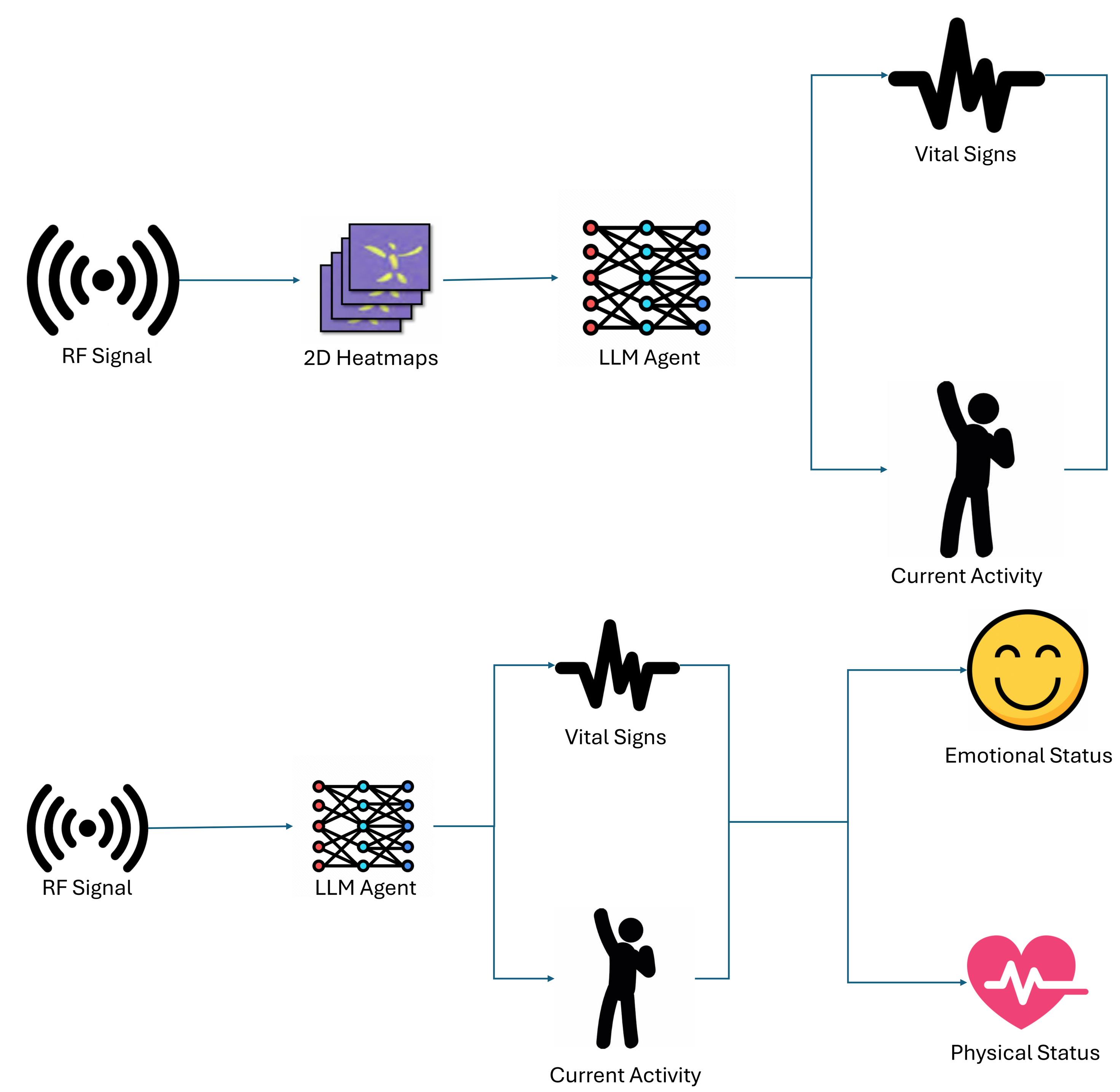
Bringing Large Language Models to Wireless Sensing

Andy Gonzalez, Jackie Ye, and Xin Liu

Introduction

Wireless networks are fundamental for communication, enabling devices to connect seamlessly. Beyond their primary role in data transmission, they also possess inherent sensing capabilities. By analyzing the interactions of wireless signals with the environment—such as reflections, attenuations, and changes caused by objects or human presence—these networks can be leveraged for a wide range of applications, including human activity recognition, environmental monitoring, and health tracking. Unlike traditional methods that rely on wearable sensors, wireless-based sensing offers a non-intrusive and scalable alternative.

However, state-of-the-art approaches for human activity recognition using wireless signals are computationally intensive and time-consuming, often requiring complex deep learning pipelines with multiple processing stages. To address these limitations, we propose an efficient methodology that integrates network sensing with Large Language Models (LLMs). By utilizing Channel State Information (CSI) to generate 2D heatmaps and video clips, which are then processed by a fine-tuned LLaVA model, we can significantly improve recognition accuracy while reducing computational overhead.



Methodology

Our approach integrates wireless sensing with a fine-tuned Large Language and Vision Model (LLaVA) to improve human activity recognition. Instead of relying on a multi-step deep learning pipeline, our method directly converts Raw Channel State Information (CSI) data into 2D heatmaps, which are then transformed into video clips. These clips are processed by LLaVA, leveraging its video encoder and multimodal capabilities for action prediction.

To assess the effectiveness of our method, we tested it on the Widar3.0 dataset, a widely used benchmark for CSI-based activity recognition. We compared its performance to the traditional deep learning approach, which processes 2D heatmaps into human mesh sequences before classifying actions. By eliminating multiple deep learning stages and integrating LLaVA, our method provides a more efficient and scalable solution for wireless-based human activity recognition.

References

Results

We evaluated our approach using the Widar3.0 dataset to measure the effectiveness of integrating wireless sensing with a fine-tuned LLaVA model. The baseline deep learning method achieved an accuracy of 6.37% for human activity recognition. After fine-tuning LLaVA for CSI-based action recognition, accuracy improved significantly to around 50%. These results demonstrate that LLM-based models can enhance performance while reducing computational complexity, making them a viable alternative to traditional deep learning pipelines.

Discussion

Our simplified approach opens new opportunities for practical deployment across diverse applications, especially in healthcare and well-being scenarios.

In healthcare settings, this technology could enable continuous, non-invasive monitoring of patients, providing caregivers with valuable insights into patient movements, behaviors, and physiological signals without the discomfort or inconvenience associated with wearable devices. Some potential applications could include:

- Behavioral and Mental Health Monitoring: Tracking patterns in daily activities or changes in routine behaviors that may indicate underlying mental health conditions or cognitive decline, thus facilitating early intervention.
- Remote Patient Monitoring: Allowing healthcare providers to remotely observe patients' vital signs and movement patterns, enhancing personalized care, especially for individuals with chronic conditions or those recovering at home.

Future research directions include refining the sensitivity of the model to subtle physiological signals, expanding the range of recognizable activities, and validating performance in clinical or real-world residential settings. Additionally, attention to privacy and data security will be essential to responsibly deploying these systems in healthcare contexts.