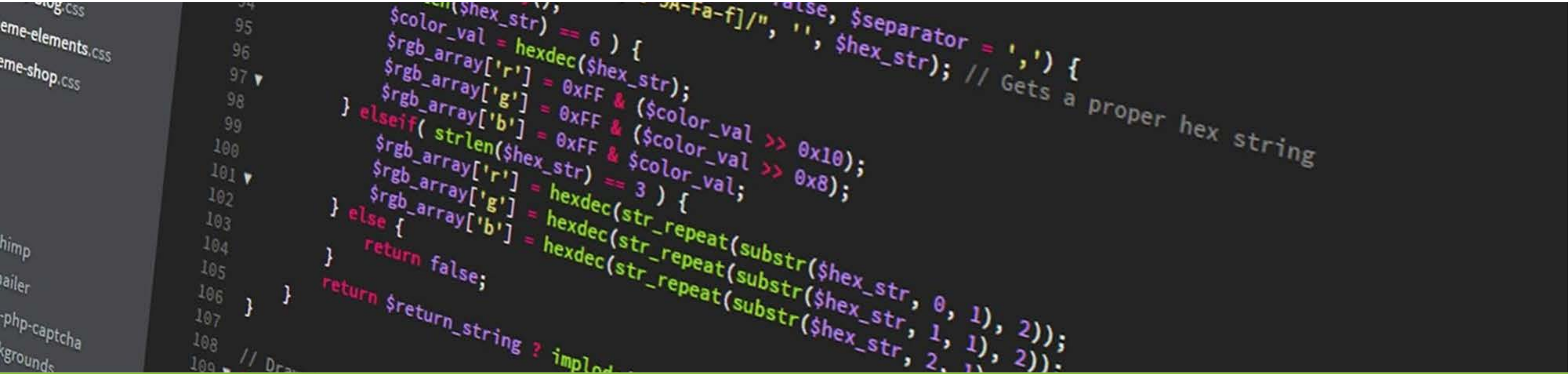


Adaptive User Interfaces

Challenges and Application of AI in HCI

CC BY SA

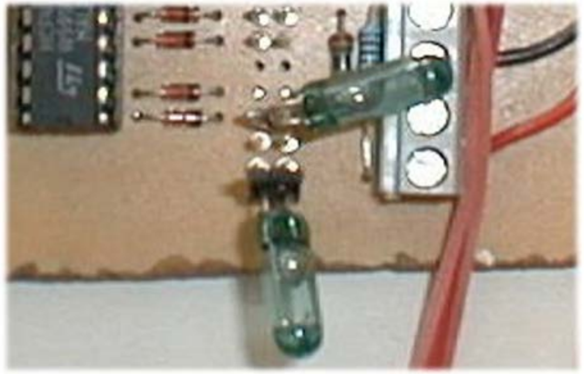
“White Kitchen Interior” by Paintzen Attribution 2.0 Generic (CC BY 2.0)

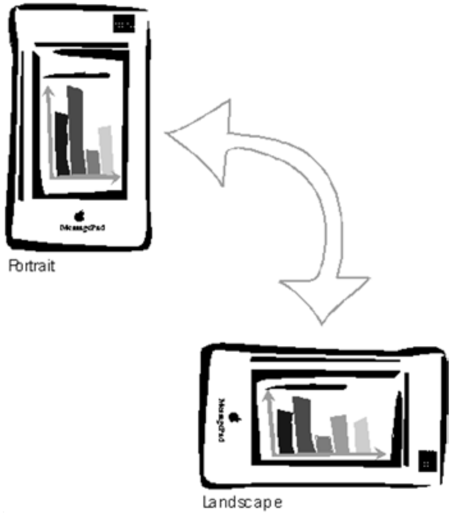
Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing Mobile Voice Assistants with WorldGaze. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–10. DOI: https://doi.org/10.1145/3313831.3376479

Sensing

Getting Physical (1) Initial Experience (1998)

- **Extremely simple, but still it creates a new experience**
- 2-Bit Input
- Not an input device
- Very specific function

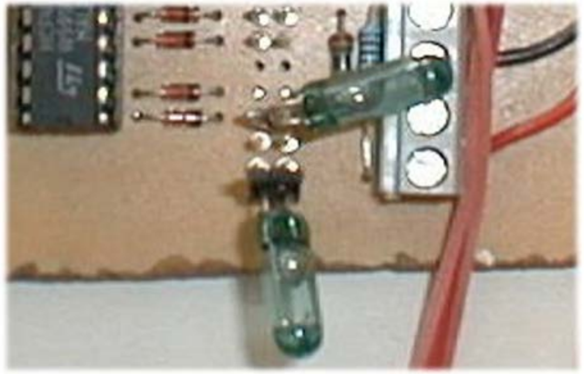


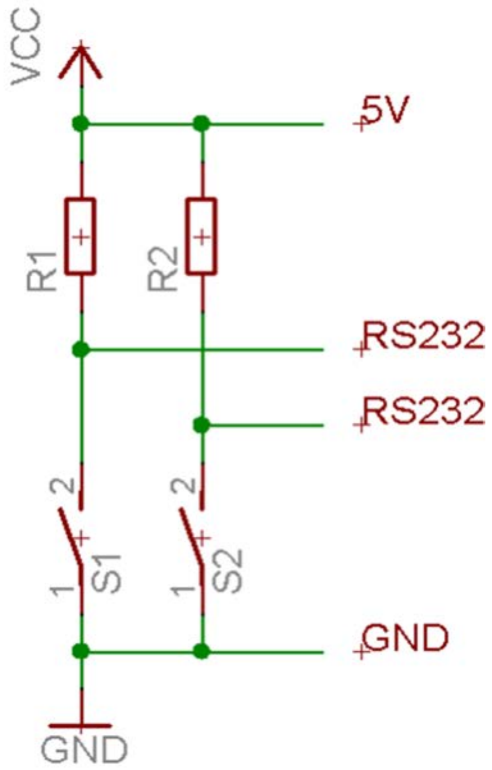


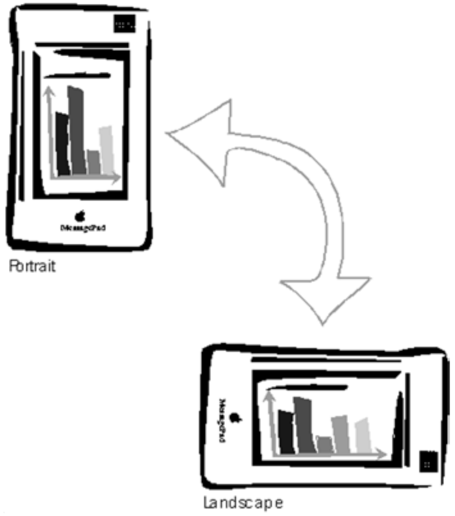
A. Schmidt, M. Beigl, H. Gellersen. There is more to context than location. Computers and Graphics, 23(6):893--901, 1999.
http://www.comp.lancs.ac.uk/~albrecht/pubs/pdf/schmidt_cug_elsevier_12-1999-context-is-more-than-location.pdf

Getting Physical (1) Initial Experience (1998)

- **Extremely simple, but still it creates a new experience**
- 2-Bit Input
- Not an input device
- Very specific function

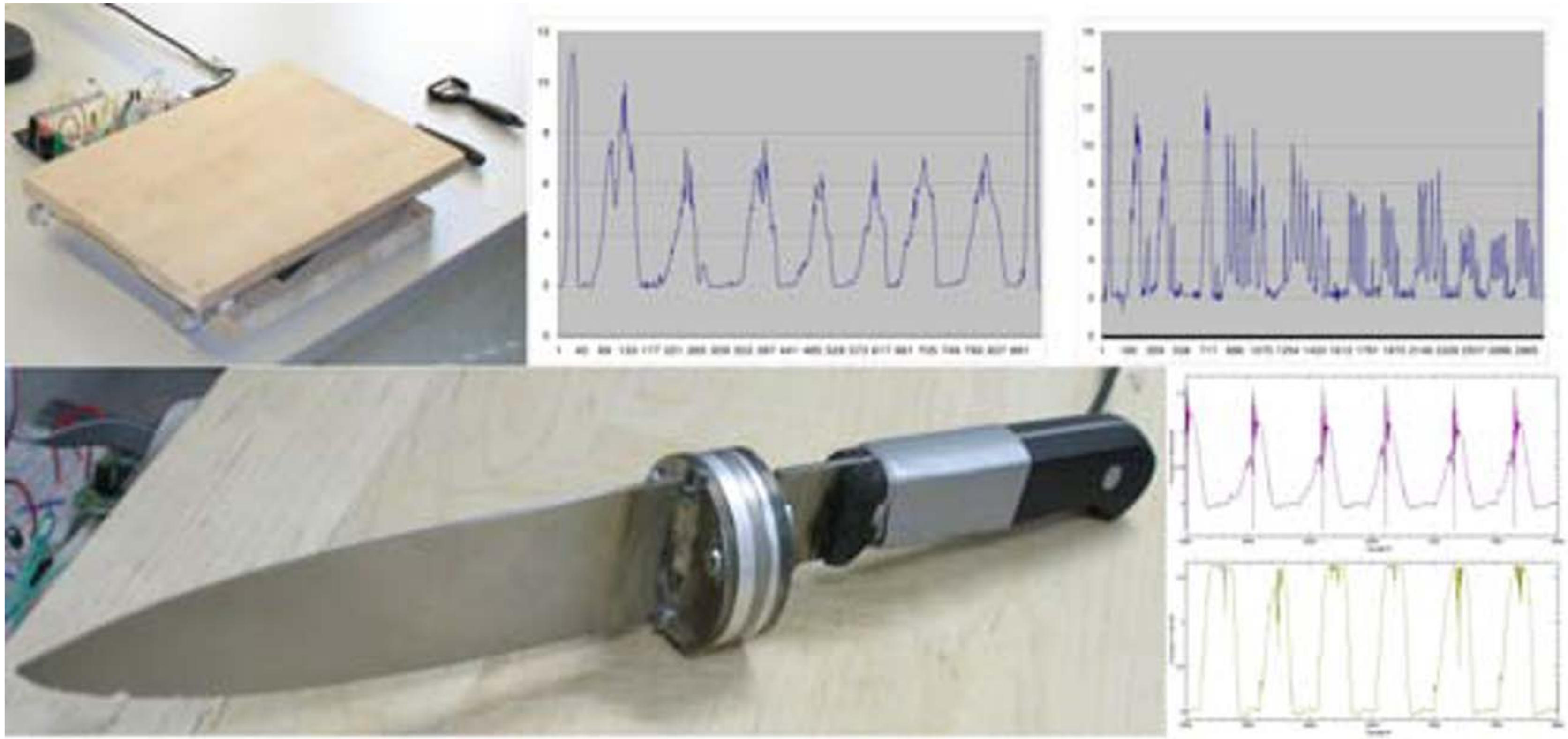






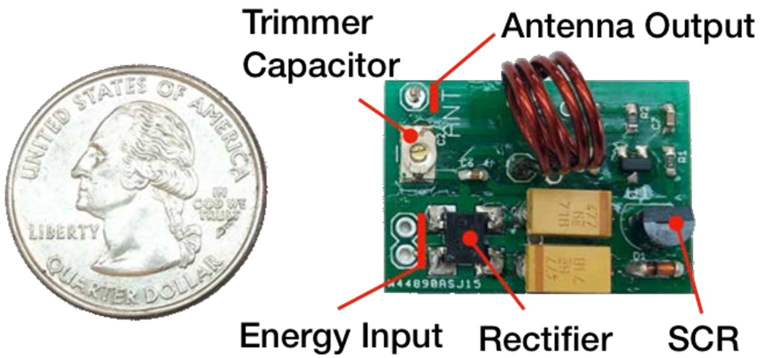
A. Schmidt, M. Beigl, H. Gellersen. There is more to context than location. Computers and Graphics, 23(6):893--901, 1999.
http://www.comp.lancs.ac.uk/~albrecht/pubs/pdf/schmidt_cug_elsevier_12-1999-context-is-more-than-location.pdf

