

Dezentrale automatisierte Gesichtserkennung

Fachvortrag im Rahmen der Themenreihe
„Künstliche Intelligenz → verständlich“

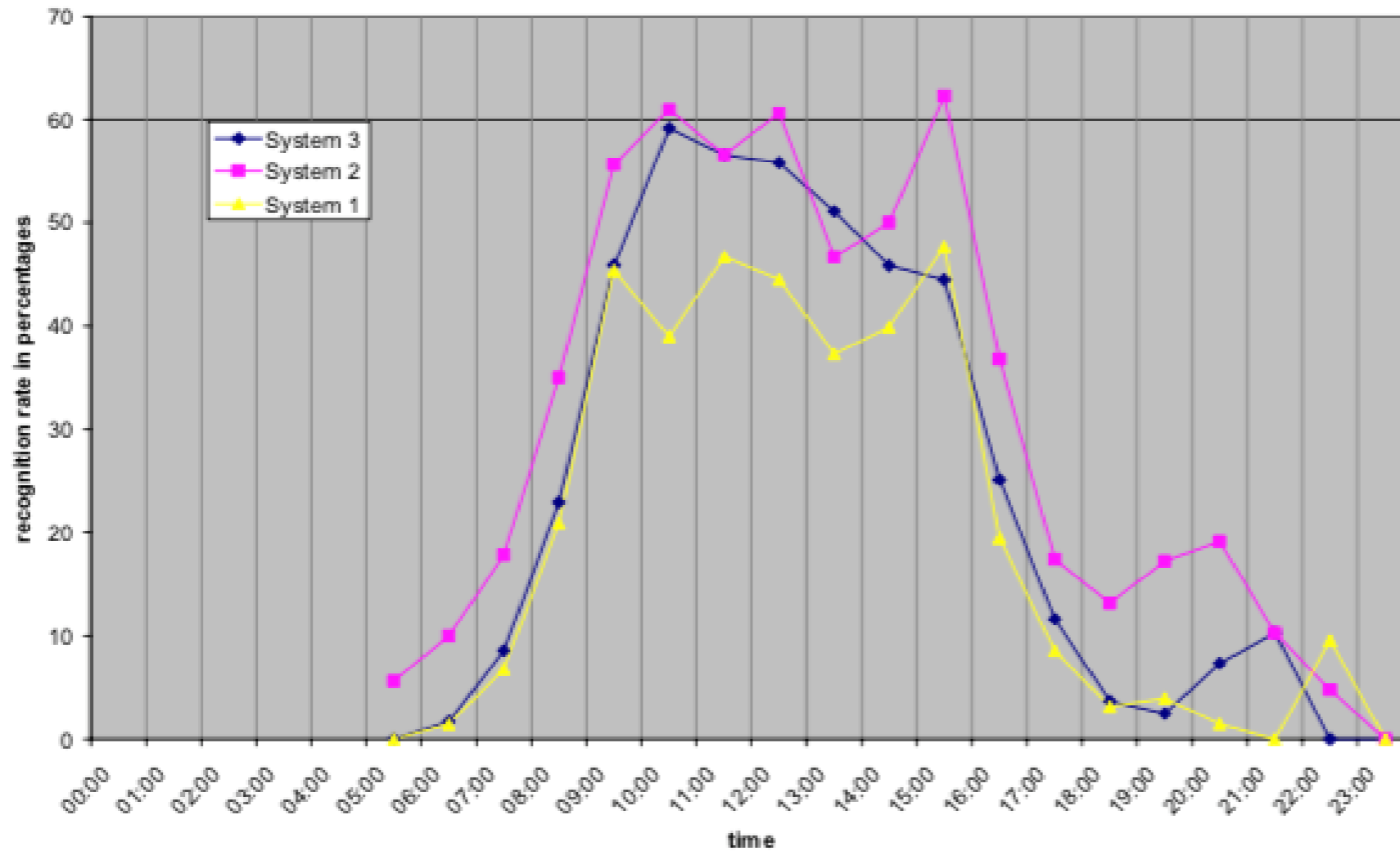
Technische Hochschule Wildau
9. März 2020

