CNN-BASED COGNITIVE ACTIVITY RECOGNITION

Bob Brockberndd

Supervisors: Guohao Lan, Lingyu Du
EEMCS, Delft University of Technology, The Netherlands.

What is Cognitive Activity Recognition?

Cognitive activity recognition is the **classification of cognitive behaviour** using human data. This research uses **eye-movement data** in particular. In other words: **Gaze-based activity recognition**

Why is *CAR* useful?

- Recognizing driver fatigue with eye-tracking, is used in modern cars to prevent accidents.[1]
- Providing context for virtual reality.
- Cognitive activity recognition provides insights for educational research.

.

What is the challenge of *CAR*?

- Current technologies require **manual feature extraction**.
- Different people have large **differences** in **eye-movement behaviour**. Which makes manual feature extraction difficult.

Convolutional neural networks can solve these challenges by **learning feature extraction**.

Research questions

Can a convolutional neural network classifier be used for gaze-based activity recognition?

- What CNN architecture is best suited for gaze-based activity recognition?
- Which CNN hyperparameters perform best for gaze-based activity recognition?
- How do the found CNN performances compare to other machine learning techniques?

Method

1. **Preprocessing data**
2. **Finding suitable CNN architectures**
3. **Tuning hyperparameters**
4. **Validation**

