MACHINE LEARNING FOR WEARABLE COMPUTING.

A CHANCE FOR DEEP LEARNING?

DR. ROBERT DÜRICHEN
(ROBERT.DUERICHEN@DE.BOSCH.COM)

BOSCH CORPORATE RESEARCH
USER TECHNOLOGIES

BOSCH

Machine Learning for Wearable Computing
Bosch

- Bosch is one of the **world's leading providers** of innovative MEMS sensors and actuator solutions tailored for smartphones, tablets, wearable devices and IoT applications.
- Start of MEMS production in 1995 - over 9.5 billion MEMS sensors produced
- 100% in-house from MEMS design to manufacturing

Inertial

Accelerometer
Detects linear motion and gravitational forces

Gyroscope
Measures the rate of rotation in space (roll, pitch, yaw)

Magnetometer
Measures the terrestrial earth's magnetic fields

Environmental

Pressure
Measures barometric pressure and altitude

Temperature
Measures ambient temperature

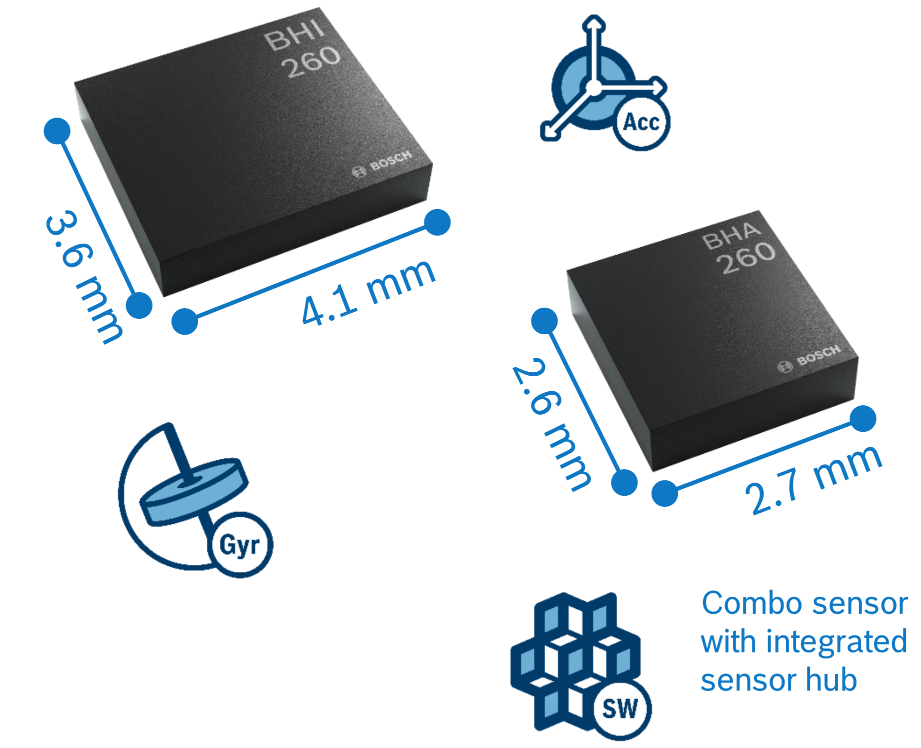
Relative Humidity
Measures relative humidity with a fast response time

Gas
Measures Volatile Organic Compounds (VOC)

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing

Combo sensors with embedded intelligence



Integrated sensor hubs BHI260 and BHA260

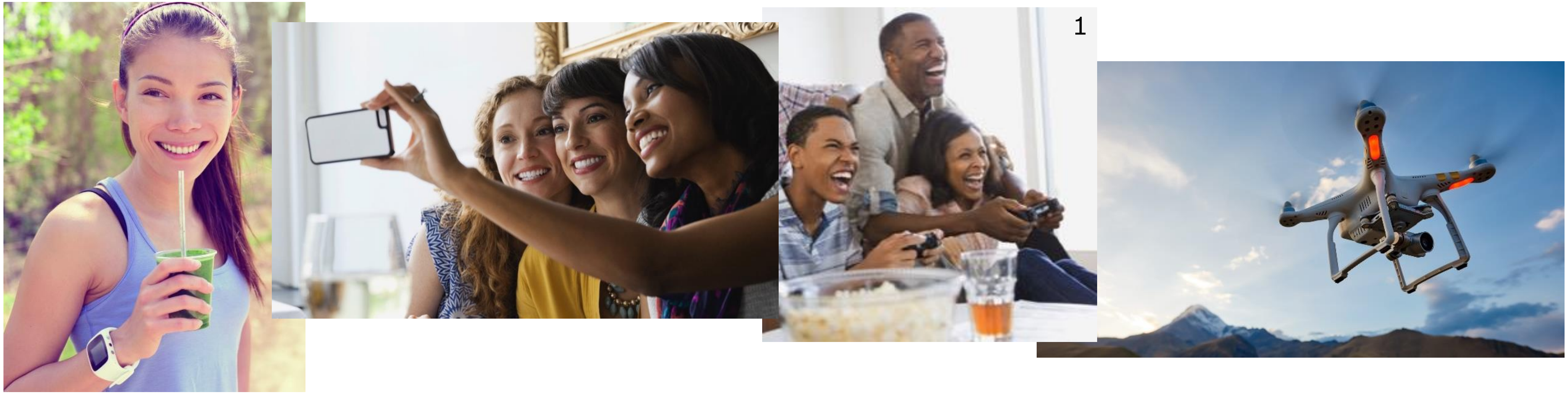
Smart Hub solutions combine Bosch Sensortec's…