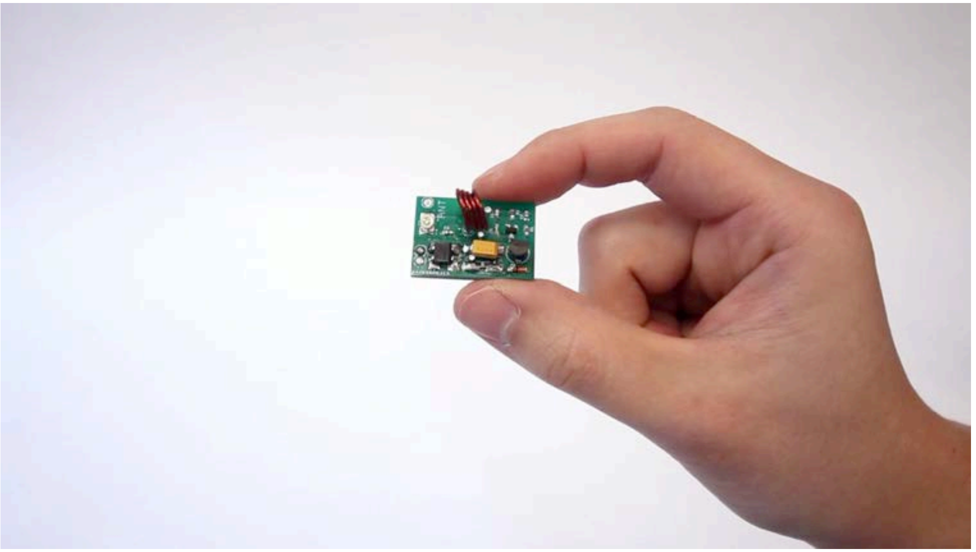
Knife that “knows” what its cuts



Matthias Kranz, Albrecht Schmidt, Alexis Maldonado, Radu Bogdan Rusu, Michael Beetz, Benedikt Hörnler, and Gerhard Rigoll. 2007. Context-aware kitchen utilities. In Proceedings of the 1st international conference on Tangible and embedded interaction (TEI '07). Association for Computing Machinery, New York, NY, USA, 213–214. DOI: https://doi.org/10.1145/1226969.1227013

Radio Tags for Activity Sensing

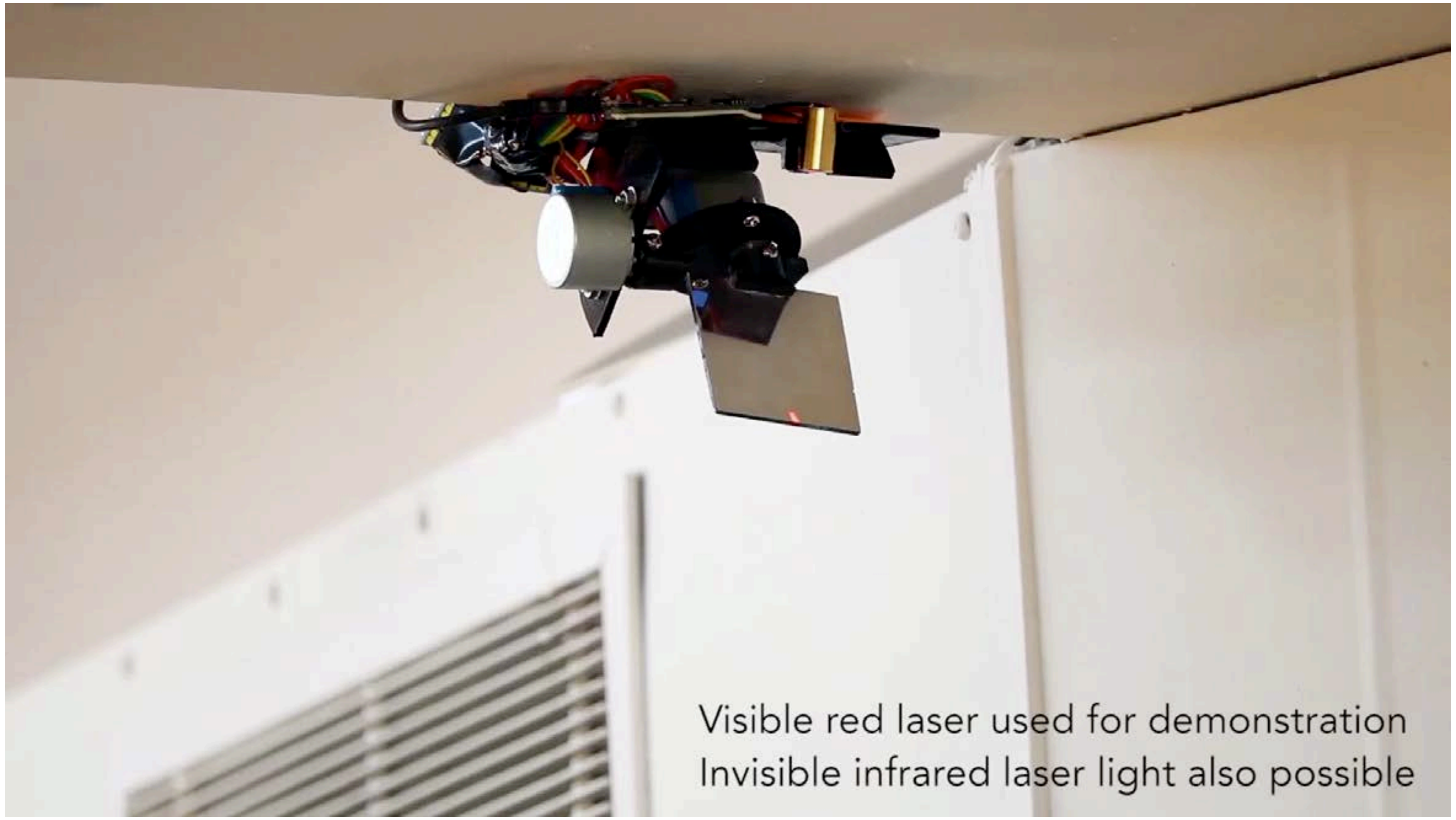




Yang Zhang, Yasha Iravantchi, Haojian Jin, Swarun Kumar, and Chris Harrison. 2019. Sozu: Self-Powered Radio Tags for Building-Scale Activity Sensing. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19). ACM, New York, NY, USA, 973–985. DOI: https://doi.org/10.1145/3332165.3347952 Video: https://youtu.be/wbq-eOOIPyw

Vibrometry for Environment Sensing

https://youtu.be/51XaZDki6yg



Yang Zhang, Gierad Laput, and Chris Harrison. 2018. Vibrosight: Long-Range Vibrometry for Smart Environment Sensing. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18). ACM, New York, NY, USA, 225–236. DOI: https://doi.org/10.1145/3242587.3242608

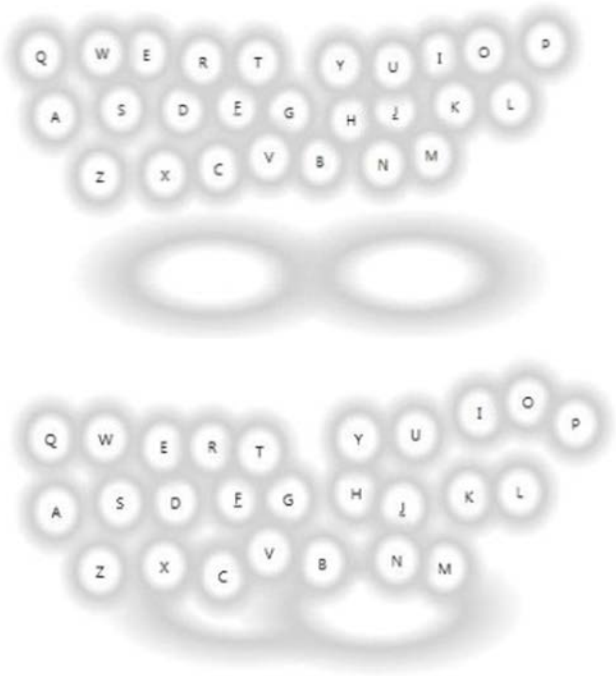
Extracting Contextual Information

- Users Location
  - GPS
  - Direction of Voice
- Users Activity
- Users Emotion
- Users Pose
- Objects Surrounding the User
- Status of Objects

AI in Adaptive Interfaces

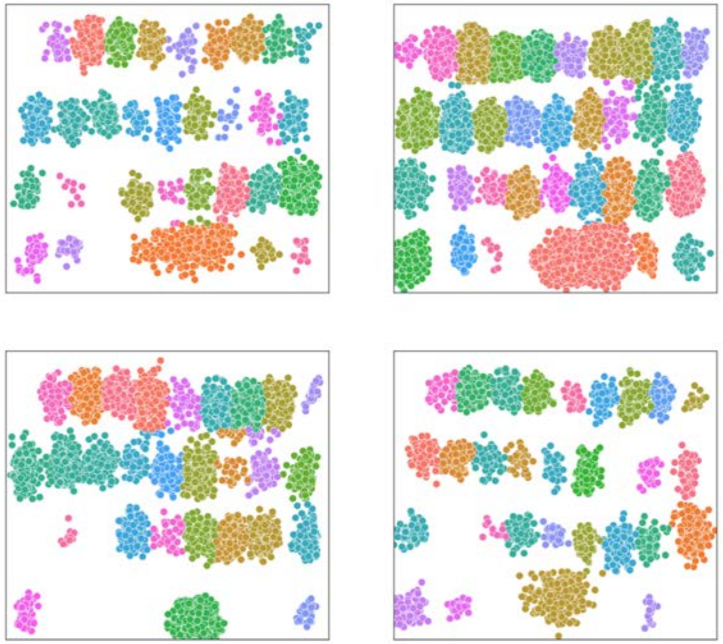
Individual Typing Behaviour

Tabletop



[Findlater and Wobbrock 2012]

Smartphone

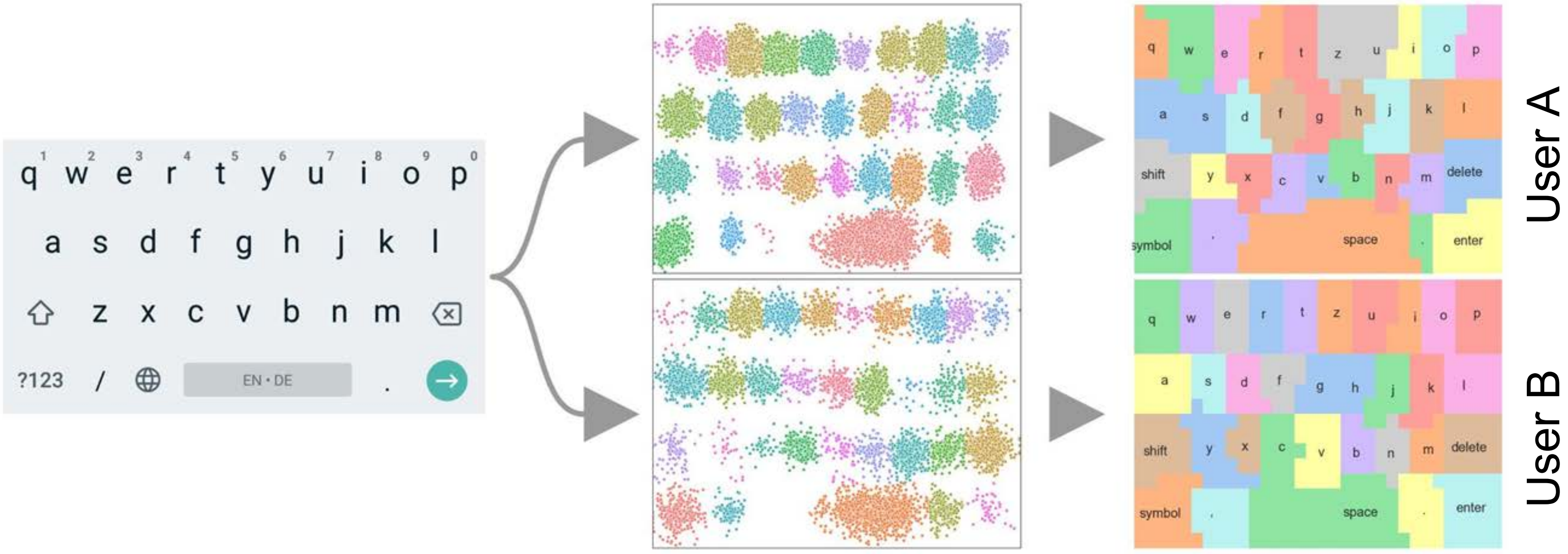


[Buschek et al. 2018]

Adapting Keyboards to Typists

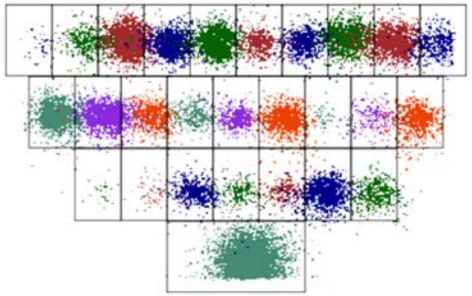
Overview

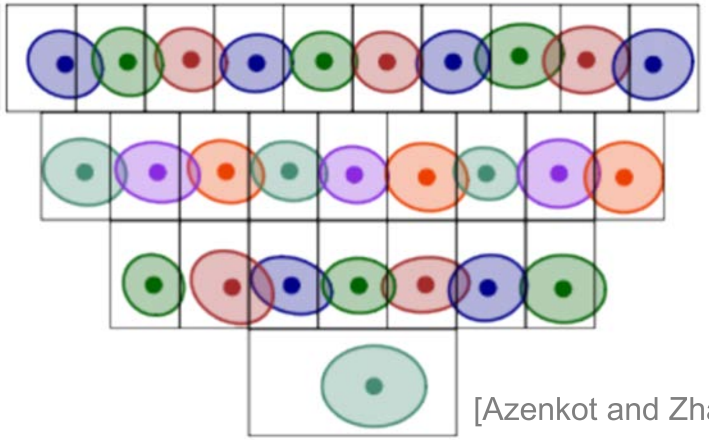
Visible keyboard | **Collect touches** | **Adapt underlying key regions**