Prof. Dr. Rainer Stollhoff

Automatisierte Gesichtserkennung am Berliner Bahnhof Südkreuz

In 2008

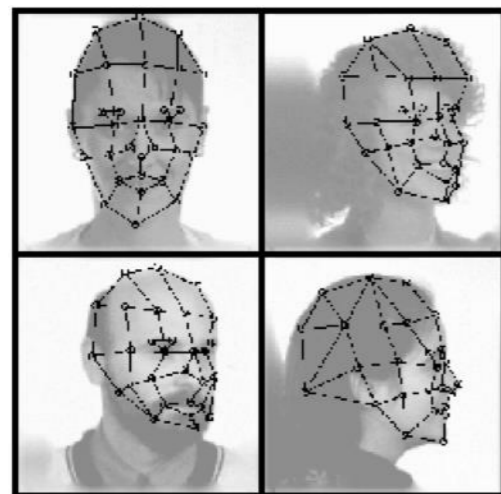
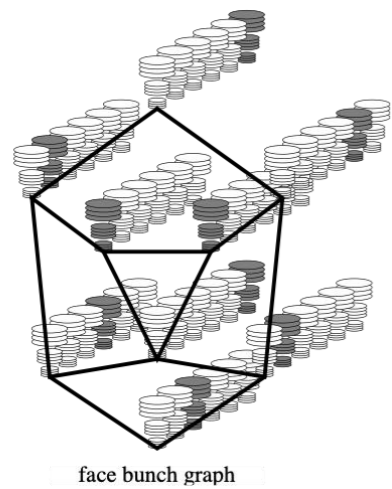
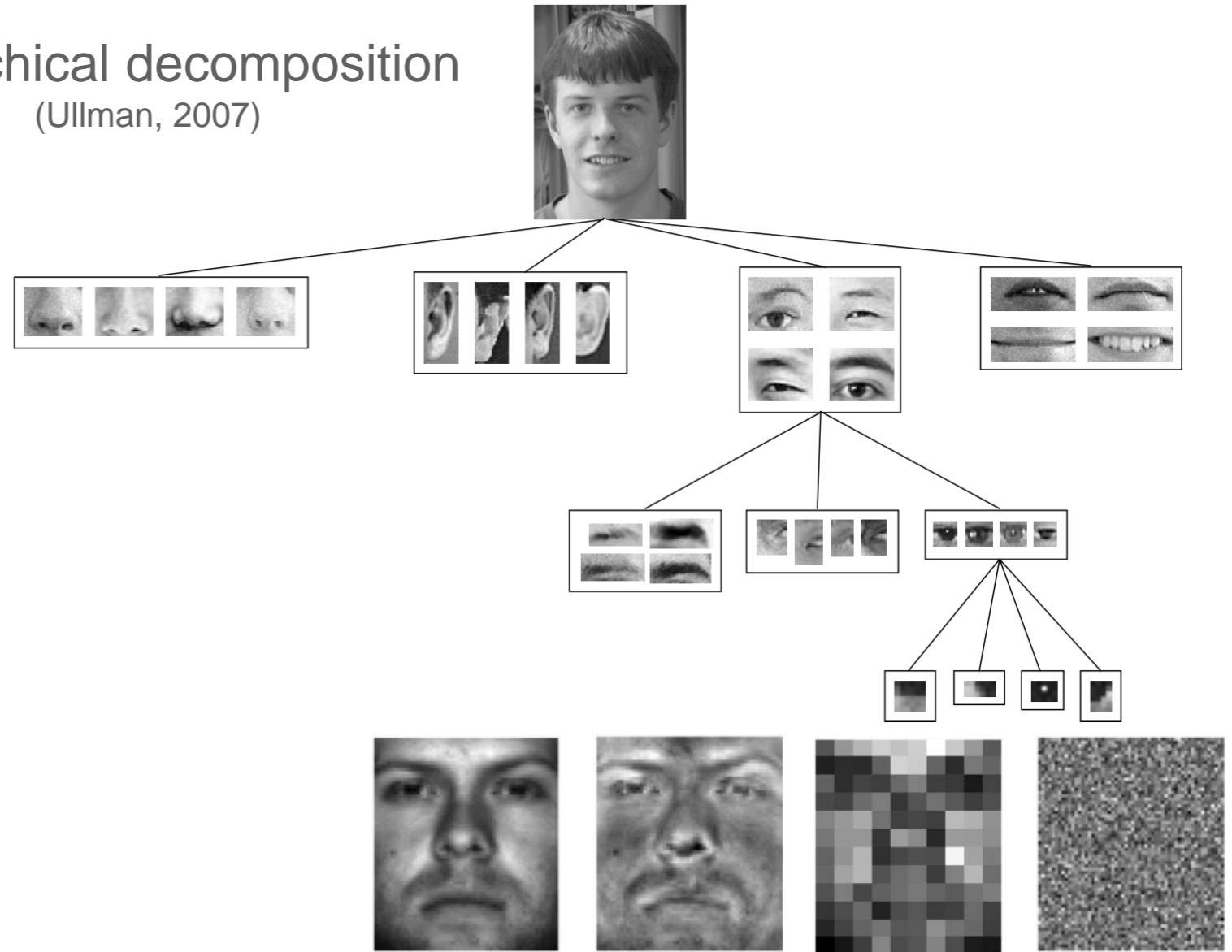
system comparison with FAR 0,1 % according to time