- 16 bit MEMS sensors with an ultra-low power, high performance coprocessor
- Best-in-class sensor data fusion software
- Application specific software:
  - Orientation and gesture recognition (for VR controllers),
  - Activity recognition (for fitness applications),
  - Pedestrian dead recognition (for navigation),
  - Context awareness

Idealy suited for demanding always-on sensor applications without compromising features or performance

BOSCH

Machine Learning for Wearable Computing
Applications


[1]

- Up-coming („medical“) applications :
  - Wearable based affect recognition (e.g. mood and stress recognition[2])
  - Eldery care (e.g. fall detection)
  - Intelligent sensor patches[3, 4]

© VivaInk

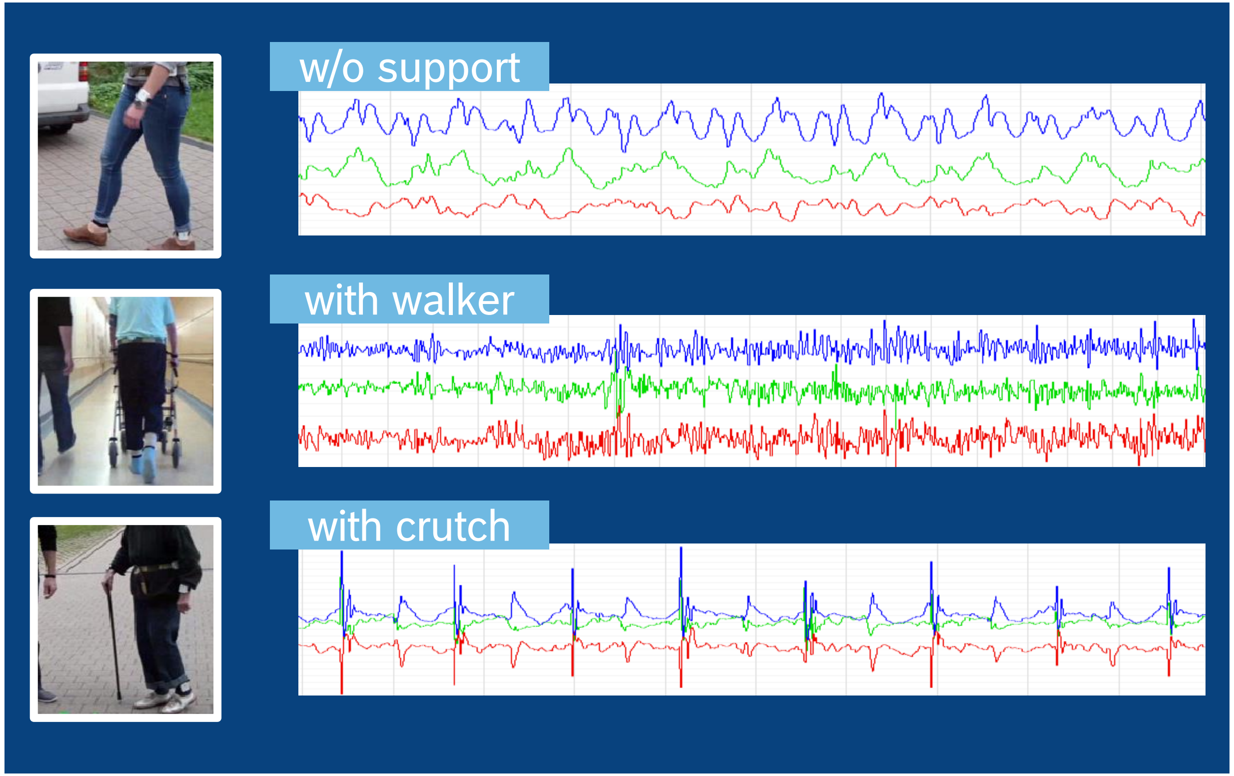
[1] Ghaffarzadegan *et al.*: Occupancy Detection in Commercial and Residential Environments Using Audio Signal (Interspeech 2017)
[2] Schmidt *et al.*: Introducing WESAD, a multimodal dataset for WEarable Stress and Affect Detection (submitted to ICMI 2018)
[3] Dürichen *et al.*: Prediction of electrocardiography features points using seismocardiography data: a machine learning approach (accepted for ISWC 2018)
[4] Humanyu et al.: Learning Front-end Filter-bank Parameters using CNNs for Abnormal Heart Sound Detection (EMBC 2018).

