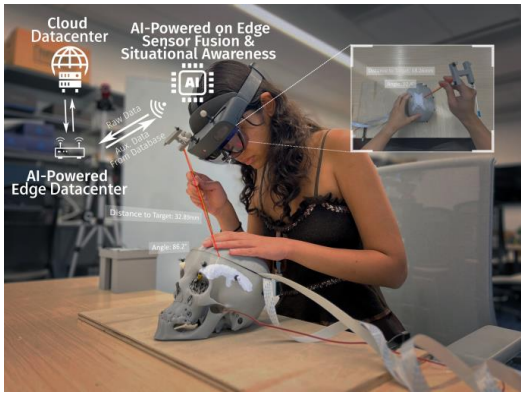
Edge and IoT-supported Intelligent Augmented Reality: Promise, Challenges, and Solutions



Maria Gorlatova

September 28, 2023

About the Speaker

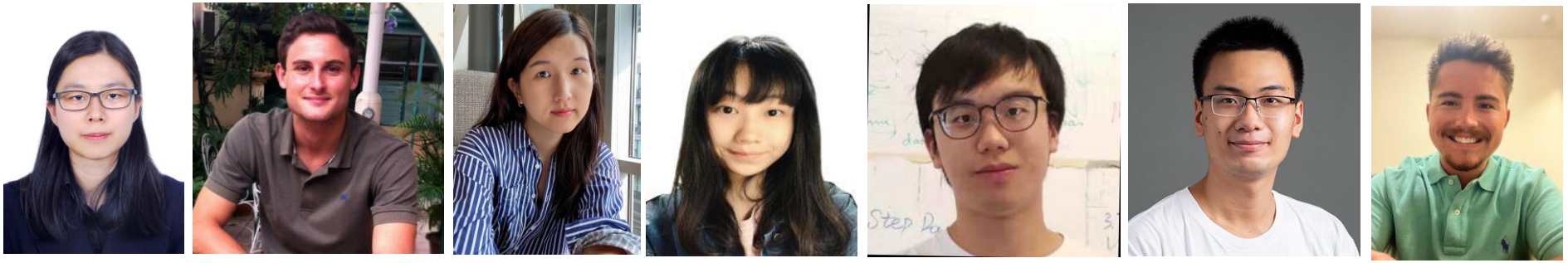
- Nortel Networks Assistant Professor, ECE/CS, Duke University
- Previously:
  - Associate Research Scholar, Princeton University, Electrical Engineering
  - Ph.D. Columbia University, Electrical Engineering
  - Industry positions:

Telcordia Technologies

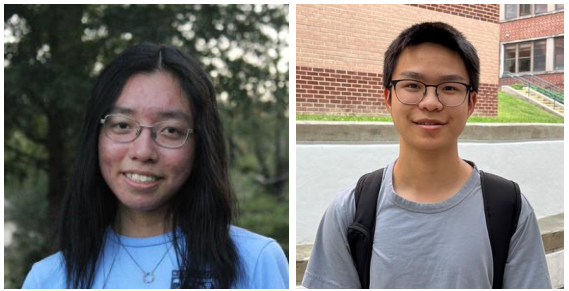
D E Shaw Research

IBM

Duke University Intelligent Interactive Internet of Things (I^3T) Lab



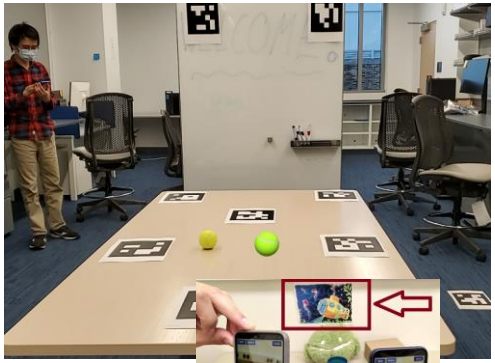
- 7 PhD students (2 graduating)
- 5-12 undergraduate students each semester
- Visiting PhD, undergraduate, high school students

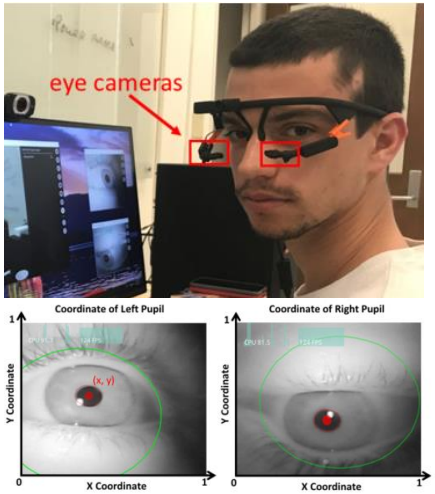


- Join us! Openings for **5 CS & ECE PhD students** and **2 postdocs**

Duke University Intelligent Interactive Internet of Things (I^3T) Lab

- Core research direction: **reliable context-aware** AR





Our Vision: Multi-Device Support for AR

ACM/IEEE IPSN'20
Best Research Artifact Award

ACM UbiComp'22
Best Poster Award

ACM SenSys'20, ACM/IEEE IPSN'22

NSF CAREER

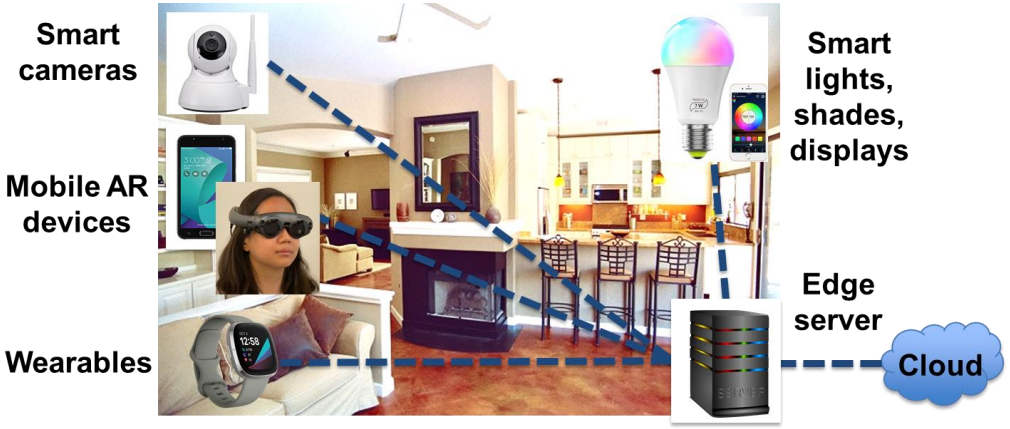
ACM CoNext'20, ACM IMWUT'22, IEEE IoTJ'22, IEEE ISMAR'22, IEEE INFOCOM'23, IEEE ISMAR'23

