# Fine-Grained Activity Detection in the Kitchen with UWB

University of Alberta

Steven Phan

08 April 2023

# Contents

# Chapter 1

# Introduction

This thesis project will investigate how to detect fine-grained action within the meal preparation activity of daily living (ADL) in the home without the use of privacy-intruding cameras. ADLs are common activities that an individual performs inside their homes. These include walking around, eating, dressing, personal hygiene, toileting, transportation, meal preparation, house cleaning, and managing medication. The meal preparation ADL was chosen as the main focus because cooking is a uniquely enjoyable activity while being procedurally dense. Meal preparation can include the following actions: opening the fridge, retrieving ingredients, cutting vegetables, and assembling the ingredients. Monitoring these actions may be used as part of a health monitoring program by enabling the assessment of the presence and length of each individual step in a goal-orientated activity. This information may help guide interventions and track the effectiveness of interventions in clinical populations such as people with dementia or frailty.

It is hypothesized that the combination of context, such as hyper accurate indoor localization (down to 30 cm), and inertial data can accurate and reliable classification of fine-grained ADLs. The system used for indoor localization is the Pozyx Creator Kit which provides a wrist mounted wearable that can obtain data at a maximum of 60 Hz [2]. This data includes position relative to a floorplan, and inertial data from a BNO055 which outputs 3D Acceleration, 3D angular velocity, 3D Linear Acceleration, and the Heading, Pitch and Roll.

Prior to any experiments related to classification of these cooking actions, the optimal configuration of the system that provides reliable position data had to be investigated. The next section details the different attempts at changing the configuration of the Pozyx Creator Kit to obtain the most reliable positioning data in X, Y, and Z at a satisfactory sampling rate.

# Chapter 2

# Literature Review

This thesis project will investigate how to detect fine-grained action within the meal preparation activity of daily living (ADL) in the home without the use of privacy-intruding cameras. ADLs are common activities that an individual performs inside their homes. These include walking around, eating, dressing, personal hygiene, toileting, transportation, meal preparation, house cleaning, and managing medication. The meal preparation ADL was chosen as the main focus because cooking is a uniquely enjoyable activity while being procedurally dense. Meal preparation can include the following actions: opening the fridge, retrieving ingredients, cutting vegetables, and assembling the ingredients. Monitoring these actions may be used as part of a health monitoring program by enabling the assessment of the presence and length of each individual step in a goal-orientated activity. This information may help guide interventions and track the effectiveness of interventions in clinical populations such as people with dementia or frailty.

It is hypothesized that the combination of context, such as hyper accurate indoor localization (down to 30 cm), and inertial data can accurate and reliable classification of fine-grained ADLs. The system used for indoor localization is the Pozyx Creator Kit which provides a wrist mounted wearable that can obtain data at a maximum of 60 Hz [2]. This data includes position relative to a floorplan, and inertial data from a BNO055 which outputs 3D Acceleration, 3D angular velocity, 3D Linear Acceleration, and the Heading, Pitch and Roll.

Prior to any experiments related to classification of these cooking actions, the optimal configuration of the system that provides reliable position data had to be investigated. The next section details the different attempts at changing the configuration of the Pozyx Creator Kit to obtain the most reliable positioning data in X, Y, and Z at a satisfactory sampling rate.

# Chapter 3

# System Tuning at the Independent Living Suite

## 3.1 System Tuning Review

The Pozyx Creator Kit comes with anchors and several tags. Anchors are mounted on the walls and are used to position the tags. Multiple tags may be positioned at the same time. The Pozyx Creator kit uses ultrawideband (UWB) signals with the two-way ranging protocol to localize the tag. The tag is mounted on custom 3D printed wearables which the participant can wear as a wrist-watch or a necklace. Through trial-and-error and consultation with the Pozyx Creator Documentation [3, 1] it was determined that the accuracy of the system depends on factors listed below:

- Number of anchors
- Position of anchors

These variables were modified to achieve satisfactory actual position error and standard deviation below the expected error of 30 cm for UWB systems. The protocol for obtaining data and evaluating the actual position error and standard deviation will be described.

## 3.2 Methodology

This protocol tests the X, Y, and Z positional accuracy of the Pozyx Creator system in the Independent Living Suite (ILS) at the Glenrose Rehabilitation

Hospital by having a participant stand at a specific location in each room. Permanent appliances or furniture such as the stove or dining table were used as much as possible to ensure that the experiment is repeatable.

### 3.2.1 Setup

Masking tape was used to mark the locations where the participant should place their feet. The following procedure was followed to place the tape:

1. Using a measuring tape, measure 1 meter out from the middle of the appliance or furniture and place a 20 cm piece of tape centered on, perpendicular to and underneath the measuring tape (the tips of the participant's toes should be 1 meter away from the appliance).

2. Place parallel tape on the sides of the tape placed in Step 1 to constrain the feet to a box. (The participant should have their toes on the tape perpendicular to the measuring tape and usually facing the appliance or furniture). Figure 3.1 outlines some examples of tape placements.

Following the tape placement guidelines outlined at the beginning of this section, tape was placed at or near the following locations. Refer to the AUTOCAD floor plan for the location of the rooms (Figure 3.2):

- The Hallway between Living Room and Kitchen facing the Dining Table.

- The Living Room facing the Desk.

- The Bedroom facing the bed.

- The Hallway between the bedroom and the bathroom, facing away from the wall.

- Bathroom facing the toilet.

- Kitchen facing the stove.

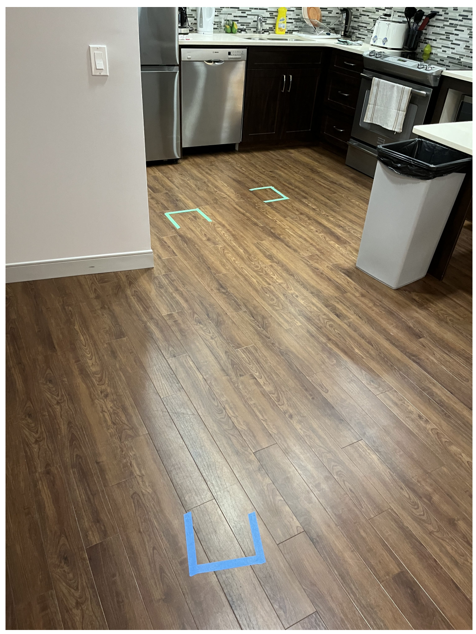


Figure 3.1: Box tape placement at the stove, fridge, and dining table. Participant's toes and sides of feet should touch the tape.

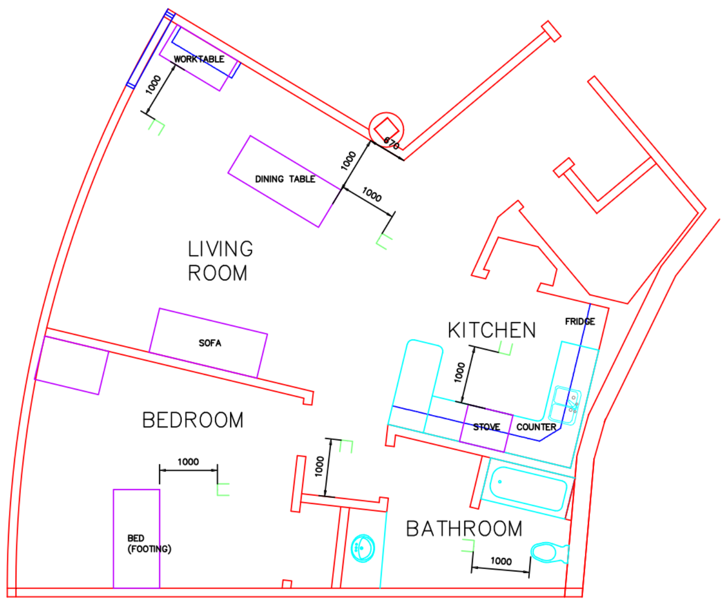



Figure 3.2: Floor plan of the ILS. Positions where the participant stood 1 meter from appliances or furniture are marked in the green "open" box on the floorplan.