BOSCH

Machine Learning for Wearable Computing
Example: Human Activity Recognition (HAR) for Elderly

- Human activity recognition for people (elderly) who rely on support is still a challenging problem

- Example of 3D acceleration data acquired at wrist position



Münzner *et al.*: CNN-based sensor fusion techniques for multimodal human activity recognition (2017)

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Example: Human Activity Recognition (HAR)

- State of the Art algorithms use „classical“ machine learning aproaches including:

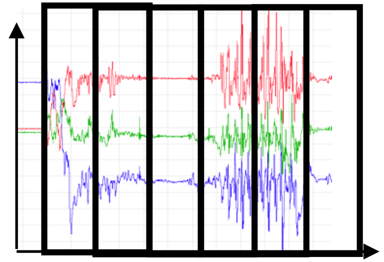
Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing

Example: Human Activity Recognition (HAR)

- State of the Art algo… ng:

Sliding Window Approach

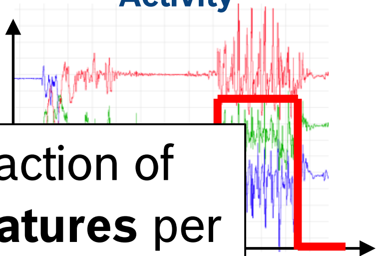


Input / Raw Data

Table 1. The features list.

Feature	Description	Feature	Description
Mean	$\mu_s = \frac{1}{n}\sum_{i=1}^{n} s_i$	*Skewness*	$\frac{1}{n\sigma_s^3}\sum_{i=1}^{n}(s_i-\mu_s)^3$
Minimum	$min(s_1, s_2, \ldots s_n)$	*Kurtosis*	$\frac{1}{n\sigma_s^4}\sum_{i=1}^{n}(s_i-\mu_s)^4$
Maximum	$max(s_1, s_2, \ldots s_n)$	*Signal Power*	$\sum_{i=1}^{n} s_i^2$
Median	$median\ (s_1, s_2, \ldots s_n)$	*Root Mean Square*	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} s_i^2}$
Standard Deviation	$\sigma_s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i-\mu_s)^2}$	*Peak Intensity*	*The number of signal peaks within a certain period of time*
Coefficients of Variation	$\frac{\sigma_s}{\mu_s}$	*Pearson's Correlation Coefficient*	$\frac{cov(a,b)}{\sigma_a\sigma_b}$
Peak-to-peak Amplitude	$max\ (s) - min(s)$	*Inter-axis Cross-Correlation*	$\frac{\sum_{i=1}^{n}(a_i-\mu_a)(b_i-\mu_b)}{\sqrt{\sum_{i=1}^{n}(a_i-\mu_a)^2\sum_{i=1}^{n}(b_i-\mu_b)^2}}$
Percentiles	$t = \frac{np_i}{100} + 0.5,\ p_i = 10, 25, 50, 75, 90$ $percentile(s, p_i) = (1-f)s_k + fs_{k+1}$ *k = integer part of t; f = fractional part of t*	*Autocorrelation*	$R(k) = \frac{1}{(n-k)\sigma_s^2}\sum_{i=1}^{n-k}(s_i-\mu)(s_{i+k}-\mu)\ \forall\ k < n$; *the height of the first and second peaks and the position of the second peak of* $R(k)$
Interquartile Range	$percentile(s, 75) - percentile(s, 25)$	*Trapezoidal Numerical Integration*	$\int_1^n s(x)dx$ *using Multiple Segment Trapezoidal Rule*
Pitch Angle	$arctan(\frac{x_i}{\sqrt{y^2+z_i^2}})$	*Signal Magnitude Area*	$\frac{1}{n}\sum_{i=1}^{n}(\lvert x_i\rvert + \lvert y_i\rvert + \lvert z_i\rvert)$
Roll Angle	$arctan(\frac{y_i}{\sqrt{x^2+z_i^2}})$	*Signal Vector Magnitude*	$\frac{1}{n}\sum_{i=1}^{n}\sqrt{x_i^2+y^2+z_i^2}$
Median Crossings	$t = s - median(s)$ $MC = \sum_{i=1}^{n} sgn(t_i.t_{i+1})$ $sgn(a,b) = \{1$ *if* $(a.b) < 0$; 0 *if* $(a.b) > 0\}$	*Power Spectral Density*	$\frac{1}{n}\sum_{i=1}^{n-1}(s_i\cos\frac{2\pi fi}{n})^2 + (s_i\sin\frac{2\pi fi}{n})^2$ *f denotes the* f^{th} *Fourier coefficient in the frequency domain; the positions and power levels of highest 6 peaks of PSD computed over the sliding window; total power in 5 adjacent and pre-defined frequency bands.*

Activity



Extraction of **176 features** per segment

*Janidarmian *et al.*: A Comprehensive Analysis of Wearable Acceleration Sensors in Human Activity Recognition (2017)

BOSCH

Machine Learning for Wearable Computing
And what about Deep Learning?