(Bundeskriminalamt, 2008)

Automatisierte Gesichtserkennung in 2008

Hierarchical decomposition
(Ullman, 2007)



Elastic bunch graph matching
(Wiskott et al., 1997)

Sparse representation
(Yang et al. 2007)



Herausforderungen in der Gesichtserkennung

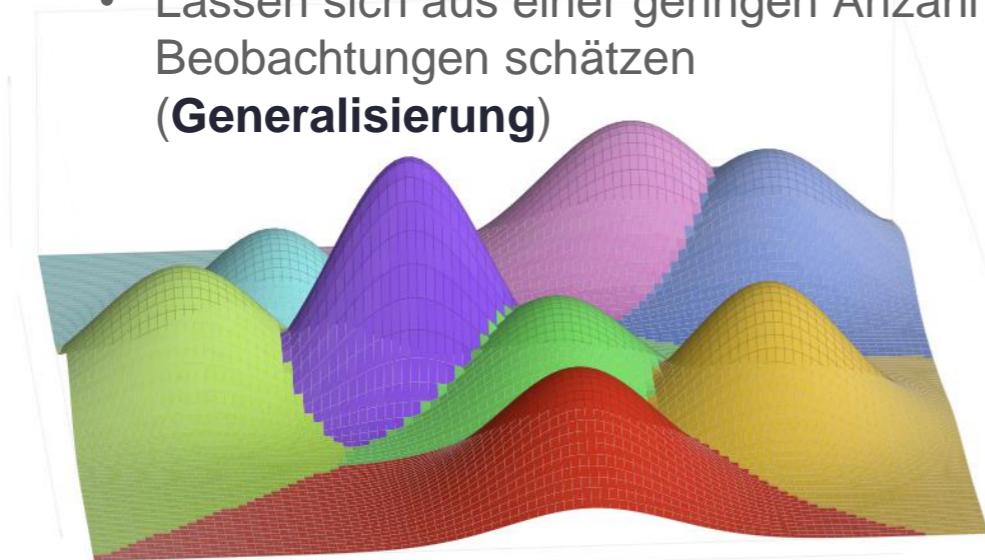
Für eine Beobachtung x weise eine Identität y zu...

... wobei nur zwischen bekannten Identitäten unterschieden wird

- bedingte Wahrscheinlichkeiten

$p(x|y)$

- Faktorisieren über identitätstypische Merkmalen (Augensabstand,...)
- Lassen sich aus einer geringen Anzahl an Beobachtungen schätzen (**Generalisierung**)