### 3.2.2 Protocol

A stopwatch python script was created with predetermined labels and used as the ground truth for positions. A single participant wore the tag on a 3D printed necklace mount (Figure 3.3). The measuring tape was used to measure the height from the ground and height when squatting. For this participant, the standing height was **144cm** and the squatting height was **68.5cm**.

Figure 3.3: The Pozyx tag mounted in a custom 3D printed necklace mount.

The protocol had the following steps:

1. At the first location (Hallway Between Living Room and Kitchen) stand still for 10 seconds
2. Squatt still for 10 seconds.

3. Move to next position.

4. Repeat steps 1-3 until all of the positions have been reached.

5. Finally return to the first position (Hallway Between Living Room and Kitchen)

There were 3 trials for each configuration. Following the guidelines from the Pozyx Creator Setup [1] anchors were staggered at heights of 1.4m and 2.4m (ceiling height) for 4, 5, 6, 8, and 10 anchors. A configuration where anchors were all low (10cm) were tested for 8 anchors and configurations where anchors were all high (2.4m) were tested for 8 and 9 anchors.

## 3.3 Results

Trials for each configuration were aggregated, transition periods were removed, data of interest was time normalized and the error and standard deviation of each location while standing and squatting were calculated. An as-built AUTOCAD file of the ILS was used to obtain the real position and used in the calculation of the error between measured versus the actual position. The results of the experiment are summarized in the heatmap tables (Figures 3.4, 3.5 and 3.6) with a minimum darkness set at 30cm and a maximum darkness set at 60cm.

X Position Error at Each Location (cm)

| | POS_X_A4 | POS_X_A5 | POS_X_A6 | POS_X_A8 | POS_X_A10 | POS_X_A8H | POS_X_A8L | POS_X_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 40.1 | 41.5 | 42.2 | 47.9 | 33.0 | 11.8 | 3.5 | 28.0 |
| **Go Hallway between kitchen and living(sit)** | 18.6 | 11.7 | 7.4 | 44.2 | 20.6 | 29.2 | 21.9 | 41.4 |
| **Living Room** | 50.4 | 57.8 | 48.3 | 7.7 | 34.5 | 20.0 | 31.3 | 41.8 |
| **Living Room(sit)** | 44.6 | 30.2 | 20.0 | 28.8 | 23.7 | 5.2 | 6.5 | 11.0 |
| **bathroom** | 49.1 | 39.8 | 33.0 | 65.3 | 15.4 | 111.9 | 120.6 | 52.0 |
| **bathroom(sit)** | 29.2 | 19.6 | 22.9 | 23.2 | 39.0 | 64.5 | 24.6 | 61.1 |
| **bedroom** | 24.1 | 2.5 | 39.4 | 58.3 | 55.8 | 99.8 | 78.3 | 49.2 |
| **bedroom(sit)** | 26.9 | 12.8 | 4.1 | 103.8 | 45.9 | 53.5 | 27.0 | 37.5 |
| **hallway between bedroom and bathroom** | 2.7 | 14.9 | 5.3 | 5.4 | 7.6 | 6.6 | 3.7 | 36.1 |
| **hallway between bedroom and bathroom(sit)** | 2.8 | 13.9 | 26.5 | 0.1 | 3.4 | 1.7 | 51.3 | 22.2 |
| **kitchen** | 23.7 | 26.8 | 18.5 | 30.1 | 31.5 | 37.3 | 29.1 | 7.7 |
| **kitchen(sit)** | 8.2 | 22.2 | 8.8 | 41.9 | 31.2 | 43.7 | 44.1 | 28.0 |

(a)

X Standard Deviation at Each Location (cm)

| | POS_X_A4 | POS_X_A5 | POS_X_A6 | POS_X_A8 | POS_X_A10 | POS_X_A8H | POS_X_A8L | POS_X_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 105.9 | 56.5 | 62.1 | 39.2 | 42.2 | 43.8 | 29.1 | 28.2 |
| **Go Hallway between kitchen and living(sit)** | 9.8 | 19.2 | 12.7 | 24.5 | 13.7 | 12.3 | 20.0 | 15.6 |
| **Living Room** | 5.4 | 7.9 | 6.4 | 18.7 | 19.1 | 17.3 | 12.0 | 12.9 |
| **Living Room(sit)** | 11.6 | 24.9 | 19.2 | 18.0 | 20.2 | 17.4 | 16.2 | 13.2 |
| **bathroom** | 44.4 | 27.3 | 19.6 | 82.5 | 43.4 | 65.3 | 90.6 | 39.1 |
| **bathroom(sit)** | 13.8 | 12.6 | 9.7 | 68.0 | 39.7 | 44.5 | 80.2 | 23.1 |
| **bedroom** | 17.7 | 17.3 | 17.6 | 21.0 | 21.6 | 36.3 | 30.5 | 24.7 |
| **bedroom(sit)** | 7.9 | 20.7 | 28.1 | 28.3 | 14.6 | 29.5 | 29.1 | 14.4 |
| **hallway between bedroom and bathroom** | 10.2 | 12.5 | 6.9 | 24.6 | 18.1 | 16.0 | 13.6 | 19.3 |
| **hallway between bedroom and bathroom(sit)** | 14.6 | 4.8 | 10.8 | 13.0 | 13.4 | 10.1 | 12.3 | 12.9 |
| **kitchen** | 16.7 | 7.8 | 10.3 | 5.1 | 5.0 | 7.9 | 7.9 | 7.8 |
| **kitchen(sit)** | 20.7 | 6.6 | 16.4 | 14.4 | 9.6 | 8.5 | 9.7 | 5.1 |

(b)

Figure 3.4: The positional error in X (a) and the standard deviation in X (b) at each location and body position

Y Position Error at Each Location (cm)

| | POS_Y_A4 | POS_Y_A5 | POS_Y_A6 | POS_Y_A8 | POS_Y_A10 | POS_Y_A8H | POS_Y_A8L | POS_Y_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 20.5 | 28.7 | 38.8 | 95.4 | 62.3 | 58.2 | 8.6 | 28.5 |
| **Go Hallway between kitchen and living(sit)** | 46.5 | 58.2 | 57.9 | 65.5 | 49.5 | 63.4 | 31.7 | 42.7 |
| **Living Room** | 11.6 | 6.2 | 3.2 | 62.3 | 63.5 | 47.1 | 37.4 | 38.0 |
| **Living Room(sit)** | 7.8 | 0.7 | 3.7 | 47.8 | 35.6 | 25.8 | 32.7 | 15.1 |
| **bathroom** | 27.9 | 0.4 | 7.8 | 44.4 | 56.1 | 47.3 | 10.3 | 38.2 |
| **bathroom(sit)** | 54.7 | 39.2 | 37.7 | 53.7 | 29.7 | 5.1 | 15.1 | 1.0 |
| **bedroom** | 22.8 | 14.5 | 9.2 | 39.5 | 16.1 | 71.7 | 58.4 | 5.7 |
| **bedroom(sit)** | 20.2 | 5.3 | 33.9 | 60.9 | 5.0 | 34.6 | 1.8 | 8.7 |
| **hallway between bedroom and bathroom** | 4.6 | 7.8 | 24.8 | 7.1 | 9.4 | 23.5 | 2.0 | 19.4 |
| **hallway between bedroom and bathroom(sit)** | 56.2 | 67.0 | 44.4 | 31.0 | 7.7 | 28.3 | 8.8 | 13.8 |
| **kitchen** | 16.8 | 2.8 | 10.8 | 28.5 | 33.7 | 39.0 | 38.0 | 31.3 |
| **kitchen(sit)** | 17.5 | 4.0 | 9.0 | 21.9 | 31.1 | 15.5 | 28.6 | 22.2 |