Machine Learning for Wearable Computing
And what about Deep Learning?

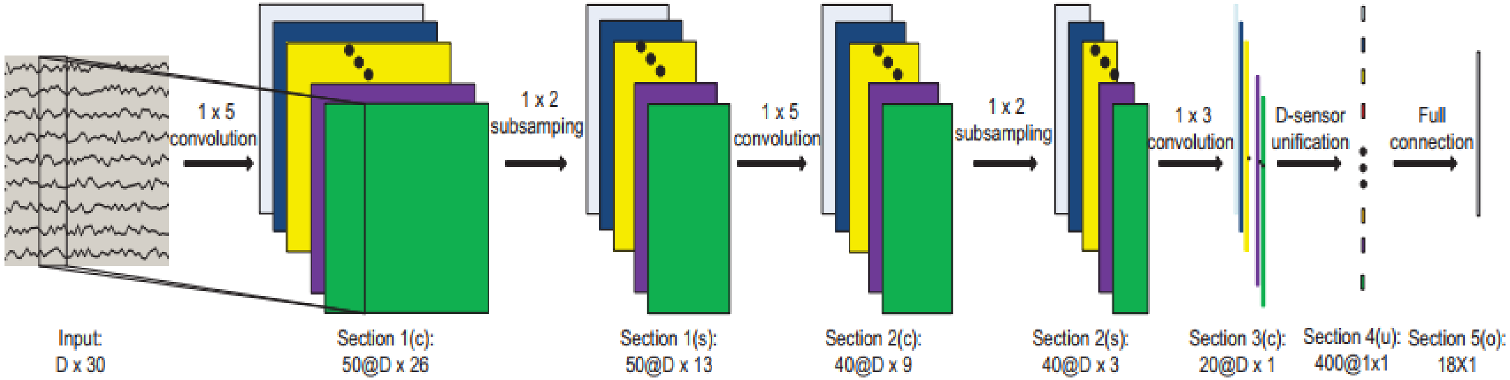
- First promising end-to-end learning results published for HAR since 2011:
  - Plötz *et al.*[1] showed that meaningful features can be learned for HAR using Random Boltzmann Machines

[1] Plötz *et al.*: Feature learning for activity recognition in ubiquitous computing (2011)

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
And what about Deep Learning?

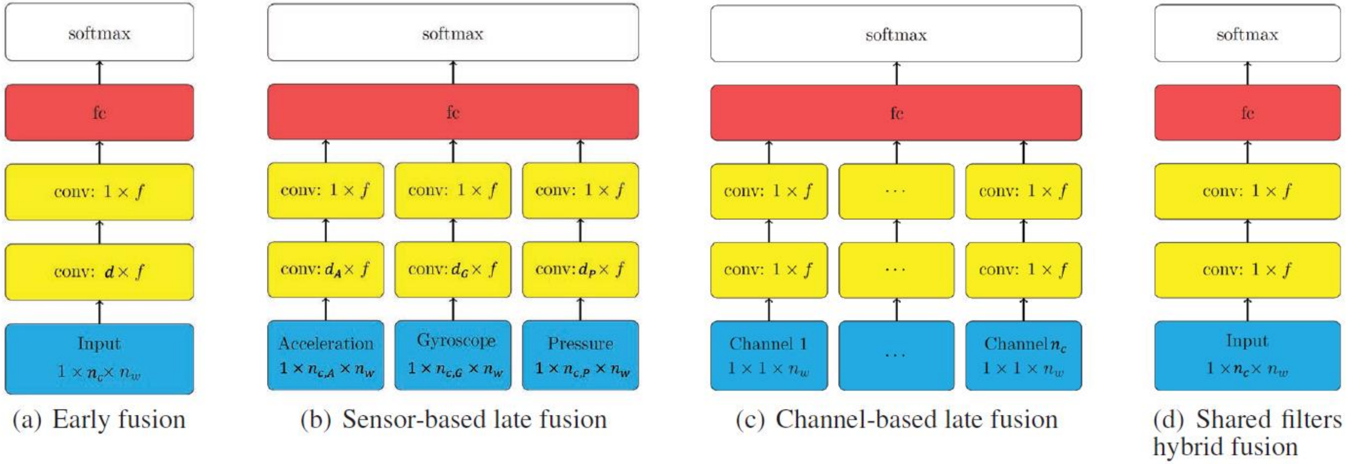
- First promising end-to-end learning results published for HAR since 2011:
  - Plötz *et al.*[1] showed that meaningful features can be learned for HAR using Random Boltzmann Machines
  - Yang *et al.*[2] presented first deep CNN network for HAR



[1] Plötz *et al.*: Feature learning for activity recognition in ubiquitous computing (2011)
[2] Yang *et al.*: Deep convolutional neural networks on multichannel time series for human activity recognition (2015)

Machine Learning for Wearable Computing
And what about Deep Learning?