- Vorhersage anhand einer Partitionierung des Merkmalsraums

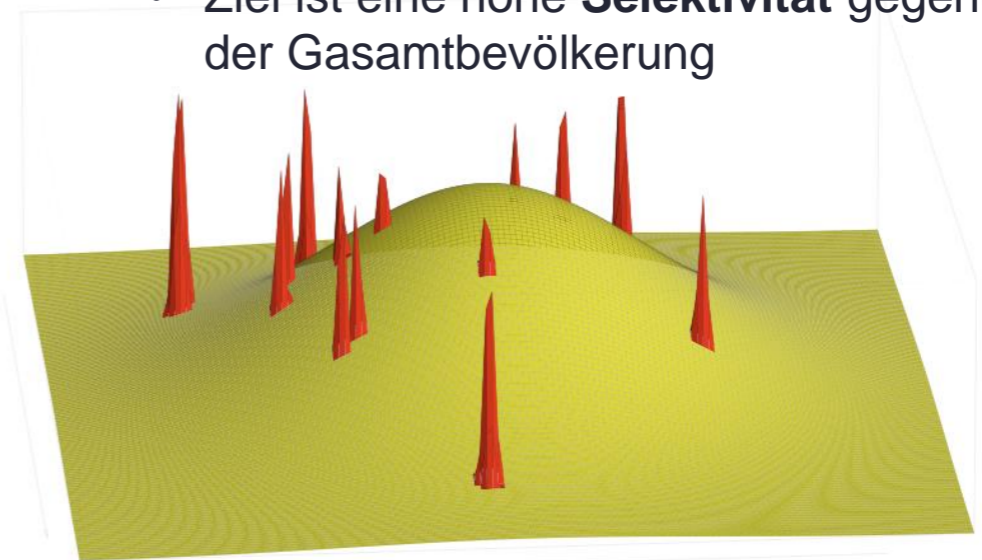
$$p(y|x) \propto p(x|y) p(y)$$

... wobei bekannte von unbekanntem Identitäten unterschieden werden

- Einzelne Punktmassen $p(x|y) = \delta(x - x_y)$

- Individuen sind Stichproben aus einer Grundgesamtheit

- Ziel ist eine hohe **Selektivität** gegenüber der Gesamtbevölkerung

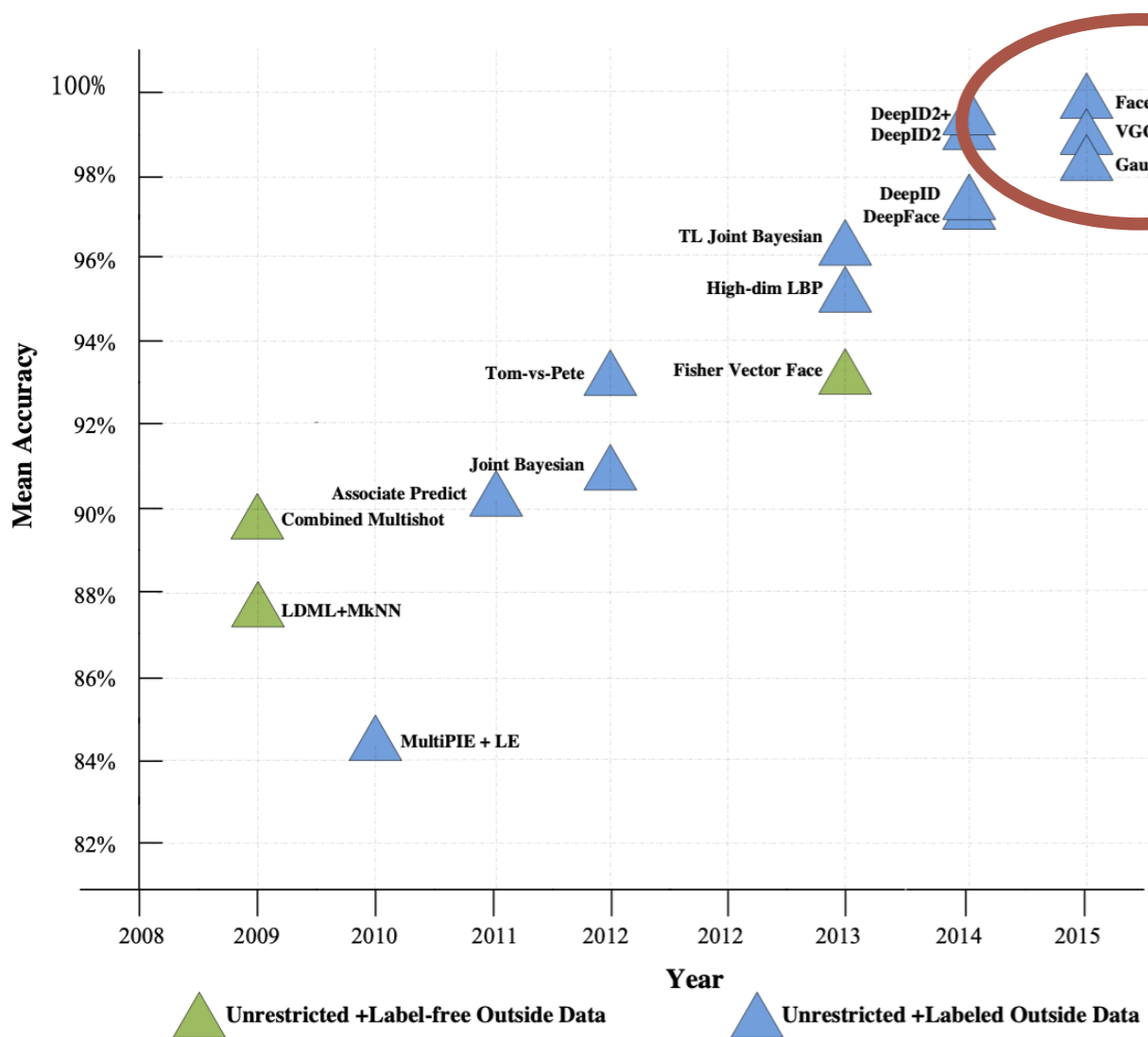


- Identität als einzigartige Gesamtheit

$$p(y|x) = \delta(x - x_y)$$

Entwicklung der automatisierten Gesichtserkennung

Performance on LFW dataset



(Liu et al., 2018)

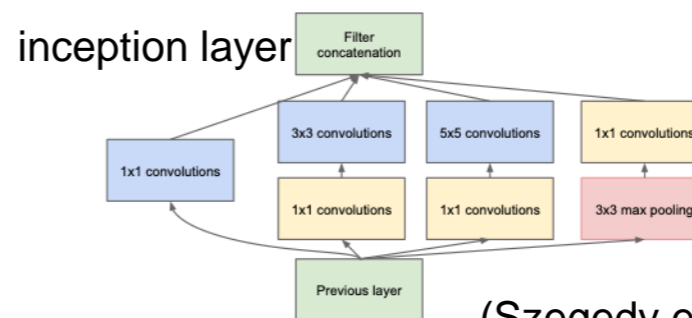
FaceNet

(Schroff et al. 2015)



type	output size	depth	#1x1	#3x3 reduce	#3x3	#5x5 reduce	#5x5	pool proj (p)	params	FLOPS
conv1 (7x7x3, 2)	112x112x64	1							9K	119M
max pool + norm	56x56x64	0						m 3x3, 2		
