Light Residual Network for Human Activity Recognition using Wearable Sensor Data

Francisco M. Calatrava Nicolás (francisco.calatrava-nicolas@oru.se)
Oscar Martinez Mozos
Örebro University, AI for Life

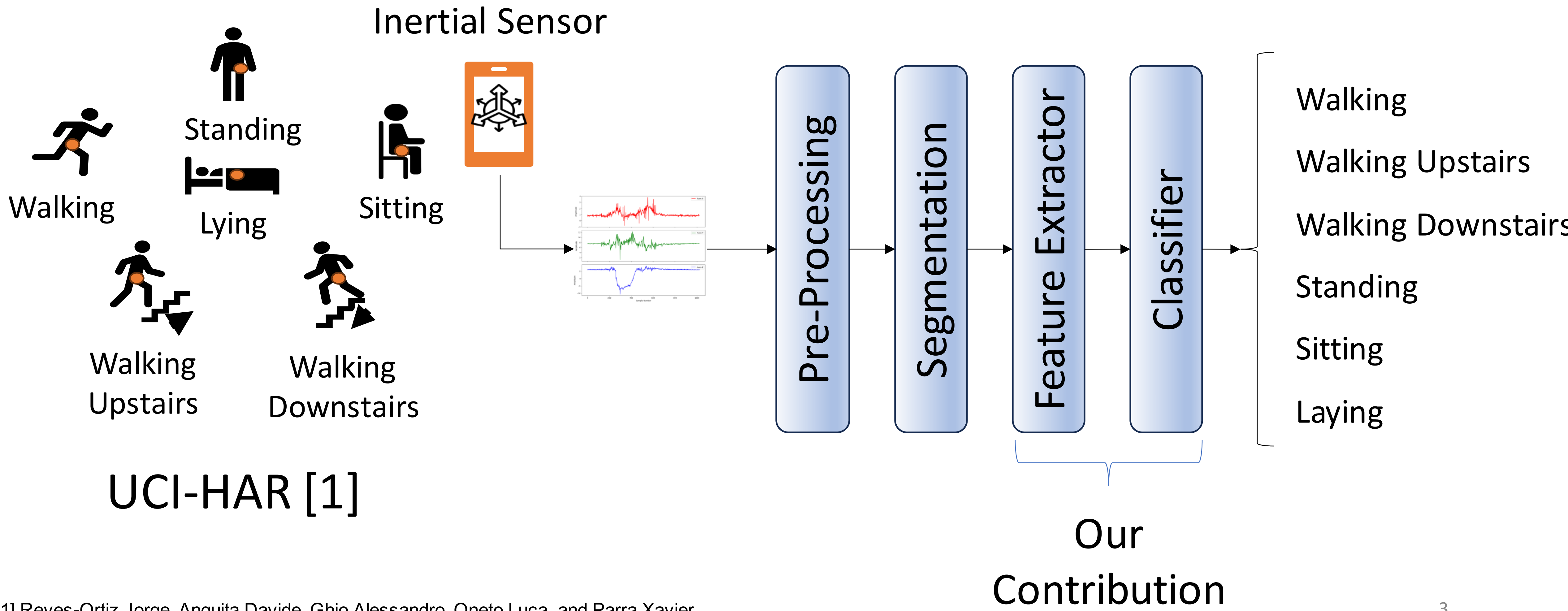
ÖREBRO UNIVERSITY
WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

Human Activity Recognition (HAR)

- It is about classifying and understanding human activities
- This could be used in many fields: health, sports, security, human robot interaction…
- Example:
    - Wellbeing monitoring of elderly at home.

Human Activity Recognition Problem

Classify a **set of human activities** using **inertial sensor data**



Inertial Sensor

Walking

Standing

Lying

Sitting

Walking Upstairs

Walking Downstairs

UCI-HAR [1]

Pre-Processing

Segmentation

Feature Extractor

Classifier

Walking

Walking Upstairs

Walking Downstairs

Standing

Sitting

Laying

Our Contribution

[1] Reyes-Ortiz,Jorge, Anguita,Davide, Ghio,Alessandro, Oneto,Luca, and Parra,Xavier. (2012). Human Activity Recognition Using Smartphones. https://doi.org/10.24432/C54S4K.

Previous Works: Gaps

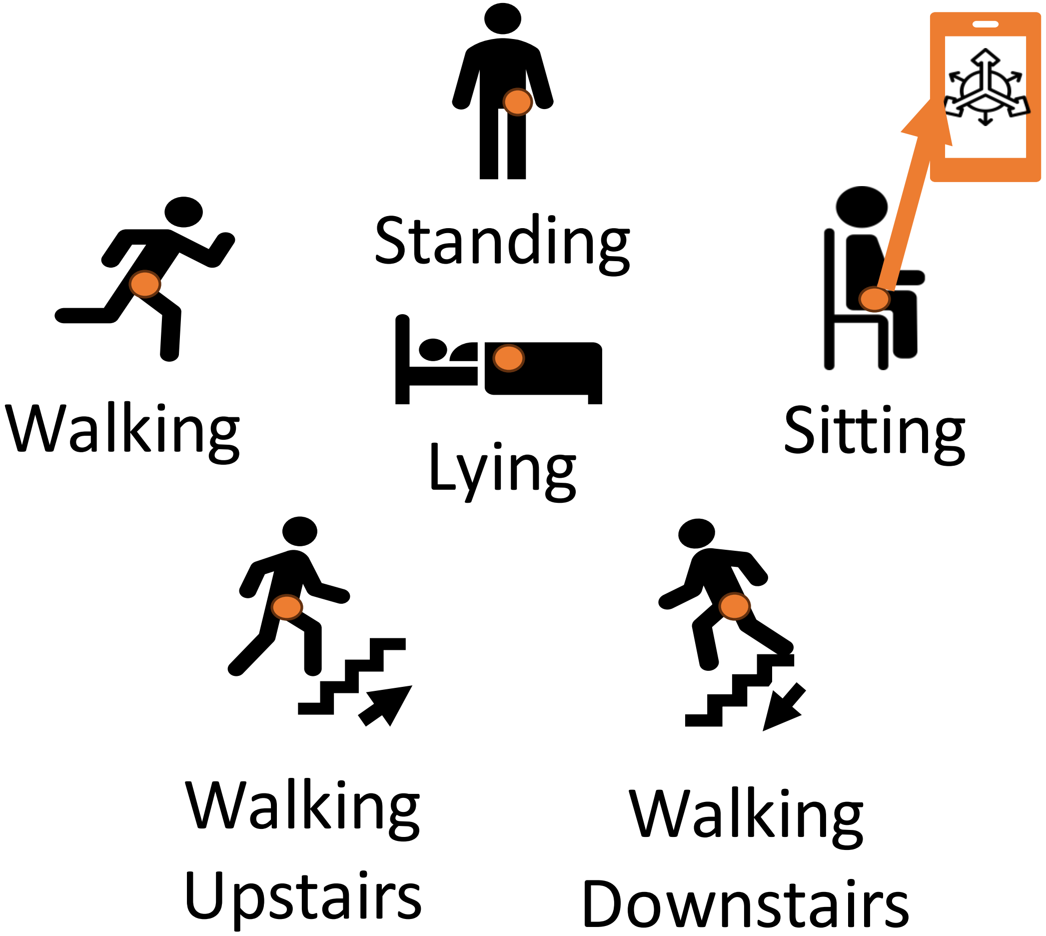
- **Classification:** Deep Learning surpass traditional machine learning methods in HAR but:
  - Increase the number of parameters[19][21][27]
  - Increase the computational cost
- **Comparisons**: Observed a benchmark gap for the UCI-HAR dataset.
  - State-of-the-art training and testing conditions differ greatly.
  - It's challenging to make a fair comparison.

