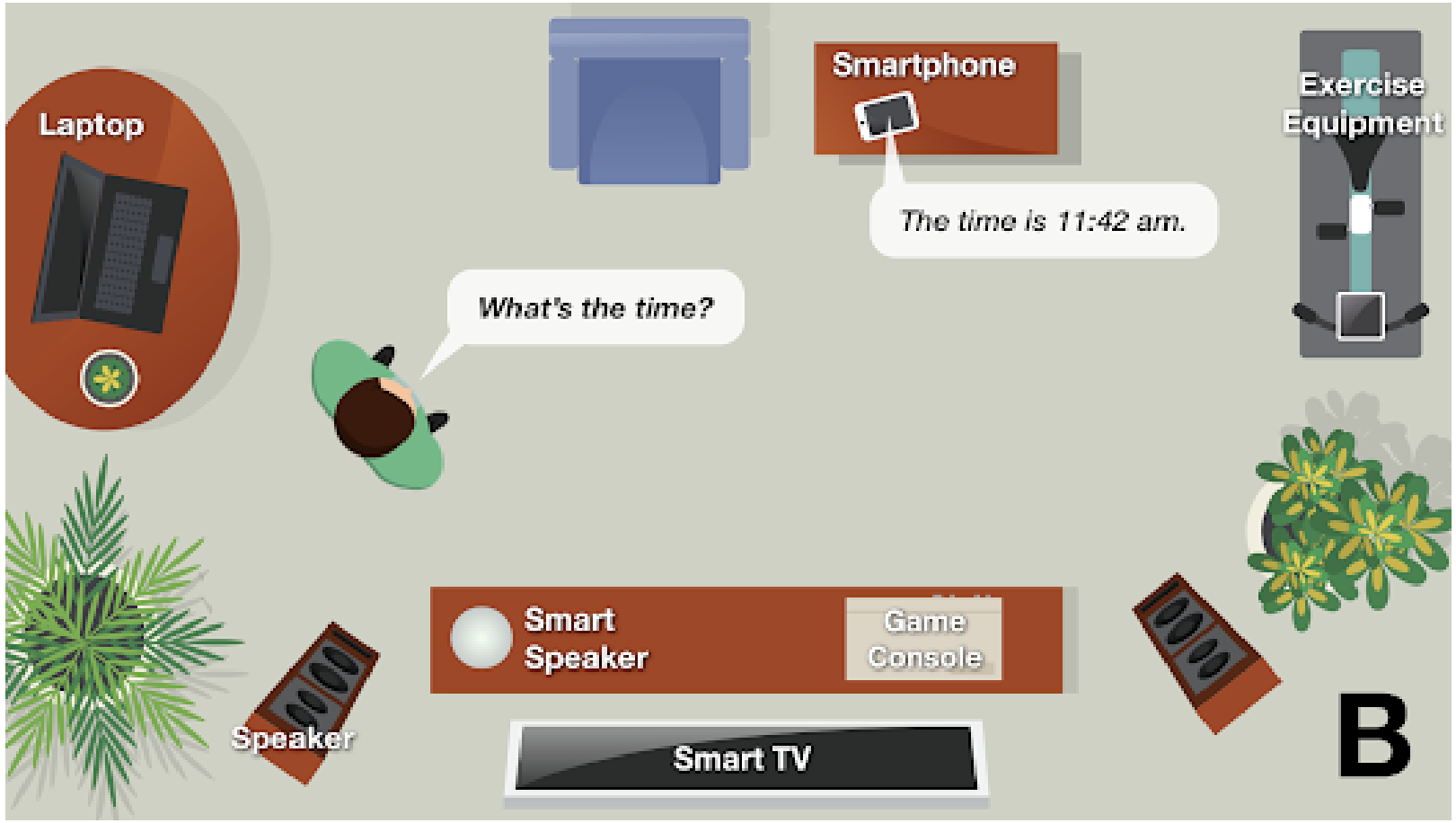
Introduction to Intelligent User Interfaces

Context Awareness Interaction in Smart Environments

Context Awareness Interaction



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

Breakout Sessions

- List as many **addition information** which would support am intelligent user interface?
- Scenario 1: Touch Screen Interaction
- Scenario 2: Navigation
- Scenario 3: Movie
- Scenario 4: Voice activated lights in a living room

