Food Intake Detection from Inertial Sensors using LSTM Networks

Kyritsis Konstantinos
Diou Christos
Delopoulos Anastasios

Multimedia Understanding Group
Information Processing Laboratory
Dpt. Electrical and Computer Engineering
Aristotle University of Thessaloniki
Greece

Dietary monitoring

Overweight and obesity are two of the most prevalent preventable causes of death, alongside smoking tobacco and sexually transmitted diseases, and are responsible for over 2.5 million deaths per annum since 2001 [1].

Monitoring and modification of dietary behaviour has shown to be a promising approach for the treatment of obesity and eating disorders [2].

[1] World Health Organization: Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organization (2009)
[2] I. Ioakimidis, M. Zandian, C. Bergh, and P. S¨odersten, "A method for the control of eating rate: a potential intervention in eating disorders," *Behavior research methods*, vol. 41, no. 3, pp. 755–760, 2009

Food intake detection

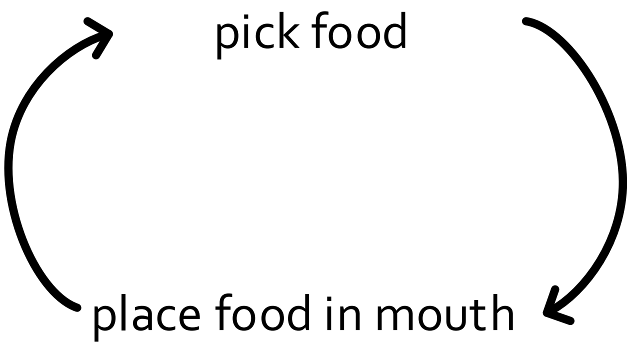
- Multiple ways to approach the problem
  - Inertial sensors
  - Microphones
  - Photoplethysmography
  - Cameras
  - Weight scales
  - Single sensor
  - Multiple same body-mounted sensors
  - Sensor fusion

- Every-day usage
  - Usability/practical restrictions
- Our approach
  - Accelerometer & gyroscope data
  - Off-the-shelf smartwatch
  - Detect intake moments during a meal

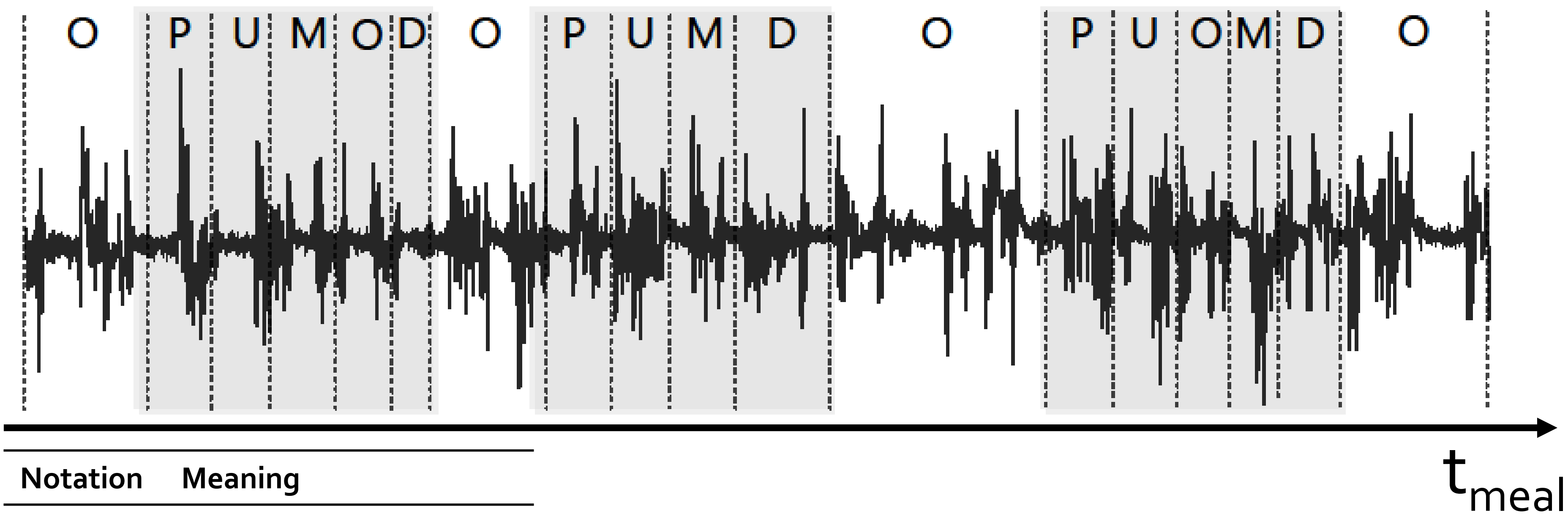
Meals, intake cycles and micro-movements

- **Micro-movement:** An eating-related and limited duration movement of the hand
- **Intake cycle:** A series of micro-movements leading to, and after a food intake
- Modelling of a meal session into intake cycles
- Modelling intake cycles as sequences of specific hand micro-movements