Our Contributions

A new light Deep Learning architecture

Unified benchmark for the UCI-HAR dataset

Dataset: UCI-HAR



30 Participants

6 Activities

Smartphone’s IMU

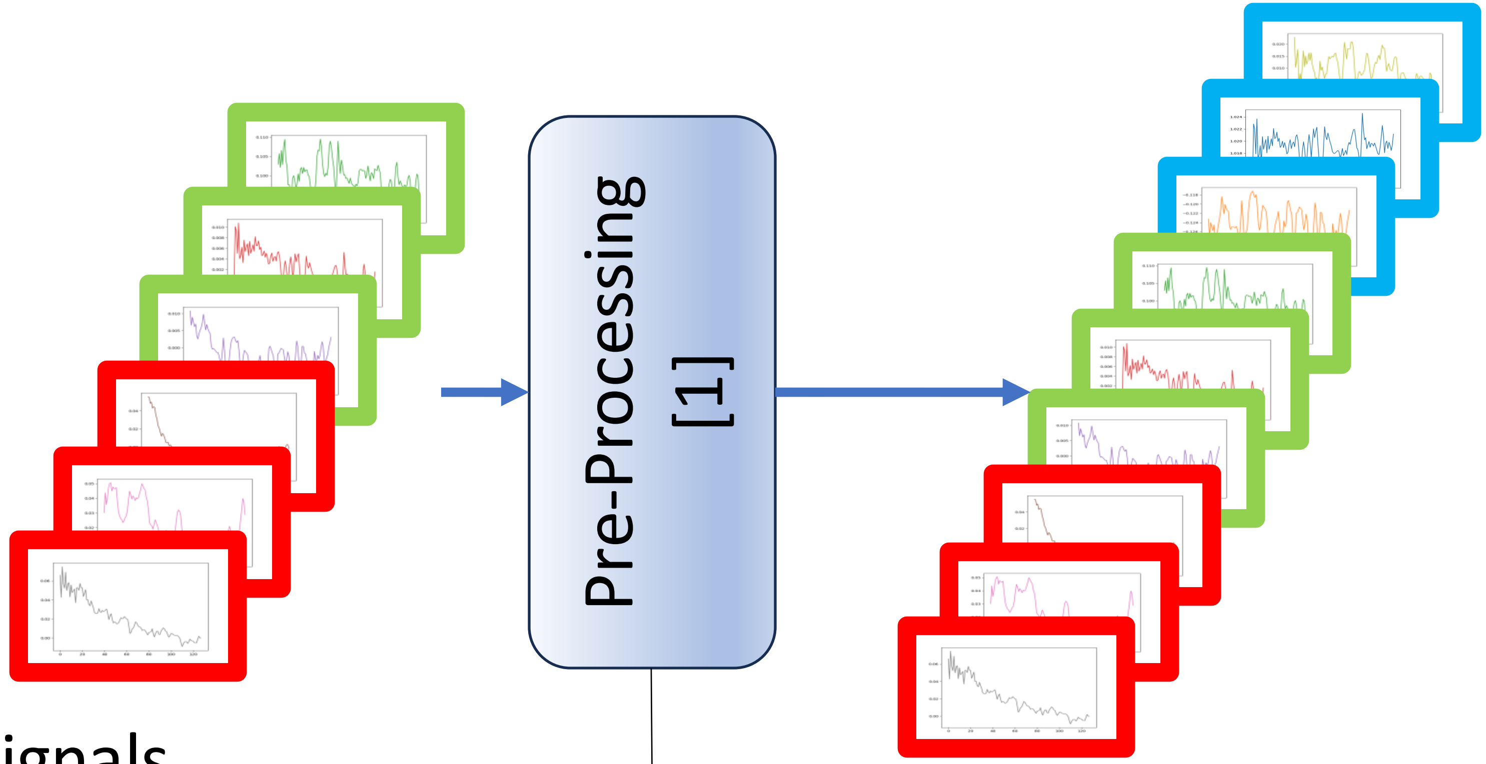
Controlled lab environment

[1] Reyes-Ortiz,Jorge, Anguita,Davide, Ghio,Alessandro, Oneto,Luca, and Parra,Xavier. Human Activity Recognition using smartphones (2012). https://doi.org/10.24432/C54S4K.