inception (2)	56x56x192	2		64	192				115K	360M
norm + max pool	28x28x192	0						m 3x3, 2		
inception (3a)	28x28x256	2	64	96	128	16	32	m, 32p	164K	128M
inception (3b)	28x28x320	2	64	96	128	32	64	L ₂ , 64p	228K	179M
inception (3c)	14x14x640	2	0	128	256,2	32	64,2	m 3x3,2	398K	108M
inception (4a)	14x14x640	2	256	96	192	32	64	L ₂ , 128p	545K	107M
inception (4b)	14x14x640	2	224	112	224	32	64	L ₂ , 128p	595K	117M
inception (4c)	14x14x640	2	192	128	256	32	64	L ₂ , 128p	654K	128M
inception (4d)	14x14x640	2	160	144	288	32	64	L ₂ , 128p	722K	142M
inception (4e)	7x7x1024	2	0	160	256,2	64	128,2	m 3x3,2	717K	56M
inception (5a)	7x7x1024	2	384	192	384	48	128	L ₂ , 128p	1.6M	78M
inception (5b)	7x7x1024	2	384	192	384	48	128	m, 128p	1.6M	78M
avg pool	1x1x1024	0								
fully conn	1x1x128	1							131K	0.1M
L2 normalization	1x1x128	0								
total									7.5M	1.6B