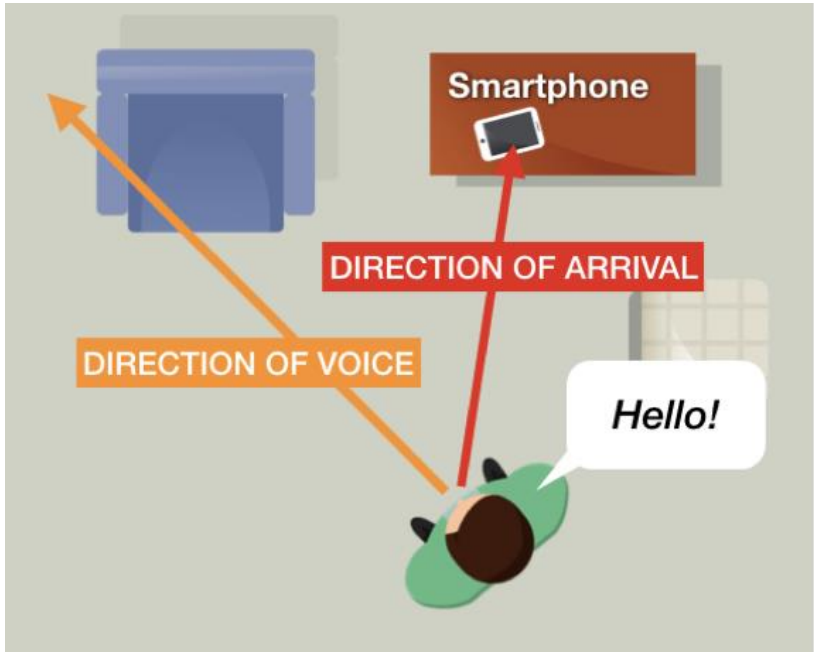
5 min

Extracting Contextual Information

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

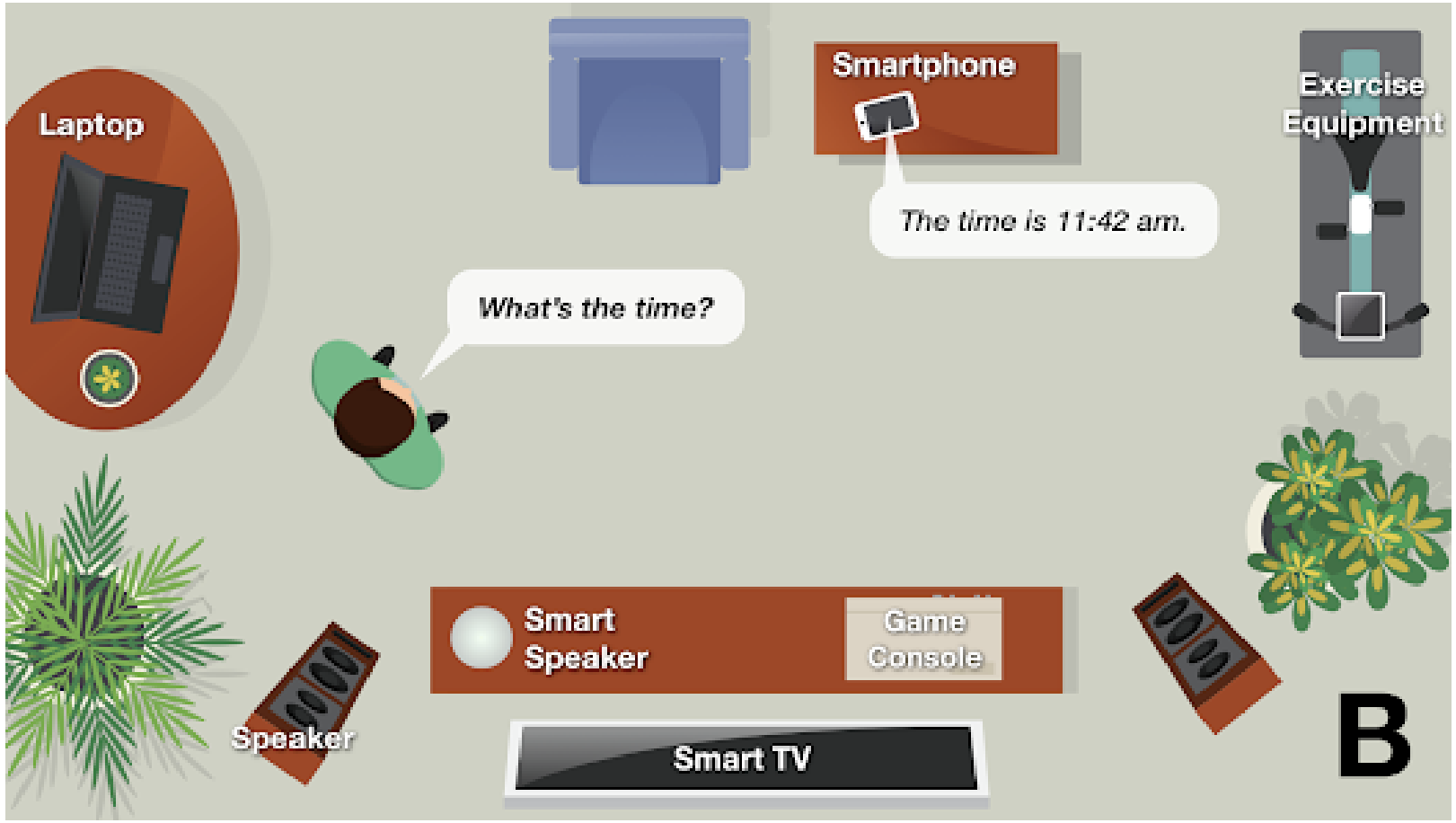
Smart Environments - Direction-of-Voice

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



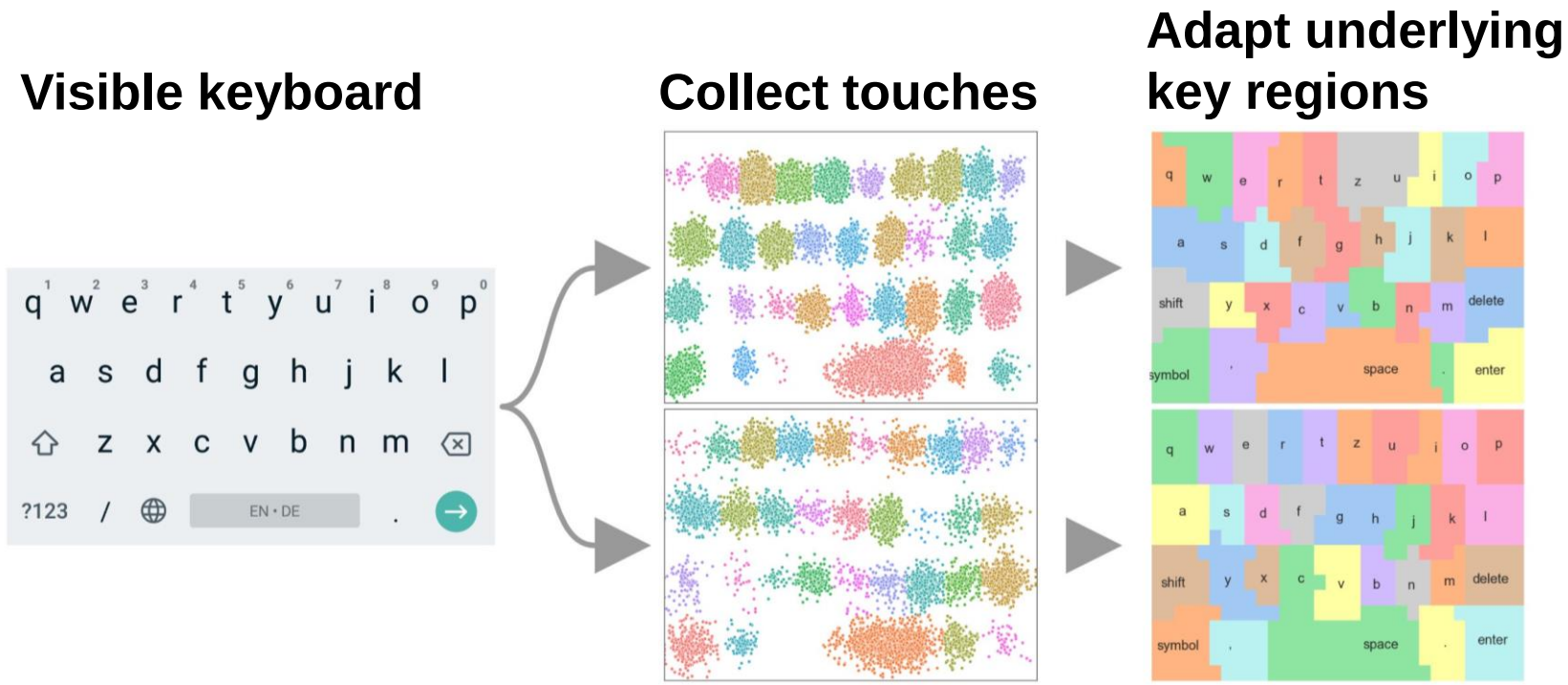
Karan Ahuja, Andy Kong, Mayank Goel, and Chris Harrison. 2020. Direction-of-Voice (DoV) Estimation for Intuitive Speech Interaction with Smart Devices Ecosystems. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20). ACM, New York, NY, USA, 1121–1131. DOI: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