(a)

Y Standard Deviation at Each Location (cm)

| | POS_Y_A4 | POS_Y_A5 | POS_Y_A6 | POS_Y_A8 | POS_Y_A10 | POS_Y_A8H | POS_Y_A8L | POS_Y_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 103.7 | 53.5 | 62.1 | 115.5 | 79.0 | 137.9 | 25.8 | 24.7 |
| **Go Hallway between kitchen and living(sit)** | 10.3 | 13.7 | 10.7 | 23.6 | 17.8 | 21.1 | 27.6 | 16.0 |
| **Living Room** | 5.7 | 8.3 | 4.8 | 13.7 | 16.1 | 17.4 | 20.8 | 14.2 |
| **Living Room(sit)** | 11.1 | 14.7 | 11.0 | 21.3 | 12.2 | 16.3 | 13.1 | 7.5 |
| **bathroom** | 35.1 | 25.1 | 21.9 | 50.8 | 24.8 | 36.2 | 39.8 | 23.4 |
| **bathroom(sit)** | 29.8 | 42.1 | 29.3 | 19.1 | 23.4 | 20.3 | 38.9 | 17.8 |
| **bedroom** | 15.6 | 17.7 | 14.2 | 18.9 | 12.2 | 18.4 | 23.7 | 9.5 |
| **bedroom(sit)** | 9.5 | 12.4 | 14.3 | 24.9 | 14.9 | 16.8 | 17.5 | 6.2 |
| **hallway between bedroom and bathroom** | 15.2 | 9.0 | 12.2 | 23.0 | 13.6 | 13.4 | 9.5 | 13.0 |
| **hallway between bedroom and bathroom(sit)** | 12.7 | 10.5 | 16.7 | 8.0 | 12.9 | 12.3 | 12.1 | 8.2 |
| **kitchen** | 17.6 | 9.1 | 6.2 | 9.9 | 7.1 | 14.0 | 9.0 | 9.8 |
| **kitchen(sit)** | 24.4 | 16.0 | 17.2 | 15.3 | 11.0 | 12.7 | 16.1 | 5.9 |

(b)

Figure 3.5: The positional error in Y (a) and the standard deviation in Y (b) at each location and body position

Z Position Error at Each Location (cm)

| | POS_Z_A4 | POS_Z_A5 | POS_Z_A6 | POS_Z_A8 | POS_Z_A10 | POS_Z_A8H | POS_Z_A8L | POS_Z_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 132.3 | 56.9 | 47.9 | 180.1 | 87.6 | 7.5 | 58.2 | 3.2 |
| **Go Hallway between kitchen and living(sit)** | 137.0 | 207.8 | 184.0 | 232.3 | 104.7 | 7.6 | 22.8 | 16.0 |
| **Living Room** | 121.5 | 36.7 | 41.7 | 99.4 | 88.5 | 58.1 | 54.8 | 57.0 |
| **Living Room(sit)** | 79.1 | 177.1 | 197.1 | 224.9 | 16.7 | 30.8 | 97.0 | 17.2 |
| **bathroom** | 84.0 | 66.7 | 114.7 | 64.2 | 27.3 | 43.6 | 147.6 | 80.3 |
| **bathroom(sit)** | 43.2 | 10.2 | 11.2 | 211.1 | 66.4 | 26.3 | 4.6 | 66.2 |
| **bedroom** | 57.5 | 65.8 | 121.6 | 192.9 | 233.5 | 155.3 | 234.6 | 37.1 |
| **bedroom(sit)** | 39.7 | 1.3 | 54.8 | 305.5 | 3.5 | 80.9 | 142.5 | 20.6 |
| **hallway between bedroom and bathroom** | 74.0 | 103.4 | 4.1 | 208.3 | 48.6 | 30.8 | 72.3 | 71.7 |
| **hallway between bedroom and bathroom(sit)** | 33.4 | 26.2 | 59.1 | 84.2 | 2.8 | 61.1 | 54.4 | 65.2 |
| **kitchen** | 70.5 | 117.0 | 40.5 | 0.1 | 38.7 | 33.5 | 73.8 | 31.5 |
| **kitchen(sit)** | 119.0 | 3.9 | 25.9 | 216.1 | 70.6 | 38.5 | 139.6 | 20.5 |

(a)

Z Standard Deviation at Each Location (cm)

| | POS_Z_A4 | POS_Z_A5 | POS_Z_A6 | POS_Z_A8 | POS_Z_A10 | POS_Z_A8H | POS_Z_A8L | POS_Z_A9H |
|---|---|---|---|---|---|---|---|---|
| **Go Hallway between kitchen and living** | 93.9 | 80.0 | 89.6 | 68.1 | 69.9 | 58.7 | 110.3 | 33.8 |
| **Go Hallway between kitchen and living(sit)** | 84.8 | 110.0 | 123.2 | 151.2 | 108.1 | 21.0 | 77.3 | 9.0 |
| **Living Room** | 46.7 | 36.1 | 23.5 | 92.1 | 25.6 | 20.0 | 144.2 | 17.1 |
| **Living Room(sit)** | 53.1 | 118.1 | 112.8 | 107.5 | 106.7 | 13.2 | 87.2 | 22.0 |
| **bathroom** | 89.9 | 105.2 | 69.6 | 99.9 | 77.3 | 79.4 | 151.9 | 40.7 |
| **bathroom(sit)** | 107.7 | 18.3 | 35.4 | 90.0 | 115.6 | 119.0 | 93.6 | 63.7 |
| **bedroom** | 65.4 | 109.7 | 67.8 | 142.2 | 58.5 | 24.2 | 246.9 | 20.0 |
| **bedroom(sit)** | 15.1 | 28.4 | 66.1 | 154.6 | 130.0 | 25.6 | 144.2 | 12.1 |
| **hallway between bedroom and bathroom** | 37.2 | 50.0 | 21.1 | 51.5 | 98.9 | 31.1 | 150.1 | 30.2 |
| **hallway between bedroom and bathroom(sit)** | 104.2 | 13.2 | 24.7 | 176.0 | 112.5 | 14.1 | 57.9 | 15.7 |
| **kitchen** | 75.2 | 62.0 | 64.5 | 61.8 | 13.2 | 24.4 | 26.3 | 17.0 |
| **kitchen(sit)** | 149.3 | 46.5 | 95.5 | 172.8 | 119.5 | 15.9 | 42.7 | 7.2 |

(b)

Figure 3.6: The positional error in Z (a) and the standard deviation in Z (b) at each location and body position

## 3.4 Discussion

### 3.4.1 X Position

Visually, the heatmap of error in the X position shows different spots where the system struggled to obtain the location based on the AUTOCAD as-builts depending on the configuration selected. For 4, 5 and 6 anchors, the errors seemed to be larger in the living room and the hallway between the kitchen and the living room. 8, 10, 8 (L)ow, 8 (H)igh, and 9H anchors seemed to struggle most around the bathroom and bedroom area. The standard deviation in X position seems to follow a similar pattern where 4, 5 and 6 anchors have higher standard deviation in hallway between the kitchen and living area and 8, 10, 8 (L)ow, 8 (H)igh, and 9H seems to struggle the most in the bathroom. Out of all of the configurations the 9H configuration has the most locations where the standard deviation is acceptable.

