Research Statement

The convergence of wearable sensing, edge computing, and foundation models has enabled a new class of intelligent systems that perceive, reason about, and act alongside humans in the physical world. State-of-the-art sensing systems, such as ambient sensors in living spaces, cameras, and wearable devices enabling physiological monitoring, a plethora of robot-mounted sensors ranging from LiDARs, depth sensing, tactile feedback, inertial sensing, acoustic perception, to localization systems, open a new frontier in human-centered intelligent systems research. My work has advanced from sensing and interpreting human behavior in real-world environments to building end-to-end, human-centered intelligent systems that (a) perceive context from multi-modal sensor data, (b) support decision-making and real-time assistance, (c) coordinate actions in shared human-machine settings, and (d) translate learning algorithms from real-world sensor deployment and data acquisition to representation learning and deployed edge inference, with applications in **Pervasive and Intelligent Health Systems, AI-Augmented Human Skill and Performance Assessment, and Collaborative Autonomous Robotics.** My long-term research goal envisions a future with (i) homes that can recognize subtle changes in health and cognition before they are clinically detected, (ii) humans receive real-time personalized feedback to improve skill and performance, (iii) robot-teams coordination across different capabilities, and (iv) humans and machines exchange information seamlessly without significant overhead. Thus, these advances lay the foundation for adaptive, proactive, and collaborative systems that deeply integrate into everyday life.

Past and Current Research

My research develops end-to-end, human-centered intelligent systems that connect real-world sensor deployment [1, 2], multi-modal representation learning [3], edge inference [4], and decision support [5] in physical environments. A central theme of my work is learning from noisy, heterogeneous, and label-scarce sensor data [1, 3, 6] collected outside controlled labs, including smart-home apartments, wearable and contactless health-monitoring platforms, instrumented sports environments, and heterogeneous edge platforms. To investigate systems under these circumstances requires deployment of sensing infrastructure, collection of longitudinal data, learning robust representations, compressing models for embedded platforms, and testing performance in real-world environments [1, 4]. The following research themes highlight my current contributions and serve as a stepping stone for building on this work to investigate my future research vision.

Pervasive and Intelligent Health Systems

My work in this thrusts began with sensing and detecting human activities in real-world environments using wearables, contactless sensors, cameras, and infrastructure sensors. This work focused on capturing Activities of Daily Living (ADLs), Instrumental Activities of Daily Living (IADLs), photoplethysmogram, facial video, and other health-related signals while addressing noise such as motion artifacts and power line interference. A key part of this work was collecting real-world data rather than relying only on controlled laboratory datasets. For activity recognition, this included longitudinal sensor data from older adults performing ADLs and IADLs in natural living environments; for rPPG, this included facial video data collected under realistic conditions for contactless physiological monitoring [1, 3, 2].

This led to the challenge of building reliable learning models when real-world health sensing data are noisy, heterogeneous, and label-scarce. In my Ph.D. dissertation, I addressed the lack of labeled data for older adults' activity recognition, especially for those living with Dementia. Since their activities are difficult to detect due to restricted motion, dementia-like symptoms, activities performed at inappropriate times, sensor limitations, class imbalance, high inter- and intra-class variations, and the lack of long-term annotators, I asked whether abundant unlabeled data can aid the classification task of detecting activities in a supervised setting. To address this, I proposed a novel pre-training phase that learns a new representation space using unlabeled data and projects labeled data onto the newly learned representation space before classification, resulting in a drastic improvement in classification performance [1].

The ability to detect activities then supported the next question: can these sensed activities provide insight into cognitive or functional health impairment? To answer this, I combined wearable and clinical assessment data collected concurrently and labeled accelerometer data with both activity labels and functional/cognitive impairment assessment labels. I then used contrastive learning loss and a biased multi-task learning methodology to study whether misclassified activity labels can be linked to the underlying cognitive or functional health impairment. This showed that studying the model's errors can help explain why the machine learning model failed and relate those failures to the underlying science.

These sensing and learning models also raised the problem of real-time deployment on resource-constrained edge devices. Since many IoT devices are constrained by memory, power usage, and latency, I developed mechanisms to compress machine learning models for low-resource platforms. Specifically, I proposed a layer-wise pruning rate optimization methodology combined with quantization to adapt compressed models for resource-constrained devices. I also co-developed the RhythmEdge system, which estimates the PPG signal from facial videos on the edge in near real-time, enabling instantaneous and contactless bio-signal estimation.

