Demo Abstract : Light and Vibration Gesture Sensing with OTTER: Embedded Data Collection and Analysis Using LLMs

Steven Waskito*
National University of Singapore
Singapore
steven.waskito@u.nus.edu

Kai Jie Leow*
National University of Singapore
Singapore
e0959130@u.nus.edu

Pramuka Medaranga*
National University of Singapore
Singapore
pramukas@comp.nus.edu.sg

Tejas Gupta*
IIT Kanpur
India
tgupta21@iitk.ac.in

Shantanu Chakrabarty
NCS Group
Singapore
shantanu.chakrabarty@ncs.com.sg

Manoj Gulati
National University of Singapore
Singapore
manojg@nus.edu.sg

Ambuj Varshney
National University of Singapore
Singapore
ambujv@nus.edu.sg

ABSTRACT

The rapid growth in wireless embedded systems is threatened by the challenges associated with programming and deploying them. In addition, there is also the complexity inherent in analyzing of the sensor data. Notably, these tasks require high levels of end-user expertise. In this way, an entry barrier is introduced to deploying wireless embedded systems. In this work, we introduce Otter, an end-to-end system designed to simplify these tasks by leveraging the emergent properties of large language models. We demonstrate that Otter allows commodity embedded platforms to capture sensor data, such as light and vibration sensors, which can then be used to identify hand gestures in a near real-time manner. This is all while being prompted using natural language prompts by the end-user. Otter is the first system of its kind and has the potential to facilitate wireless embedded systems proliferation significantly.

KEYWORDS

Wireless embedded systems, Sensor data analysis, LLMs

ACM Reference Format:
Steven Waskito, Kai Jie Leow, Pramuka Medaranga, Tejas Gupta, Shantanu Chakrabarty, Manoj Gulati, and Ambuj Varshney. 2023. Demo Abstract : Light and Vibration Gesture Sensing with OTTER: Embedded Data Collection and Analysis Using LLMs. In *The 21st ACM Conference on Embedded Networked Sensor Systems (SenSys '23), November 12–17, 2023, Istanbul, Turkiye.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3625687.3628413

*These authors contributed equally to this work and are co-primary authors.

SenSys '23, November 12–17, 2023, Istanbul, Turkiye

ACM ISBN 979-8-4007-0414-7/23/11.
https://doi.org/10.1145/3625687.3628413

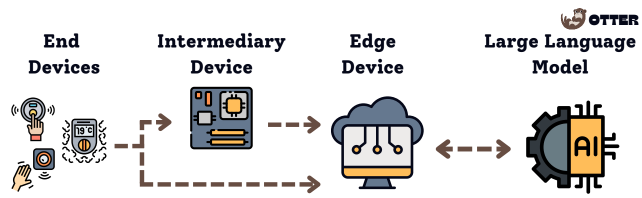


Figure 1: In Otter, end devices collect and transmit sensor data to an intermediary or edge device. This edge device conducts primary data preprocessing and then engages a large language model for sensor data analysis. The end-user only needs to provide natural language prompts. Thus, Otter greatly simplifies the deployment of WES.

1 INTRODUCTION

Wireless embedded systems (WES) are ubiquitously utilized across diverse sectors, exemplifying and extending Mark Weiser's vision of seamlessly integrated ubiquitous computing [8]. The various steps involved in a WES deployment generally involves programming devices, deployment, capturing of the data from a variety of sensors, transmitting the sensor data to an edge device, and then the comprehensive analysis of the sensor data using local or cloud-based tools [1]. All of these steps, in-particular, the analysis of the sensor data require a certain level of expertise from the end-users. This generally involves learning and employing machine learning techniques and custom models that are often trained on large datasets, which makes them less adaptable in diverse real-world scenarios, and thus hindering the wider deployment of WES.

The adaptability of the WES is the key attribute to facilitate the usage across a range of applications. In a way, the key attribute we are looking for is to be some what like a Swiss Army knife – it has a foundational design but can be adapted for various tasks without needing a separate tool for each one of the tasks.

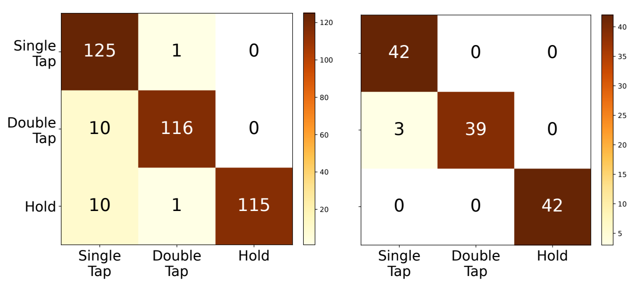
With the recent progress in LLMs, they are showcasing advanced reasoning[2, 3] and pattern recognition abilities[7], we believe that sensor data analysis using LLMs can serve as the "Swiss Army knife" in WES deployments, simplifying the deployment of WES, and thus enabling their widespread deployments.

To this end, we present OTTER: an end-to-end system that simplifies programming, sensor data collection, and subsequent data analysis in WES. OTTER collects, preprocesses, transmits sensor data, and crafts suitable prompts to interact with an LLM that runs locally on the edge device or in the cloud to analyze the sensor data. In this demonstration, we show the ability of OTTER to capture light and vibration sensor data in a near real-time manner, and then analyse the sensor data for identification of the hand gestures.