### 3.4.2 Y Position

For 4, 5, and 6 anchors, the error seems to increase in the seated position meaning that there may be some dependence on the Z position. This occurs in the hallway between the kitchen and the living room, the bathroom and the hallway between the bedroom and bathroom. There seems to be a large struggle for 8, 10 and 8H anchors to pinpoint the Y position in the hallway between the kitchen and the living, the living room and the bathroom. The 8L anchor configuration struggled when the in the bedroom, but was overall within or near the acceptable threshold of 30cm. 9H anchors overall seemed to be the best at determining the Y position with mild errors at the hallway between the kitchen and the living room, the living room and the bathroom.

In terms of standard deviation, anchor configurations 4, 5, 6, 8, 8H, and 10 had trouble at a height of 144cm, but otherwise had low standard deviation. 8L had minor issues regarding standard deviation in the bathroom but was otherwise low. The 9H configuration seemed to yield the lowest standard deviations in the Y Position.

### 3.4.3 Z Position

The Z position at many of the locations and all configurations seem to deviate from the measured heights and have high error. Only the 8H and 9H configurations have acceptable standard deviations for most of the rooms (there is still some struggle in the bathroom). Considering the inaccuracies in the Z positioning, it is recommended that the Z not be used as a absolute source of truth for height. Rather Z position should be used relative to another reference tag with the 9H configuration. For example, a necklace tag may be combined a wrist tag. When standing, the position of the wrist may be compared with the position of the necklace to determine if the wrist is above, below or at chest height.

### 3.4.4 Overall

The 9H configuration seems to provide the most reliable data when observing the standard deviations of the X, Y, and Z positions. With this configuration, each room had around 4 anchors surrounding it Figure 3.7

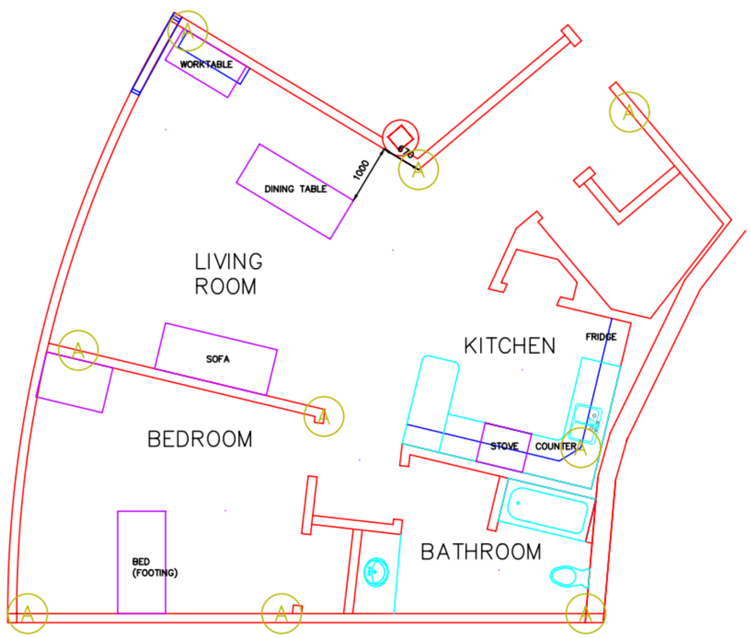



Figure 3.7: ILS Floorplan with the 9 anchors all high.

Though the inaccuracies in the hallway, living room, bathroom and bedroom may prevent the 9H configuration from using heuristics for classification at these locations, the inherent repeatability evident in the low standard deviation in each axes of the position can make the position data from the 9H configuration a candidate for machine-learning based classification.

# Chapter 4

# Pilot Testing of Detecting Activities to make a Sandwich

Continuing from the findings in Chapter 2, the 9H anchor configuration was used to perform a preliminary classification of activities in the kitchen. Steps performed in making a sandwich were broken down and organized into Setup, Preparation, Cooking, and Finishing steps, Figure 4.1.

From the actions shown in Figure 4.1, actions with distinct location or patterns were selected as classes for classification. OPENFRIDGE, OPENFREEZER, and GETPLATE were selected as classes from the Setup category, washing hands/vegetables/fruits/using the kitchen sink were grouped into a WASHHANDS category, and SLICETOMATO were selected as a class. Finally, All intermediary transitions or motionless segments were grouped into a UNDEFINED category. Single trials were performed to collect data for each of these classes.

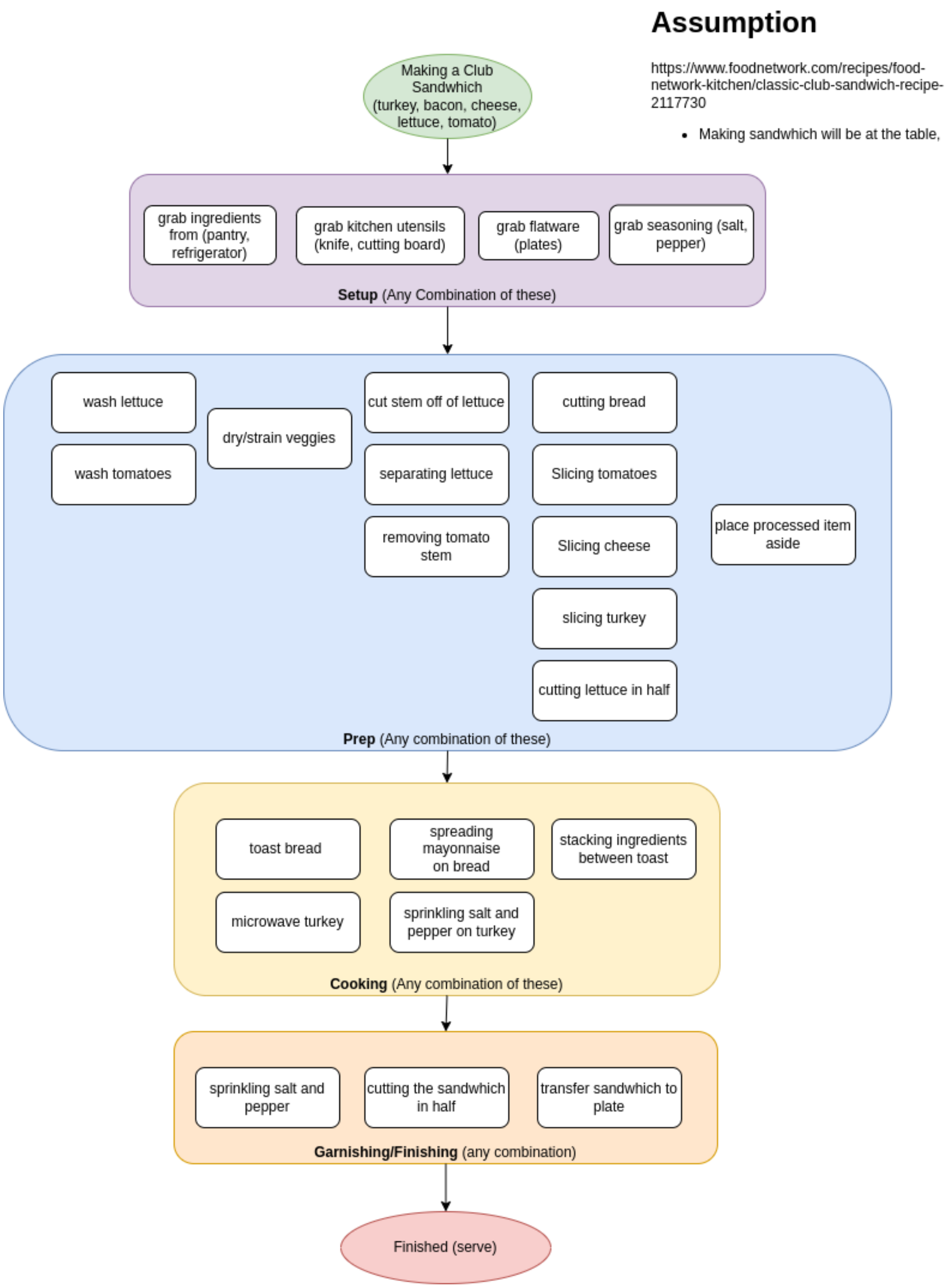



Figure 4.1: Task decomposition of making a sandwich.