OpenFog

ATHENA

Outline

- **Augmented reality (AR): current state, applications, limitations**
- Robust environmental awareness in AR
- Enhanced user context awareness in AR
- Summary and future work



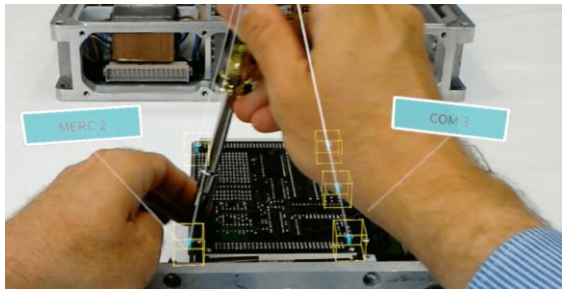
Augmented Reality (AR): Definition & Device Options

- 3D virtual objects integrated into a 3D real environment in real time

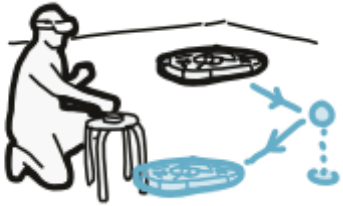
- Google ARCore (2018), Apple ARKit (2017)
  - Vast majority of modern phone models support it
- Microsoft HoloLens (2016), Magic Leap One (2018)
- Constantly expanding in devices and capabilities

AR Applications: ***Showing the Invisible***

- Bringing user the information she needs where she needs it
  - ➢ Retail, navigation, education, human-robot collaboration, …



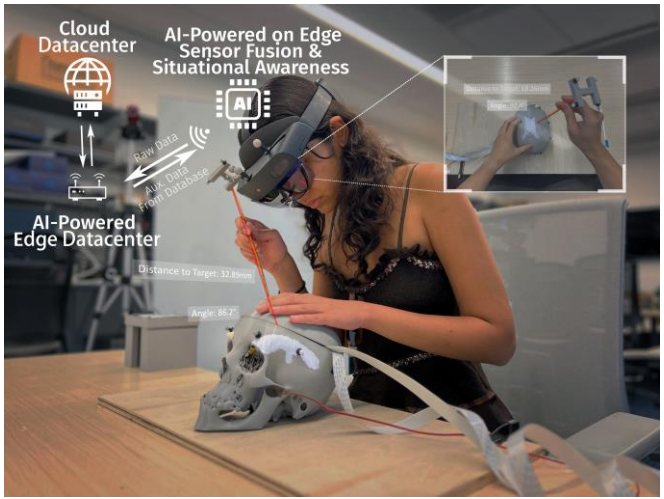




- A wide range of medical applications
  - ➢ Our own work: collaborations with clinical experts

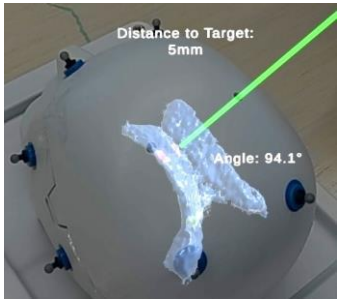
Duke Neurosurgery Duke Eye Center Duke Psychiatry & Behavioral Sciences
Duke University School of Medicine

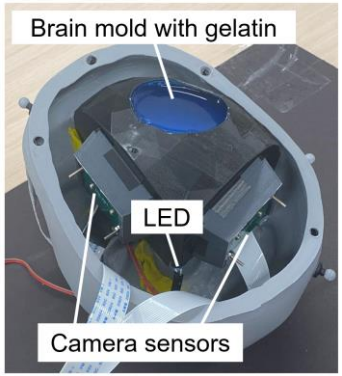
AR-Assisted Neurosurgery and Neurosurgical Training



Submitted: IEEE TVCG 2024, Neurosurgical Focus 2024, IEEE VR 2024

- **Consistent targeting accuracy improvements** under AR guidance
  - **16 – 48%**
- **Edge AI** for real-time feedback
- **Sensing-integrated** phantom model to replace cumbersome 3D scanning
  - **Sub-mm** real-time tool tracking
- >100 medical students and residents
  - External ventricular drain placement (EVD)
  - Twist-drill craniostomy





❑ S. Eom, D. Sykes, S. Rahimpour, **M. Gorlatova**, NeuroLens: Augmented Reality-based Contextual Guidance through Surgical Tool Tracking in Neurosurgery, in *Proc. IEEE ISMAR*, Oct. 2022 (21% acceptance rate).

Last Week on Capitol Hill

AR: Core Mobile Technology of the Future

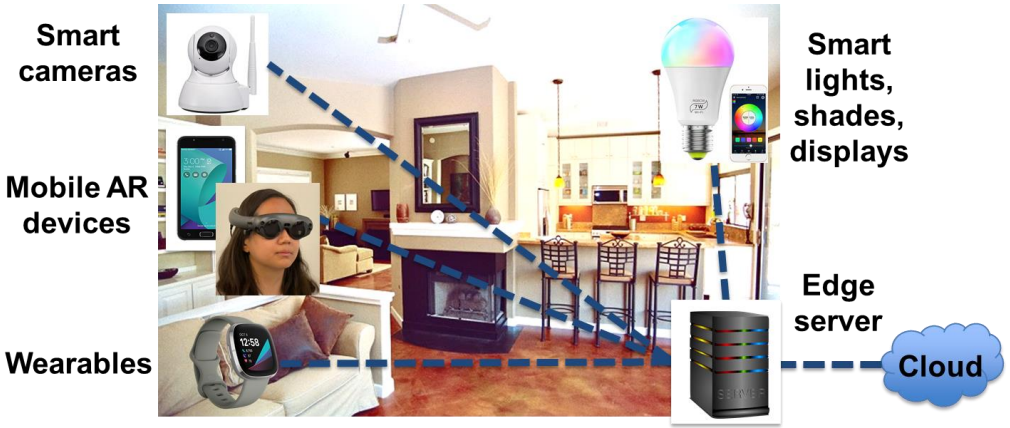
"AR will redefine our relationship with technology"

"It is the next big thing, and will pervade our entire lives"

- How far are we from this vision?
  - Among key limitations: resource consumption & headset form factor, security & privacy, **robust context awareness** in unrestricted conditions