Input data: Preprocessing [1]

Vel. Z
Vel. Y
Vel. X
Acc. Z
Acc. Y
Acc. X

6 Inertial signals



Pre-Processing [1]

- Clean Noise : Butterworth and Median Filters
- Body Estimated ACC: Butterworth Filter

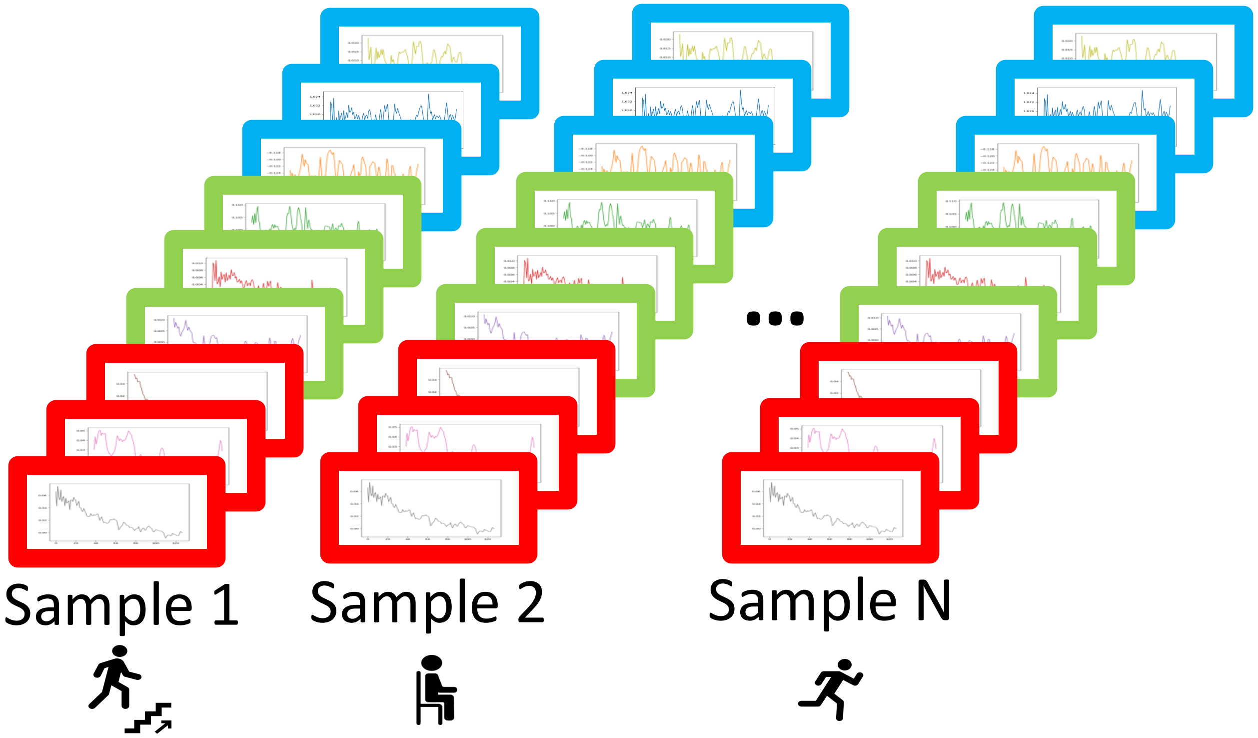
Body Estimated Acc X
Body Estimated Acc Y
Body Estimated Acc Z
Vel. Z
Vel. Y
Vel. X
Acc. Z
Acc. Y
Acc. X

9 Inertial signals

[1] Reyes-Ortiz,Jorge, Anguita,Davide, Ghio,Alessandro, Oneto,Luca, and Parra,Xavier. (2012). Human Activity Recognition Using Smartphones. https://doi.org/10.24432/C54S4K.

Fixed-Width Sliding Window [1]

Fixed-Width Sliding Window



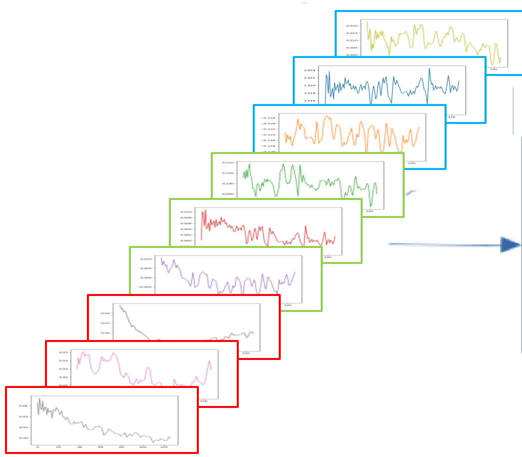
- Windows Size: 128 points (2.56 s)
- Overlapping: 50%
- A unique activity is mapped to every sample

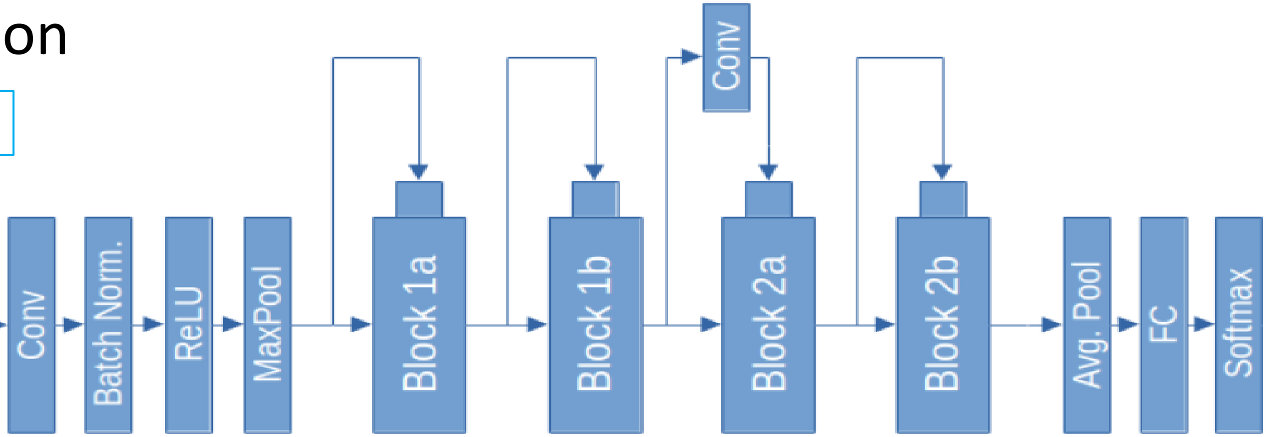
---

[1] Reyes-Ortiz,Jorge, Anguita,Davide, Ghio,Alessandro, Oneto,Luca, and Parra,Xavier. (2012). Human Activity Recognition Using Smartphones. https://doi.org/10.24432/C54S4K.

