Dear Search Committee,

My name is Jingxiao Liu, a Postdoctoral Researcher in the Senseable City Lab at the Department of Urban Studies and Planning at Massachusetts Institute of Technology (MIT). I received my Ph.D. degree in Civil and Environmental Engineering with a Ph.D. Minor in Electrical Engineering from Stanford University. I am submitting this letter to express my interest in the tenure-track faculty position at Purdue University.

The main objective of my research is to enable scalable monitoring of civil infrastructure through non-dedicated sensing and physics-informed learning. Under the evolving conditions induced by climate change, the need for a scalable infrastructure monitoring solution has become increasingly critical to ensure the long-term resilience of infrastructure. Specifically, I have been using pre-existing vehicles and telecommunication fiber-optic cables as non-dedicated sensors. I investigate the interactions between these sensors and the infrastructure systems to extract structural condition information. In collaboration with the Port Authority of Allegheny County, City of San Jose, and City of Palo Alto, I have conducted field experiments to monitor bridge health and roadway traffic using multi-modal sensing systems including vehicle vibration sensing and distributed fiber-optic sensing. My work has been published in top-tier conferences and journals in both civil and electrical engineering. Furthermore, I have received multiple awards and fellowships, including 6 best paper awards.

In the future, I plan to assess the impact of future climate on civil infrastructure, incorporating climate data into the monitoring framework to ensure long-term infrastructure resilience. Moreover, I aim to advance scalable infrastructure monitoring by developing multi-modal crowd-sensing and decentralized learning approaches. These approaches will handle heterogeneous data from diverse sensing platforms and ensure data privacy. Additionally, I plan to extend my research framework to various urban monitoring domains (e.g., mobility and air pollution), addressing complex interactions between non-dedicated sensors and sensing targets, while providing explainable data-driven models for decision-making in dynamic sensing environments.

As practical utilization of infrastructure monitoring research is of great necessity, I will seek new collaborations with Indiana Department of Transportation, Greater Lafayette Public Transportation Corporation, and City of West Lafayette to conduct experiments with real-world infrastructures. Some of the current/potential funding agencies for my research include the National Science Foundation (e.g., CIS, CPS, and CMMI programs), the Department of Transportation, the Federal Highway Administration, the Federal Railroad Administration, NVIDIA, Google, and Bentley Systems.

In addition, one of the key reasons I am determined to pursue a career in academia is my enthusiasm to learn and share what I have learned with students. My teaching philosophy is to encourage students to learn from observation and experimentation with my progressive assessment and adaptive teaching styles, thus building students' investigating and problem-solving skills. I am capable of teaching a wide variety of courses, including general engineering courses (e.g., statics and dynamics), core structural engineering courses (e.g., structural stability, structural analysis, and finite element analysis), and advanced courses in smart systems such as sensing, signal processing, data mining, and machine learning.

Last but not least, I plan to continue promoting and encouraging diverse, equal, and inclusive academic environments as an academic instructor and mentor in my future career. A few examples of specific actions I would take include seeking opportunities to actively recruit and retain minority students, developing hands-on summer courses and outreach events for K-12 students, and designing course materials that are approachable to students with diverse background.

Thank you for considering my application. I am attaching my CV, research statement, teaching statement, and diversity statement. I look forward to hearing from you.

Sincerely,
Jingxiao Liu

Jingxiao Liu

Postdoctoral Fellow at MIT Senseable City Lab
Last Updated 2023-10-30,   jingxiao@mit.edu

RESEARCH AREAS	• Scalable Structural Health Monitoring Using Vehicles as Non-Dedicated Vibration Sensors • Urban Monitoring Using Telecommunication Fiber Cables as Non-Dedicated Strain Sensors • Physics-guided Machine Learning and Signal Processing • Dynamics of Vehicle-Structure Interaction Systems

EDUCATION

Stanford University, Stanford, California, USA

Ph.D. in Civil & Environmental Engineering   June 2023
- Cumulative GPA: 4.0 / 4.0
- Advisor: Prof. Hae Young Noh
- Thesis title: Accurate and Scalable Bridge Health Monitoring Using Drive-by Vehicle Vibrations

Ph.D. Minor in Electrical Engineering   June 2023
- Cumulative GPA: 4.0 / 4.0

Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
Master of Science in Civil Engineering   Dec 2017
- Cumulative GPA: 3.9 / 4.0

Central South University, Changsha, China
Bachelor of Engineering in Civil Engineering   Jun 2016
- Cumulative GPA: 89.6 / 100.0, Valedictorian

APPOINTMENTS

Postdoctoral Fellow,   Sep 2023 - Present
Senseable City Lab, Dept. of Urban Studies and Planning, MIT, MA, USA

Research Affiliate,   Sep 2023 - Sep 2023
Dept. of Geophysics and Dept. of Civil & Envir. Engineering,
Stanford University, Stanford, CA, USA

Research Affiliate,   Jul 2023 – Sep 2023
Senseable City Lab, Dept. of Urban Studies and Planning, MIT, MA, USA

Postdoctoral Scholar,   Jun 2023 – Aug 2023
Dept. of Geophysics and Dept. of Civil & Envir. Engineering,
Stanford University, Stanford, CA, USA

Graduate Research Assistant,   Jan 2020 – Jun 2023
Dept. of Civil and Envir. Engineering, Stanford University, Stanford, CA, USA

Research Intern - IoT system, data science, fiber-optic sensing,   Jun 2021 – Sep 2021
PARC, a Xerox company, Palo Alto, CA, USA

Data Analyst,   Jun 2021 – Sep 2021
Eloque, Remote, Joint with the research intern role at PARC, a Xerox company.

Graduate Research Assistant,   Jan 2018 – Dec 2019
Dept. of Civil and Envir. Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Summer Research Intern,   Jun 2017 – Sep 2017
Dept. of Civil and Envir. Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

ACADEMIC HONORS & AWARDS

Best Journal Paper Award,   2023
ASME SHM/NDE Committee

The Gary Marsden Travel Awards for attending ACM Ubicomp 2023,   2023
ACM, Special Interest Group on Computer-Human Interaction (SIGCHI)

The First-Place Award, Student Paper Competition, 2022
Structural Health Monitoring and Control Committee, ASCE Engineering Mechanics Institute

The Second-Place Award, Student Project Competition, 2022
Second International Competition for Structural Health Monitoring, Asia-Pacific Network of Centers for Research in Smart Structures Technology (ANCRiSST)

The Third-Place Award, Student Paper Competition, 2021
Structural Health Monitoring and Control Committee, ASCE Engineering Mechanics Institute

Best Journal Paper Award (Runner-Up), 2021
ASME SHM/NDE Committee

Leavell Fellowship on Sustainable Built Environment (Merit-based) 2020
Dept. of Civil and Envir. Engineering, Stanford University

Best Presentation Award (Runner-Up), 2020
BuildSys/SenSys Joint PhD Forum, ACM BuildSys/Sensys Conference

Best Student Paper Award, 2020
Second Nurse Care Activity Recognition Challenge, ACM International Joint Conference on Pervasive and Ubiquitous Computing

Fenves Travel Grant for Attending IWSHM Conference in Stanford, CA 2019
Carnegie Mellon University

Fenves Travel Grant for Attending SPIE Conference in Denver, CO 2019
Carnegie Mellon University

Dean's Fellowship (support 1-year PhD study) 2018
College of Engineering, Carnegie Mellon University

Valedictorian of Dept. of Civil Engineering, 2016
Central South University

Outstanding Undergraduate Thesis, 2016
Central South University

Mao Yi-sheng Science and Technology Award - Star of Hope, 2015
Mao Yi-sheng Science and Technology Education Foundation

RESEARCH EXPERIENCES	**Urban System Monitoring Using Combined Vehicle On-Board Sensing (VOS) and Roadside Distributed Acoustic Sensing (DAS)** Jun 2020 – present

Stanford University, Stanford, California, USA
Research Advisor: Professor Hae Young Noh & Professor Biondo Biondi
- Exploring the value of combined VOS and DAS to achieve a cost-effective urban infrastructure monitoring.
- Achieved car position and speed estimation, underground fiber localization, and traffic-induced surface wave characterization using combined VOS and DAS.
- Collaborated with the City of Palo Alto, and the City of San Jose to demonstrate the capability of using DAS for bridge monitoring at low cost and with low maintenance.

Drive-by Bridge Health Monitoring Jan 2020 – present
Stanford University, Stanford, California, USA
Research Advisor: Professor Hae Young Noh
- Introducing signal processing and machine learning approaches to detect, localize, and quantify damage on multiple bridges using vehicle vibration responses.
- Demonstrated the feasibility of the drive-by bridge health monitoring approach through numerical simulations, lab-scale and full-scale experiments.
- Published our work on drive-by bridge health monitoring in both top-tier conferences and journals in civil engineering and electrical engineering.

Railroad Track Monitoring Using Train Vibrations   Jun 2017 – Dec 2019
Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
Research Advisors: Professor Hae Young Noh & Professor Mario Bergés
  • Introduced an anomaly detection algorithm for monitoring railroad track using train vibration responses.
  • Collaborated with Port Authority of Allegheny County to develop a train-based structural health monitoring system to help monitor railroad tracks of a 42.2-km light rail network in Pittsburgh.
  • Collected and published a comprehensive dataset from light rail vehicles, which is the first open-access dataset for vehicle-vibration-based structural health monitoring.

Study of Australia Traffic for Bridge Loading Applications   Oct 2015 – Jul 2016
Monash University, Melbourne, Victoria, Australia
Research Advisor: A/Professor Colin Caprani
  • Developed a traffic characteristic model by analyzing Weigh-in-Motion data to classify vehicles based on the physical nature of the traffic configurations.
  • Collaborated with the Roads Corporation of Victoria (VicRoads).
  • Received the outstanding student thesis award.

PROFESSIONAL EXPERIENCES

PARC, a Xerox Company, Palo Alto, California, USA   Jun 2021 – Sep 2021
Research Intern in Data Science, IoT system, Fiber-optic sensing
  • Contributed to a systematic algorithm and hardware development of a fiber-optic sensing system for civil infrastructure monitoring. The system is successfully deployed in Melbourne, Australia for superload monitoring with Department of Transportation.
  • Developed signal processing, statistical, and computer vision algorithms for utilizing and validating Fiber Bragg Grating sensors in bridge monitoring and traffic characterization applications.

PUBLICATIONS

JOURNAL PAPERS

[J11] Jingxiao Liu, Siyuan Yuan, Yiwen Dong, Biondo Biondi, Hae Young Noh, "TelecomTM: A Fine-grained and Ubiquitous Traffic Monitoring System Using Pre-Existing Telecommunication Cables as Sensors," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 7, 2, Article 64 (June 2023)*.
Received "The Gary Marsden Travel Awards" for attending Ubicomp 2023 from ACM Special Interest Group on Computer-Human Interaction (SIGCHI).

[J10] Siyuan Yuan, Jingxiao Liu, Hae Young Noh, Robert Clapp, Biondo Biondi, "Using Vehicle-induced DAS Signals for Near-surface Characterization with High Spatiotemporal Resolution," *Journal of Geophysical Research: Solid Earth (Under Review)*. DOI: 10.22541/essoar.169755571.19783911/v1

[J9] Jingxiao Liu, Susu Xu, Mario Bergés, Hae Young Noh, "HierMUD: Hierarchical Multi-task Unsupervised Domain Adaptation between Bridges for Drive-by Damage Diagnosis," *Structural health monitoring, (2022), https://doi.org/10.1177/1475921722108115*.
Received "Best Journal Paper Award" from SHM/NDE Committee, ASME.

[J8] Jingxiao Liu, Siyuan Yuan, Bin Luo, Biondo Biondi, Hae Young Noh, "Turning optical telecommunication cables into distributed acoustic sensors for vibration-based bridge health monitoring," *Structural Control and Health Monitoring, vol. 2023, Article ID 3902306, 14 pages, 2023*.

[J7] Yiwen Dong, Megan Iammarino, Jingxiao Liu, Jesse Codling, Jonathon Fagert, Mostafa Mirshekari, Linda Lowes, Pei Zhang, Hae Young Noh, "Ambient Floor Vibration Sensing Advances Accessibility of Functional Gait Assessment for Children with Muscular Dystrophies," *Submitted to Scientific Report (Under Review)*.

[J6] Daniel Cantero, Muhammad Zohaib Sarwar, Abdollah Malekjafarian, Robert Corbally, Mehrisadat Alamdari, Prasad Cheema, Hae Young Noh, Jingxiao Liu, Jatin Aggarwal, "Numerical benchmark for road bridge damage detection from passing vehicles responses applied to four data-driven methods," *Submitted to Measurements (Under review)*.

[J5] Jingxiao Liu, Yujie Wei, Bingqing Chen, Hae Young Noh, "A Hierarchical Semantic Segmentation Framework for Computer Vision-based Bridge Damage Detection," *Smart Structures and Systems, 31(4) (2023), pp.325-334.*
Received "the Second-Place Award" in the Second International Competition for Structural Health Monitoring, 2022, and invited to submit to a special issue in the Journal of Smart Structures and Systems.

[J4] Siyuan Yuan, Martijn van den Ende, Jingxiao Liu, Hae Young Noh, Robert Clapp, Cédric Richard, Biondo Biondi, "Spatial Deep Deconvolution U-Net for Traffic Analyses with Distributed Acoustic Sensing," in *IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3322355..*

[J3] Ankit Shrivastava, Jingxiao Liu, Kaushik, Dayal, Hae Young Noh, "Predicting peak stresses in microstructured materials using convolutional encoder-decoder learning," *Mathematics and Mechanics of Solids, (2022), 10812865211055504.*

[J2] Jingxiao Liu, Siheng Chen, Mario Bergés, Jacobo Bielak, James H Garrett, Jelena Kovačević, Hae Young Noh, "Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction," *Mechanical Systems and Signal Processing 136 (2020): 106454.*
Received "Best Journal Paper Award (Runner-up)" from SHM/NDE Committee, ASME.

[J1] Jingxiao Liu, Siheng Chen, George Lederman, David B Kramer, Hae Young Noh, Jacobo Bielak, James H Garrett, Jelena Kovačević, Mario Bergés, "Dynamic responses, GPS positions and environmental conditions of two light rail vehicles in Pittsburgh," *Scientific data, 6(1) (2019), pp.1-11.*

ARCHIVAL PEER-REVIEWED CONFERENCE PAPERS

[AC2] Jingxiao Liu, Bingqing Chen, Siheng Chen, Mario Bergés, Jacobo Bielak, HaeYoung Noh, "Damage-sensitive and domain-invariant feature extraction for vehicle-vibration-based bridge health monitoring," In *IEEE 45th International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2020*

[AC1] Bingqing Chen, Jingxiao Liu, Henning Lange, Mario Bergés, "Dyna-BOLT: Domain adaptative binary factorization of current waveforms for energy disaggregation," In *IEEE 45th International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2020*

OTHER CONFERENCE AND WORKSHOP PAPERS

[C9] Yiwen Dong, Jingxiao Liu, Hae Young Noh, "GaitVibe+: Enhancing Structural Vibration-based Footstep Localization Using Temporary Cameras for In-home Gait Analysis," *Accepted in the Fourth Workshop on Continual and Multimodal Learning for Internet of Things, co-located with SenSys 2022.*

[C8] Jingxiao Liu, Siyuan Yuan, Biondo Biondi, Hae Young Noh, "Vibration-Based Bridge Health Monitoring using Telecommunication Cables," *Accepted In the 8th World Conference on Structural Control and Monitoring, 2022.*

[C7] Jingxiao Liu, Susu Xu, Mario Bergés, Hae Young Noh, "A Hierarchical Domain-Adversarial and Multi-Task Learning Algorithm for Bridge Damage Diagnosis Using a Drive-by Vehicle," *International Workshop on Structural Health Monitoring, IWSHM 2021.*

[C6] Siyuan Yuan, Jingxiao Liu, Hae Young Noh, Biondo Biondi, "Urban system monitoring using combined vehicle onboard sensing and roadside distributed acoustic sensing," In *First International Meeting for Applied Geoscience & Energy, pp. 3235-3239. Society of Exploration Geophysicists, 2021.*

[C5] Yiwen Dong, Joanna Jiaqi Zou, Jingxiao Liu, Jonathon Fagert, Mostafa Mirshekari, Linda Lowes, Megan Iammarino, Pei Zhang, and Hae Young Noh, "MD-Vibe: physics-informed analysis of patient-induced structural vibration data for monitoring gait health in individuals with muscular dystrophy.," In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, UbiComp-ISWC '20.*

[C4] Yiwen Dong, Jingxiao Liu, Yitao Gao, Sulagna Sarkar, Zhizhang Hu, Jonathon Fagert, Shijia Pan, Pei Zhang, Hae Young Noh, and Mostafa Mirshekari, "A window-based sequence-to-one approach with dynamic voting for nurse care activity recognition using acceleration-based wearable sensor.," In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, UbiComp-ISWC '20.*

[C3] Jingxiao Liu, Susu Xu, Mario Bergés, Jacobo Bielak, James H. Garrett, Hae Young Noh, "An Expectation-maximization Algorithm-based Framework for Vehicle-Vibration-Based Indirect Structural Health Monitoring of Bridges," *International Workshop on Structural Health Monitoring, IWSHM 2019.*

[C2] Jingxiao Liu, Yujie Wei, Mario Bergés, Jacobo Bielak, James H Garrett Jr, HaeYoung Noh, "Detecting anomalies in longitudinal elevation of track geometry using train dynamic responses via a variational autoencoder," *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019.*

[C1] Jingxiao Liu, Mario Bergés, Jacobo Bielak, James H Garrett, Jelena Kovačević, Hae Young Noh, "A damage localization and quantification algorithm for indirect structural health monitoring of bridges using multi-task learning," In *AIP Conference Proceedings (Vol. 2102, No. 1, p. 090003). AIP Publishing LLC. QNDE 2018.*
Received "Best Paper Award" in the Second Nurse Care Activity Recognition Challenge (part of HASCA Workshop, ACM UbiComp, 2020)

POSTERS AND EXTENDED ABSTRACTS

[A6] Jingxiao Liu, Siyuan Yuan, Biondo Biondi, and Hae Young Noh, "Turning Telecommunication Cables into Distributed Acoustic Sensors for Bridge Health Monitoring," In *Engineering Mechanical Institute, ASCE, 2022.*

[A5] Jingxiao Liu, Susu Xu, Mario Bergés, and Hae Young Noh, "Hierarchical Model Transfer Between Bridges for Multi-Task Damage Diagnosis Using Drive-by Vehicles," In *Engineering Mechanical Institute, ASCE, 2021.*

[A4] Jingxiao Liu, "Scalable bridge health monitoring using drive-by vehicles: PhD forum abstract," In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems, Sensys 2020.*
Received "Best Presentation Award (Runner-up)" from BuildSys/SenSys Joint PhD Forum, ACM, 2020.

[A3] Jingxiao Liu, M. Bergés, J. Bielak, J. Garrett, J. Kovačević, and H. Noh, "A damage localization and quantification algorithm for indirect structural health monitoring of bridges using multi-task learning," in *Engineering Mechanical Institute, ASCE, 2018.*

[A2] Jingxiao Liu, Siheng Chen, Mario Bergés, Jacobo Bielak, James H Garrett, Jelena Kovačević, Hae Young Noh, "A Damage Localization and Quantification Algorithm for Indirect Structural Health Monitoring of Bridges Using Multi-Task Learning," in *Machine Learning in Science and Engineering, Pittsburgh, PA, 2018.*

[A1] Jingxiao Liu, Mario Bergés, Jacobo Bielak, James H Garrett, Jelena Kovačević, Hae Young Noh, "Poster Presentation: Damage Diagnosis Algorithms for Indirect Structural Health Monitoring of Bridges," in *PIANC-SMART Rivers Conference, Pittsburgh, PA, 2017.*

OPEN-ACCESS DATASET

[D3] <u>Jingxiao Liu</u>, Siheng Chen, George Lederman, David B Kramer, Hae Young Noh, Jacobo Bielak, James H Garrett, Jelena Kovačević, Mario Bergés, "The DR-Train dataset: dynamic responses, GPS positions and environmental conditions of two light rail vehicles in Pittsburgh," *Zenodo 2020, https://doi.org/10.5281/zenodo.1432702*

[D2] Yiwen Dong, Megan Iammarino, <u>Jingxiao Liu</u>, Jesse Codling, Jonathon Fagert, Mostafa Mirshekari, Linda Lowes, Pei Zhang, Hae Young Noh, "The MD-Vibe Dataset: Footstep-Induced Floor Vibration Data for Functional Gait Assessment of Individuals with Muscular Dystrophy," *Zenodo 2023, Zenodo. https://doi.org/10.5281/zenodo.8125744*

[D1] Xinlei Chen, Xinyu Liu, Kent X. Eng, <u>Jingxiao Liu</u>, Hae Young Noh, Lin Zhang, Pei Zhang, "The S&M-HSTPM2d5 dataset: High Spatial-Temporal Resolution PM 2.5 Measures in Multiple Cities Sensed by Static & Mobile Devices," *Zenodo 2020, http://doi.org/10.5281/zenodo.4028130*

PRESENTATIONS AND TALKS

DEPARTMENT SEMINAR

[S2] "Scalable civil infrastructure monitoring through non-dedicated sensing and physics-informed learning," at *Dept. of Civil and Environmental Engineering, stony brook university,* Nov 2022.

[S1] "Taking the pulse of the city via non-dedicated sensing of ambient vibrations," at *Dept. of Electrical Engineering and Computer Science, University of California Merced*, Sept 2022.

INVITED TALKS

[T9] "Scalable civil infrastructure monitoring through non-dedicated sensing and physics-informed learning," at *Dept. of Civil and Environmental Engineering, University of Vermont*, Feb 2023.

[T8] "DAS Monitoring of Bridges and Urban Traffic," at *Distributed Acoustic Sensing (DAS) Research Coordination Network (RCN), IRIS Earthquake Science Presentations,* Dec 2022.

[T7] "Scalable civil infrastructure monitoring through non-dedicated sensing and physics-informed learning," at *Centre for Advanced Materials and Structures, Dept. of Civil and Environmental Engineering, National University of Singapore,* Dec 2022.

[T6] "Scalable civil infrastructure monitoring through non-dedicated sensing and physics-informed learning," at *Senseable City Lab, MIT*, Nov 2022.

[T5] "Scalable civil infrastructure monitoring through non-dedicated sensing and physics-informed learning," at *Dept. of Civil and Environmental Engineering, Tufts University*, Nov 2022.

[T4] "Turning Telecommunication Fiber-cables into Distributed Acoustic Sensors for Vibration-based Bridge Health Monitoring," at *Stanford Exploration Project (SEP) group, Dept. of Geophysics, Stanford University*, May 2022.

[T3] "Accurate and Scalable Bridge Health Monitoring Using Drive-by Vehicle Vibrations," at *Advanced Infrastructure Systems group, Dept. of Civil and Envir. Engineering, Carnegie Mellon University*, March 2022.

[T2] "Scalable Bridge Health Monitoring Using Drive-by Vehicles," at *Eloque & PARC, a Xerox company*, Summer 2021.

[T1] "Diagnosis Algorithms for Indirect Structural Health Monitoring of a Bridge Model via Dimensionality Reduction," at *Dept. of Civil and Envir. Engineering, Chongqing University*, China, Summer 2019.

CONFERENCE PRESENTATIONS

[P11] Jingxiao Liu, Siyuan Yuan, Biondo Biondi, and Hae Young Noh, "Integrating Pre-Existing Telecommunication Fiber Cable Vibration Sensing and Drive-by Vehicle Vibration Sensing for Scalable Bridge Health Monitoring," in *International Workshop on Structural Health Monitoring*, 2023.

[P10] Jingxiao Liu, Siyuan Yuan, Yiwen Dong, Biondo Biondi, and Hae Young Noh, "TelecomTM: A Fine-Grained and Ubiquitous Traffic Monitoring System Using Pre-Existing Telecommunication Fiber-Optic Cables as Sensors," in *Engineering Mechanical Institute, ASCE*, 2023.

[P9] Jingxiao Liu, Siyuan Yuan, Biondo Biondi, and Hae Young Noh, "Vibration-Based Bridge Health Monitoring using Telecommunication Cables," in *8th World Conference on Structural Control and Monitoring*, 2022.

[P8] Jingxiao Liu, Siyuan Yuan, Biondo Biondi, and Hae Young Noh, "Turning Telecommunication Cables into Distributed Acoustic Sensors for Bridge Health Monitoring," in *Engineering Mechanical Institute, ASCE*, 2022.

[P7] Jingxiao Liu, Susu Xu, Mario Bergés, and Hae Young Noh, "Hierarchical Model Transfer Between Bridges for Multi-Task Damage Diagnosis Using Drive-by Vehicles," in *Engineering Mechanical Institute, ASCE*, 2021.

[P6] Jingxiao Liu and Hae Young Noh, "Scalable Bridge Health Monitoring Using Drive-by Vehicles," in *BuildSys/SenSys Doctoral Colloquium, ACM*, 2020.

[P5] Jingxiao Liu, Mario Bergés, Jacobo Bielak, and Hae Young Noh, "Damage-Sensitive and Domain-Invariant Feature Extraction for Bridge Health Monitoring Using Vehicle Vibration Responses," in *Engineering Mechanical Institude, ASCE*, 2020.

[P4] Jingxiao Liu, Susu Xu, Mario Bergés, Jacobo Bielak, James H Garrett, and Hae Young Noh, "An Expectation-maximization Algorithm-Based Framework for Vehicle-Vibration-Based Indirect Structural Health Monitoring of Bridges," in *International Workshop on Structural Health Monitoring*, 2019.

[P3] Jingxiao Liu, Yujie Wei, Mario Bergés, Jacobo Bielak, James H Garrett, and Hae Young Noh, "Detecting Anomalies in Longitudinal Elevation of Track Geometry Using Train Dynamic Responses via a Variational Autoencoder," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, SPIE*, 2019.

[P2] Jingxiao Liu, M. Bergés, J. Bielak, J. Garrett, J. Kovačević, and H. Noh, "A damage localization and quantification algorithm for indirect structural health monitoring of bridges using multi-task learning," in *Engineering Mechanical Institute, ASCE*, 2018.

[P1] Jingxiao Liu, M. Bergés, J. Bielak, J. Garrett, J. Kovačević, and H. Noh, "A damage localization and quantification algorithm for indirect structural health monitoring of bridges using multi-task learning," in *Annual Review of Progress in Quantitative Nondestructive Evaluation (QNDE), ASME*, 2018.

TEACHING EXPERIENCES	**Stanford University**, Stanford, CA, USA

Teaching Assistant   Winter 2023
Course Title: Structural Monitoring (CEE286)
Teaching Faculty: Professor Haeyoung Noh

Head Teaching Assistant   Fall 2020, Fall 2022
Course Title: Data Analytics for Physical Systems (CEE154/254)
Teaching Faculty: Professor Haeyoung Noh
  • Helped design the course, including lectures, homework, and project.
  • Held MATLAB tutorial lecture.
  • Held weekly office hours to help students develop a better understanding of the course materials.

Guest Lecturer   Winter 2023
  • Lecture Topic: Scalable Civil Infrastructure Monitoring through Non-Dedicated Sensing and Physics-Informed Learning (part of CEE286: Structural monitoring)

Carnegie Mellon University, Pittsburgh, PA, USA
Teaching Assistant   Win 2018, Win 2019
Course Title: Building Information Modeling (BIM) for Engineering, Construction and Facility Management (12-711)
Teaching Faculty: Professor Xuesong Liu
  • Ran 2-4 labs to assist students to build BIM models and explore contemporary BIM topics.
  • Held weekly office hours to help students develop a better understanding of the state-of-the-art BIM-based software systems that are being used during design and construction.

STUDENTS MENTORED	**Doyun Hwang**, Graduate Student, *Civil and Env. Engineering, Stanford University*	2023
	Peiyao Xu, Graduate Student, *Civil and Env. Engineering, Stanford University*	2023
	Jatin Aggarwal, Graduate Student, *Civil and Env. Engineering, Stanford University*	2022
	Kent Eng, Graduate Student, *Civil and Env. Engineering, Stanford University*	2021
	Aaron Appelle, Graduate Student, *Civil and Env. Engineering, Stanford University*	2020
	Yiwen Dong, Graduate Students, *Civil and Env. Engineering, Stanford University*	2020
	Joanna Zou, Graduate Students, *Civil and Env. Engineering, Stanford University*	2020

SERVICES

Workshop Chair
  • The Fourth Workshop on Continual and Multimodal Learning for Internet of Things (CML-IoT), co-located with ACM SenSys, 2022

Session Chair
  • Engineering Mechanics Institute Conference, ASCE, 2023
  • Engineering Mechanics Institute Conference, ASCE, 2022
  • International Conference on Information Processing in Sensor Networks (ACM, IPSN), 2022
  • CML-IoT workshop, co-located with ACM Ubicomp, 2020

Technical Program Committee
  • IASA 2023: 2nd ACM International Workshop on Intelligent Acoustic Systems and Applications, co-located with CPS-IoT Week 2023
  • Data: Acquisition to Analysis workshop, co-located with ACM SenSys, 2021

Reviewer
  • Structural Health Monitoring, SAGE
  • Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, ASME
  • International Conference on Systems for Energy-Efficient Buildings (ACM BuildSys), 2022
  • Annual Conference of the Prognostics and Health Management (PHM) Society
  • International Conference on Acoustics, Speech, & Signal Processing (IEEE ICASSP)
  • Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (ACM IMWUT)
  • Journal of Engineering Mechanics, ASCE
  • Mechanical Systems and Signal Processing, Elsevier
  • Measurement, Elsevier
  • Journal of Vibration and Control, SAGE
  • Structural Control & Health Monitoring, Wiley
  • Frontiers in Big Data
  • Frontiers in Built Environment
  • Canadian Journal of Civil Engineering

MEDIA

Stanford Woods Institute for the Environment: 2023
"Improving Communities: Stanford researchers design car sensors to track neighborhood quality," https://youtu.be/DD6ik5gD53Q

CBS Mission Unstoppable: 2021
"What are Geophone Sensors?" https://youtu.be/S9OwJi88Dus

CMU News Letter: 2018
"Simulations, Sensors Provide Insight into Health of Transportation Infrastructure," https://shorturl.at/ehzAR

My experiences as a student, a researcher, an instructor, and a mentor have deeply shaped my teaching philosophy that students need to learn from observations and experiments to solve real-world problems, and my teaching style should be adaptive to students from different backgrounds. In this statement, I first describe the elements of my teaching philosophy and then give an overview of how I put my philosophy into practice. Finally, I conclude with courses I am interested in and qualified to teach.

Teaching Philosophy. One thing I have learned through years of studying and researching is that knowledge can be learned better through application and experimentation. I encourage students to apply the knowledge in real-world situations through progressive assessment, from well-defined assignments to open-ended projects and experimentation. The foundation of my teaching philosophy includes four components:

- *Context-first & experiential learning.* Students would learn better if they understood the purpose of what they were learning. I use hands-on projects and experimentation to present new concepts to students so they can better understand how things work and their significance in the real world.
- *Progressive assessment.* The goals of engineering learning evolve from understanding fundamental concepts to implementing solutions in real-world situations as the students gain experience. Therefore, I try to gradually raise the level of difficulty in assessing students (e.g., from assignments to hands-on projects).
- *Encouraging self-guided learning.* The world is fast changing, and course materials may not give students all the solutions for their future careers. Thus, I guide students to resources and apply knowledge they can use to solve problems instead of directly giving them answers.
- *Adaptive teaching.* Teaching activities are open to students with unique learning styles and at various academic levels. Thus, I re-orient the course material and adapt my teaching style based on feedback from students.

Teaching Experience. I had teaching experiences across two institutes. At Carnegie Mellon University, I was a TA for one project-based course: "Building Information Modeling (BIM) for Engineering, Construction and Facility Management" (12-711). Being a TA for 12-711 gave me the opportunity to run lab sessions to assist students in building BIM models and exploring contemporary BIM topics. At Stanford University, I helped the professor develop a new course: "Data Analytics for Physical Systems" (CEE154/254), for both graduate and undergraduate students and was the head TA when the course was first offered. This experience allowed me to go through the development of a new course for students with varying levels of background in civil engineering.

In particular, as the first TA for CEE154/254, I implemented the *Context-first & experiential learning* philosophy when helping the instructor design lecture notes. We demonstrated course concepts through real-world examples. For example, I helped create lecture materials from field experimental data. I also invited students to visit our lab to show them how we measure footstep-induced floor vibrations to infer human gait patterns. Also, I implemented the *Progressive assessment* philosophy by significantly aiding the instructor in developing homework and the course project, with the level of difficulty gradually increasing. Specifically, the first two sets of homework are well-defined assignments to reinforce students' understanding of essential course materials; the third homework and the course project let the students analyze real-world air pollution data with open-ended questions.

Mentoring Experience. My teaching philosophy can be extended in many ways to mentoring students both individually and in groups. I mentored several students during my Ph.D. studies in different fields (e.g., structural and electrical engineering) at all levels (including high school, undergraduate, master's, and Ph.D.) and for various mentoring purposes (e.g., coursework projects and research). I have been a project mentor for the course: "Sensing and Data Mining for Smart Structures and Systems" at Carnegie Mellon University, the course "Mobile & Pervasive Sensing and Computing" at the University of Michigan, and the course "Structural Monitoring" (CEE286) at Stanford University. The mentoring experience in CEE286 resulted in the publication of two research papers in the Ubicomp 2020 Workshops. These papers were co-authored with students from the class, and notably, one of them received the Best Student Paper Award.

Teaching Interests. I'm enthusiastic about the prospect of teaching and passing on my knowledge and experiences to upcoming students and engineers. Based on my strong background in civil and electrical engineering and my highly interdisciplinary Ph.D. research, I am capable of teaching a wide variety of courses, including general engineering courses (e.g., statics and dynamics), core civil/structural engineering courses (e.g., structural stability, structural analysis, seismic analysis, and finite element analysis), and advanced courses in smart systems such as sensing, signal processing, data mining, and machine learning. I also look forward to the opportunities to design/develop new courses, especially project-based and interactive courses, in interdisciplinary areas combining several engineering specialties, such as structural health monitoring and smart city.

In summary, my teaching philosophy is to encourage students to learn from observation and experimentation with my progressive assessment and adaptive teaching styles, thus building students' self-guided and problem-solving skills. I look forward to the opportunity to teach and mentor the next generation.

Scalable Civil Infrastructure Monitoring under Future Climate through Non-Dedicated Sensing and Physics-Informed Learning

Climate change significantly threatens civil infrastructure assets like bridges, railways, and roads. We've seen an 83% increase in climate-related disasters, rising from 3,656 events between 1980 and 1999 to 6,681 in the past two decades. This surge places immense pressure on our extensive inventory of aging infrastructure, posing significant challenges to the economy and public safety. Effective monitoring of various infrastructure assets, including their structural health and surrounding urban environments, is an important step toward making our city safer, enhancing its resilience, and preparing for future climates. For example, detecting early-stage damage in these assets is crucial to prevent severe deterioration and potential collapse. Currently, dedicated sensors such as cameras and vibration sensors have been deployed directly on infrastructure for monitoring. Nevertheless, these dedicated sensors require on-site installation and maintenance, which is inefficient and costly for large-scale infrastructure monitoring.

To this end, my research introduces a scalable civil infrastructure monitoring system based on non-dedicated sensing and physics-informed learning. Specifically, I have been using pre-existing vehicles and telecommunication (telecom) fiber-optic cables as non-dedicated sensors. The interactions between our non-dedicated sensors and infrastructure enable us to indirectly extract infrastructure health information. For example, many modern automotive vehicles and trains come equipped with vibration sensors primarily for monitoring vehicle operational conditions and detecting accidents. Yet, these sensors can also be used to capture ambient vibrations of transportation infrastructure systems, revealing critical structural damage like cracking, spalling, and corrosion in elements such as bridges and railroad tracks. In addition, these non-dedicated sensors are widely dispersed across the city, eliminating the need for additional sensor installation and maintenance. There fore, this approach to civil infrastructure monitoring is not only cost-effective but also highly scalable.

Research Challenges. Although the non-dedicated sensing approaches for infrastructure monitoring have various benefits, real-world implementation faces several challenges:

1) *Complex interactions.* Interactions between the non-dedicated sensors and infrastructure (i.e., sensor-infrastructure interactions), such as vehicle-bridge and train-track interactions, are complex and often non-linear. The analysis of non-dedicated sensing data to extract the desired infrastructure information is challenging because the data have unpredictable noise conditions as well as many uncertainties involved.

2) *Heterogeneous interactions.* The non-dedicated measurements are sensitive to the heterogeneous properties of both the sensors and the infrastructure, such as vehicle suspensions and bridge dimensions. The monitoring model learned for one infrastructure system is hard to generalize to monitor other systems due to the data distribution shift between different interactions.

3) *Incomplete information.* Due to the non-dedicated nature of our sensing methods, infrastructure information that is captured by non-dedicated sensors may be limited or incomplete. For instance, the drive-by vehicle moves across the bridge, resulting in a high spatial resolution during the sensing time, but having limited temporal information at each coordinate on the bridge, compared to fixed sensors.

My Research Framework. To overcome these challenges, I have introduced a physics-informed learning of sensor-infrastructure interaction framework (Figure 1) that **learns**, **generalizes**, and **integrates** infrastructure monitoring models. To address the *complex interaction* challenge, the first module learns the infrastructure monitoring model by informing data-driven models of sensing data with physical properties of the sensor-infrastructure interaction formulation (e.g., non-linearity of the vehicle-bridge interaction). To address the *heterogeneous interaction* challenge, the second module generalizes the infrastructure monitoring model from one infrastructure to other infrastructure systems by introducing a transfer learning approach that integrates physical constraints of monitoring tasks (e.g., different task difficulties). To address the *incomplete information* challenge, the third module integrates multiple sensing modalities for providing complementary information about the infrastructure.
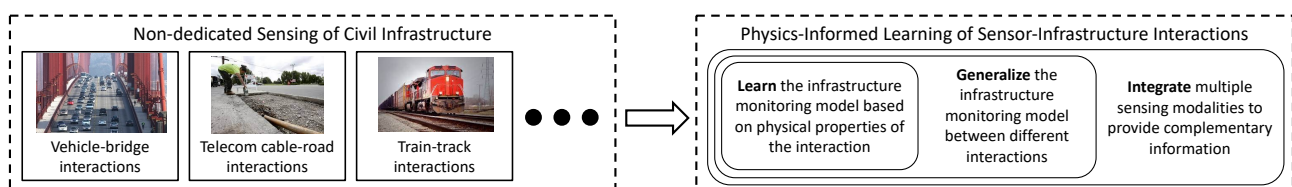

Figure 1: Overview of my research framework. Ambient vibration data of infrastructure captured by non-dedicated sensing systems is input to the Physics-informed Learning of Sensor-Infrastructure Interaction framework.

In collaboration with Port Authority of Allegheny County, City of San Jose, and City of Palo Alto, I have conducted real-world deployments on multiple infrastructure systems, including multiple bridges and a 42.2-km light rail system for 6 years. Moreover, I have applied my framework to other civil infrastructure monitoring problems, including track geometry monitoring, traffic monitoring, and underground telecom infrastructure localization using vehicle and telecom fiber responses. My work has been published in both top-tier conferences and journals in civil engineering and electrical and computer engineering. I have received various best paper and presentation awards from ASCE, ASME, and ACM conferences.

Future Research Plans. I plan to advance the scalable infrastructure monitoring with my non-dedicated sensing and physics-informed learning framework in three aspects:

1) **Multi-Modal Crowd-Sensing and Decentralized Learning of Civil Infrastructure Systems.** An increasing number of urban sensing platforms, such as vehicles, fibers, drones, and smartphones, have been connected to the Internet. They can provide complementary information to enhance the accuracy of infrastructure monitoring. Thus, ensuring scalability for integrating multiple sensing modalities is important. Two primary challenges arise in this process: a) data recorded from different sensing platforms exhibit heterogeneity in terms of physical units, reliability, and data quality, and b) aggregating crowd-sensing data sharing may lead to data privacy issues and network congestion.

   To tackle these challenges, I plan to fuse information from multiple sensing modalities in a decentralized manner, guided by the physical understanding of each sensor-infrastructure interaction. I plan to develop multi-modal crowd-sensing and decentralized learning (e.g., federated learning and gossip learning) approaches that train models collaboratively on crowd-sensing edge devices and upload only the updated model, low-dimensional features, or inference results to a cloud server without exchanging the data. The server then aggregates the updates/results to conduct a feature/decision-level fusion of multi-modal information.

2) **Service Life Assessment under Future Climate.** With climate change intensifying, it is important to assess whether our built civil infrastructure can withstand future climate. My second research direction is to understand the impacts of future climate on our existing civil infrastructure. One main goal is to conduct service life assessments that consider the dynamic interactions between environmental factors and civil infrastructure under future climate.

   To achieve this objective, I plan to incorporate climate data, including temperature fluctuations, precipitation patterns, and the occurrence of extreme weather events, into my monitoring framework. I will evaluate the relationships between climate data and structural responses and create digital twin models for infrastructure systems. These models will incorporate the collected structural response data to predict how infrastructure systems respond to a changing climate. They can further inform proactive maintenance and adaptation strategies, ensuring the long-term resilience of infrastructure in an unpredictable climate.

3) **Expanding Monitoring beyond Civil Infrastructure to the Urban Environment.** In addition to monitoring civil infrastructure systems, I plan to extend my framework to other aspects of urban environments, such as mobility, noise emission, and air pollution. This expansion creates new research challenges. The relationships between non-dedicated sensors and these sensing targets are more complex and indirect than those for infrastructure systems like bridges and tracks. Moreover, these sensing targets involve and are affected by human activities. Therefore, it is important to develop data-driven models that are accurate, as well as explainable and interpretable. This will enable us to gain insights into the underlying reasons governing these complex urban phenomena.

   To this end, I plan to develop explainable data-driven approaches, such as using Causal Inference, to characterize the interactions between sensors and the various sensing targets. These approaches will enable humans to understand the decision-making processes of the data-driven models and provide interpretations of the sensing target conditions.