2 DESIGN

Figure 1 presents an overview of the OTTER. It comprises: (i) end devices that interface with sensors to capture and transmit data wirelessly; (ii) intermediate devices that receive and process data for compatibility from diverse end-devices; (iii) an edge device that selectively processes and forwards pertinent data as prompts to the LLM and delivers relevant responses from it to the end users.

The intermediate device may support communication through various protocols like Wi-Fi, BLE, ZigBee, and LoRa, as well as low-power techniques such as backscatter [4, 6], ensuring extensive compatibility with end devices. The edge device communicates with intermediate devices using the standardized MQTT protocol over a Wi-Fi connection. Varied sensors can be integrated into OTTER due to its modular architecture. In order to enhance the accuracy of LLM inference, the model can also be fine-tuned using prompts and expected results. Fine tuning also reduces the token usage.



(a) 252 train / 378 test, 94.2% accuracy

(b) 504 train / 126 test, 97.6% accuracy

Figure 2: We can identify hand gestures with high accuracy. Fine-tuning of the language model with the gesture data significantly improves the accuracy.

To fine-tune the language model, we utilized a dataset consisting of 630 instances of light-based gesture data. This dataset incorporated data gathered under varying light conditions and included information from seven participants across distinct light levels (low: 100-200 lux, medium: 600-750 lux, and high: 1500-1600 lux) and two distances (close: 2-4 cm and far: 8-10 cm). The dataset covered three gestures: Single Tap, Double Tap, and Hold. The data split for training and validation was as follows: (i) 252 Train vs. 378 Test Data and (ii) 504 Train vs. 126 Test Data. The results are depicted in Figure 2. The size of sample data used for fine-tuning is minimal when compared to traditional machine learning models employed.

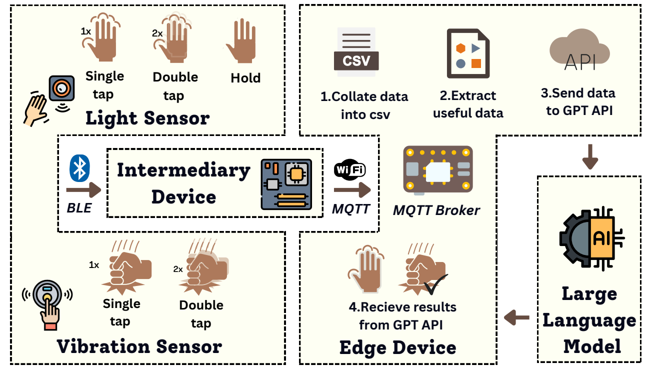


Figure 3: In the demonstration, we place a commodity embedded platform on the desk. It captures data from the light and vibration sensors. The captured data corresponds to the hand gestures. Next, these gestures are transmitted to an LLM, which is prompted by the end-user to detect the gestures. To improve accuracy, we also employ fine-tuning.

3 DEMONSTRATION

We demonstrate OTTER's capability to analyze light and vibration sensor data for gesture identification [5]. The demonstration setup employs a light sensor (APDS9960) and IMU (LSM9DS1) on the Arduino Nano 33 BLE Sense Lite to capture light and vibration signals, respectively. This data is transmitted via BLE to an ESP32, which acts as an intermediary and functions as an MQTT broker. It relays the collected data to the edge computer, a Raspberry Pi 4, that formulates prompts and sends them, along with the sensor data, to the LLM for analysis.Users can perform one or more of the specified gestures, and the OTTER end-to-end system analyzes these gestures using the LLM to determine the specific gesture performed.

ACKNOWLEDGMENTS

This work is supported through a startup grant (A-8000277-00-00), and a grant from the NUS-NCS centre (A-0008542-02-00) at the National University of Singapore.

REFERENCES

[1] Joshua Adkins et al. 2018. The Signpost Platform for City-Scale Sensing. In *ACM/IEEE IPSN 2018*.
[2] Md Tahmid Rahman Laskar et al. 2023. A Systematic Study and Comprehensive Evaluation of ChatGPT on Benchmark Datasets. arXiv:2305.18486 [cs.CL]
[3] Sébastien Bubeck et al. 2023. Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv:2303.12712 [cs.CL]
[4] Ambuj Varshney et al. [n. d.]. LoRea: A Backscatter Architecture That Achieves a Long Communication Range. In *ACM SenSys 2017*.
[5] Ambuj Varshney et al. 2017. Battery-Free Visible Light Sensing. In *ACM VLCS 2017*.
[6] Ambuj Varshney et al. 2022. Judo: Addressing the Energy Asymmetry of Wireless Embedded Systems through Tunnel Diode Based Wireless Transmitters. In *ACM MobiSys 2022*.
[7] Taylor Webb, Keith J Holyoak, and Hongjing Lu. 2023. Emergent analogical reasoning in large language models. *Nature Human Behaviour* (2023), 1–16.
[8] Mark Weiser. 1999. The Computer for the 21st Century. *SIGMOBILE Mob. Comput. Commun. Rev.* 3, 3 (jul 1999), 3–11. https://doi.org/10.1145/329124.329126