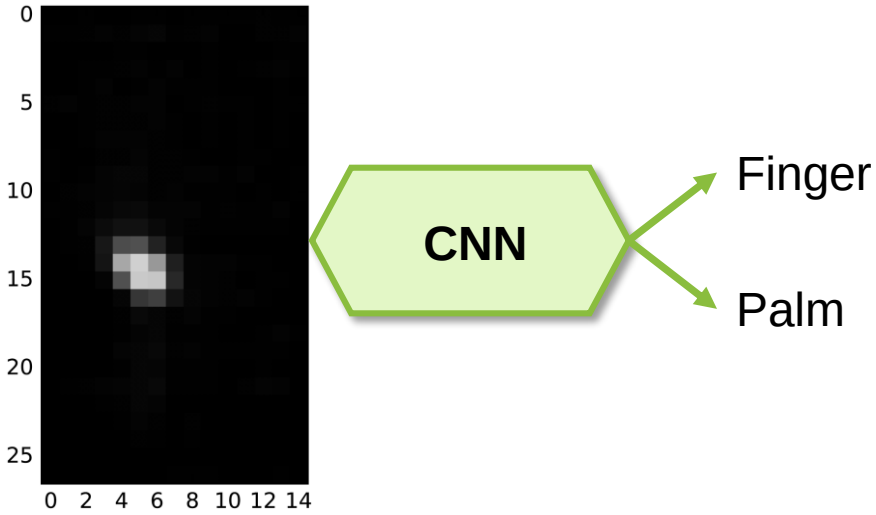
Context-Aware Keyboards



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

Palm Detection

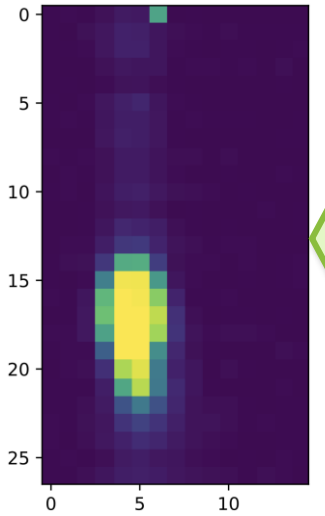
- Convolutional Neural Network
- Classification
- Representation Learning

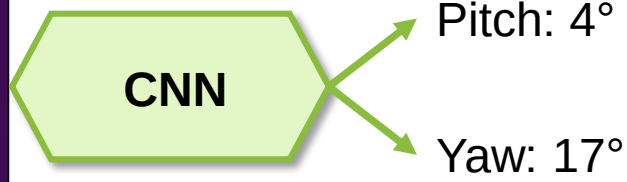


Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. 2018. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Paper 360, 1–13. DOI: https://doi.org/10.1145/3173574.3173934

Finger Orientation

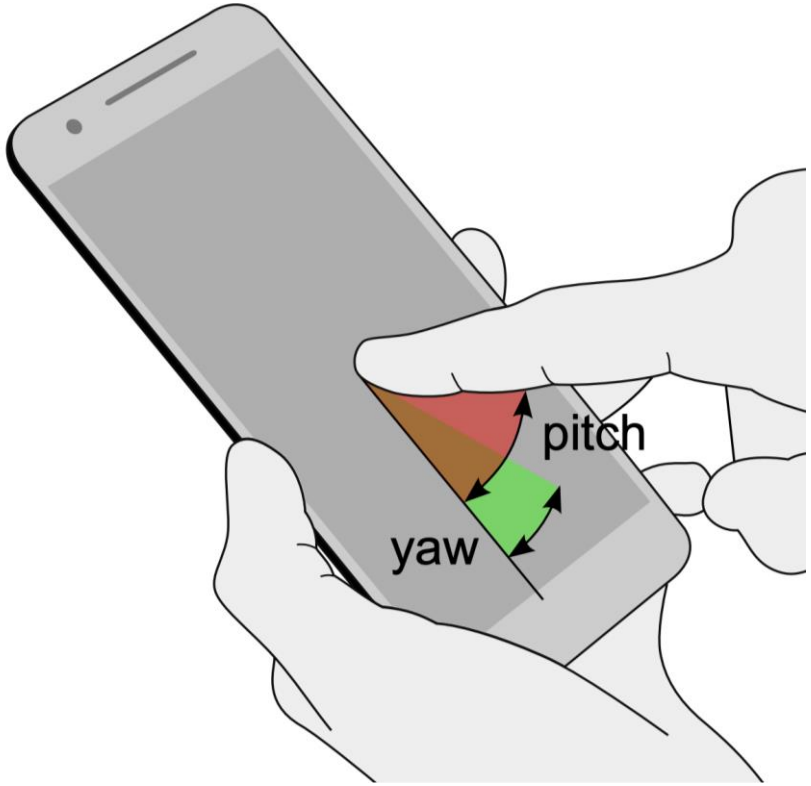
- Convolutional Neural Network
- Regression
- Representation Learning