As practical utilization of infrastructure monitoring research is of great necessity, I will continue collaborating with companies and authorities such as **Palo Alto Research Center, Bosch Center for AI, OptaSense, a Luna Company, City of Palo Alto, City of San Jose**, and **Port Authority of Allegheny County** and seek new collaborations with **Indiana Department of Transportation**, **Greater Lafayette Public Transportation Corporation**, and **City of West Lafayette** to conduct experiments and investigations of my work with real-world infrastructures. Some of the current/potential funding agencies for my research include the National Science Foundation (e.g., CIS, CPS, and CMMI programs), the Department of Transportation, the Federal Highway Administration, the Federal Railroad Administration, NVIDIA, Google, and Bentley Systems.

WILEY | Hindawi

Research Article

Turning Telecommunication Fiber-Optic Cables into Distributed Acoustic Sensors for Vibration-Based Bridge Health Monitoring

Jingxiao Liu [ID],[1] **Siyuan Yuan** [ID],[2] **Bin Luo** [ID],[2] **Biondo Biondi** [ID],[2] **and Hae Young Noh** [ID][1]

[1]*Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, USA*
[2]*Department of Geophysics, Stanford University, Stanford, CA, USA*

Correspondence should be addressed to Jingxiao Liu; liujx@stanford.edu

We introduce a nondedicated bridge health monitoring (BHM) system that turns pre-existing telecommunication fiber-optic cables into distributed acoustic sensors to collect bridge dynamic strain responses. Due to extensively installed telecommunication fiber cables in the cities, our telecommunication cable-based system enables efficient and low-cost BHM without the requirement of on-site sensor installation and maintenance; however, it is challenging to extract bridge damage-sensitive information (e.g., natural frequencies and mode shapes) from this nondedicated strain data as it has large measurement noise and error propagation. To overcome the challenge, we develop a physics-guided system identification method that models strain mode shapes based on physics-guided parametric mode shape functions derived from bridge dynamics. We then estimate the displacement mode shape function by analytically double-integrating the modeled strain mode shape. Our method improves the accuracy of estimating bridge damage-sensitive features and reduces error propagation by constraining strain and displacement mode shapes with bridge dynamics. We evaluated our system on a concrete continuous three-span bridge in San Jose, California. Our system successfully identified the first three natural frequencies and reconstructed strain and displacement mode shapes in a meter-scale resolution.

1. Introduction

Bridges are vital components of transportation infrastructure that link people, roadways, railways, and more. However, around 140,000, or 22% of the more than 617,000 bridges in the United States are considered structurally deficient or functionally obsolete in 2019 [1], which require immediate maintenance and regular monitoring. Therefore, an efficient and effective bridge health monitoring (BHM) approach is essential for keeping our bridges safe and reliable.

Manual inspection [2] and dedicated sensor instrumentation [3] are currently used to monitor bridge conditions. Manual inspection by skilled inspectors is the primary strategy for BHM on the majority of the bridges, but such an approach is labor-intensive, time-consuming, and potentially hazardous [4]. To address such disadvantages, dedicated sensor-based BHM approaches that install sensing devices directly on bridges to collect structural performance data autonomously and continuously were introduced; however, these approaches are also costly and inefficient as they require on-site installation and maintenance of the dedicated sensing equipment and instruments on every bridge of interest [5]. To improve scalability and efficiency, many researchers recently proposed mobile sensing methods for BHM. For example, capturing visual and dynamic information by scanning the bridge using vehicles (e.g., cars and unmanned aerial vehicles) in a nondedicated manner [6–10]. Although such mobile sensing methods can capture high-spatial-resolution information of multiple bridges, they have limited temporal information at each coordinate of bridges due to their mobile sensing nature, restricting their ability to continuously infer and diagnose bridge conditions.

To this end, we introduce a novel BHM system that uses pre-existing telecommunication fiber cables as a nondedicated distributed acoustic sensing (DAS) system to collect bridge dynamic responses. DAS is based on the ϕ

-OTDR technique [11] that uses a standard fiber cable as long-offset virtual dynamic strain sensor arrays [12–14]. In particular, a DAS system consists of an interrogator unit and a standard optical fiber, where the DAS interrogator unit injects short laser pulses into the optical fiber and receives the Rayleigh backscattering induced by inhomogeneities in the optical fiber [15]. The scattering centers could phase shift as a result of a strain perturbation in the fiber's surroundings. Therefore, by repeatedly measuring the phase shift, the strain variations in different fiber sections can be obtained. Due to the unique capacity of DAS, it was initially proposed for intrusion detection for important infrastructure [16]. Later, it was introduced in many important applications, including acoustic source localization [17], industrial downhole monitoring [18], near-surface imaging [12, 19], seismic monitoring [14], and road traffic monitoring [13, 20].

Using pre-existing telecommunication fiber cables for BHM is low-cost, scalable, and efficient. In modern society, millions of kilometers of fiber-optic cables are deployed for telecommunication around the world. For instance, in 2017, the length of the fiber-optic cable network in China alone was more than 37 million kilometers [21]. Also, most telecommunication infrastructure utilizes pipes and conduits installed on bridges to span rivers, roads, or other obstacles. DAS responses collected from cables bound to bridges reflect bridge dynamic responses that contain damage-sensitive information. Such information can be used for diagnosing bridge damages and providing early alarm about bridge health. In addition, our system only requires connecting an optoelectronic instrument called the interrogator unit to one end of the fiber. The interrogator unit used by our system can record strain data from a telecommunication fiber cable up to 100 km long in a high spatial-temporal resolution (up to 250 Hz and 1-meter channel spacing) [13]. By taking advantage of unlit dark fibers (i.e., fibers that are not used for data transmission), our BHM system can continuously record data for years without any interference to regular telecommunication signals or any on-site sensor installation and maintenance [19, 22].

In this work, we extract three damage-sensitive features using DAS responses from telecommunication cables, namely, bridge natural frequencies, strain mode shapes, and displacement mode shapes. Natural frequencies and displacement mode shapes are sensitive to bridge global behavior, so they have been used as indicators for damage detection [23, 24]. Strain mode shapes reflect the longitudinal strain distribution along the bridge and are more sensitive to bridge local behavior [25]. Hence, they have been used for damage localization and quantification [23–25].

However, estimating the above-given features using this novel system is challenging due to its nondedicated nature. In particular, telecommunication fiber cables were originally deployed and optimized for data transmission as opposed to strain sensing. Due to distinct properties of the cable deployments, such as fiber conduit types, cable materials, coupling conditions between the cable and the bridge, and so on, the cable responses have larger measurement noises and

uncertainties than direct measurements from dedicated and well-calibrated bridge sensors (e.g., strain gauges, geophones, and accelerometers). Especially, the error of strain mode shape estimation propagates when estimating displacement mode shapes as it requires the double integration of the estimated strain mode shapes numerically or analytically. As the error propagates, numerical double integration of strain mode shape becomes unstable; analytical double integration of strain mode shape with conventional basis functions, such as polynomial functions, produces inaccurate results [26, 27].

To overcome the challenge, we introduce a physics-guided damage-sensitive feature extraction method. Our method improves the estimation accuracy of bridge damage-sensitive features and reduces error propagation by constraining strain and displacement mode shapes with physical principles of bridge dynamics. In the first module, we formulate an elemental strain state-space model that considers DAS responses from telecommunication cables as observations of the actual bridge strain dynamics. Our model effectively represents the multiple-input-multiple-output (MIMO) system of the telecommunication cable-bridge interaction in the time domain. We then estimate the bridge's natural frequencies and strain mode shapes using the data-driven stochastic subspace identification (SSI-data) algorithm [28–30]. SSI-data are used because it is an output-only method that works for operational system identification [31]. It estimates modal parameters from only vibration response data obtained in operational rather than laboratory conditions. In addition, based on reliable numerical algorithms of QR decomposition and the singular value decomposition (SVD), it does not need complex (nonlinear) optimization techniques required by other methods such as prediction error minimization [32]. In the second module, we fit a parametric shape function that is derived from bridge dynamics to the estimated strain mode shape. Finally, we perform an analytical double integration on the fitted function to estimate the displacement mode shape. This method avoids the instability problem that may be encountered by numerical double integration methods and improves estimation accuracy by physically constraining the mode shape functions.

We evaluated our system with real-world field experiments on a 50-meter-long concrete continuous three-span bridge in San Jose, California, with telecommunication fiber cables running in a conduit beneath the deck. Our system effectively identified the first three natural frequencies with up to 1.2% mean absolute percentage difference compared to those extracted from accelerometers deployed on the bridge deck. In addition, our system estimated the first three strain and displacement mode shapes in a meter-scale resolution and achieved a 0.800 modal assurance criterion compared to those from accelerometers. Overall, our evaluation results show the applicability and feasibility of turning telecommunication fiber cables into large-scale, cost-effective, and high-spatial-temporal-resolution BHM systems.

In summary, the three main contributions of the study are as follows:

(1) introduce a nondedicated BHM system through DAS with pre-existing telecommunication fiber cables to capture the dynamic strain responses of bridges. This system could achieve cost-effective, efficient, and large-scale BHM since it eliminates the need for on-site sensor installation and maintenance.

(2) introduce a new physics-guided system identification method that overcomes the large measurement noise and uncertainty challenge by constraining the mode shapes with bridge dynamics. This method can accurately estimate damage-sensitive features of the bridge, including bridge natural frequencies and strain and displacement mode shapes.

(3) validated this new system on real-world experiments. Our experimental results show that DAS responses from cables attached to the bridge contain damage-sensitive features, which can be directly used for model-based BHM or indirectly analyzed for data-driven BHM.

The remainder of this paper is divided into five sections. Section 2 provides a background of the DAS with pre-existing telecommunication fiber cables. Section 3 presents our vibration-based BHM system using telecommunication cable responses. Section 4 describes the evaluation of our system with field experiments, followed by Section 5 which presents the evaluation results. Section 6 concludes this paper and discusses the important future works of using DAS for BHM.

2. Distributed Acoustic Sensing with Pre-Existing Telecommunication Fiber Cables

To provide a background understanding of our BHM system that uses pre-existing telecommunication cables for distributed acoustic sensing (DAS), we begin with describing the principles of DAS. DAS is an emerging technology that uses a long-range fiber-optic cable as distributed virtual sensors to measure the ground vibration along the cable with high temporal and spatial sampling rates. A DAS system comprises an interrogator unit and a standard optical fiber, where the DAS interrogator unit injects short laser pulses into the optical fiber and receives the Rayleigh backscattering induced by naturally occurring inhomogeneities in the optical fiber [15]. By means of optical reflectometry, the optical phase change between the outgoing and backscattered laser pulses is measured, which can be linearly converted to the strain or strain rate measurements exerted on the fiber-optic cable [19]. DAS has received increasing attention in geotechnical applications [12, 14, 19] and urban monitoring [13] in recent years thanks to its advantages over traditional single-point sensors. Modern DAS can achieve meter-scale spatial resolution on a fiber-optic cable of tens of kilometers long, resulting in tens of thousands of sensing channels [22], whereas point sensors, such as geophones, fiber Bragg gratings (FBGs), and accelerometers, have limited spatial resolution due to their high per unit cost.

We can rapidly establish a large-scale DAS system by connecting the pre-existing telecommunication fiber cable with a DAS interrogator unit stored in a secured room for meter-scale spatial resolution monitoring at the cost of a single interrogator unit. The well-established telecommunication fiber-optic network can be repurposed by DAS as strain sensors for infrastructure monitoring. By taking advantage of unlit dark fibers (i.e., spare fibers that are not used for data transmission), DAS can continuously record data for years without any interference to regular telecommunication signals or any on-site sensor installation and maintenance [19, 22]. On the other hand, the installation and maintenance of conventional point sensors are logistically challenging, especially since their usage in large-scale and long-period measurements is limited in urban environments.

Therefore, by leveraging pre-existing along-bridge telecommunication fiber, DAS offers significant potential for BHM. The fiber-optic cable becomes an array of numerous densely spaced virtual sensors attached to the bridge that records the local vibration of the bridge induced by the traffic (Figures 1(a) and 1(b)). DAS array data are therefore in a 2-D format, with axes in space and time (Figure 1(c)). After DAS system-related preprocessing [33], vehicles traveling along the road instrumented by DAS fiber can be easily identified in the space-time domain, and their interactions with bridges generate characteristic DAS signals that can be used for BHM. Bridge dynamic properties and damage-sensitive features, such as natural frequencies and mode shapes, can be extracted using DAS responses from along-bridge telecommunication cables. These features can be used for bridge damage diagnosis.

Although using pre-existing telecommunication fiber cables enables low-cost and low-maintenance BHM, challenges still exist due to the larger measurement noise and uncertainty of DAS responses from nondedicated fibers compared to that of well-calibrated and dedicated sensors. The large measurement noise and uncertainty affect the estimation accuracy of bridge dynamic properties and damage-sensitive features. To overcome this challenge, we develop an elemental strain state-space model followed by a physics-guided analytical double integration method to extract bridge damage-sensitive features (Section 3). We overcome the large measurement noise and uncertainty challenge by constraining the estimation of mode shapes with bridge dynamics.

3. Vibration-Based BHM System Using DAS with Pre-Existing Telecommunication Cables

This section introduces the details of our system (as shown in Figure 2) that uses DAS responses from telecommunication cables for vibration-based BHM. Our system contains two modules: 1) an elemental strain state-space model identification module and 2) a physics-guided displacement mode shape estimation module. In the following subsections, we present each module in detail.
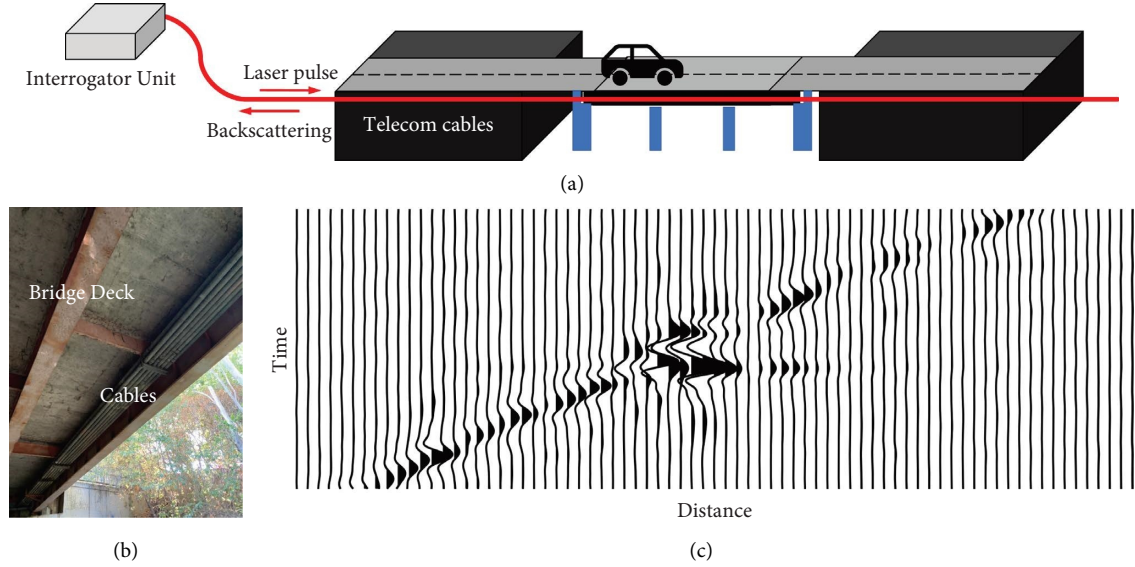
FIGURE 1: An illustration of the distributed acoustic sensing system with pre-existing telecommunication (telecom) fiber cables. (a) Distributed acoustic sensing with pre-existing telecommunication fiber cables. (b) Telecom cables attached to a bridge. (c) An example of DAS measurement.
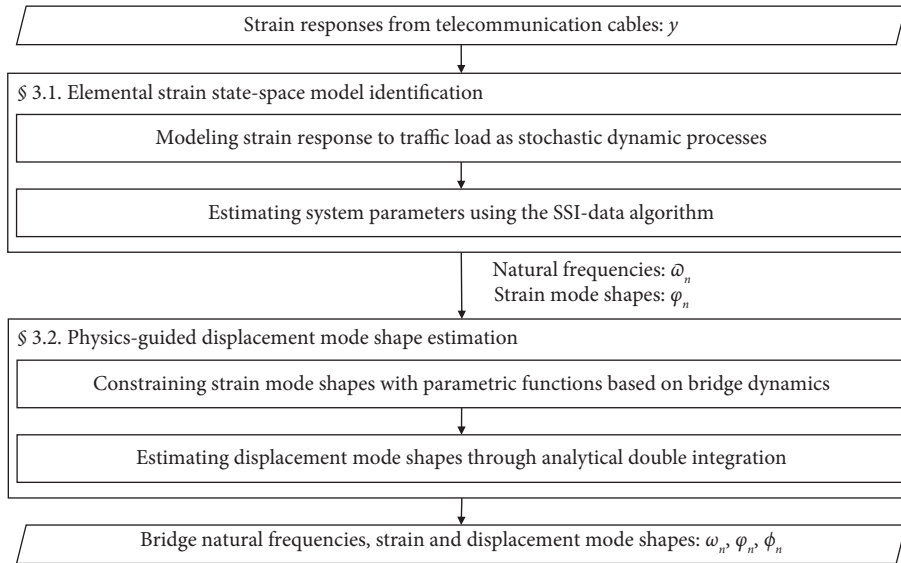

FIGURE 2: The flowchart of our bridge health monitoring system using DAS with pre-existing telecommunication fiber cables.

3.1. Elemental Strain State-Space Model Identification.

The elemental strain state-space model identification module contains two steps: (1) modeling strain response to traffic load as stochastic dynamic processes and (2) estimating system parameters using the SSI-data algorithm. The estimated system parameters are then used to estimate bridge damage-sensitive features. We describe the details of the two steps in the following paragraphs.

First, an elemental strain state-space model is employed to mathematically formulate bridge dynamics with DAS responses from telecommunication cables. A dynamical system can be represented conventionally in four different ways, namely, the state-space model, the differential equation model, the impulse response model, and the transfer function model [28]. We use the state-space model because it (1) models both the physics and measurement systems using the state and output equations, respectively and (2) is a more compact and convenient representation for multiple-input-multiple-output (MIMO) systems compared to other representations of dynamical systems.

For the bridge dynamics in the finite element formulation with N degrees of freedom, the differential equation based on elemental strain is

$$M_\varepsilon \ddot{\varepsilon}(t) + C_\varepsilon \dot{\varepsilon}(t) + K_\varepsilon \varepsilon(t) = Q^{-T} f(t), \tag{1}$$

where $M_\varepsilon = Q^{-T}MQ^{-1}$, $C_\varepsilon = Q^{-T}CQ^{-1}$, $K_\varepsilon = Q^{-T}KQ^{-1}$; M, C, and K are the mass, damping, and stiffness matrices of the bridge structure, respectively; $\varepsilon(t)$ is the nodal strain vector; $f(t)$ is the input load on the bridge due to traffic traveling; Q is the transformation matrix that converts the nodal displacement vector to the nodal strain vector [34], and Q^{-T} is the transpose of Q^{-1}.

Traffic excitation is the primary source of the input load, $f(t)$, applied to the bridge, which may not be measured in practice. Thus, the input loads have to be identified together with the system. Such identification problem is called the output-only system identification that requires a pre-processing to convert the unmeasured input loads into stochastic or harmonic loads [31]. Signals produced by traffic excitation (i.e., moving vehicles) that are recorded by telecommunication cables have mainly two components: quasi-static signals ($<$ 1 Hz for street traffic) resulting from the ground deformation due to the vehicle weight and surface waves (3 to 20 Hz) caused by the dynamic vehicle-road interaction because of the roughness profile of the road [12]. From our observation, most of the vehicles don't excite evident 1 ~ 3 Hz energy, which is relatively weak unless other excitation sources (e.g., earthquakes) exist, so we ignore this band signal in our analysis. Since there are many cars, trucks, bicycles, people, etc., all with different dynamic characteristics, positions, and speeds, acting on the bridge during the measurement, such traffic loads can be approximately modeled in a stochastic way after removing their contributions to the quasi-static signals [31].

Therefore, in the first step, a high-pass filter with a 1-Hz cutoff frequency is applied to DAS responses from telecommunication cables to attenuate quasi-static signals due to traffic traveling. In this way, we model the bridge strain response to traffic load as stochastic dynamic processes using the following elemental strain state-space model:

$$\dot{\hat{x}}(t) = A_c x(t) + w(t),$$
$$y(t) = C_c x(t) + v(t). \tag{2}$$

where $x(t) = \{\varepsilon_d(t)^T, \dot{\varepsilon}_d(t)^T\}$ is the system state; $\varepsilon_d(t)^T$ and $\dot{\varepsilon}_d(t)^T$ are dynamic nodal strain and strain rate of the bridge, respectively; $y(t)$, $w(t)$, and $v(t)$ are the high-pass filtered output observation, state noise and stochastic input load, and observation noise at time t, respectively; and

$$A_c = \begin{bmatrix} 0, & I \\ -M_\varepsilon^{-1}K_\varepsilon, & -M_\varepsilon^{-1}C_\varepsilon \end{bmatrix}, C_c = [I, 0], \tag{3}$$

are the system and observation matrices, respectively. Our elemental strain state-space model considers DAS responses from the telecommunication cable as observations, $y(t)$, of the actual bridge strain dynamics, $x(t)$.

The continuous-time state-space model is converted to discrete-time with the time step Δt by holding each sample value for one sample interval (i.e., zero-order hold [35]),

$$\dot{x}_k = A x_{k-1} + w_k,$$
$$y_k = C x_k + v_k, \tag{4}$$

where $A = e^{A_c \Delta t}$ and $C = C_c$.

Then, in the second step, we adopt the data-driven stochastic subspace identification algorithm (SSI-data) [29, 31] to estimate the dynamic properties of the elemental strain state-space model (equation (4)). The outputs of this step are the estimations of bridge natural frequencies, ω_n, and strain mode shapes, φ_n. Strain mode shapes then become the input to the next module for estimating displacement mode shapes.

The SSI-data algorithm estimates the stochastic linear system given a set of output measurements. It proceeds with projecting the row space of the future output Y_f into the row space of the past output Y_p, where Y_f and Y_p are defined in a block Hankel matrix:

$$H = \frac{1}{\sqrt{j}} \left(\begin{array}{cccc} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_j \\ \cdots & \cdots & \cdots & \cdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \\ y_i & y_{i+1} & \cdots & y_{i+j-1} \\ y_{i+1} & y_{i+2} & \cdots & y_{i+j} \\ \cdots & \cdots & \cdots & \cdots \\ y_{2i-1} & y_{2i} & \cdots & y_{2i+j-2} \end{array} \right) = \frac{Y_p}{Y_f}, \tag{5}$$

where i and j are user-defined numbers of block rows and columns, respectively. The projection is defined as

$$\mathscr{P}_i = Y_f \cdot Y_p^T \cdot \left(Y_p Y_p^T\right)^\dagger \cdot Y_p, \tag{6}$$

where $(\cdot)^\dagger$ denotes the pseudoinverse of a matrix. This projection retains all the information in the past that is useful to predict the future and can be factorized as the production of the observability matrix O_i and the Kalman filter state sequence $\hat{X}_i$ [36]:

$$\mathscr{P}_i = \left(\begin{array}{c} C \\ CA \\ CA^2 \\ \cdots \\ CA^{i-1} \end{array} \right) (\hat{x}_i, \hat{x}_{i+1}, \ldots, \hat{x}_{i+j-1}) = O_i \hat{X}_i, \tag{7}$$

O_i can be obtained by applying SVD to the projection matrix:

$$\mathscr{P}_i = USV^T = (U_1, U_2) \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}, \tag{8}$$

$$O_i = U_1 S_1^{1/2}.$$

Furthermore, we can obtain the estimated system matrices $\hat{C}$ and $\hat{A}$ from the estimated observability matrix, $\hat{O}_i$. The dynamic properties of the system are characterized by performing an eigenvalue decomposition of the system matrix $\hat{A} = \Psi\Lambda\Psi^{-1}$:

$$\omega_n = \left| \frac{\ln \lambda_n}{\Delta t} \right|, \Phi = \hat{C}\Psi, \tag{9}$$

where ω_n is the eigenfrequency of mode n; $\lambda_n = \Lambda_{nn}$ is the n-th eigenvalue; Φ is the mode shape matrix with each column ϕ_n being the strain mode shape of mode n. Note that, the physical modes of the structure appear at nearly the same eigenfrequency when the model order is overspecified [31]. Therefore, the same eigenfrequency being identified for models with different model orders would be one of the structural natural frequencies.

Since DAS measures axial strain responses of the cable along the bridge, we only estimate the strain mode shapes with the SSI-data algorithm, not the displacement mode shape. Displacement mode shape is an important global bridge property that shows deflection patterns related to a particular natural frequency and represents the relative displacement of all components of the bridge. To this end, in the next module, a physics-guided double integration method that estimates displacement mode shapes based on bridge dynamics and beam theory is introduced.

3.2. Physics-Guided Displacement Mode Shape Estimation.

The physics-guided displacement mode shape estimation module contains two steps: (1) constraining strain mode shapes with parametric functions based on bridge dynamics and (2) estimating displacement mode shapes through analytical double integration. The details of the two steps and explanations are provided in the following paragraphs.

We first revisit the bridge dynamics to determine the relationship between the displacement mode shape and the strain mode shape. Specifically, assuming an Euler–Bernoulli beam whose span is greater than ten times the height of its cross-section so that the effect of shear force on the deflection is negligible. We also assume that all the strain measurement points have the same distance to the neutral axis of the beam so that the longitudinal strain is proportional to the beam curvature. According to the beam theory [37], the relationship between the vertical displacement and longitudinal strain is

$$u''(l,t) = -\frac{1}{d}\varepsilon(l,t), \tag{10}$$

where $u(l,t)$ and $\varepsilon(l,t)$ are the vertical displacement and strain of position l along the beam at time t, respectively; d is the distance from the strain measurement point to the beam neutral axis. Furthermore, according to beam dynamics [38], the beam displacement and strain can be expressed as the following form of mode shapes superimposed with the modal coordinates:

$$u(l,t) = \sum_{n=1}^{N} \phi_n(l)q_n(t),$$

$$\varepsilon(l,t) = \sum_{n=1}^{N} \varphi_n(l)q_n(t), \tag{11}$$

where $\phi_n(l)$ and $\varphi_n(l)$ are the n-th displacement and strain mode shapes of position l, respectively; $q_n(t)$ is the n-th modal coordinate at time t. Combining equations (10) and (11), we obtain the relationship between the n-th displacement and strain mode shape as

$$\phi_n''(l) = -\frac{1}{d}\varphi_n(l). \tag{12}$$

Therefore, estimating the displacement mode shapes requires double integrating the strain mode shapes estimated in the first module of our method. This double integration can be performed numerically or analytically. However, due to the large measurement noise and uncertainty of DAS responses from telecommunication cables, conventional numerical and analytical double integration methods (e.g., the trapezoidal rule [39] and analytical integration with a polynomial basis [26]) would produce inaccurate results as the error propagates in the integration steps.

To this end, we estimate the displacement mode shape by first fitting a physics-guided shape function to the strain mode shape and then double integrating the fitted strain mode shape. In this way, we do not have the instability problem caused by numerical double integration and improve the estimation accuracy of analytical double integration by physically constraining the strain mode shape.

In detail, the physics-guided shape function is derived based on the beam dynamics. According to the homogeneous solution of the beam vibration equation [38], for a multispan continuous beam with P spans, the general solution for the n-th mode's displacement mode shape of the p-th beam span is

$$\begin{aligned} \phi_{n,p}(l) = {}& C_{1,n,p}\sin\left(\beta_{n,p}l\right) + C_{2,n,p}\cos\left(\beta_{n,p}l\right) \\ & + C_{3,n,p}\sinh\left(\beta_{n,p}l\right) + C_{4,n,p}\cosh\left(\beta_{n,p}l\right), \\ & 0 \le l \le l_p, \end{aligned} \tag{13}$$

where $C_{1,n,p}$, $C_{2,n,p}$, $C_{3,n,p}$, and $C_{4,n,p}$ are constant coefficients; $\beta_{n,p}$ is a variable that depends on beam properties such as material density, Young's modulus, moment of inertia, and span length; and l_p is the length of the p-th beam span. Then, by combining equations (12) and (13), the strain mode shape of the p-th beam span is defined as

$$\varphi_{n,p}(l) = -d\beta_{n,p}^2\left(-C_{1,n,p}\sin\left(\beta_{n,p}l\right) - C_{2,n,p}\cos\left(\beta_{n,p}l\right) + C_{3,n,p}\sinh\left(\beta_{n,p}l\right) + C_{4,n,p}\cosh\left(\beta_{n,p}l\right)\right), \text{for } 0 \le l \le l_p. \tag{14}$$

Constant coefficients, $C_{1,n,p}$, $C_{2,n,p}$, $C_{3,n,p}$, and $C_{4,n,p}$, are evaluated from the boundary conditions of the beam span:

(1) Zero displacement values at support locations:

$$\phi_{n,p}(0) = \phi_{n,p}(l_p) = 0. \tag{15}$$

(2) Continuous angle of rotation at support locations:

$$\begin{aligned} \phi'_{n,p}(0) &= \phi'_{n,p-1}(l_{p-1}), \\ \phi'_{n,p}(l_p) &= \phi'_{n,p+1}(0). \end{aligned} \tag{16}$$

(3) Continuous bending-moment at support locations:

$$\begin{aligned} \phi''_{n,p}(0) &= \phi''_{n,p-1}(l_{p-1}), \\ \phi''_{n,p}(l_p) &= \phi''_{n,p+1}(0). \end{aligned} \tag{17}$$

Subsequently, our method fits the physics-guided shape function defined in equation (14) (subject to equation (15)–(17)) to the strain mode shapes estimated using DAS responses and performs an analytical double integration on the fitted shape functions to estimate the displacement mode shapes.

In summary, our method estimates bridge natural frequencies and strain mode shapes by analyzing telecommunication cables' dynamic responses using the SSI-data algorithm in the first module. Then, by physically constraining the strain mode shape function with bridge dynamics, the second module of our method effectively estimates the displacement mode shapes that provide the necessary understanding of bridge global behaviors.

4. Field Evaluation

This section presents our field evaluation of the introduced system with field experiments on a 50-meter-long concrete continuous three-span bridge in San Jose, California. We first show the experimental setup, followed by a description of our data.

4.1. Experimental Setup. The testbed bridge (the Coyote Creek bridge) is a concrete girder continuous three-span bridge that carries roadway traffic and walkway over Coyote Creek in San Jose, California. The Coyote Creek bridge and a representation of our sensing deployment on the bridge are shown in Figure 3. Telecommunication fiber cables run in a conduit beneath the bridge deck. An Optasense QuantX DAS interrogator [40] that was installed around 2 kilometers away from the bridge performed distributed acoustic sensing on the bridge with a 250 Hz sampling rate. The Optasense QuantX interrogator is based on the phase-sensitive Optical Time-Domain Reflectometry ϕ-OTDR technique [11] with phase demodulation. We chose a 10-meter gauge length and a 1-meter channel spacing in our system to obtain fine-grained bridge dynamic information with an adequate signal-to-noise ratio for accurate bridge health monitoring. Furthermore, four PCB 354C03 accelerometers [41] were installed on the bridge deck to measure the vertical acceleration of the bridge, providing validation signals to our BHM system. Specifically, three of them were installed in the middle of each span, and the last one was installed in the third quarter of the middle main span.

4.2. Data Description. During the experiment, we collected ten sets of telecommunication cable DAS responses and accelerometer data on two consecutive days in June 2021. Each set contains around eight minutes of data. Specifically, the ten datasets include two sets collected from 6:49 pm to 7:03 pm on June 18, 2021, and eight sets collected from 5:19 pm to 6:39 pm on June 19, 2021. Figure 4 shows an example of a 100-second telecommunication cable DAS response and accelerometer data collected from the Coyote Creek bridge on June 18, 2021. We can observe the dynamic responses of the bridge caused by traffic during the measurement. Each vertical darker line in Figure 4(a) and impulse response in Figure 4(b) (e.g., the signal in the red boxes) are the telecommunication cable DAS responses and the acceleration signals induced by a passing-by vehicle on the bridge, respectively. Vehicles traveling in the lane closer to the telecommunication cable and accelerometers induce larger DAS and acceleration responses. For this 100-second sample data in Figure 4, 11 vehicles were passing through the bridge in the lane closer to the telecommunication cable, and six vehicles were in another lane further from the cable.

5. Results and Discussions

This section presents evaluation results and discussions of our system performance for telecommunication cable-based BHM, including bridge natural frequencies identification and strain and displacement mode shapes estimation.

5.1. Natural Frequency Identification Results. We first evaluate the natural frequency identification results of our system using the stabilization diagram. A stabilization diagram is a standard tool in modal analysis to display and identify the modes of a structure [30]. It plots identified modes in a model order versus eigen frequency diagram. At each model order, the SSI-data algorithm is used to estimate the eigen frequencies and mode shapes. The mode of a structure is stable when the stabilization criteria in the estimated modal parameters between two consecutive model orders are smaller than certain threshold values; otherwise, it is an unstable mode. For the eigen frequencies, the stabilization criterion is defined as a mean percentage absolute difference, and for the mode shapes, the criterion is defined as one minus their modal assurance criterion value. We use the same stability criteria values as in the work of Van Overschee et al. [29]: 1% for eigen frequencies and 2% for mode shapes. Figure 5 shows the stabilization diagrams for the accelerometer and our telecommunication cable DAS data. The solid red curve is the complex mode indicator function (CMIF) [42] which is the singular values of the MIMO system's frequency response matrix (equation in the paper by Shih et al. [42]). Local peaks on the CMIF plot indicate damped natural frequencies of the structure. Cross
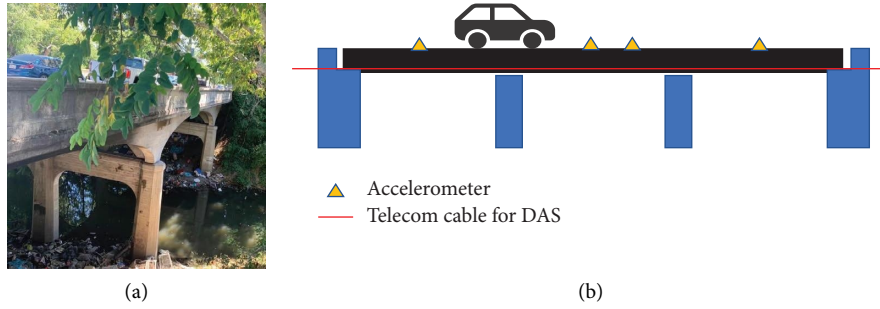
FIGURE 3: Our experimental setup: (a) the Coyote Creek bridge testbed and (b) a representation of our sensing deployment on the bridge.
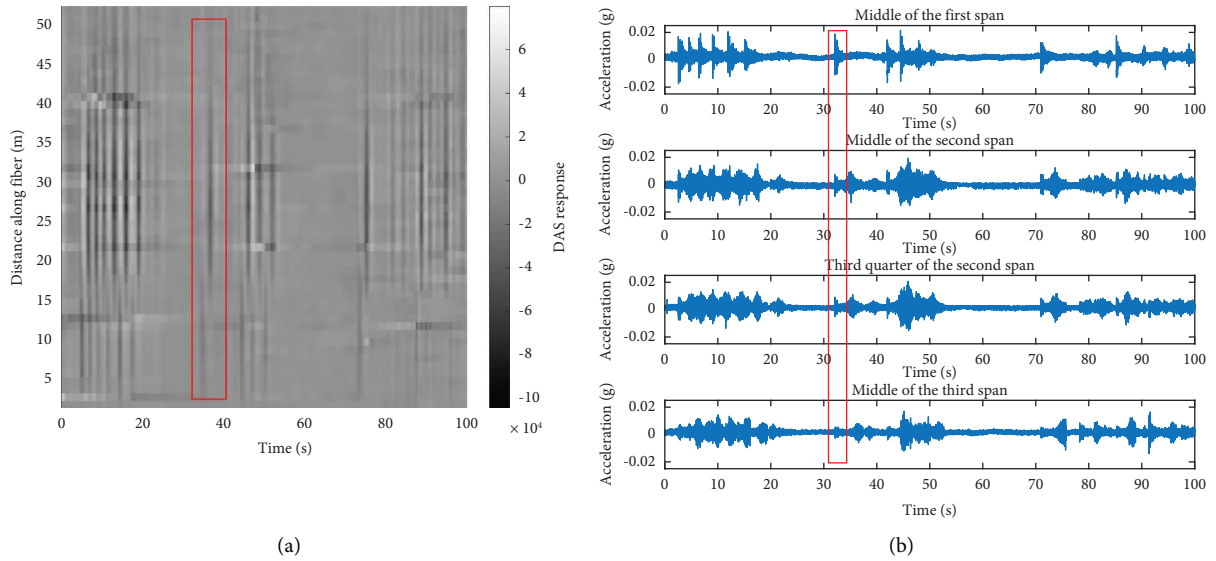

FIGURE 4: An example of 100-seconds of (a) telecommunication cable DAS response data and (b) accelerometer data of the Coyote Creek bridge on June 18, 2021. Red boxes indicate an example of vehicle-induced signals.

markers and dots indicate stable and unstable modes, respectively. We identify the eigen frequencies with stable modes across model orders (i.e., vertically aligned blue crosses in Figure 5) as structural modes. Also, the identified structural modes should be at the local peaks of the CMIF (red curve). Both telecommunication cable DAS data and accelerometer data show that the structural modes are at 4.6, 6.3, and 8.9 Hz (indicated by red ovals in Figure 5).

For the ten datasets we collected, our system successfully identified the first three natural frequencies with a 0.055 Hz mean absolute difference (MAD) or up to 1.2% mean absolute percentage difference compared to those identified from accelerometer data. Table 1 presents the mean values of the identified bridge natural frequencies using the accelerometer data and telecommunication cable DAS data and the MAD values between them. Furthermore, boxplots in Figure 6 show the distributions of absolute differences between identified frequencies of the bridge using accelerometer data and telecommunication cable DAS data. Each box in the boxplot presents five values, including minimum, maximum, median, the first quartile, and the third quartile. Specifically, for the first two modes, our identification results have a 0.110 Hz maximum absolute difference and

a 0.030 Hz MAD. Due to the smaller signal-to-noise ratio for higher modes, the maximum and mean absolute differences for identifying the third mode are 0.272 Hz and 0.107 Hz, respectively, which are larger than those of the first two modes.

5.2. Strain Mode Shape Estimation Results. Figure 7 presents the estimated strain mode shapes using the SSI-data algorithm (indicated as "Measured") and two fitted parametric strain mode shapes using (1) our physics-guided shape function (indicated as "Ours") and (2) a baseline method based on the second-order piecewise polynomial function (indicated as "Poly"). We can observe that compared to our physics-guided method, the polynomial fitting overfits the estimated strain mode shape that includes propagated error from telecommunication cable DAS data.

We evaluate the consistency of the estimated strain mode shapes by computing the modal assurance criterion (MAC) between each pair of mode shapes estimated using the ten datasets. The MAC is a statistical indicator that is calculated as the normalized scalar product of two sets of vectors [43]. It is often used to evaluate the consistency of mode shapes
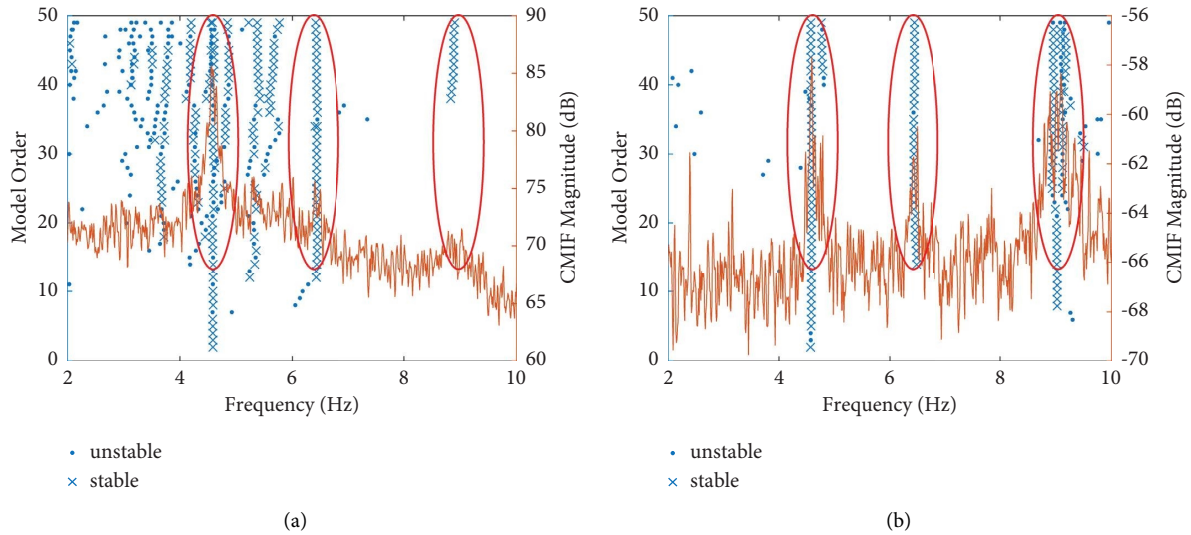
(a)

(b)

FIGURE 5: Stabilization diagrams of (a) telecommunication cable DAS data and (b) accelerometer data collected from 18:57 to 19:05 on June 18, 2021. We identify the same set of modes at 4.6, 6.3, and 8.9 Hz (red ovals) using telecommunication cable DAS data and accelerometer data.

TABLE 1: Mean values of identified bridge natural frequencies using accelerometer data (mean freq-acc) and telecommunication cable DAS data (mean freq-DAS), and the MAD between them. Our system successfully identifies the first three natural frequencies with a 0.055 Hz mean absolute difference (MAD) compared to those identified from accelerometer data.

