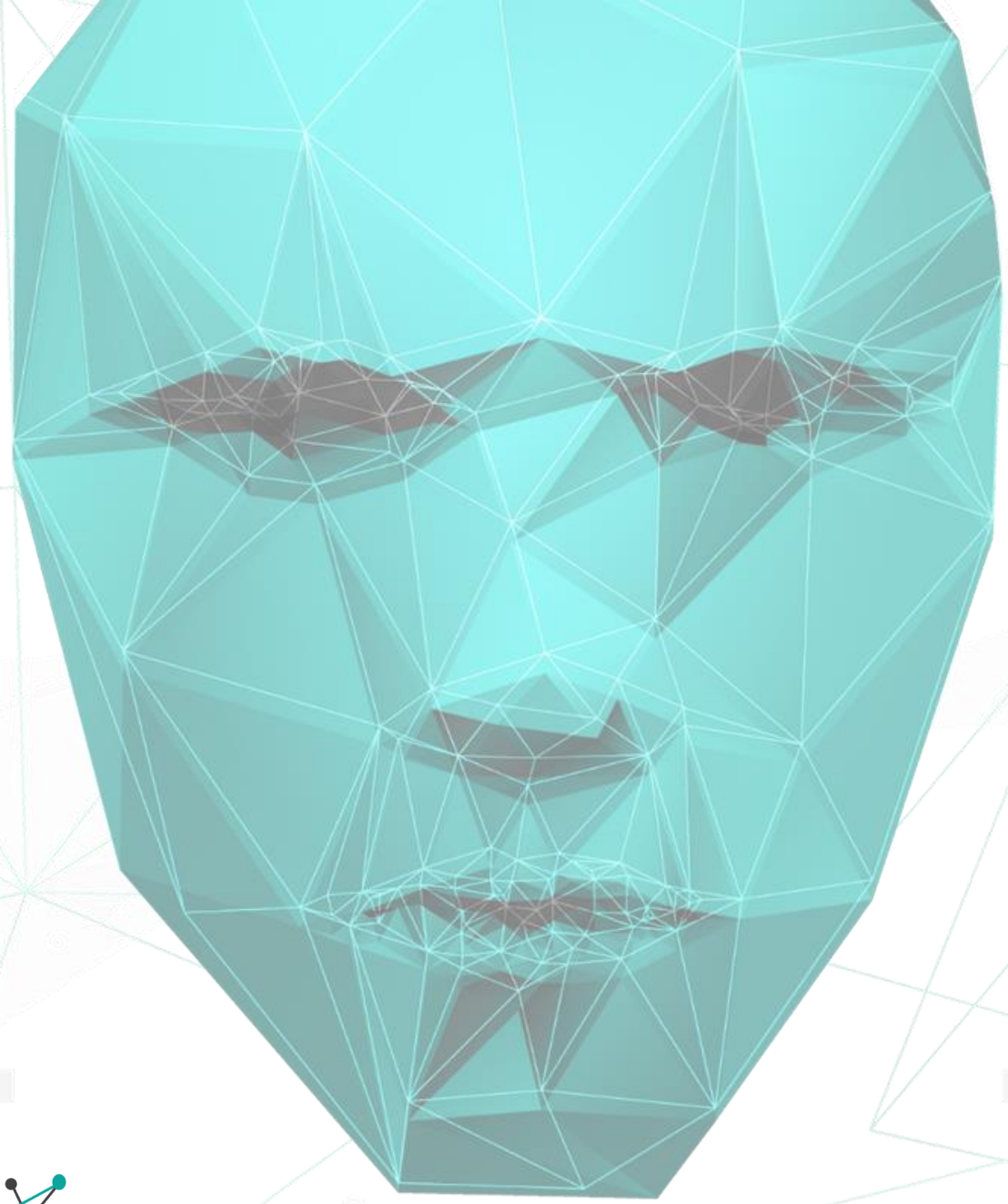


Vortrag 07.07.2021

COMPUTER VISION

Ein Einblick in die faszinierende Welt des maschinellen Sehens

Manuel & Michael Wäger

deeprotype

Agenda

01 **Computer Vision**
Begriffserklärung, Einsatz

02 **Deep Learning**
Neuronale Netze & CNN
Aufbau Objektdetektion

03 **Gedankenspiel**
Von der Bildklassifizierung bis hin zur Bildmanipulation mittels GAN

ca. 45 min.

04 **Fragen**
Hier kannst du deine Fragen stellen

ca. 15 min.