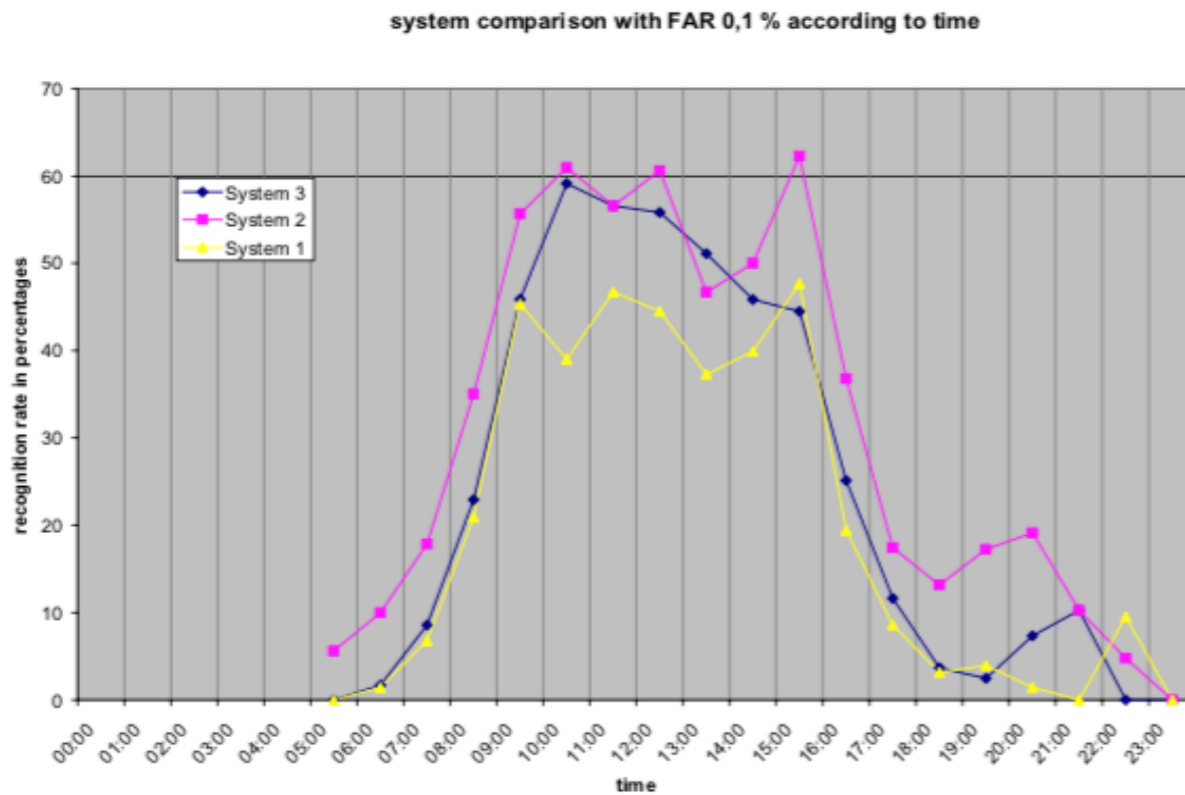
7.5 M parameters to map 224x224 image to 128 dim embedding space



(Szegedy et al., 2015)