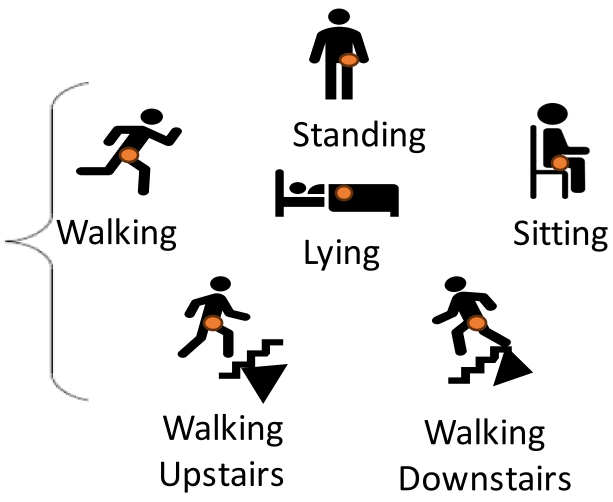
Our Approach: Light Architecture

Inertial Signals after Preprocessing & Segmentation & Normalization

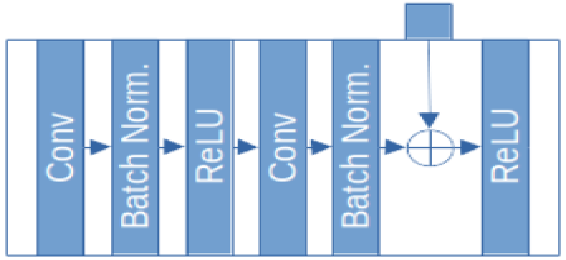




Labels



Single Block Architecture:



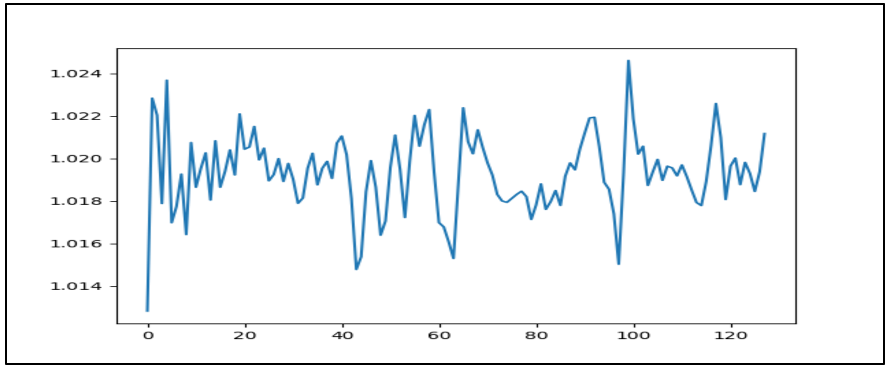
Input Shape: (9, 1, 128)

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, arXiv:1512.03385

Our Approach: Light Architecture

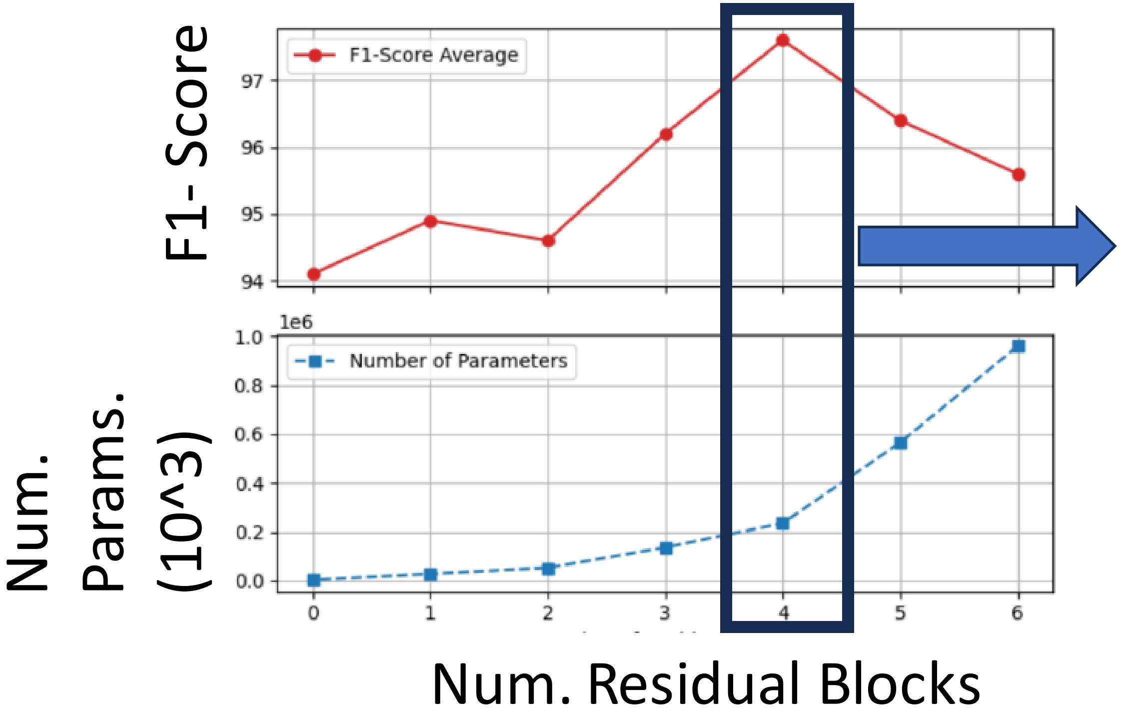
Inspired in ResNet 18 [1]. We have modified:

1-D Kernels



Example 1D inertial signal

Only 4 Residua Blocks



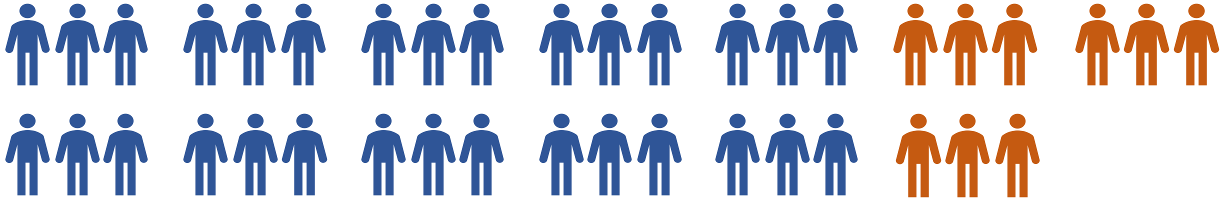
4 Residual Blocks

As a result, we reduced to 234 950 Parameters

(ResNet18 is around 11 million)