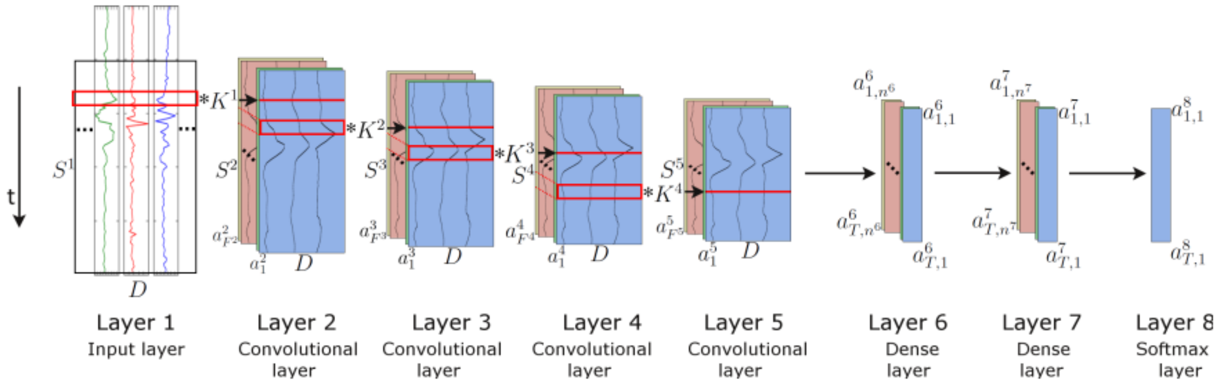
- First promising end-to-end learning results published for HAR since 2011:
  - Plötz *et al.*[1] showed that meaningful features can be learned for HAR using Random Boltzmann Machines
  - Yang *et al.*[2] presented first deep CNN network for HAR
  - Münzner *et al.*[3] detailed investigation multimodal sensor fusion for CNN networks



[1] Plötz *et al.*: Feature learning for activity recognition in ubiquitous computing (2011)
[2] Yang *et al.*: Deep convolutional neural networks on multichannel time series for human activity recognition (2015)
[3] Münzner *et al.*: CNN-based sensor fusion techniques for multimodal human activity recognition (2017)

Machine Learning for Wearable Computing
And what about Deep Learning?

- First promising end-to-end learning results published for HAR since 2011:
  - Plötz *et al.*[1] showed that meaningful features can be learned for HAR using Random Boltzmann Machines
  - Yang *et al.*[2] presented first deep convolutional neural networks (CNN) for HAR
  - Münzner *et al.*[3] detailed investigation multimodal sensor fusion for CNN networks
  - Ordóñez *et al.*[4] used of a deep CNN and RNN which outperformed classical methods by 4% on average



[1] Plötz *et al.*: Feature learning for activity recognition in ubiquitous computing (2011)
[2] Yang *et al.*: Deep convolutional neural networks on multichannel time series for human activity recognition (2015)
[3] Münzner *et al.*: CNN-based sensor fusion techniques for multimodal human activity recognition (2017)
[4] Ordóñez *et al.*: Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition (2016)

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Challenges – Deep Learning

1

Lack of large datasets – ImageNet for HAR is missing

- Only small public datasets available (starting to change*)
- Labelling of data is cumbersome and time consuming
- Cannot be solved easily by e.g. using crowd sourcing

2

Very resource intensive due to high number of computations

„DL approaches are far too complex to be executed on wearables like a smart watch!"

* Doerty *et al.*: Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank Study (2017)

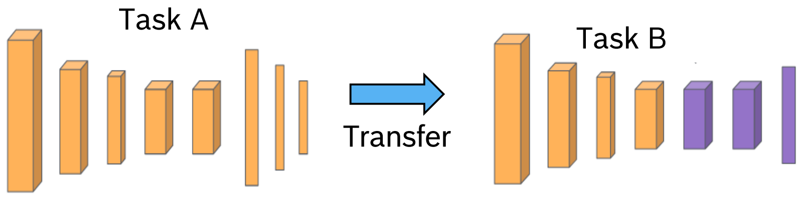
Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Challenges – Deep Learning

1 Lack of large datasets – ImageNet for HAR is missing

- DL approaches offer new possibilities to merge different datasets (e.g. different sensor positions) and to integrate unlabeled data