Modelling Touchscreen Keypresses

From x,y touch points to one Gaussian per key

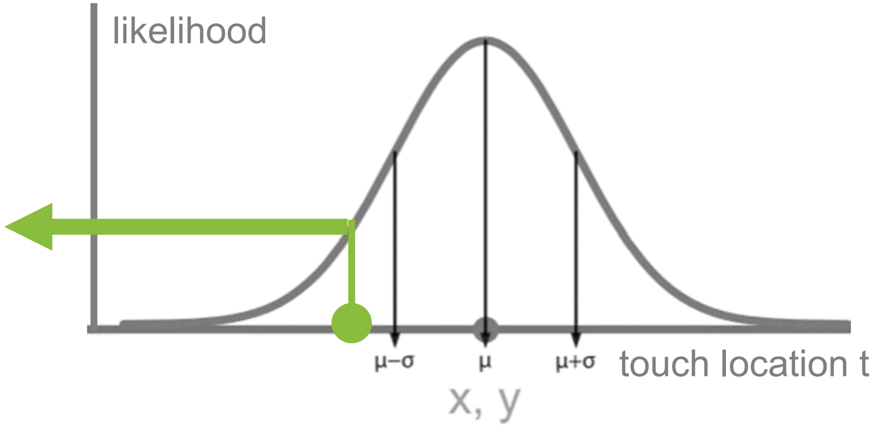




[Azenkot and Zhai 2012]

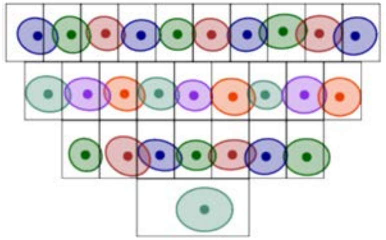
Gaussian key model:
(shown in 1D here)

$$p(t|k) = \mathcal{N}(\mu_k, \sigma_k^2)$$



Probabilistic Keyboard Model

Which key does the user intend to press? i.e. „input decoding“



Gaussian key model

Language model

e.g. based on last five characters:

$$p(k_n|k_{n-1}..k_{n-5})$$

$$k' = \underset{k\in K}{\operatorname{argmax}} \underbrace{\left(\frac{p(t|k)p(k)}{p(t)}\right)}_{\text{Bayes' rule } = p(k|t)} = \underset{k\in K}{\operatorname{argmax}} \left(p(t|k)p(k)\right)$$

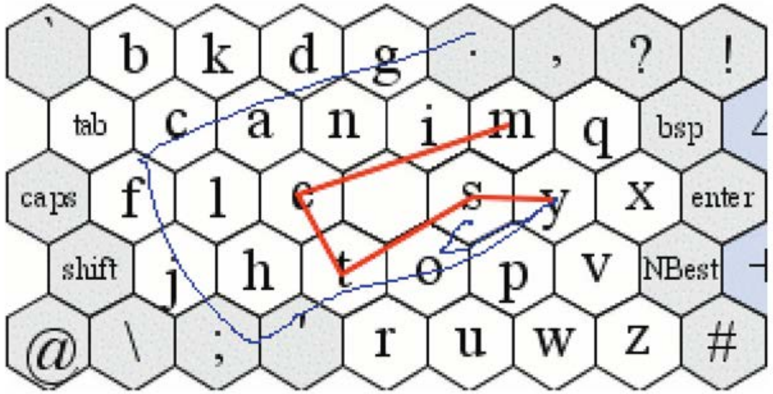
predicted keypress

Evaluate and colour ***k'*** for each pixel

[Yin et al. 2013]

Gesture-based Decoding

Infer intended word from shape of finger trace on the keyboard

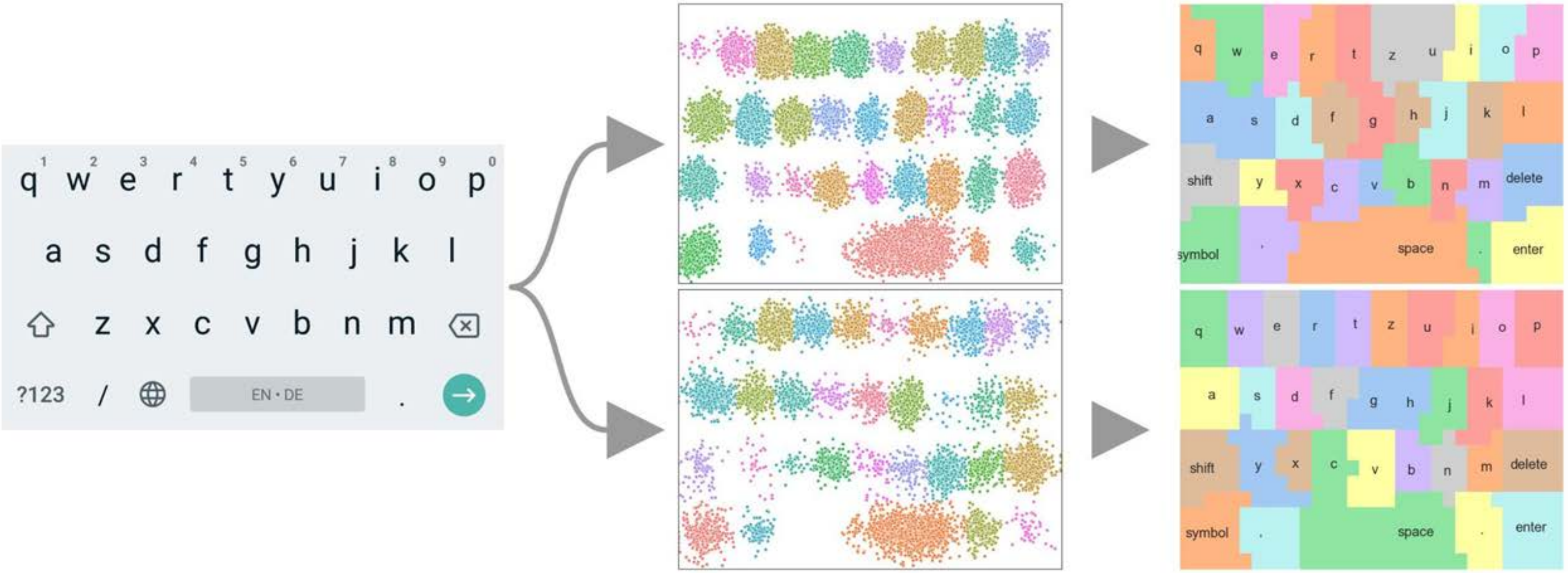


„SHARK²" [Kristensson and Zhai 2004]

Microsoft SwiftKey (screenshot Nov 2020)

Context-Aware Keyboards

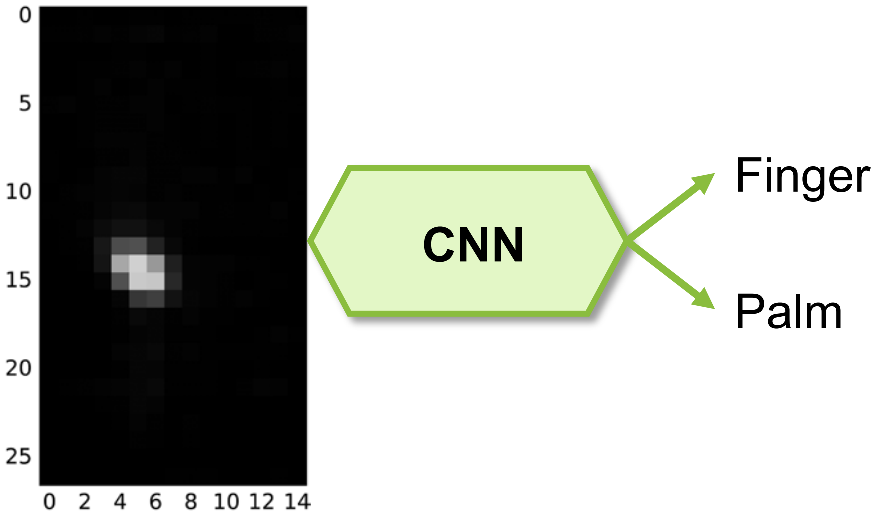
Visible keyboard **Collect touches** **Adapt underlying key regions**



- Palm detection – input rejection
- Finger identification, Finger Orientation – model improvement
- Grip detection – model selection

Palm Detection

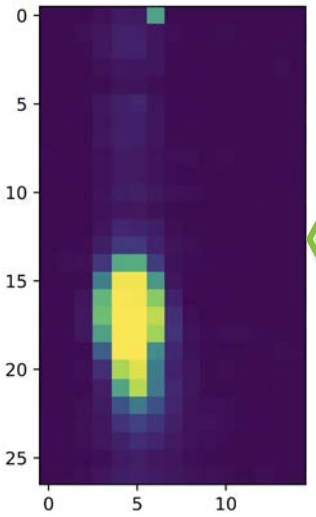
- Convolutional Neural Network
- Classification
- Representation Learning



Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. 2018. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Paper 360, 1–13. DOI: https://doi.org/10.1145/3173574.3173934

Finger Orientation

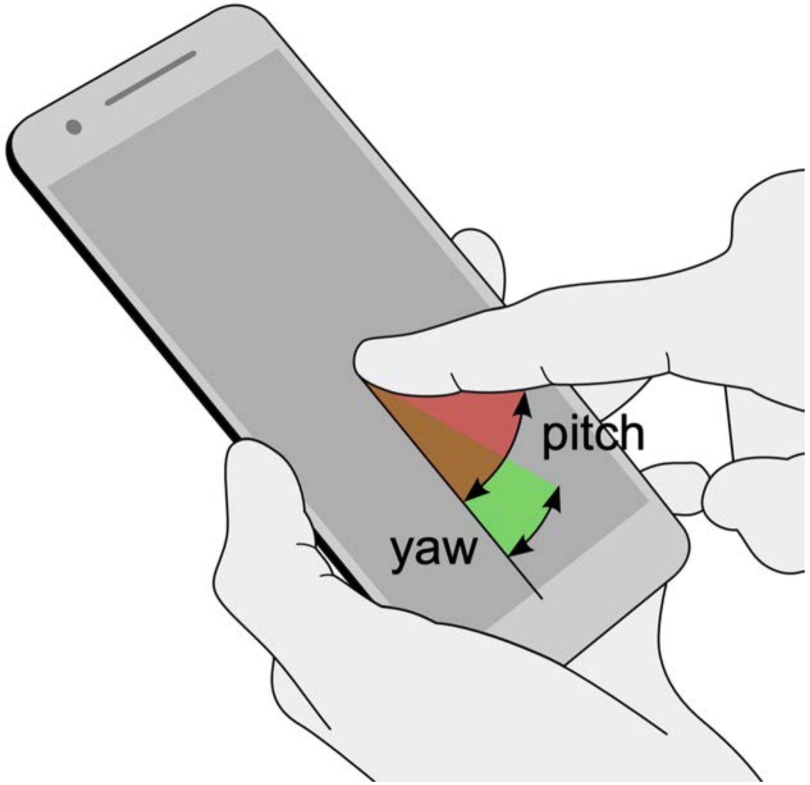
- Convolutional Neural Network
- Regression
- Representation Learning