## 4.1 Experimental Protocol

### 4.1.1 Setup

Several points were enforced to ensure that the training dataset captures the variation in action sufficiently when classifying data from right-handed individuals.

- Pozyx Tag is mounted on the right wrist (Figure 4.2).
- Initial position for each of the single trials are not marked. Participant will be able to choose a location from which they can perform the action comfortably without moving their feet.
- An action starts when the individual contacts the appliance or furniture. For SLICETOMATO the action starts when an individual starts slicing the tomato and ends when they stop slicing the tomato. Motions such as picking up the knife and getting in position to slice were considered transitions and labelled as UNDEFINED.

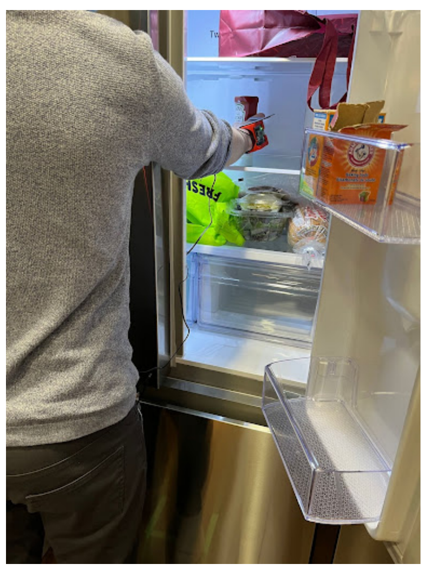


Figure 4.2: Pozyx tag mounted on the wrist. The participant is performing the OPENFRIDGE task

### 4.1.2 Data Collection

Custom Python stopwatch scripts were created to accurately label periods of transitions (quiet standing + getting into position for the action) and the action. An example of the data collected is shown in Figure 4.3. For each action there is a quiet standing period at the beginning and end. OPENFRIDGE, OPENFREEZER, OPENPLATE, WASHHANDS each had 5 repetitions for each trial. SLICETOMATO contained 3 slices to conserve the amount of tomato. Each action had a total of 5 trials.

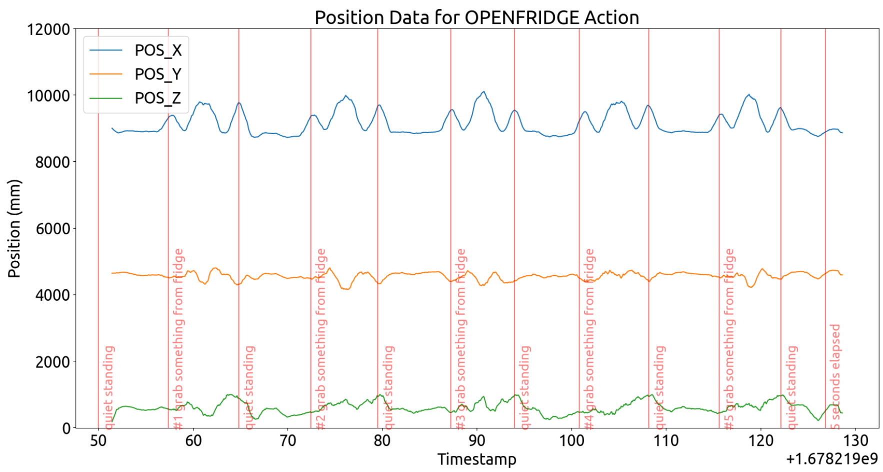



Figure 4.3: Labelled position data of the OPENFRIDGE action. Note that the "quiet standing" periods do not consist entirely of quiet standing, but also include traces of transitions from getting into the correct position to perform the action.

Since the Pozyx Tag contained a BNO055 chip, in addition to 3D Position data, the tags were able to capture inertial data including Accelerometer Data, Linear Accelerometer Data, Angular Velocity Data, and the orientation.

Data for each of the actions that relate to making a sandwich were collected from 2 participants.

## 4.2 Feature Extraction

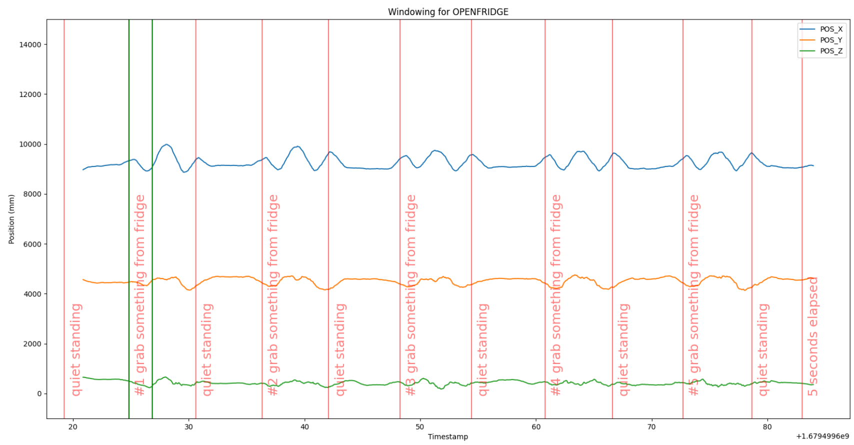
An initial sliding window with a width of 2 seconds and a stride length of 1 second was used to ensure that enough feature vectors could be extracted from the SLICETOMATO dataset. An example of a windows taken for the OPENFRIDGE action and UNDEFINED action are shown in Figure 4.4.

From each window, basic statistical measures over the entire window were taken. These measures include the MEAN, MEDIAN, MODE (to 5cm for position), MAX, MIN, and STD of the entire window. From each window of data, there were a total of 3 (axes) * 5 (types of data) * 6 (statistical measures) = 90 Features

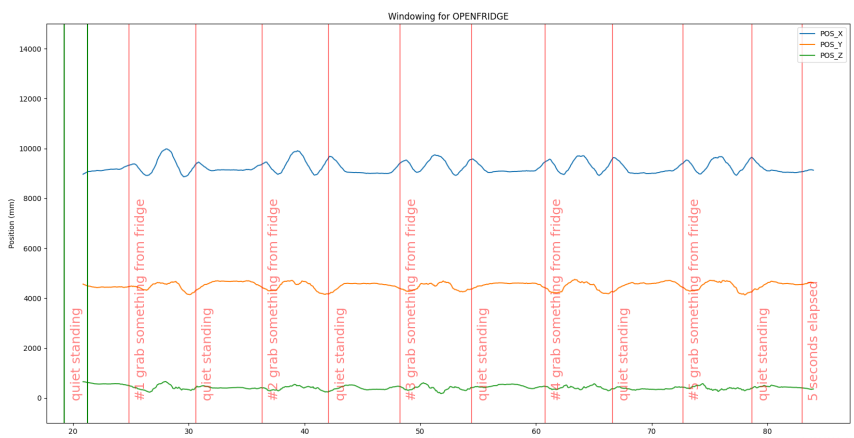
From the entire timeseries dataset, 2773 feature vectors were extracted. Refer to Table 4.1 for the breakdown of counts for each label.