Mode #	Mean freq-acc (Hz)	Mean freq-DAS (Hz)	MAD (Hz)
1	4.572	4.606	0.034
2	6.331	6.323	0.025
3	8.937	8.848	0.107
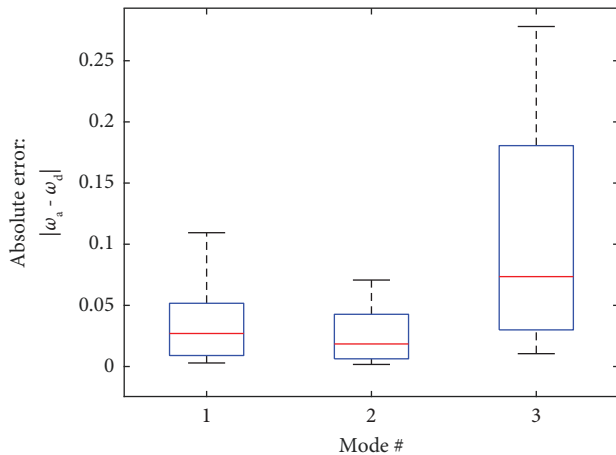
FIGURE 6: Absolute differences between the identified bridge natural frequencies using accelerometer data and telecommunication cable DAS data.

derived from two different scenarios (e.g., two different experiments, sensing modalities, or models such as analytical vs. empirical). It is bounded between 0 and 1, with 1 indicating fully consistent mode shapes. Figure 8 presents a MAC value matrix of estimated strain mode shapes using telecommunication cable DAS data for the entire bridge. The diagonal elements of the matrix in Figure 8 are MAC values of modes compared to themselves, and the off-diagonal elements are MAC values of modes compared to different modes. We can observe that for the same mode comparison (diagonal elements), strain mode shapes of the first two modes are on average 20% more consistent between each pair of the ten datasets than those of the third mode. It is because the third mode response has a smaller magnitude and lower signal-to-noise ratio than the first two mode responses. We also observe that the off-diagonal elements in the red ovals of Figure 8 are close to one, meaning the first strain mode shape is not clearly distinguishable from the second and the third strain mode shapes.

To further validate our strain mode shape estimation, we calculate the MAC value matrices for the three bridge spans (as shown in Figure 9) to show the differences in mode shapes of each bridge span. It is observed that the off-diagonal elements of the first and the third spans' MAC value matrix are small (as shown in Figures 9(a) and 9(c)), meaning that the first three modes' strain mode shapes of the first and the third spans are distinct. On the other hand, the off-diagonal elements of the second span's MAC value matrix are close to 1 (as shown in Figure 9(b)), meaning that the first three modes' strain mode shapes of the second bridge span are similar to one another. It could be because the one-dimensional measurements of telecommunication cables are insufficient to distinguish between independent two- or three-dimensional mode shapes (e.g., mode shapes in torsional directions) for the second bridge span.

5.3. Displacement Mode Shape Estimation Results. We compare the estimations of displacement mode shapes using our method (DAS-ours) with those using two
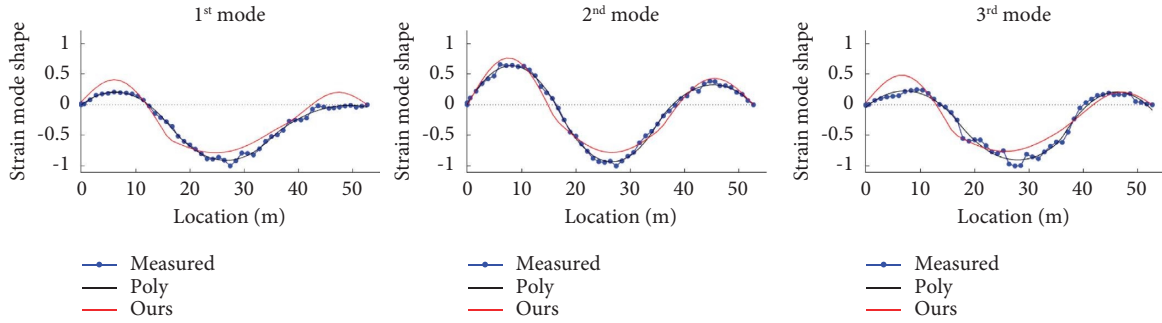
FIGURE 7: Measured and fitted strain mode shapes using telecommunication cable DAS data for the three identified mode shapes using data collected from 18:57 to 19:05 on June 18, 2021.
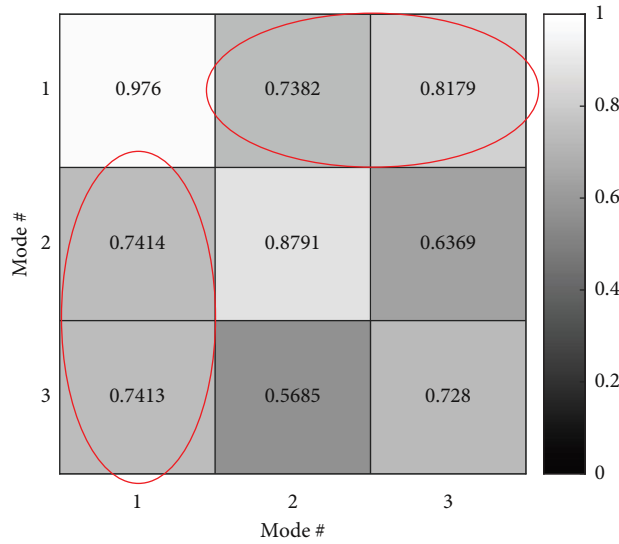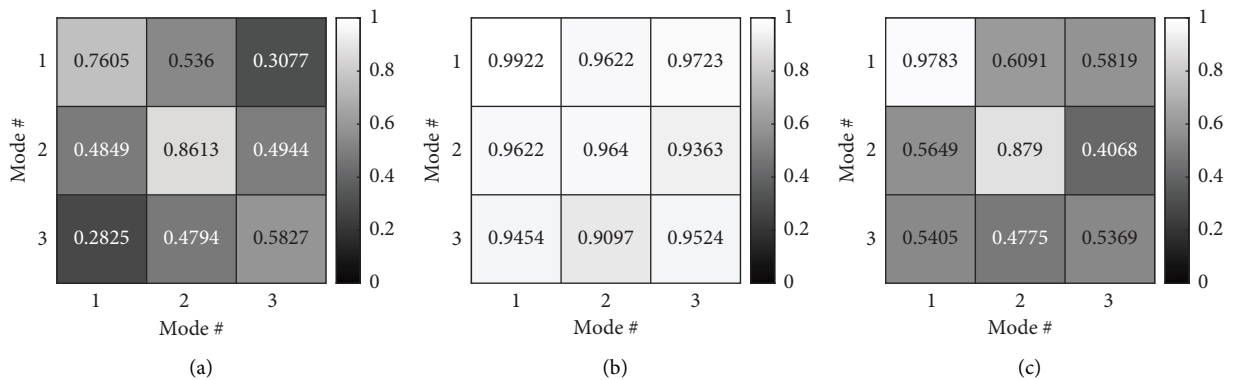

FIGURE 8: MAC value matrix of estimated strain mode shapes using telecommunication cable DAS data. Strain mode shapes of the first two modes are on average 20% more consistent between each pair of the ten datasets than those of the third mode.

FIGURE 9: MAC value matrices of estimated strain mode shapes using telecommunication cable DAS data for different bridge spans. Our estimations of the second bridge span's first three strain mode shapes are similar, but those of the first and the third spans are different. (a) First span. (b) Second span. (c) Third span.
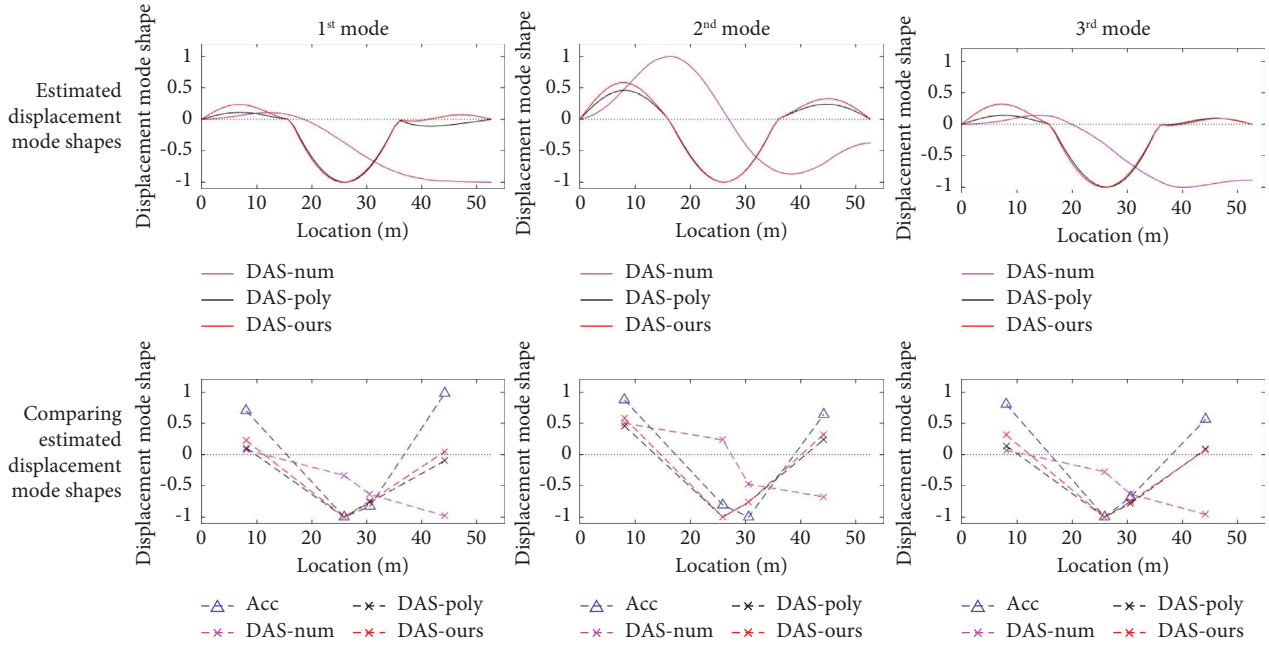
FIGURE 10: Estimated displacement mode shapes (plots in the first row) and comparison of displacement mode shapes between using accelerometer data and telecommunication cable DAS data (plots in the second row) for the three identified mode shapes based on data collected from 18:57 to 19:05 on June 18, 2021.

baseline methods, namely, a numerical method (DAS-num) that double integrates the measured strain mode shapes using the trapezoidal rule and an analytical method (DAS-poly) that double integrates the strain mode shapes fitted using a piecewise polynomial basis function. Figure 10 shows the estimated displacement mode shapes using the three different double integration methods and the comparison of displacement mode shapes between using accelerometer data and telecommunication cable DAS data, respectively.

To evaluate our results, we compute MAC values of displacement mode shapes estimated using accelerometers versus using telecommunication cable DAS data for different double integration methods. Table 2 and Figure 11 present the mean values and distributions of the computed MAC, respectively. Our method achieves the highest consistency with the validation system having an average 0.800 MAC value for the three modes and a 72% and an 11% improvement compared to the numerical method and the polynomial-based method, respectively. Moreover, the displacement mode shapes of the second modes are 6.5% and 7.8% more consistent with the validation system compared to the first and the third modes, respectively. Note that, both the strain and displacement mode shapes estimated using telecommunication cable DAS data have a one-meter spatial resolution, which is very beneficial for detecting local bridge damages and achieving a high-spatial-resolution BHM.

We also evaluate the consistency of the estimated displacement mode shapes between the two days of data

TABLE 2: Mean MAC values of displacement mode shapes estimated using numerical double integration (mean MAC-num), analytical double integration with piecewise polynomial basis (mean MAC-poly), and with our physics-guided shape function (mean MAC-ours)-accelerometer data vs. telecommunication cable DAS data. Our method achieves the highest consistency with the validation system with an average 0.800 MAC value for the three modes.

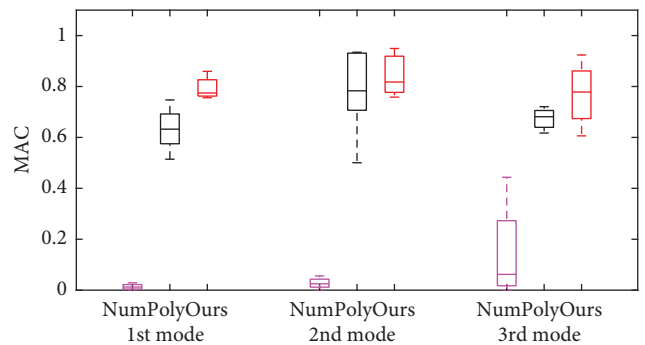| Mode #	Mean MAC-num	Mean MAC-poly	Mean MAC-ours
1	0.013	0.633	0.783
2	0.039	0.783	0.848
3	0.188	0.657	0.770

FIGURE 11: MAC value distributions of displacement mode shapes estimated using accelerometer data vs. using telecommunication cable DAS data. "Num," "Poly," and "Ours" refer to methods using the numerical double integration, analytical double integration with piecewise polynomial basis, and with our physics-guided shape function.
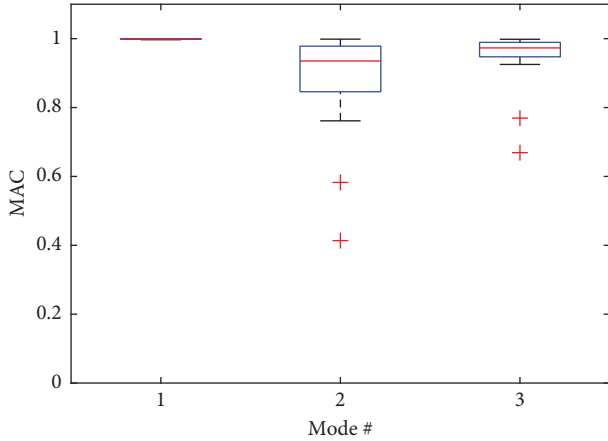
FIGURE 12: MAC values of estimated displacement mode shapes using telecommunication cable DAS data of two different days. Mean MAC values of each mode are 0.999, 0.874, and 0.934, respectively.

collection. Figure 12 shows MAC values of estimated displacement mode shapes using June 18 data versus June 19 data. Cross markers in the figure are outlier values that are more than 1.5× interquartile range away from the top or bottom of the box. The displacement mode shapes estimated using telecommunication cable DAS data have strong consistency (i.e., mean MAC values are greater than 85%) between the two days.

In summary, our evaluation shows promising results for using DAS data from pre-existing telecommunication cables to achieve cost-effective and high-spatial-resolution BHM, including bridge natural frequencies identification and strain and displacement mode shapes estimation.

6. Conclusion

In this work, we introduce a nondedicated BHM system that turns pre-existing telecommunication cables into distributed acoustic sensors to capture bridge dynamic strain responses. The system successfully estimates damage-sensitive dynamic properties of bridges, including bridge natural frequencies and strain and displacement mode shapes. It enables an efficient and low-cost BHM as it does not require on-site installation and maintenance of sensors and equipment by taking advantage of extensively installed telecommunication fiber cables. Specifically, we developed an elemental strain state-space formulation followed by a physics-guided analytical double integration method to overcome the challenge of inaccurate estimations for the dynamic properties due to the large measurement noise of DAS responses from telecommunication cables. The system was evaluated with a concrete three-span bridge and validated with a conventional system using accelerometers. Our system effectively identified the first three bridge natural frequencies with a 0.055 Hz mean absolute difference compared to the natural frequencies identified using a conventional accelerometer system. It also

estimates the strain and displacement mode shapes with a 0.800 modal assurance criterion, resulting in 72% and 11% improvements compared to the two baseline methods that use conventional numerical and analytical double integration, respectively. The experimental results discussed in the result section have validated the feasibility and the effectiveness of the introduced system to extract bridge dynamic information for achieving an efficient BHM.

7. Future Work

To the best of our knowledge, this work is the first one that uses pre-existing telecommunication fiber cables for BHM. Our vision of future works is provided as follows:

(1) Characterizing the noise of using telecommunication fiber cables for DAS. Since the introduced system is a nondedicated sensing system, it would have more uncertainties and noise sources and larger noise magnitudes compared with the dedicated and well-calibrated sensors. For example, the telecommunication cable at different DAS channels along the bridge often has different coupling conditions to the bridge structure, resulting in different levels of sensitivity of distributed acoustic sensors. In the future, we intend to study how different coupling conditions affect the DAS responses and characterize different types of noise and uncertainty in this new system.

(2) Developing a bridge damage diagnosis framework. The introduced BHM system in this paper presents an efficient way to extract bridge dynamic properties using DAS responses from pre-existing telecommunication fiber cables. It enables model-based BHM that diagnoses bridge damage through detecting and quantifying changes in the estimated dynamic properties such as modal parameters. Also, damage-sensitive features can be extracted from DAS responses for bridge damage diagnosis using data-driven methods such as dimensionality reduction techniques, classification, and regression methods. In the future, we plan to develop a framework to detect, localize, and quantify bridge damage using DAS responses from telecommunication cables. The framework would include DAS data collection, damage-sensitive feature extraction, damage diagnosis and inference, and BHM decision-making.

(3) Validating on comprehensive and complex field experiments. The evaluation of our system in this paper was conducted on an operational bridge that carries roadways and pedestrian walkways. The experimental time and the allowed number of validation sensors are limited to not interfere with regular traffic. Especially, as there is a limited number of accelerometers being used in the validation system, it may not be sufficient to distinguish the independent mode shapes from each other and validate the high-spatial-resolution

mode shapes estimated using DAS. In the future, we plan to look for a testing bridge that would allow us to deploy more sensors and even induce structural changes (or damages) to further validate the capacity of the introduced system for assessing bridge health and detecting damages.

Data Availability

The telecommunication cable DAS data and accelerometer data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] ASCE, "Bridges-infrastructure report card," 2021, https://infrastructurereportcard.org/wp-content/uploads/2020/12/Bridges-2021.pdf.

[2] R. A. Hartle, T. W. Ryan, and E. Mann, *Bridge Inspector's Reference Manual: Volume 1 and Volume 2*, Federal Highway Administration, Washington, DC, USA, 2002.

[3] S. Sony, S. Laventure, and A. Sadhu, "A literature review of next-generation smart sensing technology in structural health monitoring," *Structural Control and Health Monitoring*, vol. 26, no. 3, Article ID e2321, 2019.

[4] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah, "Review of bridge structural health monitoring aided by big data and artificial intelligence: from condition assessment to damage detection," *Journal of Structural Engineering*, vol. 146, no. 5, Article ID 04020073, 2020.

[5] J. Liu, S. Chen, M. Bergés et al., "Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction," *Mechanical Systems and Signal Processing*, vol. 136, Article ID 106454, 2020.

[6] J. Liu, B. Chen, S. Chen, M. Bergés, J. Bielak, and H. Noh, "Damage-sensitive and domain-invariant feature extraction for vehicle-vibration-based bridge health monitoring," *IEEE*, 2020.

[7] J. Liu, S. Xu, M. Bergés, and H. Y. Noh, "HierMUD: hierarchical multi-task unsupervised domain adaptation between bridges for drive-by damage diagnosis," *Structural Health Monitoring*, Article ID 147592172210811, 2022.

[8] J. Liu, "Scalable bridge health monitoring using drive-by vehicles: PhD forum abstract," *ACM*, 2020.

[9] B. F. Spencer Jr, V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, 2019.

[10] Y. B. Yang and J. P. Yang, "State-of-the-art review on modal identification and damage detection of bridges by moving test vehicles," *International Journal of Structural Stability and Dynamics*, vol. 18, no. 2, Article ID 1850025, 2018.

[11] S. V. Shatalin, V. N. Treschikov, and A. J. Rogers, "Interferometric optical time-domain reflectometry for distributed optical-fiber sensing," *Applied Optics*, vol. 37, no. 24, pp. 5600–5604, 1998.

[12] S. Yuan, A. Lellouch, R. G. Clapp, and B. Biondi, "Near-surface characterization using a roadside distributed acoustic sensing array," *The Leading Edge*, vol. 39, no. 9, pp. 646–653, 2020.

[13] S. Yuan, J. Liu, H. Y. Noh, and B. Biondi, *Urban System Monitoring Using Combined Vehicle Onboard Sensing and Roadside Distributed Acoustic Sensing*Society of Exploration Geophysicists, Houston, Texas, USA, 2021.

[14] B. Biondi, R. G. Clapp, S. Yuan, and F. Huot, *Scaling up to City-wide Dark-Fiber Seismic Arrays: Lessons from Five Years of the Stanford DAS Array Project*Society of Exploration Geophysicists, Houston, Texas, USA, 2021.

[15] P. G. E. Lumens, *Fibre-optic Sensing for Application in Oil and Gas wells*, Technische Universiteit Eindhoven, Eindhoven, Netherlands, 2014.

[16] H. F. Taylor and C. E. Lee, "Apparatus and method for fiber optic intrusion sensing," *US Patent*, vol. 5, 1993, https://patents.google.com/patent/US5194847A/en.

[17] L. Jiajing, W. Zhaoyong, L. Bin et al., "Distributed acoustic sensing for 2D and 3D acoustic source localization," *Optics Letters*, vol. 44, no. 7, pp. 1690–1693, 2019.

[18] A. Lellouch, B. Luo, F. Huot et al., "Microseismic analysis over a single horizontal distributed acoustic sensing fiber using guided waves," *Geophysics*, vol. 87, no. 3, pp. KS83–KS95, 2022.

[19] N. J. Lindsey, H. Rademacher, and J. B. Ajo-Franklin, "On the broadband instrument response of fiber-optic DAS arrays," *Journal of Geophysical Research: Solid Earth*, vol. 125, no. 2, Article ID e2019JB018145, 2020.

[20] S. Yuan, E. Mvd, and J. Liu, "Spatial deep deconvolution u-net for traffic analyses with das," 2022, https://arxiv.org/abs/2212.03936.

[21] Length, "Length of optical fiber cable network in china from 2012 to 2017," 2022, https://www.statista.com/statistics/898868/china-optical-fiber-cable-network-length.

[22] J. B. Ajo-Franklin, S. Dou, N. J. Lindsey et al., "Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection," *Scientific Reports*, vol. 9, no. 1, pp. 1328–1414, 2019.

[23] N. Maia, J. Silva, E. Almas, and R. Sampaio, "Damage detection in structures: from mode shape to frequency response function methods," *Mechanical Systems and Signal Processing*, vol. 17, no. 3, pp. 489–498, 2003.

[24] W. Hong, J. Zhang, G. Wu, and Z. Wu, "Comprehensive comparison of macro-strain mode and displacement mode based on different sensing technologies," *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 563–579, 2015.

[25] A. Pandey, M. Biswas, and M. Samman, "Damage detection from changes in curvature mode shapes," *Journal of Sound and Vibration*, vol. 145, no. 2, pp. 321–332, 1991.

[26] H. Xu, W. X. Ren, and Z. C. Wang, "Deflection estimation of bending beam structures using fiber Bragg grating strain sensors," *Advances in Structural Engineering*, vol. 18, no. 3, pp. 395–403, 2015.

[27] D. Inaudi and S. Vurpillot, "Monitoring of concrete bridges with long-gage fiber optic sensors," *Journal of Intelligent Material Systems and Structures*, vol. 10, no. 4, pp. 280–292, 1999.

[28] K. J. Keesman and K. J. Keesman, *System identification: an introduction*, Springer, Heidelberg, Germany, , 2011

[29] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory–Implementation–Applications*, Springer Science & Business Media, Heidelberg, Germany, 2012.

[30] C. Rainieri and G. Fabbrocino, *Operational Modal Analysis of Civil Engineering Structures*, Springer, Heidelberg, Germany, 2014.

[31] E. Reynders, "System identification methods for (operational) modal analysis: review and comparison," *Archives of Computational Methods in Engineering*, vol. 19, no. 1, pp. 51–124, 2012.

[32] T. Katayama, *Subspace Methods for System Identification. 1*, Springer, Heidelberg, Germany, 2005.

[33] A. Lellouch, R. Schultz, N. J. Lindsey, B. L. Biondi, and W. L. Ellsworth, "Low-magnitude seismicity with a downhole distributed acoustic sensing array-examples from the FORGE geothermal experiment," *Journal of Geophysical Research: Solid Earth*, vol. 126, no. 1, Article ID e2020JB020462, 2021.

[34] P. Ren and Z. Zhou, "Strain response estimation for the fatigue monitoring of an offshore truss structure," *Pacific Science Review*, vol. 16, no. 1, pp. 29–35, 2014.

[35] K. C. Pohlmann, *Principles of Digital Audio*, Sams, Bentonville, AR, USA, 1989.

[36] B. Peeters and G. De Roeck, "Reference based stochastic subspace identification in civil engineering," *Inverse Problems in Engineering*, vol. 8, no. 1, pp. 47–74, 2000.

[37] O. A. Bauchau and J. I. Craig, *Euler-Bernoulli Beam Theory* Springer, Heidelberg, Germany, 2009.

[38] A. K. Chopra, *Dynamics of Structures*, Pearson Education India, Noida, India, 2007.

[39] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Courier Corporation, Chelmsford, Ma, USA, 2007.

[40] Optasense, "Quantx das interrogator: optasense," 2022, https://www.optasense.com/technology/quantx.

[41] P. C. B. Piezotronics, "PCB Model 354C03," 2022, https://www.pcb.com/products?m=354c03.

[42] C. Shih, Y. Tsuei, R. Allemang, and D. Brown, "Complex mode indication function and its applications to spatial domain parameter estimation," *Mechanical Systems and Signal Processing*, vol. 2, no. 4, pp. 367–377, 1988.

[43] M. Pastor, M. Binda, and T. Harčarik, "Modal assurance criterion," *Procedia Engineering*, vol. 48, pp. 543–548, 2012.

[44] J. Liu, S. Yuan, B. Biondi, and H. Y. Noh, "Turning telecommunication cables into distributed acoustic sensors for bridge health monitoring," 2022, https://www.emi-conference.org/sites/emi-conference.org/2022/files/inline-files/EMI%202022%20Book%20of%20Abstracts.pdf.

TelecomTM: A Fine-Grained and Ubiquitous Traffic Monitoring System Using Pre-Existing Telecommunication Fiber-Optic Cables as Sensors

JINGXIAO LIU, Stanford University, USA
SIYUAN YUAN, Stanford University, USA
YIWEN DONG, Stanford University, USA
BIONDO BIONDI, Stanford University, USA
HAE YOUNG NOH, Stanford University, USA

We introduce the *TelecomTM* system that uses pre-existing telecommunication fiber-optic cables as virtual strain sensors to sense vehicle-induced ground vibrations for fine-grained and ubiquitous traffic monitoring and characterization. Here we call it a virtual sensor because it is a software-based representation of a physical sensor. Due to the extensively installed telecommunication fiber-optic cables at the roadside, our system using redundant dark fibers enables to monitor traffic at low cost with low maintenance. Many existing traffic monitoring approaches use cameras, piezoelectric sensors, and smartphones, but they are limited due to privacy concerns and/or deployment requirements. Previous studies attempted to use telecommunication cables for traffic monitoring, but they were only exploratory and limited to simple tasks at a coarse granularity, e.g., vehicle detection, due to their hardware constraints and real-world challenges. In particular, those challenges are 1) unknown and heterogeneous properties of virtual sensors and 2) large and complex noise conditions. To this end, our *TelecomTM* system first characterizes the geographic location and analyzes the signal pattern of each virtual sensor through driving tests. We then develop a spatial-domain Bayesian filtering and smoothing algorithm to detect, track, and characterize each vehicle. Our approach uses the spatial dependency of multiple virtual sensors and Newton's laws of motion to combine the distributed sensor data to reduce uncertainties in vehicle detection and tracking. In our real-world evaluation on a two-way traffic road with 1120 virtual sensors, *TelecomTM* achieved 90.18% vehicle detection accuracy, 27× and 5× error reduction for vehicle position and speed tracking compared to a baseline method, and ±3.92% and ±11.98% percent error for vehicle wheelbase and weight estimation, respectively.

CCS Concepts: • **Computer systems organization** → **Embedded and cyber-physical systems**.

Additional Key Words and Phrases: Traffic monitoring, intelligent transportation, distributed acoustic sensing, ubiquitous sensing

Authors' addresses: Jingxiao Liu, liujx@stanford.edu, Stanford University, Stanford, CA, USA, 94305; Siyuan Yuan, syyuan@stanford.edu, Stanford University, Stanford, CA, USA, 94305; Yiwen Dong, ywdong@stanford.edu, Stanford University, Stanford, CA, USA, 94305; Biondo Biondi, biondo@sep.stanford.edu, Stanford University, Stanford, CA, USA, 94305; Hae Young Noh, noh@stanford.edu, Stanford University, Stanford, CA, USA, 94305.

1 INTRODUCTION

A traffic monitoring system, which automatically and continuously detects, tracks, and characterizes vehicles in moving traffic, is important for urban management, maintenance, and planning. For instance, a traffic monitoring system can track and predict traffic patterns to help reduce traffic congestion [30, 32, 34] and manage safety and emergency situations [11, 21, 41]. Also, with a detailed understanding of individual vehicle characteristics (e.g., vehicle number, size, weight) and the use of roads and bridges, we can image the near-surface (tens of meters under the ground) seismic properties of urban areas [56, 57], monitor critical transportation infrastructure [15, 29, 35, 51] and efficiently determine the needs of future transportation projects [9, 14, 53].

There are several existing technologies for traffic monitoring, such as vision-based systems [6, 40, 42] and pavement sensing systems (e.g., inductive loops [4, 16, 18], piezoelectric sensors [16, 23, 60], and fiber optic sensors [49, 58]). However, these systems bring several drawbacks: Vision-based systems are perceived as privacy-invasive and are sensitive to reduced visibility caused by weather conditions. Pavement sensing systems only capture traffic information at specific locations as they are point sensing. Due to the high cost in installations and maintenance, it is difficult to scale up and achieve fine-grained monitoring using existing vision-based and pavement sensing systems. Furthermore, crowd-sensing approaches that use mobile phone data from the drivers/passengers [17, 43, 61] have been developed to enable cost-effective and high-resolution traffic monitoring. However, they only capture the vehicle position and speed information and have also raised privacy concerns.

To this end, we introduce the *TelecomTM* system that uses pre-existing roadside telecommunication (telecom) fiber-optic cables as virtual strain sensors to sense vehicle-induced ground vibrations for fine-grained and ubiquitous traffic monitoring and characterization. In particular, *TelecomTM* achieves vehicle detection, tracking, speed and position estimation, weigh-in-motion, and wheelbase estimation. Our system is based on the distributed acoustic sensing (DAS) technology, specifically, the Phase Sensitive Optical Time Domain Reflectometry (ϕ-OTDR) [31, 38]. A DAS channel is called a "virtual sensor" because it is a software-based representation of a physical sensor. It uses the readings of the backscattered light in the fiber to calculate the strain responses of the telecom cable. *TelecomTM* is built on the idea that vehicle motion creates unique vibration patterns on the ground and near-surface structures. When vehicles move on the road, the ground deforms due to the vehicles' self-weight [19, 56]. When vehicles pass by a structure (e.g., a roadway structure), forces applied by their wheels induce the structure to vibrate [25–27, 54]. These vibrations carrying information about vehicle characteristics (e.g., size and weight) are transmitted to the roadside telecom fiber conduits coupled to the earth and road structures. Our *TelecomTM* system senses these vehicle-induced telecom fiber vibrations to monitor traffic ubiquitously and infer vehicle activities with fine-grained spatial resolution.

TelecomTM is a scalable, efficient, and cost-effective system. There are millions of kilometers of telecom fiber cables deployed around the world that can be utilized for ubiquitous traffic monitoring. For instance, in 2017, the length of the optical fiber cable network in China alone was more than 37 million kilometers [48]. Most telecom infrastructure utilizes pipes and conduits several meters under the ground and along the roadways to distribute around the urban area, which can capture fine-grained traffic information. It only requires connecting an optoelectronic instrument called the interrogator unit to one end of the fiber. The interrogator unit used by *TelecomTM* can record strain data from a telecom fiber cable up to 100 km long in a high spatial-temporal resolution (up to 250 Hz and 1-meter channel spacing) [28, 57]. In addition, by taking advantage of unlit dark fibers (i.e., fibers that are not used for data transmission), *TelecomTM* can continuously record data for years without any interference to regular telecommunication signals or any on-site sensor installation and maintenance [1, 24].

Researchers have explored the idea of using dark fibers for traffic monitoring before; however, due to hardware constraints and real-world challenges, existing works are limited to exploratory simple tasks at coarse-granularity, e.g., vehicle detection and traffic speed estimation [33, 37, 50, 52, 55, 57]. For example, incoherent OTDR [13] only has intensity measurement at a lower spatial resolution (e.g., 10-meter channel spacing). Importantly, existing

works lack a systematic approach to cope with real-world challenges, which prevented them from achieving fine-grained traffic monitoring and individual vehicle characterization. In particular, the key research challenges are:

- **Unknown and heterogeneous properties of virtual sensors**. *TelecomTM* measures the dynamic strain of the fiber around each virtual sensor. Due to cable spooling in cabinets and manholes, each virtual sensor's actual geographic location (geo-location) is unknown. Therefore, the surrounding conditions (e.g., coupling between the cable and the conduit and between the conduit and the earth, near-surface soil properties, etc.) of each virtual sensor is unidentified, resulting in high uncertainties and heterogeneity in signal properties, including signal pattern and the ratio between the virtual sensor's response and vehicle-induced forces (i.e., transmissibility). Without a prior understanding of the geo-location and signal properties, we cannot accurately model vehicle-induced telecom fiber responses at each virtual sensor. Vehicle detection and tracking could also fail due to inaccurate vehicle position and speed estimations.
- **Large and complex noise conditions**. Fiber cables were originally deployed for data transmission as opposed to strain sensing. Telecom fiber responses are indirect measurements of the vehicle-induced ground vibrations, which have larger and more complex noises and uncertainties than direct measurements from dedicated and well-calibrated sensors (e.g., piezoelectric sensors in the pavement). The noise signals created by non-vehicle vibrations (e.g., environmental changes) may be falsely recognized as the vehicle-induced vibration signals or overwhelm the vehicle signals, resulting in wrong or missing detection of vehicles. These large and complex noise conditions can further affect the accuracy of vehicle tracking and characterization.

TelecomTM addresses the above two challenges through a System Characterization module and a Bayesian Analysis using Distributed Sensors module. In the first module, we characterize the system through driving tests that use a car with a GPS antenna to drive across the road. We estimate the geographic position of each virtual sensor by matching the vehicle's GPS signals with the induced telecom fiber responses. Virtual sensor properties, including signal patterns and transmissibility, are also learned from the driving tests to help design the vehicle detection method and determine its model parameters. In the second module, the arrival times of vehicles at each virtual sensor channel are first estimated through a prominence-based peak detection method. Then, a spatial-domain Bayesian filtering and smoothing algorithm is developed to address the challenge of large and complex noise conditions. It estimates the posterior probability of vehicle arrival time recursively over the space (in the direction of vehicle motion) through fusing spatial-dependent vehicle detection results across multiple virtual sensors. It uses the spatial dependency of distributed sensors and Newton's laws of motion to combine the distributed sensor data to reduce vehicle detection uncertainties and estimate vehicle motion states (positions and moving speed). Vehicle tracking is achieved by converting the estimated arrival times and their derivatives into vehicle positions and speeds. Furthermore, the time differences between responses induced by vehicle wheels and the magnitude of vehicle-induced quasi-static strain are calculated for estimating wheelbase length and vehicle weight, respectively.

We evaluated *TelecomTM* through comprehensive field experiments on an approximately 900-meter road with regular traffic. *TelecomTM* achieves a 90.18% two-way traffic detection accuracy, 27× and 5× error rate reductions for vehicle position tracking and speed tracking, respectively, compared to a baseline method without geo-localization, ±3.92% percent error (95% confidence interval) for wheelbase estimation, and ±11.98% percent error (95% confidence interval) for weight estimation.

The main contributions of this work are:

- We introduce the *TelecomTM* system that uses pre-existing telecom fiber responses induced by vehicle vibrations to enable fine-grained and ubiquitous detection, tracking, and characterization of each vehicle.
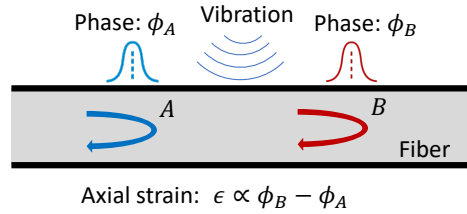
Fig. 1. Illustration of the principle of DAS. A strain perturbation affecting the optical fiber caused by vibrations between A and B produces linearly proportional variations in the phase of the backscattered light.

- We analyze the telecom fiber's dynamic strain responses to address the challenges of high uncertainties in sensor properties and complex noise conditions through a Bayesian analysis approach. Our approach efficiently characterizes the system and integrates the spatial dependency of distributed sensors to improve vehicle detection, tracking, and characterization accuracy.
- We evaluate our system through real-world experiments and characterize the system's performance with various traffic conditions, vehicle types, traveling directions, and speeds.

The rest of the paper is organized as follows: Section 2 describes the physical foundations enabling *TelecomTM*. We introduce our system in Section 3. Our real-world evaluation, its results, and the characterization of our system's performance are described in Section 4. Section 5 discusses related work and the differences between our work and previous research. In Section 6, we conclude our work.

2 PHYSICAL FOUNDATIONS OF TELECOMTM

To provide a background understanding of the *TelecomTM* system, we begin with describing its physical foundations, including the principles of distributed acoustic sensing (DAS) and an exhibition of vehicle-induced telecom fiber vibration.

2.1 Principles of Distributed Acoustic Sensing

DAS based on the ϕ-OTDR technique [46] that uses a standard fiber cable as virtual axial strain sensors. Specifically, an optoelectronic instrument called the interrogator unit repeatedly injects a laser pulse into a fiber cable. An optical interferometry system measures the Rayleigh-backscattered light. The arrival time of the segmented backscattered light can be mapped to the distance along the fiber because the speed of light in the fiber is known. A strain perturbation in the fiber's surroundings may cause a phase shift of the scattering centers. The phase shift is quasi-linearly proportional to the total strain along fiber [12]. Therefore, the strain variations in different fiber sections can be obtained by repeatedly measuring the phase shift. Figure 1 illustrates the principle of DAS. By sending the laser pulses at a high frequency (e.g., 250 Hz for *TelecomTM*), the dynamic strain profile along the fiber can be determined [22].

Besides the sampling rate, there are two important specifications of DAS systems: gauge length and channel spacing. Gauge length is the length over which the phase shifts are measured. Channel spacing is the distance between each virtual sensor. Selecting the gauge length and channel spacing is a trade-off: a longer gauge length has a higher signal-to-noise ratio but a lower spatial resolution and vice versa. A finer channel spacing can improve the spatial resolution but would have more overlapping sensing areas (i.e., requiring larger data storage) for virtual sensors if the channel spacing is smaller than the gauge length. We chose a 10-meter gauge length and a 1-meter channel spacing in our system to obtain fine-grained traffic information with an adequate signal-to-noise
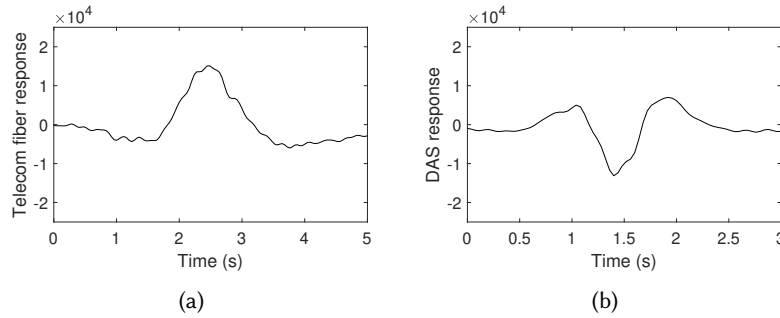
Fig. 2. Vehicle-induced telecom fiber response in two virtual sensors having (a) bell-shaped response and (b) polarity-flipped response.

ratio for accurate vehicle detection. In other words, we convert the telecom fiber cable into virtual sensors spatially distributed every meter to measure the strain over each 10-meter section of the fiber.

2.2 Vehicle-Induced Telecom Fiber Vibration

When a vehicle passes over virtual sensors of the roadside telecom fiber cable, the interaction between the vehicle and the road structure induces the telecom fiber cable to vibrate. The signal pattern of vehicle-induced telecom fiber vibrations depends on the vehicle characteristics, fiber conduit properties, the cable surrounding conditions, etc. There are mainly two components of signals produced by moving vehicles that are recorded by roadside distributed acoustic sensors: 1) quasi-static signals (< 1 Hz) resulting from the ground deformation due to the vehicle's weight, and 2) surface waves (3 to 20 Hz) caused by the dynamic vehicle-road interaction due to the roughness of the road (e.g., bumps). Note that vehicle-induced surface-waves are usually the strongest between 3~20 Hz. From our observation, vehicles don't excite evident 1~3 Hz energy, which is relatively weak unless other sources (e.g., earthquakes) exist. Previous studies [24, 56] have found that the quasi-static component dominates the energy of vehicle-induced telecom fiber vibration and is theoretically described by the Flamant-Boussinesq approximation [8, 10]. As a vehicle approaches the virtual sensor, ground deformation above the sensor increases, and the fiber coupled to the earth is stretched, resulting in increased tension in the fiber. As the vehicle moves away, ground deformation near the virtual sensor and the fiber tension decreases. As a result, the vehicle motion creates a bell-shaped response when it passes a virtual sensor. Figure 2 (a) shows an example of the bell-shaped response of a virtual sensor to a passing car that matches the theoretical telecom response. The signal peaks at around 2.5 second when the car reaches the sensor. Positive amplitude indicates the fiber beneath the car is under tension. Since the quasi-static signal created by a moving vehicle dominates the signal energy and is easy to recognize, traffic volume and speed estimation approaches have been developed by detecting, extracting, and localizing these quasi-static signals with the surface-wave component filtered out [8, 24, 47, 50, 57].