AI-Augmented Human Performance and Coaching

This thrust develops systems for assessing, coaching, and scoring skilled human performance. My work in this area moves beyond recognizing what action a person is performing toward understanding how well the action is performed and how feedback can improve future performance. Using skeleton sequences, visual motion patterns, instrumented-floor sensing, and game data, this work has addressed skilled performance assessment in domains such as badminton, dance, and sports reporting [7, 8, 9, 10]. Across these studies, I developed learning-based methods for stroke-quality classification, per-limb performance scoring, novice-expert comparison, data-driven stance assessment, movement-quality estimation, and dense natural-language performance narration [9, 11, 5, 12, 8].

A central contribution of this thrust is the development of models that quantify movement quality rather than only classify activity labels. These systems combine motion sensing, representation learning, error modeling, and feedback generation to connect human motion patterns with performance scores and corrective guidance [8]. The work also explored human-AI collaboration by comparing novice and expert movement representations, identifying body segments or movement phases that contribute to performance errors, and generating interpretable feedback that can support coaching and skill improvement [8]. My current research focuses on making human performance assessment more domain-agnostic, real-time, and closed-loop rather than limiting these methods to sport-specific and offline settings. I aim to extend them toward broader human-centered environments where intelligent systems can assess skill, fatigue, error patterns, and task progression while working alongside people. This will allow performance assessment to serve as a foundation for future human-machine teaming systems that can decide when to guide, assist, intervene, or step back.

Collaborative Autonomous Robotics

This thrust extends my work in sensing, edge intelligence, and human-performance assessment into robotic systems that can perceive, reason, coordinate, and act in shared physical environments. The focus is on building collaborative autonomous systems that operate under real-world constraints, including limited compute, heterogeneous sensors, distributed platforms, latency, power usage, and human-in-the-loop decision-making [13, 10].

My work in this area began with resource-aware and heterogeneous edge intelligence for collaborative autonomous systems. In *HeteroEdge* [13], we addressed computational asymmetry across platforms with different compute budgets, such as Jetson Nano and Xavier devices, by profiling device resources, compressing transmitted data using ROI masking, and scheduling tasks through a constraint-aware allocation policy. *HeteroSys* [10] extended this line of work to heterogeneous and collaborative sensing in the wild by integrating sensor fusion, MQTT-based middleware, and clock synchronization across distributed platforms. These contributions provide the foundation for multi-robot systems that must coordinate sensing, inference, and task execution under physical and computational constraints.

Building on this foundation, my current work moves toward embodied agents that collaborate with humans in real time. In *SVLM-FeedbackBench* [14], we developed an end-to-end small vision-language model-based robot advisory agent pipeline for human manipulation tasks. The motivating scenario is a robotic copilot that monitors a human partner during tasks such as assembly, repair, or troubleshooting and provides precise spoken guidance when it detects an error. This work evaluates the quality of narration, mistake detection, grounded instruction generation, latency, memory footprint, throughput, and power consumption. Through this work, we demonstrated that carefully configured small VLMs can support real-time human task guidance on resource-constrained robotic platforms, such as the NVIDIA Jetson Orin AGX.

This thrust also includes another work on multi-agent coordination and a security-aware autonomous system. *MARSys* introduces a multi-agent recommendation system with knowledge graph coordination, adaptive strategy blending, and Bayesian trust modeling [15]. While developed in a recommendation setting, its core ideas in multi-agent coordination, trust propagation, and composable agent design transfer to robotic systems that require trust-calibrated decision fusion. *VIGIL*, on the other hand, is security-oriented, but it contributes to this thrust by studying inertial sensing, LLM-based reasoning, and VLM-in-the-loop control for physical systems [16]. Rather than treating this as only a cybersecurity problem, this work informs how autonomous systems can reason over sensor traces, detect abnormal behavior, and remain robust when deployed in safety-critical environments.

Overall, this thrust integrates perception, reasoning, coordination, and action under real-world constraints. My future direction is to build collaborative autonomous systems that are computationally efficient, physically grounded, trust-aware, and capable of working with humans rather than simply operating around them.

Future Work

Thrust 1: Human-Centered Sensing and Embodied Intelligence

As sensors become cheaper, smaller, and more pervasive, including wearables on wrists, passive infrared arrays in hallways, and cameras that estimate heart rate without contact, it becomes possible to monitor how people live, move, and recover in their own homes rather than only in clinics. My prior work has addressed foundational sensing challenges: learning from label-scarce sensor streams through self-taught and contrastive pre-training, compressing inference pipelines for real-time operation on edge hardware through structured pruning and quantization, and fusing activity classification with health-state prediction through multi-task architectures. These problems, including label efficiency, edge deployment, and single-task health inference, are now mature enough to build on.