Table 4.1: Count of the occurrences of each action.

| Action | Count |
|---|---|
| UNDEFINED | 1518 |
| SLICETOMATO | 197 |
| WAHSHANDS | 316 |
| OPENFRIDGE | 239 |
| OPENFREEZER | 203 |
| GETPLATE | 300 |




(a)




(b)

Figure 4.4: Obtaining windows from the OPENFRIDGE dataset. The green vertical lines section off a 2-second window. (a) A window labelled OPENFRIDGE. (b) A window labelled UNDEFINED.

## 4.3 Model Selection

A 60:40 split was used to train and test the model selected. Several models were chosen including Linear Support Vector Machine, Radial Support Vector Machine, K-Nearest Neighbors, Decision Trees and Random Forests. As this was a pilot study in determining the feasibility of classification of the fine-grained actions involved in making a sandwich, rigorous parameter tuning and feature selection were neglected and the defaults from the sklearn Python package were used.

## 4.4 Results

The confusion matrices from each model are output in Figures 4.5-4.10. Total accuracy was reported as well as the sensitivity, specificity, and precision of each class were reported. These measures are calculated as follows:

$$\text{Accuracy} = \frac{\text{All}\, TP}{N} \tag{4.1}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{4.2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.3}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4.4}$$

Where $N$ is the number of samples $TP$ are True Positives, $TN$ are true negatives, $FP$ are false positives and $FN$ are false negatives.

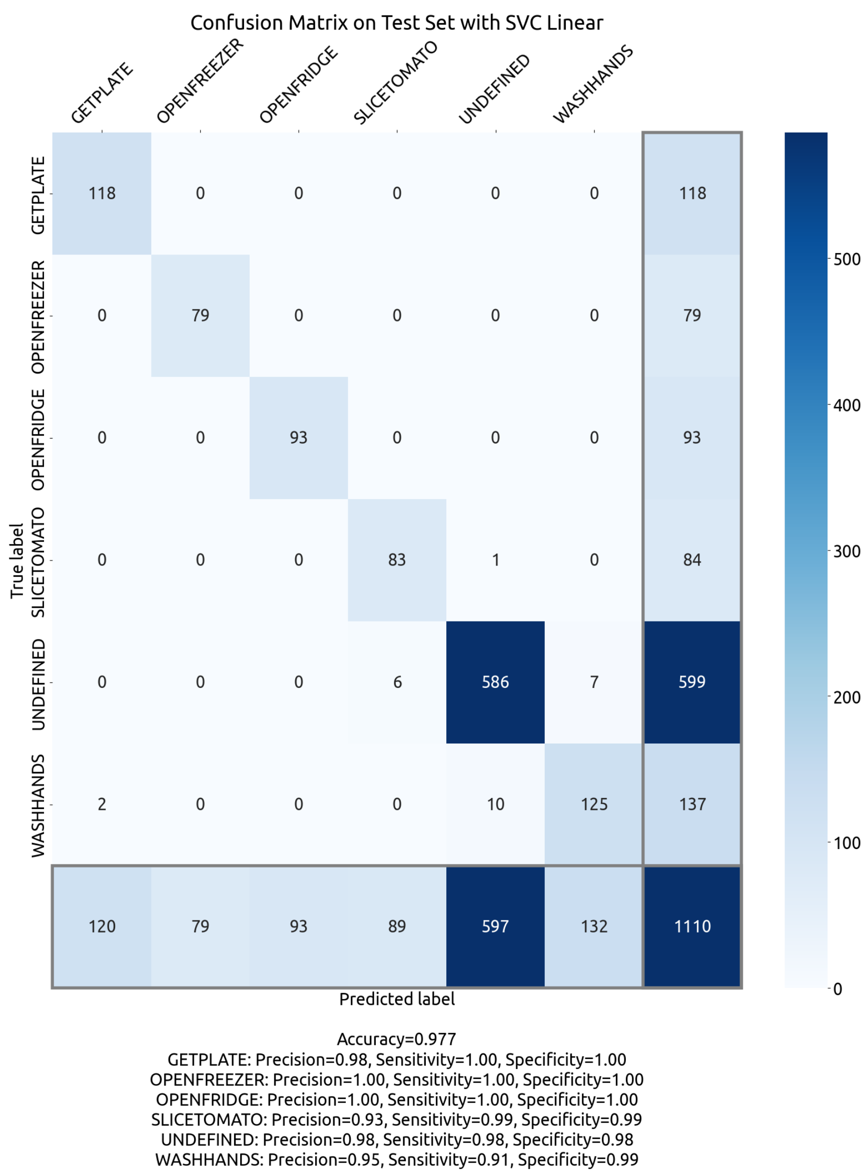



Figure 4.5: Test confusion matrix using the Support Vector Classifier with a Linear Kernel

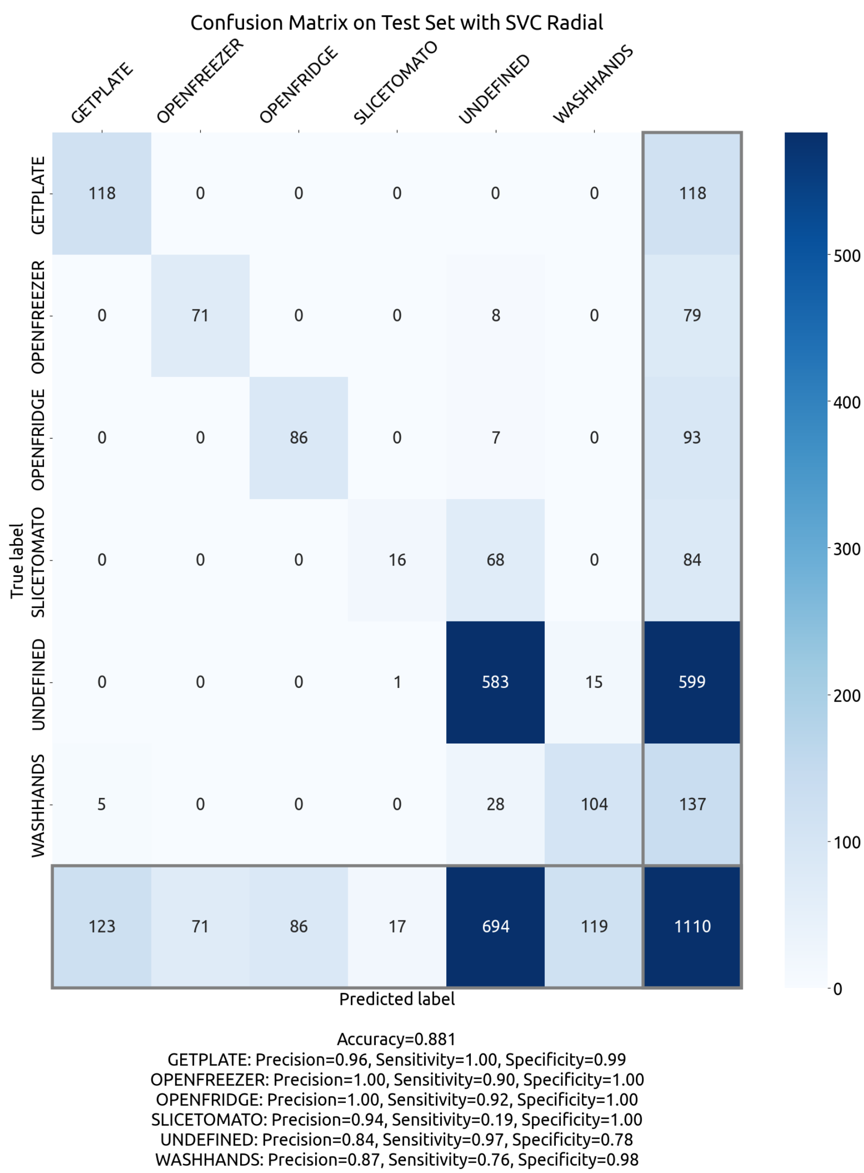



Figure 4.6: Test confusion matrix using the Support Vector Classifier with a Radial Kernel