However, we observe that due to the unknown sensor properties and complex noise condition challenges mentioned in Section 1, field data do not always match the theoretical vehicle-induced telecom fiber response. First, moving vehicle locations estimated using quasi-static signals do not match the designed locations of virtual sensors based on the maps/drawings of urban fiber installations. It is because these maps are often inaccurate and do not take into account the slack in the fiber cable accumulated underground (e.g., spools in cabinets and manholes). Figure 3 shows an example photo of fiber spooling in a manhole. There are around 100 meters of fibers spooled in this manhole, resulting in a large estimation error of vehicle positions. We also observe from field data that a polarity-flipped response (Figure 2 b) consistently occurs at some fiber subsections. The polarity flipping phenomenon could be explained by the near-surface heterogeneity and/or fiber conduit property

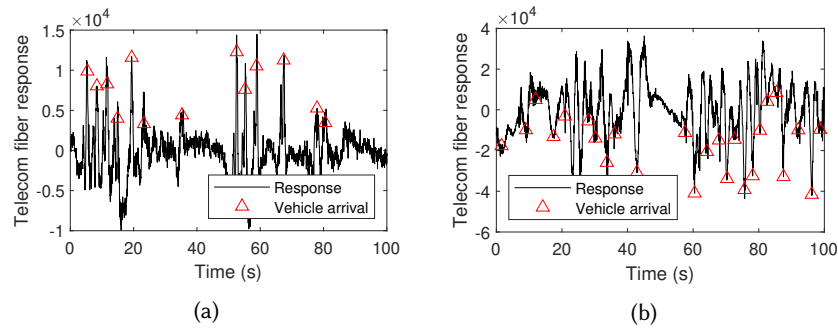Fig. 3. An example photo showing fiber spooling.

(a)

(b)

Fig. 4. Telecom fiber signal examples of (a) bell-shaped and (b) polarity-flipped response. Red triangle markers indicate arrival times of vehicles. Vehicles are detected at local maximums for the bell-shaped response and at local minimums for the polarity-flipped response.

changes that could lead to stress concentration reversing the fiber's response from tension to compression. In addition, telecom fiber responses have high-frequency and irregular signal backgrounds/trends due to noise and non-vehicle-induced perturbations (e.g., environmental changes) in the surroundings of the telecom fiber cable. Figure 4 shows two telecom fiber signals having (a) bell-shaped and (b) polarity-flipped responses. Red triangle markers indicate the arrival of vehicles. We can observe that because of the polarity flipping and irregular signal background phenomenons, conventional peak detection methods, which find local maxima or minima by comparison of neighboring values, or baseline subtraction methods [7], which remove low-frequency or harmonic signal background, may not work well for detecting and tracking vehicles.

To this end, we introduce our *TelecomTM* system in the next section, which geo-localizes every virtual sensor and analyzes their signal patterns (bell-shaped or polarity-flipped response) by matching the vehicle's position signal with the corresponding vehicle-induced quasi-static signal in telecom fiber. *TelecomTM* overcomes the challenge of large and complex noise conditions by fusing spatially dependent vehicle detection results across multiple virtual sensors. It uses Newton's laws and the spatial dependency of vehicle motion to reduce vehicle detection uncertainty and estimate vehicle positions and speeds.
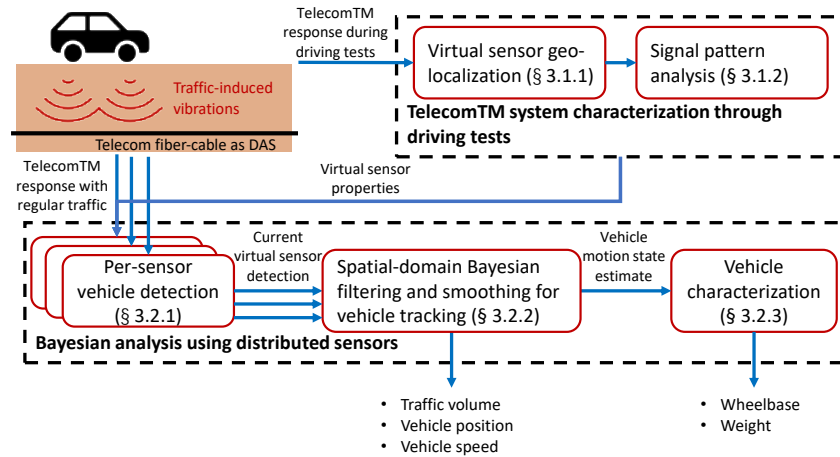
Fig. 5. *TelecomTM* system overview.

3 TELECOMTM: TRAFFIC MONITORING SYSTEM USING PRE-EXISTING ROADSIDE TELECOM FIBER CABLES

The main goal of *TelecomTM* is to develop a fine-grained and ubiquitous traffic monitoring system that (1) detects and tracks vehicle motion and (2) characterizes vehicles (e.g., vehicle position, speed, wheelbase, and weight) from telecom fiber responses.

TelecomTM consists of two modules, as shown in Figure 5. The *TelecomTM* System Characterization Module (Section 3.1) conducts driving tests to estimate the geographic position and analyze the signal pattern of each virtual sensor by matching the testing vehicle's GPS recording with the vehicle-induced quasi-static signal in the telecom fiber. In the second module (Section 3.2), telecom fiber responses are inputted to detect vehicles at each virtual sensor using a prominence-based detection method whose parameters are determined using the virtual sensor properties. Then, our system integrates vehicle detection results from spatially-distributed sensors through Bayesian analysis. We estimate the posterior probability of vehicle arrival time using a spatial-domain Bayesian filtering and smoothing algorithm. Our method combines multiple sensors' information to reduce vehicle detection uncertainties and estimate vehicle motion states (positions and speeds) based on the spatial dependencies of the virtual sensors and Newton's laws of motion. Finally, vehicle characteristics, including weight and wheelbase, are estimated using the vehicle-induced quasi-static signals and the time difference between vehicle wheel-induced responses.

3.1 TelecomTM System Characterization

The first module of our system is to characterize each virtual sensor to estimate their geographic locations and analyze their signal patterns.

3.1.1 Virtual Sensor Geographic Localization. Accurate mapping of the virtual sensors to their geographic locations is essential for all traffic monitoring applications discussed in the paper. We address the unknown geo-location challenge by analyzing the telecom fiber response generated by a car with onboard GPS, which is an easily scalable tool when the fiber cables are installed in proximity of public streets [57] (i.e., roadside). In particular, we first conducted "tap" tests in the interrogator room (underground and without GPS signal) by hitting the fiber for temporal synchronization between the *TelecomTM* system and the onboard GPS receiver.
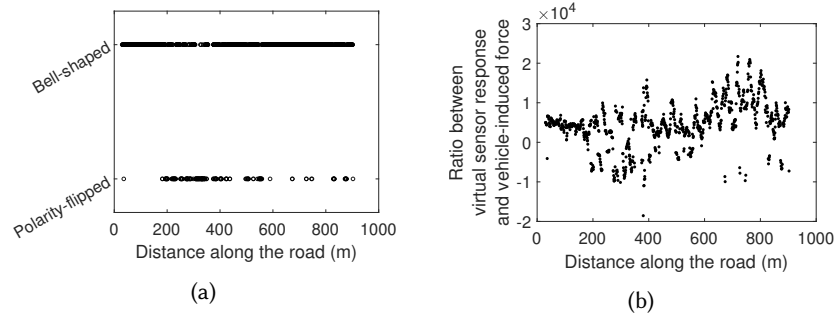
Fig. 6. Spatial-variations of telecom fiber signal characteristics: (a) signal pattern and (b) transmissibility (ratio between virtual sensor response and vehicle's weight). Signal properties of distributed virtual sensors have large spatial variations.

We then drove the testing car on the road under regular traffic. The most predominant quasi-static signal peaks when the car is the closest to a virtual sensor. Based on the peaking time and the onboard GPS recording, we can retrieve the geographic positions of every roadside virtual sensor. Also, to reduce GPS error, we run the driving test several times and average the estimated virtual sensors' geo-locations of all the tests. Note that this characterization process can be done with any city maintenance vehicles, public transit vehicles, mail delivery trucks, etc., which need to go around the city every day. Therefore, it is time- and cost-effective and does not have any effect on regular traffic and telecommunication signals.

3.1.2 Vehicle-induced Telecom Fiber Signal Pattern Analysis. As mentioned in Section 2.2, telecom fiber responses have various signal patterns, which require different methods and model parameters for detecting vehicles at different virtual sensors. For the bell-shaped response, vehicle motion creates prominence of a peak and should be detected at local maximums; for the polarity-flipped response, vehicle motion creates prominence of a valley and should be detected at local minimums. Also, since the ratio between the virtual sensor's response and vehicle-induced force (i.e., transmissibility) is unknown, a universal vehicle detection method with the same parameters (e.g., peak detection threshold) would fail to detect vehicles accurately. The signal pattern and transmissibility are randomly varying along the fiber, and their distributions are difficult to predict. For example, Figure 6 shows the spatially varied signal pattern and transmissibility along an around 900-meter long road section, having 1120 virtual sensors. 923 of them have bell-shaped responses; 197 of them have polarity-flipped responses. The absolute value of transmissibility varies between 264 and 21704 of telecom fiber response magnitude per ton. To address this heterogeneous and unknown signal characteristics challenge, during our driving tests, we also estimate the transmissibility of each virtual sensor. We define the transmissibility, T_k, of the k-th sensor as the ratio between the prominence amplitude of the quasi-static signal and the testing vehicle weight. For bell-shaped responses, $T_k > 0$, and for polarity-flipped responses, $T_k < 0$. In addition, by using the signal of our testing vehicle passing over the fiber on the nearest lane and the physical model that describes the distribution of subsurface stresses (Boussinesq's theory), we can estimate the signal transmissibilities at further lanes without the need for additional driving tests.

3.2 Bayesian Analysis Using Distributed Sensors

The second module of *TelecomTM* is a three-step Bayesian analysis approach. The per-sensor detection step estimates the arrival times of vehicles at each virtual sensor by detecting the prominence of peaks or valleys in time-domain signals. The second step integrates spatially dependent information across various adjacent virtual sensors to reduce false positive detections and estimate vehicle motion states (positions and speeds). The third step estimates the wheelbase lengths and weights of the tracked vehicles.

3.2.1 Per-sensor Vehicle Detection. In this step, we detect vehicles by estimating vehicles' arrival times at each virtual sensor from the time-domain telecom fiber responses. As we discussed in Section 2.2, the quasi-static component due to the vehicle's weight dominates the energy of the signal. Therefore, we use a prominence-based detection method to detect vehicle occurrence. The prominence of a peak measures how much the peak stands out due to its intrinsic height and its relative location to other peaks. We use the prominence-based detection method because it is more robust for detecting local minima or maxima of a signal having high-frequency and irregular signal background. Specifically, we first smooth the data using the locally weighted smoothing (LOESS) [5] with a smoothing span of a one-second window. Here we remove outliers and high-frequency noise that could be incorrectly detected as prominence. In addition, our per-sensor detection method determines the prominence detection threshold of each sensor using its transmissibility. The larger the transmissibility, the larger the prominence threshold. It allows our method to be adaptive to sensors with heterogeneous properties. Specifically, for sensors with $T_k > 0$, the prominence of a peak larger than $r_0 \times (|T_k|/T_0)$ within a one-second window is detected as a vehicle, and for sensors with $T_k \le 0$, prominence of a valley is detected. T_0 is the minimum absolute transmissibility ($T_0 = \min(|T_k|)$, for $k = 1, \cdots, K$). K is the total number of virtual sensors. r_0 is the prominence threshold for vehicle detection at the sensor with the transmissibility of T_0.

3.2.2 Spatial-Domain Bayesian Filtering and Smoothing for Vehicle Tracking. Although an efficient system characterization is conducted through the driving tests mentioned in the first module of *TelecomTM*, the large and complex noise conditions challenge would still cause false or missing detection of vehicles using only the per-sensor detection method. For instance, the noise signals may be incorrectly recognized as vehicle-induced telecom fiber signals or overwhelm the vehicle signals. Wrong or missing vehicle detection results can further affect the accuracy of vehicle tracking and characterization.

To overcome the challenge, we develop a spatial-domain Bayesian filtering and smoothing algorithm that fuses spatial-dependent vehicle detection results of distributed virtual sensors to improve vehicle detection and tracking accuracy. Since vehicle motion is continuous in the spatial domain, our algorithm estimates vehicle motion states based on Newton's laws of motion. In particular, we first formulate our vehicle detection and tracking problem as a spatial-domain state-space model. State-space model is a more compact and convenient representation for multiple-input multiple-output dynamic systems compared to other representations, such as the transfer function model and impulse response function model [20]. In addition, to use the spatial dependency of distributed virtual sensors, our formulation considers the arrival time of a vehicle at every virtual sensor as measurement. It is different from the conventional formulation of Bayesian filtering and smoothing for motion tracking or navigation [39, 44, 59] whose measurement/observation is the position of the object at every timestep and does not take into account the spatial dependency information. The vehicle arrival time at the k-th virtual sensor, t, and its derivative, $\dot{t}$, are described by the linear state space $\mathbf{t}_k = [t, \dot{t}]^T$. The derivative of arrival time, $\dot{t}$, is the time for the vehicle to travel one meter. Based on Newton's laws of motion, the state-space model is

$$
\begin{aligned}
\mathbf{t}_k &= \mathbf{A}_k \mathbf{t}_{k-1} + \mathbf{w}_k \\
z_k &= \mathbf{C}\mathbf{t}_k + v_k
\end{aligned}
\tag{1}
$$

$$
\mathbf{A}_k = \begin{bmatrix} 1 & \Delta x_k \\ 0 & 1 \end{bmatrix}, \ \mathbf{C} = [1, \ 0], \ \mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k), \ v_k \sim \mathcal{N}(0, \sigma_z)
$$

$$
\mathbf{Q}_k = \begin{bmatrix} \frac{1}{4}\Delta x_k^4 & \frac{1}{2}\Delta x_k^3 \\ \frac{1}{2}\Delta x_k^3 & \Delta x_k^2 \end{bmatrix} \sigma_{\ddot{t}}^2,
$$

where Δx_k is the distance between the $(k-1)$-th and the k-th virtual sensor projected on the road; z_k is the observed vehicle arrival time at virtual sensor k; $\mathbf{w}_k$ is the process noise; $\mathbf{v}_k$ is the observation noise; and $\sigma_{\ddot{t}}$ and σ_z are the standard deviations of arrival time's second derivative and that of the measurement noise, respectively.
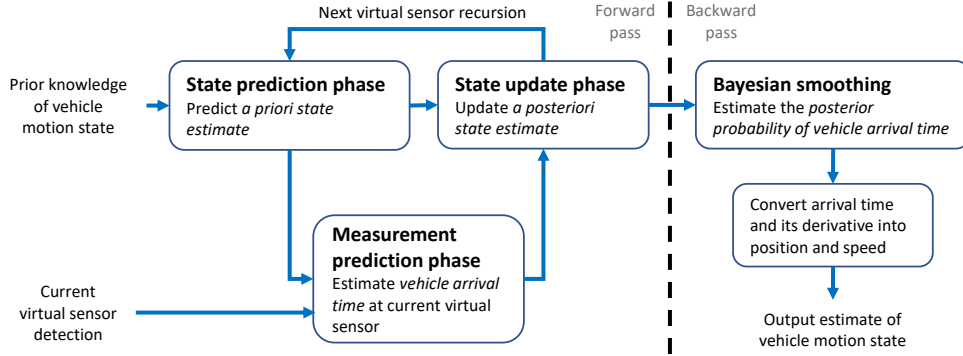
Fig. 7. Our Bayesian filtering and smoothing algorithm.

To estimate the vehicle motion state (described in the above state-space model) from multiple sensors' noisy data, we develop a Bayesian filtering and smoothing algorithm (as shown in Figure 7 and Algorithm 1). It consists of a forward filtering pass and a backward smoothing pass. The forward pass estimates the filtered posterior probability of the state estimate recursively using past observations; the backward pass computes the smoothed posterior probability using all observations [3, 44]. Specifically, the forward pass of our algorithm estimates the posterior probability of vehicle arrival time over space using vehicle detection results of previous and current virtual sensors: $p(\mathbf{t}_k|z_{1:k})$. The forward pass has three phases: 1) state prediction phase, 2) measurement prediction phase, and 3) state update phase. The state prediction phase uses the state estimate from the previous virtual sensors to produce the state estimate at the current virtual sensor. Since we don't have direct measurement/observation (i.e., the arrival time of the vehicle) at each virtual sensor, our algorithm has a measurement prediction phase that predicts the measurement, z_k, by finding the detection in the current virtual sensor having the largest probability of the predicted arrival time:

$$\hat{z}_k = \arg\max_{z \in \mathbf{D}_k} f(z; \hat{t}_{k|k-1}, \sigma_{t,k|k-1}) \tag{2}$$

where $\mathbf{D}_k$ are per-sensor vehicle detection results for the virtual sensor k; $f(z)$ is the probability density function of the predicted arrival time following the Gaussian distribution: $\mathcal{N}(z; \hat{t}_{k|k-1}, \sigma_{t,k|k-1}^2)$; $\hat{t}_{k|k-1}$ and $\sigma_{t,k|k-1}$ are the predicted arrival time estimate and its standard deviation, respectively. In other words, our algorithm tracks the vehicle arriving at each virtual sensor by finding the closest vehicle detection to the predicted arrival time estimate. Then, the state update phase updates the state estimate (i.e., a posteriori state estimate) using a weighted average of the predicted state estimates and the measurement. The backward pass of our algorithm adopts the Rauch-Tung-Striebel smoother [39] to estimate the posterior probability of vehicle arrival time using vehicle detection results at all virtual sensors: $p(\mathbf{t}_k|z_{1:K})$. Here our algorithm reduces detection uncertainties by utilizing the spatial dependency of the distributed sensors. Finally, vehicle motion state estimation or tracking is achieved by converting arrival time and its derivative at every virtual sensor, $[t, \dot{t}]$, into vehicle position and speed, $[x, v]$ using the virtual sensors' geographic locations estimated in our module one.

3.2.3 Vehicle Characterization. TelecomTM estimates vehicle wheelbase lengths and weights using the vehicle speed estimation results and the quasi-static component of the vehicle-induced signals. To estimate the wheelbase length, we use the high-frequency telecom fiber response (≥3 Hz) created by a vehicle's two wheels passing a bump, joint, or pothole. It creates clear impulse responses due to the interaction between the vehicle wheels and the abrupt change in road profiles. Figure 8 shows an example of a virtual sensor response and its highpass filtered signal (≥3 Hz) at a bridge-road joint. We can observe the vibration components created by the two wheels

---

Algorithm 1 Spatial-domain Bayesian filtering and smoothing

Input: Vehicle detection results at every virtual sensor, $\mathbf{D}_k$, for $k = 1, \cdots, K$; Prior knowledge of standard deviations, $\sigma_{\tilde{t}}$ and
    σ_z

Output: Posterior probability of vehicle state, $p(\mathbf{t}_k|z_{1:K})$

1: #*Forward filtering pass:*
2: **Initialization**: initial guess of state estimate and estimate covariance, $\mathbf{t}_{0|0}$ and $\mathbf{P}_{0|0}$
3: **for** Virtual sensor $k = 1, \cdots, K$ **do**
4:     #*State prediction phase:*
5:     Calculate the predicted state estimate,
$$\hat{\mathbf{t}}_{k|k-1} = \mathbf{A}_k\hat{\mathbf{t}}_{k-1|k-1}$$
6:     Calculate the predicated state covariance,
$$\mathbf{P}_{k|k-1} = \mathbf{A}_k\mathbf{P}_{k-1|k-1}\mathbf{A}_k^T + \mathbf{Q}_k$$
7:     Predict the state probability given previous measurement,
$$p(\mathbf{t}_k|z_{1:k-1}) = \mathcal{N}(\mathbf{t}_k; \hat{\mathbf{t}}_{k|k-1}, \mathbf{P}_{k|k-1})$$
8:     #*Measurement prediction phase:*
9:     Find the vehicle detection of the k-th virtual sensor having the largest probability of the predicted state using Equation (2)
10:     #*State update phase:*
11:     Calculate the updated state estimate,
$$\hat{\mathbf{t}}_{k|k} = \hat{\mathbf{t}}_{k|k-1} + \mathbf{P}_{k|k-1}\mathbf{C}^T$$
$$\times (\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \sigma_z)^{-1}(z - \mathbf{C}\hat{\mathbf{t}}_{k|k-1})$$
12:     Calculate the updated state covariance,
$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \sigma_z)^{-1}\mathbf{C})\mathbf{P}_{k|k-1}$$
13:     Update the state probability given previous and current measurement,
$$p(\mathbf{t}_k|z_{1:k}) = \mathcal{N}(\mathbf{t}_k; \hat{\mathbf{t}}_{k|k}, \mathbf{P}_{k|k})$$
14: **end for**
15: #*Backward smoothing pass:*
16: **Initialization**: Smoothed state estimate and covariance at the last virtual sensor, $\hat{\mathbf{t}}_{K|1:K} = \hat{\mathbf{t}}_{K|K}$, $\mathbf{P}_{K|1:K} = \mathbf{P}_{K|K}$
17: **for** Virtual sensor $k = K - 1, \cdots, 1$ **do**
18:     Calculate the smoothed state estimate,
$$\hat{\mathbf{t}}_{k|1:K} = \hat{\mathbf{t}}_{k|k} + \mathbf{P}_{k|k}\mathbf{A}_{k+1}^T\mathbf{P}_{k+1|k}^{-1}(\hat{\mathbf{t}}_{k+1|1:K} - \hat{\mathbf{t}}_{k+1|k})$$
19:     Calculate the smoothed state covariance,
$$\mathbf{P}_{k|1:K} = \mathbf{P}_{k|k} + \mathbf{P}_{k|k}\mathbf{A}_{k+1}^T\mathbf{P}_{k+1|k}^{-1}$$
$$\times (\mathbf{P}_{k+1|1:K} - \mathbf{P}_{k+1|k})(\mathbf{P}_{k|k}\mathbf{A}_{k+1}^T\mathbf{P}_{k+1|k}^{-1})^T$$
20:     Estimate the posterior state probability given all sensors' measurement,
$$p(\mathbf{t}_k|z_{1:K}) = \mathcal{N}(\mathbf{t}_k; \hat{\mathbf{t}}_{k|1:K}, \mathbf{P}_{k|1:K})$$
21: **end for**

---

have a repeating impulse response pattern. The time difference between the two wheel-induced responses can be estimated by using the auto-correlation function to find the repeating patterns. Then, we can estimate the wheelbase length by multiplying the time difference with the vehicle speed, $\tau_m \times v$, where $\tau_m = \arg\max_\tau R_{zz}(\tau)$ is the time difference between wheel-induced telecom fiber response; $R_{zz}(\tau)$ is the auto-correlation function at
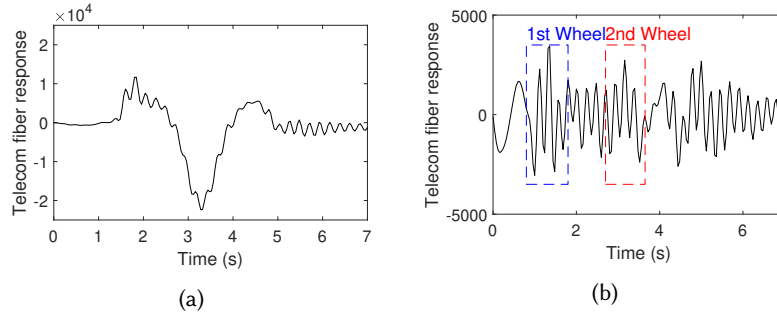
Fig. 8. (a) Vehicle-induced telecom fiber responses and (b) its highpass filtered response at a bridge-road joint. Vehicle wheel-induced vibration responses have a repeating impulse response pattern.
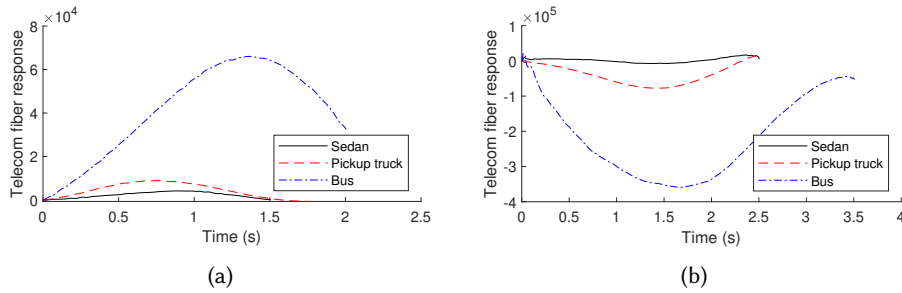

Fig. 9. Quasi-static signals of (a) bell-shaped and (b) polarity-flipped responses for three different types of vehicle. Heavier vehicle creates larger prominence amplitude.

lag τ of the highpass filtered response of the vehicle-induced telecom fiber vibration; and v is the moving speed of the vehicle.

Furthermore, by assuming the linear elasticity of the road and the near-surface structure, the prominence amplitude of the quasi-static signal is approximately proportional to the moving vehicle's weight. Figure 9 shows the telecom fiber responses having bell-shaped and polarity-flipped signal patterns created by three different types of vehicles moving in the same lane and measured by the same virtual sensor. We can observe that the heavier the vehicle, the larger the prominence amplitude. Therefore, we estimate the moving vehicle's weight by computing the weighted average of quasi-static signal prominence, $\frac{1}{K} \sum_{k=1}^{K} (P_k/|T_k|)$, where P_k is the prominence amplitude of the vehicle's quasi-static signal at virtual sensor k.

4 REAL-WORLD EVALUATION

In this section, we describe the experimental setup and results evaluating the performance of *TelecomTM*, followed by a characterization of the system's performance.

4.1 Experimental Setup

We evaluate *TelecomTM* on a two-way two-lane road with pre-existing telecom fiber cables in San Jose, California. *TelecomTM* responses were collected using an OptaSense QuantX interrogator [36] (as shown in Figure 10 a) at a 250 Hz sampling rate, 10-meter gauge length, and 1-meter virtual sensor spacing. The interrogator recorded dynamic strain responses from a roadside telecom fiber cable (an example photo of fiber conduits is shown in Figure 10 b) along an around 900-meter road section of the East Julian St, San Jose, CA. Figure 11 shows the
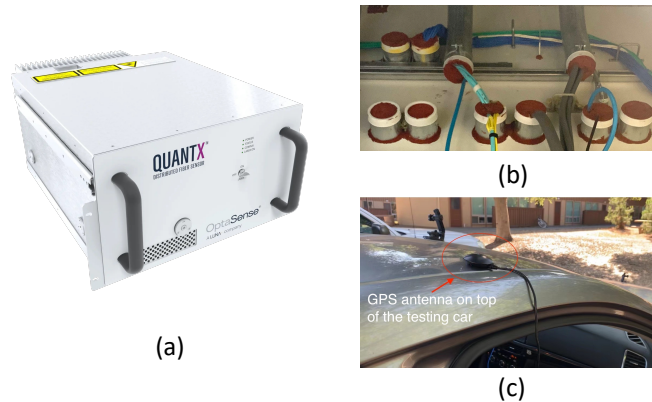
Fig. 10. Experimental setup: (a) the QUANTX interrogator [36], (b) an example photo of fiber conduits, and (c) GPS receiver on the testing car.
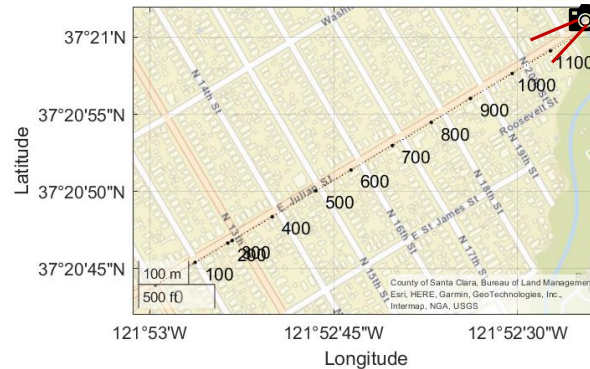

Fig. 11. Locations of virtual sensors. The camera icon indicates the location of a camera. The dot line and numbers indicate virtual sensors' locations

location of the telecom fiber cable that has 1120 virtual sensors. The outbound traffic (with virtual sensor number increasing in Figure 11) travels away from the interrogator, and is closer to the telecom fiber cable than inbound direction traffic.

To estimate geographic locations of virtual sensors and provide ground-truth of vehicle positions, we installed a GPS receiver that records its geographical position every second on a testing sedan vehicle (1.47 tons weight), as shown in Figure 10 (c). We also logged the vehicle speed at a 100 Hz sampling rate through the Controller Area Network (CAN) bus. In addition, a camera was placed close to the last virtual sensor (as shown in Figure 11) to acquire ground-truth information on vehicle arrival times and vehicle models. We obtain wheelbase lengths and weights of vehicles with the vehicle model information. With the supervision of staff from the City of San Jose, we conducted two daytime experiments with 388 vehicles recorded in a 31-minute video and seven nighttime experiments with our testing vehicle running through the testing road.

Our system was evaluated for vehicle detection, tracking, and wheelbase and weight estimations. It achieves overall 90.18% accuracy for vehicle detection, 27× error rate reduction (average mean absolute error (MAE) reduced
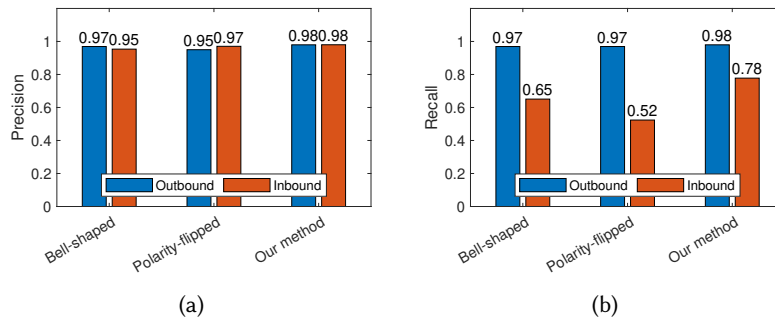
Fig. 12. (a) Precision and (b) recall of detection results using only bell-shaped responses, polarity-flipped responses, or using our method with all distributed sensors. Our method improves vehicle detection recall by fusing spatial dependency information from distributed virtual sensors.

from 140.09 m to 5.16 m) for vehicle position estimation, 5× error rate reduction (average MAE reduced from 17.18 km/h to 3.57 km/h) for vehicle speed estimation compared to a baseline method that does not geo-localize the virtual sensors through the driving test. The baseline method directly applies our spatial-domain Bayesian analysis algorithm introduced in Section 3.2 to the telecom cable responses. In addition, *TelecomTM* achieves ±3.92% accuracy for wheelbase estimation, and ±11.98% percent error for weight estimation. The performance of *TelecomTM* was found to remain consistent during evaluations conducted on weekdays and weekends, as well as during daytime and nighttime, despite variations in temperature. However, the evaluations were limited to sunny weather conditions only. We discuss our results in detail in the following subsections.

4.2 Vehicle Detection Results

We compare the vehicle detection results using *TelecomTM* with the ground truths captured using the camera. Figure 12 shows the average precision and recall of detection results using only bell-shaped responses, polarity-flipped responses, or using our *TelecomTM* method. All three cases achieve ≥ 95% precision for both inbound and outbound directions and ≥ 95% recall for the outbound direction. Recall for the inbound direction is lower than that for the outbound direction mainly due to the long distance from the inbound traffic to the telecom fiber cable, which will be discussed in Section 4.5.1. The average recall for vehicle detection using the bell-shaped response is 13% higher than that using the polarity-flipped response. It is because there are multiple factors causing the polarity-flipping phenomenon, which has high uncertainties, resulting in the distinct signal patterns of the polarity-flipped responses across various virtual sensors. Our *TelecomTM* method has been successful in addressing this issue. It has a 1.6× improvement in error rate reduction (false negative rate reduced from 35% to 22%) by considering spatially dependent vehicle detection information of distributed virtual sensors.

4.3 Vehicle Tracking Results

We evaluated the effectiveness of *TelecomTM* for vehicle tracking by comparing the estimated travel distance and speed with onboard GPS and CAN bus recordings. Figure 13 shows the telecom fiber response to the testing vehicle plotted with the y-axis being the fiber distance to virtual sensors (i.e., before geo-localization) or the distance along the road (i.e., after geo-localization). It can be observed that there is an around 100-meter gap between virtual sensors 200 and 300. The gap in the data indicates the lack of fiber-soil coupling due to fiber spooling, which can be corrected after geo-localization, as shown in Figure 13 (b). Figure 14 shows the ground truth of vehicle location and speed and their estimations before and after geo-localization. Here we can observe the big difference between ground truth and the estimations without geo-localization. To quantify the error,

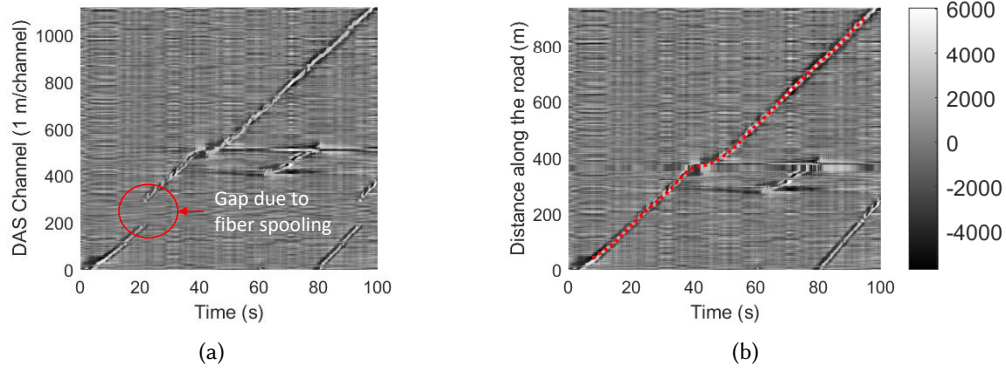(a)   (b)

Fig. 13. Telecom fiber response with the y-axis being (a) the fiber distance to virtual sensors or (b) the distance along the road after geo-localization. Red dot curve in (b) shows our vehicle tracking result. Our method removed the signal gap (red circle) due to fiber spooling and estimated the geographic locations of virtual sensors by matching the testing vehicle signal and the corresponding telecom fiber response.

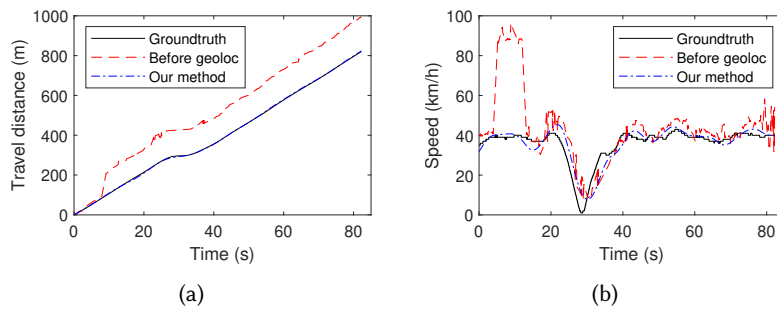(a)   (b)

Fig. 14. Vehicle (a) locations and (b) speeds ground truth (black curve) and estimations before (blue curve) and after (red curve) geo-localization (geoloc).

(a)   (b)
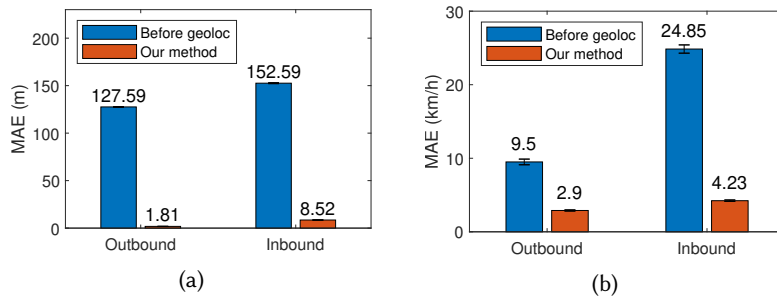
Fig. 15. Mean absolute value bar chart with a 95% confidence interval for vehicle (a) location and (b) speed estimations before and after geo-localization (geoloc). The number above each bar indicates the mean absolute value. Our method improves the vehicle location and speed estimation accuracy by geo-localizing each virtual sensor.

Figure 15 shows the mean absolute errors between ground truth and estimations of vehicle location and speed before and after geo-localization. Our method has a 70× and an 18× error rate reduction for the outbound and
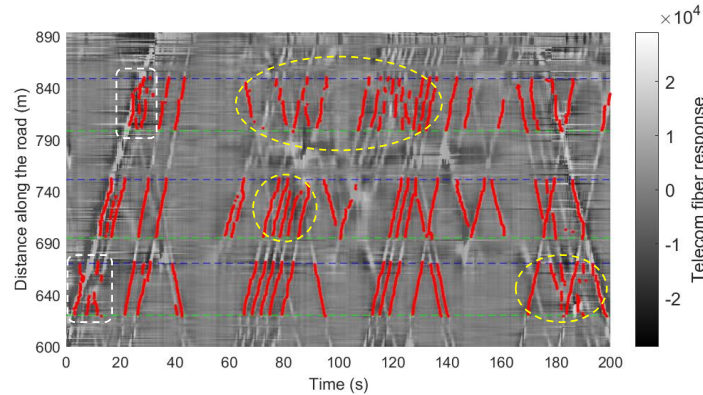
Fig. 16. Multiple vehicle tracking. Red curves indicate vehicles tracked by our algorithm. Horizontal dash lines segment road sections by intersections. Dashed line boxes and circles are two types of tracking errors identified and discussed in Section 4.3.

inbound direction vehicle location estimations, respectively; and a 3× and a 6× error rate reduction for the two-direction vehicle speed estimations, respectively.

TelecomTM performs well for tracking multiple vehicles with daily and busy traffic conditions. Figure 16 shows an example of our multiple vehicle tracking results. Since vehicle motions in intersections are complex and easy to be overlapping, we segment the road into building blocks divided by intersections along the road, which are indicated by the horizontal dash lines in the figure. We successfully tracked the two-way traffic with various traffic patterns, including multiple vehicles traveling in opposite directions and trains of vehicles passing by in the same direction. Here we also want to discuss two main types of inaccurate tracking results and the reasons causing them. The first type is caused by heavy vehicles marked by white dashed line boxes. When a heavy car is followed by several light vehicles moving across the instrumented road sections, a large quasi-static signal caused by the heavy vehicle masks the quasi-static signals of nearby vehicles, making light vehicles' tracking difficult. We also identified the second type of inaccurate tracking marked by the yellow circles. When multiple vehicles traveling from two directions meet, quasi-static signals created by them have overlapping and more complicated superpositions, causing tracking errors. Especially, there are more missing tracking results in the inbound direction as the traffic is further away from the telecom fiber cable compared to the outbound direction.

4.4 Vehicle Characterization Results

We evaluated *TelecomTM*'s vehicle characterization performance by comparing our wheelbase and weight estimations with vehicles' specifications. We identified the vehicle models using camera images and searched online for their specifications. Figure 17 shows the high-frequency ($> 3Hz$) responses of three different types of vehicles, a sedan, a pickup truck, and a bus. It can be observed that with a similar moving speed, the larger the vehicle size, the longer the time difference between the wheel-induced vibrations. Overall, our system achieves a ±3.92% percent error (95% confidence interval) for wheelbase length estimation, which has a 2× improvement in error rate reduction (error reduced from ±8% to ±3.92%) compared to the commercial piezoelectric pavement sensing system. Furthermore, since the actual weights of vehicles were not allowed to be measured during our experiments, we compared our weight estimation with the vehicles' curb weight which is the weight of the vehicle, including a full tank of fuel and all standard equipment. Although it may not be equal to the actual weight of the moving vehicle, it could reflect our estimation accuracy because there were no heavy trucks during
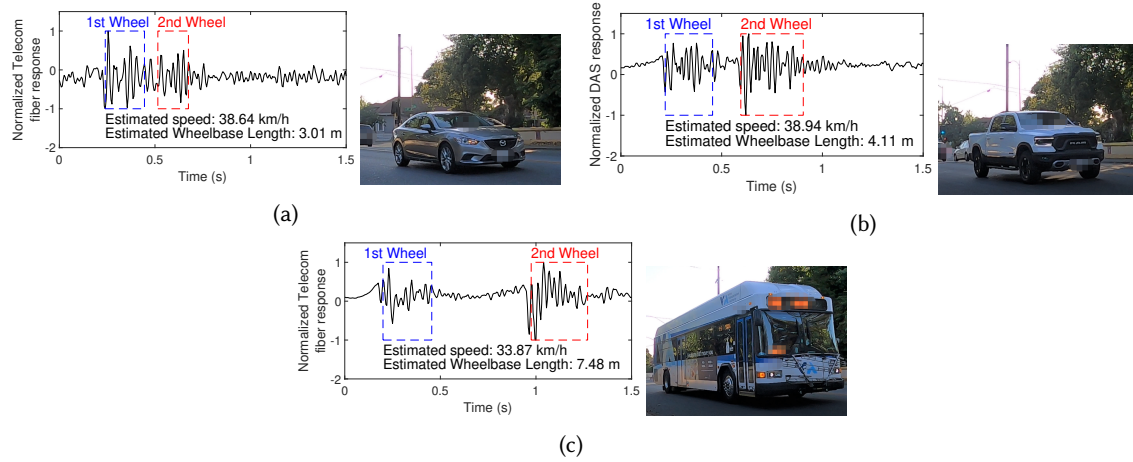
(a)

(b)

(c)

Fig. 17. High-frequency responses (≥ 3 Hz) and the estimated wheelbase lengths of (a) a sedan, (b) a pickup truck, and (c) a bus. Our system estimates wheelbase length by estimating the vehicle speed and time difference between the wheel-induced vibration responses.

the experiments, and the payloads of the recorded vehicles are considered much smaller than their curb weights. Overall, our system achieves a ±11.98% percent error (95% confidence interval) for weight estimation, which has a 3% improvement in error rate reduction compared to the commercial piezoelectric pavement sensing systems.

4.5 System Characterization

In this subsection, we characterize our *TelecomTM* system's performance by examining the effects of sensing distance, vehicle types, crosstalking event, and vehicle moving speed.

