2022 CSDI Workshop
Paris, 4 April 2022

SHARE
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

Automatic Scoring of Cognition Drawings

Arne Bethmann
SHARE Germany
TU Munich / MEA MPISOC

mea MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY
Munich Center for the Economics of Aging

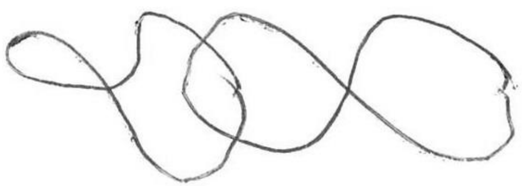
This project has received funding from the European Union under grant agreement SOCPL No 101052589 and the European Union's Horizon 2020 research and innovation programme under grant agreements No 870628, No 101015924.

SPONSORED BY THE
Federal Ministry of Education and Research

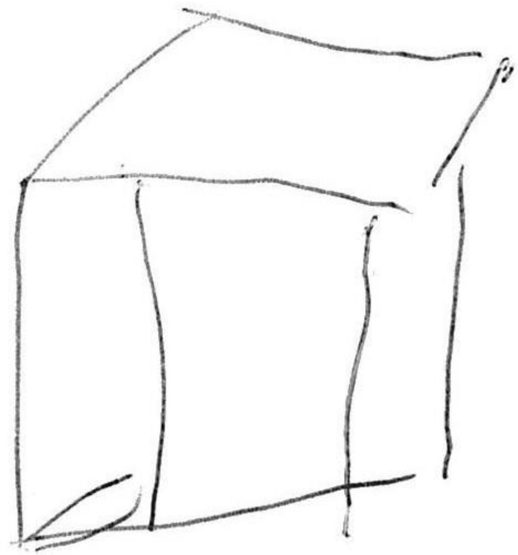
NIH National Institute on Aging

Drawing Exercises

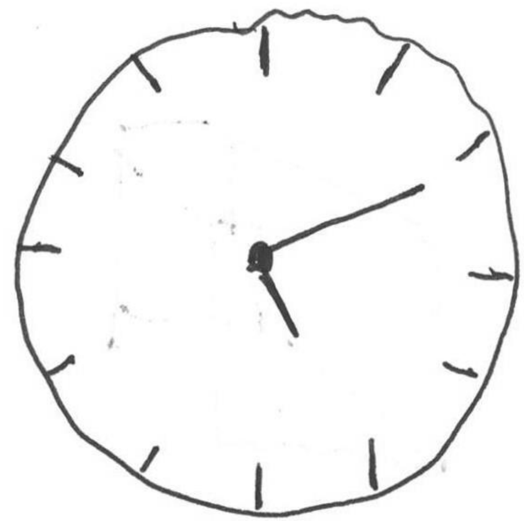
- Since wave 8 SHARE contains three „constructional praxis tests“ (Wagner & Douhou, 2021)



