How to maximize the capabilities of in-mouth sensors for human activity recognition?

Maosheng Jiang (M.Jiang-5@student.tudelft.nl)
Responsible Professor: Koen Langendoen
Supervisor: Hayley Hung, Vivian Dsouza, Stephanie Tan

1. Introduction

Human Activity Recognition (HAR):
Automatic detection of human behavior using data sources such as sensors and camera footage.

Beneficial for humans:

- Health monitoring [1].
- Human-computer interaction [2].

Problems?

- Practicality issues with wearable sensors [3].
- Privacy concerns with camera footage [4].

Objective:
Performing HAR using a more practical approach without privacy concerns.

Research Question:
How to maximize the capabilities of in-mouth sensors for human activity recognition?

2. Methodology

- **Clean up** provided dataset.
- **Investigate** what kind of behavior can be recognized using an in-mouth sensor.
- **Extract features** from the sensor data to create training data and **balance** the training data.
- **Train classifier models:** Decision Trees (DT), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Logistic Regression (LR), and Random Forests (RF).
- **Implement classifier on the embedded microcontroller (MCU):** Implementing a Decision Tree classifier directly on a STM32F103C6T6 MCU.

3. Results

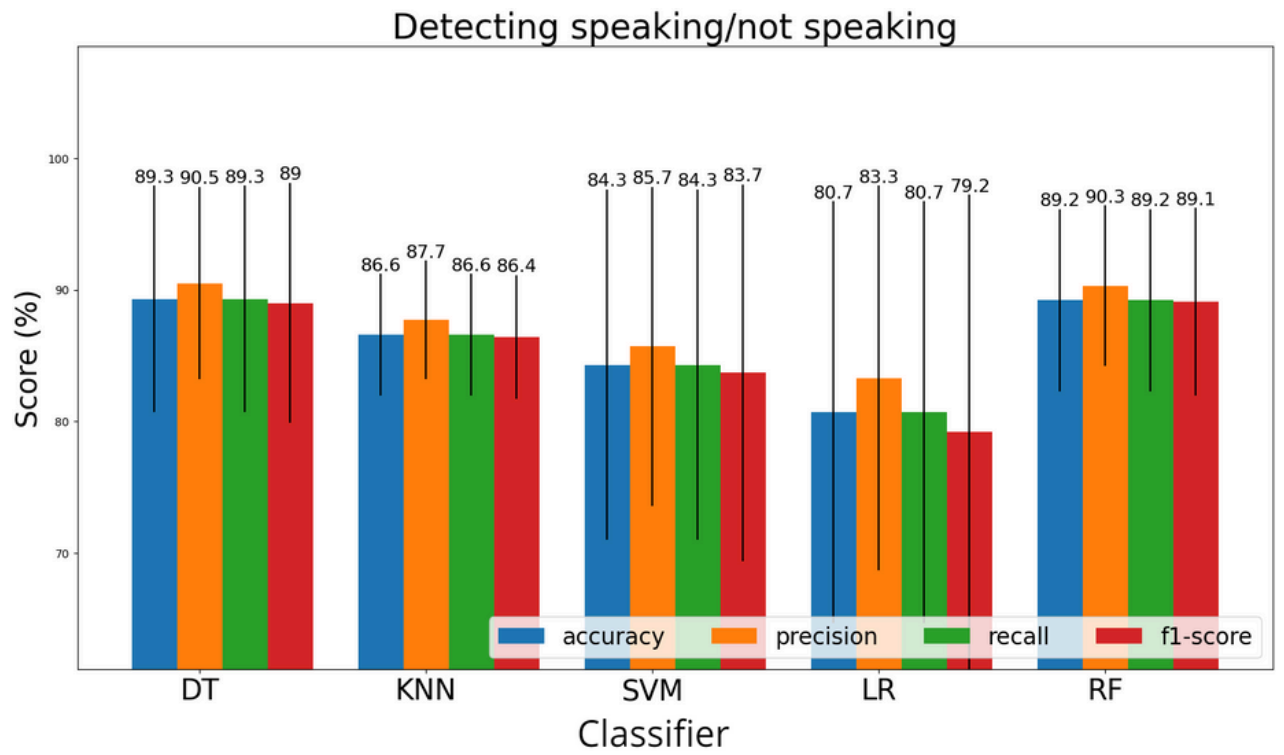

Figure 1:
Performance F1-score for detecting whether the user of the in-mouth sensor is speaking or not. Results are obtained after parameter tuning and applying a sliding window approach.