CNN

Pitch: 4°

Yaw:
17°



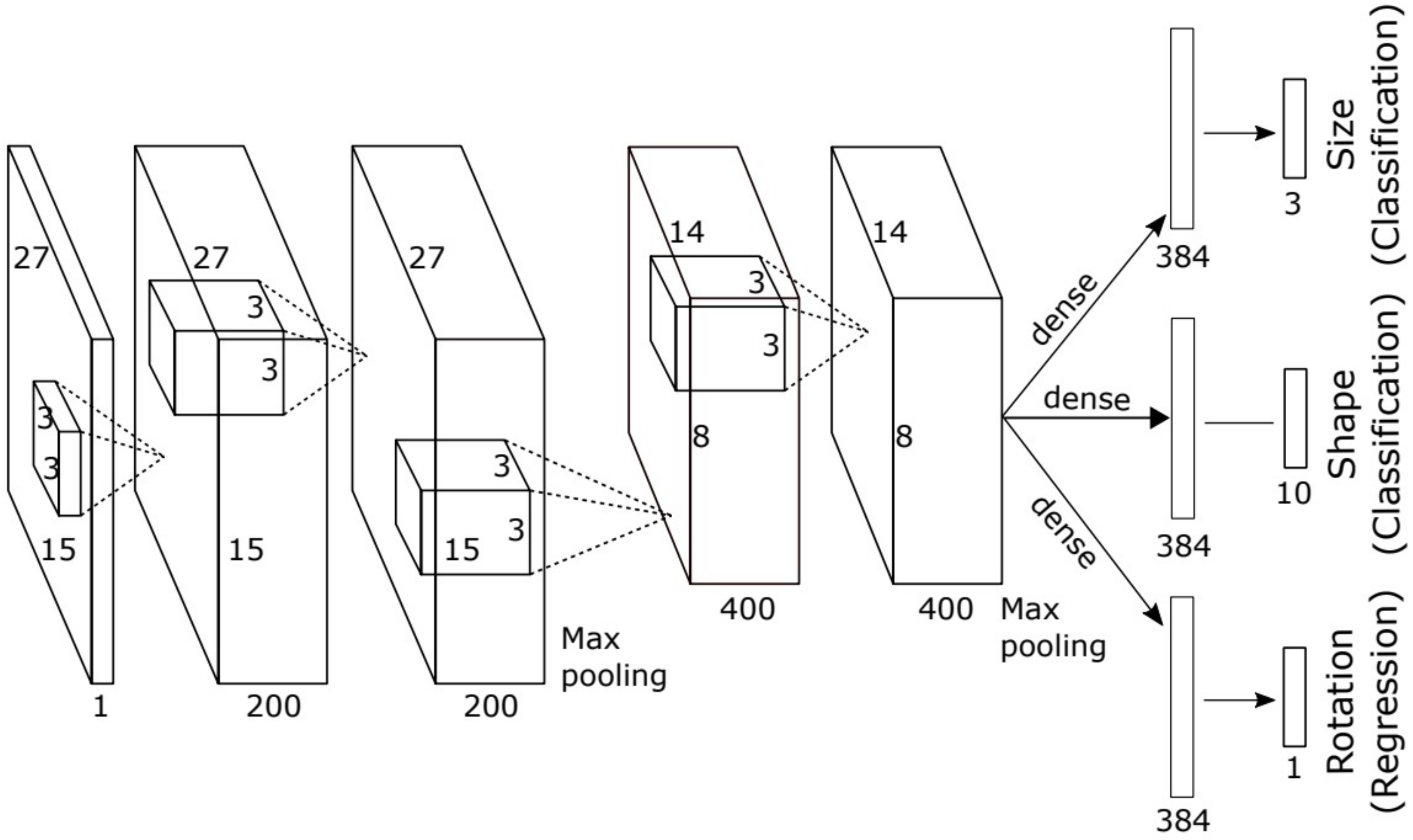
Sven Mayer, Huy Viet Le, and Niels Henze. 2017. Estimating the Finger Orientation on Capacitive Touchscreens Using Convolutional Neural Networks. In Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17). ACM, New York, NY, USA, 220–229. DOI: https://doi.org/10.1145/3132272.3134130

Itsy-Bits: Fabrication and Recognition of 3D-Printed Tangibles with Small Footprints on Capacitive Touchscreens

Figure 1: Itsy-Bits recognizes 3D-printed tangibles as small as a fingertip via the capacitive image of an embedded conductive shape (A). This opens up a variety of tangible user interfaces on the most common form factors of touchscreens, such as individualized interactive board games (B) or a more tangible learning experience (C).

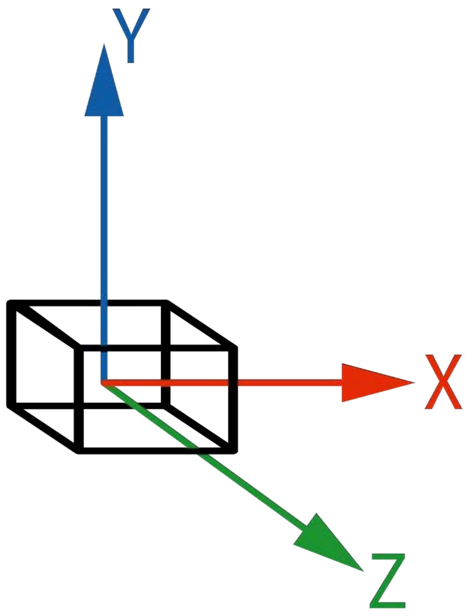
Martin Schmitz, Florian Müller, Max Mühlhäuser, Jan Riemann, and Huy Viet Viet Le. 2021. Itsy-Bits: Fabrication and Recognition of 3D-Printed Tangibles with Small Footprints on Capacitive Touchscreens. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 419, 1–12. DOI: https://doi.org/10.1145/3411764.3445502

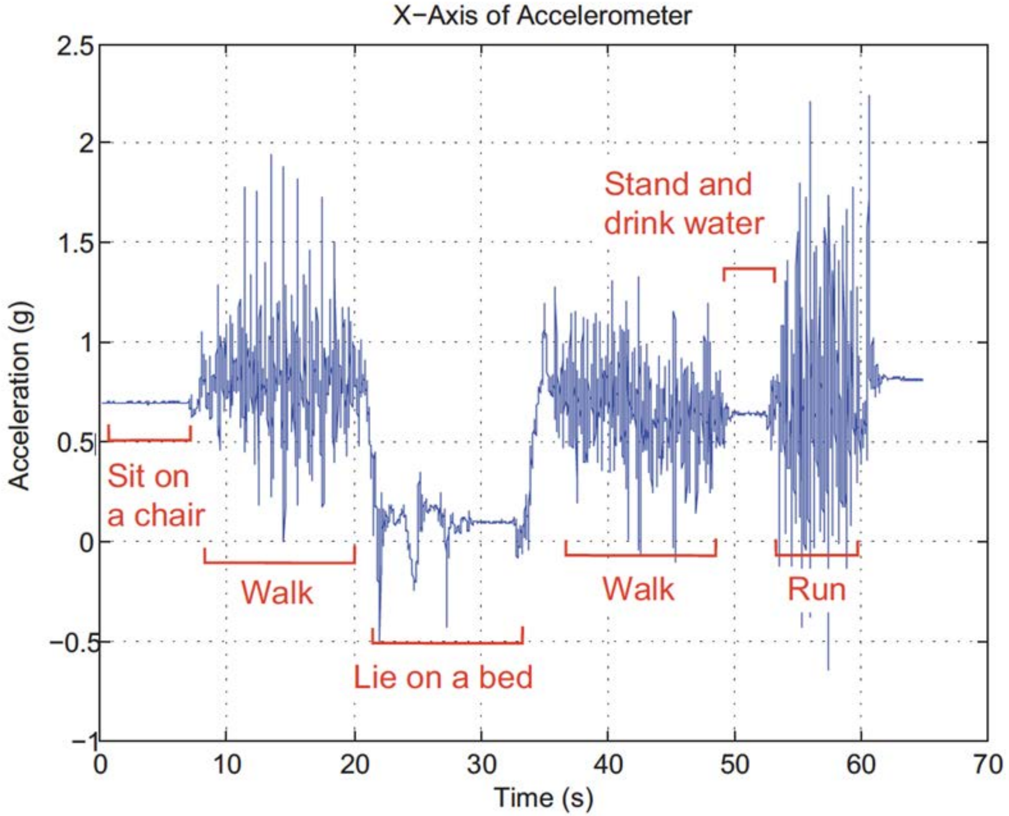
Itsy-Bits: Fabrication and Recognition of 3D-Printed Tangibles with Small Footprints on Capacitive Touchscreens



Martin Schmitz, Florian Müller, Max Mühlhäuser, Jan Riemann, and Huy Viet Viet Le. 2021. Itsy-Bits: Fabrication and Recognition of 3D-Printed Tangibles with Small Footprints on Capacitive Touchscreens. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 419, 1–12. DOI: https://doi.org/10.1145/3411764.3445502

Activity Recognition – Accelerometers





Mi Zhang and Alexander A. Sawchuk. 2012. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12). Association for Computing Machinery, New York, NY, USA, 1036–1043. DOI: https://doi.org/10.1145/2370216.2370438

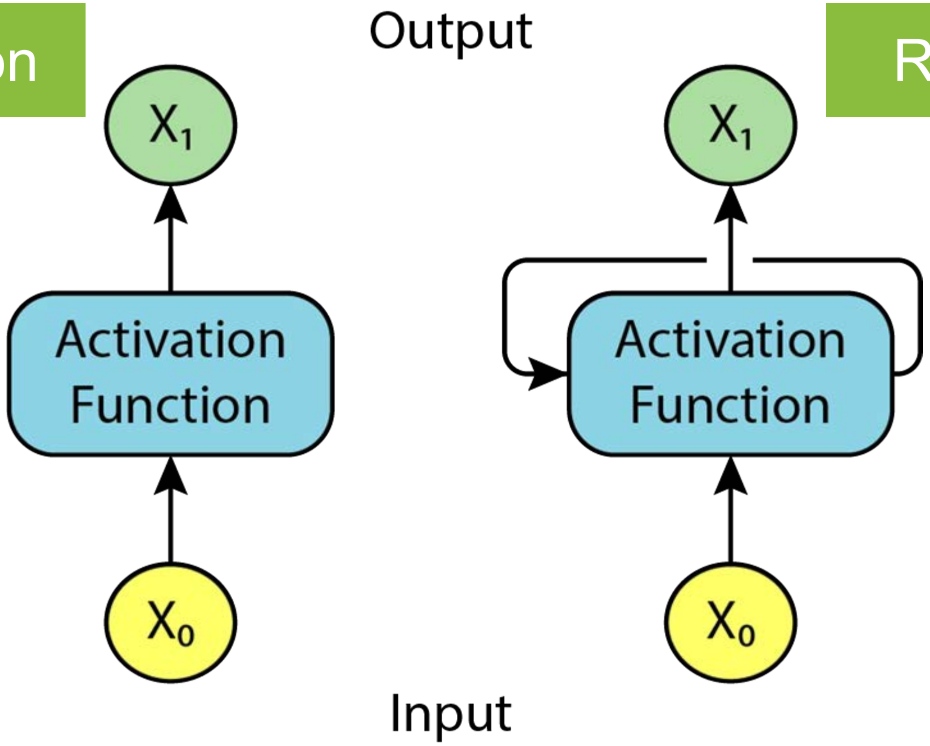
Deep Learning Methods for Multimodal Human Activity Recognition, Paul Lukowicz, DFKI/TU Kaiserslautern (Monday, 11 October 10:00)

Neuronal Network With Timeseries Data

Recurrent Neural Network

Perceptron

RNN Unit

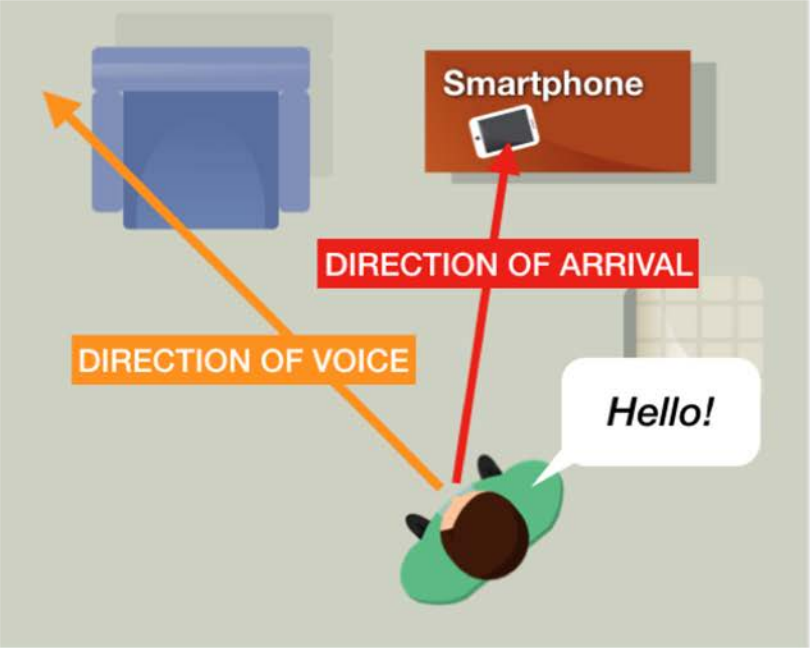


Long short-term memory (**LMST**)

In depth LSTM tutorial: http://colah.github.io/posts/2015-08-Understanding-LSTMs/
Code Examples: https://github.com/cwi-dis/mobile-har-tutorial

