Remote Inference for Microcontrollers at the Edge

Jingtao Leo Zhang

Overall Project Goals and Specific Aims

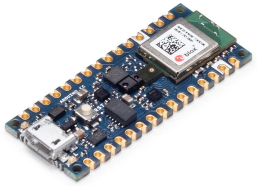
Objective

- Evaluate effectiveness of using networked CNN edge processing unit to improve ML performance on embedded devices.
- Metrics: accuracy, latency, throughput, power

Deliverables

1. Capability to send data from Arduino to Coral Dev Board, run inference on complex model, then return result.
2. Capability to run inference with simple model on Arduino
3. Performance metrics for 1 and 2.

Technical Approach

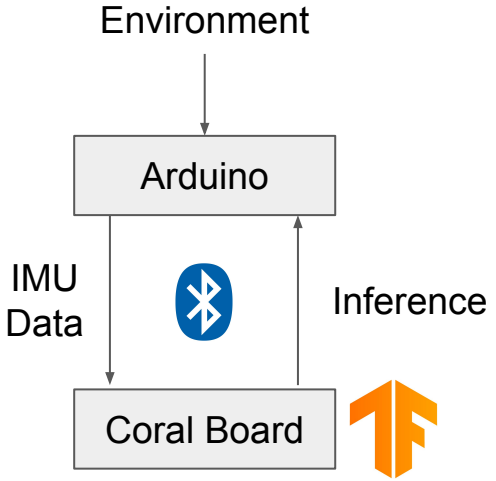


Hardware
- Arduino Nano 33 BLE Sense
  - IMU Sensor
- Coral Dev Board
  - Edge TPU

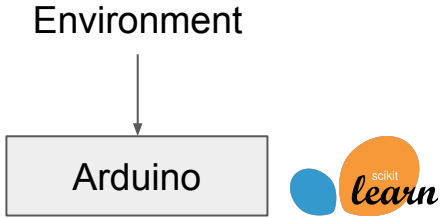
Software
- Bluetooth Low Energy
- CNN
  - TensorFlow Lite
  - Human Activity Recognition using IMU data
- SVM
  - Scikit

Data
- UCI Human Activities and Postural Transitions Data Set

Setup 1



Setup 2



Current Status

- Data collection using Arduino IMU functional
- BLE communication between Coral Dev Board and Arduino functional
- HAR CNN trained and converted to Edge TPU compatible TFlite model

Next Steps

- Integrate sensing, communication, and inference with target hardware
- Train and run SVM on Arduino
- Performance Measurement