Infinity loops



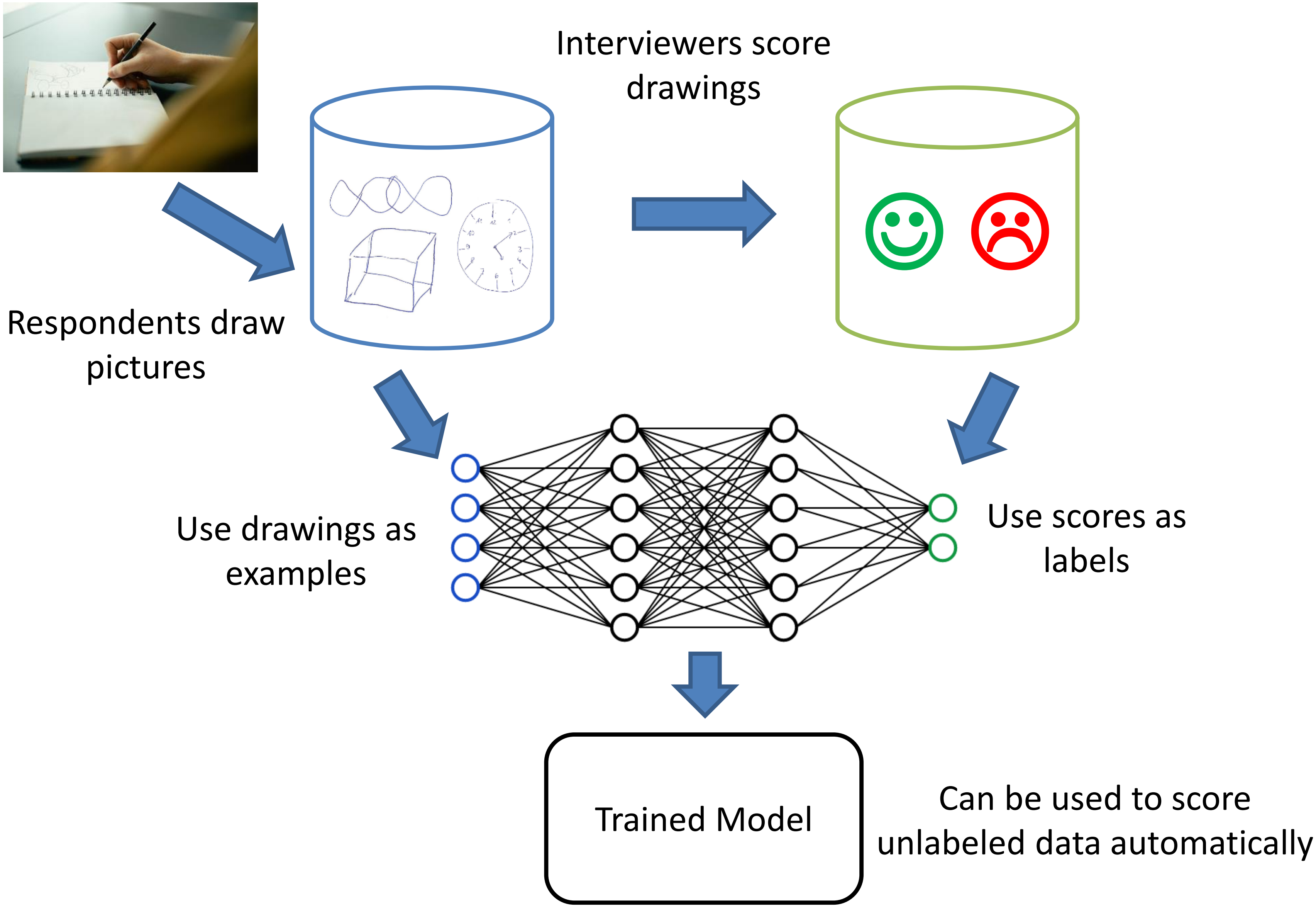
Cube



Clock

- Part of several screening tests (e.g. MoCA) for early signs of cognitive decline

Overview



Interviewers score drawings
Respondents draw pictures
Use drawings as examples
Use scores as labels
Trained Model
Can be used to score unlabeled data automatically

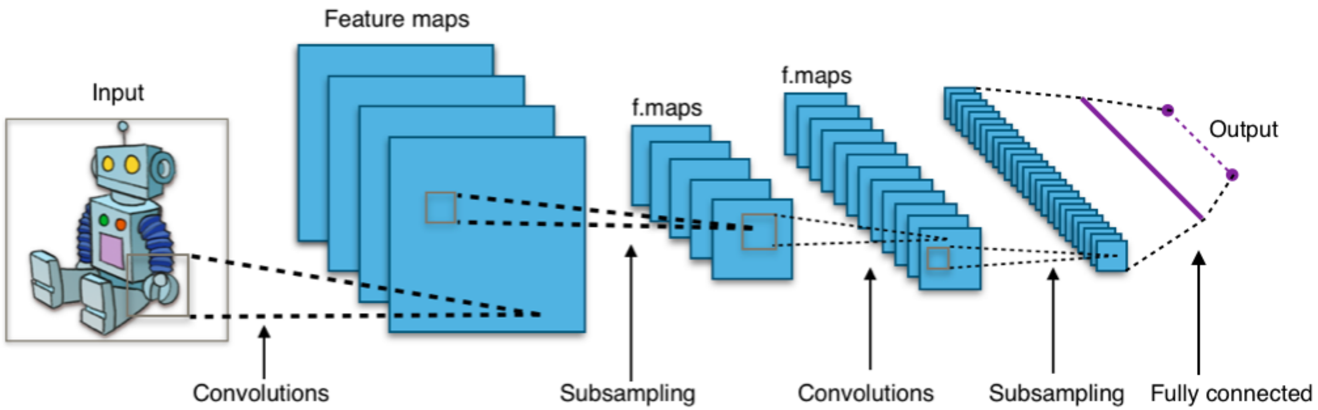
Progress so far

- Scanned 2,109 SHARE DE recording booklets
- Linked to survey data
- Processed and enhanced images

Thank you, Amany, Julia, Pandora, Claudia and Charlotte!