The next step is to advance these systems toward a clinical-grade, personalized health assessment. Current approaches typically map one sensing modality to one health outcome, such as accelerometry to fall risk, rPPG to heart rate, or sleep sensors to insomnia severity. However, clinicians do not diagnose this way. Cognitive decline, for instance, manifests through a combination of symptoms: subtle changes in gait cadence, sleep fragmentation, slower ADL transitions, reduced social engagement, and altered physiological rhythms. No single sensor stream captures this full picture. I will develop multimodal fusion architectures that combine ambient, wearable, and camera-derived signals to jointly detect these symptoms. Critically, the symptom signatures detected by these models will be defined in collaboration with clinicians, including geriatricians, neuropsychologists, and rehabilitation specialists, so that what the system flags as abnormal reflects validated clinical criteria rather than statistical artifacts of sensor data.

The end goal is to close the loop from multi-modal monitoring to intervention. Rather than generating passive alerts, the system must reason about what it observes and decide when and how to act: whether to prompt the patient to take a break, adjust rehabilitation exercises based on gait fatigue, notify a caregiver about worsening sleep fragmentation, or escalate to a clinician when behavioral changes match a known prodromal pattern. These are sequential decisions under uncertainty because the system cannot directly observe the patient's true health state; it can only infer it from noisy sensor streams, and each action has consequences over time. I will formulate this as a partially observable decision process, where the belief state fuses multi-modal sensor evidence into a probabilistic estimate of the patient's condition, and the action space is grounded in clinician-defined care protocols rather than model-generated choices. Intervention policies will be learned from logged clinical sessions using offline methods, avoiding the ethical and safety concerns of trial-and-error learning on real patients, with explicit safety constraints ensuring that high-risk actions, such as medication changes, always route through a clinician. The result is a system that moves from passive, single-symptom dashboards to holistic, clinician-backed health intelligence.

Thrust 2: Adaptive Human-Machine Teaming and Skill Augmentation

This thrust advances human performance assessment and coaching as a foundation for adaptive human-machine teaming. My prior work in AI-augmented human performance developed systems that assess skilled movement, compare novice and expert behavior, score performance across body segments, and generate feedback for improvement. While these systems were mainly studied in sports and movement-quality settings, the core idea extends to human-machine teaming: for a machine to work effectively with a human, it must understand *what the human is doing*, *how well the human is doing it*, *when the human needs help*, and *what type of assistance would be most useful.*

A key direction is to develop domain-general representations of human skill that transfer across activities with different biomechanical and task constraints. This requires models that learn from expert and novice demonstrations and capture *coordination, timing, stability, fatigue, uncertainty, and error patterns* using a variety of sensing platforms, including visual pose, wearable inertial signals, physiological indicators, VR headsets, gesture tracking, gaze sensing, and task context. I will investigate self-supervised learning, contrastive learning, and few-shot adaptation to separate transferable motor patterns from task-specific biomechanics.

The long-term goal is to move from offline skill scoring to real-time, closed-loop teaming. Once a system can estimate skill, fatigue, and error patterns, it can decide whether to *provide a verbal cue*, *adjust task difficulty*, *demonstrate the next step*, *take over a subtask*, or *step back when the human is performing well.* I will study these decisions as sequential decision-making problems under uncertainty, balancing assistance, autonomy, trust, workload, and safety. This will enable machines to act as adaptive collaborators that can coach, assist, and team with people in shared physical and immersive environments.

Thrust 3: Collaborative Autonomy and Distributed Robotic Intelligence

This thrust focuses on building collaborative robotic systems that can perceive human activity, infer context, operate on constrained hardware, and coordinate actions in shared physical environments. Future robot teams will include drones, ground vehicles, manipulators, mobile sensors, and edge servers with different sensing, computing, mobility, battery, and communication capabilities. The goal is to develop distributed robotic intelligence that operates under real-world constraints, including limited bandwidth, limited power, uncertain environments, and human-in-the-loop decision-making.

A key direction is heterogeneous robot coordination under changing mission conditions. I will develop resource-aware coordination methods that decide which agent should *sense*, *compute*, *communicate*, or *act* when conditions change, such as bandwidth drops, battery limits, new objects in the scene, or changes in a human partner's task. This requires jointly optimizing *sensing coverage*, *inference allocation*, *communication routing*, and *action assignment* under latency, power, and safety constraints.