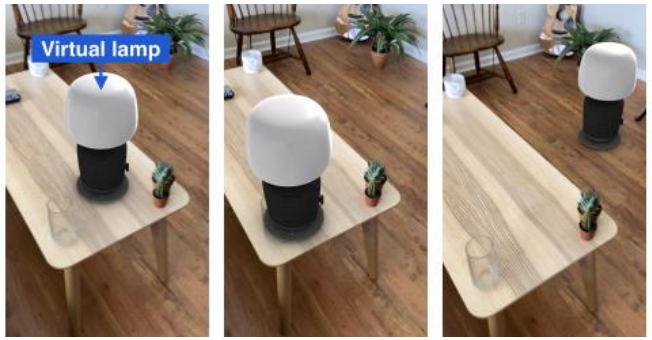
Outline

- AR: current state, applications, limitations
- **Robust environmental awareness in AR**
- Enhanced user context awareness in AR
- Summary and future work



Spatial Awareness: Critical Foundational Capability for Mobile AR

- AR requires accurate knowledge of device **position** and **orientation** (*pose*)
  - ➢ Uninterrupted 6DoF localization on **sub-cm level**
- Typically achieved with Visual-Inertial (VI) SLAM
  - ➢ Complex, resource-intensive
- Poor SLAM performance leads to gross AR artifacts



Unintended motion (*drift*)

Incorrect scale

❑ T. Scargill, G. Premsankar, J. Chen, **M. Gorlatova**, Here To Stay: A Quantitative Comparison of Virtual Object Stability in Markerless Mobile AR, in *Proc. IEEE/ACM CPHS Workshop*, 2022.

Semantic Scene Understanding for Context-Aware AR

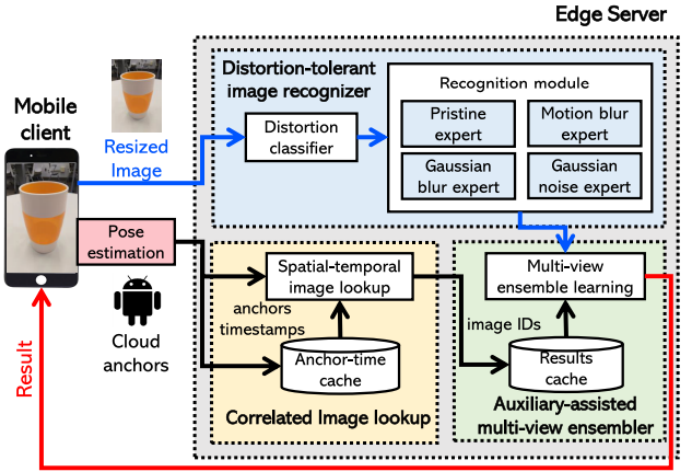
- Types and location of objects and surfaces in the space around the user

"*Even though object recognition rates can exceed 90% on many datasets, in the real world the results from these algorithms are* ***almost unusable***" – Huynh et al, *IEEE VR'19*

- Multiple types of ***domain mismatches***
  - Image quality, camera pose
- ***Open-set*** conditions
- Resource limitations, latency constraints
  - Edge computing: de-facto standard solution

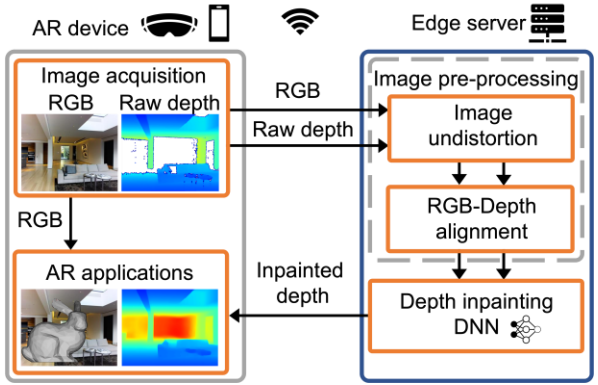
❑ Z. Liu, G. Lan, J. Stojkovic, Y. Zhang, C. Joe-Wong, **M. Gorlatova**, CollabAR: Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, in Proc. *IEEE/ACM IPSN*, Apr. 2020. **IEEE/ACM IPSN Best Research Artifact Award.**

Edge Computing Support: Key to Context-Aware AR



ACM/IEEE IPSN'20

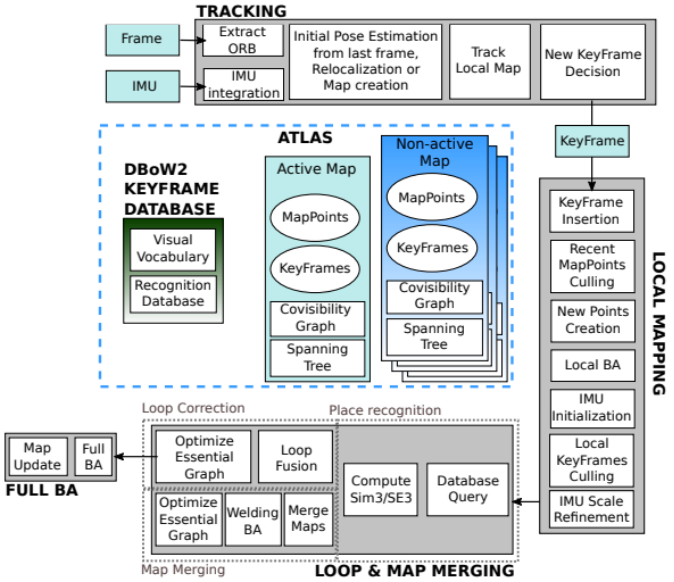
Best Research Artifact Award



ACM IMWUT'22

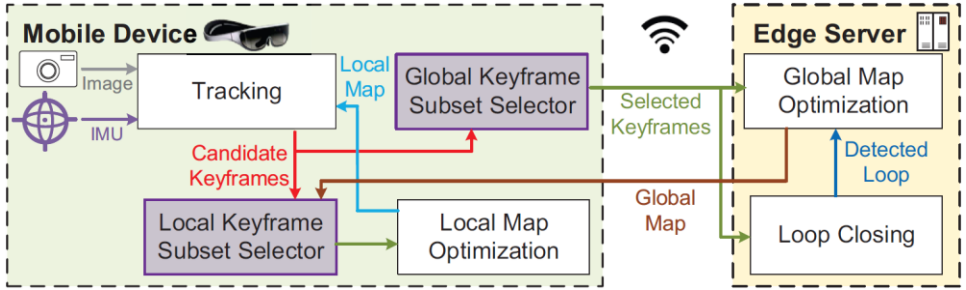
- Key to enabling advanced capabilities
- Latency-accuracy co-optimization across communications & computing