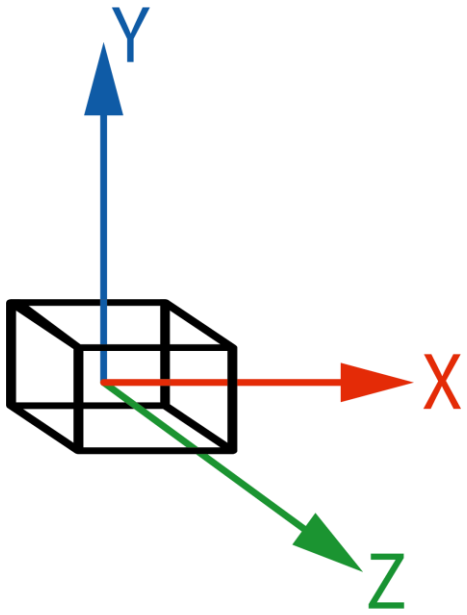
Pitch: 4°

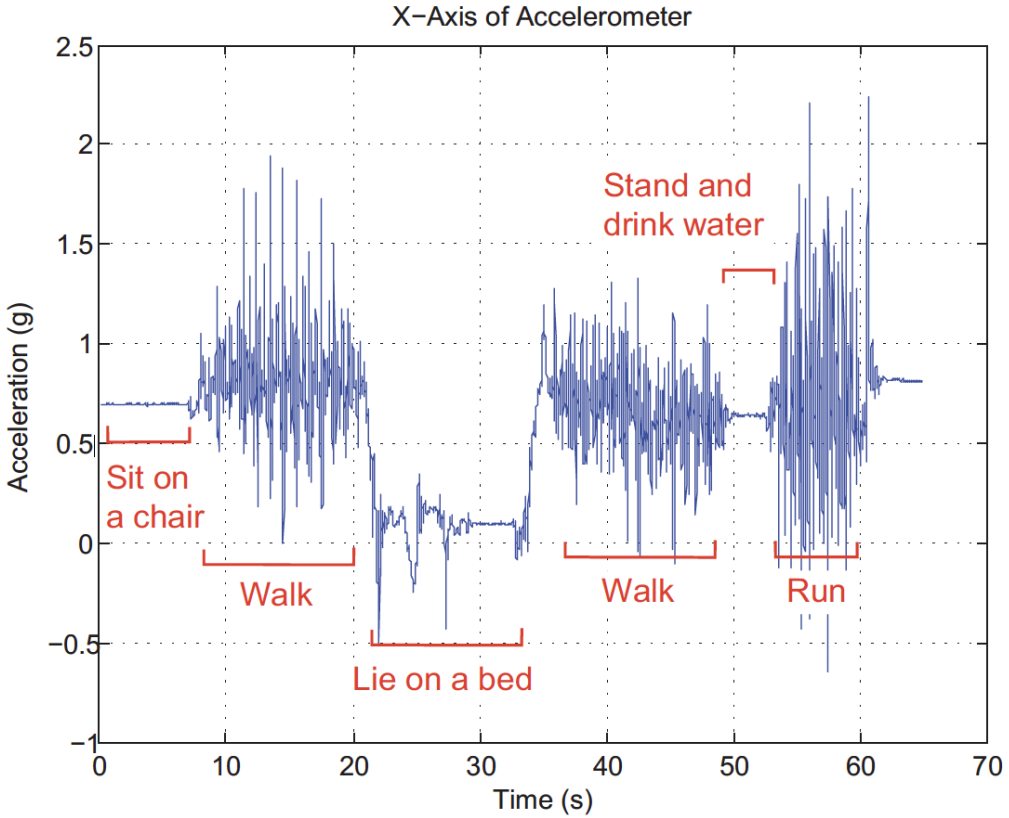
Yaw: 17°



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

Activity Recognition – Accelerometers



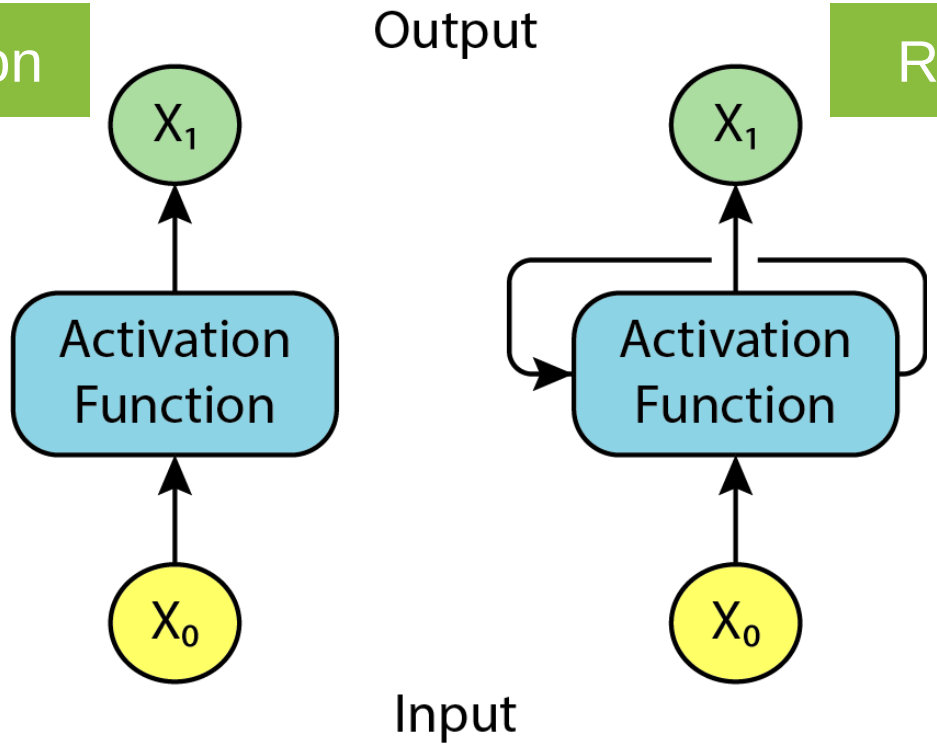


Mi Zhang and Alexander A. Sawchuk. 2012. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12). Association for Computing Machinery, New York, NY, USA, 1036–1043. DOI: https://doi.org/10.1145/2370216.2370438