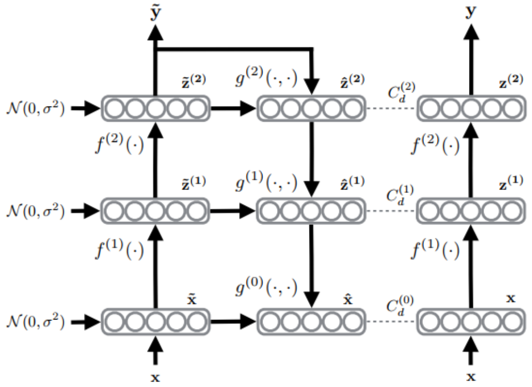
Transfer Learning



- *“Transfer Learning will be the next driver of ML success”* – Andrews Ng, NIPS (2016)
- Ordóñez Morales *et al.*[1] showed that transfer learning is possible between users, application domains, sensor modalities and sensor locations

Semi-Supervised Learning

- End-to-end learning offers possibility to combine labeled and unlabeled data in one optimization function.
- Example of semi-supervised ladder network presented by Rasmus et al.[2] (NIPS, 2016).



General Adversarial Networks…?

[1] Ordóñez Morales *et al.*: Deep Convolutional Feature Transfer Across Mobile Activity Recognition Domains, Sensor Modalities and Locations (2016)
[2] Rasmus *et al.*: Semi-supervised learning with ladder networks (2015)

Machine Learning for Wearable Computing

Challenges – Deep Learning

2 Resource intensity

- **Cry for better hardware …**
  - Apple A11 Bionic chip
  - HUAWEI Kirin 970 with neural processing unit
  - ARM – Project Trillium
  - …

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 **Resource intensity**

- Cry for better hardware ...
- **Algorithmic & architecture adaptations:**

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 Resource intensity

- Cry for better hardware ...
- Algorithmic & architecture adaptations:
  - **“Simple architectural solutions”**
    - Model architecture selection based on the architecture the platform can afford
    - Replace feed-forward layers with single shallow classifier,

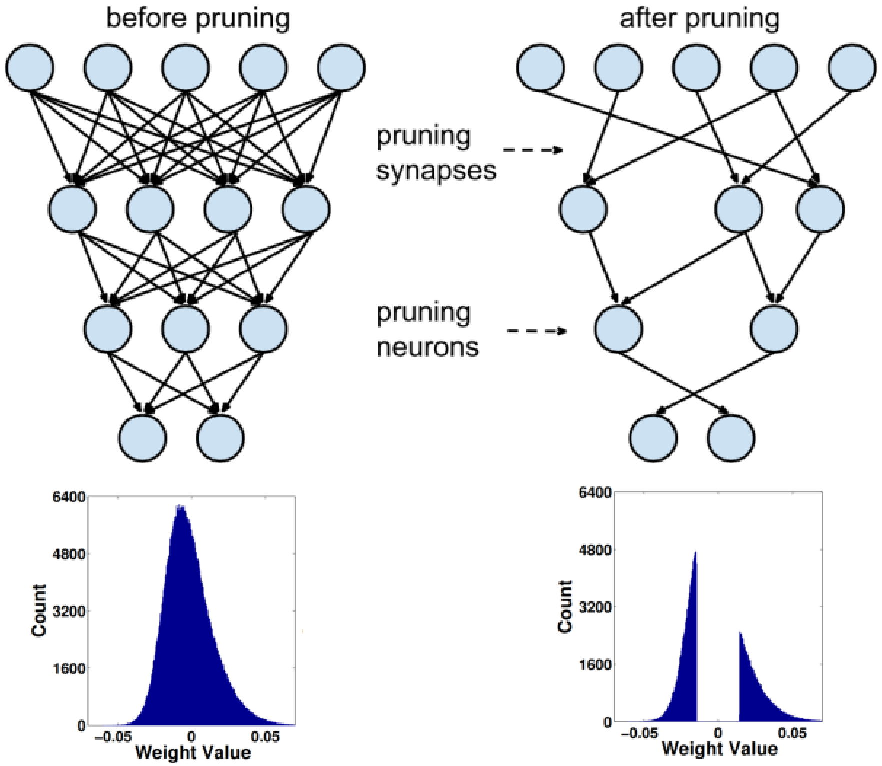
    ...

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 **Resource intensity**

- Cry for better hardware ...
- Algorithmic & architecture adaptations:
  - "Simple architectural solutions"
  - **Node pruning**