Edge Support for AR: Edge-Supported VI-SLAM



ORB-SLAM3

- Edge-supported architecture with local and global map optimization

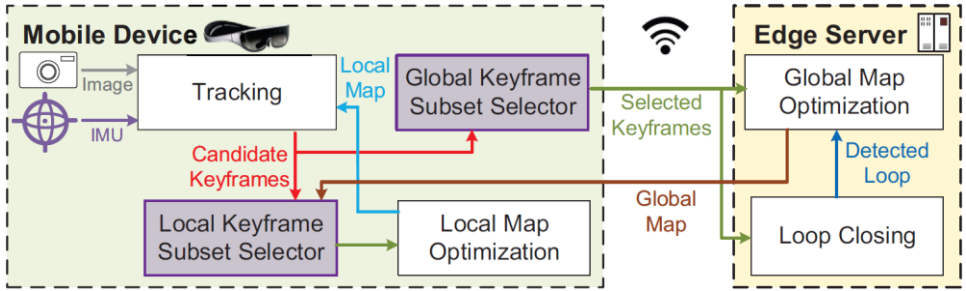


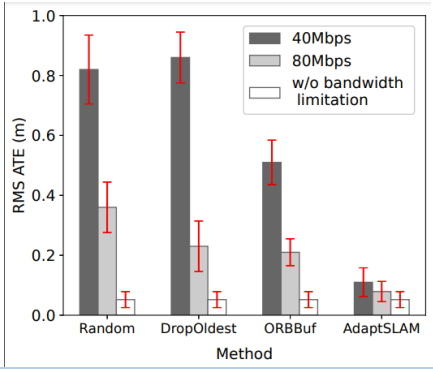
- Use a **theoretically grounded method** for the optimal adaptive offloading **under resource constraints**

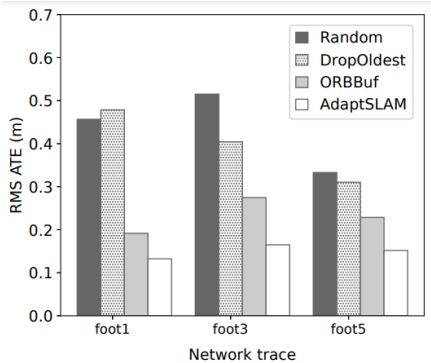
❑ Y. Chen, H. Inaltekin, **M. Gorlatova**, AdaptSLAM: Edge-Assisted Adaptive SLAM with Resource Constraints via Uncertainty Minimization, in *Proc. IEEE INFOCOM,* Hoboken, NJ, May 2023 (19% acceptance rate). **Code available via GitHub.**

Resource-efficient Edge-Supported VI-SLAM via Online Quality of Service (QoS) Metrics

- **Map uncertainty quantification** to select best keyframes to use in local and global maps under computation and communication constraints
- **Low-complexity algorithms** to enable it to run in real-time
  - ➢ Local keyframe selection & map optimization latency: original ORB-SLAM3 556ms, simple heuristics 133-149ms, AdaptSLAM 163ms

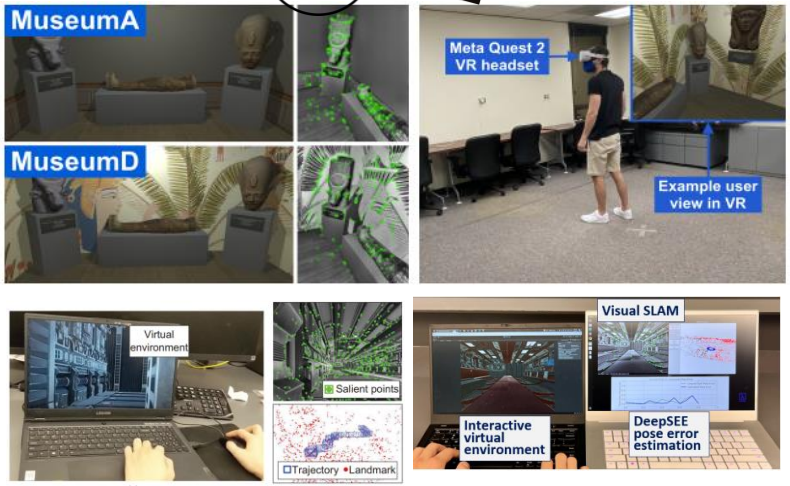






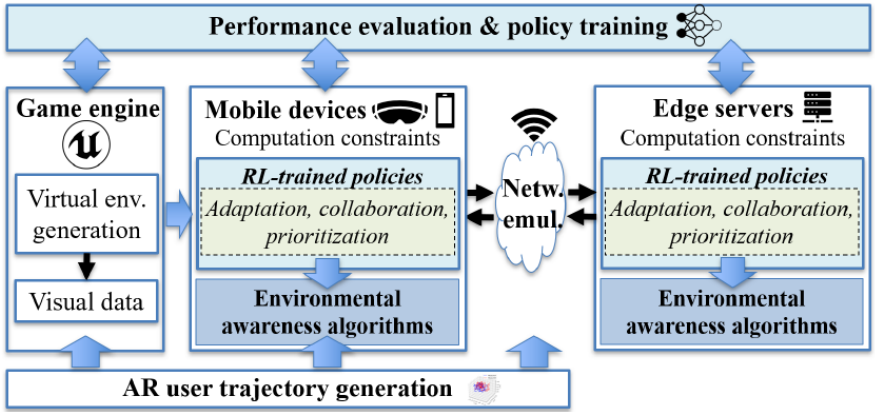
❑ Y. Chen, H. Inaltekin, **M. Gorlatova**, AdaptSLAM: Edge-Assisted Adaptive SLAM with Resource Constraints via Uncertainty Minimization, in *Proc. IEEE INFOCOM, 2023,* Hoboken, NJ, May 2023 (19% acceptance rate). **Code available via GitHub.**

AR Environmental Awareness Evaluation & Design: Role of Game Engines



IEEE ISMAR'22, submitted IEEE TMC'24, submitted IEEE ICRA'24

Multi-user AR-EMU



Carnegie Mellon University

Fall 2023 – Fall 2027

NSF

CISCO

MICHIGAN

❑ T. Scargill, Y. Chen, N. Marzen, **M. Gorlatova**, Integrated Design of Augmented Reality Spaces Using Virtual Environments, in *Proc. IEEE ISMAR'22,* Oct. 2022 (21% acceptance rate). **Code available via GitHub.**