1. Preprocessing

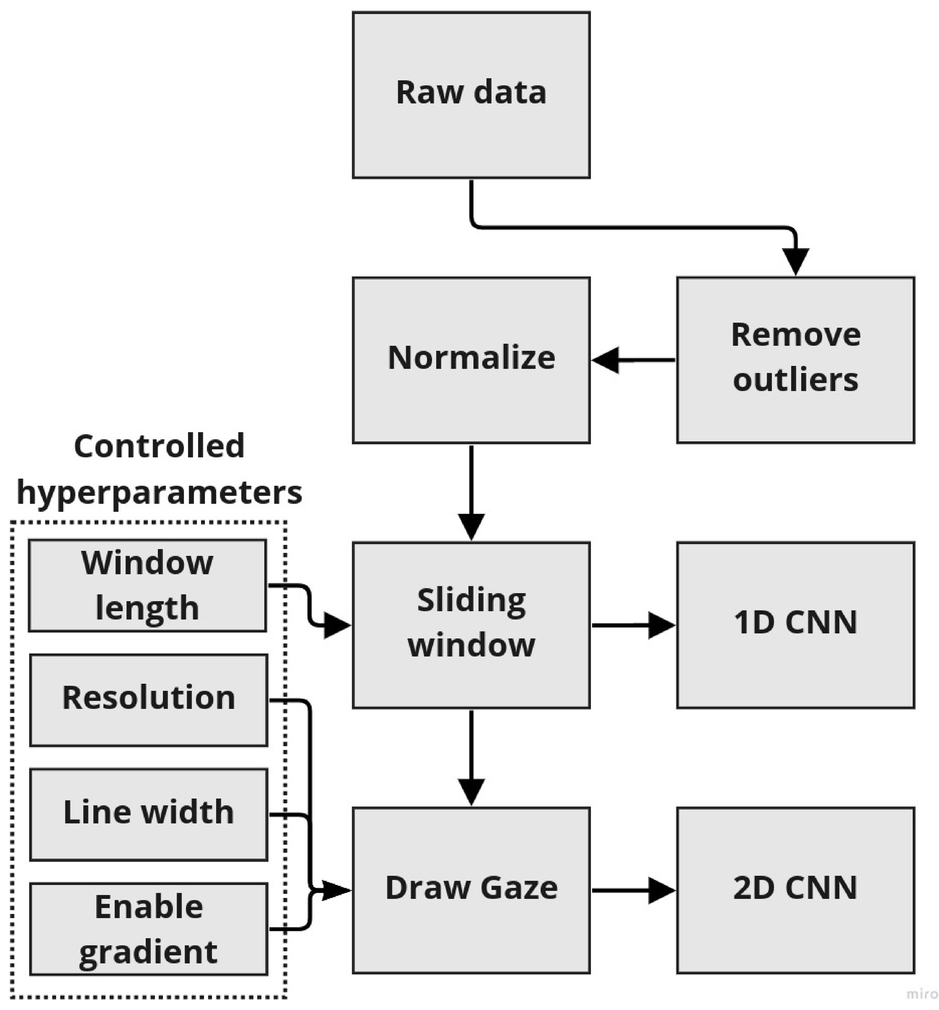

Fig 1: preprocessing pipeline

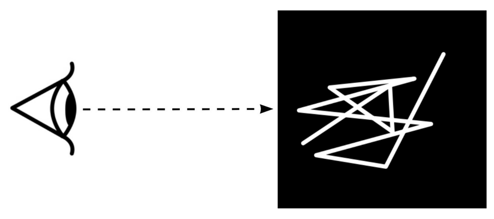

Fig 2: Drawing a gaze

2. Baseline CNN architectures

1D Convolution

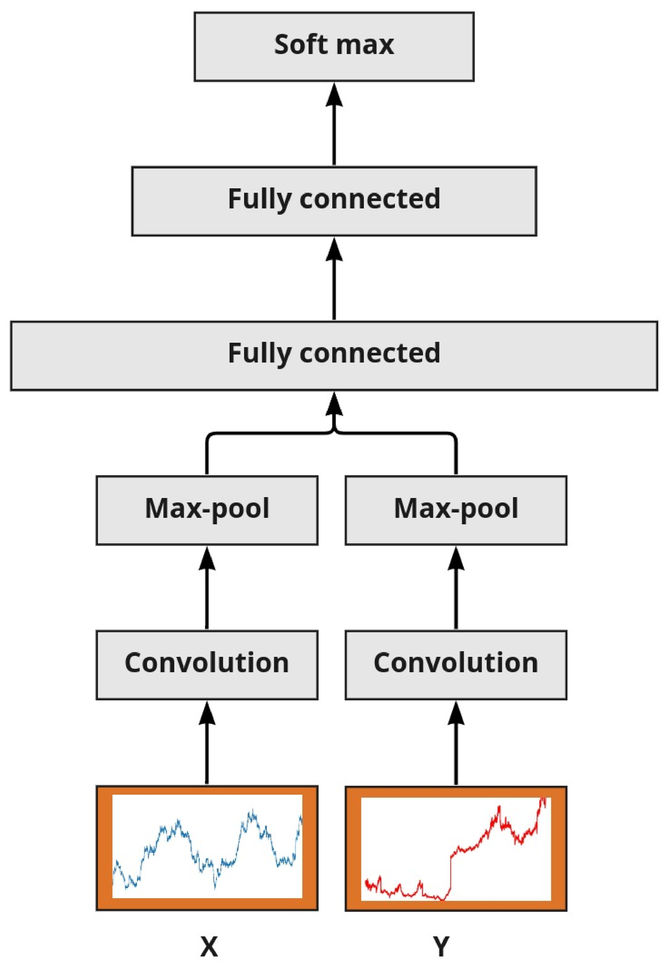

Fig 3: 1D convolutional neural network based on [2]

2D convolution

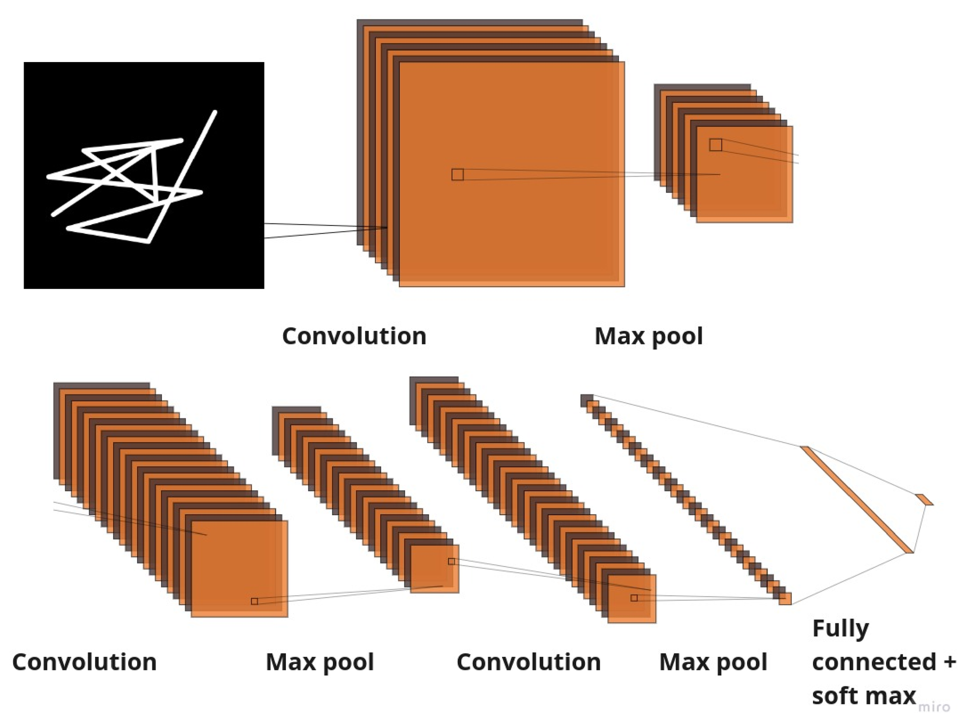

Fig 4: The LeNet5 convolutional neural network has proven to be effective in recognizing patterns in black and white images. [3]

3. Hyperparameter tuning

Hyperparameters: framelength, image resolution, line thickness, filter count, filter size, drop out rate, fully connected size, extra layers, enable gradient.