- Collected ~10,000 recording booklets from 9 additional countries

- Convolutional Neural Networks
  - Work horse for image recognition tasks
  - Initially used to digitize handwritten numbers (LeCun et al. 1989)
  - Made dramatical improvements in recent years and is the basis of the Deep Learning „hype“



Preliminary results

			% classified correctly		
Model architecture	Layers	Trainable parameters (millions)	**Clock / Cube / Loops** 3 classes	**Cube: (partially) correct / wrong** 2 classes	**Cube: correct / partially / wrong** 3 classes
AlexNet[1]	8	62	99.8	84.9	68.4
VGG (BN)[2]	19	144	99.8	86.6	71.2
ResNet[3]	18	11	99.8	86.7	70.3
	50	26	99.8	86.2	68.4
			N_{Clock} = 1,907 N_{Cube} = 1,882 N_{Loops} = 1,891	$N_{(Partially)\ Correct}$ = 1,417 (80.5%) N_{Wrong} = 343	$N_{Correct}$ = 1,088 (61.8%) $N_{Partially}$ = 329 N_{Wrong} = 343

Best out of 20 epochs; Validation set = 30%; All models pretrained on ImageNet (https://www.image-net.org/)
[1] Krizhevsky, Sutskever & Hinton (2012); [2] Simonyan & Zisserman (2015); [3] He et al. (2015)
Software: PyTorch 1.10.0, fastai 2.5.3; Hardware: NVIDIA GeForce RTX 3070 Ti

Ways to improve

- # Fiddle with models

  longer training, hyperparameter tuning, newer architectures, class imbalances, data augmentation, better pre-training datasets …

- # Get more data

  scan available recording booklets, collect wave 9 drawings

- # Improve scoring

  Prediction performance could be impaired by bad scoring due to:

  - Interviewers not sticking to the rules
  - Rules being sub-optimal
  - Scoring not lending itself well to exact rules

  → Investigate error sources and get closer to bayes error rate

Ideas for investigating error sources

- Make „gold standard“ scoring
  - Check performance of interviewers compared to gold standard
  - Use as improved labels for the training data

- Visualize predictive features in drawings

  (cf. e.g. Zeiler & Fergus 2013)
  - Check whether interviewers stick to rules: features indicated by rules shoud be predictive of interviewers‘ scores
  - Check if rules work as intended: try to predict other criteria of cognitive decline with drawings and see if features indicated by rules are predictive

Next steps

- Phase 1 (2020-2022):
  - Scan initial batch of SHARE DE drawings (~2,000)
  - Make proof-of-concept models
  - Assess feasibility and sketch further plans
- Phase 2 (2022-2023):
  - Find funding
  - Scan all available wave 8 drawings (~10,000)
  - First substantial publications
- Phase 3 (2023-2024):
  - Scan wave 9 drawings (~60,000)
  - More data intensive analyses (e.g. clock draws)
  - Release as well curated, strictly anonymous open dataset
  - Release access restricted linked dataset

SHARE

SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

THANK YOU!

With any questions/issues please don't hesitate to contact me:

bethmann@mea.mpisoc.mpg.de

mea | MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

Munich Center for the Economics of Aging

This project has received funding from the European Union under grant agreement SOCPL No 101052589 and the European Union's Horizon 2020 research and innovation programme under grant agreements No 870628, No 101015924.

SPONSORED BY THE

Federal Ministry of Education and Research

NIH National Institute on Aging

References

- He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE. pp. 770–778. arXiv:1512.03385 https://arxiv.org/abs/1512.03385
- Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey E. (2012). "ImageNet classification with deep convolutional neural networks" (PDF). Communications of the ACM. 60 (6): 84–90. doi:10.1145/3065386.
- LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L. D. (December 1989). "Backpropagation Applied to Handwritten Zip Code Recognition". Neural Computation. 1 (4): 541–551. doi:10.1162/neco.1989.1.4.541.
- Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556. https://arxiv.org/abs/1409.1556
- Wagner, M. and Douhou, S. (2021). Cognitive Measures. In: Bergmann, M. and A. Börsch-Supan (Eds.). SHARE Wave 8 Methodology: Collecting Cross-National Survey Data in Times of COVID-19. Munich: MEA, Max Planck Institute for Social Law and Social Policy.
- Zeiler M.D., Fergus R. (2014) Visualizing and Understanding Convolutional Networks. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer, Cham. https://doi.org/10.1007/978-3-319-10590-1_53