Simple example



Food intake cycles as micro-movement sequences: our approach

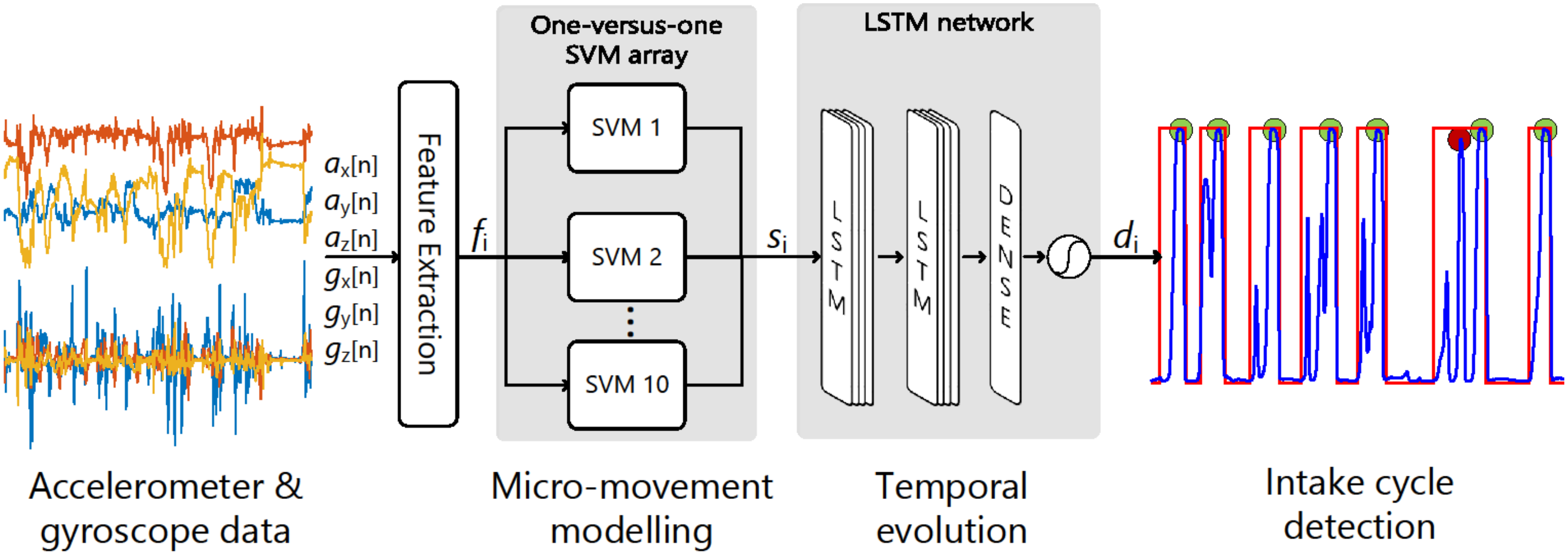


Notation	Meaning
P	Pick Food
U	Upwards
M	Mouth
D	Downwards
N	No hand movement
O	Other hand movement

We **define** the food intake cycle as:

a sequence that starts with **P**, contains at least **M** and ends with **D**

Proposed intake detection approach



Pre-processing and feature extraction

- Remove gravitational effect from acceleration
- 5th order median filter for each accelerometer and gyroscope channel

- Sliding window approach
  - length w_l = 0.2 seconds
  - step w_s = 0.1 seconds

- Extract feature vector $\boldsymbol{f}$ in each window
  1. # of zero-crossings
  2. mean
  3. std
  4. min/max values
  5. range of values
  6. normalized energy
  7. the first $\left\lfloor \frac{w_l}{2} + 1 \right\rfloor$ DFT coefficients
  8. simple moving average (SMA)

(1-7 for every $a_x, a_y, a_z, g_x, g_y, g_z$)
(8 individually for the accelerometer and gyroscope)

Learning the micro-movements

- Train multiclass SVM with RBF kernel
  - features $\vec{\boldsymbol{f}}$, labels $\boldsymbol{l} \in$ {pick food, upwards, mouth, downwards, no movement}
- Features from “other movement” micro-movement not included in the training process
  - high inner class variance
- Transform each $\vec{\boldsymbol{f}}$ into $\vec{\boldsymbol{s}}$
  - $\vec{\boldsymbol{s}} \in \boldsymbol{R}^{10}$
  - contains the pairwise prediction scores of the 10 one-versus-one SVMs

Modelling time evolution with LSTM

- Why LSTM?
  - solves the long term dependency and vanishing gradient problems
  - retain information over a long period
    - ❖ efficiently model food intake sequences that greatly differ from the ideal intake sequence

- Network architecture
  1. Two consecutive LSTM layers with 128 hidden cells each
     - hard sigmoid for reccurent steps
  2. One fully connected output layer
     - sigmoid function

- Compact notation: LSTM(128)-LSTM(128)-D(1)

Modelling time evolution with LSTM (cont.)