4.5.1 Effect of Sensing Distance. Vehicle-induced vibrations attenuate as the distance between the vehicle and fiber cable increases. Traffic in the lane farther from the telecom fiber cable induces smaller responses and has a lower signal-to-noise ratio. To study the effect of sensing distance on vehicle detection and tracking performance, we computed the vehicle detection accuracy, location and speed estimation results for different traffic directions. The outbound traffic travels away from the interrogator, and it is closer to the telecom fiber cable than inbound direction traffic. During our experiments, we recorded 182 inbound vehicles and 206 outbound vehicles. As shown in Figure 12, detecting vehicles that travel inbound and outbound has similar precision but different recall. The recall for inbound vehicle detection has an up to 45% accuracy reduction rate compared to that for outbound vehicle detection. Also, Figure 15 shows the MAE of vehicle location and speed estimations for different traffic directions. The location and speed estimations for outbound traffic have a 4.7× and a 1.5× improvement in error rate reduction compared to that for inbound traffic. We found that due to the signal-to-noise ratio change, the closer the distance between the traffic and the telecom fiber cable, the more accurate the vehicle detection and tracking performance. To improve the accuracy for detecting and tracking vehicles in further traffic lanes or roads, we can de-noise the vehicle-induced telecom cable responses to reduce interference among closely traveling vehicles. For instance, works in [50, 55] de-convolve the simulated quasi-static strain response induced by the vehicle from the telecom cable responses to compress the signals into sharp pulses and remove the background noises.

4.5.2 Effect of Vehicle Types. Vehicles with different sizes, shapes, and weights induce different vibration signals in virtual sensors. We categorize vehicles in our experiments into four types based on their size: 223
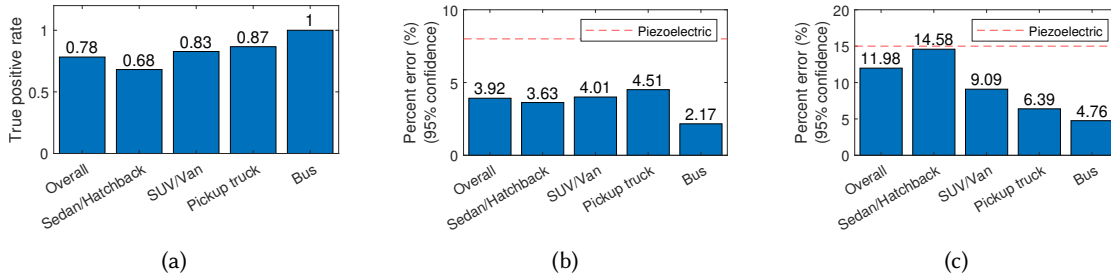
(a)   (b)   (c)

Fig. 18. Effect of vehicle types. (a) shows the true positive rate (TPR) of vehicle detection for different vehicle types. The larger the size of a vehicle, the better the vehicle detection result. (b) and (c) show the percent error (95% confidence interval) for wheelbase and vehicle weight estimations. Red dash lines indicate the percent error of current commercial piezoelectric sensors.
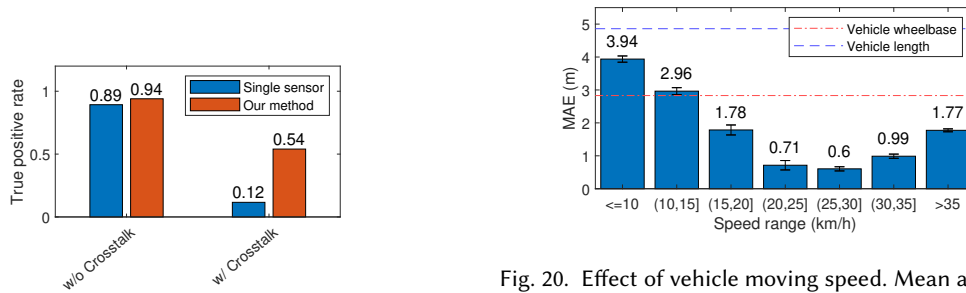

Fig. 19. Effect of crosstalking events. True positive rate of vehicle detection with or without crosstalking using a single sensor or our method with distributed sensors.

Fig. 20. Effect of vehicle moving speed. Mean absolute value bar chart with a 95% confidence interval for vehicle location estimation with different speed ranges. The number above each bar indicates the mean absolute value. Red and blue dash lines indicate vehicle wheelbase and length, respectively. Lower and higher speed ranges have larger location estimation errors.

sedans/hatchbacks, 114 SUVs/vans, 57 pickup trucks, and 4 buses. Figure 18 (a) shows the true positive rate (recall) of vehicle detection for different vehicle types moving in the inbound direction. We can observe that the larger the vehicle size, the better the vehicle detection accuracy. It is because larger and heavier vehicles create larger quasi-static responses and have a higher signal-to-noise ratio.

Furthermore, we also studied the effect of vehicle types on vehicle characterization performance. Figure 18 (b) and (c) show the percent error (95% confidence interval) of wheelbase length and vehicle weight estimations for different vehicle types. We observe that our system has similar wheelbase estimation errors for different vehicle types. For vehicle weight estimation, we found that the estimation accuracy increases as the vehicle size and weight increase since heavier vehicles at the same location would induce larger telecom cable response (i.e., higher signal-to-noise ratios).

4.5.3 Effect of Crosstalking Events. TelecomTM monitors traffic with one-dimensional responses, the axial dynamic strain of the telecom fiber. Telecom fiber responses induced by multiple vehicles from different lanes passing a virtual sensor at the same time overlap and would be detected as one vehicle event. We define this overlapping as the crosstalking event that mainly affects the inbound (longer sensing distance) vehicle detection performance. During our experiments, 40% of inbound traffic are crosstalking events. Figure 19 shows the true positive rate of vehicle detection with or without the crosstalking effect using a single sensor or using our *TelecomTM* method.
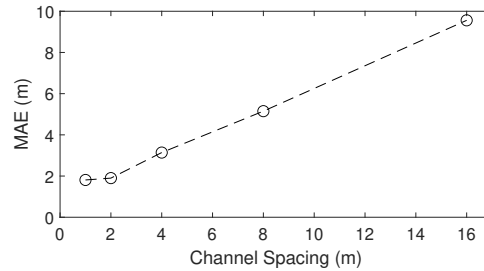
Fig. 21. Effect of channel spacing. Multiple vehicle tracking. Mean absolute errors of vehicle location estimation with different channel spacing values.

We can observe that for detection using a single sensor, crosstalking events reduce the true positive rate by 77%. By leveraging the results from multiple adjacent virtual sensors, we mitigate the issue and improve the TPR from 12% to 54% for detecting vehicles having crosstalking.

4.5.4 Effect of Vehicle Moving Speed. We studied the effect of vehicle moving speed by investigating the vehicle location estimation errors with respect to different speed ranges, as shown in Figure 20. Here we see that lower and higher speeds have large MAE for vehicle location estimation. Vehicle moving at speed between 25 and 30 km/h has the smallest MAE, 0.6 m. The larger error for tracking slower vehicles would be because, with the 10-meter gauge length, the vehicle-induced vibration signals are smoothed out, which affects the sensitivity of peak detection-based arrival time estimation for locating slower vehicles. The larger error for tracking faster vehicles would be because these vehicles create larger dynamic responses that reduce the signal-to-noise ratio. Horizontal dash lines in the figure indicate the wheelbase and total length of the testing sedan. Here we see that the location estimations for all the speed ranges have smaller errors than the vehicle length, and the estimations for the speed > 15 km/h have smaller errors than the wheelbase length. Our method has enough location estimation resolution for tracking small vehicles, such as sedans, and for estimating wheelbases with high vehicle speed.

4.5.5 Effect of Channel Spacing. We studied the effect of DAS channel spacing by calculating the mean absolute errors of vehicle location estimation with different channel spacing values, as shown in Figure 21. We can observe that the vehicle tracking error increases as the channel spacing increases. In our work, we first set the channel spacing to be 1 meter, which is the finest spatial interval we can achieve. Then, we chose the gauge length that makes sure vehicle-induced telecom vibrations are clearly visible in the DAS signals.

5 RELATED WORK

This section provides a review of related works in roadway traffic monitoring, including the state-of-the-art of traffic sensing technologies and existing studies for traffic monitoring using pre-existing telecom fiber cables.

5.1 The State-of-the-art of Traffic Sensing Technologies

Much work has been done on traffic monitoring with sensing technologies, including traffic volume estimation, speed estimation, vehicle classification, and weight estimation. The most common technologies are pavement sensing systems, vision-based sensing systems [6, 40, 42], and crowd-sensing systems [17, 43, 61]. Pavement sensing systems, such as inductive loops [4, 16, 18], piezoelectric sensors [16, 23, 60], and fiber optic sensors [49, 58], capture traffic information through measuring the changes on the pavement, such as inductance of the coil, pressure fluctuation, vibration, etc. Although such pavement sensing systems have been used for various traffic

monitoring applications, their installation and maintenance require complicated field works that interrupt regular traffic [2]. On the other hand, installing cameras to monitor road networks is cheaper and less disruptive than pavement sensing systems. Also, many cameras are already installed on roadways for surveillance purposes [45]. However, such vision-based systems have line-of-sight restrictions, are sensitive to weather conditions, and bring privacy concerns. Crowd-sensing approaches that use cell phone data or Vehicle-to-everything (V2X) data to enable low-cost and efficient traffic monitoring by detecting the drivers'/passengers' phones or the vehicles in motion. However, such crowd-sensing systems cannot provide vehicle characterization information and raise privacy concerns. In summary, the drawbacks of existing traffic monitoring systems hamper their effectiveness in achieving fine-grained and ubiquitous traffic monitoring.

5.2 Traffic Monitoring Using Pre-existing Telecom Fiber Cables

Existing studies have explored traffic monitoring using telecom fiber cable in three main aspects, including 1) human mobility characterization, 2) vehicle detection and counting, and preliminary work in 3) vehicle speed estimation.

For human mobility characterization, previous studies have demonstrated the strong correlation between the telecom fiber signal variations and the intensity of human activities, including footsteps, traffic, and construction activities [24, 47]. When there are fewer activities on the sensing site (e.g., during the COVID-19 or spring break of a school), a decrease of amplitude in low-frequency signals is observed. These studies validate the trend by comparing the ground truth mobility data with the seismic noise level (indicated by the signal energy in low-frequency bands) from telecom fiber signals.

For vehicle detection and counting, previous studies have introduced template-matching algorithm and deep deconvolution models to capture the telecom fiber signal variations induced by drive-by vehicles [8, 24, 50, 52, 58]. These models capture the amplitude changes in the signals to predict the time when the vehicle presents. The template-matching algorithm detects vehicles by comparing the signal with the standard pattern of a signal as the vehicle occurs. The deep deconvolution model deconvolves the vehicle impulse responses from the quasi-static distributed acoustic sensing recordings. Then, a beamforming algorithm is applied to the deconvolved signal rather than the original DAS signal shows improvements in terms of the resolution in vehicle speed estimation, and the detection accuracy.

In addition, preliminary work for vehicle speed estimation estimates the speed of a single-vehicle based on the time interval between two adjacent locations of the same detected vehicle [57]. This preliminary work characterizes the traffic-induced surface waves recorded by telecom fiber and compares it with vehicle onboard sensors to validate the vehicle observation in telecom fiber signals.

While the previous work has validated the feasibility of using pre-existing telecom fiber-optic cables to monitor traffic, these methods are exploratory and mainly focus on coarse-grained traffic flow and speed estimation. In our work, we introduce a systematic approach for traffic monitoring using telecom fiber, which provides a quantitative analysis of traffic data for accurate and fine-grained vehicle detection, tracking, and characterization.

6 CONCLUSION AND FUTURE WORKS

In conclusion, we introduced the *TelecomTM* system that uses pre-existing roadside telecom fiber cables to achieve fine-grained and ubiquitous traffic monitoring at low cost with low maintenance. *TelecomTM* uses the distributed acoustic sensing technique to turn telecom fiber cables into virtual strain sensors in a meter-scale spatial resolution. It senses vehicle-induced near-surface vibrations to detect and track vehicles and estimate vehicle positions, speeds, wheelbase lengths, and weights. To overcome the unknown and heterogeneous sensor properties challenge of using this non-dedicated sensing system, we first estimated the geographic location and analyzed the signal pattern of each virtual sensor by matching a testing vehicle's position and quasi-static

signals in the telecom fiber induced by its motion. Further, to overcome the challenge of large and complex noise conditions, we developed a spatial-domain Bayesian filtering and smoothing algorithm that estimates the posterior probability of vehicle arrival time at each virtual sensor. Our system can accurately track vehicle motion by converting the estimated arrival time into vehicle positions and speeds. Vehicle wheelbase and weight are estimated by analyzing the time difference between vehicle wheel-induced responses and the quasi-static strain.

We evaluated our system through real-world experiments and extensively characterized its performance with different traffic conditions, vehicle types, sensing distance, and vehicle speeds. Our system detects two-way traffic with 90.18% accuracy, tracks vehicles with 5.16-meter position estimation error and 3.57 km/h speed estimation error, estimates vehicle wheelbase length with ±3.92% percent error, and estimates vehicle weight with ±11.98% percent error.

Our vision of future works to generalize TelecomTM to multiple roadways with various environmental and operational conditions is

1) **Understanding Environmental Influences:** Environmental influences, including the temperature, humidity, and human activity disturbances on the road surfaces, vary over time and significantly affect the resultant telecom fiber responses. To understand how these environmental factors influence traffic monitoring performance, we plan to collect data under various situations in the long term. Further, we will explore methods to reduce these influences to improve the robustness of our *TelecomTM* system.

2) **Characterizing Noise and Uncertainties:** The noise and uncertainties in the telecom fiber responses come from various sources, including different types of conduits, coupling conditions, and soil properties around the cable. To further improve the signal quality and prediction accuracy, we plan to characterize the noise and uncertainties caused by each source by conducting controlled experiments with different source combinations.

3) **Exploring Complex Roadways:** In real-life scenarios, the roadways may have more complex configurations than that of our evaluation site. Such configurations include curved and sloped road sections, multiple lanes, etc. For future work, we will explore the capability of our system on these roadways and develop solutions for more complex traffic settings.

ACKNOWLEDGMENTS

REFERENCES

[1] Jonathan B Ajo-Franklin, Shan Dou, Nathaniel J Lindsey, Inder Monga, Chris Tracy, Michelle Robertson, Veronica Rodriguez Tribaldos, Craig Ulrich, Barry Freifeld, Thomas Daley, et al. 2019. Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection. *Scientific reports* 9, 1 (2019), 1–14.

[2] Marcin Bernas, Bartłomiej Płaczek, Wojciech Korski, Piotr Loska, Jarosław Smyła, and Piotr Szymała. 2018. A survey and comparison of low-cost sensing technologies for road traffic monitoring. *Sensors* 18, 10 (2018), 3243.

[3] Arthur E Bryson and Yu-Chi Ho. 2018. *Applied optimal control: optimization, estimation, and control.* Routledge.

[4] Tom Cherrett, Hugh Bell, and Mike McDonald. 2000. Traffic Management Parameters from Single Inductive Loop Detectors. *Transportation Research Record* 1719, 1 (2000), 112–120. https://doi.org/10.3141/1719-14

[5] William S Cleveland and Susan J Devlin. 1988. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American statistical association* 83, 403 (1988), 596–610.

[6] Sokèmi René Emmanuel Datondji, Yohan Dupuis, Peggy Subirats, and Pascal Vasseur. 2016. A Survey of Vision-Based Traffic Monitoring of Road Intersections. *IEEE Transactions on Intelligent Transportation Systems* 17, 10 (2016), 2681–2698. https://doi.org/10.1109/TITS.2016.2530146

[7] Shireen Y Elhabian, Khaled M El-Sayed, and Sumaya H Ahmed. 2008. Moving object detection in spatial domain using background removal techniques-state-of-art. *Recent patents on computer science* 1, 1 (2008), 32–54.

[8] Martijn Van Den Ende, André Ferrari, Anthony Sladen, and Cédric Richard. 2021. Deep Deconvolution for Traffic Analysis with Distributed Acoustic Sensing Data. (2021).

[9] Federal Highway Administration. 2001. Traffic Monitoring Guide. (2001).

[10] YC Fung. 1965. Foundations of solid mechanics Prentice-Hall. *Inc, New Jersey* (1965).

[11] Amnesh Goel, Sukanya Ray, and Nidhi Chandra. 2012. Intelligent traffic light system to prioritized emergency purpose vehicles based on wireless sensor network. *International Journal of Computer Applications* 40, 12 (2012), 36–39.

[12] KTV Grattan, LS Grattan, and BT Meggitt. 2000. *Optical fiber sensor technology: advanced applications-Bragg gratings and distributed sensors.* Vol. 5. Springer Science & Business Media.

[13] P Healey. 1986. Instrumentation principles for optical time domain reflectometry. *Journal of Physics E: Scientific Instruments* 19, 5 (1986), 334.

[14] Qiuyang Huang, Yongjian Yang, Yuanbo Xu, Funing Yang, Zhilu Yuan, and Yongxiong Sun. 2021. Citywide road-network traffic monitoring using large-scale mobile signaling data. *Neurocomputing* 444 (2021), 136–146. https://doi.org/10.1016/j.neucom.2020.07.150

[15] Ana Mandić Ivanković, Dominik Skokandić, Aleš Žnidarič, and Maja Kreslin. 2019. Bridge performance indicators based on traffic load monitoring. *Structure and Infrastructure Engineering* 15, 7 (2019), 899–911. https://doi.org/10.1080/15732479.2017.1415941

[16] Neeraj Kumar Jain, RK Saini, and Preeti Mittal. 2019. A review on traffic monitoring system techniques. *Soft Computing: Theories and Applications* (2019), 569–577.

[17] Andreas Janecek, Danilo Valerio, Karin Anna Hummel, Fabio Ricciato, and Helmut Hlavacs. 2015. The Cellular Network as a Sensor: From Mobile Phone Data to Real-Time Road Traffic Monitoring. *IEEE Transactions on Intelligent Transportation Systems* 16, 5 (2015), 2551–2572. https://doi.org/10.1109/TITS.2015.2413215

[18] Shin-Ting Jeng and Lianyu Chu. 2014. A high-definition traffic performance monitoring system with the Inductive Loop Detector signature technology. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC).* 1820–1825. https://doi.org/10.1109/ITSC.2014.6957957

[19] Philippe Jousset, Thomas Reinsch, Trond Ryberg, Hanna Blanck, Andy Clarke, Rufat Aghayev, Gylfi P Hersir, Jan Henninges, Michael Weber, and Charlotte M Krawczyk. 2018. Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features. *Nature communications* 9, 1 (2018), 1–11.

[20] Karel J Keesman and Karel J Keesman. 2011. *System identification: an introduction.* Vol. 2. Springer.

[21] Amine Kherraki and Rajae El Ouazzani. 2022. Deep convolutional neural networks architecture for an efficient emergency vehicle classification in real-time traffic monitoring. *IAES International Journal of Artificial Intelligence* 11, 1 (2022), 110.

[22] A Lellouch, S Yuan, Z Spica, B Biondi, and WL Ellsworth. 2019. Seismic velocity estimation using passive downhole distributed acoustic sensing records: Examples from the San Andreas fault observatory at depth. *Journal of Geophysical Research: Solid Earth* 124, 7 (2019), 6931–6948.

[23] Zhong-Xian Li, Xiao-Ming Yang, and Zongjin Li. 2006. Application of cement-based piezoelectric sensors for monitoring traffic flows. *Journal of transportation engineering* 132, 7 (2006), 565–573.

[24] Nathaniel J. Lindsey, Siyuan Yuan, Ariel Lellouch, Lucia Gualtieri, Thomas Lecocq, and Biondo Biondi. 2020. City-Scale Dark Fiber DAS Measurements of Infrastructure Use During the COVID-19 Pandemic. *Geophysical Research Letters* 47, 16 (2020), 1–8. https://doi.org/10.1029/2020GL089931

[25] Jingxiao Liu, Bingqing Chen, Siheng Chen, Mario Bergés, Jacobo Bielak, and HaeYoung Noh. 2020. Damage-sensitive and domain-invariant feature extraction for vehicle-vibration-based bridge health monitoring. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 3007–3011.

[26] Jingxiao Liu, Siheng Chen, Mario Bergés, Jacobo Bielak, James H Garrett, Jelena Kovačević, and Hae Young Noh. 2020. Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction. *Mechanical Systems and Signal Processing* 136 (2020), 106454.

[27] Jingxiao Liu, Susu Xu, Mario Bergés, and Hae Young Noh. 2023. HierMUD: Hierarchical multi-task unsupervised domain adaptation between bridges for drive-by damage diagnosis. *Structural Health Monitoring* 22, 3 (2023), 1941–1968. https://doi.org/10.1177/14759217221081159

[28] Jingxiao Liu, Siyuan Yuan, Bin Luo, Biondo Biondi, Hae Young Noh, et al. 2023. Turning Telecommunication Fiber-Optic Cables into Distributed Acoustic Sensors for Vibration-Based Bridge Health Monitoring. *Structural Control and Health Monitoring* 2023 (2023).

[29] Myra Lydon, Su E Taylor, Desmond Robinson, A Mufti, and EJO Brien. 2016. Recent developments in bridge weigh in motion (B-WIM). *Journal of Civil Structural Health Monitoring* 6, 1 (2016), 69–81.

[30] Pallavi A Mandhare, Vilas Kharat, and CY Patil. 2018. Intelligent road traffic control system for traffic congestion: a perspective. *International Journal of Computer Sciences and Engineering* 6, 07 (2018), 2018.

[31] Yonas Muanenda. 2018. Recent advances in distributed acoustic sensing based on phase-sensitive optical time domain reflectometry. *Journal of Sensors* 2018 (2018).

[32] T. Nadeem, S. Dashtinezhad, Chunyuan Liao, and L. Iftode. 2004. TrafficView: a scalable traffic monitoring system. In *IEEE International Conference on Mobile Data Management, 2004. Proceedings. 2004.* 13–26. https://doi.org/10.1109/MDM.2004.1263039

[33] Chaitanya Narisetty, Tomoyuki Hino, Ming-Fang Huang, Ryusuke Ueda, Hitoshi Sakurai, Akihiro Tanaka, Takashi Otani, and Toru Ando. 2021. Overcoming challenges of distributed fiber-optic sensing for highway traffic monitoring. *Transportation Research Record* 2675, 2 (2021), 233–242.

[34] Kapileswar Nellore and Gerhard P. Hancke. 2016. A Survey on Urban Traffic Management System Using Wireless Sensor Networks. *Sensors* 16, 2 (2016). https://doi.org/10.3390/s16020157

[35] Eugene J. OBrien and Bernard Enright. 2013. Using Weigh-in-Motion Data to Determine Aggressiveness of Traffic for Bridge Loading. *Journal of Bridge Engineering* 18, 3 (2013), 232–239. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000368

[36] Optasense, a Luna company. 2022. Quantx das Interrogator: Optasense. https://www.optasense.com/technology/quantx/

[37] Optasense, a Luna company. 2022. Traffic Monitoring | Road & Highway Incident Detection. https://www.optasense.com/transportation/traffic-monitoring/

[38] Yunjiang Rao, Zinan Wang, Huijuan Wu, Zengling Ran, and Bing Han. 2021. Recent advances in phase-sensitive optical time domain reflectometry (Φ-OTDR). *photonic sensors* 11, 1 (2021), 1–30.

[39] Herbert E Rauch, F Tung, and Charlotte T Striebel. 1965. Maximum likelihood estimates of linear dynamic systems. *AIAA journal* 3, 8 (1965), 1445–1450.

[40] Peter Reinartz, Marie Lachaise, Elisabeth Schmeer, Thomas Krauss, and Hartmut Runge. 2006. Traffic monitoring with serial images from airborne cameras. *ISPRS Journal of Photogrammetry and Remote Sensing* 61, 3 (2006), 149–158. https://doi.org/10.1016/j.isprsjprs.2006.09.009 Theme Issue: Airborne and Spaceborne Traffic Monitoring.

[41] Kapseong Ro, Jun-Seok Oh, and Liang Dong. 2007. Lessons learned: Application of small uav for urban highway traffic monitoring. In *45th AIAA aerospace sciences meeting and exhibit*. 596.

[42] Kostia Robert. 2009. Video-based traffic monitoring at day and night vehicle features detection tracking. In *2009 12th International IEEE Conference on Intelligent Transportation Systems*. 1–6. https://doi.org/10.1109/ITSC.2009.5309837

[43] Geoff Rose. 2006. Mobile Phones as Traffic Probes: Practices, Prospects and Issues. *Transport Reviews* 26, 3 (2006), 275–291. https://doi.org/10.1080/01441640500361108

[44] Simo Särkkä. 2013. *Bayesian filtering and smoothing*. Number 3. Cambridge university press.

[45] Christopher John Setchell. 1998. *Applications of computer vision to road-traffic monitoring*. Ph.D. Dissertation. Citeseer.

[46] Sergey V Shatalin, Vladimir N Treschikov, and Alan J Rogers. 1998. Interferometric optical time-domain reflectometry for distributed optical-fiber sensing. *Applied optics* 37, 24 (1998), 5600–5604.

[47] Junzhu Shen and Tieyuan Zhu. 2021. Correlation between seismic noise variation and COVID-19 pandemic measures using recordings from Penn State FORESEE array. *SEG Technical Program Expanded Abstracts* 2021-September (2021), 3316–3320. https://doi.org/10.1190/segam2021-3584263.1

[48] Statista. 2022. Length of optical fiber cable network in China from 2012 to 2017. https://www.statista.com/statistics/898868/china-optical-fiber-cable-network-length/.

[49] Mustafa Tekinay, Tim Sylvester, Matthew Brunton, and Thiagarajan Ganesh. 2022. Applications of Fiber Optic Sensors in Traffic Monitoring: A Review. (2022).

[50] Martijn van den Ende, Andre Ferrari, Anthony Sladen, and Cedric Richard. 2022. Next-Generation Traffic Monitoring with Distributed Acoustic Sensing Arrays and Optimum Array Processing. (2022), 1104–1108. https://doi.org/10.1109/ieeeconf53345.2021.9723373

[51] Jinfang Wang and Mingguang Wu. 2004. An overview of research on weigh-in-motion system. In *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No. 04EX788)*, Vol. 6. IEEE, 5241–5244.

[52] Xin Wang, Zhongwen Zhan, Ethan F Williams, Miguel González Herráez, Hugo Fidalgo Martins, and Martin Karrenbach. 2021. Ground vibrations recorded by fiber-optic cables reveal traffic response to COVID-19 lockdown measures in Pasadena, California. *Communications Earth & Environment* 2, 1 (2021), 1–9.

[53] Myounggyu Won. 2020. Intelligent traffic monitoring systems for vehicle classification: A survey. *IEEE Access* 8 (2020), 73340–73358.

[54] Yeong-Bin Yang, JD Yau, Zhongda Yao, and YS Wu. 2004. *Vehicle-bridge interaction dynamics: with applications to high-speed railways*. World Scientific.

[55] Siyuan Yuan, Martijn van den Ende, Jingxiao Liu, Hae Young Noh, Robert Clapp, Cédric Richard, and Biondo Biondi. 2022. Spatial Deep Deconvolution U-Net for Traffic Analyses with DAS. *arXiv preprint arXiv:2212.03936* (2022).

[56] Siyuan Yuan, Ariel Lellouch, Robert G Clapp, and Biondo Biondi. 2020. Near-surface characterization using a roadside distributed acoustic sensing array. *The Leading Edge* 39, 9 (2020), 646–653.

[57] Siyuan Yuan, Jingxiao Liu, Hae Young Noh, and Biondo Biondi. 2021. Urban system monitoring using combined Vehicle Onboard Sensing and roadside Distributed Acoustic Sensing. *SEG Technical Program Expanded Abstracts* 2021-September, 2020 (2021), 3235–3239. https://doi.org/10.1190/segam2021-3584136.1

[58] Kivilcim Yuksel, Damien Kinet, Karima Chah, and Christophe Caucheteur. 2020. Implementation of a mobile platform based on fiber bragg grating sensors for automotive traffic monitoring. *Sensors* 20, 6 (2020), 1567.

[59] Paul Zarchan and H Musoff. 2000. Fundamentals of Kalman filtering: a practical approach. *Progress in Astronautics and Aeronautics, American Institute of Aeronautics and Astronautics (AIAA)* 190 (2000).

[60] Jinrui Zhang, Youyuan Lu, Zeyu Lu, Chao Liu, Guoxing Sun, and Zongjin Li. 2015. A new smart traffic monitoring method using embedded cement-based piezoelectric sensors. *Smart Materials and Structures* 24, 2 (2015), 025023.

[61] Weida Zhong, Qiuling Suo, Abhishek Gupta, Xiaowei Jia, Chunming Qiao, and Lu Su. 2021. MetaTP: Traffic prediction with unevenly-distributed road sensing data via fast adaptation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–28.

HierMUD: Hierarchical multi-task unsupervised domain adaptation between bridges for drive-by damage diagnosis

Jingxiao Liu[1] , Susu Xu[2], Mario Bergés[3] and Hae Young Noh[1]

Abstract
Monitoring bridges through vibration responses of drive-by vehicles enables efficient and low-cost bridge maintenance by allowing each vehicle to inspect multiple bridges and eliminating the needs for installing and maintaining sensors on every bridge. However, many existing drive-by monitoring approaches are based on supervised learning models that require massive labeled data from every bridge. It is expensive and time-consuming, if not impossible, to obtain these labeled data. Furthermore, directly applying a supervised learning model trained on one bridge to new bridges would result in low accuracy due to the shift between different bridges' data distributions. Moreover, when we have multiple tasks (e.g., damage detection, localization, and quantification), the distribution shifts become more challenging than having only one task because different tasks have distinct distribution shifts and varying task difficulties. To this end, we introduce **HierMUD**, the first **Hier**archical **M**ulti-task **U**nsupervised **D**omain adaptation framework that transfers the damage diagnosis model learned from one bridge to a new bridge without requiring any labels from the new bridge. Specifically, our framework learns a hierarchical neural network model in an adversarial way to extract features that are informative to multiple tasks and invariant across multiple bridges. To match distributions over multiple tasks, we design a new loss function based on a newly derived generalization risk bound to adaptively assign higher weights to tasks with more shifted distributions. To learn multiple tasks with varying task difficulties, we split them into easy-to-learn and hard-to-learn tasks based on their distributions. Then, we formulate a feature hierarchy to utilize more learning resources to improve the hard-to-learn tasks' performance. We evaluate our framework with experimental data from 2 bridges and 3 vehicles. We achieve up to 2X better performance than baseline methods, including average accuracy of 95% for damage detection, 93% for localization, and 0.38 lbs mean absolute error for quantification.

Keywords
Bridge health monitoring, indirect structural health monitoring, vehicle scanning method, vehicle-bridge interaction, transfer Learning, unsupervised domain adaptation, domain adversarial learning, multi-task learning

Introduction

Bridges are key components of transportation infrastructure. Aging bridges all over the world pose challenges to the economy and public safety. According to the 2016 National Bridge Inventory of the Federal Highway Administration, 139,466 of 614,387 bridges in the U.S. are structurally deficient or functionally obsolete.[1] The state of aging bridges demands researchers to develop efficient and scalable approaches for monitoring a large stock of bridges.

Currently, bridge maintenance is based on manual inspection,[2] which is inefficient, incurs high labor costs, and fails to detect damages in a timely manner. To address these challenges, structural health monitoring techniques,[3] where structures are instrumented using sensors (e.g., strain gauge,

accelerometers, cameras, etc.) to collect structural performance data, have been developed to achieve continuous and autonomous bridge health monitoring (BHM). Yet, such sensing methods are hard to scale up as they require on-site installation and maintenance of sensors on every bridge and

[1]Stanford University, CA, USA
[2]Stony Brook University, NY, USA
[3]Carnegie Mellon University, PA, USA

Corresponding author:
Jingxiao Liu, Department of Civil & Environmental Engineering, Stanford University, 473 Via Ortega Room 311, Stanford, CA 94305, USA.
Email: liujx@stanford.edu

cause interruptions to regular traffic for running tests and maintaining instruments.[4]

To address the drawbacks of current BHM, drive-by BHM approaches were proposed to use vibration data of a vehicle passing over the bridge for diagnosing bridge damage. Drive-by BHM is also referred to as the vehicle scanning method[5] or indirect structural health monitoring.[6–10] Vehicle vibrations contain information about the vehicle-bridge interaction (VBI) and thus can indirectly inform us of the dynamic characteristics of the bridge for damage diagnosis.[10,11] This is a scalable sensing approach with low-cost and low-maintenance requirements because each instrumented vehicle can efficiently monitor multiple bridges. Also, there is no need for direct installations and on-site maintenance of sensors on every bridge.

Previous drive-by BHM focuses on estimating bridge modal parameters (e.g., fundamental frequencies,[11–13] mode shapes,[14,15] and damping coefficients[13,16]) that can be used for detecting and localizing bridge damage. However, since the VBI system is a complex physical system involving many types of noise and uncertainties (e.g., environmental noise, vehicle operational uncertainties, etc.), such modal analysis methods for drive-by BHM are susceptible to them.[10] This makes the modal parameters' estimation inaccurate and limits the ability of drive-by BHM to diagnose bridge damage (e.g., localize and quantify damage severity).

More recently, data-driven approaches use signal processing and machine learning techniques to extract informative features from the vehicle acceleration signals.[6,7,10,13,17–23] The extracted features are more robust to noise, enabling more sophisticated diagnoses such as damage localization and quantification. However, such data-driven approaches generally use supervised learning models developed on available labeled data (i.e., a set of bridges with known damage labels) that is expensive, if not impossible, to obtain. This labeled data requirement is further exacerbated by having multiple diagnostic tasks (e.g., damage detection, localization, and quantification). Furthermore, the standard supervised learning-based approaches learned using vehicle vibration data collected from one bridge are inaccurate for monitoring other bridges because data distributions of the vehicle passing over different bridges are shifted. Having to re-train for each new bridge in multiple bridge monitoring is time-consuming and costly. Therefore, for multiple bridge monitoring with multiple diagnostic tasks, one needs to transfer or generalize the multi-task damage diagnostic model learned from one bridge to other bridges, in order to eliminate the need for requiring training labeled data from every bridge in multiple tasks.

To this end, we introduce **HierMUD**, a new **Hier**archical **M**ulti-task **U**nsupervised **D**omain adaptation framework that transfers a model learned from one bridge data to predict multiple diagnostic tasks (including damage detection, localization, and quantification) in another bridge in an unsupervised way. Specifically, HierMUD makes a prediction for the target bridge, without any labels from the target bridge in any of the tasks. We achieve this goal by extracting features that are 1) informative to multiple tasks (i.e., task-informative) and 2) invariant across the source and target domains (i.e., domain-invariant).

Our framework is inspired by an unsupervised domain adaptation (UDA) approach that has been developed in the machine learning community to address the data distribution shift between two different domains, namely source and target domains.[24–29] In our work, we denote the bridge with labeled data as "Source Domain," and the new bridge of interest without any labels as "Target Domain." UDA focuses on the unsupervised learning tasks in the target domain, which transfers the model learned using source domain data and labels (e.g., vehicle vibration data with the corresponding bridge damage labels) to predict tasks in the target domain without labels (e.g., vehicle vibration data from other bridges without knowing damage labels). In particular, domain adversarial learning algorithm extracts a feature that simultaneously maximizes the damage diagnosis performance for the source domain based on source domain labeled data while minimizing the performance of classifying which domain the feature came from using both source and target domain data.[29–32] This algorithm can better match complex data distributions between the source and target domains by learning feature representations through neural networks, when compared to other conventional UDA approaches.

However, directly matching different domains' data distributions over each diagnostic task separately through domain adversarial learning limits the overall performance of drive-by BHM, which has multiple diagnostic tasks. This is because these tasks are coupled/related with each other and share damage-sensitive information. The prediction performance of each task depends on that of other coupled tasks.

Therefore, our algorithm integrates multi-task learning (MTL) with domain-adversarial learning to fuse information from different bridges and multiple diagnostic tasks for effectively improving diagnostic accuracy and scalability of drive-by BHM. MTL is helpful because it simultaneously solves multiple learning tasks to improve the generalization performance of all the tasks, when compared to training the models for each task independently or sequentially (i.e., independent or sequential task learning).[33–42]

Yet, when integrating MTL and domain-adversarial learning, two particular research challenges exist:

1) **Distinct distribution shift:** When we have multiple damage diagnostic tasks, the distribution shift problem becomes more challenging than having

only one task because different tasks have distinct shifted distributions between the source and target domains. Therefore, it is important to develop efficient optimization strategies to find an optimal trade-off for matching different domains' distributions over multiple tasks.

2) **Varying task difficulty:** The learning difficulties of different tasks vary with the complexity of mappings from the vehicle vibration data distribution to the damage label distributions. Some tasks (e.g., damage quantification task) would have highly non-linear mappings between data and label distributions and, therefore, be more difficult to learn than other tasks (i.e., hard-to-learn). Therefore, we need to distribute learning resources (e.g., representation capacities) according to the learning difficulty of tasks.

To address the distinct distribution shift challenge, we introduce a new loss function that prioritizes domain adaptations for tasks having more severe distribution shifts between different domains through a soft-max objective function. To formulate this new loss function, we first derive a new generalization risk bound for multi-task UDA, by optimizing which the new loss function is designed. Specifically, we minimize our loss function to jointly optimize three components: 1) feature extractors, 2) task predictors, and 3) domain classifiers. The parameters of task predictors are optimized to predict task labels in the source domain training set, which ensures that the extracted features are task-informative. The parameters of feature extractors are optimized with the domain classifiers in an adversarial way, such that the best trained domain classifier cannot distinguish which domains the extracted features come from. During the optimization, our new loss function adaptively weighs more on minimizing the distribution divergence of tasks having more shifted distributions between different domains. In this way, the model is optimized to automatically find a trade-off for matching different domains' distributions over multiple tasks.

Further, to address the varying task difficulty challenge, we develop hierarchical feature extractors to allocate more learning resources to hard-to-learn tasks. We first split the multiple tasks into easy-to-learn tasks (e.g., damage detection and localization) and hard-to-learn tasks (e.g., damage quantification) based on their degrees of learning difficulty. Specifically, we model the learning difficulty of each task to be inversely proportional to performance in the source domain (e.g., damage localization and quantification accuracy in supervised learning settings). Then, the hierarchical feature extractors learn two-level features: task-shared and task-specific features. To achieve high prediction accuracy for multiple tasks without creating a complex model, we extract task-shared features from input data for easy-to-learn tasks and then extract task-specific features

from the task-shared features for only the hard-to-learn tasks. In this way, we allocate more learning resources (learning deeper feature representations) to learn hard-to-learn tasks, which improves the overall performance for all the tasks.

We evaluate our framework on the drive-by BHM application using lab-scale experiments with two structurally different bridges and three vehicles of different weights. We introduce damage by attaching a mass to the bridge as a damage proxy. By modifying the location and size of the attached mass, we can non-destructively assess the sensitivity of a diagnostic algorithm to changes in the bridge. This damage proxy has been used by researchers for validating damage diagnostic algorithms prior to their use in actual systems.[43–45] In the evaluation, we train our framework using labeled data collected from a vehicle passing over one bridge to diagnose damage in another bridge with unlabeled vehicle vibration data. Our framework outperforms five baselines without UDA, MTL, the new loss function, or hierarchical structure.

In summary, this paper has three main contributions:

1) We introduce HierMUD, a new multi-task UDA framework that transfers the model learned from one bridge to achieve multiple damage diagnostic tasks in another bridge without any labels from the target bridge in any of the tasks. To the best of our knowledge, this new framework is the first domain adaptation framework for multi-task bridge monitoring. We have since released a PyTorch[46] implementation of HierMUD at https://github.com/jingxiaoliu/HierMUD.

2) We derive a generalization risk bound that provides a theoretical guarantee to achieve domain adaptation on multiple learning tasks. This work bridges the gaps between the theories and algorithms for multi-task UDA. Based on this bound, we design a new loss function to find a trade-off for matching different domains' distributions over multiple tasks, which addresses the distinct distribution shift challenge.

3) We develop a hierarchical architecture for our multi-task and domain-adversarial learning algorithm. This hierarchical architecture ensures that the framework accurately and efficiently transfers the model for predicting multiple tasks in the target domain, which addresses the varying task difficulty challenge.

The remainder of this paper is divided into six sections. In section 2, we study the MTL and data distribution shift challenges in the drive-by BHM application. Section 3 derives the generalization risk bound for multi-task UDA. Section 4 presents our HierMUD framework, which

includes the description of our framework, loss function, and algorithm design. Section 5 describes the evaluation of our framework on the drive-by BHM application, following by Section 6 that presents the evaluation results. Section 7 concludes our work and provides discussions about future work.

Data distribution shift and multi-task learning challenges for the drive-by bridge health monitoring

In this section, we describe the physical insights that enable our drive-by BHM and explain the associated challenges in achieving this scalable BHM approach. We first characterize the structural dynamics of the VBI system. Next, we study the multi-task learning and data distribution shift challenges, respectively, by proving the error propagation between multiple damage diagnostic tasks and characterizing the shifting of the joint distribution of vehicle vibration and damage labels.

Characterizing the structural dynamics of the VBI system

To provide physical insights of drive-by BHM, we model the VBI system, as shown in Figure 1, as a sprung mass (representing the vehicle) traveling with a constant speed on a simply supported beam (representing the bridge). We assume the beam is of the Euler-Bernoulli type with a constant cross section. We also assume that there is no friction force between the 'wheel' and the beam. The damage is simulated by attaching a mass (magnitude/severity level: q) at location (l) on the beam. The added mass changes the mass of the bridge and its dynamic characteristics.[47] Modifying the weight of the attached mass is a non-destructive way of creating physical changes to the VBI system to mimic structural damage.[43–45]

In our prior works,[10,42] we have derived the theoretical formulation of the VBI system in the frequency domain, which is summarized in the following paragraphs.

The equations of motion for the vehicle and bridge in the time domain are first derived as

$$m_v \ddot{u}_{dy}(t) + k_v \left[u_{dy}(t) - y_{dy}(x = vt, t) - y_{st}(x = vt) \right]$$
$$+ c_v \left[\dot{u}_{dy}(t) - \dot{y}_{dy}(x = vt, t) - \dot{y}_{st}(x = vt) \right] = 0 \quad (1)$$

$$\rho A \ddot{y}_{dy}(x, t) + q \ddot{y}_{dy}(l, t) \delta(x - l) + EI y_{dy}''''(x, t)$$
$$= \left\{ -m_v g + k_v \left[u_{dy}(t) - y_{dy}(x = vt, t) - y_{st}(x = vt) \right] \right.$$
$$\left. + c_v \left[\dot{u}_{dy}(t) - \dot{y}_{dy}(x = vt, t) - \dot{y}_{st}(x = vt) \right] \right\} \delta(x - vt)$$
$$(2)$$

where m_v, k_v, c_v, and u are the mass, stiffness, damping coefficient, and total displacement of the vehicle, respectively; ρ, A, E, I, and y_{st} are the density, sectional area, Young's modulus, moment of inertia, and the static displacement of the bridge, respectively; $\delta(x - vt)$ is the Dirac delta function; and $u_{dy}(t)$ and $y_{dy}(t)$ are the dynamic displacements of vehicle and bridge, respectively.