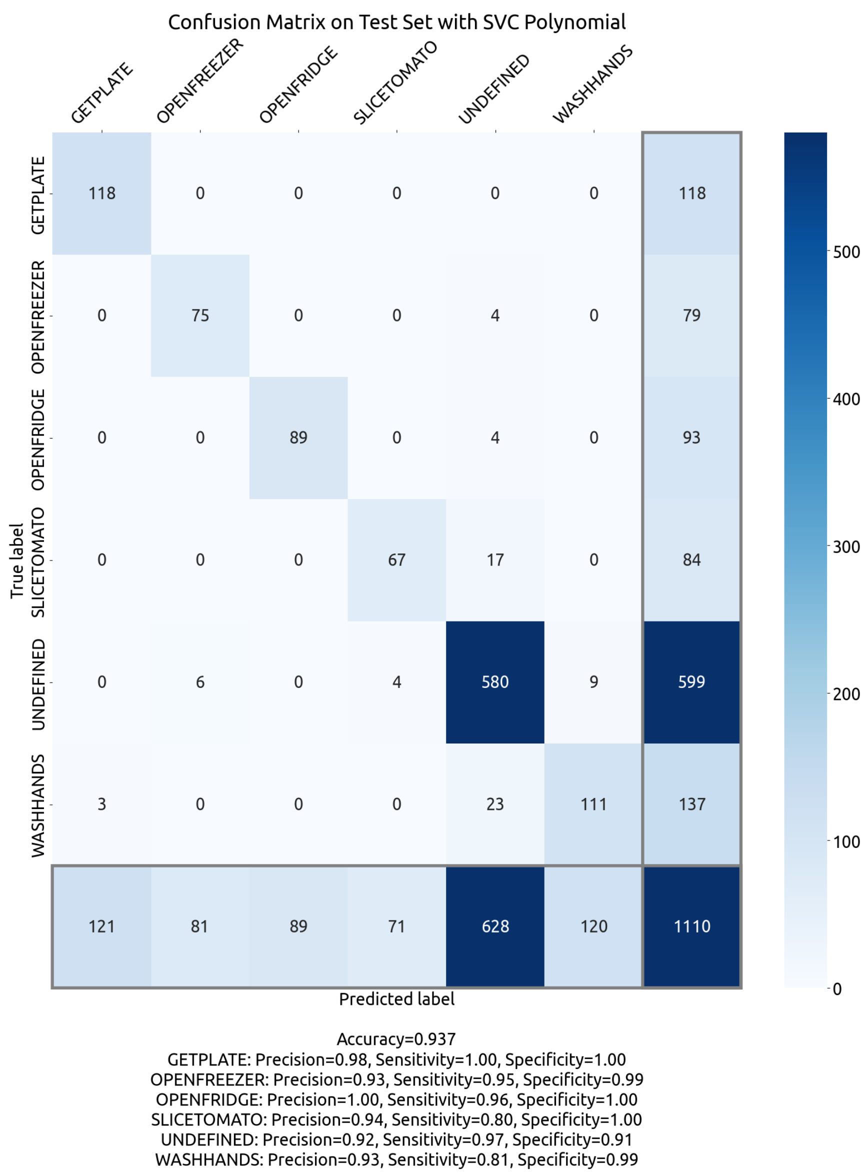



Figure 4.7: Test confusion matrix using the Support Vector Classifier with a Polynomial Kernel