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015, arXiv:1512.03385

UCI-HAR Original Splitting (70%-30%) [1]



🟦 Train – 70% of Participants (21)

🟧 Test – 30% of Participants (9)

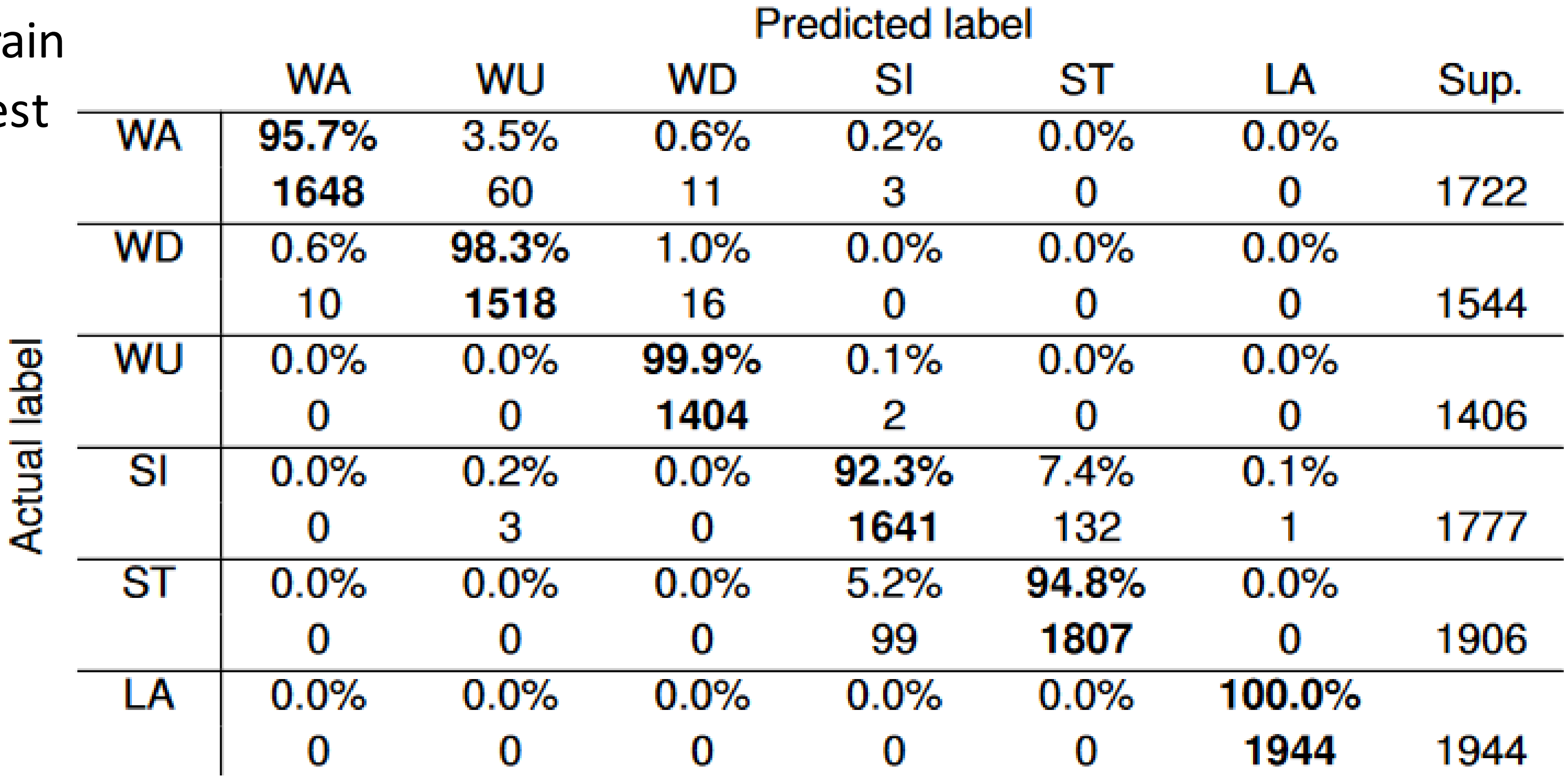
Approaches	F1-Score (%)	Accuracy (%)	Params.
GRU [19] (2023)	89.2	89.2	-
CNN-LSTM Self-Att. [26] (2022)	90.9	93.1	634,188
Res-BiLSTM [22] (2018)	91.5	91.6	-
CNN-LSTM [23] (2020)	-	92.1	-
Bi-LSTM [21] (2019)	92.7	92.7	-
Stacked LSTM [24] (2019)	93.1	93.1	-
CNN [25] (2019)	93.5	-	-
iSPLInception [27](2021)	95.0	95.1	1,327,754
GRU+Attention [19] (2023)	95.8	96.0	1,600,000
CNN-DCT [20] (2023)	97.1	-	930,000
ResNet-DCT [20] (2023)	97.5	-	3,540,000
Our model	**97.6**	**97.6**	**234,950**

[1] Reyes-Ortiz,Jorge, Anguita,Davide, Ghio,Alessandro, Oneto,Luca, and Parra,Xavier. (2012). Human Activity Recognition Using Smartphones. https://doi.org/10.24432/C54S4K.