Ziel

- Ansätze verstehen & Herausforderungen von Computer Vision in der Bildverarbeitung
- Kennenlernen der Funktionsweise und Idee von Deep Learning in Computer Vision

deeprotype

Vorstellung

Michael Wäger

Entwickler & Gründer deeprotype

- Elektrotechnik Dual (FH-Vorarlberg)
- Daten- und Prozessleittechnik (illwerke vkw AG)
- Datenauswertung NLP und Data Mining (illwerke vkw AG)
- Spezialisierung: Datenauswertung NLP und Data Mining

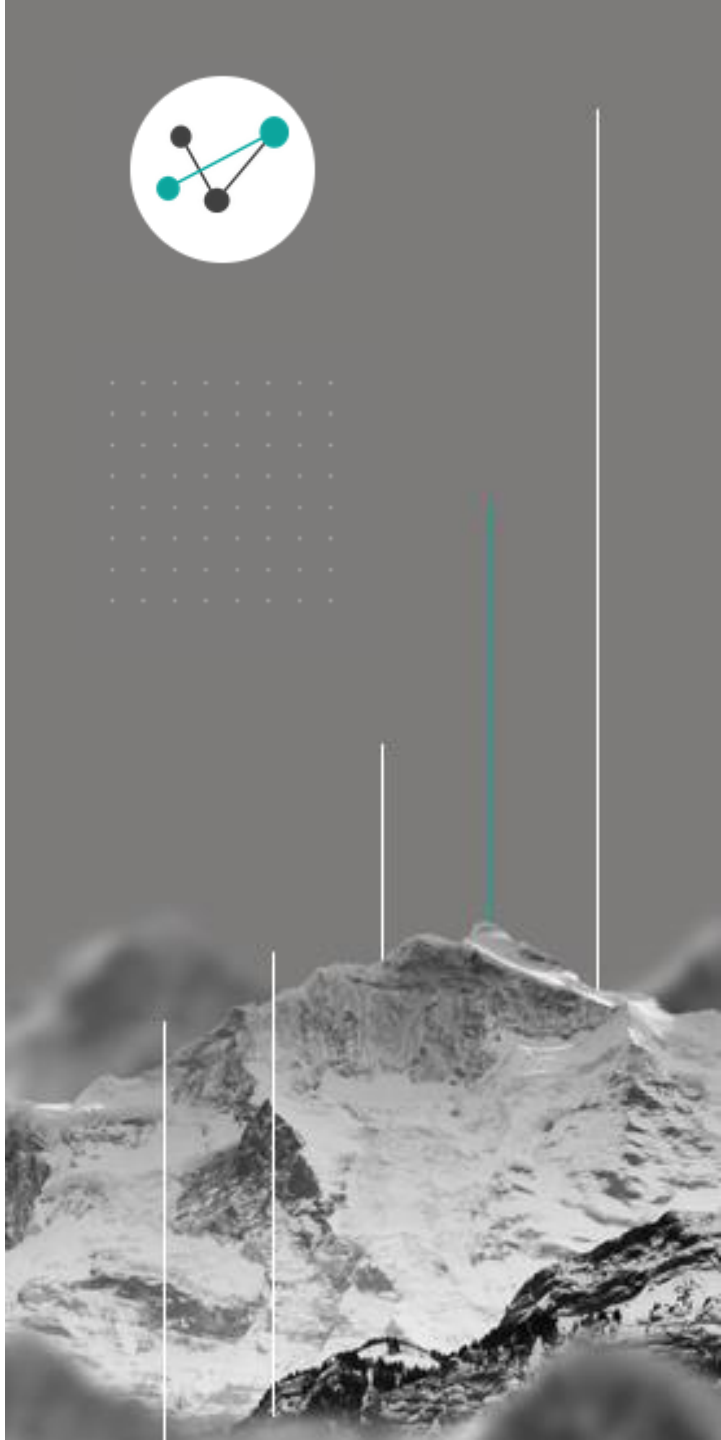
michael.waeger@deeprotype.com

Manuel Wäger

Entwickler & Gründer deeprotype

- Mechatronik & Robotik (FH-Technikum Wien)
- Konstrukteur (TAUROB GmbH)
- Data Scientist für ind. CV (Julius Blum GmbH)
- Spezialisierung: Deep Learning und Computer Vision