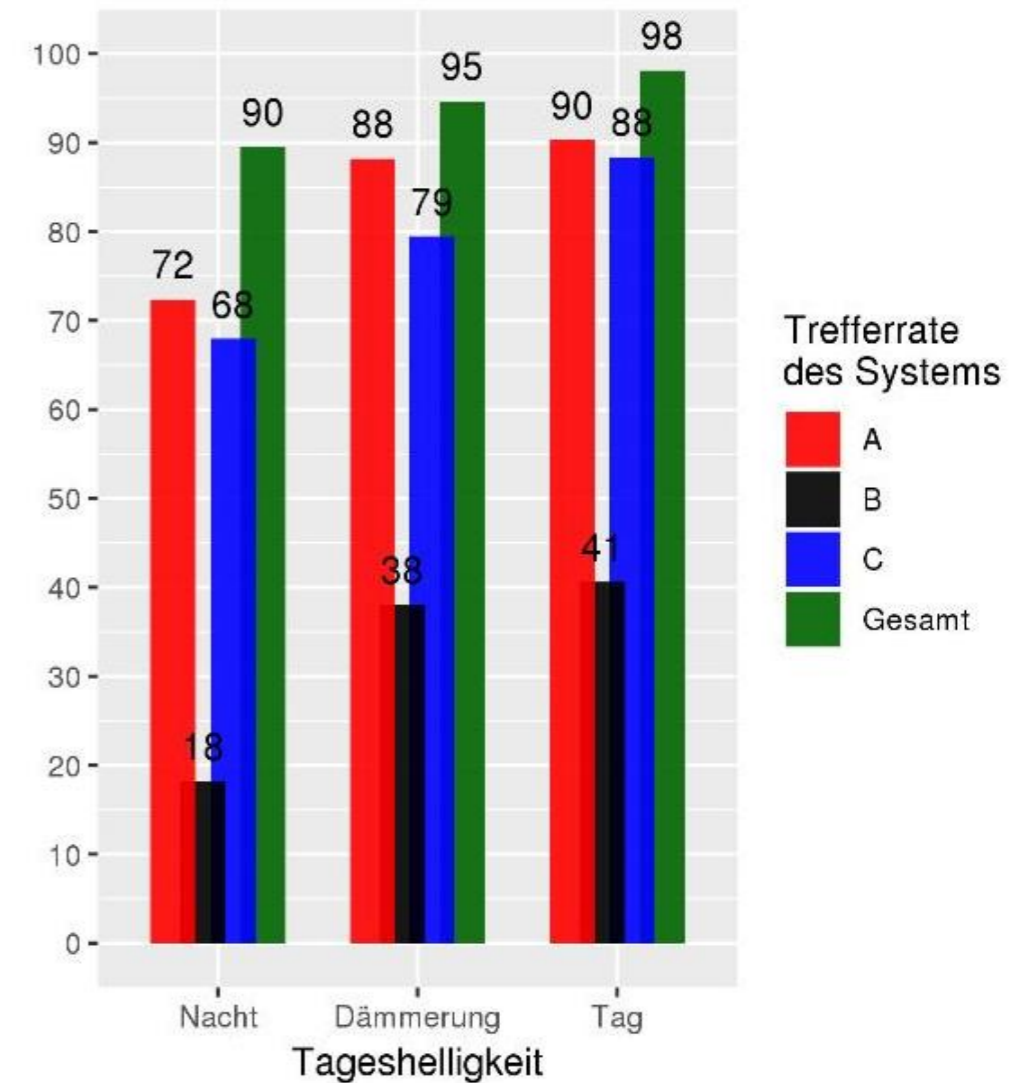
Automatisierte Gesichtserkennung am Berliner Bahnhof Südkreuz

In 2008



(Bundeskriminalamt, 2008)

In 2018



(Bundespolizeipräsidium, 2018)

Datenschutz und Datensparsamkeit als zentrale Herausforderungen der automatisierten Gesichtserkennung

- Großflächiger Einsatz automatisierter Gesichtserkennung wird zunehmend kritisch hinterfragt, insbesondere zentralisierte Lösungen mit
 - umfassender Videoüberwachung
 - umfassenden Datenbanken
 - automatisiertem Abgleich mittels KI
- Alternativen zu umfassenden zentralisierten Lösungen
 - Anlassbezogene und ortsgebundene Kontrollen
 - Geschlossene Systeme
 - Abgleich Gesichtsbild mit biometrischen Daten in Ausweisdokument
 - Dezentrale Systeme mit lokalen Datenbanken und KI-Systemen
 - Vollständig lokale Systeme:
eigene Trainingsdaten, eigene Parameter, eigene Architektur, bekannte Algorithmen
 - Transfer Learning:
eigene Trainingsdaten, fremde Parameter, fremde Architektur, bekannte Algorithmen

References

- Bundeskriminalamt. (2008). *Face recognition as a search tool: "Foto-Fahndung" Final report*. Bundeskriminalamt. www.bka.de/kriminalwissenschaften/fotofahndung/pdf/fotofahndung_final_report.pdf
- Bundespolizeipräsidium Potsdam. (2018). *Abschluss bericht Teilprojekt 1 „Biometrische Gesichtserkennung“*. Bundespolizeipräsidium Potsdam.
https://www.bundespolizei.de/Web/DE/04Aktuelles/01Meldungen/2018/10/181011_abschlussbericht_gesichtserkennung_down.pdf
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4401–4410.
- Liu, X., Kan, M., Wu, W., Shan, S., & Chen, X. (2017). VIPLFaceNet: an open source deep face recognition SDK. *Frontiers of Computer Science*, 11(2), 208–218.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823.
- Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). *Going deeper with convolutions*. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Sciences*, 11(2), 58–64.
- Wiskott, L., Fellous, J. M., Kruger, N., & Malsburg, C. von der. (1997). Face recognition by elastic bunch graph matching. *IEEE*, 19(7), 775–779.
- Yang, A., Wright, J., Ma, Y., & Sastry, S. (2007). Feature selection in face recognition: A sparse representation perspective. *Submitted to IEEE Transactions Pattern Analysis and Machine Vdots*.