Leave-One-(Person)-Out Cross-Validation

🟦 Train

🟧 Test



Predicted label

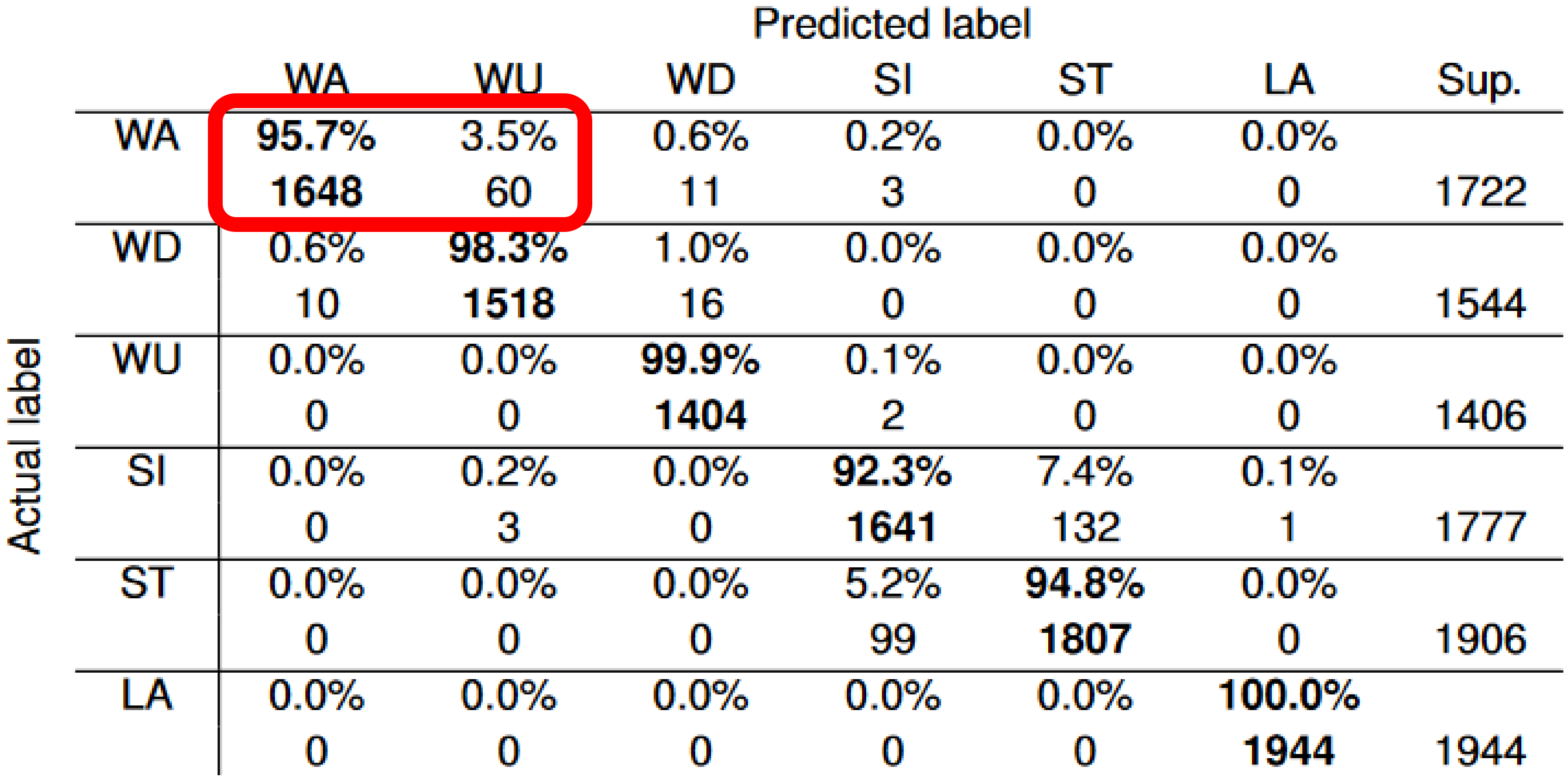
Actual label	WA	WU	WD	SI	ST	LA	Sup.
WA	**95.7%** **1648**	3.5% 60	0.6% 11	0.2% 3	0.0% 0	0.0% 0	1722
WD	0.6% 10	**98.3%** **1518**	1.0% 16	0.0% 0	0.0% 0	0.0% 0	1544
WU	0.0% 0	0.0% 0	**99.9%** **1404**	0.1% 2	0.0% 0	0.0% 0	1406
SI	0.0% 0	0.2% 3	0.0% 0	**92.3%** **1641**	7.4% 132	0.1% 1	1777
ST	0.0% 0	0.0% 0	0.0% 0	5.2% 99	**94.8%** **1807**	0.0% 0	1906
LA	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **1944**	1944

WA – Walking
WU – Walking Upstairs
WD – Walking Downstairs
SI – Sitting
ST – Standing
LA – Laying

Leave-One-(Person)-Out Cross-Validation

🟦 Train

🟧 Test



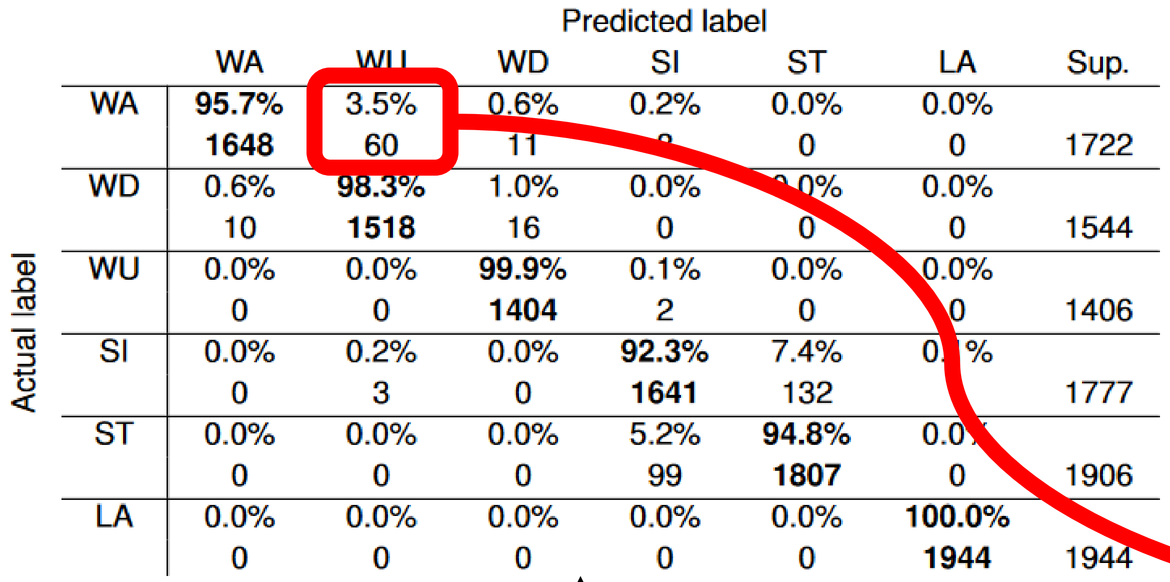
Predicted label

Actual label

	WA	WU	WD	SI	ST	LA	Sup.
WA	**95.7%** **1648**	3.5% 60	0.6% 11	0.2% 3	0.0% 0	0.0% 0	1722
WD	0.6% 10	**98.3%** **1518**	1.0% 16	0.0% 0	0.0% 0	0.0% 0	1544
WU	0.0% 0	0.0% 0	**99.9%** **1404**	0.1% 2	0.0% 0	0.0% 0	1406
SI	0.0% 0	0.2% 3	0.0% 0	**92.3%** **1641**	7.4% 132	0.1% 1	1777
ST	0.0% 0	0.0% 0	0.0% 0	5.2% 99	**94.8%** **1807**	0.0% 0	1906
LA	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **1944**	1944