Then the n-th mode frequency response of the vehicle's acceleration is

$$\ddot{U}_{dy,n}(\omega) = \ddot{U}_{dy,n}\left(\omega - \frac{2n\pi v}{L} \right)$$

$$- \frac{i\pi^4 EI - i\left(\omega - \frac{n\pi v}{L} \right)^2 L^3 \left(\rho AL + 2q\sin^2\left(\frac{n\pi l}{L} \right) \right)}{2\pi m_v L^3 \left(\omega - \frac{n\pi v}{L} \right)^2 \sin\left(\frac{n\pi}{2} \right)}$$

$$\times \ddot{Y}_{dy,n}\left(\frac{L}{2}, \omega - \frac{n\pi v}{L} \right)$$

$$- \sqrt{2\pi} g \left[\Delta\left(\omega - \frac{2n\pi v}{L} \right) - \Delta(\omega) \right]$$
$$(3)$$

where i is the imaginary number; $Y_{dy,n}(x, \omega)$ is the n-th mode frequency response function (FRF) of the bridge element acceleration at the location x; and $U_{dy,n}(\omega)$ and $\Delta(\omega)$
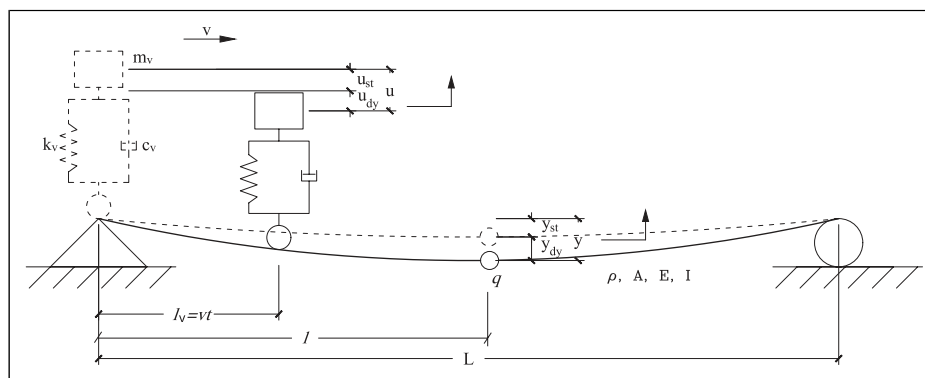

Figure 1. The vehicle-bridge interaction system with a surrogate damage simulated by attaching a mass having magnitude q at location l.

are the n-th mode FRF of the vehicle acceleration and the FRF of the Dirac delta function, respectively.

From Equation (3), we obtain the following important physical understandings of the drive-by vehicle vibration:

1) **Non-linear property**: Vehicle acceleration $(\ddot{U}_{dy,n}(\omega))$ is a high-dimensional signal that has a complex non-linear relationship with bridge properties (ρ, A, E, I) and damage parameters (q, l). Thus, it is difficult to infer damage states for different bridges by directly analyzing the raw vehicle signals. It is important to model features that can represent the non-linearity of the VBI system.

2) **Coupled diagnostic tasks**: Different damage locations (l) and severity levels (q) only vary the term $q \sin^2\left(\frac{n\pi l}{L}\right)$. Let's define the damage information as $d = q \sin^2\left(\frac{n\pi l}{L}\right)$, representing structural dynamic characteristic changes due to damage. The damage localization and quantification tasks are coupled with each other through the same damage information d, and thus the estimation of them depends on each other's estimation.

We incorporate these physical insights of the VBI system to develop our multi-task UDA framework. The non-linear property instructs us to use non-linear models or extract non-linear features from the vehicle vibrations for estimating bridge damage. In this work, we use a neural network-based model to non-linearly extract task-informative features from vehicle vibration data.

Moreover, the coupled tasks property of the VBI system informs us to use or extract the shared damage information (e.g., task-shared features) instead of independently learning multiple diagnostic tasks. This is further discussed in the following subsection.

Error propagation between multiple tasks for a VBI system

The shared information among multiple tasks can be learned simultaneously from multiple tasks or learned sequentially from one task to the next. In this section, we illustrate that simultaneous learning (i.e., MTL) is more accurate than sequential learning through a theoretical study of the VBI system. The study shows that the sequential learning method results in a significant error propagation from previous task to the next.

For instance, if we localize and quantify the bridge damage sequentially (localize the damage first and then quantify the severity of the damage at the obtained damage location), the estimation of damage severity q is

$d / \sin^2\left(\frac{n\pi\widehat{l}}{L}\right)$, where $\widehat{l}$ is the estimated damage location. Then, the propagation of error from the damage location estimation to the severity estimation is

$$\sigma_q = \pm \sqrt{\sigma_d^2 \frac{1}{\sin^4\left(n\pi\widehat{l}\Big/L\right)} + \sigma_l^2 \frac{4n^2\pi^2 d^2 \cos^2\left(n\pi\widehat{l}\Big/L\right)}{L^2 \sin^6\left(n\pi\widehat{l}\Big/L\right)}}$$

(4)

where $\sigma_q, \sigma_l, \sigma_d$ are errors of severity, location, and damage information estimations, respectively. $\sin^4(n\pi\widehat{l}/L)$ and $\sin^6(n\pi\widehat{l}/L)$ in the denominators are smaller than 1, which makes the estimation error of q very large as their values decrease. Especially, when the damage is close to the bridge supports (i.e., $|\widehat{l} - L/2| \to L/2$), it leads to $\sin(n\pi\widehat{l}/L) \to 0$ and $\sigma_q \to \infty$. Thus, damage location estimation error propagates, which results in a very inaccurate estimation of damage severity level.

To this end, we solve multiple tasks simultaneously, which can improve the overall accuracy by minimizing error propagation and learning the shared information (e.g., task-shared feature representations) from the coupled tasks.

Besides simultaneously learning multiple tasks, a scalable drive-by BHM approach needs to work for multiple domains (i.e., bridges). The following subsection discusses the data distribution shift challenge for drive-by monitoring of multiple bridges.

Data distribution shift for VBI systems

The joint distributions of vehicle vibrations and damage labels are shifted as the vehicle passes by different bridges. If we consider the process of the VBI system as a stochastic process, according to Equation (3), the joint distributions of the vehicle accelerations ($\ddot{U}_{dy,n}(\omega)$) and damage labels (q or l) changes non-linearly with bridge properties (e.g., ρ, A, E, I). An example of the shifting of these joint data distributions for different bridges is visualized in a low-dimensional space in Figure 2. It shows a two-dimensional t-Distributed Stochastic Neighbor Embedding (tSNE)[48] visualization of vehicle vibration data distributions for Bridges#1 and #2. Each vibration signal is collected from a vehicle passing over an 8-feet bridge model with a damage at the location (l) of 2 feet, 4 feet, or 6 feet. The data for the two structurally different bridges (Bridge#1 and #2) are represented by filled and unfilled markers, respectively. More details of this experiment and dataset are in Evaluation Section. We can observe from Figure 2 that directly applying the model learned from one bridge's dataset (e.g., Bridge#1) to
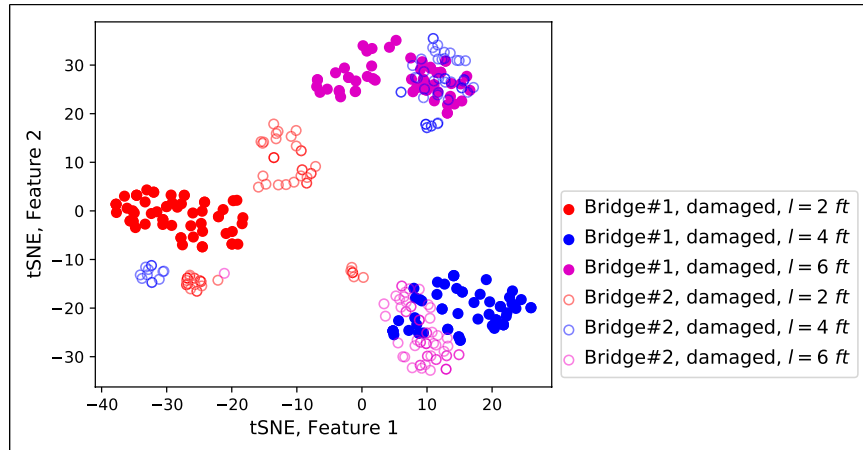
Figure 2. The 2D tSNE visualization of vibration data distributions of a vehicle passing over different bridges. The two coordinates are the first and the second low-dimensional embeddings of the vehicle accelerations calculated by tSNE. Different colors represent different damage locations on the bridge. Filled markers indicate bridge#1 data, and unfilled markers indicate bridge#2 data. This figure shows that directly utilizing the model trained on one bridge's (e.g., Bridge#1) dataset to predict damage locations of another bridge (e.g., Bridge#2) can result in very low accuracy due to data distribution shift.

localize damage on the other bridge (e.g., Bridge#2) results in a low prediction accuracy because of the joint distribution shift.

To address the distribution shift challenge and achieve a scalable drive-by BHM that is invariant across multiple bridges (i.e., it can predict damage without requiring training data from every bridge), we introduce a new multi-task UDA approach. In the next section, we first investigate the multi-task UDA problem theoretically through deriving a generalization risk bound.

A generalization risk bound for multi-task unsupervised domain adaptation

In this section, we derive the upper bound of the generalization risk for multi-task UDA problems to investigate the theoretical guarantee of its performance on target domain unseen data. The generalization risk (or error) of a model is the difference between the empirical loss on the training set and the expected loss on a test set, as defined in statistical learning theory.[49] In other words, it represents the ability of the trained model to generalize from the training dataset to a new unseen dataset. In our problem, the generalization risk is defined to represent how accurately a classifier trained using source domain labeled data and target domain unlabeled data predicts class labels in the target domain. Therefore, deriving the upper bound of the generalization risk provides insights on how to develop learning algorithms to efficiently optimize it.

We first derive a generalization risk bound for UDA and then integrate it with the risk bound for MTL. Next, we

characterize the newly derived generalization risk bound to provide insights to our multi-task UDA problem.

A generalization risk bound for unsupervised domain adaptation

We first derive a new generalization risk bound for UDA by representing the original data distribution in an intermediate feature space, which has been ignored by the existing risk bounds for UDA.[31,32,50] Having a feature space enables the modeling of task-shared feature representation when we have multiple tasks. This results in a tighter generalization risk bound for multi-task UDA than independently estimating each task's generalization risk bound.[51] Yet, this feature space requires us to estimate the discrepancy between the marginal feature distributions of the source and target domains for obtaining the generalization risk bound, which is introduced in this section.

We consider a classification task that labels input X as belonging to different classes Y. We also consider mappings

$$X \xrightarrow{w} Z \xrightarrow{h} Y,$$

where X, Z, and Y are random variables of input, feature representation, and class label, which are taken from the input, feature, and output space $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{Y}$, respectively. The function $w : \mathcal{X} \to \mathcal{Z}$ is a k-dimension feature transformation, and the function $h : \mathcal{Z} \to \mathcal{Y}$ is a hypothesis on the feature space (i.e., a labeling function). Then, we have a predictor $h \circ w$, that is, $(h \circ w)(x) = h(w(x))$, for every $x \in \mathcal{X}$.

Further, we define a domain as a distribution on the input space $\mathcal{X}$ and the output space $\mathcal{Y}$. UDA problems involve two domains, a source domain $\mathcal{D}_S$ and a target domain $\mathcal{D}_T$. We denote $\mathcal{D}_S^X$, $\mathcal{D}_T^X$, $\mathcal{D}_S^Y$, and $\mathcal{D}_T^Y$ as the marginal data (X) and label (Y) distributions in the source ($\mathcal{D}_S$) and target ($\mathcal{D}_T$) domains, respectively. Note that we also have feature representation distributions, $\mathcal{D}_S^Z$ and $\mathcal{D}_T^Z$, in the source and target domains. Mathematically, an unsupervised domain adaptation algorithm has independent and identically distributed (i.i.d.) labeled source samples $<X_S, Y_S>$ drawn from $\mathcal{D}_S$ and i.i.d. unlabeled target samples X_T drawn from $\mathcal{D}_T^X$, as shown below

$$<X_S, Y_S> = \{x_i, y_i\}_{i=1}^{n_S} \sim \mathcal{D}_S;$$

$$X_T = \{x_i\}_{i=1}^{n_T} \sim \mathcal{D}_T^X$$

where n_S and n_T are the number of samples in the source and target domains, respectively. The goal of UDA is to learn h and w with a low target domain risk under distribution $\mathcal{D}_T$, which is defined as: $\epsilon_T(h \circ w, f_T) = \Pr_{x \sim \mathcal{D}_T^X}(h \circ w(x) \neq f_T(x))$, where f_T is the ground truth labeling function and $y = f_T(x)$ for $(x, y) \sim \mathcal{D}_T$.

Since we do not have labeled data in the target domain, we cannot directly compute the target domain risk. Therefore, the upper bound of the target domain risk is estimated by the source domain risk and the discrepancy between the marginal data distributions of the source and target domains, $\mathcal{D}_S^X$ and $\mathcal{D}_T^X$. The discrepancy between $\mathcal{D}_S^X$ and $\mathcal{D}_T^X$ is quantified through the $\mathcal{H}\Delta\mathcal{H}$-divergence[52] that measures distribution divergence with finite samples of unlabeled data from $\mathcal{D}_S^X$ and $\mathcal{D}_T^X$. It is defined as

$$d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_S^X, \mathcal{D}_T^X) = 2\sup_{h, h' \in \mathcal{H}}\left|\Pr_{x \sim \mathcal{D}_S^X}\left[h(x) \neq h'(x)\right] - \Pr_{x \sim \mathcal{D}_T^X}\left[h(x) \neq h'(x)\right]\right|, \quad (5)$$

where $\mathcal{H}$ is a hypothesis space, and $\mathcal{H}\Delta\mathcal{H}$ is the symmetric difference hypothesis; h' and h are two samples of hypotheses randomly drawn from the hypothesis set $\mathcal{H}$.

Then, we can derive the generalization bound for UDA in Theorem 1.

Theorem 1. *Let $\mathcal{W}$ be a hypothesis space on $\mathcal{X}$ with Vapnik-Chervonenkis dimension (VC dimension) d_W and $\mathcal{H}$ be a hypothesis space on $\mathcal{Z}$ with VC dimension d_H, where VC dimension is a measure of the complexity of a set of functions that a classification algorithm can learn.[53] If X_S and X_T are samples of size N from $\mathcal{D}_S^X$ and $\mathcal{D}_T^X$, respectively, and Z_S and Z_T follow distributions $\mathcal{D}_S^Z$ and $\mathcal{D}_T^Z$, respectively, then for any $\delta \in (0, 1)$ with probability at least $1 - \delta$, for every $h \in \mathcal{H}$ and $w \in \mathcal{W}$*

$$\epsilon_T(h \circ w; f_T) \leq \epsilon_S(h \circ w; f_S) + 2\epsilon_S(h \circ w^*, f_S)$$
$$+ \frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}(Z_T, Z_S)$$
$$+ \mathcal{O}\left(\sqrt{\frac{2d_H\log(2N) + \log(2/\delta)}{N}}\right)$$
$$+ \frac{1}{2}\sup_{\widehat{h} \in \mathcal{H}}\left[\widehat{d_{h, \mathcal{W}\Delta\mathcal{W}}}(X_S, X_T)\right] \quad (6)$$
$$+ \mathcal{O}\left(\sqrt{\frac{2d_W\log(2N) + \log(2/\delta)}{N}}\right)$$
$$+ \epsilon_T(h^* \circ w^*; f_T) + \epsilon_S(h^* \circ w^*; f_S),$$

where

$$w^*, h^* = \operatorname*{argmin}_{w^* \in \mathcal{W}, h^* \in \mathcal{H}} \epsilon_S(h \circ w; f_S) + \epsilon_T(h \circ w; f_T),$$

$$d_{h, \mathcal{W}\Delta\mathcal{W}}(\mathcal{D}_S^X, \mathcal{D}_T^X)$$

$$= 2\sup_{w, w' \in \mathcal{W}}\left|\Pr_{x \sim \mathcal{D}_T^X}\left[h \circ w(x) \neq h \circ w'(x)\right]\right.$$

$$\left. - \Pr_{x \sim \mathcal{D}_S^X}\left[h \circ w(x) \neq h \circ w'(x)\right]\right| \quad (7)$$

$$\leq d_{h, \mathcal{W}\Delta\mathcal{W}}(X_S, X_T)$$

$$+ \mathcal{O}\left(\sqrt{\frac{2d_W\log(2N) + \log(2/\delta)}{N}}\right).$$

Proof. See Appendix A.

We prove in Theorem 1 that the upper bound of the target domain risk consists of five components:

1) The source domain risk

$$\epsilon_S(h \circ w; f_S) + 2\epsilon_S(h \circ w^*; f_S),$$

which quantifies the error for estimating class labels in the source domain.

2) The minimal risk

$$\epsilon_T(h^* \circ w^*; f_T) + \epsilon_S(h^* \circ w^*; f_S),$$

which quantifies the error for estimating class labels using the ideal joint hypothesis over the source and target domains. It is the smallest error we can achieve using the best predictor in the hypothesis set.

3) The empirical symmetric divergence between marginal feature distributions

$$\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}(Z_T,Z_S),$$

which quantifies the distribution difference between the source and target domain marginal feature distributions.

4) The supremum of empirical symmetric divergence set between marginal data distributions

$$\frac{1}{2}\sup_{\widehat{h}\in\mathcal{H}}\left[d_{\widehat{h},\mathcal{W}\Delta\mathcal{W}}(X_S,X_T)\right],$$

which quantifies the distribution difference between the source and target domain marginal data distributions.

5) The Big-O terms that measure the complexity of the estimation of divergence.

Next, in the following subsection, we derive a generalization risk bound for multi-task UDA problems by considering the feature space $\mathcal{Z}$ being the task-shared feature space for multiple tasks.

Integrating multi-task learning bound with the unsupervised domain adaptation bound

We first consider multiple classification tasks that label input X_m as belonging to different classes Y_m, for $m = 1, 2, \ldots, M$, where M is the total number of tasks. The mappings for this multi-task learning problem becomes

$$X_m \xrightarrow{w} Z_m \xrightarrow{h_m} Y_m, \text{ for } m = 1,2,\ldots,M,$$

where X_m, Z_m, and Y_m are the m-th task's random variables of input, feature representation, and class label, respectively, which are taken from the input, feature, and output space $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{Y}$, respectively. The function w is a task-shared k-dimensional feature transformation, and the function h_m is a task-specific hypothesis for the m-th task on the feature space. For each task, we have a predictor $h_m{\circ}w$. We define the task-averaged true risk under the joint distribution $\prod_{m=1}^{M}\mathcal{D}_m$ as

$$\varepsilon_{avg}(w,h_1,\cdots,h_M;f_1,\ldots,f_M)$$

$$= \frac{1}{M}\sum_{m=1}^{M}\Pr_{x \sim \mathcal{D}_m^X}(h_m \circ w(x) \neq f_m(x)), \quad (8)$$

where f_m is the ground truth labeling function for the m-th task and $y = f_m(x)$ for $(x,y) \sim \mathcal{D}_m$. We also define the task-averaged empirical risk as

$$\epsilon_{avg}\left(\widehat{w},\widehat{h}_1,\cdots,\widehat{h}_M;\overline{Y}\right)$$

$$= \frac{1}{NM}\sum_{m=1}^{M}\sum_{i=1}^{N}\mathbb{I}\left[\widehat{h}_m\left(\widehat{w}(x_{m,i})\right) \neq y_{m,i}\right], \quad (9)$$

where N is the total number of samples in each task, which is assumed to be the same for each task, and $\mathbb{I}(\cdot)$ is the indicator function. We define that $(\overline{X},\overline{Y}) = (X_1,\ldots,X_M,Y_1,\ldots,Y_M)$ are i.i.d. samples drawn from the joint distribution $\prod_{m=1}^{M}\mathcal{D}_m$. For the m-th task

$$(X_m,Y_m) = \left\{x_{m,i},y_{m,i}\right\}_{i=1}^{N} \sim \mathcal{D}_m.$$

Further, we consider a multi-task UDA problem, which has labeled samples $(\overline{X}_S,\overline{Y}_S) = (X_{S,1},\ldots,X_{S,M},Y_{S,1},\ldots,Y_{S,M})$ drawn from a joint source domain $\prod_{m=1}^{M}\mathcal{D}_{S,m}$ and unlabeled samples $\overline{X}_T = (X_{T,1},\ldots,X_{T,M})$ drawn from a joint target domain $\prod_{m=1}^{M}\mathcal{D}_{T,m}^{X}$. The goal of multi-tasks UDA is to learn $h_1, \ldots, h_M$ and w with a low target domain task-averaged risk under the joint distribution $\prod_{m=1}^{M}\mathcal{D}_{T,m}^{X}$, which is defined as: $\epsilon_{avg,T}(w, h_1, \ldots, h_M; f_{T,1}, \ldots, f_{T,M})$.

Our generalization risk bound for multi-task UDA is built on our Theorem 1 by combining it with the risk bound for multi-task learning. The multi-task learning risk bound was introduced in the work of Maurer et al.,[51] which showed that the upper bound of the task-averaged risk consists of the task-averaged empirical risk, the complexity measure relevant to the estimation of the representation, and the complexity measure of estimating task-specific predictors. Specifically, with probability at least $1 - \delta$, where $\delta \in (0, 1)$, in the draw of $(\overline{X},\overline{Y}) = (X_1,\ldots,X_M,Y_1,\ldots,Y_M) \sim \prod_{m=1}^{M}\mathcal{D}_m$, it holds for every $w \in \mathcal{W}$ and every $h_1,\ldots,h_M \in \mathcal{H}$ that

$$\epsilon_{avg}(w,h_1,\cdots,h_M;f_1,\ldots,f_M) \leq \epsilon_{avg}\left(\widehat{w},\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}\right)$$

$$+c_1\frac{L\widehat{G}\left(\mathcal{W}\left(\overline{X}\right)\right)}{NM} + c_2\frac{Q\sup_{w\in\mathcal{W}}\|w(\overline{X})\|}{N\sqrt{M}}$$

$$+\sqrt{\frac{9\log(2/\delta)}{2NM}}, \quad (10)$$

where L is the Lipschitz constant for $h \in \mathcal{H}$; c_1 and c_2 are universal constants; $\widehat{G}(\mathcal{W}(\overline{X}))$ is the Gaussian average that measures the empirical complexity relevant to the estimation of the feature representation; and $Q = \sup_{z\neq z'\in\mathcal{Z}}\frac{1}{\|z-z'\|}\mathbb{E}\sup_{h\in\mathcal{H}}\sum_{i=1}^{N}\gamma_i(h(z_i) - h(z_i'))$, γ_i's are independent standard Gaussian random variables.

Now, by integrating the multi-task learning risk bound in Equation (10) with the unsupervised domain adaptation risk bound in Theorem 1, we obtain the following theorem:

Theorem 2. *Let $\mathcal{D}_{S,1}, \cdots \mathcal{D}_{S,M}$ and $\mathcal{D}_{T,1}^X, \cdots \mathcal{D}_{T,M}^X$ be probability measure on $(\mathcal{X}, \mathcal{Y})$. Let $\mathcal{W}$ be a hypothesis space on $\mathcal{X}$ with VC dimension d_W and $\mathcal{H}$ be a hypothesis space on $\mathcal{Z}$ with VC dimension d_H. Let $\delta \in (0, 1)$. With probability at least $1 - \delta$ in the draw of $(\overline{X}_S, \overline{Y}_S) \sim \prod_{m=1}^M \mathcal{D}_{S,m}$, $\overline{X}_T \sim \prod_{m=1}^M \mathcal{D}_{T,m}^X$ (i.e., $(X_{S,m}, Y_{S,m}) \sim \mathcal{D}_{S,m}$ and $X_{T,m} \sim \mathcal{D}_{T,m}^X$ for $m = 1, \ldots, M$), and $\overline{Z}_S, \overline{Z}_T$ that follow distributions $\prod_{m=1}^M \mathcal{D}_{S,m}^Z$ and $\prod_{m=1}^M \mathcal{D}_{T,m}^Z$, it holds for every $w \in \mathcal{W}$ and every $h_1, \ldots, h_M \in \mathcal{H}$ that*

$$\epsilon_{avg,T}\left(w,h_1,\ldots,h_M;f_{T,1},\ldots,f_{T,M}\right) \leq \epsilon_{avg,S}\left(\widehat{w},\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right)$$

$$+c_1\frac{L\widehat{G}\left(\mathcal{W}\left(\overline{X}_S\right)\right)}{NM} + c_2\frac{Q\sup_{w\in\mathcal{W}}\|w\left(\overline{X}_S\right)\|}{N\sqrt{M}}$$

$$+2\epsilon_{avg,S}\left(w^*,\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right) + \frac{2\sqrt{2\pi}\widehat{G}\left(\mathcal{H}\left(\overline{Z}_S\right)\right)}{NM}$$

$$+\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}\left(\overline{Z}_S,\overline{Z}_T\right) + \mathcal{O}\left(\sqrt{\frac{2d_H\log(2N)+\log(2/\delta)}{N}}\right)$$

$$+ \sup_{\hat{h}_1\cdots\hat{h}_M\in\mathcal{H}}\left[\frac{1}{2}d_{\hat{h}_1,\ldots,\hat{h}_M,\mathcal{W}\Delta\mathcal{W}}\left(\overline{X}_S,\overline{X}_T\right)\right]$$

$$+\mathcal{O}\left(\sqrt{\frac{2d_W\log(2N)+\log(2/\delta)}{N}}\right)$$

$$+\epsilon_{avg,S}\left(w^*,h_1^*,\ldots,h_M^*;\overline{Y}_S\right)$$

$$+\epsilon_{avg,T}\left(w^*,h_1^*,\ldots,h_M^*;f_{T,1},\ldots,f_{T,M}\right),$$

$$(11)$$

where

$$w^*,h_1^*,\cdots,h_M^*$$

$$= \underset{w^*\in\mathcal{W},h_1^*,\ldots,h_M^*\in\mathcal{H}}{\operatorname{argmin}}\left[\epsilon_{avg,S}\left(w,h_1,\ldots,h_M;\overline{Y}_S\right)\right. \qquad (12)$$

$$\left.+\epsilon_{avg,T}\left(w,h_1,\ldots,h_M;f_{T,1},\ldots,f_{T,M}\right)\right].$$

We show in Theorem 2 that the upper bound of task-averaged target domain risk contains seven components:

1) The source domain empirical risks

$$\epsilon_{avg,S}\left(\widehat{w},\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right)$$

$$+2\epsilon_{avg,S}\left(w^*,\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right), \qquad (13)$$

2) the task-averaged minimal risks

$$\epsilon_{avg,S}\left(w^*,h_1^*,\ldots,h_M^*;\overline{Y}_S\right)$$

$$+\epsilon_{avg,T}\left(w^*,h_1^*,\ldots,h_M^*;f_{T,1},\ldots,f_{T,M}\right), \qquad (14)$$

3) the complexity measure relevant to the estimation of the representation

$$c_1\frac{L\widehat{G}\left(\mathcal{W}\left(\overline{X}_S\right)\right)}{NM},$$

4) the complexity measure of estimating task-specific predictors

$$c_2\frac{Q\sup_{w\in\mathcal{W}}\left\|w\left(\overline{X}_S\right)\right\|}{N\sqrt{M}} + \frac{2\sqrt{2\pi}\widehat{G}\left(\mathcal{H}\left(\overline{Z}_S\right)\right)}{NM},$$

5) the empirical symmetric divergence between marginal feature distributions

$$\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}\left(\overline{Z}_S,\overline{Z}_T\right),$$

6) the supremum of empirical symmetric divergence set (for multiple tasks) between marginal data distributions

$$\sup_{\hat{h}_1\cdots\hat{h}_M\in\mathcal{H}}\left[\frac{1}{2}d_{\hat{h}_1,\ldots,\hat{h}_M,\mathcal{W}\Delta\mathcal{W}}\left(\overline{X}_S,\overline{X}_T\right)\right],$$

7) the Big-O complexity measures of the estimation of divergence.

Once we determine the hypothesis sets $\mathcal{H}$ and $\mathcal{W}$, the task-averaged minimal risk and complexity terms are fixed.[50] Therefore, we can minimize the target domain risk bound (Equation 11) by minimizing the sum of the source domain empirical risks, the empirical divergence between marginal data distributions, and the empirical divergence between marginal feature distributions

$$\text{minimize}\left(\epsilon_{avg,S}\left(\widehat{w},\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right)\right.$$

$$+2\epsilon_{avg,S}\left(w^*,\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right)$$

$$+\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}\left(\overline{Z}_S,\overline{Z}_T\right)$$

$$\left.+\sup_{\widehat{h}_1\cdots\widehat{h}_M\in\mathcal{H}}\left[\frac{1}{2}d_{\widehat{h}_1,\ldots,\widehat{h}_M,\mathcal{W}\Delta\mathcal{W}}\left(\overline{X}_S,\overline{X}_T\right)\right]\right) \tag{15}$$

To efficiently solve Equation (15) that minimizes the generalization risk bound for multi-task UDA, we interpret and characterize the bound in the following subsection.

Characterizing the derived risk bound

We can learn two main insights of the new generalization risk bound for multi-task UDA from Equation (15):

1) **Feature divergence minimization:** Some UDA methods[30, 32] minimize the empirical symmetric divergences between marginal feature distributions (i.e., $\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}(\overline{Z}_S,\overline{Z}_T)$) to make the task-specific classifiers (i.e., $h_1, \ldots, h_M$) invariant across domains. However, these methods are not scalable as the number of tasks grows because they require every task-specific classifier to be domain-invariant. Therefore, our multi-task UDA approach avoids directly minimizing the feature divergence, which requires adapting classifiers of each task separately.

2) **Data divergence minimization:** Some UDA methods[31, 54] minimize the empirical symmetric divergence between marginal data distributions (i.e., $\sup_{\widehat{h}_1\cdots\widehat{h}_M\in\mathcal{H}^M}\left[\frac{1}{2}d_{\widehat{h}_1,\ldots,\widehat{h}_M,\mathcal{W}\Delta\mathcal{W}}(\overline{X}_S,\overline{X}_T)\right]$) to extract domain-invariant features. Such methods would be successful and scalable if the feature distributions $\mathcal{D}_S^Z$ and $\mathcal{D}_T^Z$ are matched because in this case the empirical symmetric divergence between marginal feature distributions (i.e., $\frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}(\overline{Z}_S,\overline{Z}_T)$) would be also very small or even be zero. However, directly minimizing this supremum divergence is difficult and data inefficient under the MTL setting because different tasks have distinct distribution shifts between the source and target domains. Therefore, we introduce a new efficient optimization strategy to find an optimal trade-off for minimizing the data divergence over multiple tasks.

In summary, to develop an algorithm that is scalable to the number of tasks, we need to minimize the empirical divergence between marginal data distributions to learn a feature mapping that matches feature distributions between different domains. Therefore, we can rewrite Equation (15) as

$$\text{minimize}\left(\epsilon_{avg,S}\left(\widehat{w},\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right)\right.$$

$$+2\epsilon_{avg,S}\left(w^*,\widehat{h}_1,\ldots,\widehat{h}_M;\overline{Y}_S\right) \tag{16}$$

$$\left.+\sup_{\widehat{h}_1\cdots\widehat{h}_M\in\mathcal{H}}\left[\frac{1}{2}d_{\widehat{h}_1,\ldots,\widehat{h}_M,\mathcal{W}\Delta\mathcal{W}}\left(\overline{X}_S,\overline{X}_T\right)\right]\right)$$

For a classification problem, minimizing the first two terms in Equation (16) can be achieved by minimizing the cross-entropy loss between predicted labels and ground truth labels in source domain: $\widehat{\epsilon}_{avg,S} =$

$$-\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}_{(x,y)\sim\left(X_{S,m},Y_{S,m}\right)}\sum_{c=1}^{C_m}\mathbb{I}(y=c)\log\widehat{h}_m\left(\widehat{w}(x)\right),$$

where C_m is the number of classes for the m-th tasks.

Further, if we consider $\widehat{h}_1,\ldots,\widehat{h}_M$ are hypotheses independently drawn from the hypothesis class $\mathcal{H}$, we can write the last term in Equation (16) as

$$\max_{m\in[M]}\sup_{\widehat{h}_m\in\mathcal{H}}\left[\frac{1}{2}d_{\widehat{h}_m,\mathcal{W}\Delta\mathcal{W}}(X_{S,m},X_{T,m})\right], \tag{17}$$

without loss of generality. This means that we can minimize the maximum divergence between marginal feature distributions over all the tasks to achieve the minimization of the divergence term in Equation (16). An approximation of the empirical symmetric divergence between distributions is computed by learning a domain discriminator ($h_m \circ w$) that distinguishes samples from different domains[31]

$$d_{h_m,\mathcal{W}\Delta\mathcal{W}}(X_{S,m},X_{T,m})$$

$$=2\left(1-\min_{w\in\mathcal{W}\Delta\mathcal{W}}\left(\frac{1}{N}\sum_{x\sim X_{S,m}}\mathbb{I}(h_m(w(x)=1)\right.\right.$$

$$\left.\left.+\frac{1}{N}\sum_{x\sim X_{T,m}}\mathbb{I}(h_m(w(x)=0)\right)\right). \tag{18}$$

To this end, there are three ways to minimize Equation (16):

1) *Hard-max objective*: directly minimizing the maximum divergence over all M tasks

$$\text{minimize}\left(\widehat{\epsilon}_{avg,S}\right.$$

$$\left.+\max_{m\in[M]}\sup_{\widehat{h}_m\in\mathcal{H}}\left[\frac{1}{2}d_{\widehat{h}_m,\mathcal{W}\Delta\mathcal{W}}(X_{S,m},X_{T,m})\right]\right); \tag{19}$$

2) *Average objective*: minimizing the average divergence:

$$
\text{minimize} \left(\widehat{\epsilon}_{avg,S} \right.
$$
$$
\left. + \frac{1}{2M} \sum_{m=1}^{M} \sup_{\widehat{h_m} \in \mathcal{H}} \left[d_{\widehat{h_m}, \mathcal{W} \Delta \mathcal{W}} (X_{S,m}, X_{T,m}) \right] \right);
$$

(20)

3) *Soft-max objective*: minimizing a soft maximum, that is, a LogSumExp,[55] of Equation (18)

$$
\text{minimize} \left(\widehat{\epsilon}_{avg,S} + \frac{1}{2} \log \sum_{m=1}^{M} \exp \right.
$$
$$
\left. \left(\sup_{\hat{h}_m \in \mathcal{H}} \left[d_{\hat{h}_m, \mathcal{W} \Delta \mathcal{W}} (X_{S,m}, X_{T,m}) \right] \right) \right).
$$

(21)

The hard-max objective is data inefficient because the gradient of the max function is only non-zero for h_m with the maximum divergence, and the algorithm only updates its parameters based on the gradient from one of the M tasks. The average objective updates algorithm parameters based on the average gradient from all M tasks. However, this objective considers each task as equally contributing to updating the algorithm parameters, which may not allocate enough computational and learning resources to optimize tasks with larger divergence. The soft-max objective combines the gradients from all the tasks and adaptively assigns the loss of task that have a larger divergence with a heavier weight.[50] In this way, the model automatically applies larger gradient magnitudes to tasks having more shifted distributions between different domains. As a result, we propose to use the soft-max objective (Equation 21) to optimize the multi-task UDA problem.

HierMUD: A hierarchical multi-task unsupervised domain adaptation framework

We now proceed to introduce our HierMUD framework that transfers the model learned from a source domain to predict multiple tasks on a target domain without any labels from the target domain in any of the tasks. In our drive-by BHM application, the two domains are vibration data and damage information collected from a vehicle passing over two structurally different bridges, and the multiple learning tasks are damage diagnostic tasks, such as damage detection, localization, and quantification. The overview flowchart of our framework is shown in Figure 3. The framework contains three modules: 1) a data pre-processing module, 2) a multi-task UDA module, and 3) a target domain prediction

module. In the following subsections, we present each module in detail.

Data pre-processing module

The data pre-processing module contains two steps: data augmentation and initial processing. In the first step, to avoid overfitting and data biases while providing sufficient information of each class, we expand the size of the dataset and introduce data variability by conducting data augmentation, which improves the robustness of the learned multi-task UDA model. Some widely utilized data augmentation procedures include adding white noise, randomly cropping matrix size, or randomly erasing samples from the original data for each batch of training data.[56] The procedure is selected based on the prediction performance on the validation set (e.g., for our drive-by BHM application, we add white noise to augment the vibration data).

In the second step, we create the input to our multi-task UDA module, including the source domain data with the corresponding labels and the target domain data. Feature transforms are applied to the raw input data to provide information in other feature space. For example, a Fast Fourier Transform can be used to convert the signal from its original domain (time or space) to the frequency domain. Short-Time Fourier Transform (STFT) or wavelet transform can be used to convert the time or space domain signal to the time–frequency domain. Specifically, in our drive-by BHM application, we conduct data augmentation by adding white noise to vehicle vibration signals. Then, we compute the STFT of each vertical acceleration record of the vehicle traveling over the bridge to preserve the time–frequency domain information.

Multi-task unsupervised domain adaptation module

In this module, we introduce our hierarchical multi-task and domain adversarial learning algorithm (as shown in Figure 4) that exploits the derived generalization risk bound for multi-task UDA based on the theoretical study in the previous section. This algorithm integrates domain adversarial learning and hierarchical multi-task learning to achieve an optimal trade-off between domain-invariance and task-informativeness.

Our algorithm consists of three components: hierarchical feature extractors (orange blocks), task predictors (blue blocks), and domain classifiers (red blocks). Domain adversarial learning utilizes the domain classifier to minimize the domain discrepancy through an adversarial objective (e.g., Equation 18) for training against the feature extractors, which encourages the extracted feature to be domain-invariant.[24] The architectures of each component are presented in the following paragraphs.
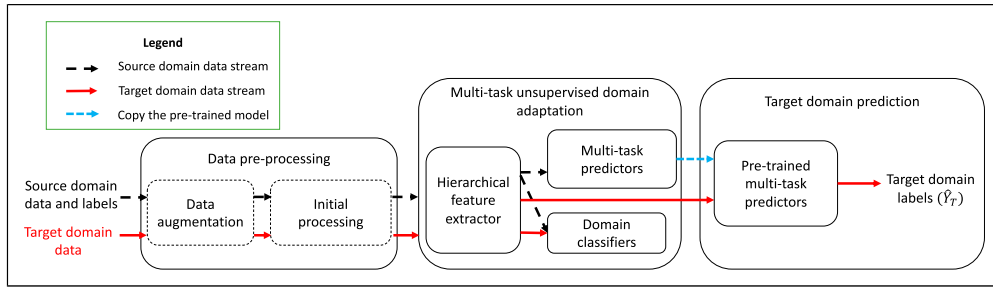
Figure 3. The flowchart for our HierMUD framework. The input to the framework is source domain data with the corresponding labels and the unlabeled target domain data, and the framework outputs the predicted target domain labels. Black dash lines indicate source domain data stream, solid red lines indicate target domain data stream, and the blue dash line indicates that we copy the multi-task predictors pre-trained using source domain data to predict target domain tasks.
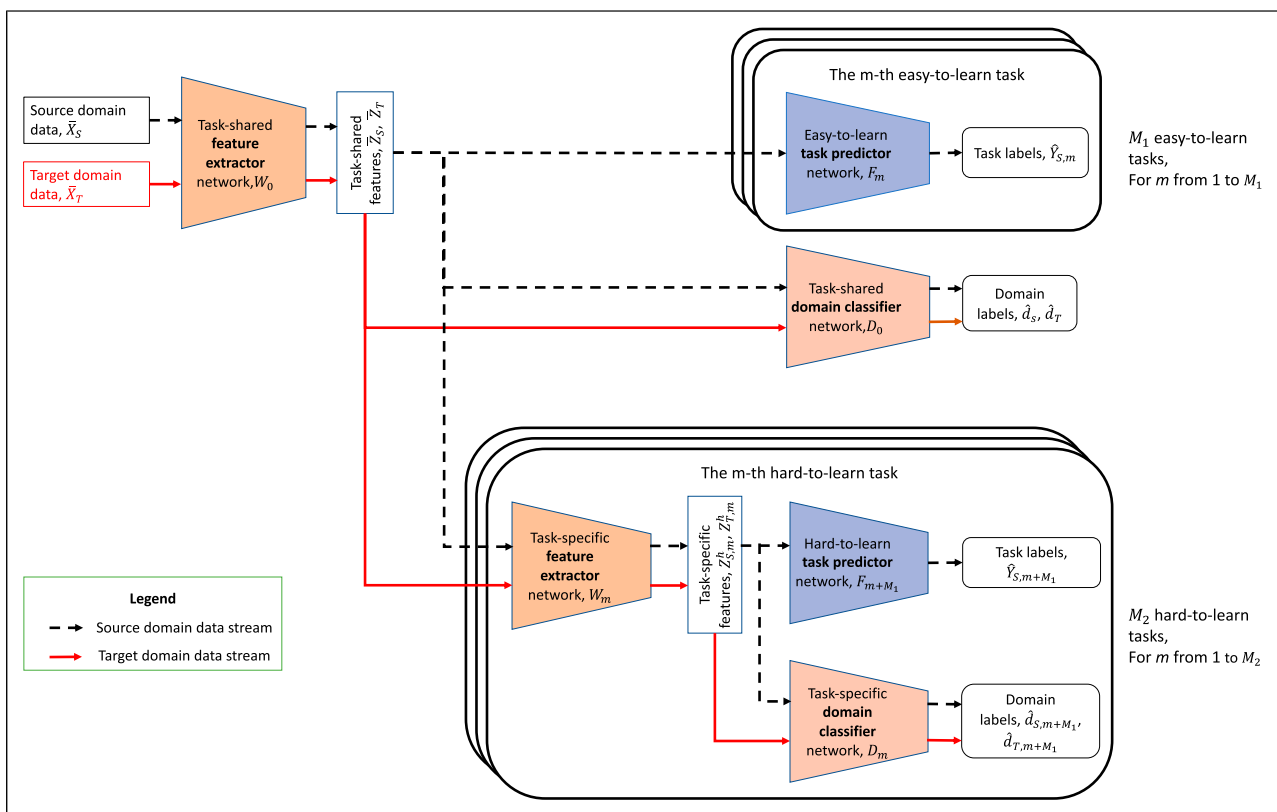
Figure 4. The architecture of our hierarchical multi-task and domain-adversarial learning algorithm. The red and black arrows between blocks represent source and target domain data stream, respectively. Orange blocks are feature extractors, blue blocks are task predictors, and red blocks are domain classifiers.