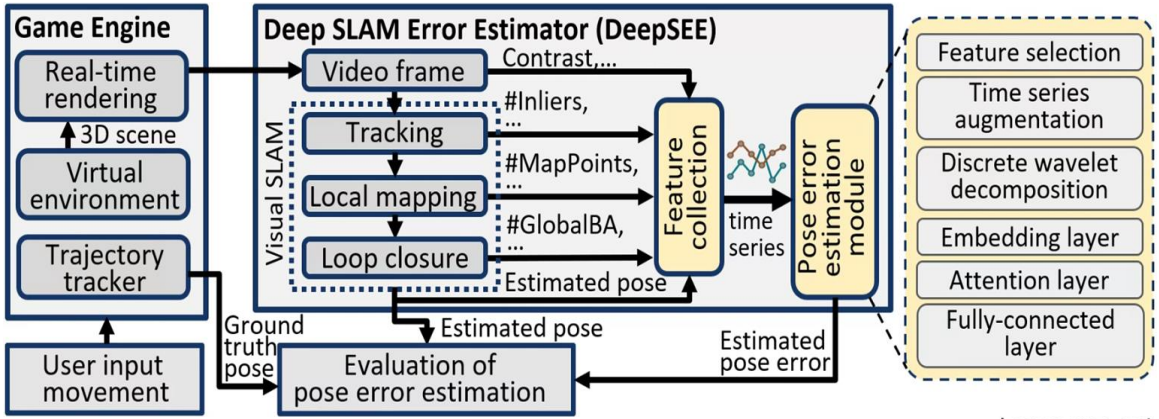
Online Quality of Service (QoS) Metrics for AR

- How well is environmental awareness operating?
- Important to know for **system optimizations** and for **user guidance**
- Ground truth not available!

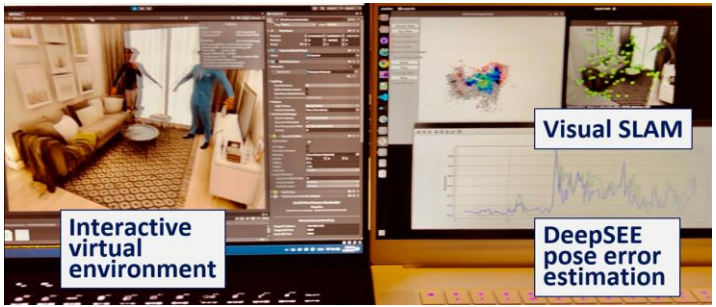
- Dramatically under-studied
  - Traditional multimedia quality of service (QoS) and quality of experience (QoE) metrics not a fit
  - Need to be quantitative & operate at low latency on resource-constrained devices

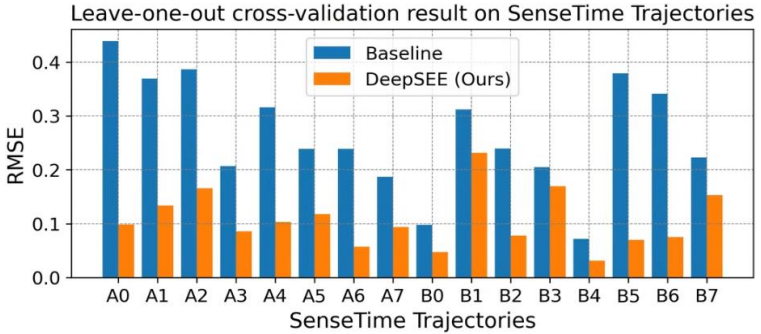
- Join us! Openings for **5 CS & ECE PhD students** and **2 postdocs**

Ongoing Work: DNN-based SLAM Tracking Error Online Estimation



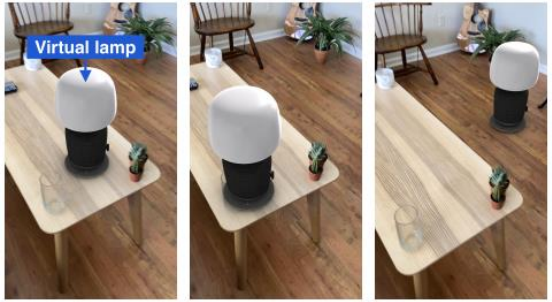
- Fast and accurate SLAM error magnitude estimation





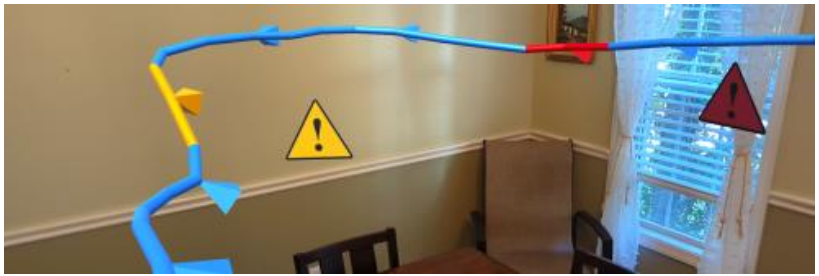
❑ T. Hu, T. Scargill, Y. Chen, G. Lan, **M. Gorlatova**, Demo: DNN-based SLAM Tracking Error Online Estimation, to appear in P*roc. ACM MobiCom'23*, Madrid, Spain, Oct. 2023.

Situated Trajectory Analysis for In-the-Wild Pose Error Estimation



- If I place virtual content *here*, is it likely to appear stable to the users?