Neuronal Network With Timeseries Data

Recurrent Neural Network

Perceptron



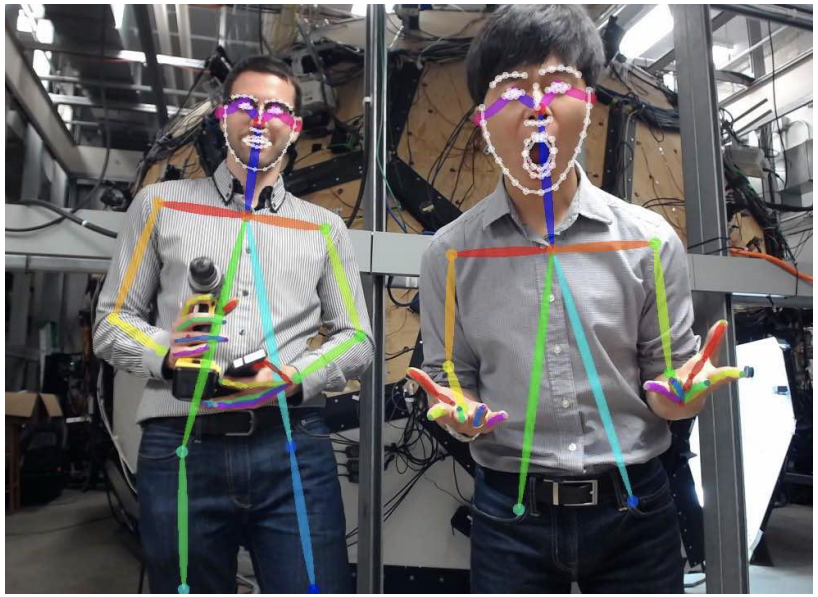
RNN Unit

Side note:
Long short-term memory (**LMST**) is an advanced RNN

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

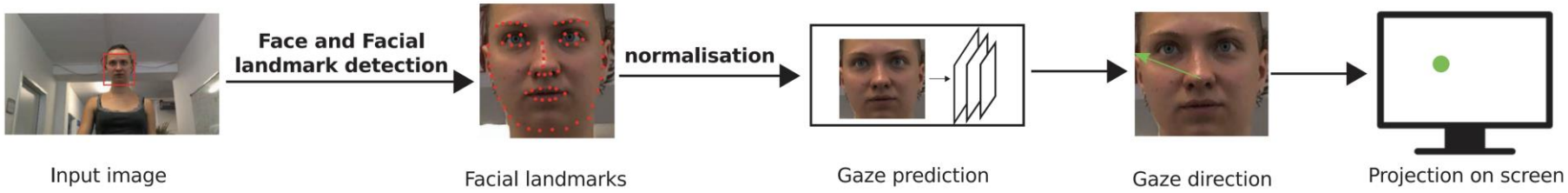
Human Pose Detection

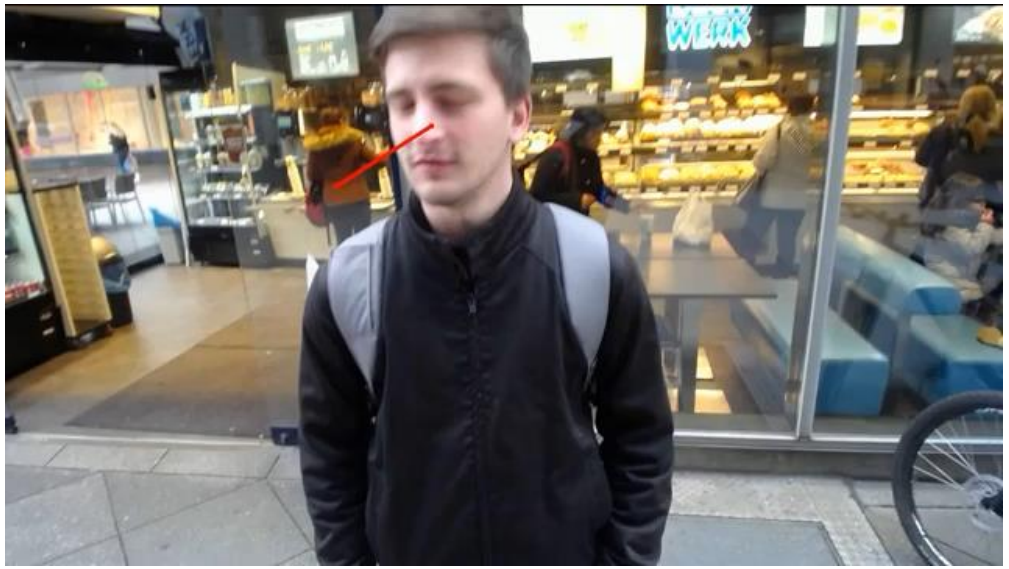
- Keypoint Estimation
- Multi-stage CNN



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



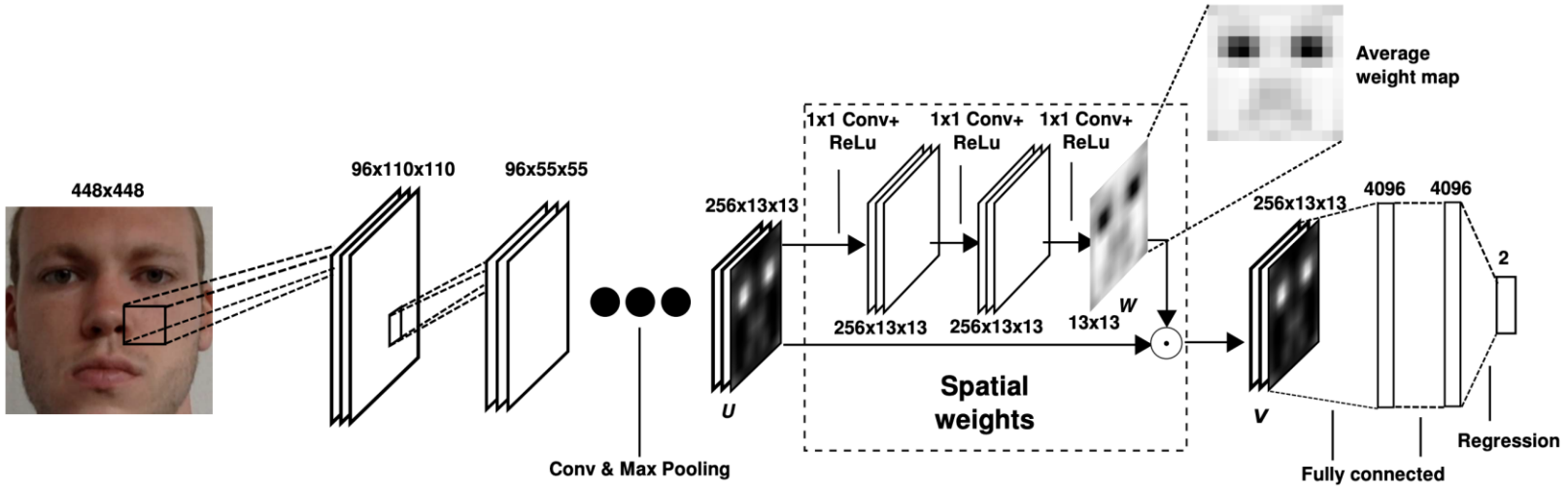