Architecture of hierarchical feature extractors, task predictors, and domain classifiers. Hierarchical feature extractors. The hierarchical feature extractors extract domain-invariant and task-informative features. To ensure domain-invariance, the parameters of the extractors are optimized with domain classifiers in an adversarial way to extract features that cannot be differentiated by the domain classifiers (i.e., domain-invariant) while the domain classifiers are

optimized to best distinguish which domain the extracted features come from.

To ensure task-informativeness, we implement hierarchical feature extractors that learn task-shared and task-specific feature representations for tasks with different learning difficulties. Inspired by human learning[57,58] and the work of Guo et al.,[59] we separate the total of M tasks into M_1 easy-to-learn and M_2 hard-to-learn tasks based on the

task difficulty, which is inversely proportional to the learning performance (e.g., prediction accuracy) in the source domain. In particular, we train, respectively, M classifiers with the same model complexity using source domain data of the M tasks, and obtain testing accuracy values, $\{p_1, p_2, \ldots, p_M\}$, for the M tasks. A threshold p_t is determined based on our domain knowledge and empirical observations of the problem. We then consider tasks with $p_m \geq p_t$ as easy-to-learn tasks, and tasks with $p_m \leq p_t$ as hard-to-learn tasks. The choice of p_t is application different and relies on domain knowledge. A smaller p_t would categorize more tasks as hard-to-learn ones and distribute more learning resources to these tasks, increasing the model complexity, and vice versa. For example, in our drive-by BHM application, the value of p_t is 0.9, damage detection and localization are considered as easy-to-learn tasks, and damage quantification is considered as a hard-to-learn task.

Furthermore, we implement two types of feature extractors: one task-shared feature extractor and M_2 task-specific feature extractors. We denote $W_0 (\cdot)$ as the task-shared feature extractor with parameters θ_{W_0}, and $W_m (\cdot)$ as the task-specific feature extractor for the m-th hard-to-learn task with parameters θ_{W_m}, where $m \in \{1, 2, \ldots, M_2\}$. The source and target domain data after being processed in the data pre-processing module, $(\overline{X}_S, \overline{X}_T)$, are input to the task-shared feature extractor to extract task-shared features, $\overline{Z}_S$ and $\overline{Z}_T$, for the source and target domain, respectively. For each hard-to-learn task, task-specific features, $Z_{S,m}^h$ and $Z_{T,m}^h$, are extracted from the task-shared features using the corresponding task-specific feature extractor.

The task-shared feature extractor is implemented as a deep convolutional neural network (CNN) that combines convolutional layers and pooling layers to extract feature representations. We utilize CNN to extract features because it has an excellent performance in understanding spatial hierarchies and structures of features in various resolutions.[55] Further, task-specific feature extractors are implemented as deep fully connected neural networks, consisting of multiple fully connected layers that map task-shared features to task-specific features.

Task predictors. Task predictors are trained to ensure that the extracted features from the hierarchical feature extractors are task-informative. They are implemented as deep fully connected neural networks that map features to task labels. There are M task predictors for all the M learning tasks. We denote $F_m (\cdot)$ as the task predictor for the m-th task with parameters θ_{F_m}, where $m \in \{1, 2, \ldots, M\}$. In the training phase, the input to the task predictor of each easy-to-learn task is the task-shared feature in the source domain. The input to the task predictor of each hard-to-learn task is the corresponding task-specific feature in the source domain. Each predictor outputs the predicted source domain labels, $\widehat{Y}_{S,m}$, in each task.

Domain classifiers. Domain classifiers are trained to distinguish which domain the extracted features are

from. We also have two types of domain classifiers: one task-shared domain classifier and M_2 task-specific domain classifiers. We denote $D_0 (\cdot)$ as the task-shared domain classifier with parameters θ_{D_0}, and $D_m (\cdot)$ as the task-specific domain classifier for the m-th hard-to-learn task with parameters θ_{D_m}, where $m \in \{1, 2, \ldots, M_2\}$. The task-shared domain classifier takes the task-shared features in all M tasks from the source domain or the target domain as input and predicts if the feature sample comes from the source domain or not (i.e., a binary classification). Each task-specific domain classifier takes the task-specific features for each task from the source domain or the target domain as input and also classifies the feature sample into two classes (as the source or the target domain).

We implement the domain adversarial learning by back-propagation, inspired by Ganin et al.,[31] using the gradient reversal layer (GRL). We use GRL because it can be easily incorporated into any existing neural network architecture that can handle high-dimensional signals and multiple learning tasks. In particular, each domain classifier is connected to the corresponding feature extractor via a GRL that multiplies the gradient by a negative constant during back-propagation updating. With GRL, feature extractors and domain classifiers are trained in an adversarial way, such that the extracted features are as indistinguishable as possible for even well-trained domain classifiers.

Domain classifiers are implemented as deep fully-connected neural networks that map feature to domain labels. One should note that the architecture of domain classifiers is simpler than that of task predictors to avoid overusing learning resources to train domain classifiers over task predictors, which could reduce task-informativeness of the extracted features.[31]

Loss function for hierarchical multi-task and domain-adversarial learning algorithm. In this subsection, we present the loss function for our hierarchical multi-task and domain-adversarial learning algorithm, which minimizes the objective function in Equation (21).

After considering the hierarchical structure, we can rewrite the objective function of the optimization in Equation (21) as

$$
\min_{\substack{\theta_{W_0}, \theta_{W_1} \cdots, \theta_{W_{M_2}}, \\ \theta_{F_1}, \ldots, \theta_{F_M},}} \left[\sum_{m=1}^{M_1} \lambda_m \mathcal{L}_{e,m}(\theta_{W_0}, \theta_{F_m}) \right.
$$

$$
+ \frac{1}{M_2} \sum_{m=1}^{M_2} \lambda_{m+M_1} \mathcal{L}_{h,m}\left(\theta_{W_0}, \theta_{W_m}, \theta_{F_{m+M_1}}\right)
$$

$$
+ \lambda_{D_0} \log \sum_{m=1}^{M} \exp\left(-\min_{\theta_{D_0}} \mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0})\right)
$$

$$
\left. - \min_{\theta_{D_1}, \ldots, \theta_{D_M}} \sum_{m=1}^{M_2} \lambda_{D_m} \mathcal{L}_{D_m}(\theta_{W_0}, \theta_{W_m}, \theta_{D_m}) \right], \tag{22}
$$

where $\mathcal{L}_{e,m}(\theta_{W_0}, \theta_{F_m}) = -\mathbb{E}_{(x,y) \sim (X_{S,m}, Y_{S,m})} \sum_{c=1}^{C_m} \mathbb{I}(y = c)$ $\log \mathbf{F}_m(\mathbf{W}_0(x))$ is the cross-entropy loss for the m-th easy-to-learn task; C_m is the number of classes for the m-th task; $\mathcal{L}_{h,m}(\theta_{W_0}, \theta_{W_m}, \theta_{F_{m+M_1}}) = -\mathbb{E}_{(x,y) \sim (X_{S,m}, Y_{S,m})}[\sum_{c=1}^{C_{m+M_1}} \mathbb{I}(y = c)$ $\log \mathbf{F}_{m+M_1}(\mathbf{W}_m(\mathbf{W}_0(x)))]$ is the cross-entropy loss for the m-th hard-to-learn tasks; $\mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}) = -\mathbb{E}_{x \sim X_{S,m}}[\log \mathbf{D}_0(\mathbf{W}_0(x))] - \mathbb{E}_{x \sim X_{T,m}}[\log(1 - \mathbf{D}_0(\mathbf{W}_0(x)))])$ is the task-shared domain classifier loss for the m-th tasks ($m \in 1, 2, \ldots, M$); $\mathcal{L}_{D_m}(\theta_{W_0}, \theta_{W_m}, \theta_{D_m}) = -\mathbb{E}_{x \sim X_{S,m}}[\log \mathbf{D}_m(\mathbf{W}_m(\mathbf{W}_0(x)))] - \mathbb{E}_{x \sim X_{T,m}}[\log(1 - \mathbf{D}_m(\mathbf{W}_m(\mathbf{W}_0(x))))]$ is the task-specific domain classifier loss; $\lambda_1, \ldots, \lambda_M$ are hyper-parameters to represent the trade-off between easy-to-learn tasks and hard-to-learn tasks weights; and $\lambda_{D_0}, \lambda_{D_1}, \ldots, \lambda_{D_{M_2}}$ are hyper-parameters to represent the trade-off between domain-invariance and task-informativeness of features.

The minimax optimization problem in Equation (22) is solved by finding the saddle point

$$\widehat{\theta}_{W_0}, \widehat{\theta}_{W_1}, \ldots, \widehat{\theta}_{W_{M_2}}, \widehat{\theta}_{F_1}, \ldots, \widehat{\theta}_{F_M}, \widehat{\theta}_{D_0}, \widehat{\theta}_{D_1}, \ldots, \widehat{\theta}_{D_M},$$

such that

$$\left(\widehat{\theta}_{W_0}, \widehat{\theta}_{W_1}, \ldots, \widehat{\theta}_{W_{M_2}}, \widehat{\theta}_{F_1}, \ldots, \widehat{\theta}_{F_M} \right)$$
$$= \underset{\substack{\theta_{W_0}, \theta_{W_1} \cdots, \theta_{W_{M_2}}, \\ \theta_{F_1}, \ldots, \theta_{F_M}}}{\operatorname{argmin}} \left[\frac{1}{M_1} \sum_{m=1}^{M_1} \lambda_m \mathcal{L}_{e,m}(\theta_{W_0}, \theta_{F_m}) \right.$$
$$+ \lambda_h \frac{1}{M_2} \sum_{m=1}^{M_2} \lambda_{m+M_1} \mathcal{L}_{h,m} \left(\theta_{W_0}, \theta_{W_m}, \theta_{F_{m+M_1}} \right) \qquad (23)$$
$$+ \lambda_{D_0} \log \sum_{m=1}^{M} \exp \left(-\min_{\theta_{D_0}} \mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}) \right)$$
$$\left. - \min_{\theta_{D_1}, \ldots, \theta_{D_M}} \sum_{m=1}^{M_2} \lambda_{D_m} \mathcal{L}_{D_m}(\theta_{W_0}, \theta_{W_m}, \theta_{D_m}) \right],$$

$$\widehat{\theta}_{D_0} = \underset{\theta_{D_0}}{\operatorname{argmax}} \, \lambda_{D_0} \log \sum_{m=1}^{M} \exp(-\mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0})), \quad (24)$$

$$\widehat{\theta}_{D_m} = \underset{\theta_{D_m}}{\operatorname{argmin}} \, \lambda_{D_m} \mathcal{L}_{D_m} \left(\widehat{\theta}_{W_0}, \widehat{\theta}_{W_m}, \theta_{D_m} \right), \quad (25)$$

A saddle point defined by Equations (23)–(25) can be found as a stationary point of the following gradient updates: for updating hierarchical feature extractors,

$$\theta_{W_0} \leftarrow \theta_{W_0} - \mu \left(\frac{1}{M_1} \sum_{m=1}^{M_1} \lambda_m \frac{\partial \mathcal{L}_{e,m}}{\partial \theta_{W_0}} + \frac{1}{M_2} \sum_{m=1}^{M_2} \lambda_{m+M_1} \frac{\partial \mathcal{L}_{h,m}}{\partial \theta_{W_0}} \right.$$
$$\left. - \lambda_{D_0} \sum_{m=1}^{M} w_m \frac{\partial \mathcal{L}_{D_0,m}}{\partial \theta_{W_0}} - \sum_{m=1}^{M_2} \lambda_{D_m} \frac{\partial \mathcal{L}_{D_m}}{\partial \theta_{W_0}} \right),$$
$$(26)$$

$$\theta_{W_m} \leftarrow \theta_{W_m} - \mu \left(\frac{\lambda_{m+M_1}}{M_2} \frac{\partial \mathcal{L}_{h,m}}{\partial \theta_{W_m}} - \lambda_{D_m} \frac{\partial \mathcal{L}_{D_m}}{\partial \theta_{W_m}} \right), \quad (27)$$

for updating the m-th easy-to-learn task predictor, where $m \in \{1, \ldots, M_1\}$

$$\theta_{F_m} \leftarrow \theta_{F_m} - \mu \frac{\lambda_m}{M_1} \frac{\partial \mathcal{L}_{e,m}}{\partial \theta_{F_m}}, \quad (28)$$

for updating the m-th hard-to-learn task predictor, where $m \in \{1, \ldots, M_2\}$

$$\theta_{F_{m+M_1}} \leftarrow \theta_{F_{m+M_1}} - \mu \frac{\lambda_{m+M_1}}{M_2} \frac{\partial \mathcal{L}_{e,m}}{\partial \theta_{F_{m+M_1}}}, \quad (29)$$

for updating domain classifiers

$$\theta_{D_0} \leftarrow \theta_{D_0} - \mu \lambda_{D_0} \sum_{m=1}^{M} w_m \frac{\partial \mathcal{L}_{D_0,m}}{\partial \theta_{D_0}}, \quad (30)$$

$$\theta_{D_m} \leftarrow \theta_{D_m} - \mu \lambda_{D_m} \frac{\partial \mathcal{L}_{D_m}}{\partial \theta_{D_m}}, \quad (31)$$

where μ is the learning rate;

$$w_m = \frac{\exp \left(-\min_{\theta_{D_0}} \mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}) \right)}{\sum_{m=1}^{M} \exp \left(-\min_{\theta_{D_0}} \mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}) \right)}$$

is the adaptive weight for the m-th task. Task with larger distribution divergence has larger weight. The updates of Equation (28)–(31) are similar to those of deep neural network models using stochastic gradient descent (SGD). For Equation (26) and (27), the difference between the updates of them and SGD updates is that the gradients from the domain classifiers are subtracted. This subtracted gradient is accomplished by inserting the aforementioned GRL between feature extractors and domain classifiers. Specifically, the forward propagation of the GRL is the same as an identity transformation, and during back-propagation, the GRL changes the sign of the gradient (i.e., multiply it by a negative constant) before passing it to the preceding layer.[31] The optimization of our hierarchical multi-task and domain-adversarial learning algorithm is summarized in Algorithm 1.

Algorithm 1. Hierarchical multi-task and domain-adversarial learning algorithm

---

Input: - Training iterations: P;
     - Batch size: K;
     - Number of tasks: M;
     - The first M_1 tasks are easy-to-learn tasks, the last M_2 tasks are hard-to-learn tasks, and $M = M_1 + M_2$.
     - Number of classes for the tasks: C_m for $m \in 1, \ldots, M$;
     - Hyper-parameters: $\lambda_1, \ldots, \lambda_M, \lambda_{D_0}, \lambda_{D_1}, \ldots, \lambda_{D_{M_2}}$.
Output: Neural network: $\{\mathbf{W}_0, \mathbf{D}_0, \mathbf{W}_1 \cdots, \mathbf{W}_{M_2}, \mathbf{F}_1, \ldots, \mathbf{F}_M, \mathbf{D}_1, \ldots, \mathbf{D}_{M_2}\}$
1: Randomly initialize network parameters: $\theta_{W_0}, \theta_{D_0}, \theta_{W_1} \cdots, \theta_{W_{M_2}}, \theta_{F_1}, \ldots, \theta_{F_M}, \theta_{D_1}, \ldots, \theta_{D_{M_2}}$
2: **for** p from 1 to P **do**
3:   # Forward pass
4:   **for** m from 1 to M **do**
5:      Sample K data points and label from the source domain $\mathcal{D}_{S,m}$ for each task: $(X_{S,m}, Y_{S,m})$.
6:      Sample K data points from the target domain $\mathcal{D}_t$ for each task: $X_{T,m}$.
7:      Compute the task-shared feature for the source domain data: $Z_{S,m} = \mathbf{W}_0(X_{S,m})$.
8:      Compute the task-shared feature for the target domain data: $Z_{T,m} = \mathbf{W}_0(X_{T,m})$.
9:   **end for**
10: **for** m from 1 to M_2 **do**
11:     Compute the task-specific feature for the source domain hard-to-learn tasks: $Z_{S,m}^h = \mathbf{W}_m(Z_{S,m+M_1})$.
12:     Compute the task-specific feature for the target domain hard-to-learn tasks: $Z_{T,m}^h = \mathbf{W}_m(Z_{T,m+M_1})$.
13: **end for**
14: # Backward pass
15: Update the feature extractors and task predictors using equation (22):

$$
\left(\widehat{\theta}_{W_0}, \widehat{\theta}_{W_1}, \ldots, \widehat{\theta}_{W_{M_2}}, \widehat{\theta}_{F_1}, \ldots, \widehat{\theta}_{F_M}\right)
$$

$$
= \underset{\substack{\theta_{W_0}, \theta_{W_1} \cdots, \theta_{W_{M_2}}, \\ \theta_{F_1}, \ldots, \theta_{F_M}}}{\operatorname{argmin}} \left[\frac{1}{M_1} \sum_{m=1}^{M_1} \lambda_m \mathcal{L}_{e,m}(\theta_{W_0}, \theta_{F_m}) + \frac{1}{M_2} \sum_{m=1}^{M_2} \lambda_{m+M_1} \mathcal{L}_{h,m}\left(\theta_{W_0}, \theta_{W_m}, \theta_{F_{m+M_1}}\right) \right.
$$

$$
+ \lambda_{D_0} \log \sum_{m=1}^{M} \exp\left(-\min_{\theta_{D_0}} \mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}) \right)
$$

$$
\left. - \min_{\theta_{D_1}, \ldots, \theta_{D_M}} \sum_{m=1}^{M_2} \lambda_{D_m} \mathcal{L}_{D_m}(\theta_{W_0}, \theta_{W_m}, \theta_{D_m}) \right],
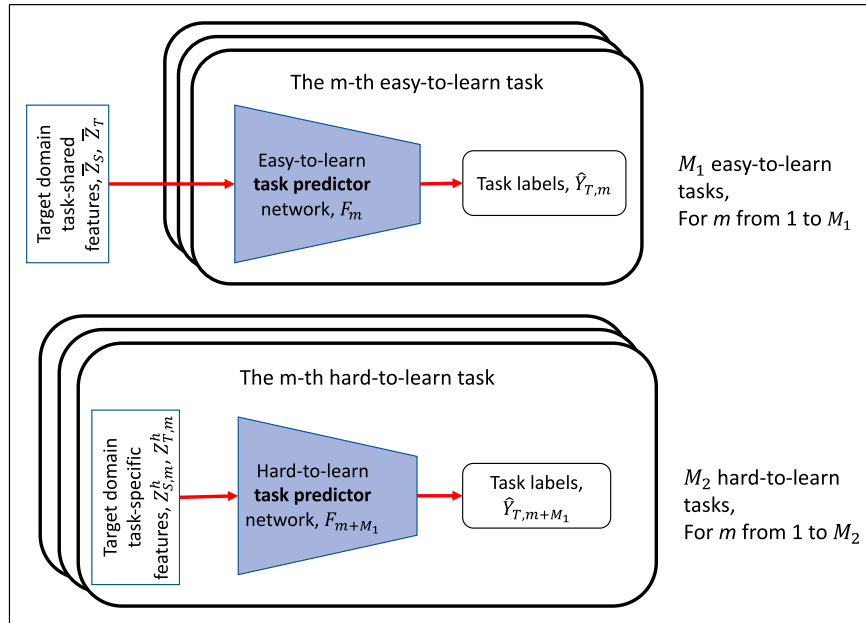$$

(32)

16: Update the task-shared domain classifier: $\widehat{\theta}_{D_0} = \operatorname{argmax}_{\theta_{D_0}} \log \sum_{m=1}^{M} \exp(-\mathcal{L}_{D_0,m}(\theta_{W_0}, \theta_{D_0}))$.
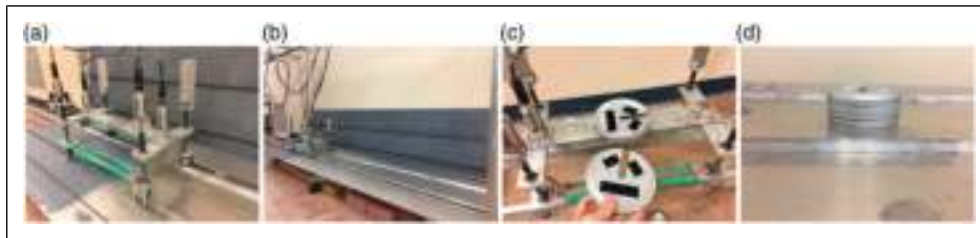17: **for** m in $1:M_2$ **do**
18:     Update the task-specific domain classifier: $\widehat{\theta}_{D_m} = \operatorname{argmin}_{\theta_{D_m}} \lambda_{D_m} \mathcal{L}_{D_m}(\widehat{\theta}_{W_0}, \widehat{\theta}_{W_m}, \theta_{D_m})$.
19: **end for**
20: **end for**

Figure 5. The architecture of our hierarchical multi-task and domain-adversarial learning algorithm in the target domain prediction module.

Figure 6. (a) A vehicle (V1) moving at a controlled speed and (b) a bridge (B1) that the vehicle passes. Hook-and-loop fasteners are used to attach the mass on (c) the vehicle and (d) the bridge.

Target domain prediction module

The architecture of the target domain prediction module is shown in Figure 5. In this module, the extracted target domain task-shared features are input to the pre-trained easy-to-learn predictors to predict target domain labels in the easy-to-learn tasks, and the extracted target domain task-specific features are input to the pre-trained hard-to-learn predictors to predict target domain labels in the hard-to-learn tasks.

Evaluation

In this section, we evaluate our HierMUD framework for drive-by BHM using data collected from lab-scale experiments. The experiments are conducted on two structurally different bridges using three vehicles of different weights.

Experimental setup and data description

A lab-scale VBI system, as shown in Figure 6, was employed to create the dataset. The collected dataset is subjected to environmental noise (e.g. different temperatures and humidity), electrical noise from the sensing system, and the noise caused by the friction between vehicle wheels and bridge surface. The experiments involved two 8-foot bridges (B1 and B2) with different weights of 34.2 lb and 43.0 lb, different dominant frequencies of 5.9 Hz and 7.7 Hz, and damping ratios of 0.13 and 0.07, respectively. The data are collected from three small-scale vehicles (V1, V2, and V3) with different weights of 10.6 lb, 11.6 lb, and 12.6 lb, respectively, that were driven over the bridge. Vertical acceleration signals were collected from four accelerometers mounted on each vehicle (front chassis, back chassis, front wheel, and back wheel) while they moved

individually across the bridge at a constant speed (0.75 m/s). The sampling rate of all sensors is 1600 Hz.

Damage proxy is introduced by adding mass at different locations of the bridge. Hook-and-loop fasteners are used to attach the mass on the vehicle and the bridge (as shown in Figure 6). For varying damage severity, the magnitude of the attached mass for each run ranged from 0.5 lb to 2.0 lb with an interval of 0.5 lb. A heavier mass means more severe damage since it induces more significant structural change from the initial condition (i.e., healthy state). Each damage severity level was induced at three different damage locations (l is every quarter of the bridge span, i.e., every 2 feet along the span). For each damage severity and location scenario for each vehicle and bridge combination, the experiments were repeated 30 times (i.e., 30 trials of a vehicle passing a bridge). In total, the dataset includes 2 (bridges) × 3 (vehicles) × [3 (damage locations) × 4 (damage severity levels) + 1 (undamaged case)] × 30 (iterations) = 2340 (trials), which results in 2340 (trials) × 4 (sensors) = 9360 (records). Details of the experimental instrumentation can be found in the work of Liu et al.[10]

In summary, our drive-by BHM problem has three tasks: binary damage detection, 3-class damage localization, and 4-class damage quantification. For each damage diagnostic task, two model transfers, from B1 to B2 and from B2 to B1, using signals collected from each of the three vehicles (V1, V2, and V3), are conducted, making a total of six evaluations.

Setup of our HierMUD framework

In this subsection, we describe our data pre-processing procedure and the setup of our HierMUD framework. We first pre-processed the input data by conducting data augmentation. Data augmentation adds white noise to vehicle acceleration signals. The white noise has zero mean and variance of mean squared magnitude of the vehicle acceleration signal. Then, we compute the short-time Fourier transform (STFT) representation of each signal to preserve the time–frequency domain information. We use the STFT representation because the frequency spectrum is sensitive to bridge characteristics and STFT allows us to preserve the location information of the moving vehicle and the bridge damage through its time–frequency domain representation.[10] The size of each input data is $C \times W \times H$, where C is the number of sensor channels on the vehicle; W and H are respectively the number of time segments and the sample frequencies of the STFT representation. Particularly, in our system, the original time-domain acceleration signal's dimension is 4 × 4000, and the processed STFT data's dimension is 4 × 64 × 64. Figure 7 shows an example of the vehicle acceleration signal and its corresponding STFT representation.

We consider damage detection and localization as easy-to-learn tasks and damage quantification as a hard-to-learn task. This is because we obtained higher supervised prediction accuracy for damage detection and localization tasks than that for the quantification task, and we observed that the data distributions for different damage locations are more separable than that for different damage severity levels. We then develop the hierarchical architecture that extracts task-shared features for all the tasks and further task-specific features for the quantification task. Moreover, the existence of damage (for damage detection task) is represented by introducing an additional label within the location predictor, instead of creating an additional binary damage detection classifier that increases the model complexity. After all, the overall architecture of the neural network modules used in our HierMUD model are shown in Figure 8 and Table 1. We optimize our HierMUD model using a stochastic gradient descent optimizer with weight decay, which encourages smaller weights of the networks by adding an L2-regularization term to the loss to avoid overfitting. The network architecture and hyper-parameters are determined using the unsupervised hyper-parameter selection method, which is described below. Based on our experiments, the architecture shown in Table 1 provides us with the best damage diagnosis results.

Baseline methods

We compare the performance of HierMUD with five baseline methods described below:

1) MCNN is a multi-task convolutional neural network model that directly applies the model trained using source domain data to the target bridge without applying any domain adaptation. The architecture of MCNN lacks the red boxes (domain classifiers) in Figure 8.
2) iUD is an independent task learning model with UDA. It predicts damage location and severity levels using two independent UDA models. The architecture of iUD does not have the task-specific feature extractor, and instead it uses two independent domain adversarial neural networks (DANN)[31] for damage localization and quantification.
3) sUD is a sequential task learning model with UDA. It predicts damage location and severity level step-by-step instead of using multi-task learning. Specifically, sUD uses one DANN for damage localization, and for each predicted damage location, it uses one DANN for damage quantification.
4) MUD directly combines MTL and UDA. It predicts damage location and severity level simultaneously, but does not use the hierarchical structure and the
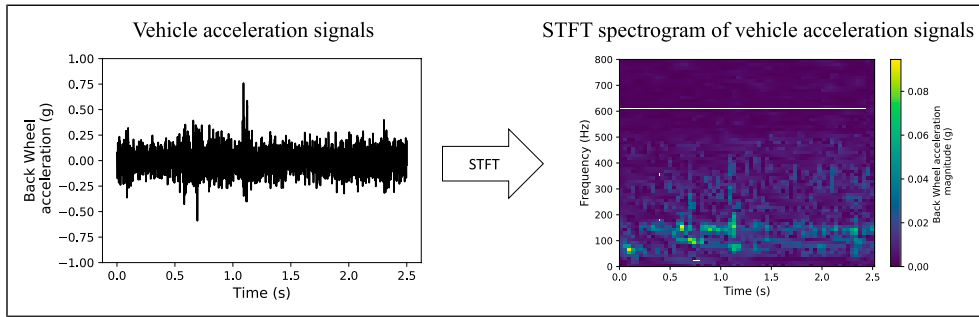
Figure 7. An example of the raw acceleration signal collected from a vehicle passing over a lab-scale bridge and its corresponding STFT representation.
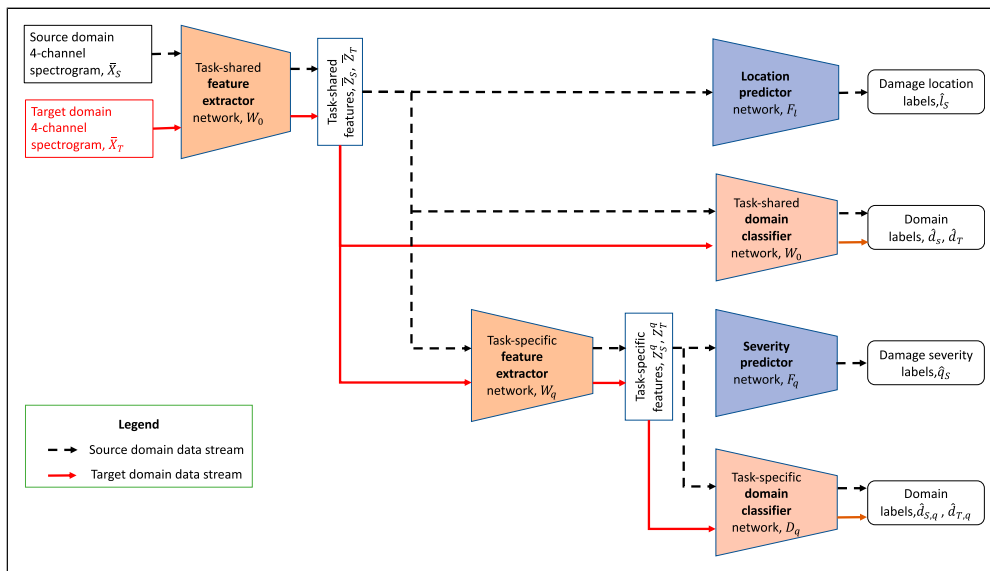
Figure 8. The architecture of our multi-task and domain adversarial learning algorithm applied on the drive-by BHM.

Table 1. Detailed network architectures of HierMUD.

Network	Layer	Patch size	Input size	Activation
Task-shared feature extractor	Convolution (2D)	64 × 5 × 5	4 × 64 × 64	
	Max pooling	2 × 2	64 × 60 × 60	LeakyReLU
	Convolution (2D)	50 × 5 × 5	64 × 30 × 30	
	Max pooling	2 × 2	64 × 26 × 26	LeakyReLU
	Convolution (2D)	50 × 3 × 3	50 × 13 × 13	
	Max pooling	2 × 2	50 × 11 × 11	LeakyReLU
Task-specific feature extractor	Flatten		50 × 5 × 5	
	Full connection	1250 × 1250	1250 × 1	ReLU
Location predictors	Flatten		50 × 5 × 5	
	Full connection	1250 × 100	100 × 1	ReLU
	Full connection	100 × 4	100 × 1	Soft-max
Severity predictors	Flatten		50 × 5 × 5	
	Full connection	1250 × 100	100 × 1	ReLU
	Full connection	100 × 5	100 × 1	Soft-max
Task-shared and task-specific domain classifiers	Flatten		50 × 5 × 5	
	Full connection	1250 × 2	1250 × 1	Soft-max

Table 2. The comparison of our method and baseline methods. MTL, ITL, STL, and UDA stand for multi-task learning, independent task learning, sequential task learning, and unsupervised domain adaptation, respectively.

Method	MTL	ITL	STL	UDA	Hierarchical structure	Soft-max objective
MCNN	✓					
iUD		✓		✓		
sUD			✓	✓		
MUD	✓			✓		
HierMUD-a	✓			✓	✓	
HierMUD (ours)	✓			✓	✓	✓

soft-max objective. The architecture of MUD can be found in Liu et al.[60]

5) HierMUD-a further adds the hierarchical structure to MUD, but it optimizes the average objective—Equation 20.

The differences of these methods are also summarized in Table 2. Note that to have a fair comparison, components (feature extractors, task predictors, and domain classifiers) of each method have the same architecture as HierMUD method. We also use the same strategy to train these methods, which is described in the following section.

Unsupervised hyper-parameter selection

In this section, we describe an approach to fine-tune hyper-parameters, including the learning rate, network architecture, and trade-off hyper-parameters (as shown in Equation 22), for performing our multi-task and domain-adversarial learning algorithm. Because there is no labeled data in the target domain during the training phase, we select hyper-parameters by conducting a reverse validation.[31,61] In short, the labeled source samples $\overline{X}_S$ and unlabeled target samples $\overline{X}_T$ are split into training sets ($\overline{X}'_S$ and $\overline{X}'_T$) which contain 90% of $\overline{X}_S$ and $\overline{X}_T$ (i.e., 8424 labeled data from the source domain and 8424 unlabeled data from the target domain) and the validation sets ($\overline{X}^V_S$ and $\overline{X}^V_S$), where $\overline{X}' \cup \overline{X}^V = \overline{X}$ and they are mutually exclusive. The training sets $\overline{X}'_S$ and $\overline{X}'_T$ are used to learn a target domain prediction model $\eta(\cdot)$. We then learn another prediction model $\eta_r(\cdot)$ using unlabeled target domain data, $\overline{X}'_T$, with the predicted labels ($\eta(x)$ for $x \in \overline{X}'_T$). The reverse classifier $\eta_r(\cdot)$ is evaluated on the source domain validation set $\overline{X}^V_S$. This evaluation process is conducted for several times with different values of hyper-parameters.

Tuning the trade-off hyper-parameters ($\lambda_1, \ldots, \lambda_M, \lambda_{D_0}, \lambda_{D_1}, \ldots, \lambda_{D_{M_2}}$) is important because they affect the convergence of HierMUD's training and damage diagnosis accuracy by changing the magnitude of gradient from each component of HierMUD. In particular, we observe that using relatively large task-informativeness trade-off hyper-parameters (λ_m, for $m \in 1, \ldots, M$) and relatively small domain-invariance trade-off hyper-parameters (λ_{D_0} and λ_{D_m}, for $m \in 1, \ldots, M_2$) speeds up the convergence of HierMUD's training but reduces the target domain prediction accuracy, and vice versa.

In our drive-by BHM problem, we set the hyper-parameter for quantification loss λ_2 to be one and tune the hyper-parameter for localization loss λ_1 by searching its optimal value in the set {0.01, 0.1, 0.5, 1} that provides the best reverse validation accuracy. For λ_{D_0} and λ_{D_1}, we tune them by searching the values of them in the same set {0.01, 0.1, 0.5, 1}. Thus, in total, we have $4 \times 4 \times 4 = 64$ possible combinations of hyper-parameters. We also conduct 10-fold cross-validation with each set of hyper-parameters and choose the hyper-parameters that give the highest cross-validation accuracy for the reverse classifier on the source domain validation set. For our experiment, the best choices of $\lambda_1, \lambda_2, \lambda_{D_0}$, and λ_{D_1} are 0.5, 1, 0.1, and 0.5, respectively.

Results and discussion

In this section, we present our evaluation results and discuss our findings. Table 3 presents the results for knowledge transfer from B1 to B2 and from B2 to B1 using each of the three vehicles' data. For the binary damage detection, we report F1-scores as the performance metric because the numbers of damaged and undamaged data samples are imbalanced; For the 3-class damage localization, we report classification accuracy (the number of correct predictions divided by the total number of predictions made) as the performance metric; For the damage quantification task, we report the mean absolute error (MAE) between predictions and ground truth damage severity. Note that we first obtain damage severity estimation through the 4-class classification and then compute MAE by taking the absolute difference between the predicted label and the ground truth. MAE is chosen because it shows the degree of error rather than classification accuracy (i.e., only showing correct vs incorrect). There are six evaluations in total (for 2 bridges and 3 vehicles), and each evaluation was conducted 10 times with different random seeds for splitting train and validation sets and initializing model parameters. Therefore, we have total 60 tests. Each column in Table 3 shows the performance for each of the six evaluations. The last column of the table provides the overall performance of all six

Table 3. The performance of baseline methods and our method on the lab-scale VBI dataset. "A → B" indicates that we transfer knowledge from the source domain A to the target domain B. The numbers in each cell are: average performance (confidence interval). We bold the number that has the best result in each task.

Task (Metric)	Method	Vehicle 1		Vehicle 2		Vehicle 3		Overall
		B1 → B2	B2 → B1	B1 → B2	B2 → B1	B1 → B2	B2 → B1	
Damage Detection (F1-score)	MCNN	0.46 (±0.02)	0.54 (±0.08)	0.45 (±0.05)	0.53 (±0.07)	0.44 (±0.04)	0.51 (±0.08)	0.49 (±0.03)
	iUD	0.57 (±0.16)	0.83 (±0.19)	0.61 (±0.19)	0.98 (±0.01)	0.94 (±0.07)	0.84 (±0.19)	0.80 (±0.08)
	sUD	0.53 (±0.15)	**0.97 (±0.03)**	0.76 (±0.20)	0.97 (±0.03)	0.76 (±0.20)	0.84 (±0.19)	0.81 (±0.07)
	MUD	0.91 (±0.05)	0.92 (±0.02)	0.98 (±0.02)	0.95 (±0.03)	0.92 (±0.07)	0.95 (±0.02)	0.94 (±0.02)
	HierMUD-a	**0.94 (±0.06)**	0.93 (±0.04)	**0.99 (±0.01)**	0.94 (±0.03)	0.95 (±0.03)	**0.97 (±0.02)**	**0.96 (±0.01)**
	HierMUD	0.92 (±0.04)	0.92 (±0.05)	**0.99 (±0.01)**	0.98 (±0.01)	0.96 (±0.02)	0.96 (±0.03)	0.95 (±0.01)
Damage Localization (Accuracy)	MCNN	0.35 (±0.05)	0.34 (±0.04)	0.38 (±0.07)	0.36 (±0.07)	0.40 (±0.04)	0.26 (±0.02)	0.35 (±0.02)
	iUD	0.62 (±0.13)	0.58 (±0.21)	0.61 (±0.15)	0.92 (±0.15)	0.90 (±0.09)	0.53 (±0.20)	0.69 (±0.08)
	sUD	0.70 (±0.08)	0.72 (±0.24)	0.49 (±0.14)	0.72 (±0.24)	0.86 (±0.11)	0.59 (±0.21)	0.68 (±0.08)
	MUD	0.85 (±0.12)	0.86 (±0.04)	0.84 (±0.10)	0.87 (±0.07)	0.95 (±0.04)	0.95 (±0.02)	0.89 (±0.03)
	HierMUD-a	0.89 (±0.05)	**0.88 (±0.05)**	0.82 (±0.11)	0.91 (±0.04)	0.97 (±0.03)	0.90 (±0.09)	0.90 (±0.03)
	HierMUD	**0.92 (±0.03)**	0.87 (±0.05)	**0.92 (±0.06)**	**0.93 (±0.04)**	**0.98 (±0.01)**	**0.95 (±0.04)**	**0.93 (±0.02)**
Damage Quantification (MAE, lbs)	MCNN	0.70 (±0.12)	0.75 (±0.11)	0.76 (±0.11)	0.71 (±0.07)	0.73 (±0.06)	0.77 (±0.11)	0.74 (±0.04)
	iUD	0.59 (±0.11)	0.60 (±0.08)	0.52 (±0.07)	0.55 (±0.08)	0.47 (±0.04)	0.48 (±0.08)	0.53 (±0.03)
	sUD	0.65 (±0.07)	0.63 (±0.06)	0.56 (±0.06)	0.57 (±0.08)	0.52 (±0.05)	0.58 (±0.06)	0.58 (±0.03)
	MUD	0.53 (±0.10)	0.55 (±0.10)	0.46 (±0.06)	0.49 (±0.11)	0.42 (±0.07)	0.41 (±0.12)	0.47 (±0.04)
	HierMUD-a	0.54 (±0.10)	0.47 (±0.07)	0.38 (±0.05)	0.44 (±0.07)	0.40 (±0.11)	0.46 (±0.08)	0.44 (±0.04)
	HierMUD	**0.40 (±0.10)**	**0.45 (±0.09)**	**0.37 (±0.07)**	**0.43 (±0.09)**	**0.29 (±0.09)**	**0.36 (±0.07)**	**0.38 (±0.04)**

evaluations. The numbers in each cell of the table are: average performance (± 95% confidence interval). To show the relationship between true damage values and predictions, we also present normalized confusion matrices of our approach for the damage quantification and localization tasks in Figures 15 and 16 (in Appendix B), respectively.

Performance comparison with baseline methods

Except for the damage detection task for the evaluation using V1 response to transfer model from B2 to B1, our method (HierMUD) outperforms other baselines. Overall, HierMUD has the best damage localization accuracy, which is 93% on average. HierMUD also achieves the smallest damage quantification error, which is 0.38 lbs on average that is within one damage severity level of 0.5 lbs. The average objective (HierMUD-a) has the best damage detection F1-score (96%), followed by HierMUD (95%). HierMUD is about twice as good as the baseline without domain adaptation (MCNN) in all three diagnostic tasks and 1.5 times as good as other baselines (iUD, sUD, and MUD), in the hard-to-learn quantification task. HierMUD achieves the best 99% average F1-score (up to 100% in the best test) in the damage detection task using signals collected from V2 for model transfer from B1 to B2, the best 98% average accuracy (up to 100% in the best test) in the damage localization task,

and the smallest 0.29 lbs average MAE (0.21 lbs in the best test) in the damage quantification task using signals collected from V3 for model transfer from B1 to B2.