- Train network with sequences of $\vec{\boldsymbol{s}}$ and labels $\boldsymbol{l}' \in \{0,1\}$
  - $\boldsymbol{l}' = 1$
    - ❖sequence is an intake cycle (starts with P, contains at least one M and ends with D)
  - $\boldsymbol{l}' = 0$
    - ❖sequence isn't an intake cycle (typicaly found between consecutive intake cycles)
- Design choises
  - minimize: Binary cross entropy loss
  - RMSprop Optimizer with Learning rate, $lr = 10^{-3}$
  - batch size, $M = 32$ for 5 epochs
- Output
  - For each sequence of $\vec{\boldsymbol{s}}$, the network outputs the propability $\boldsymbol{d}$ that the given sequence is an intake cycle

Detection

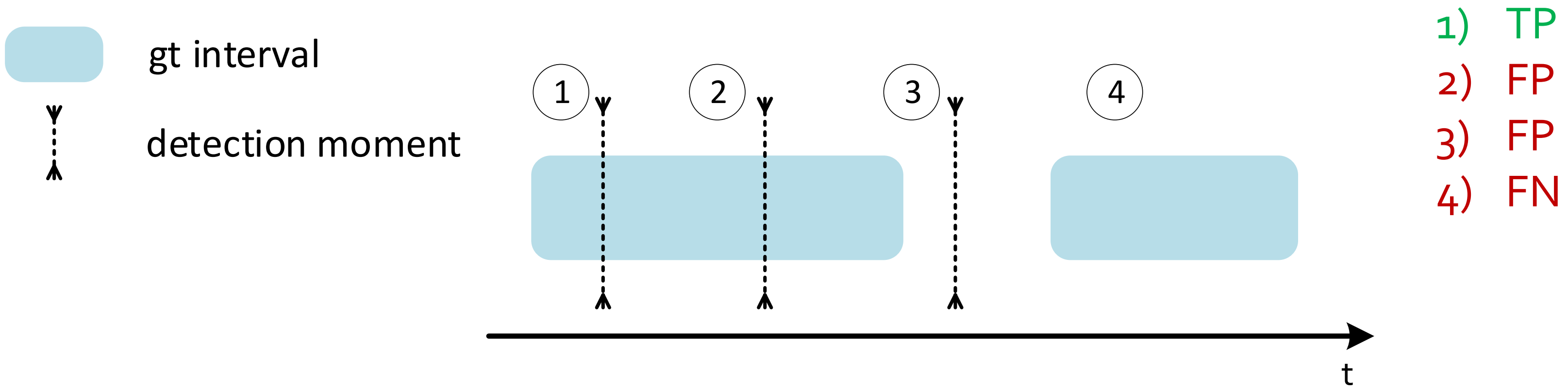
- Given the:
  - sequence of $\vec{\boldsymbol{s}}$ representing a complete meal
  - trained LSTM model
- Select frames with sliding window
  - length w'_l = 3 seconds
  - step w'_s = 0.2 seconds
- Network outputs a propability $\boldsymbol{d}$ for each windowed data frame
  - form 1-D signal $d[n]$ for the complete meal session
- Intake moments are detected by local maxima search in $d[n]$
  - minimum distance of 3 seconds between peaks

Evaluation Dataset

- The **F**ood **I**ntake **C**ycle (FIC) Dataset
- Recordings of 10 unique subjects performing unscripted eating activities
- Accelerometer and gyroscope recordings at 62 Hz
  - sensor: Microsoft Band 2
- A Go Pro Hero 5 recorded the activites
  - used to generate the ground truth at micro-movement level
- Publicly available at https://mug.ee.auth.gr/intake-cycle-detection/

Evaluation methods

- Leave One Subject Out (LOSO) cross validation scheme
- Repeat the network training for 10 times
  - smooth results produced by the stochastic training nature of the network
- Calculate TP, FP and FN



Results

- Compare performance with 2 approaches found in the literature [3][4]
  - data from FIC
  - evaluation scheme mentioned earlier

Method	TP	FP	FN	Prec	Rec	F1
Proposed	623.7	89	60.3	0.875	0.911	0.892
Dong et al. [3]	508	683	176	0.426	0.742	0.541
Kyritsis el al. [4]	603	193	81	0.757	0.881	0.814

[3] Dong Yujie, Hoover Adam, Scisco Jenna and Muth Eric. A new method for measuring meal intake in humans via automated wrist motion tracking. *Applied psychophysiology and biofeedback*, 2012
[4] Kyritsis Konstantinos, Tatli Christina Lefkothea, Diou Christos and Delopoulos Anastasios. Automated Analysis of in Meal Eating Behavior using a Commercial Wristband IMU Sensor. *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017

Conclusions

- Proposed a novel approach for intake cyce detection during the course of a meal
- Explicit modeling of intake cycles as sequences of micro-movements outperforms direct detection approaches
- Combination of micro-movement SVMs and LSTM network for score sequence classification is highly effective

Thank you

kokirits@mug.ee.auth.gr

https://mug.ee.auth.gr