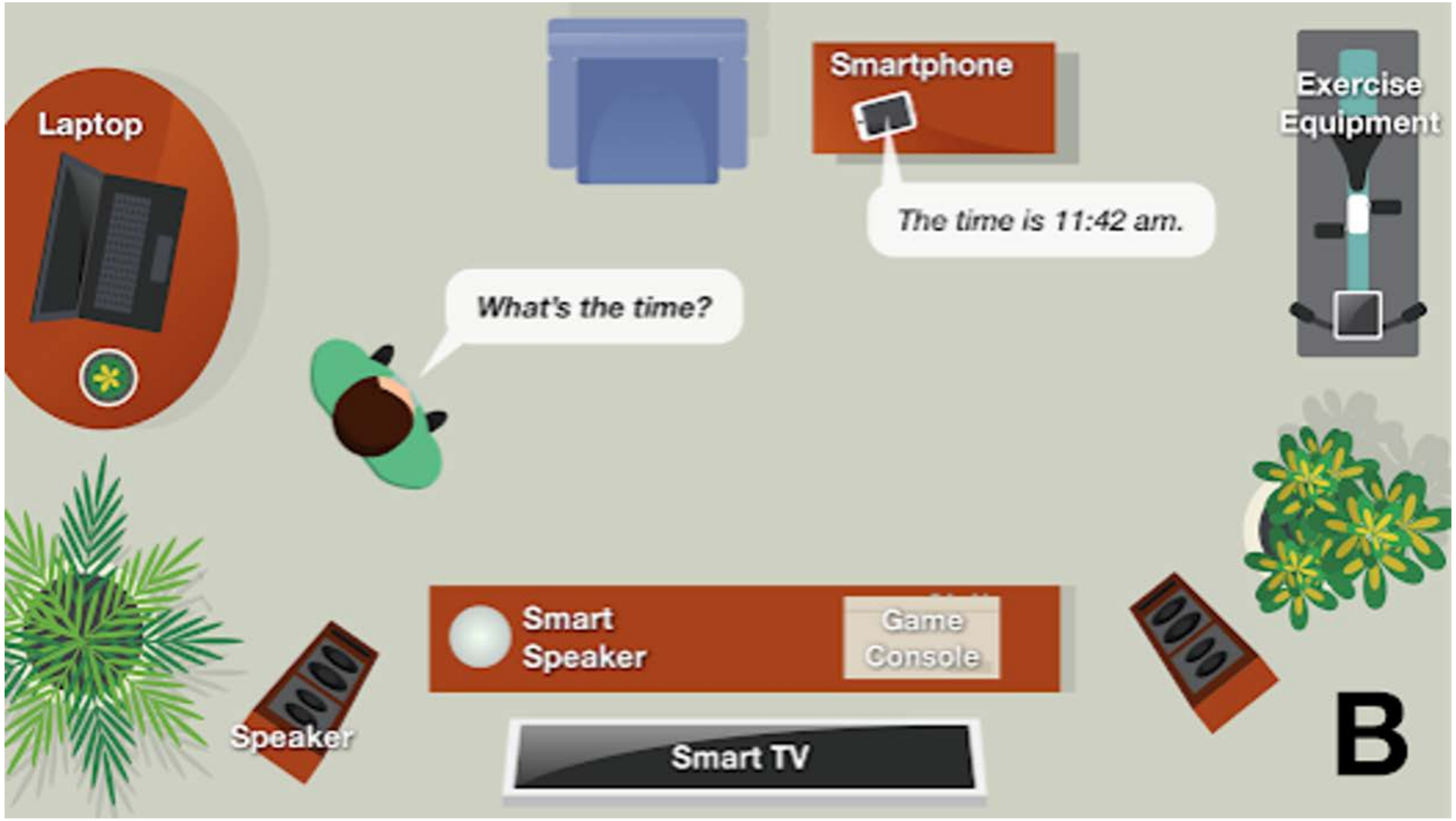
Smart Environments - Direction-of-Voice

- Feature Extraction e.g. volume, speech frequency ratio
- Machine Learning e.g. ensemble-based decision trees



Karan Ahuja, Andy Kong, Mayank Goel, and Chris Harrison. 2020. Direction-of-Voice (DoV) Estimation for Intuitive Speech Interaction with Smart Devices Ecosystems. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20). ACM, New York, NY, USA, 1121–1131. DOI: https://doi.org/10.1145/3379337.3415588

Smart Environments - Direction-of-Voice



Karan Ahuja, Andy Kong, Mayank Goel, and Chris Harrison. 2020. Direction-of-Voice (DoV) Estimation for Intuitive Speech Interaction with Smart Devices Ecosystems. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20). ACM, New York, NY, USA, 1121–1131. DOI: https://doi.org/10.1145/3379337.3415588

LSMTs in Adaptive UIs

RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor

Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 5, 1–14. DOI: https://doi.org/10.1145/3411764.3445367

LSMTs in Adaptive UIs

RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor

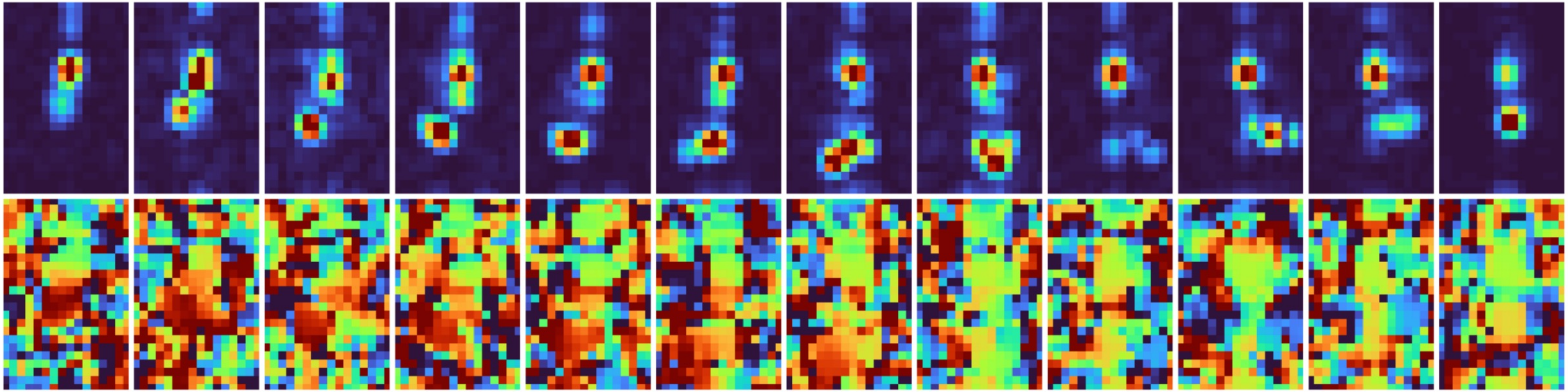


Figure 1: A time series visualization of a left swipe motion with absolute range Doppler maps (top) and interferometry maps (bottom). A hand moving close to a Soli chip round 6th frame, then, moving away. In the interferometry maps, the color of cells s at the same position as hand above changed from red to green, and to blue, showing the hand moving right to left.

Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 5, 1–14. DOI: https://doi.org/10.1145/3411764.3445367

LSMTs in Adaptive UIs

RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor

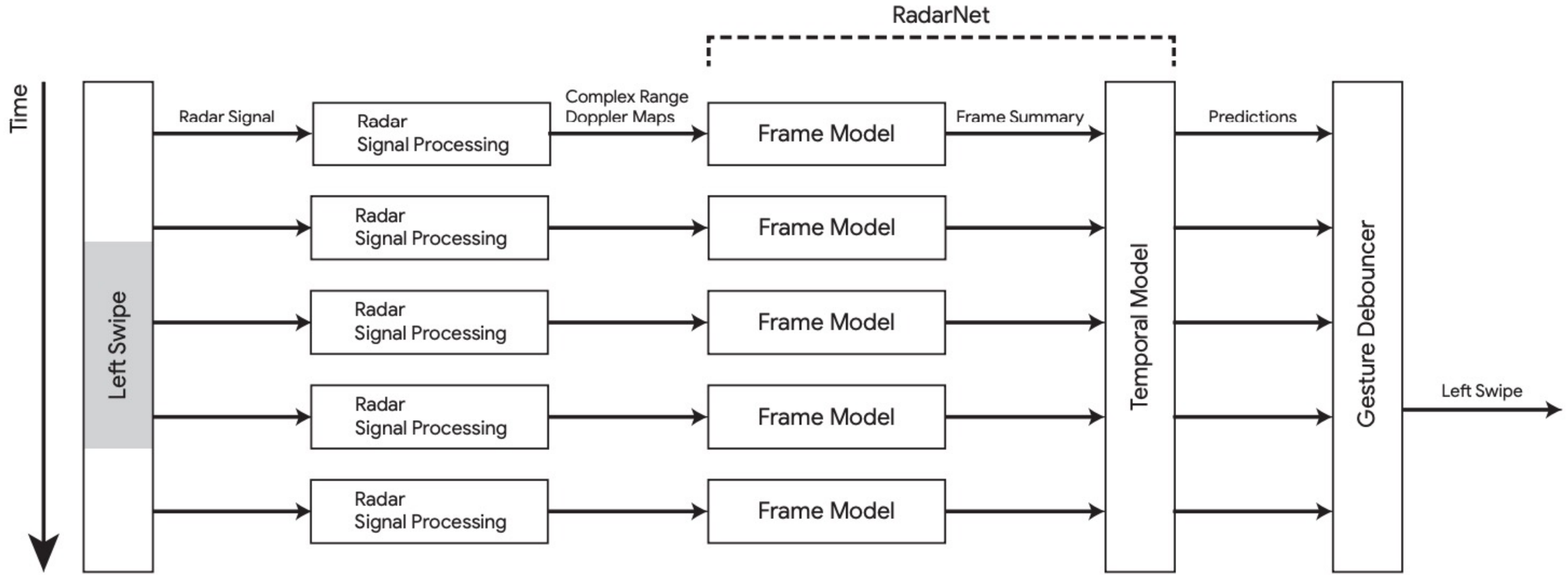
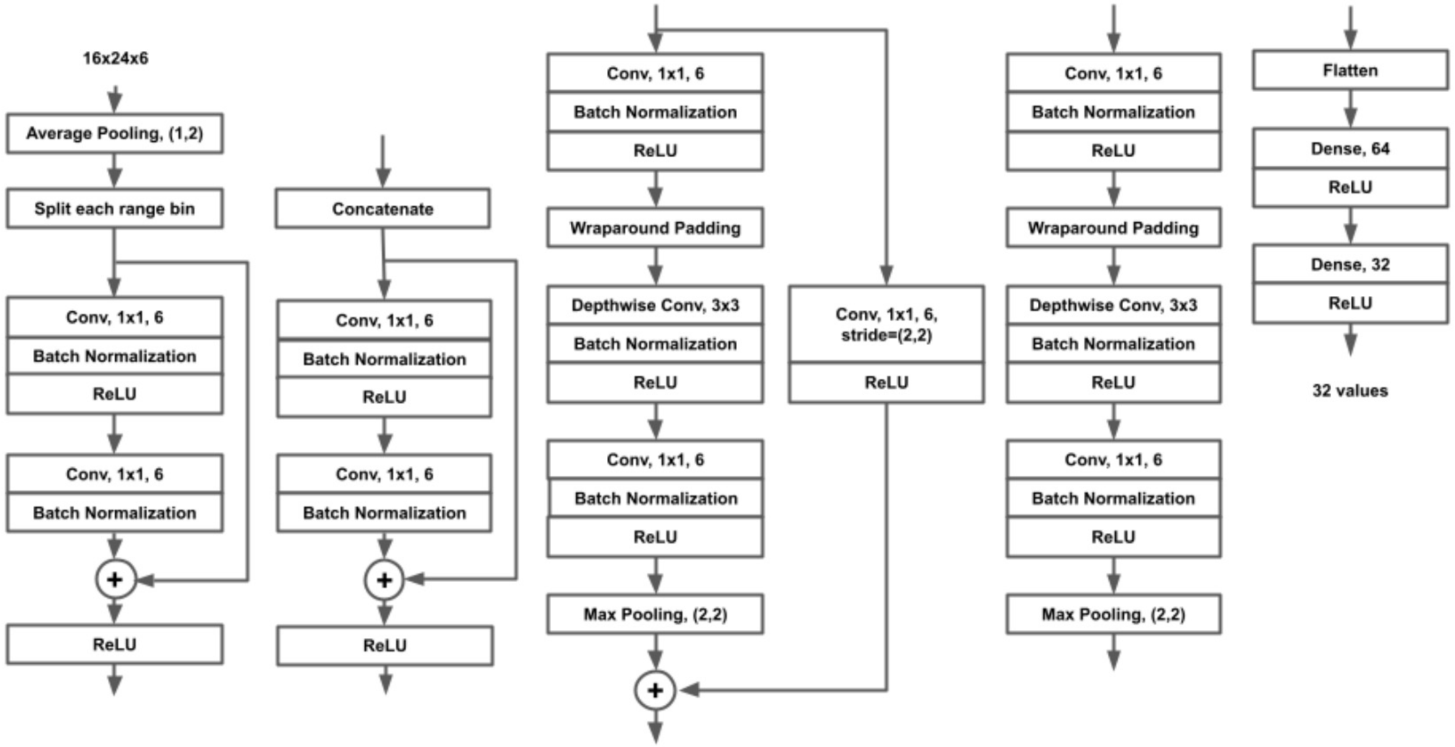


Figure 6: The pipeline applies signal processing algorithms to the radar signals to compute complex range Doppler maps for each burst. The complex range Doppler maps are converted into a summary consisting of 32 values with the frame model of the RadarNet. Then the summaries in the last 12 frames are processed by the temporal model of the RadarNet to make portrait, landscape, and omni predictions. Finally, the gesture debouncer output recognizes gestures.

Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 5, 1–14. DOI: https://doi.org/10.1145/3411764.3445367

LSMTs in Adaptive UIs

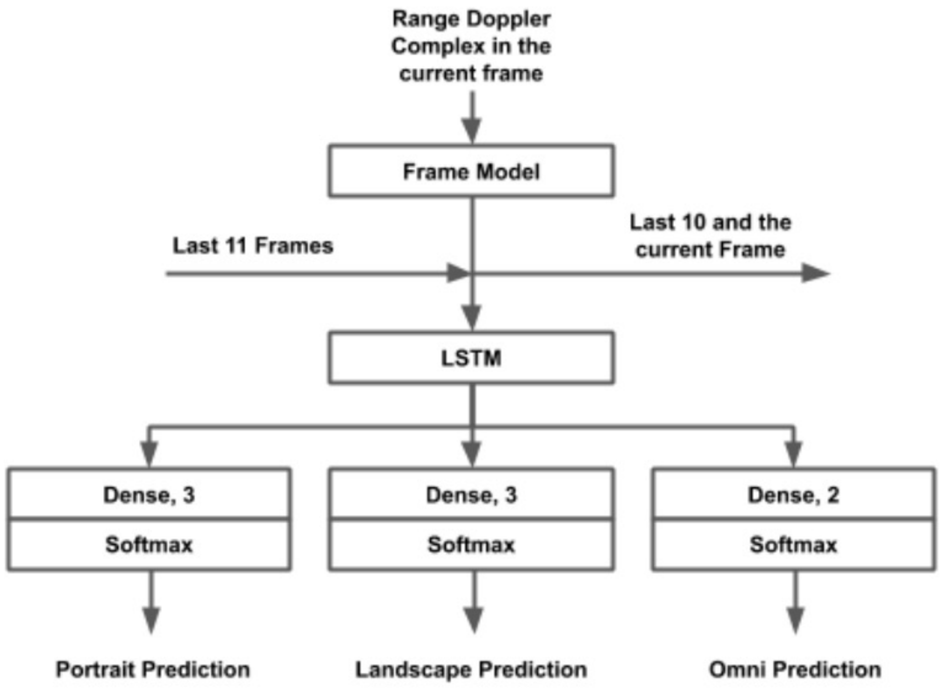
RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor



Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 5, 1–14. DOI: https://doi.org/10.1145/3411764.3445367

LSMTs in Adaptive UIs

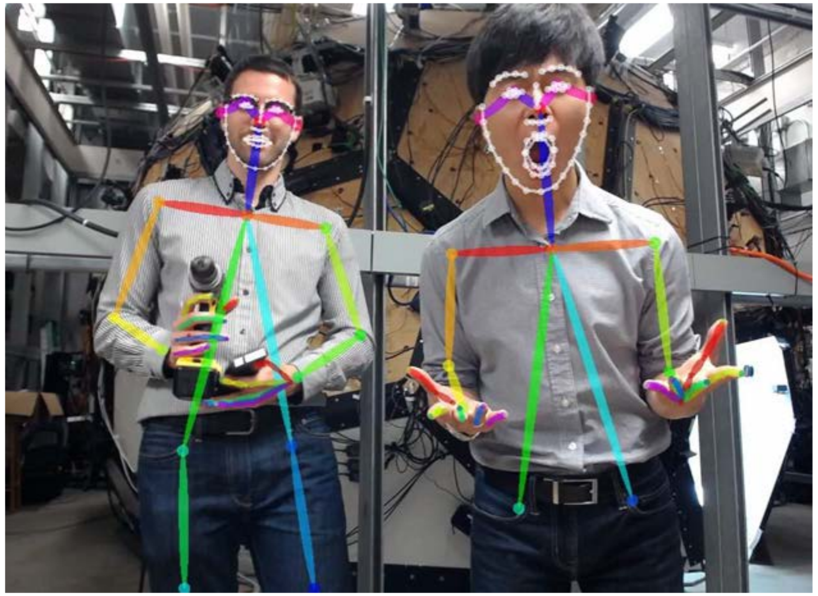
RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor



Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 5, 1–14. DOI: https://doi.org/10.1145/3411764.3445367

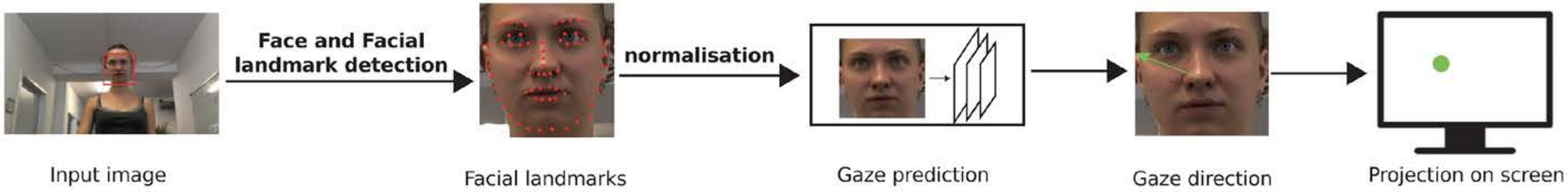
Human Pose Detection

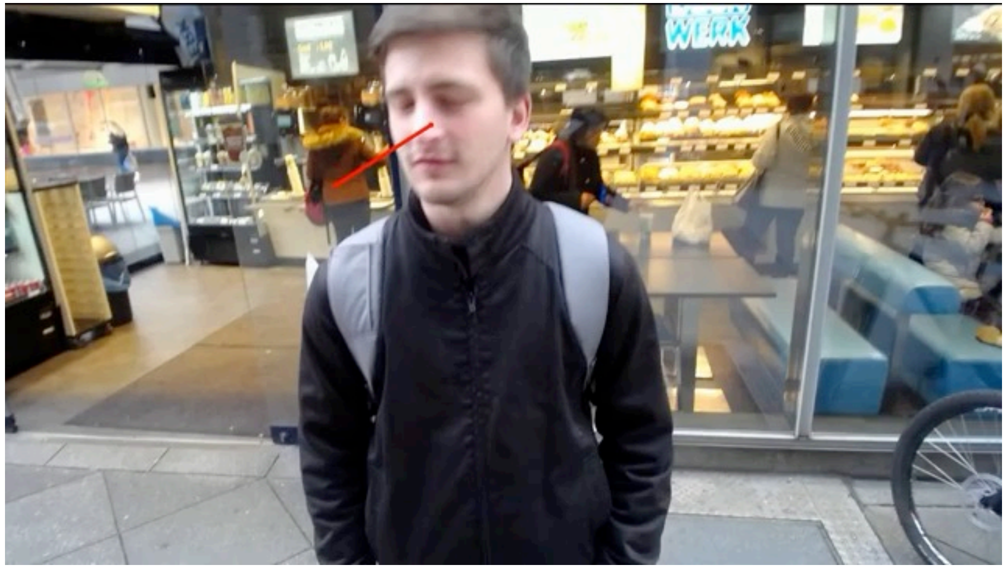
- Keypoint Estimation
- Multi-stage CNN



Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. IEEE transactions on pattern analysis and machine intelligence 43, no. 1 (2019): 172-186. DOI: https://doi.org/10.1109/TPAMI.2019.2929257

Camera based Gaze Estimation



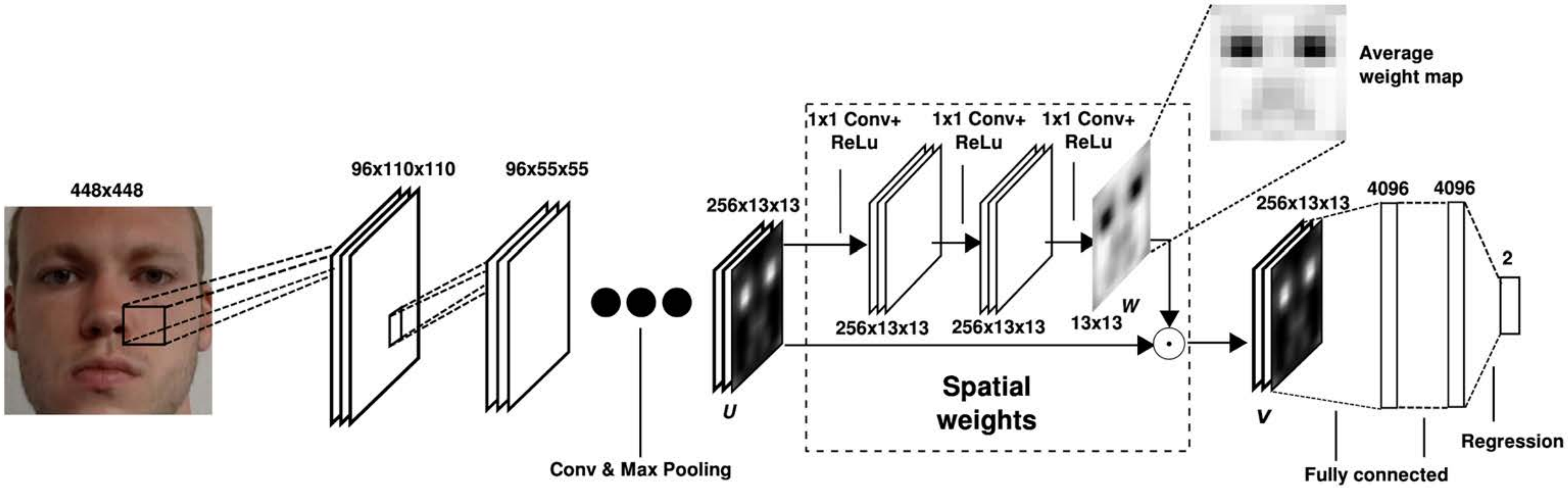


https://youtu.be/9Lujg3beiYI

Xucong Zhang, Yusuke Sugano, and Andreas Bulling. 2019. Evaluation of Appearance-Based Methods and Implications for Gaze-Based Applications. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). ACM, New York, NY, USA, Paper 416, 1–13. DOI: https://doi.org/10.1145/3290605.3300646 URL: http://www.opengaze.org/