https://youtu.be/9Lujg3beiYI

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/

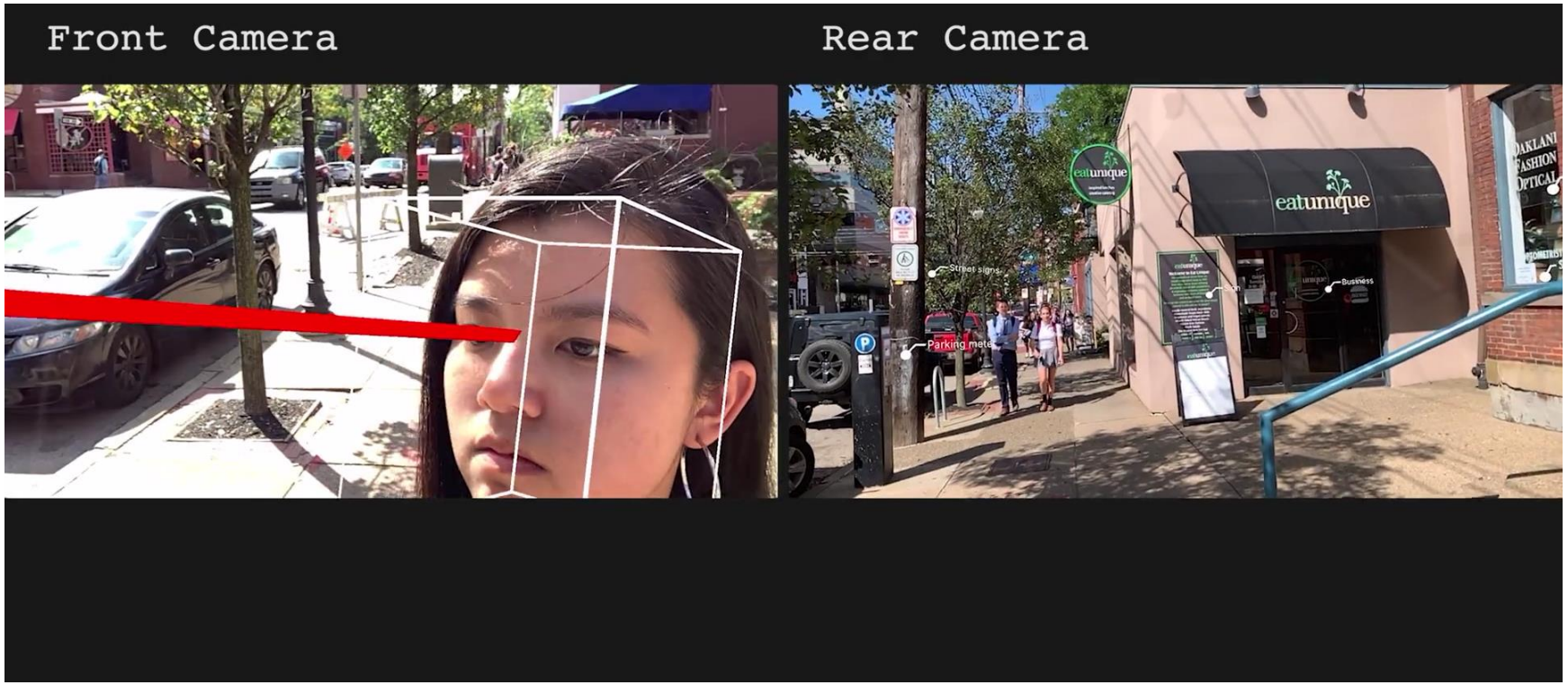
Camera based Gaze Estimation

- Convolutional Neural Network
- Regression – x/y cornindates



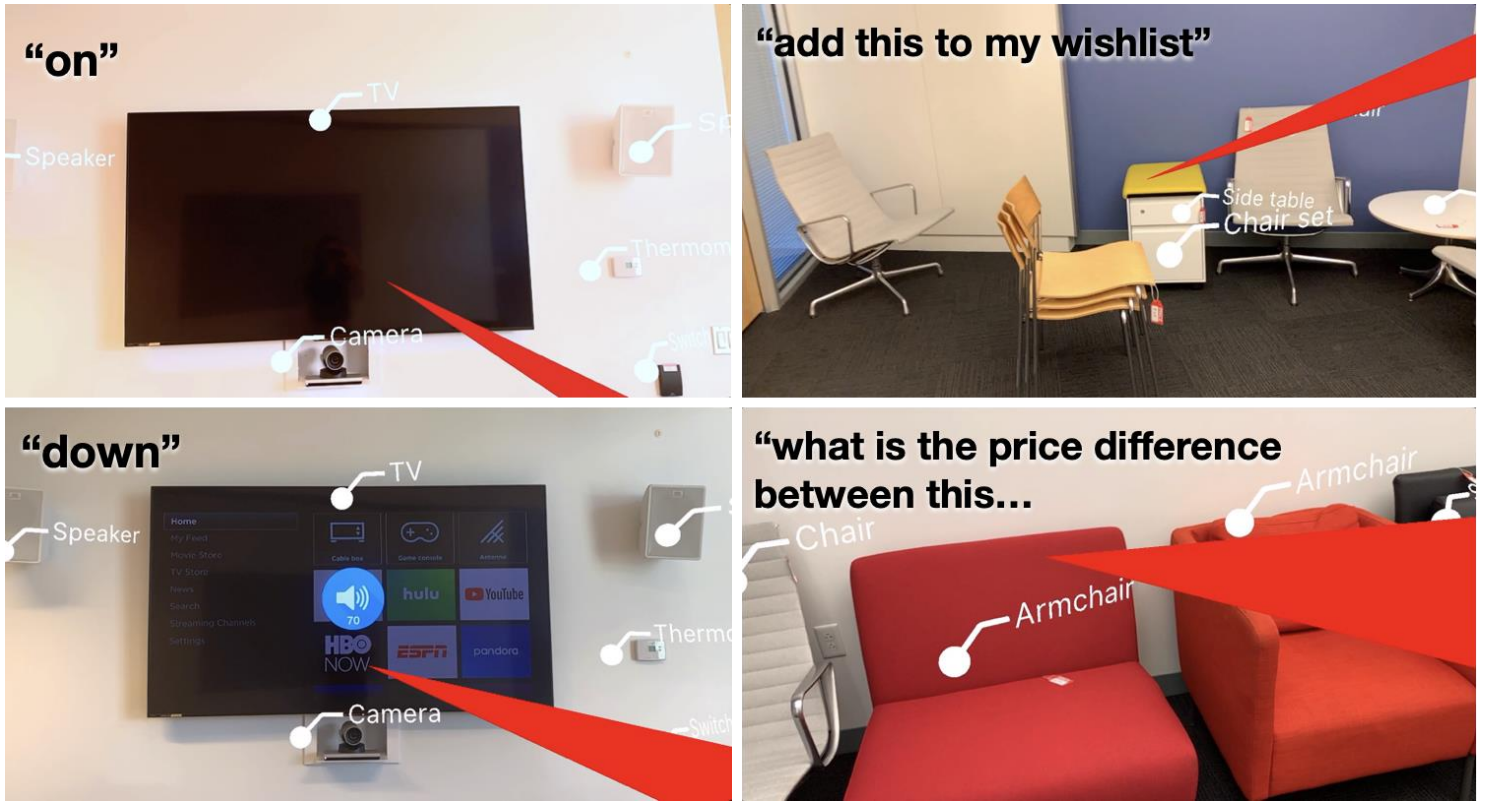
Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2017. It's written all over your face: Full-face appearance-based gaze estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 51-60. 2017. DOI: https://doi.org/10.1109/CVPRW.2017.284