manuel.waeger@deeprotype.com



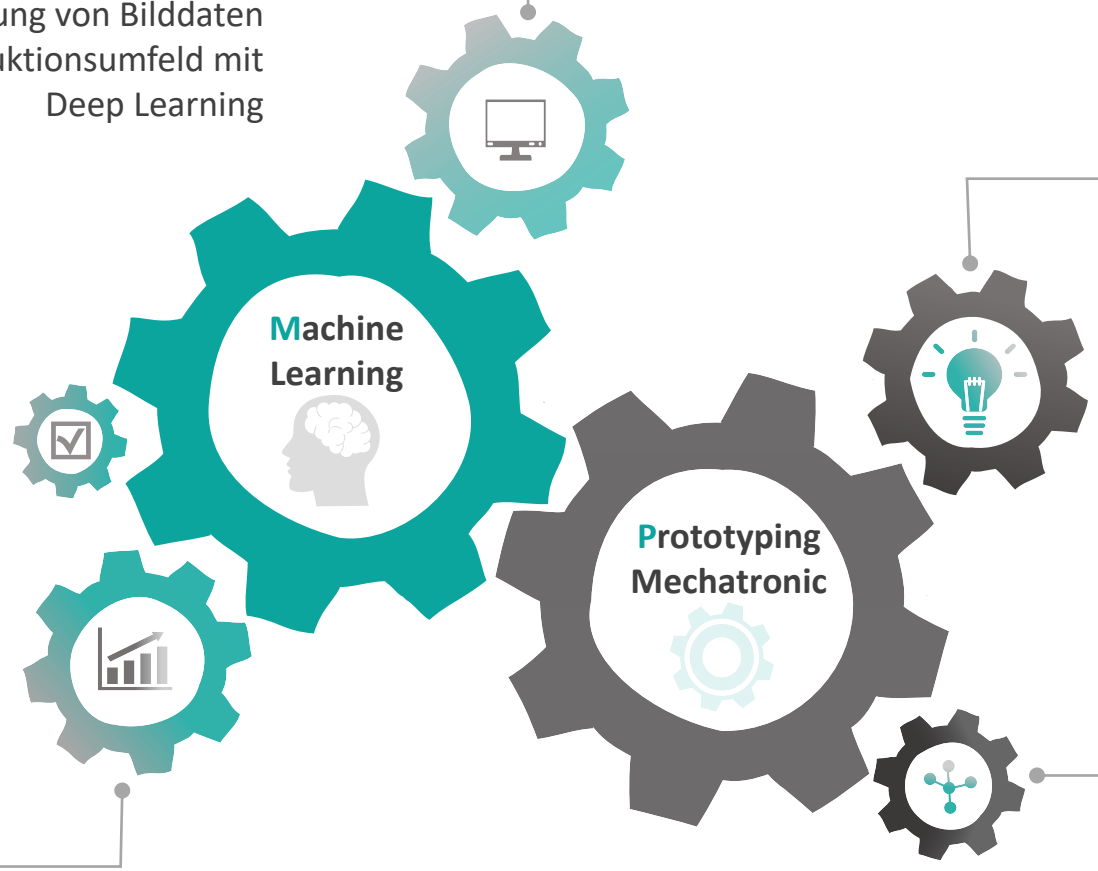
Dienstleistung

Computer Vision

Auswertung von Bilddaten im Produktionsumfeld mit Deep Learning

Mechatronic

Kundenspezifische Problemstellungen lösen (Hardware → Software)



Natural Language Processing (NLP)

Auswertung von produktionsnahen Maßnahmentexten

Condition Based Monitoring (CBM)

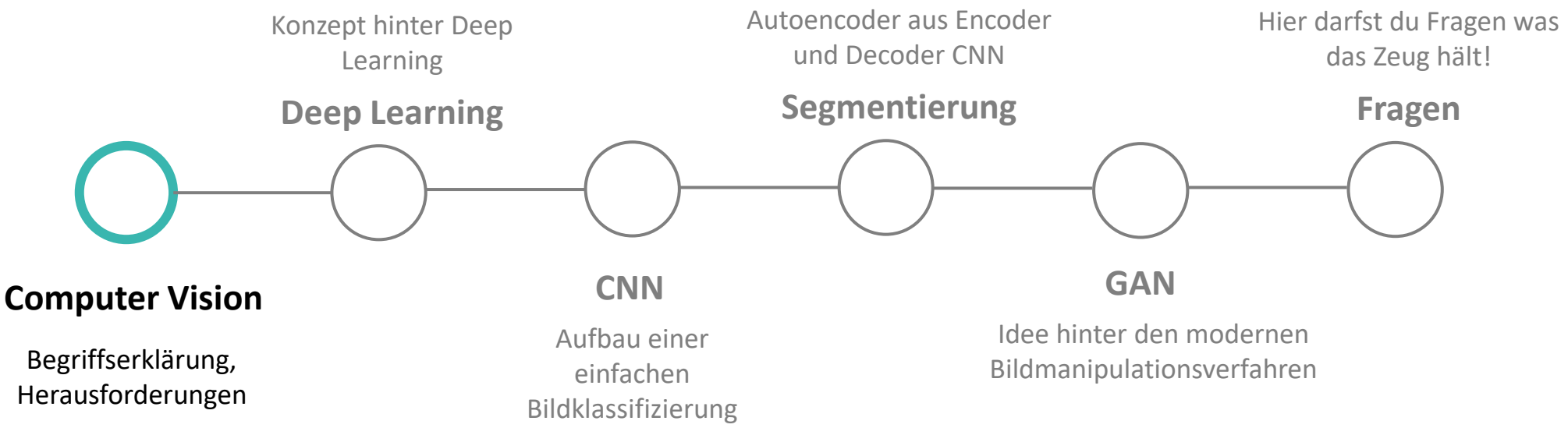
Auswertung von Maschinendaten zur Verschleißgraderkennung von Werkzeugen mittels Machine Learning