Figures 9, 10, and 11 show the boxplots of performance metric in damage detection, localization, and quantification tasks, respectively. In each figure set, boxplot (a) shows the overall 60 test results of our method and baselines. Boxplots (b) and (c) present 30 test results for model transfer from B1 to B2 and from B2 to B1, respectively. Each box in the boxplot shows five values, including minimum, maximum, median, the first quartile, and the third quartile. In addition to the results from Table 3, we observe that the results of our method have smaller variance (i.e., smaller box in boxplots) than that of baseline methods, which indicates a more stable performance with different random initialization. Furthermore, for the damage quantification task using our method, results for model transfer from B1 to B2 are better than those for model transfer from B2 to B1. This difference is coincidental with or due to the fact that different model transfer directions have distinct model generalizability (i.e., the model learned from B1 data is more generalizable and easy to transfer than that learned from B2).

In summary, our method outperforms baselines in comprehensive laboratory evaluations, which shows the effectiveness of combining domain adversarial, multi-task learning with the hierarchical architecture and the soft-max objective.
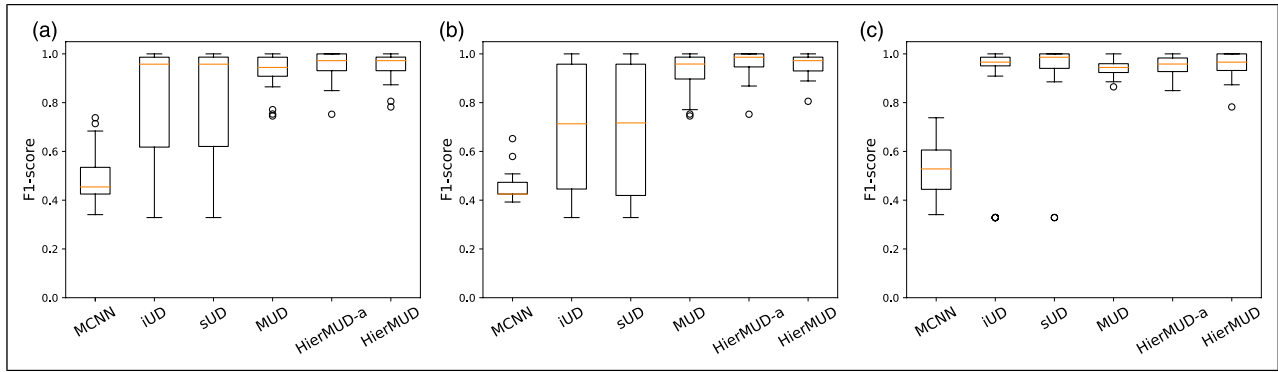
Figure 9. The performance ((a) overall performance; (b) B1 to B2; (c) B2 to B1) of baseline methods and our method for damage detection task on the lab-scale VBI dataset.
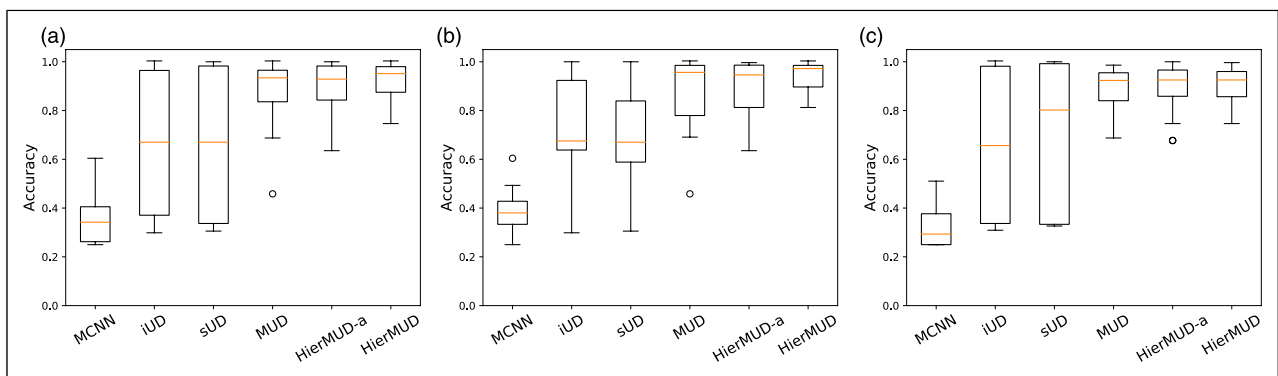
Figure 10. The performance ((a) overall performance; (b) B1 to B2; (c) B2 to B1) of baseline methods and our method for damage localization task on the lab-scale VBI dataset.
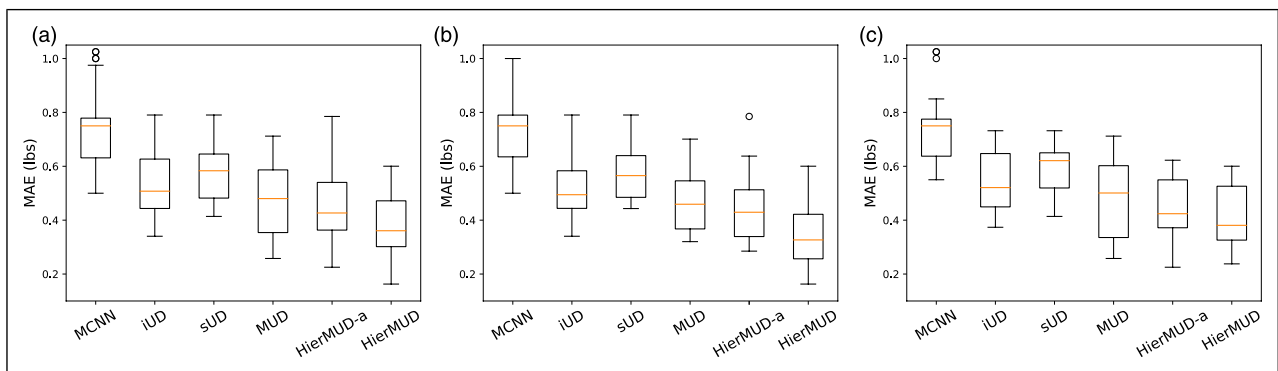

Figure 11. The performance ((a) overall performance; (b) B1 to B2; (c) B2 to B1) of baseline methods and our method for damage quantification task on the lab-scale VBI dataset. Note that our damage severity levels are 0.5, 1.0, 1.5, and 2.0 lbs.

Characterizing the extracted feature distributions

In this section, we characterize the extracted features to show the effect of domain adaptation on the distribution of the extracted features. Figure 12 and 13 shows the feature distributions plotted using tSNE to project feature distributions at different feature layers of the network into a two-dimensional feature space. Figure 12 (a) shows the tSNE projection of the task-shared features ($\overline{Z}_S$ and $\overline{Z}_T$) after domain adaptation for the damage localization task; Figure 12 (b) shows the tSNE project of the source and target domains' task-specific features (Z_S^h and Z_T^h) after domain adaptation for the damage quantification task. Different colors indicate different
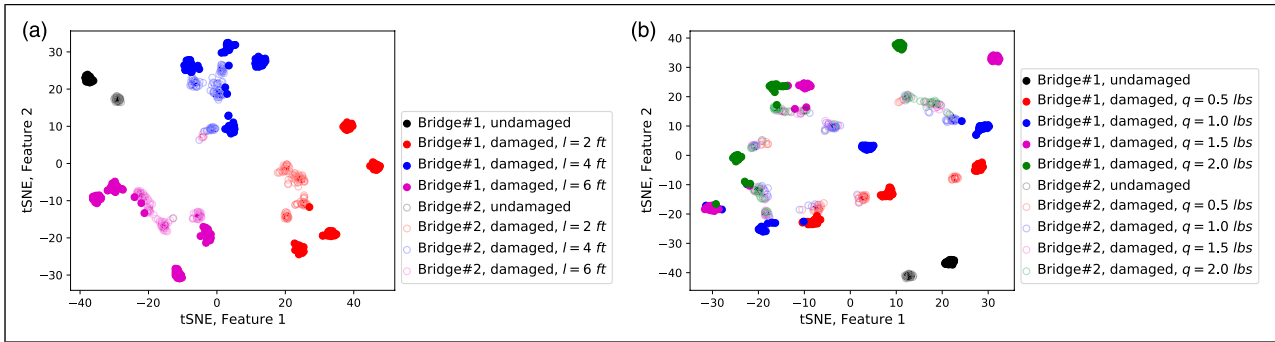
Figure 12. The 2D tSNE visualization of (a) the task-shared features after domain adaptation in the damage localization task and (b) the task-specific features after domain adaptation in the damage quantification task for the evaluation of model transfer from Bridge#1 (B1) to Bridge#2 (B2).
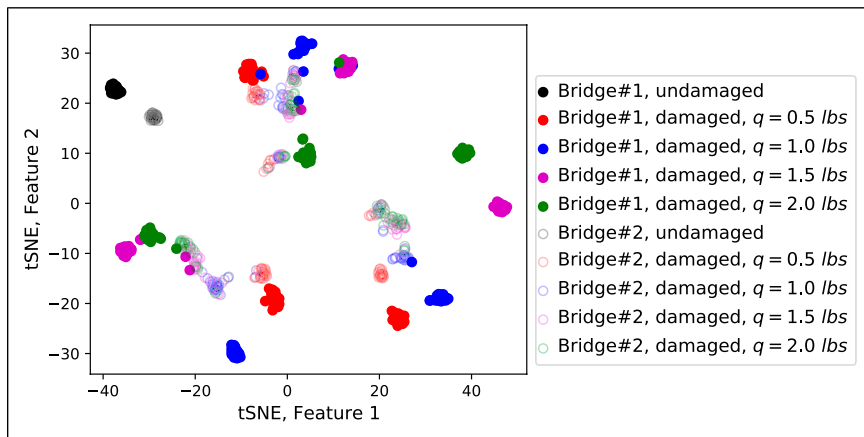
Figure 13. The 2D tSNE visualization of the task-shared features after domain adaptation in the damage quantification task for the evaluation of model transfer from Bridge#1 (B1) to Bridge#2 (B2).
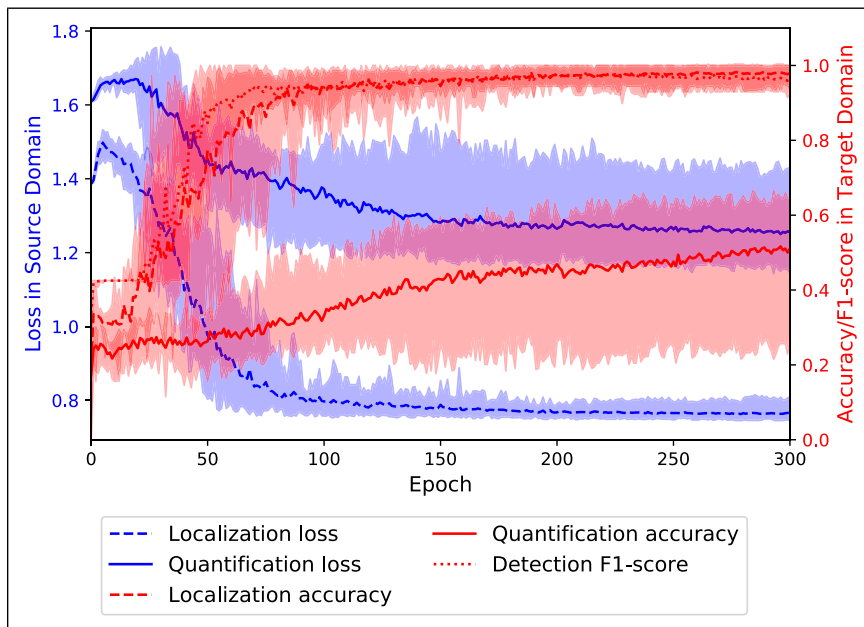
Figure 14. The losses of damage localization and quantification in the source domain, and the accuracy/F1-score of damage detection, localization, and quantification in the target domain.

damage labels. Filled and unfilled markers represent features of source and target domains, respectively. Comparing the distributions of features before and after domain adaptation in Figure 2 and 12, we observe that our domain adaptation method successfully transformed the feature distributions of the two domains to be much more similar to each other. Also, comparing Figures 12 (a) and (b), the distributions of the features are more similar and thus better matched for the easy-to-learn damage localization task than that for the hard-to-learn damage quantification task.

Further, Figure 13 shows the same tSNE projection of the task-shared features ($\overline{Z}_S$ and $\overline{Z}_T$) but color-coded with different damage severity levels. The distributions of task-shared features having the same damage severity label in Figure 13 (i.e., the markers with the same color) are more spread out than those in Figure 12 (b), making features in Figure 13 more difficult to be correctly classified. Therefore, task-specific features are more task-informative than task-shared features for the quantification task, which corresponds to our results that using the task-specific hierarchical features outperforms using the task-shared features for damage quantification.

Characterizing training process

To characterize the training process, we show the training losses and testing accuracy for model transfer from B1 to B2 using V2 data in Figure 14. Blue curves are average losses of the two task predictors in the source domain, and red curves are average accuracy/F1-score of damage detection, localization, and quantification in the target domain. The envelope covers loss or accuracy for the 10 tests with different random initialization. The figure shows that the two losses fluctuate and even increase during early training steps (i.e., when the number of epochs is small) due to domain adversarial learning. They converge to small loss values after around 200 epochs. We also observe that the quantification loss and accuracy have a large area of the envelope because it is difficult to find optimized saddle points of model parameters for this hard-to-learn quantification task.

Conclusion and future work

In this work, we introduce HierMUD, a new multi-task unsupervised domain adaptation framework that uses drive-by vehicle acceleration responses to diagnose structural damage of multiple bridges without requiring labeled data from every bridge. Our framework jointly optimizes hierarchical feature extractors, damage predictors, and domain classifiers in an adversarial way to extract features that are task-informative and domain-invariant. To match distinct data distributions over multiple tasks, we introduce a novel loss function based on a new provable generalization risk bound to adaptively provide a larger gradient magnitude on matching tasks having more shifted distribution between the source and target domains. To learn multiple tasks with

varying task difficulties, the feature extractor is designed to formulate a feature hierarchy, which learns two-level features: task-shared features are used for learning easy-to-learn tasks, and task-specific features are further extracted from task-shared features to learn hard-to-learn tasks. To the best of our knowledge, this is the first framework to transfer the model learned from one bridge to detect, localize, and quantify damage to another (target) bridge without any labels of the target bridge. We evaluate our framework on experimental data collected from three vehicles passing over two bridges individually. Our framework is up to 2X better than baseline methods and has smaller prediction variance. It achieves average accuracy values of 95% for damage detection, 93% for damage localization, and MAE 0.38 lbs for damage quantification.

In future work, we plan to investigate how the model transfer direction affects the domain adaptation performance. For instance, we will investigate what characteristics of source domain bridges lead to a better performance when the model is transferred to a target domain. In addition, we intend to study how to transfer our multi-task damage diagnosis model between bridges with different structural forms (e.g., truss bridge, suspension bridge, etc.). Moreover, we believe this framework can be generalized to other multi-task UDA problems beyond structural health monitoring. Thus, we plan to apply our framework in other applications (e.g., visual recognition, machine monitoring, etc.) to further evaluate its robustness and effectiveness.

ORCID iD

Jingxiao Liu https://orcid.org/0000-0001-5559-1627

References

1. ASCE. Bridges-infrastructure report card, https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Bridges-Final.pdf. (2017).
2. Hartle RA, Ryan TW, Mann E, et al. Bridge inspector's reference manual: Volume 1, volume 2. *United States: Federal Highway Administration, 2002. Technical report*.
3. Sun L, Shang Z, Xia Y, et al. Review of bridge structural health monitoring aided by big data and artificial intelligence:

from condition assessment to damage detection. _J Struct Eng_ 2020; 146(5): 04020073.

4. Yang Y, Wang ZL, Shi K, et al. State-of-the-art of the vehicle-based methods for detecting the various properties of highway bridges and railway tracks. _Int J Struct Stab Dyn_ 2020a: 2041004.

5. Yang YB, Yang JP, Zhang B and Wu Y. _Vehicle Scanning Method for Bridges_. Wiley Online Library, 2020b.

6. Cerda F, Chen S, Bielak J, et al. Indirect structural health monitoring of a simplified laboratory-scale bridge model. _Smart Structures Syst_ 2014; 13(5): 849–868.

7. Lederman G, Wang Z, Bielak J, et al (2014) Damage quantification and localization algorithms for indirect shm of bridges. In: _Proc. Int. Conf. Bridge Maint., Safety Manag._, Shanghai, China.

8. Xu S, Zhang L, Zhang P, et al (2016) An indirect traffic monitoring approach using building vibration sensing system. In: Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM. pp. 374–375.

9. Xu S, Zhang L, Zhang P, et al. An information-theoretic approach for indirect train traffic monitoring using building vibration. _Front Built Environ_ 2017; 3: 22.

10. Liu J, Chen S, Bergés M, et al. Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction. _Mech Syst Signal Process_ 2020c; 136: 106454.

11. Yang YB, Lin C and Yau J. Extracting bridge frequencies from the dynamic response of a passing vehicle. _Journal of Sound and Vibration_ 2004; 272(3-5): 471–493.

12. Lin C and Yang Y. Use of a passing vehicle to scan the fundamental bridge frequencies: An experimental verification. _Eng Structures_ 2005; 27(13): 1865–1878.

13. Liu J, Xu S, Bergés M, et al. An Expectation-Maximization Algorithm-Based Framework for Vehicle-Vibration-Based Indirect Structural Health Monitoring of Bridges. _Structural Health Monitoring_, 2019b.

14. Yang Y, Li Y and Chang KC. Constructing the mode shapes of a bridge from a passing vehicle: a theoretical study. _Smart Structures Syst_ 2014; 13(5): 797–819.

15. Malekjafarian A and OBrien EJ. Identification of bridge mode shapes using short time frequency domain decomposition of the responses measured in a passing vehicle. _Eng Structures_ 2014; 81: 386–397.

16. González A, OBrien EJ and McGetrick P. Identification of damping in a bridge using a moving instrumented vehicle. _J Sound Vibration_ 2012; 331(18): 4115–4131.

17. Nguyen KV and Tran HT. Multi-cracks detection of a beam-like structure based on the on-vehicle vibration signal and wavelet analysis. _J Sound Vibration_ 2010; 329(21): 4455–4465.

18. McGetrick PJ and Kim CW. A parametric study of a drive by bridge inspection system based on the morlet wavelet. in: Key Engineering Materials, volume 569. Trans Tech Publ, 2013, pp. 262–269.

19. Liu J. Scalable bridge health monitoring using drive-by vehicles: Phd forum abstract. In: Proceedings of the 18th Conference on Embedded Networked Sensor Systems 2020: pp. 817–818.

20. Eshkevari SS and Pakzad SN. Signal reconstruction from mobile sensors network using matrix completion approach. In: _Topics in Modal Analysis & Testing_, Volume 8. Springer; 2020, pp. 61–75.

21. Liu J, Chen B, Chen S, et al. Damage-sensitive and domain-invariant feature extraction for vehicle-vibration-based bridge health monitoring. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020b, pp. 3007–3011.

22. Mei Q, Gül M and Boay M. Indirect health monitoring of bridges using mel-frequency cepstral coefficients and principal component analysis. _Mech Syst Signal Process_ 2019; 119: 523–546.

23. Sadeghi Eshkevari S, Pakzad SN, Takáč M, et al. Modal identification of bridges using mobile sensors with sparse vibration data. _J Eng Mech_ 2020; 146(4): 04020011.

24. Zhang L and Gao X (2019) Transfer adaptation learning: A decade survey. _arXiv preprint arXiv:1903.04687_.

25. Jiang J and Zhai C. _Instance Weighting for Domain Adaptation in Nlp_. ACL, 2007.

26. Pan SJ, Tsang IW, Kwok JT, et al. Domain adaptation via transfer component analysis. _IEEE Trans Neural Networks_ 2010; 22(2): 199–210.

27. Cao Y, Long M and Wang J. Unsupervised domain adaptation with distribution matching machines. In: Proceedings of the AAAI Conference on Artificial Intelligence 2018; volume 32.

28. Luo YW, Ren CX, Dao-Qing D, et al. Unsupervised Domain Adaptation via Discriminative Manifold Propagation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

29. Xu S and Noh HY. Phymdan: Physics-informed knowledge transfer between buildings for seismic damage diagnosis through adversarial learning. _Mech Syst Signal Process_ 2021; 151: 107374.

30. Saito K, Watanabe K, Ushiku Y, et al. Maximum classifier discrepancy for unsupervised domain adaptation. In: Proceedings of the IEEE conference on computer vision and pattern recognition 2018. pp. 3723–3732.

31. Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks. _J Machine Learn Res_ 2016; 17(1): 2096–2030.

32. Zhang Y, Liu T, Long M, et al (2019) Bridging theory and algorithm for domain adaptation. _arXiv preprint arXiv: 1904.05801_.

33. Caruana R. Multitask learning. _Machine Learning_ 1997; 28(1): 41–75.

34. Luong MT, Le QV, Sutskever I, et al (2015) Multi-task sequence to sequence learning. _arXiv preprint arXiv:1511.06114_.

35. Augenstein I, Ruder S and Søgaard A (2018) Multi-task learning of pairwise sequence classification tasks over disparate label spaces. _arXiv preprint arXiv:1802.09913_.

36. Dong D, Wu H, He W, et al. Multi-task learning for multiple language translation. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Volume 1: *Long Papers*; 2015. pp. 1723–1732.

37. Hashimoto K, Xiong C, Tsuruoka Y, et al (2016) A joint many-task model: Growing a neural network for multiple nlp tasks. *arXiv preprint arXiv:1611.01587*.

38. Yang Y and Hospedales T (2016) Deep multi-task representation learning: A tensor factorisation approach. *arXiv preprint arXiv:1605.06391*.

39. Misra I, Shrivastava A, Gupta A, et al. Cross-stitch networks for multi-task learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition 2016. pp. 3994–4003.

40. Jou B and Chang SF. Deep cross residual learning for multitask visual recognition. In: *Proceedings of the 24th ACM international conference on Multimedia*; 2016, pp. 998–1007.

41. Wan HP and Ni YQ. Bayesian multi-task learning methodology for reconstruction of structural health monitoring data. *Struct Health Monit* 2019; 18(4): 1282–1309.

42. Liu J, Bergés M, Bielak J, et al. A damage localization and quantification algorithm for indirect structural health monitoring of bridges using multi-task learning. In: AIP Conference Proceedings, volume 2102. AIP Publishing LLC, 2019a, p. 090003.

43. Deraemaeker A and Worden K. A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mech Syst Signal Process* 2018; 105: 1–15.

44. Taddei T, Penn JD, Yano M, et al. Simulation-based classification; a model-order-reduction approach for structural health monitoring. *Arch Comput Methods Eng* 2018; 25(1): 23–45.

45. Nasrollahi A, Deng W, Ma Z, et al. Multimodal structural health monitoring based on active and passive sensing. *Struct Health Monit* 2018; 17(2): 395–409.

46. Paszke A, Gross S, Chintala S, et al. *Automatic Differentiation in Pytorch*, 2017.

47. Malekjafarian A, McGetrick PJ and OBrien EJ. A review of indirect bridge monitoring using passing vehicles 2015. *Shock and Vibration* 2015.

48. Van der Maaten L and Hinton G. Visualizing data using t-sne. *J Machine Learning Research* 2008; 9(11).

49. Jakubovitz D, Giryes R and Rodrigues MR. Generalization error in deep learning. In: *Compressed Sensing and Its Applications*. Springer; 2019, pp. 153–193.

50. Zhao H, Zhang S, Wu G, et al. Adversarial multiple source domain adaptation. In: Advances in neural information processing systems 2018. pp. 8559–8570.

51. Maurer A, Pontil M and Romera-Paredes B. The benefit of multitask representation learning. *J Machine Learn Res* 2016; 17(1): 2853–2884.

52. Ben-David S, Blitzer J, Crammer K, et al. A theory of learning from different domains. *Machine Learning* 2010; 79(1–2): 151–175.

53. Blumer A, Ehrenfeucht A, Haussler D, et al. Learnability and the vapnik-chervonenkis dimension. *J ACM (JACM)* 1989; 36(4): 929–965.

54. Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks. In: International conference on machine learning. PMLR, 2015, pp. 97–105.

55. Goodfellow I, Bengio Y, Courville A, et al. *Deep Learning*, volume 1. Cambridge: MIT press, 2016.

56. Shorten C and Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data* 2019; 6(1): 1–48.

57. Kenny J, Fluck A and Jetson T. Placing a value on academic work: The development and implementation of a time-based academic workload model. *Aust Universities' Rev* 2012; 54(2): 50–60.

58. Kenny JD and Fluck AE. The effectiveness of academic workload models in an institution: a staff perspective. *J Higher Edu Pol Manag* 2014; 36(6): 585–602.

59. Guo M, Haque A, Huang DA, et al (2018) Dynamic task prioritization for multitask learning. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 270–287.

60. Liu J, Bergés M, Bielak J, et al (2020a) Knowledge transfer between bridges for drive-by monitoring using adversarial and multi-task learning. *arXiv preprint arXiv:2006.03641*.

61. Zhong E, Fan W, Yang Q, et al. Cross validation framework to choose amongst models and datasets for transfer learning. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2010, pp. 547–562.

Appendix A

Theorem 3. *Let $\mathcal{W}$ be a hypothesis space on $\mathcal{X}$ with VC dimension d_W and $\mathcal{H}$ be a hypothesis space on $\mathcal{Z}$ with VC dimension d_H. If X_S and X_T are samples of size N from $\mathcal{D}_S^X$ and $\mathcal{D}_T^X$, respectively, and Z_S and Z_T follow distributions $\mathcal{D}_S^Z$ and $\mathcal{D}_T^Z$, respectively,, then for any $\delta \in (0, 1)$ with probability at least $1 - \delta$, for every $h \in \mathcal{H}$ and $w \in \mathcal{W}$*

$$\varepsilon_T(h \circ w; f_T) \leq \varepsilon_S(h \circ w; f_S) + 2\varepsilon_S(h \circ w^*, f_S)$$
$$+ \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(Z_T, Z_S)$$
$$+ \mathcal{O}\left(\sqrt{\frac{2d_H \log(2N) + \log(2/\delta)}{N}} \right)$$
$$+ \frac{1}{2} \sup_{\hat{h} \in \mathcal{H}} \left[d_{\hat{h}, \mathcal{W} \Delta \mathcal{W}}(X_S, X_T) \right]$$
$$+ \mathcal{O}\left(\sqrt{\frac{2d_W \log(2N) + \log(2/\delta)}{N}} \right)$$
$$+ \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_S(h^* \circ w^*; f_S), \tag{33}$$

Proof. This proof relies on the triangle inequality for classification error

$$\varepsilon_T(h \circ w; f_T) \leq \varepsilon_T(h \circ w; h^* \circ w^*) + \varepsilon_T(h^* \circ w^*; f_T)$$
$$\leq \varepsilon_T(h \circ w; h \circ w^*) + \varepsilon_T(h \circ w^*; h^* \circ w^*) + \varepsilon_T(h^* \circ w^*; f_T)$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_T(h \circ w; h \circ w^*) + \varepsilon_S(h \circ w^*; h^* \circ w^*) + |\varepsilon_T(h \circ w^*; h^* \circ w^*) - \varepsilon_S(h \circ w^*; h^* \circ w^*)|$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_T(h \circ w; h \circ w^*) + \varepsilon_S(h \circ w^*; h^* \circ w^*) + \sup_{h, h^* \in \mathcal{H}} |\varepsilon_T(h \circ w^*; h^* \circ w^*) - \varepsilon_S(h \circ w^*; h^* \circ w^*)|$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_T(h \circ w; h \circ w^*) + \varepsilon_S(h \circ w^*; h^* \circ w^*) + \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(\mathcal{D}_S^Z, \mathcal{D}_T^Z)$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_S(h \circ w^*; h^* \circ w^*) + \varepsilon_S(h \circ w; h \circ w^*) + \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(\mathcal{D}_S^Z, \mathcal{D}_T^Z) + |\varepsilon_T(h \circ w; h \circ w^*) - \varepsilon_S(h \circ w; h \circ w^*)|$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_S(h^* \circ w^*; f_S) + \varepsilon_S(h \circ w; f_S) + 2\varepsilon_S(h \circ w^*, f_S) + \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(\mathcal{D}_S^Z, \mathcal{D}_T^Z)$$
$$+ \sup_{w, w^* \in \mathcal{W}} |\varepsilon_T(h \circ w; h \circ w^*) - \varepsilon_S(h \circ w; h \circ w^*)|$$
$$\leq \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_S(h^* \circ w^*; f_S) + \varepsilon_S(h \circ w; f_S) + 2\varepsilon_S(h \circ w^*, f_S) + \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(\mathcal{D}_S^Z, \mathcal{D}_T^Z) + \sup_{h \in \mathcal{H}} \left[\frac{1}{2} d_{h, \mathcal{W} \Delta \mathcal{W}}(\mathcal{D}_S^X, \mathcal{D}_T^X) \right]$$
$$\leq \varepsilon_S(h \circ w; f_S) + 2\varepsilon_S(h \circ w^*, f_S) + \frac{1}{2} d_{\mathcal{H} \Delta \mathcal{H}}(Z_T, Z_S) + 4\sqrt{\frac{2d_H \log(2N) + \log(2/\delta)}{N}} + \frac{1}{2} \sup_{\hat{h} \in \mathcal{H}} \left[d_{\hat{h}, \mathcal{W} \Delta \mathcal{W}}(X_S, X_T) \right]$$
$$+ 4\sqrt{\frac{2d_W \log(2N) + \log(2/\delta)}{N}} + \varepsilon_T(h^* \circ w^*; f_T) + \varepsilon_S(h^* \circ w^*; f_S).$$

Appendix B

Figures 15 and 16 show normalized confusion matrices of our HierMUD approach for the damage localization and quantification tasks, respectively.
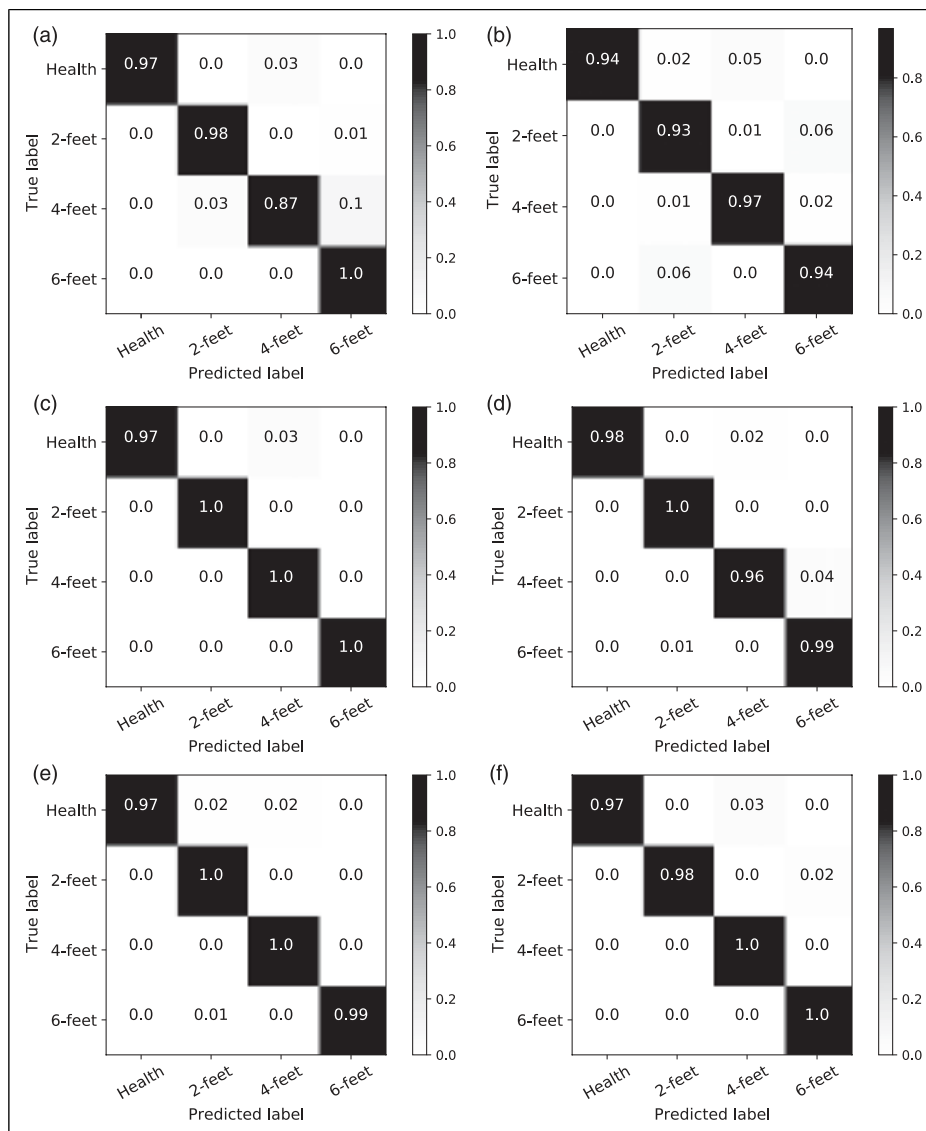
Figure 15. Normalized confusion matrices of our HierMUD approach for the damage localization task. (a) V1, B1→B2, (b) V1, B2→B1, (c) V2, B1→B2, (d) V2, B2→B1, (e) V3, B1→B2, (f) V3, B2→B1.
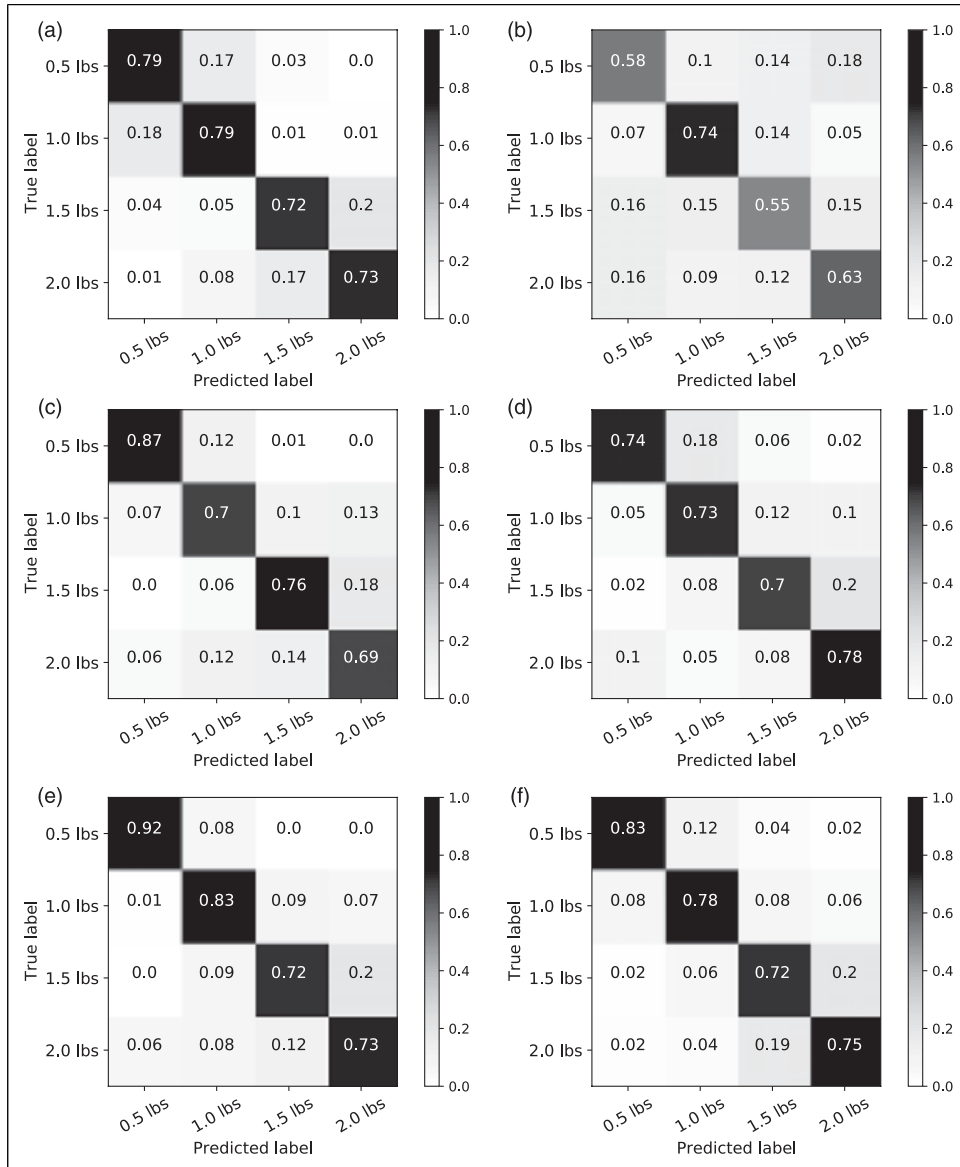
Figure 16. Normalized confusion matrices of our HierMUD approach for the damage quantification task. (a) V1, B1→B2, (b) V1, B2→B1, (c) V2, B1→B2, (d) V2, B2→B1, (e) V3, B1→B2, (f) V3, B2→B1.

To me, diversity is inclusion; everyone, regardless of ethnicity, religion, or beliefs, has equal access to opportunities and resources in every aspect, especially in education. I remember when I first met with my undergraduate final year project advisor at Monash University, Australia during my first experience of living abroad. He warmly welcomed me into the lab and introduced the life and work resources on campus, which mitigated my feeling of nervousness and overwhelmingness that many international students may also have when starting their abroad life. Since then, I have had the great opportunity to study and work in multiple countries and continents with people having diverse backgrounds. These people form my extended family, making me feel a sense of belonging to my community and inspiring me to help create diverse and inclusive environments where everyone is welcomed. In this statement, I describe my commitment to diversity, equity, and inclusion in my past experiences. Then, I outline my plan to encourage diverse and inclusive environments as an academic advisor and instructor in my future career.

I had a well-rounded experience working with amazing people from many backgrounds throughout my time as a graduate student. In my Ph.D. program, I had the great fortune to study at two universities and join a diverse research group. My advisor is a female engineer and professor who also serves as the Diversity, Equity, and Inclusion (DEI) chair of the CEE department. During my Ph.D. study, I am the only Chinese male in the 10+ people group. Being part of this diverse group allowed me to gain valuable advice and insight from my labmates into my research and graduate study. I also have enjoyed the friendship with my labmates as I introduced them to Chinese culture and food. At the same time, I learned about Iranian culture, Caribbean food, Korean drama, "American football", etc. In my future career, I look forward to having many relationships like those during my Ph.D. study because they contribute significantly to what makes the academic environment diverse and inclusive.

My research has contributed to improving the reliability, safety, and integrity of civil infrastructure using vibration-based structural health monitoring. In collaboration with transportation authorities in Pittsburgh and San Francisco Bay Area, I have developed bridge and railroad monitoring systems that can efficiently detect structural damage at low cost with low maintenance, making them widely applicable in both developed and developing regions. In addition, I have used my research skills to help improve the quality of life of people from under-represented groups. For example, in collaboration with Nationwide Children's Hospital, I helped develop in-home gait health monitoring systems for children in Amish communities diagnosed with Muscular Dystrophy.

I have helped support and promote diversity and inclusion education through my teaching and mentoring experiences. Specifically, as a Teaching Assistant at CMU and Stanford University, I encouraged students to share their feedback during office hours in order to adapt my teaching style to students with diverse backgrounds and learning styles. In addition, I have mentored several students during my Ph.D. studies in different fields at all levels (including high school, undergraduate, master's, and Ph.D.). Three of the six students I have mentored are female, and three are international students. As a mentor, I view each mentee as an equal individual and adapt my mentoring style to each student with unique learning preferences. For example, when one of my students who is in high school communicated with me about his stress because he needed to do our research while preparing for upcoming Advanced Placement exams, to respect his time, I suggested rescheduling our meeting and revisiting our project after completing his exams. Also, because he does not have research experience, I started the learning process with step-by-step instructions and hands-on tasks. Then, I improved his understanding of the problem through discussions.

I have also joined and organized outreach events for K-12 students to improve the participation

of the younger generation in STEM. For example, I have joined an outreach event to share the state-of-the-art research about Information Theory at Ravenswood Middle School, Palo Alto, CA. During this outreach event, our team introduced to students the concept of Pictograms and how Pictogram symbols transmit information through visual forms. I have also helped organize a Federal Railroad Administration (FRA) outreach event that introduces railroad engineering to K-12 students aiming to improve future diversity in the rail transportation industry. Moreover, my research has been filmed and spotlighted in the CBS Mission Unstoppable television series, which is the first STEM-centric television series created and funded by, for and about women aims to inspire young girls to open their eyes to STEM careers.

I plan to continue promoting and encouraging diverse, equal, and inclusive environments as an academic instructor and mentor in my future career. The followings are a few examples of actions I plan to take:

1. As an academic advisor and researcher,

   – I will seek opportunities to actively recruit and retain minority students or students from underrepresented, and under-served groups to encourage them in academic careers and the STEM fields. These opportunities include graduate programs, summer programs, outreach, etc. For example, I plan to give seminar talks at high schools and historically black colleges and universities (HBCUs) to share my research with students, and build connections with faculties at HBCUs so that they can send students to join the summer programs and outreach events.

   – I plan to continue improving diversity and inclusion by providing opportunities for underrepresented students to inspire them in the STEM community. In particular, I am interested in working with the university and the local community to develop hands-on summer courses and outreach events for K-12 students. I will also find opportunities to collaborate with schools, hospitals, and nonprofit organizations to develop new ideas using my research skills and domain knowledge to improve diversity, equality, and inclusion.

2. As an instructor, I will incorporate diversity, equality, and inclusion into my teaching materials and methods. Some examples include:

   – I will actively re-orient my course materials and adapt my teaching style based on evaluation and anonymous feedback from students. To implement this, I will hold in-class discussions to build connections between students and understand their backgrounds, preferred learning styles, and interests. I will also conduct course surveys to provide opportunities for students to reflect and provide feedback on their own learning and my teaching.

   – I will design multiple types of course materials (including comprehensive lecture notes, hands-on projects, and demonstrations) that are approachable to students with diverse backgrounds. My course will be developed using experiential learning with progressive assessment so that students may better understand how things work and gradually learn knowledge for solving real-world problems.

In summary, I am committed to creating a diverse and inclusive academic environment that treats each student as an equal individual. I am excited for opportunities to incorporate my ideas in my teaching, advising, mentoring, and service.