

Poster Abstract: Trade-off Analysis of Inference Accuracy and Resource Usage for Energy-Positive Activity Recognition

Minh Tuan Tran
Queensland University of Technology
Brisbane, Australia
minhtuan.tran@connect.qut.edu.au

Muhammad Moid Sandhu
University of Queensland and
CSIRO's Data61
Brisbane, Australia
m.sandhu@uqconnect.edu.au

Sara Khalifa
CSIRO's Data61
Brisbane, Australia
sara.khalifa@data61.csiro.au

Gowri Ramachandran
Queensland University of Technology
Brisbane, Australia
g.ramachandran@qut.edu.au

Raja Jurdak
Queensland University of Technology
Brisbane, Australia
r.jurdak@qut.edu.au

ABSTRACT

Energy-positive activity recognition classifies human activities, including walking, running, and sitting, while harvesting kinetic energy from such activities. In this setting, the device's lifetime depends on the user's activity profile and the resources needed to run inference to classify activities. Thus, the selection of machine learning classification models for energy-positive activity recognition must consider both model's classification accuracy and energy consumption compared to the harvested energy from human activities. In this paper, we study the trade-off between accuracy and resource usage of a neural network model when different feature extraction techniques are used. Our results indicate that an on-board scheduling algorithm can be used to dynamically switch between the optimal feature input tuned for accuracy and energy consumption.

KEYWORDS

batteryless IoT, deep learning, energy-positive, activity recognition

1 INTRODUCTION

Batteryless sensing enables users to scavenge energy from solar, kinetic, thermal, and other environmental sources. Energy-positive human activity recognition [3] falls in the category, wherein the device both classifies human activities such as walking, running and sitting while harvesting energy from those activities. Such methods run machine learning inferencing algorithms on resource-constrained microcontrollers to classify human activities.

In batteryless sensing platforms, the device operates using only the harvested energy. For example, the user of a wearable device must perform physical activities such as running and walking to generate kinetic energy to enable energy-positive human activity recognition [3]. However, when the user is idle for a prolonged duration, the device will gradually run out of power. Thus, the batteryless sensing methodologies must proactively schedule sensing and inferencing tasks to extend the device's lifetime.

Existing literature investigates the resource requirements of inferencing on microcontrollers and proposes novel techniques for resource-efficient on-board inferencing [1]. Such approaches are beneficial for machine learning, but the trade-off analysis between inferencing accuracy and resource usage is inadequately addressed in the context of energy-positive activity recognition, which is the

focus of this work. Such an analysis allows the development of a scheduling algorithm to dynamically modify the inferencing model, signal sampling frequency, and accuracy requirements based on human activity prediction and future energy availability.

2 APPROACH

In this paper, we study the trade-off between accuracy and resource usage of a neural network model when different feature extraction techniques are employed. For this, we use Edge Impulse¹ which is a leading development platform for machine learning on edge devices. Below, we explain the details of the used datasets, feature extraction techniques, and the neural network model.

2.1 Datasets

Our experiments rely on two independent datasets with a total duration of over 8 hours covering six activities, including walking, sitting, standing, running, climbing stairs, and lying. One dataset is publicly available [2] which includes activity data from 30 volunteers wearing a smartphone with built-in sensors that sample an accelerometer at 50 Hz. The other dataset is collected for this study from one user using an Arduino Nano 33 BLE sense connected to the Edge Impulse backend, which collects data from the built-in accelerometer sensor at 62.5 Hz.

2.2 Feature extraction techniques

We employ 4 different feature extraction techniques. **Spectral analysis:** Edge Impulse extracts frequency and power properties of raw data with optional pre-filtering. After performing the Fast Fourier Transform, a filtered signal is used to extract spectral power peaks, including magnitude and frequency. **Flatten:** This approach employs statistical analysis of seven features for each axis, including average, minimum, maximum, root mean square, standard deviation, skewness, and kurtosis. **Raw (Deep Learning):** Raw data is used to create windows without any pre-processing (or feature extraction). **Spectral analysis and flatten:** A combination of features is extracted using both spectral features and flatten techniques.