Tuning steps:

1. Defining search space with **hyperband**
2. Exploring search space with **bayesian optimization**
3. **Finalizing** hyperparameters

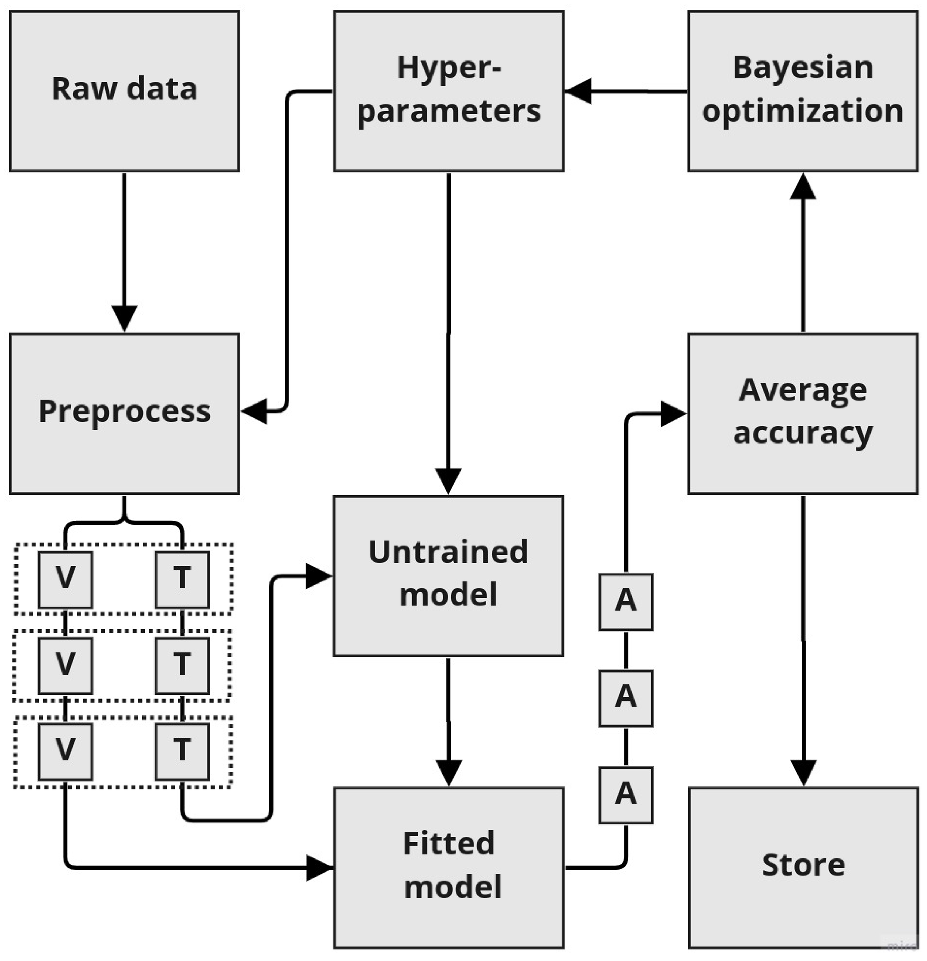

Fig 5: Search space exploration pipeline

4. Validation

Validation is done with **k-fold crossvalidation** on the **accuracy** metric. The CNNs can be validated using **known subjects or unknown subjects**.

Results

Table 1: Best CNN per dataset validated with unknown subjects compared to other machine learning techniques.

Algorithm	Reading	Sedentary	Desktop
RF	0.67	0.65	0.58
SVM	0.75	0.52	0.60
k-NN	0.71	0.48	0.54
LSTM	0.31	0.67	0.32
CNN	0.67	0.69	0.40

Table 2: Best CNN per dataset validated with known subjects compared to other machine learning techniques.

Algorithm	Reading	Sedentary	Desktop
RF	0.96	0.94	0.92
SVM	0.85	0.86	0.95
k-NN	0.91	0.77	0.84
LSTM	0.98	0.98	0.95
CNN	1.00	1.00	0.99

Conclusion

For known subjects, convolutional neural networks perform significantly better than current methods. However, further research is needed to improve performance for unknown subjects.

References

[1] P. Norloff, "Eye tracking technology is making new cars safer," Sep 2019. [Online]. Available: https://eyegaze.com/eye-tracking-technology-is-making-new-cars-safer/

[2] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," in *6th international conference on mobile computing, applications and services*. IEEE, 2014, pp. 197–205.

[3] Y. LeCun *et al.*, "Lenet-5, convolutional neural networks," *URL: http://yann. lecun. com/exdb/lenet*, vol. 20, no. 5, p. 14, 2015.

TU Delft