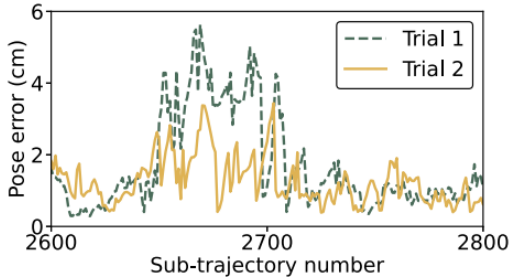
- The first situated trajectory analysis system for AR that **incorporates pose error estimates**

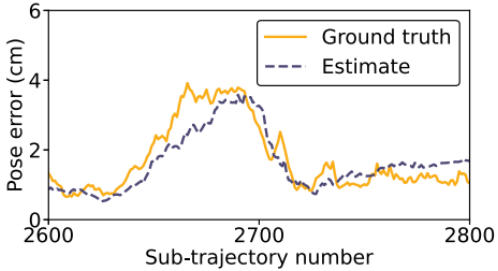


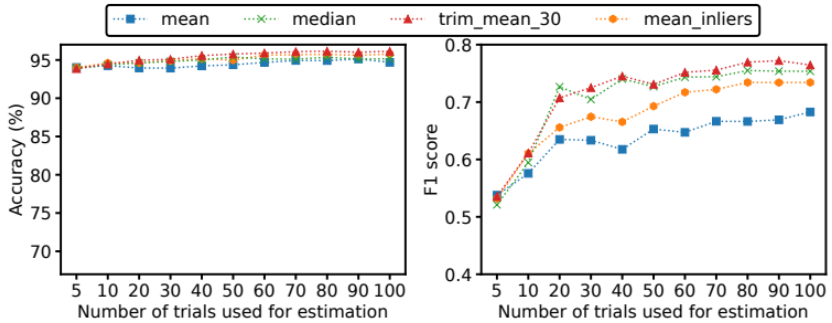
❑ T. Scargill, Y. Chen, T. Hu, **M. Gorlatova**, SiTAR: Situated Trajectory Analysis for In-the-Wild Pose Error Estimation, to appear in *Proc. IEEE ISMAR'23*, Sydney, Australia, Oct. 2023. **Code available via GitHub.**

AR QoS Metric:
Pose Error Estimation via Uncertainty Propagation

- Cannot get the ground truth … but *can* run pose estimator repeatedly
  - Observation: [all] accurate trajectory estimates resemble one another, each inaccurate trajectory estimate is inaccurate in its own way
  - **Uncertainty propagation** methodology



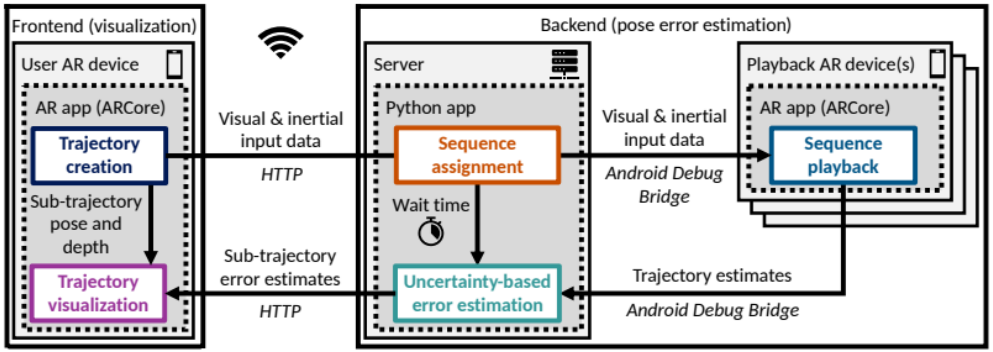


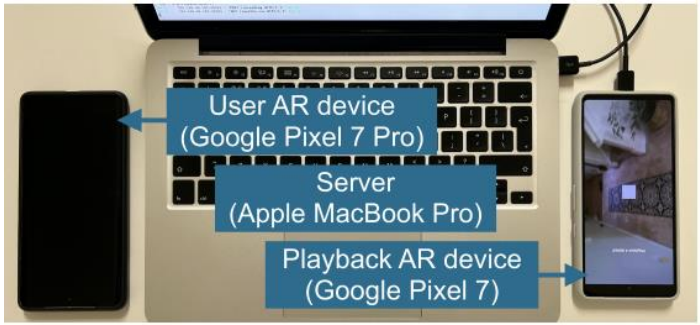


- Applies to open-source and **black-box SLAM engines**

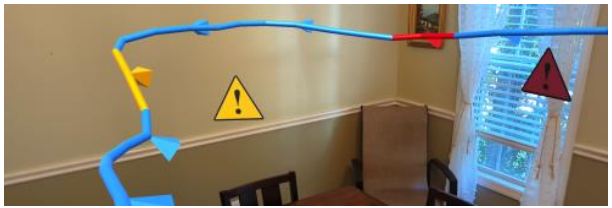
❑ T. Scargill, Y. Chen, T. Hu, **M. Gorlatova**, SiTAR: Situated Trajectory Analysis for In-the-Wild Pose Error Estimation, to appear in *Proc. IEEE ISMAR'23*, Sydney, Australia, Oct. 2023. **Code available via GitHub.**

Pose Error Estimation via Uncertainty Propagation: An Implementation for ARCore





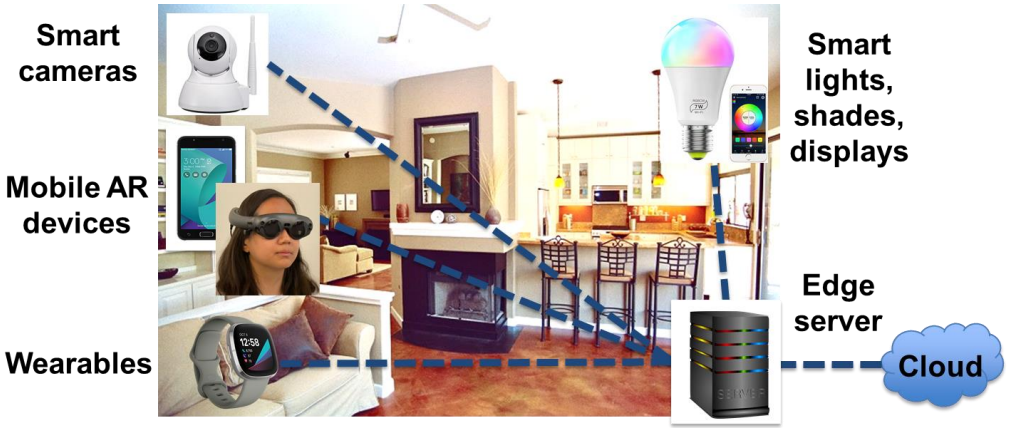
- End-to-end latency: for 15s trajectory, 5 trials, 1 playback device: **89s** on average
- Evaluated with 15 participants in 13 environments



❑ T. Scargill, Y. Chen, T. Hu, **M. Gorlatova**, SiTAR: Situated Trajectory Analysis for In-the-Wild Pose Error Estimation, to appear in *Proc. IEEE ISMAR'23*, Sydney, Australia, Oct. 2023. **Code available via GitHub.**

