multimedia Retrieval, June, 2018, Yokohama, Japan.

· **RGB2LIDAR: Towards Solving Large-Scale Cross-Modal Visual Localization**, *Best Paper Session Presentation* in ACM Multimedia Conference, October, 2020, Seattle, U.S.

· **Embodied Vision-Language Navigation with Semantic Spatio-Temporal Reasoning**, *Winning Funding Proposal Presentation* in SRI CVT SharkTank 2019, Dec., 2019, Princeton, NJ.

· **Learning Multimedia Retrieval Models with Limited Labeled Data**, *Invited Talk in The College of New Jersey (TCNJ) Computer Science Colloquium Series- Fall 2020*, Oct 2020, Ewing, NJ

· **Learning Visual Recognition Models with Limited Supervision**, *Seminar Talk* in Center for Vision Technology, SRI International, September, 2018, Princeton, NJ, U.S.

· **Object Detection from Depth Sensors on Embedded GPUs**, *Seminar Talk* in Bosch Research Center Pittsburgh, September, 2017, Pittsburgh, PA, U.S.

· **Deep Learning for Computer Vision**, *Guest Lecture* in Course EE243: Advanced Computer Vision, UC Riverside, April, 2018, Riverside, CA, U.S.

· **Learning Long-Term Invariant features for Vision-based Localization**, *Oral Presentation* in IEEE Winter Conf. on Computer Vision, March, 2018, Lake Tahoe, NV/CA, U.S.

· **Webly Supervised Image-Text Embedding with Noisy Tag Refinement**, *Oral Presentation* in International Conf. on Pattern Recognition, Jan 2021.

HONORS AND AWARDS

· **Best Paper Award**, ACM International Conference of Multimedia Retrieval (2018).

· **Winner of SRI CVT SharkTank** Awarded SRI Internal Investment for winning CVT SharkTank. (2020)

· **Best Paper Nomination/Candidate** (Top 4 out of 1698 paper), ACM Multimedia Conference. (2020)

· **Dean's Distinguished Fellowship**, University of California, Riverside. (2014)

· **63rd rank in Admission Test** of Bangladesh University of Engineering & Technology (2005)

· **Deans List Scholarship** in Bangladesh University of Engineering and Technology (2007-2010)

· **Merit Scholarship** from Dept. of EEE, Bangladesh University of Engineering and Technology (2008 & 2010)

· **Doctoral Consortium Award**, IEEE Winter Conference on Applications of Computer 2018

· **Student Travel Grant**, ACM Multimedia Conference 2018

· **Bangladesh Government Education Board Merit Scholarship** in Higher Secondary, Secondary and Junior level (2000-2003; 2003-2005 & 2006-2009 consecutively)

PROFESSIONAL SERVICES

· **Member:** IEEE, ACM, CVF

· **Journal Reviewer:** IEEE Trans. Pattern Analysis and Machine Intelligence, IEEE Trans. Multimedia, IEEE Trans. Circuits and Systems for Video Technology, IEEE Trans. Intelligent Transportation Systems, Pattern Recognition Journal, Pattern Recognition Letters, IEEE Access, Multimedia Systems Journal, ACM Trans. Asian and Low-Resource Language Information Processing, Journal of Electronic Imaging

· **Conference Reviewer:** IEEE/CVF Conf. Computer Vision and Pattern Recognition, ACM International Conf. Multimedia, International Conf. Computer Vision, European Conf. Computer Vision, AAAI Conference, IEEE Winter Conf. Applications of Computer Vision, IEEE International Conf. Image Processing