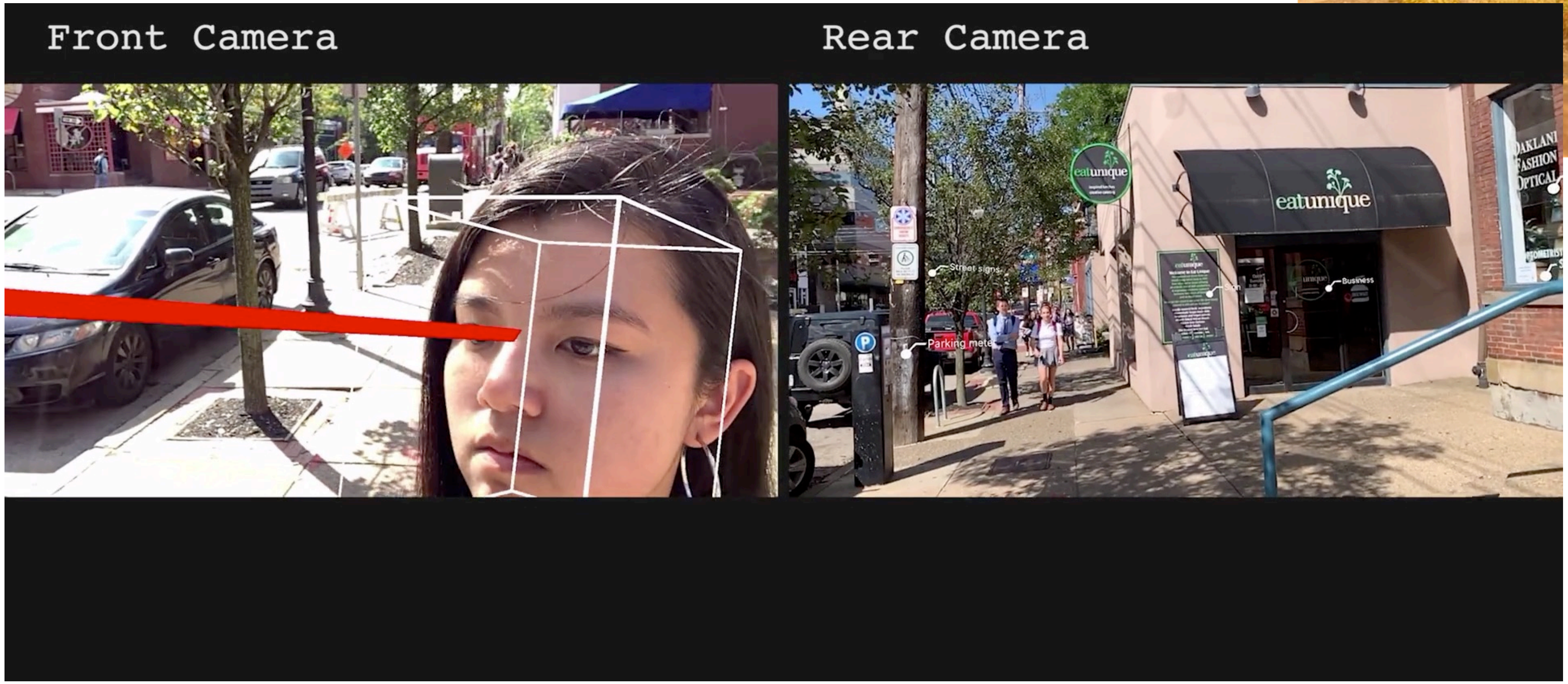
Camera based Gaze Estimation

- Convolutional Neural Network
- Regression – x/y cornindates



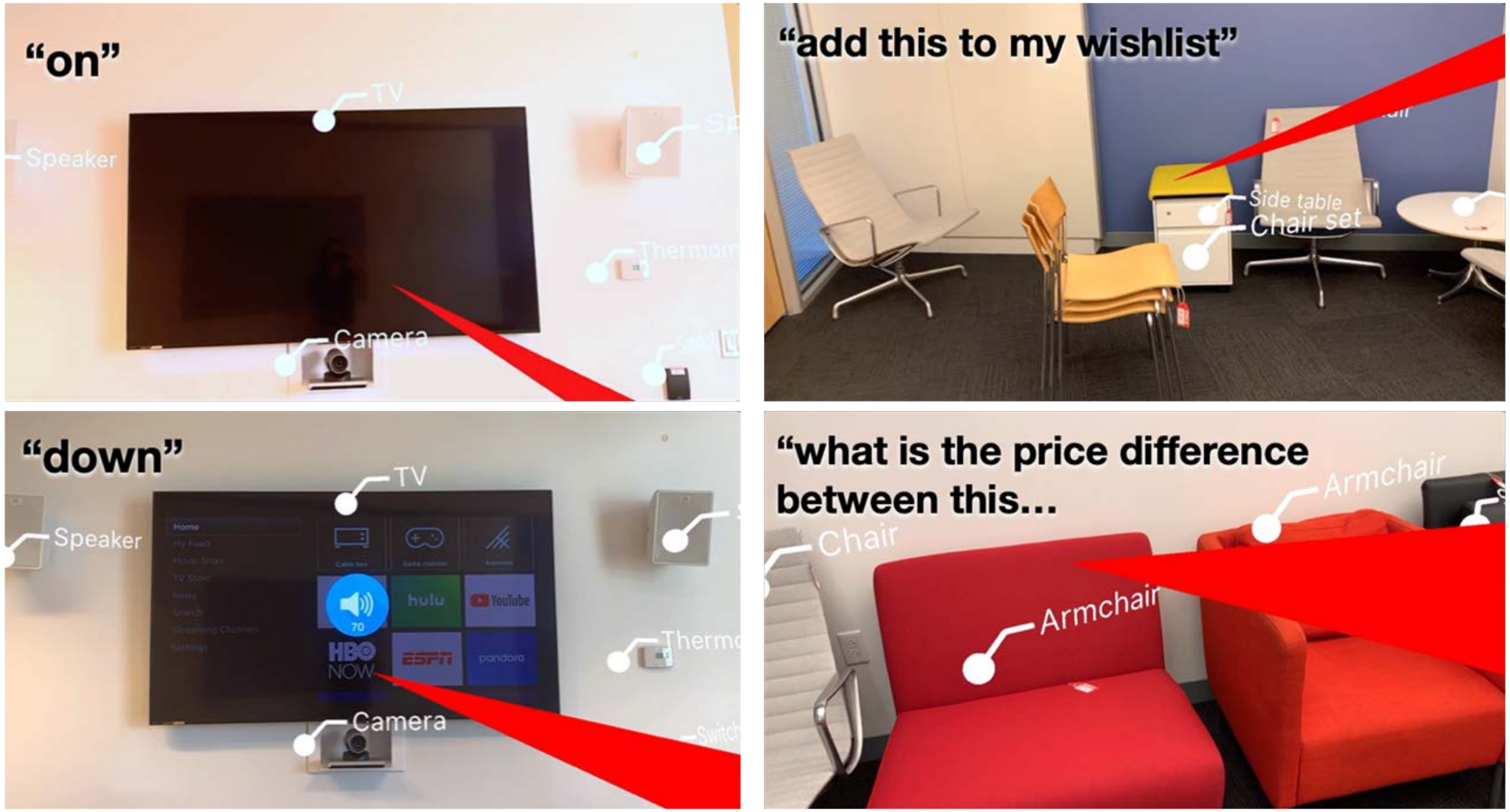
Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2017. It's written all over your face: Full-face appearance-based gaze estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 51-60. 2017. DOI: https://doi.org/10.1109/CVPRW.2017.284

Enhanced Voice Assistants



Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing Mobile Voice Assistants with WorldGaze. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–10. DOI: https://doi.org/10.1145/3313831.3376479

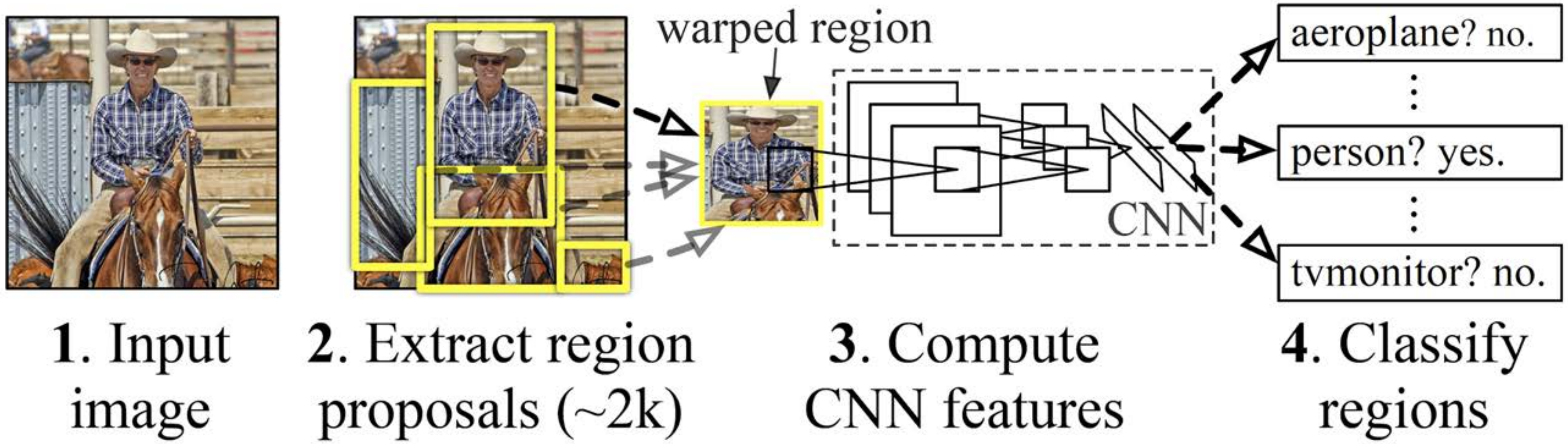
Object Detection



Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing Mobile Voice Assistants with WorldGaze. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–10. DOI: https://doi.org/10.1145/3313831.3376479

Object Detection

R-CNN: Regions with CNN feature

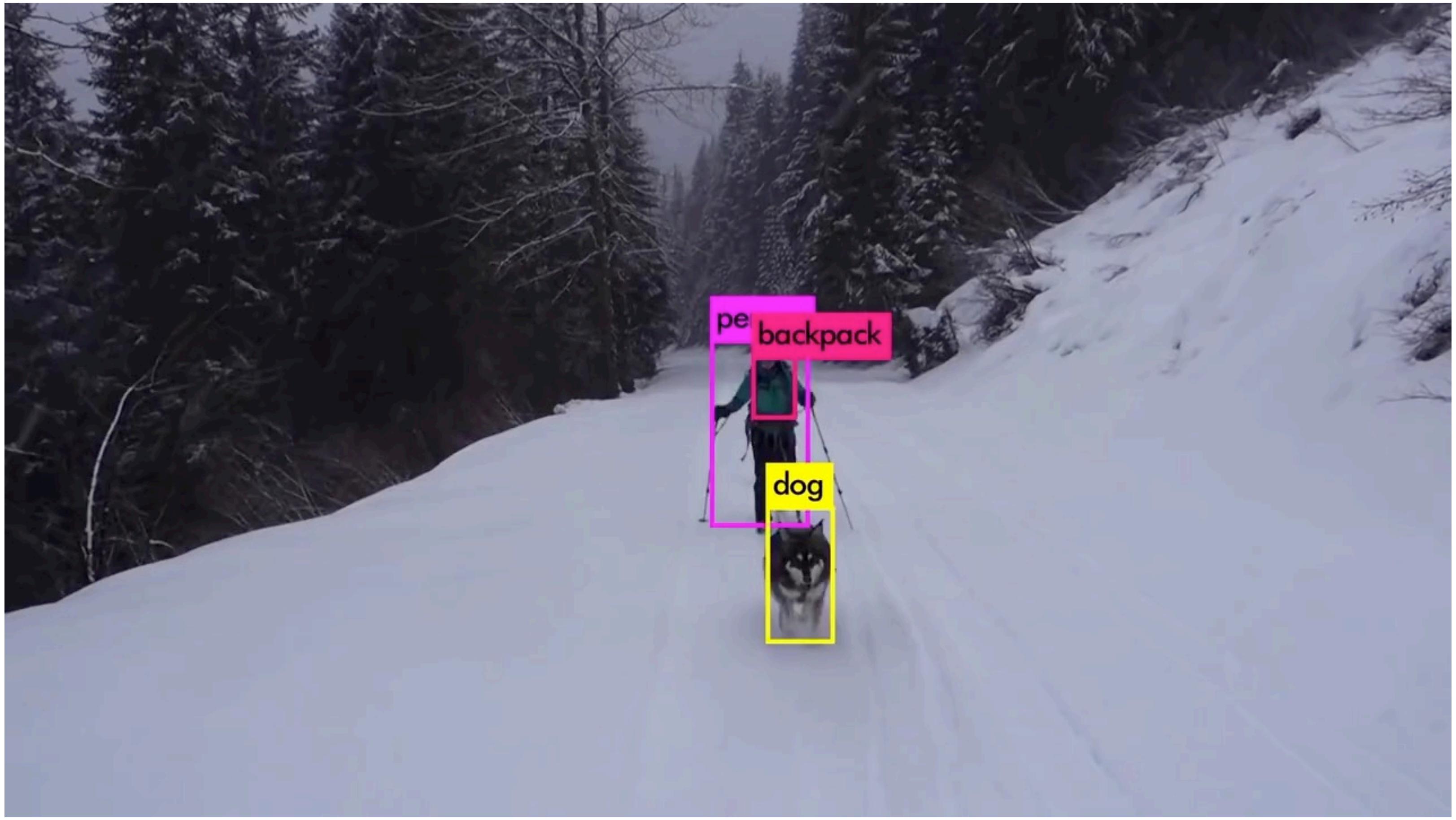


Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation (2014) In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587. DOI: https://doi.org/10.1109/CVPR.2014.81 Source Code: https://github.com/rbgirshick/rcnn