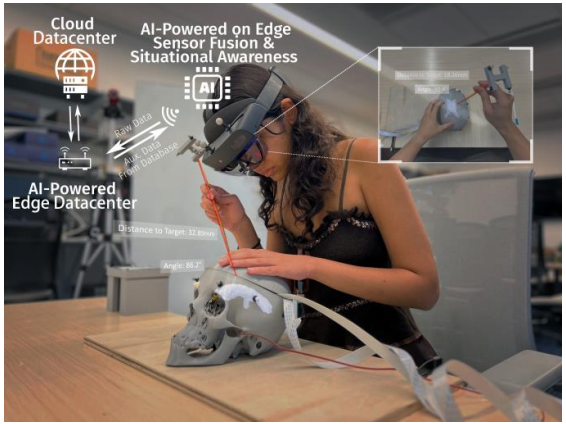
Outline

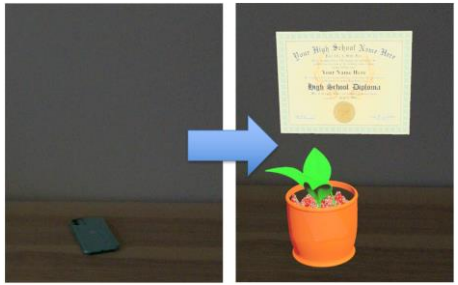
- AR: current state and applications
- Robust environmental awareness in AR
- **Enhanced user context awareness in AR**
- Summary and future work



AR User Context Awareness

- Understanding what the user is doing and how the user is feeling
- Automatically assessing user's level of expertise, concentration, mental and physical fatigue
- AR: immersive experiences tailored to the state of the user





Duke | DEPARTMENT OF Biomedical Engineering

Duke Psychiatry & Behavioral Sciences
Duke University School of Medicine

IEEE Metabuild'22, ACM DigiBiom'22

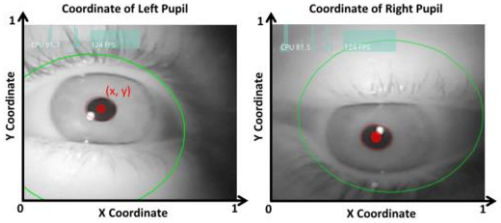
User Context Awareness: Opportunities and Challenges

- Intimately connected with the user
  - AR **sees what you see and hears what you hear**
  - Eye tracking: looks deep inside your soul
- Wearables add new opportunities

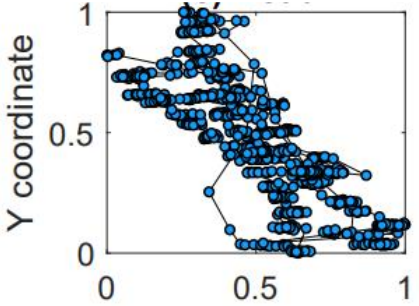
- Challenges: training **personalized** algorithms
  - Population-level models do not work well on individual users
  - Can’t collect too much data → Few-shot learning
  - Personal data can be highly sensitive

Eye Tracking-based Cognitive Context Sensing

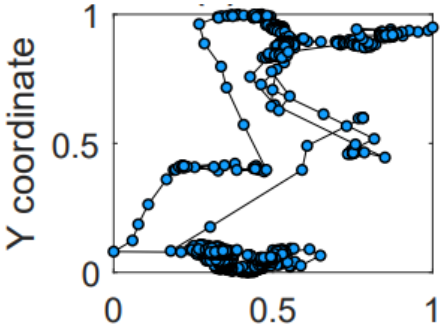
- Important part of AR and VR hardware
- Powerful source of information about the user
  - Both ***where*** and ***how*** the user looks



- Challenges:
  - Need to work in few-shot conditions
  - Highly heterogenous across subjects, stimuli, eye tracking devices

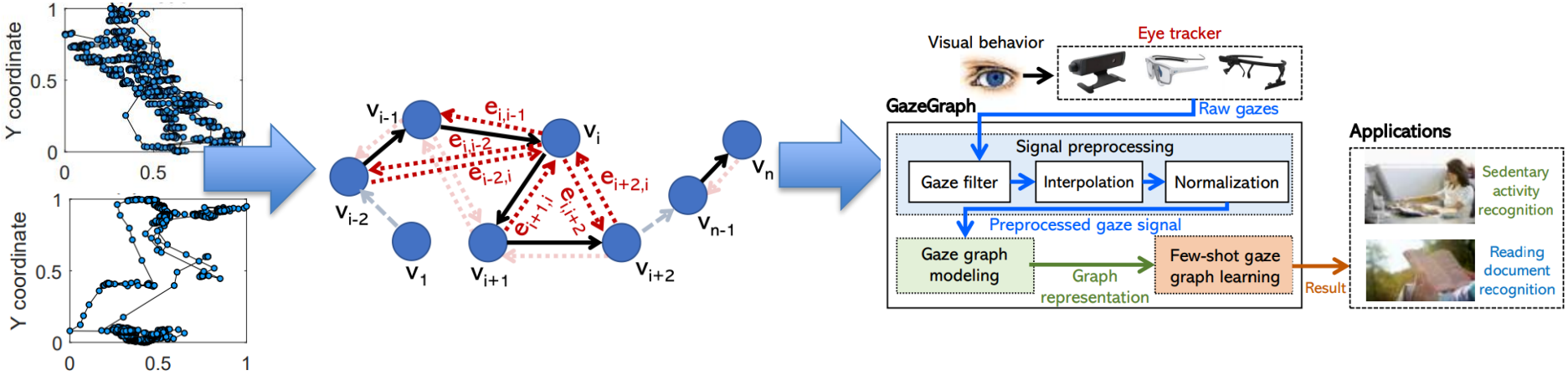


Reading text

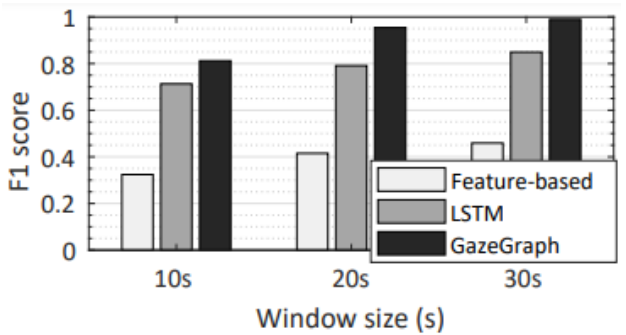


Watching a movie with subtitles

Cognitive Context Detection with GazeGraph



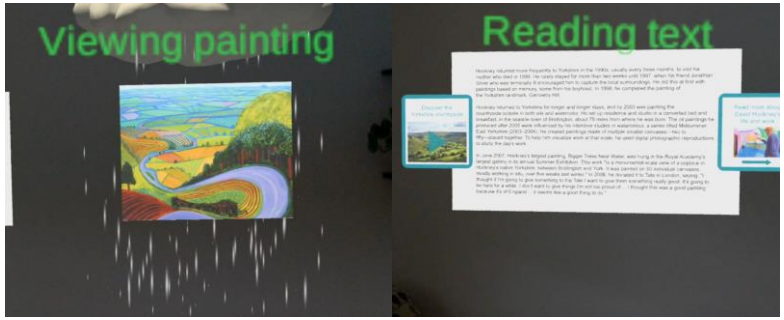
- First gaze-based cognitive context sensing method that *models human visual behavior as graphs*
- Outperforms feature-based methods by 37-54%, LSTM-based methods by 10-23%