- WA – Walking
- WU – Walking Upstairs
- WD – Walking Downstairs
- SI – Sitting
- ST – Standing
- LA – Laying

Leave-One-(Person)-Out Cross-Validation

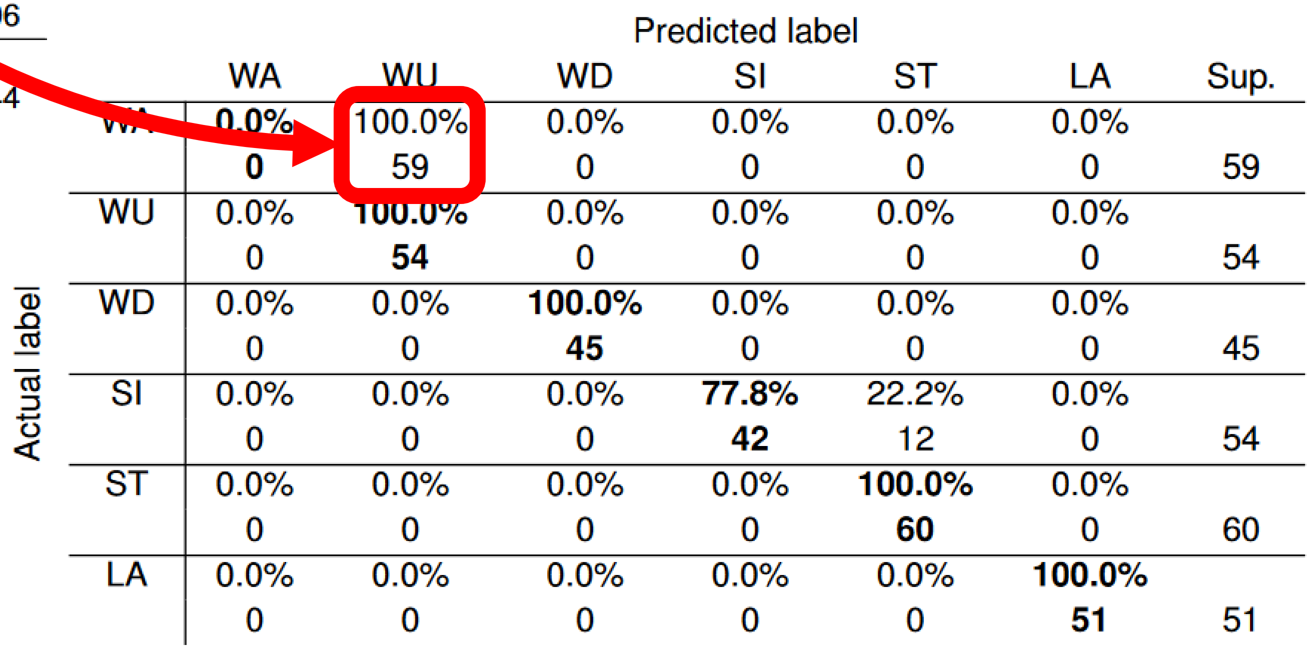


Predicted label

Actual label	WA	WU	WD	SI	ST	LA	Sup.
WA	**95.7%** **1648**	3.5% 60	0.6% 11	0.2% [illegible]	0.0% 0	0.0% 0	1722
WD	0.6% 10	**98.3%** **1518**	1.0% 16	0.0% 0	0.0% 0	0.0% 0	1544
WU	0.0% 0	0.0% 0	**99.9%** **1404**	0.1% 2	0.0% 0	0.0% 0	1406
SI	0.0% 0	0.2% 3	0.0% 0	**92.3%** **1641**	7.4% 132	[illegible]	1777
ST	0.0% 0	0.0% 0	0.0% 0	5.2% 99	**94.8%** **1807**	0.0% 0	1906
LA	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **1944**	1944