A second direction is grounded reasoning and adaptive shared autonomy. I will develop planning architectures where vision-language models break mission goals into robot-specific subtasks that are *grounded* in verified action primitives, *robust* against adversarial attacks, *trust-calibrated* using agent reliability estimates, and *human-adaptive* based on task criticality, robot confidence, environmental uncertainty, and the human partner's fatigue, workload, and skill level. The broader goal is to build collaborative autonomous systems that know when to *assist*, *take over*, *request confirmation*, *defer to a human*, or *escalate to a larger model or expert* using formal safety constraints rather than hardcoded thresholds.

References

[1] Ramamurthy, S. R., I. Ghosh, A. Gangopadhyay, E. Galik, and N. Roy, "STAR-Lite: A light-weight scalable self-taught learning framework for older adults' activity recognition," *Pervasive and Mobile Computing*, vol. 87, p. 101698, 2022.

[2] Z. Hasan, Ramamurthy, S. R., and N. Roy, "CamSense: A camera-based contact-less heart activity monitoring," *Smart Health*, vol. 23, p. 100240, 2022.

[3] Ramamurthy, S. R., S. Chatterjee, E. Galik, A. Gangopadhyay, N. Roy, B. Mitra, and S. Chakraborty, "CogAx: Early assessment of cognitive and functional impairment from accelerometry," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2022.

[4] Z. Hasan, E. Dey, Ramamurthy, S. R., N. Roy, and A. Misra, "RhythmEdge: Enabling contactless heart rate estimation on the edge," in *IEEE International Conference on Smart Computing (SMARTCOMP)*, 2022.

[5] I. Ghosh, Ramamurthy, S. R., A. Chakma, and N. Roy, "DeCoach: Deep learning-based coaching for badminton player assessment," *Pervasive and Mobile Computing*, vol. 83, p. 101608, 2022.

[6] A. Z. M. Faridee, Ramamurthy, S. R., H. M. S. Hossain, and N. Roy, "HappyFeet: Recognizing and assessing dance on the floor," in *ACM International Workshop on Mobile Computing Systems & Applications (HotMobile)*, 2018.

[7] A. Z. Md Faridee, Ramamurthy, S. R., H. M. Sajjad Hossain, and N. Roy, "HappyFeet: Recognizing and assessing dance on the floor," in *HotMobile 2018 - Proceedings of the 19th International Workshop on Mobile Computing Systems and Applications*, vol. 2018-Febru, pp. 49–54, Association for Computing Machinery, Inc, Feb. 2018.

[8] I. Ghosh, A. Chakma, M. S. Anwar, Ramamurthy, S. R., and N. Roy, "SkillNet: Human actions assessment via human-ai collaboration," *ACM Transactions on Multimedia Computing, Communications and Applications*, 2026.

[9] I. Ghosh, Ramamurthy, S. R., and N. Roy, "StanceScorer: A Data Driven Approach to Score Badminton Player," in *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 1–6, IEEE, 2020.

[10] I. Ghosh, A. Goldstein, A. Chakma, J. Freeman, T. Gregory, N. Suri, Ramamurthy, S. R., and N. Roy, "HeteroSys: Heterogeneous and collaborative sensing in the wild," in *IEEE International Conference on Distributed Computing in Smart Systems and the IoT (DCOSS-IoT)*, 2023.

[11] I. Ghosh, Ramamurthy, S. R., and N. Roy, "StanceScorer: A data driven approach to score badminton player," in *IEEE International Conference on Smart Computing (SMARTCOMP)*, 2021.

[12] I. Ghosh, A. Chakma, Ramamurthy, S. R., N. Roy, and N. Waytowich, "PerMTL: A multi-task learning framework for skilled human performance assessment," in *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2022.

[13] M. S. Anwar, E. Dey, M. K. Devnath, I. Ghosh, N. Khan, J. Freeman, T. Gregory, N. Suri, K. Jayarajah, Ramamurthy, S. R., and N. Roy, "HeteroEdge: Addressing asymmetry in heterogeneous collaborative autonomous systems," in *IEEE International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2023.

[14] Ramamurthy, S. R., A. Manjitha, M. Radhakrishnan, and K. Jayarajah, "SVLMFeedbackBench: Benchmarking real-time robotic feedback using small vision-language models for human manipulation." Manuscript under review, 2026.

[15] M. Brown, S. Olivers, S. Srinivasan, and Ramamurthy, S. R., "MARSys: Multi-agent recommendation system with knowledge graph coordination and bayesian trust." Manuscript under review, 2026.

[16] D. Trivedi, R. Agada, and Ramamurthy, S. R., "VIGIL: Securing vlm controlled robots against cyberattacks." Manuscript under review, 2026.