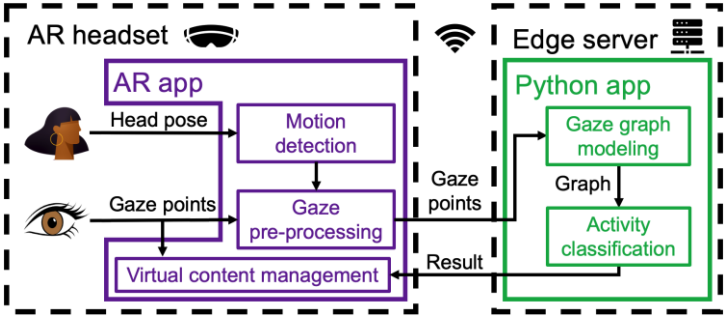


❑ G. Lan, B. Heit, T. Scargill, **M. Gorlatova**, GazeGraph: Graph-based Few-Shot Cognitive Context Sensing from Human Visual Behavior, in *Proc. ACM SenSys*, Nov. 2020. **Code & data available via GitHub.**

GazeGraph in an AR Application

- Demonstrated *first gaze-based activity recognition in an AR app*



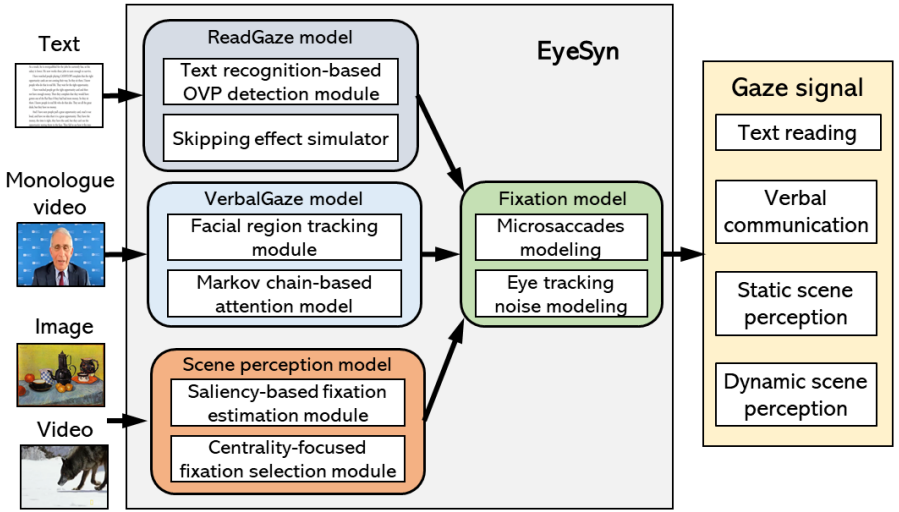


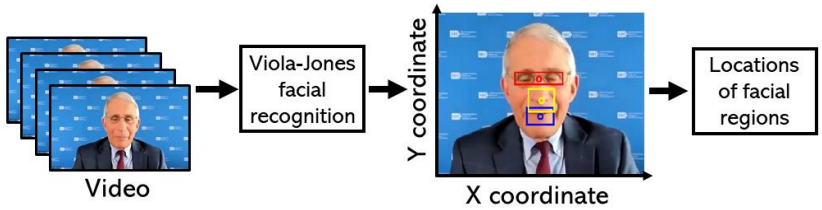
- Pre-processing, graph construction, and DNN inference time: under 15ms
- Overall round-trip latency: under 100ms

- ❑ T. Scargill, G. Lan, **M. Gorlatova**, Demo: Catch My Eye: Gaze-Based Activity Recognition in an Augmented Reality Art Gallery, in *Proc. ACM/IEEE IPSN*, May 2022.
- ❑ G. Lan, B. Heit, T. Scargill, **M. Gorlatova**, GazeGraph: Graph-based Few-Shot Cognitive Context Sensing from Human Visual Behavior, in *Proc. ACM SenSys*, Nov. 2020. **Code & data available via GitHub.**

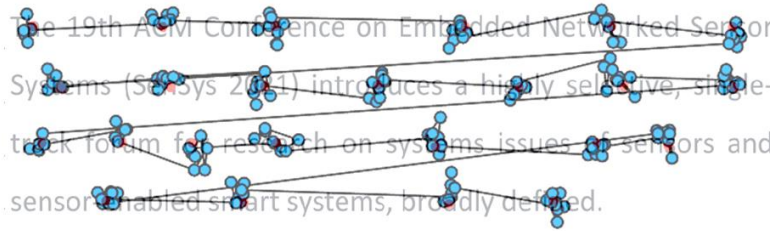
Eye Movement Synthesis for Context Detection

- Psychology-inspired generative models for *synthesizing* arbitrarily large realistic eye movement datasets





Human visual attention in a conversation



Human visual attention when reading

❑ G. Lan, T. Scargill, **M. Gorlatova**, EyeSyn: Psychology-inspired Eye Movement Synthesis for Gaze-based Activity Recognition, in *Proc. IEEE/ACM IPSN*, May 2022. **Code & data available via GitHub**. **Selected media coverage: vice.com Motherboard, hackster.io. Highlighted in the NSF-wide Discoveries newsletter.**

Summary and Future Work

- Towards reliable generation of contextual AR experiences: necessary improvements in **spatial and semantic awareness**
  - ➢ New tools, edge computing, online QoS metrics
- Exciting opportunities for **user context-aware AR** applications

- Join us! Openings for **5 CS & ECE PhD students** and **2 postdocs**
- maria.gorlatova@duke.edu

Acknowledgements