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

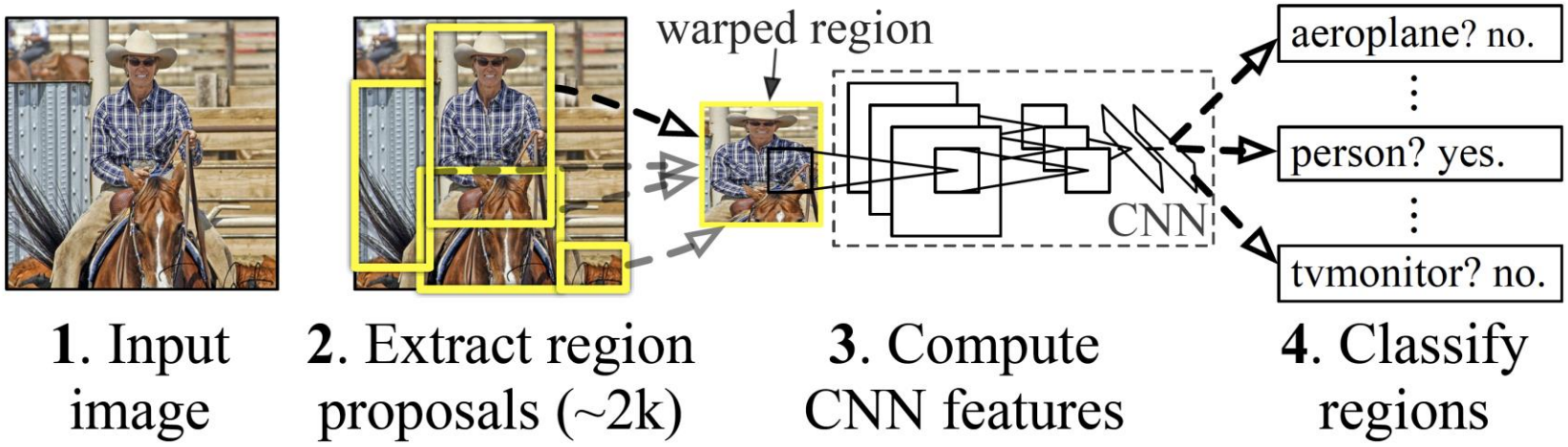
Object Detection



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

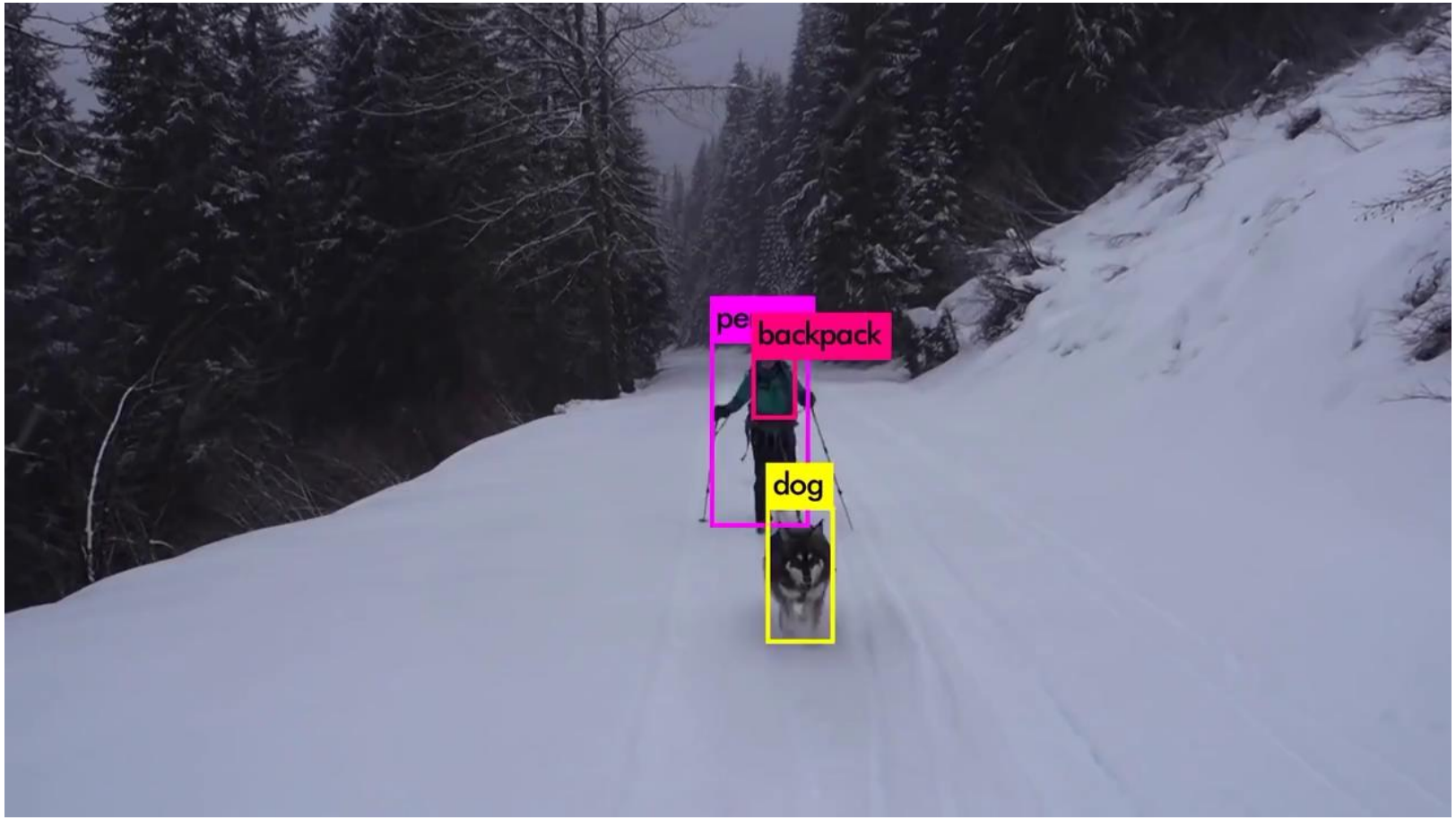
R-CNN: Regions with CNN feature



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

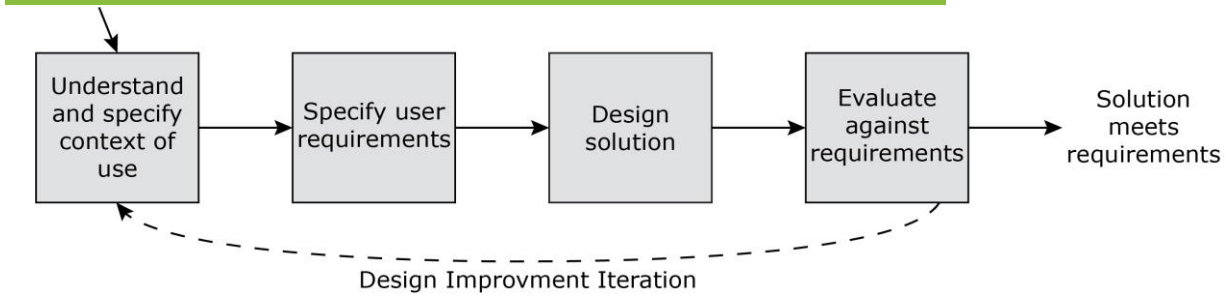


https://youtu.be/MPU2HistivI

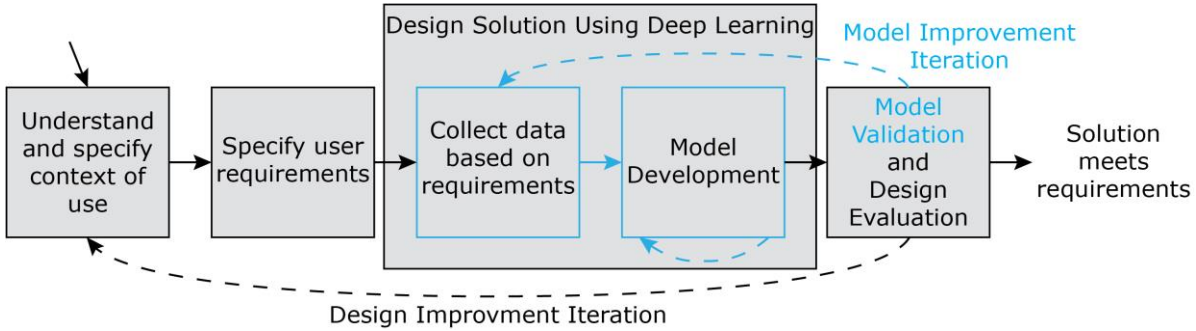
Joseph Redmon, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
URL: https://pjreddie.com/darknet/yolo/

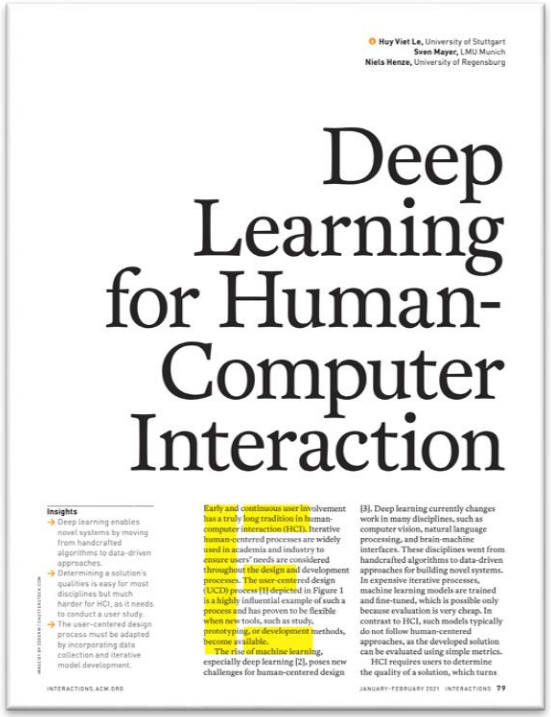
Machine Learning for HCI

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

Insights

- Deep learning enables novel systems by moving from handcrafted algorithms to data-driven approaches.
- Determining a solution's qualities is easy for most disciplines but much harder for HCI, as it needs to conduct a user study.
- The user-centered design process must be adapted by incorporating data collection and iterative model development.

Early and continuous user involvement has a truly long tradition in human-computer interaction (HCI). Iterative human-centered processes are widely used in academia and industry to ensure users' needs are considered throughout the design and development processes. The user-centered design (UCD) process [1] depicted in Figure 1 is a highly influential example of such a process and has proven to be flexible when new tools, such as study, prototyping, or development methods, become available.

The rise of machine learning, especially deep learning [2], poses new challenges for human-centered design [3]. Deep learning currently changes work in many disciplines, such as computer vision, natural language processing, and brain–machine interfaces. These disciplines went from handcrafted algorithms to data-driven approaches for building novel systems. In expensive iterative processes, machine learning models are trained and fine-tuned, which is possible only because evaluation is very cheap. In contrast to HCI, such models typically do not follow human-centered approaches, as the developed solution can be evaluated using simple metrics.

HCI requires users to determine the quality of a solution, which turns

INTERACTIONS.ACM.ORG JANUARY–FEBRUARY 2021 INTERACTIONS 79

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

Neuronal Network Concepts

Introduced concepts

- Classification vs. Regression
- Feature Extraction vs. Representation Learning
- Model Structures
  - Neuronal Network
  - Convolutional Neural Network
  - RNN
  - Long short-term memory (LSTM)