Object Detection

YOLO – You Only Look Once

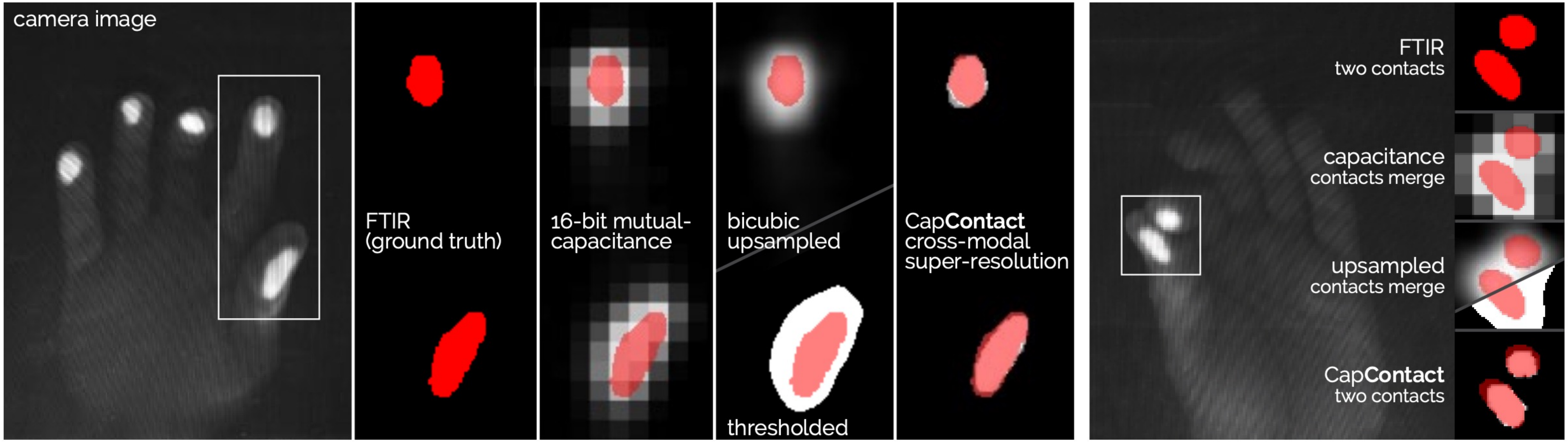
https://youtu.be/MPU2Histivl



Joseph Redmon, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
URL: https://pjreddie.com/darknet/yolo/

Data Generation: GAN Super-Resolution

CapContact: Super-resolution Contact Areas from Capacitive Touchscreens



Paul Streli and Christian Holz. 2021. CapContact: Super-resolution Contact Areas from Capacitive Touchscreens. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 289, 1–14. DOI: https://doi.org/10.1145/3411764.3445621

Data Generation: GAN Super-Resolution

CapContact: Super-resolution Contact Areas from Capacitive Touchscreens

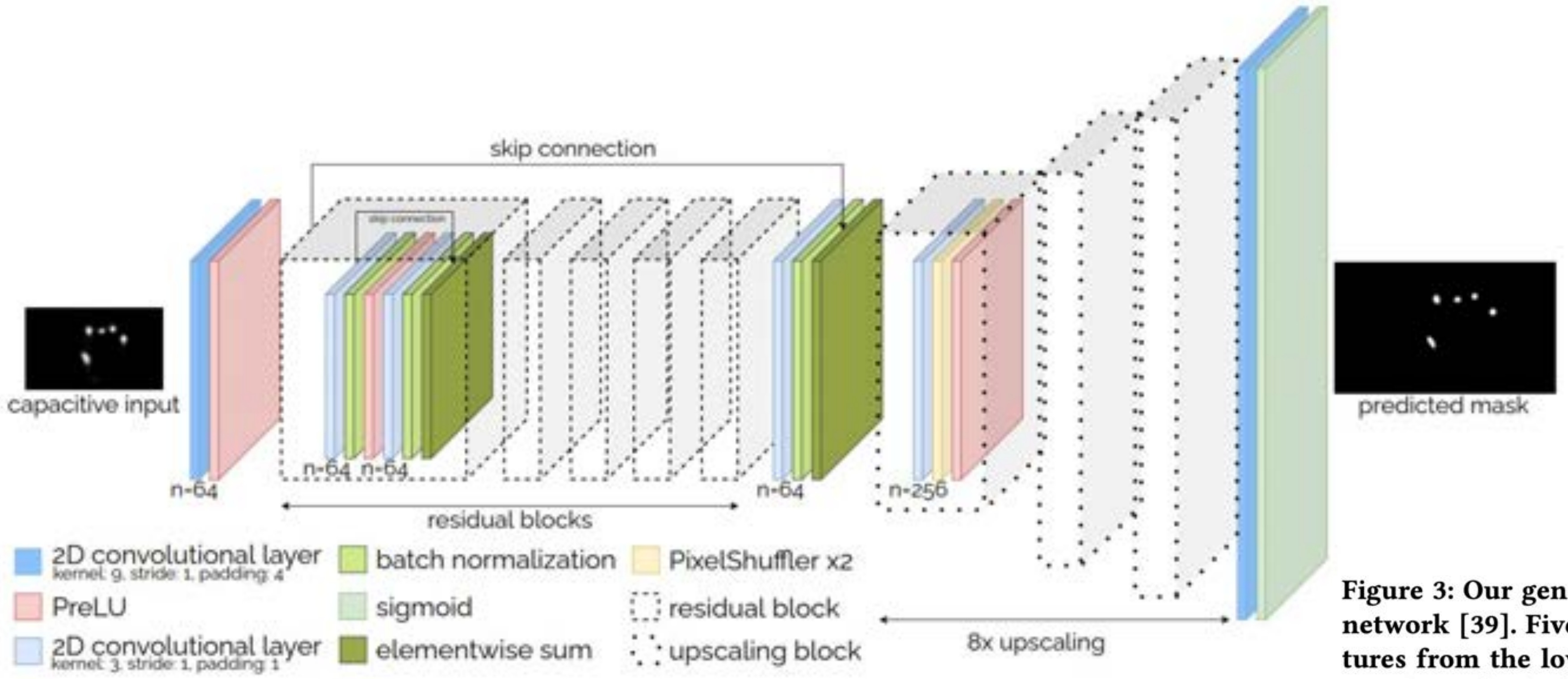
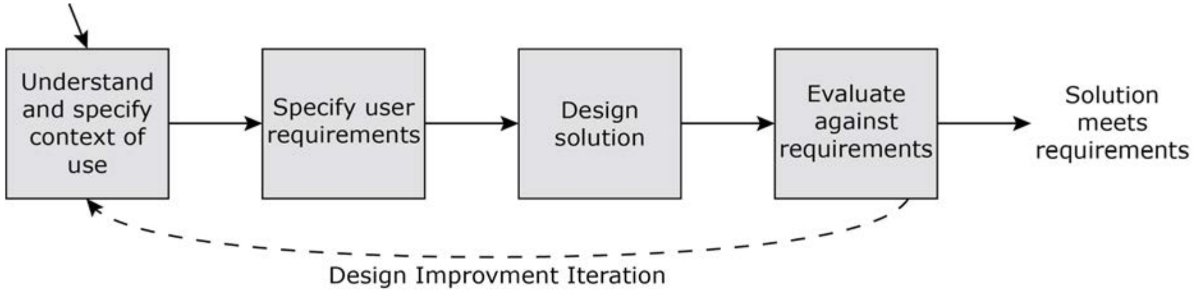


Figure 3: Our generator is a VGG-style fully-convolutional network [39]. Five residual blocks extract the relevant features from the low-resolution capacitive input. Three consecutive sub-pixel convolution layers then upsample it 8×.

Paul Streli and Christian Holz. 2021. CapContact: Super-resolution Contact Areas from Capacitive Touchscreens. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, Article 289, 1–14. DOI: https://doi.org/10.1145/3411764.3445621

Machine Learning for HCI

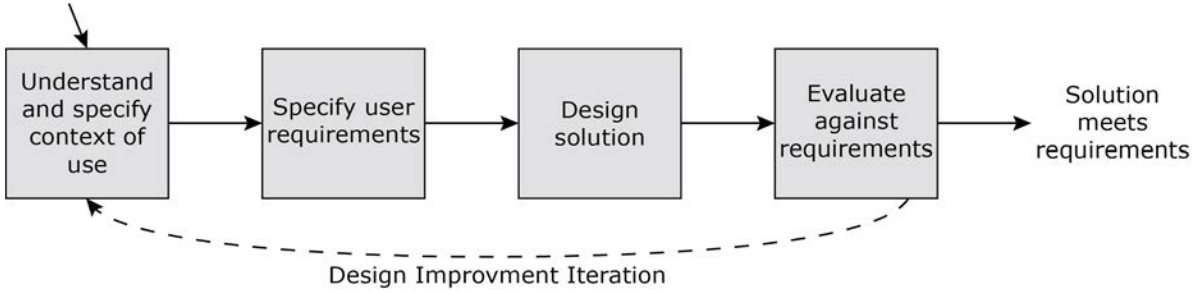
User Centered Design Cycle ISO 9241



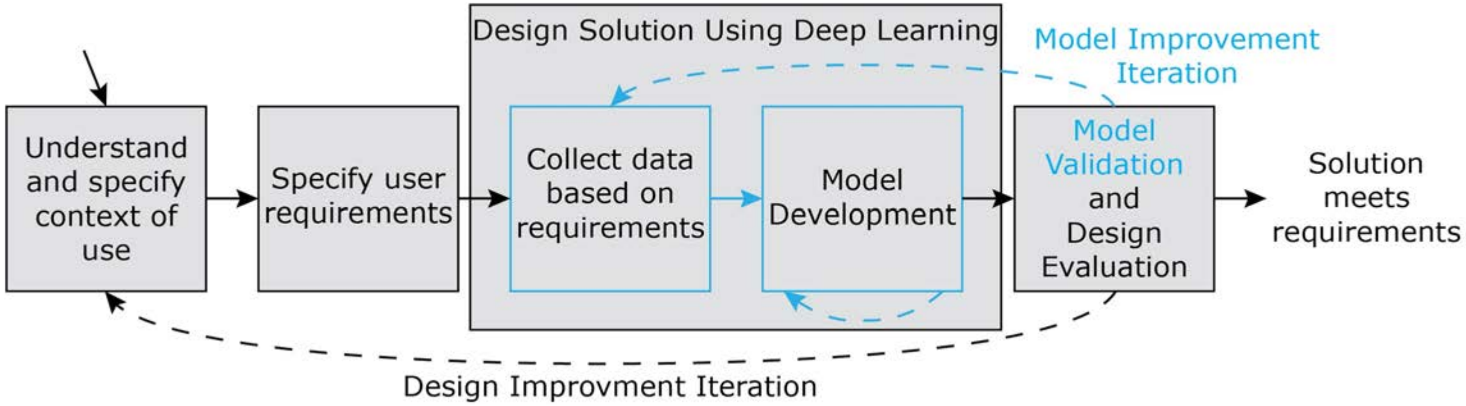
Huy Viet Le, Sven Mayer, and Niels Henze. 2020. Deep learning for human-computer interaction. interactions 28, 1 (January - February 2021), 78–82. DOI: https://doi.org/10.1145/3436958

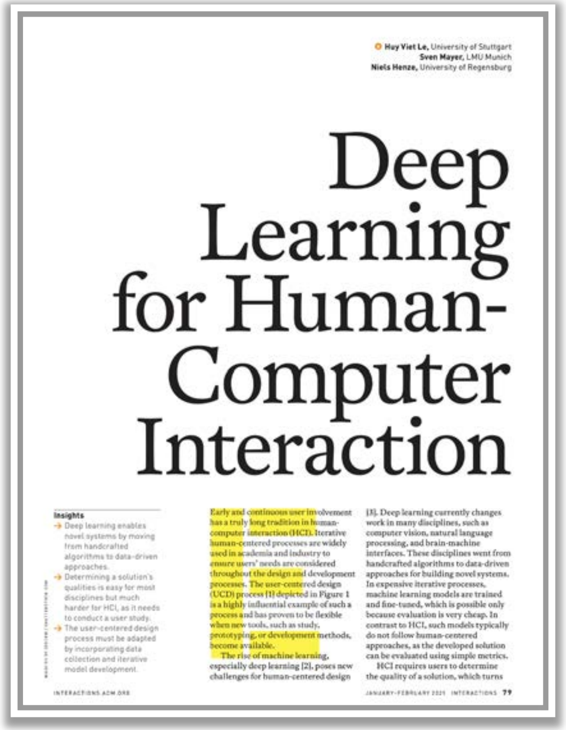
Machine Learning for HCI

User Centered Design Cycle ISO 9241



Adaptation for Machine Learning




Huy Viet Le, University of Stuttgart
Sven Mayer, LMU Munich
Niels Henze, University of Regensburg

Deep Learning for Human-Computer Interaction

Huy Viet Le, Sven Mayer, and Niels Henze. 2020. Deep learning for human-computer interaction. interactions 28, 1 (January - February 2021), 78–82. DOI: https://doi.org/10.1145/3436958

License

This file is licensed under the Creative Commons Attribution-Share Alike 4.0 (CC BY-SA) license:

https://creativecommons.org/licenses/by-sa/4.0

Attribution: Albrecht Schmidt, Sven Mayer, and Daniel Buschek