Global Confusion Matrix from Leave-One(Person)-Out Cross-Validation

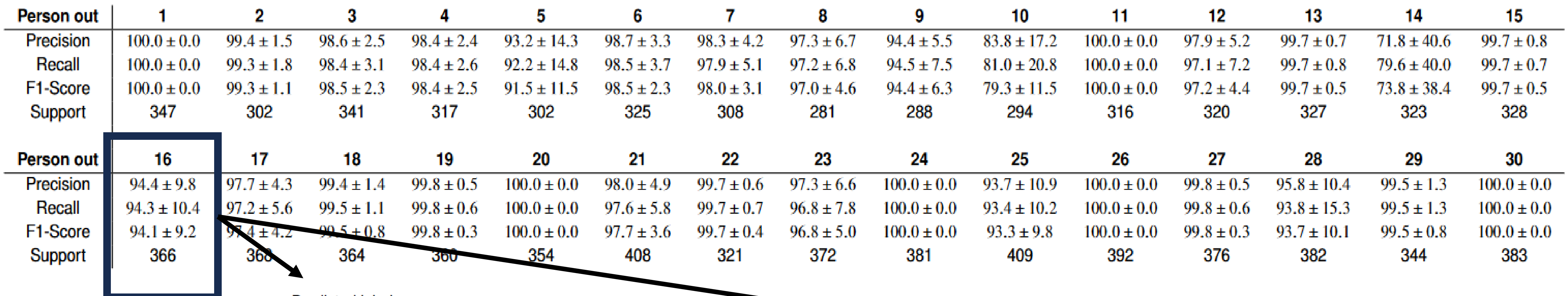
Confusion Matrix Participant 14



Predicted label

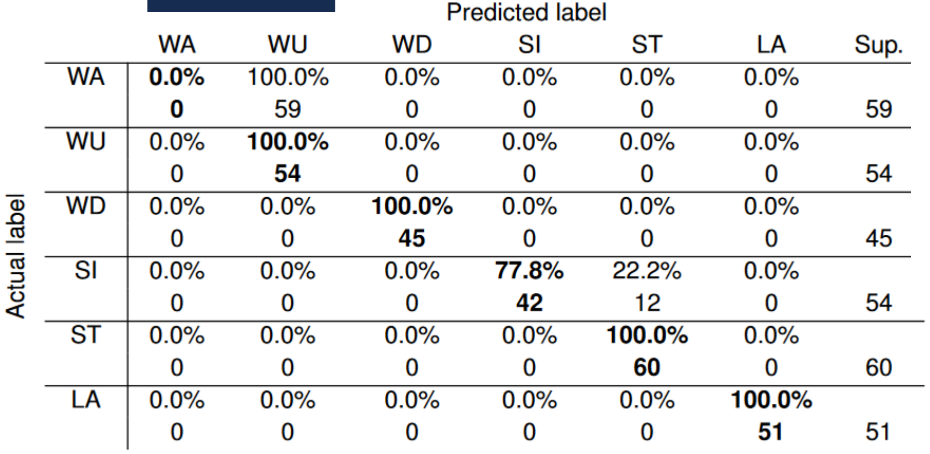
Actual label	WA	WU	WD	SI	ST	LA	Sup.
WA	**0.0%** **0**	100.0% 59	0.0% 0	0.0% 0	0.0% 0	0.0% 0	59
WU	0.0% 0	**100.0%** **54**	0.0% 0	0.0% 0	0.0% 0	0.0% 0	54
WD	0.0% 0	0.0% 0	**100.0%** **45**	0.0% 0	0.0% 0	0.0% 0	45
SI	0.0% 0	0.0% 0	0.0% 0	**77.8%** **42**	22.2% 12	0.0% 0	54
ST	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **60**	0.0% 0	60
LA	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **51**	51

Additional Results from our Benchmark



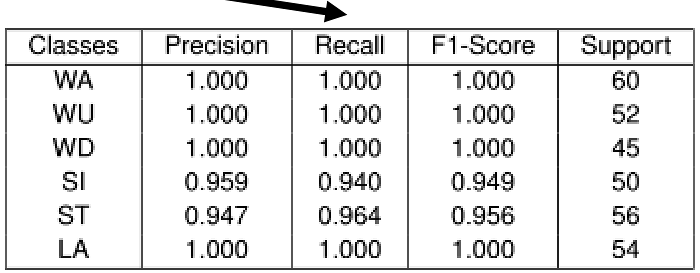
Person out	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Precision	100.0 ± 0.0	99.4 ± 1.5	98.6 ± 2.5	98.4 ± 2.4	93.2 ± 14.3	98.7 ± 3.3	98.3 ± 4.2	97.3 ± 6.7	94.4 ± 5.5	83.8 ± 17.2	100.0 ± 0.0	97.9 ± 5.2	99.7 ± 0.7	71.8 ± 40.6	99.7 ± 0.8
Recall	100.0 ± 0.0	99.3 ± 1.8	98.4 ± 3.1	98.4 ± 2.6	92.2 ± 14.8	98.5 ± 3.7	97.9 ± 5.1	97.2 ± 6.8	94.5 ± 7.5	81.0 ± 20.8	100.0 ± 0.0	97.1 ± 7.2	99.7 ± 0.8	79.6 ± 40.0	99.7 ± 0.7
F1-Score	100.0 ± 0.0	99.3 ± 1.1	98.5 ± 2.3	98.4 ± 2.5	91.5 ± 11.5	98.5 ± 2.3	98.0 ± 3.1	97.0 ± 4.6	94.4 ± 6.3	79.3 ± 11.5	100.0 ± 0.0	97.2 ± 4.4	99.7 ± 0.5	73.8 ± 38.4	99.7 ± 0.5
Support	347	302	341	317	302	325	308	281	288	294	316	320	327	323	328

Person out	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Precision	94.4 ± 9.8	97.7 ± 4.3	99.4 ± 1.4	99.8 ± 0.5	100.0 ± 0.0	98.0 ± 4.9	99.7 ± 0.6	97.3 ± 6.6	100.0 ± 0.0	93.7 ± 10.9	100.0 ± 0.0	99.8 ± 0.5	95.8 ± 10.4	99.5 ± 1.3	100.0 ± 0.0
Recall	94.3 ± 10.4	97.2 ± 5.6	99.5 ± 1.1	99.8 ± 0.6	100.0 ± 0.0	97.6 ± 5.8	99.7 ± 0.7	96.8 ± 7.8	100.0 ± 0.0	93.4 ± 10.2	100.0 ± 0.0	99.8 ± 0.6	93.8 ± 15.3	99.5 ± 1.3	100.0 ± 0.0
F1-Score	94.1 ± 9.2	97.4 ± 4.2	99.5 ± 0.8	99.8 ± 0.3	100.0 ± 0.0	97.7 ± 3.6	99.7 ± 0.4	96.8 ± 5.0	100.0 ± 0.0	93.3 ± 9.8	100.0 ± 0.0	99.8 ± 0.3	93.7 ± 10.1	99.5 ± 0.8	100.0 ± 0.0
Support	366	368	364	360	354	408	321	372	381	409	392	376	382	344	383



Predicted label

Actual label	WA	WU	WD	SI	ST	LA	Sup.
WA	**0.0%** **0**	100.0% 59	0.0% 0	0.0% 0	0.0% 0	0.0% 0	59
WU	0.0% 0	**100.0%** **54**	0.0% 0	0.0% 0	0.0% 0	0.0% 0	54
WD	0.0% 0	0.0% 0	**100.0%** **45**	0.0% 0	0.0% 0	0.0% 0	45
SI	0.0% 0	0.0% 0	0.0% 0	**77.8%** **42**	22.2% 12	0.0% 0	54
ST	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **60**	0.0% 0	60
LA	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	**100.0%** **51**	51



Classes	Precision	Recall	F1-Score	Support
WA	1.000	1.000	1.000	60
WU	1.000	1.000	1.000	52
WD	1.000	1.000	1.000	45
SI	0.959	0.940	0.949	50
ST	0.947	0.964	0.956	56
LA	1.000	1.000	1.000	54

- LOOCV Confusion Matrices
- LOOCV Metrics
- 10-fold cross validation
- Ablation Study (Residual Blocks)

ÖREBRO UNIVERSITY

WASP
WALLENBERG AI,
AUTONOMOUS SYSTEMS
AND SOFTWARE PROGRAM

Thank you for your attention

Contact Details:
Name: Francisco Miguel Calatrava Nicolás
Email: francisco.calatrava-nicolas@oru.se