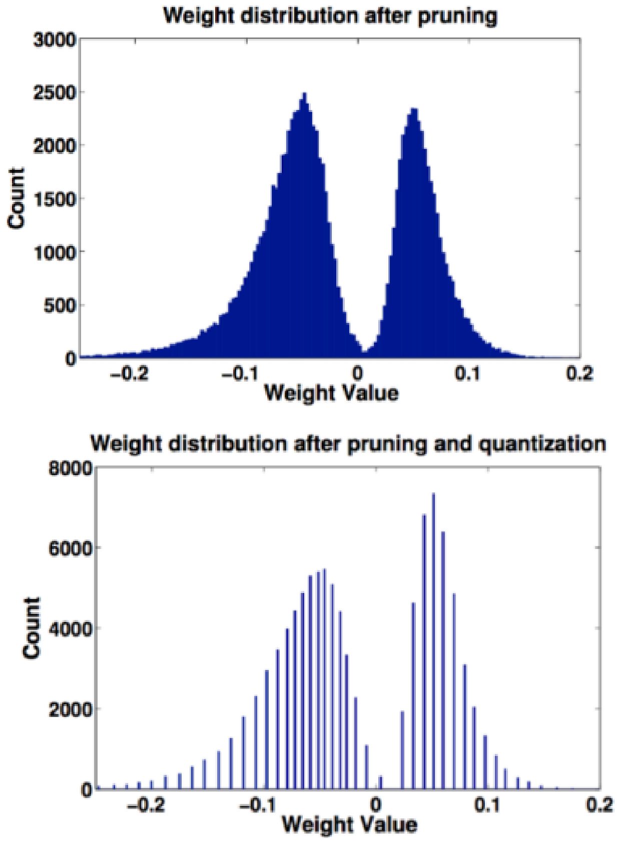
Han S.: "Learning both Weights and Connections for Efficient Neural Networks" (2015)

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 Resource intensity

- Cry for better hardware ...
- Algorithmic & architecture adaptations:
  - “Simple architectural solutions”
  - Node pruning
  - **Quantization**

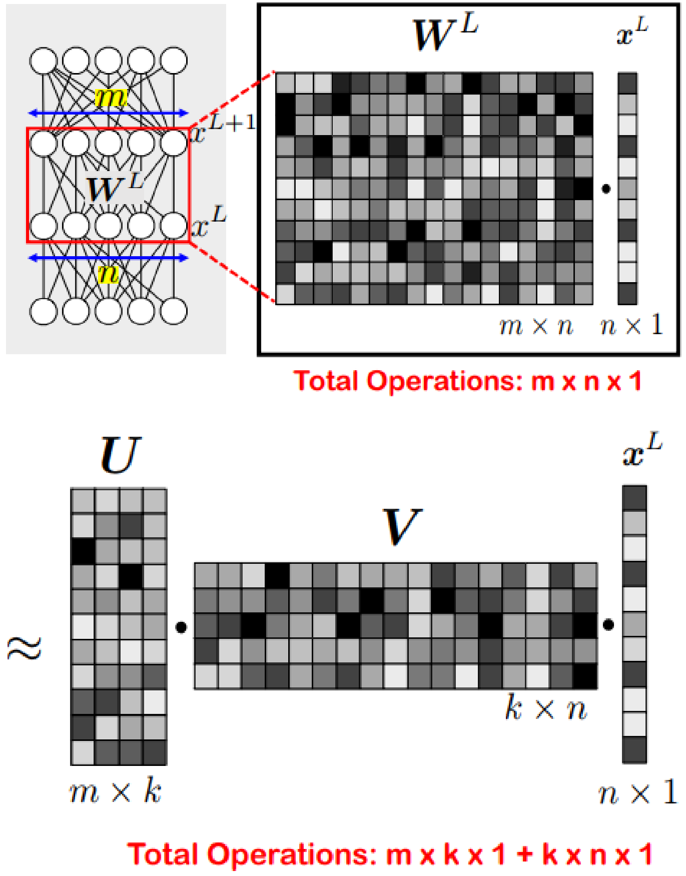


Han S.: “Deep Compression: Compressing Deep Neural Networks with pruning, trained quantization and huffman coding” (2015)

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 Resource intensity

- Cry for better hardware ...
- Algorithmic & architecture adaptations:
  - “Simple architectural solutions”
  - Node pruning
  - Quantization
  - **Low rank approximations and sparse coding of weight matrices**



Plötz *et al.*: „Deep Learning for UBICOMP“, tutorial @ISWC2017

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

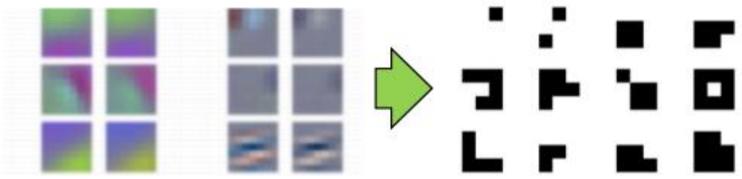
BOSCH

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 Resource intensity

- Cry for better hardware ...
- Algorithmic & architecture adaptations:
  - “Simple architectural solutions”
  - Node pruning
  - Quantization
  - Low rank approximations and sparse coding of weight matrices
  - **Low precision and binary networks**