¹<https://www.edgeimpulse.com/>

Dataset	Results	Spectral Analysis	Flatten	Raw (Deep Learning)	Spectral Analysis and Flatten
	Features	33	21	$2[\text{seconds}]/(1/\text{freq.})^*3$	54
Public Dataset - 360 minutes	Validation	60%	91.6%	93.5%	93.2%
	Testing	31.6%	82.76%	86.5%	86.5%
	On device Inference Performance	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.3KB	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.0KB	Inferencing time: 1ms Peak RAM usage: 1.9KB Flash Usage: 22.2KB	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.6KB
Data using Arduino Nano 33 BLE Sense - 140 minutes	Validation	79.7%	96.9%	94.5%	98%
	Testing	63.57%	90.91%	87.36%	91.48%
	On device Inference Performance	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.3KB	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.0KB	Inferencing time: 2ms Peak RAM usage: 2KB Flash Usage: 25.9KB	Inferencing time: 1ms Peak RAM usage: 1.7KB Flash Usage: 19.7KB

Table 1: Human activity recognition performance using various feature extraction techniques

2.3 Neural Network model

We employ a neural network model that consists of an input layer (extracted features), an output layer (six classes that need to be classified), and two middle layers. The first and second intermediate layers consist of 20 and 10 fully connected neurons respectively. We calculate the power requirement for inferencing on Arduino Nano 33, assuming the inference runtime of 1ms, and compare it with the harvested power from human activities (see Fig. 1).

3 RESULTS

Accuracy vs resource usage: Table 1 shows the trade-off between accuracy and resource usage for various feature extraction techniques. For the public dataset, the raw feature achieves the best accuracy of 93.5%, which is close to that of the Flatten’s accuracy of 93.2%. However, raw feature demands more resources while exhibiting longer inferencing time. With Arduino Nano 33 data, Flatten outperforms the others that applied spectral analysis and raw feature in terms of accuracy levels. This technique also requires the least resources considering the fewer features. To increase the accuracy, spectral analysis and flatten feature extraction can be used together, but it demands higher flash memory.

Energy comparison: In our prior work [3], we employed a wrist-wearable tiny kinetic energy harvester with 21 adult and healthy human subjects² to measure the energy harvested from running, walking, and sitting activities. Given Arduino Nano 33’s electrical characteristics from its datasheet (3.7V and 19mA during active mode), Fig. 1 (right) plots the average power consumption for running inference. We observe that average kinetic harvested power from a wearable device during various human activities is not sufficient to run the inference on the embedded device. Therefore, multi-source energy harvesters can be used to harvest higher power and a dynamic scheduler [4] can be employed to run inference on the embedded device using time-varying and dynamic harvested power. Our future work will use the results from Table 1 and Fig. 1 to develop a dynamic and energy-aware scheduler for managing the run-time trade-off in energy-positive activity recognition.

²Ethical approval has been granted from CSIRO [106/19] for carrying out this experiment.

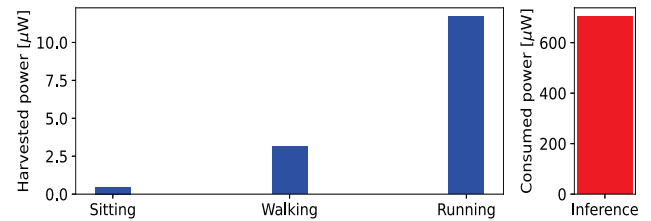


Figure 1: (Left) Average kinetic harvested power from various activities. (Right) Average inferencing power.

4 CONCLUSION

We have analysed the trade-off between classification accuracy and resource usage of different feature extraction techniques for a deep learning model using open-source and Arduino Nano 33 datasets. Our experimental results from the Edge Impulse platform show that a scheduling algorithm can be employed on microcontrollers to dynamically switch between resource-efficient and accurate inferencing, potentially extending the device lifetime. Our future work includes the development of such an energy-aware scheduling algorithm that can dynamically modify the inferencing model, signal sampling frequency, and sensing accuracy based on human activity prediction and corresponding energy production. Besides, we used average harvested power as a metric to compare different activities in this work. We acknowledge that different activities and their duration determine the harvested energy. In our future work, we consider conducting experiments of human activities in a typical day to calculate harvested energy compared to inference energy consumption.

REFERENCES

- [1] Seulki Lee and Shahriar Nirjon. 2019. Neuro.ZERO: A Zero-Energy Neural Network Accelerator for Embedded Sensing and Inference Systems. In *17th ACM SenSys* (New York, New York) (*SenSys '19*). 138–152.
- [2] Jorge-L Reyes-Ortiz et al. 2016. Transition-aware human activity recognition using smartphones. *Neurocomputing* 171 (2016), 754–767.
- [3] Muhammad Moid Sandhu et al. 2021. SolAR: Energy Positive Human Activity Recognition using Solar Cells. In *2021 IEEE PerCom*. 1–10.
- [4] Muhammad Moid Sandhu et al. 2021. Task Scheduling for Energy-Harvesting-Based IoT: A Survey and Critical Analysis. *IEEE IoT Journal* 8, 18 (2021), 13825–13848.