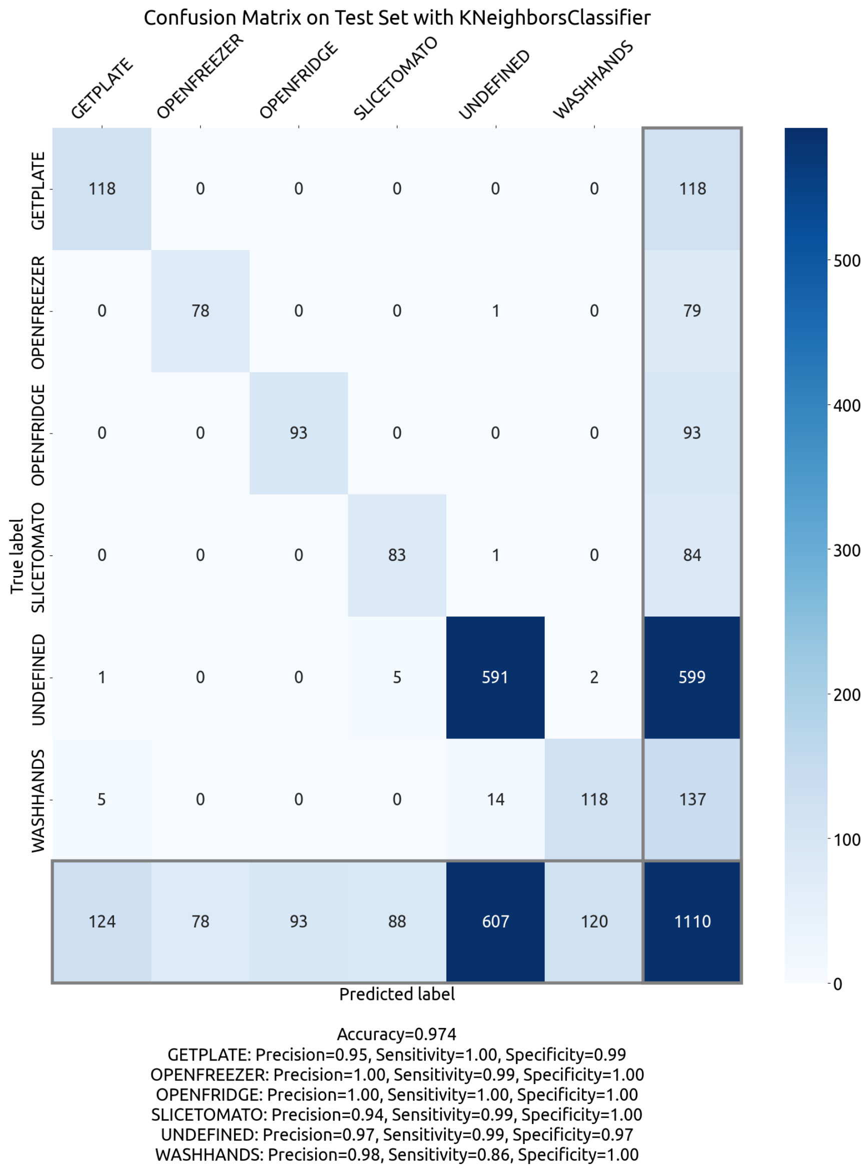



Figure 4.8: Test confusion matrix using the K-Nearest Neighbors Classifier

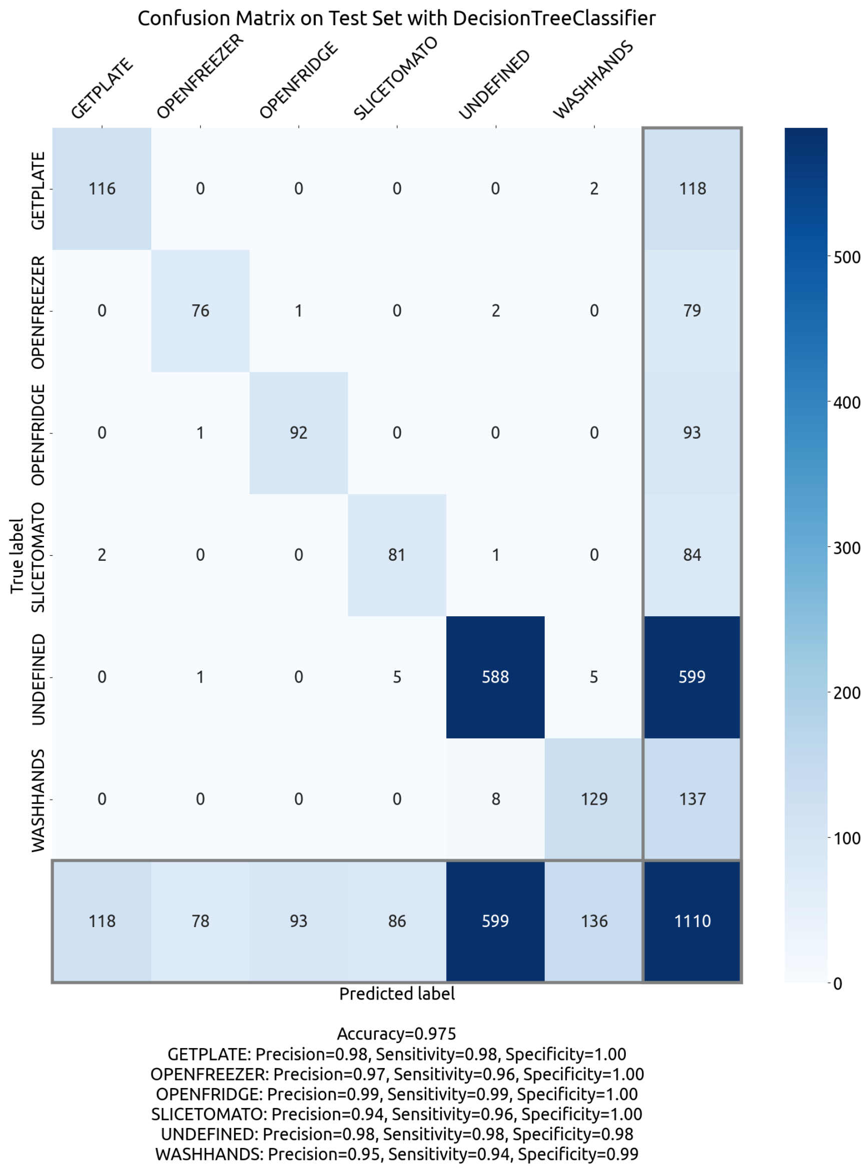



Figure 4.9: Test confusion matrix using the Decision Tree Classifier

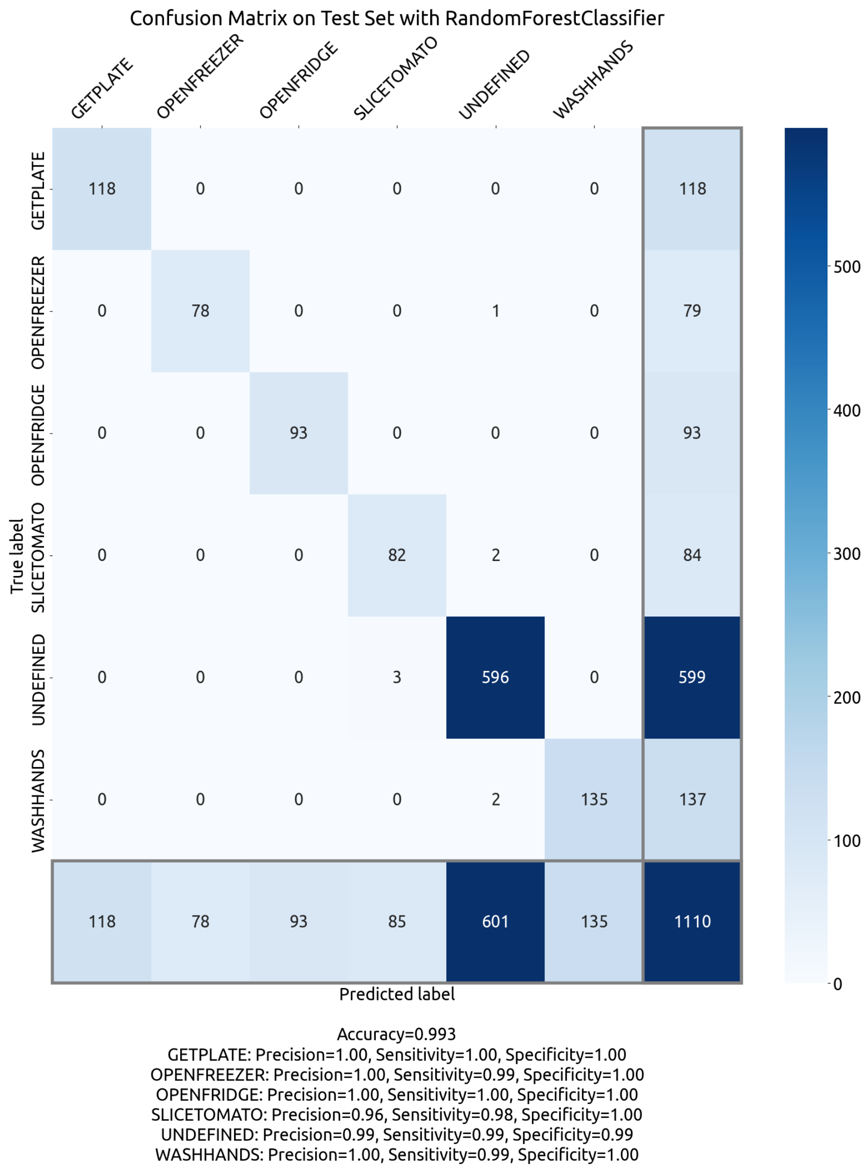



Figure 4.10: Test confusion matrix using the Random Forests Classifier

## 4.5 Discussion

Table 4.2 summarizes the accuracies obtained from each model.

Table 4.2: Accuracy of each Model

| Model Name | Accuracy (%) |
|---|---|
| SVM Linear | 97.7 |
| SVM Radial | 88.1 |
| SVM Polynomial | 93.1 |
| kNN | 97.4 |
| Decision Tree | 97.5 |
| Random Forests | 99.3 |

With the exception of the Radial SVM, all of the models perform well achieving an accuracy of somewhere in the high 90s. In day-to-day activities, there is a disproportionately higher number of the UNDEFINED class compared to the other "action" classes signifying the presence of class imbalance. If a classifier guesses all UNDEFINED it can obtain an accuracy of 599/1110 = 54%. Thus, accuracies taken around 54% should be interpreted with caution. Other metrics such as the Sensitivity, Precision and Specificity have been provided to address this class imbalance. Sensitivity is the rate at which the classifier predicts a $TP$, Precision is the fraction of predictions that are actually true, and Specificity is the rate at which the classifier predicts a $TN$. Of all the models, the Random Forests Classifier at the default settings seem to the best in terms of Accuracy and Precision, Sensitivity, and Specificity for all classes.

The performance of these models in the real-time will need to be tested and quantified before any conclusions can be made. A high accuracy is promising, but may also be indicative of overfitting which means that the model will not be able to generalize variation experienced in the real world. In later sections, more fine-grained actions will be considered, models will be more rigorously tuned, and the performance in real-time will be investigated.

# Bibliography

[1] Configuration of the UWB parameters (Arduino).

[2] Creator One Kit for research and prototyping - Pozyx.

[3] Hardware setup.