Binary Kernels



	Network Variations	Operations used in Convolution	Memory Saving (Inference)	Computation Saving (Inference)	Accuracy on ImageNet (AlexNet)
Standard Convolution	Real-Value Inputs: 0.11 -0.21 ... -0.34 / -0.25 0.61 ... 0.52; Real-Value Weights: 0.12 -1.2 ... 0.41 / -0.2 0.5 ... 0.68	+ , − , ×	1x	1x	%56.7
Binary Weight	Real-Value Inputs: 0.11 -0.21 ... -0.34 / -0.25 0.61 ... 0.52; Binary Weights: 1 -1 ... 1 / -1 1 ... 1	+ , −	~32x	~2x	%56.8

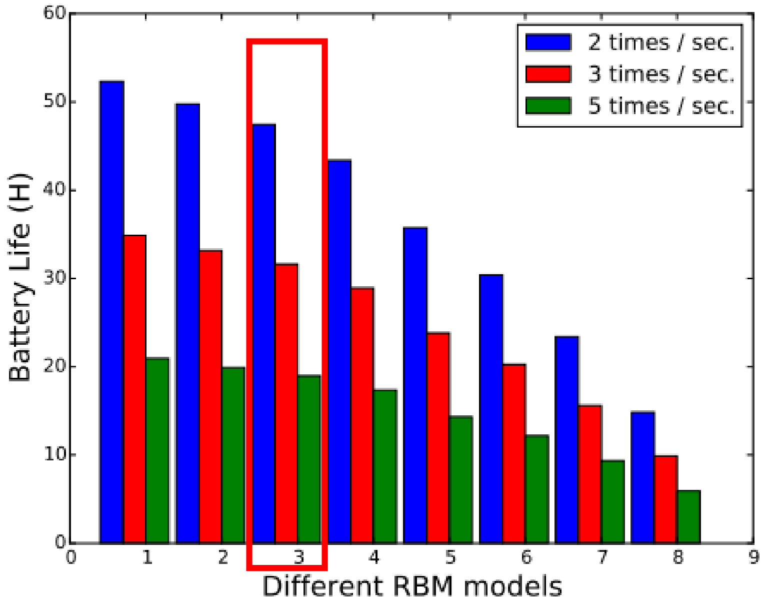
M. Rastegari: „XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks“ (2015)

BOSCH

Machine Learning for Wearable Computing
Challenges – Deep Learning

2 Resource intensity

- First results running a DL model directly on a smart watch by Bhattacharya & Lane*
  - Implemented on Snapdragon 400 (used in LG G smart watch R)
  - Investigated tasks: gesture, activity tracking and indoor/outdoor
  - Model with 3 hidden layer outperforms state of the art algorithms
  - Runtime of 20 – 48h possible (400mAh battery)



*Bhattacharya and Lane: From Smart to Deep: Robust Activity Recognition on Smartwatches using Deep Learning (2016)

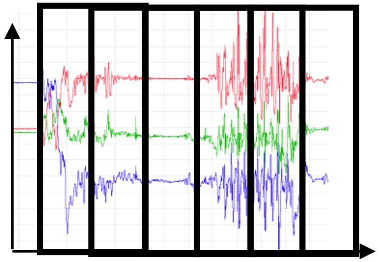
Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

BOSCH

Machine Learning for Wearable Computing
New Chances with Deep Learning

- Bringing standard machine learning algorithms on an embedded device is cumbersome and time consuming (due to manual feature enginerring)

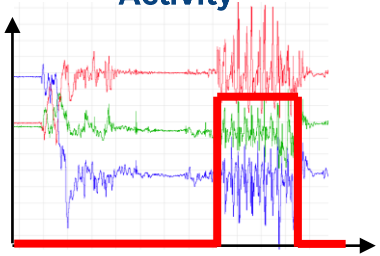
Sliding Window Approach



Input / Raw Data

⇨ **Preprocessing (Filtering)** ⇨ **Feature Extraction** ⇨ **Machine Learning Classifier** ⇨

Activity



Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing
New Chances with Deep Learning

- Bringing standard machine learning algorithms on an embedded device is cumbersome and time consuming (due to manual feature engineering)
- Fast success of deep learning approaches due to large and open frameworks
- Development of special library for fast transfer for trained DL models onto target platform:
  - Qualcomm Snapdragon Neural Processing Engine SDK
  - ARM NN SDK

Chance

- Faster deployment due to end-to-end learning with standardized building blocks and well supported frameworks

Machine Learning for Wearable Computing
Conclusions

- State of the art for edge computing is still “classical” machine learning with manual feature engineering and sliding window approach

ML algorithms for wearables will change quickly

“Deep learning on constrained devices, such as phones, watches, and even embedded sensors, is already well on its way to becoming mainstream. “

Nicolas D. Lane

- Very active research field - increasing community of industrial and academic partners (new hardware, continuous algorithm improvements…)
- Chance for faster deployment of DL models on embedded targets due to standardized building blocks (CNNs, RNNs, MLPs) and well supported frameworks

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04

Machine Learning for Wearable Computing.
A Chance for Deep Learning!

Questions?

Dr. Robert Dürichen (CR/AEU2) | 2018-07-04