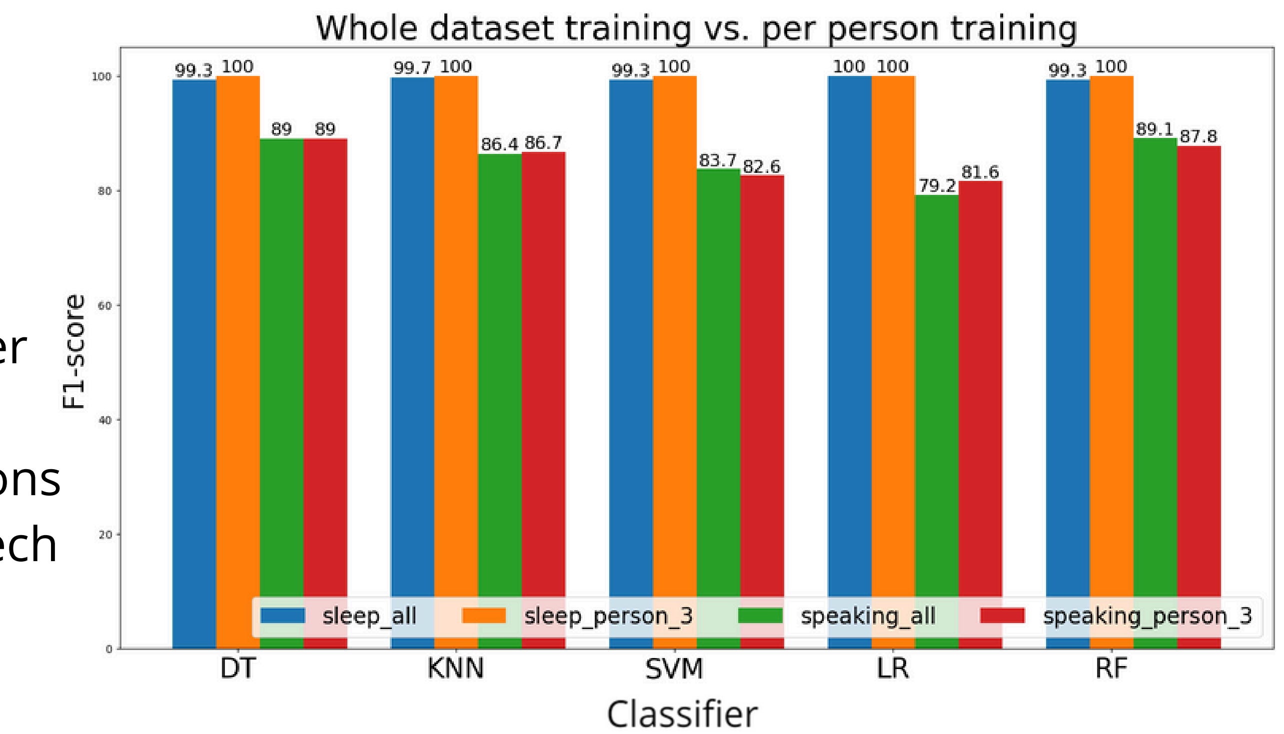

Figure 2:
Comparing models trained on the entire dataset with models trained individually per subject (subject-3) for detecting head positions during sleep and speech detection.

Classification	flash	RAM
Head up / left / right	80	0
Speaking / not speaking	1040	60
Walking / stationary	408	20
Inside / outside mouth	56	0
Mouth open / closed	360	0

Table 1:
Additional required flash memory and RAM for the implementation of a Decision Tree onto the MCU for the recognition of various human behavior.

TU Delft

4. Conclusions

- Using an in-mouth device for human activity recognition demonstrates promising results, achieving F1-scores exceeding 80%.
- Training models per person improves performance slightly for predicting certain behaviors, but the gains are not significantly greater than training on the entire dataset.
- Implementing a Decision Tree directly on a low-resource MCU is feasible within the space constraints.

5. Future Work

- Implement Decision Tree classifier onto the MCU in future versions of the in-mouth sensor.
- Investigate the applications of the ability to predict human behavior using an in-mouth sensor.
- Improve hardware of in-mouth sensor to be able to capture and recognize more complex human behavior.

References

[1] Sakorn Mekruksavanich and Anuchit Jitpattanakul. Fallnext: A deep residual model based on multi-branch aggregation for sensor-based fall detection. ECTI Transactions on Computer and Information Technology (ECTI-CIT), 16:352–364, 09 2022.

[2] Pablo Gallego Casćon, Denys J.C. Matthies, Sachith Muthukumarana, and Suranga Nanayakkara. Chewit. an intraoral interface for discreet interactions. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.

[3] Ilana Stolovas, Santiago Suarez, Diego Pereyra, Francisco De Izaguirre, and Varinia Cabrera. Human activity recognition using machine learning techniques in a low-resource embedded system. In 2021 IEEE URUCON, pages 263–267, 2021.

[4] Chirag Raman, Jose Vargas Quiros, Stephanie Tan, Ashraful Islam, Ekin Gedik, and Hayley Hung. Conflab: A data collection concept, dataset, and benchmark for machine analysis of free standing social interactions in the wild, 2022.