IoT Sensoric

Erfassung von verschiedenen Zustandsgrößen (CBM)

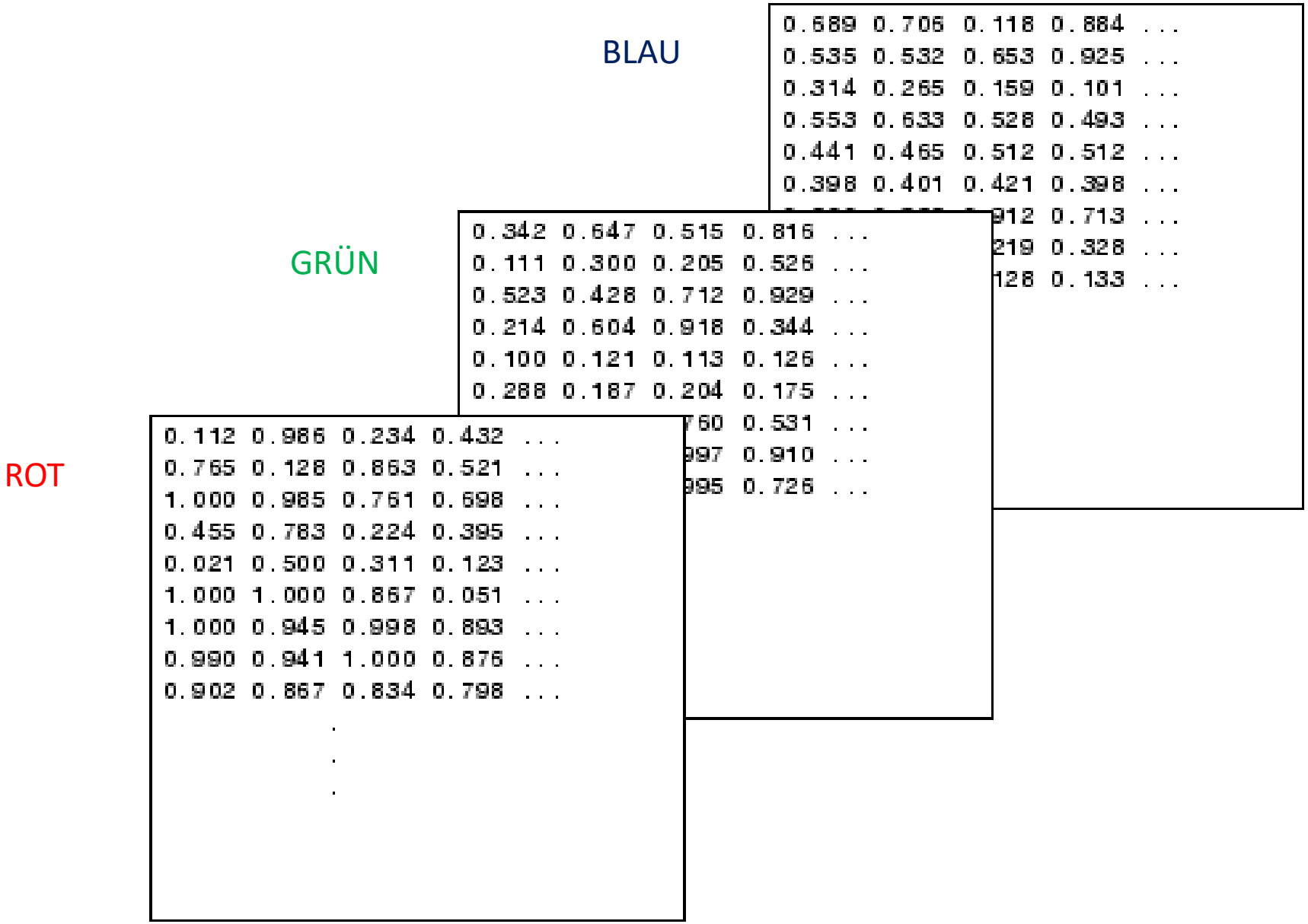
deeprotype

Fortschritt



deeprotype

Quelle: (hamburger-tierschutzverein, 26.05.2021)

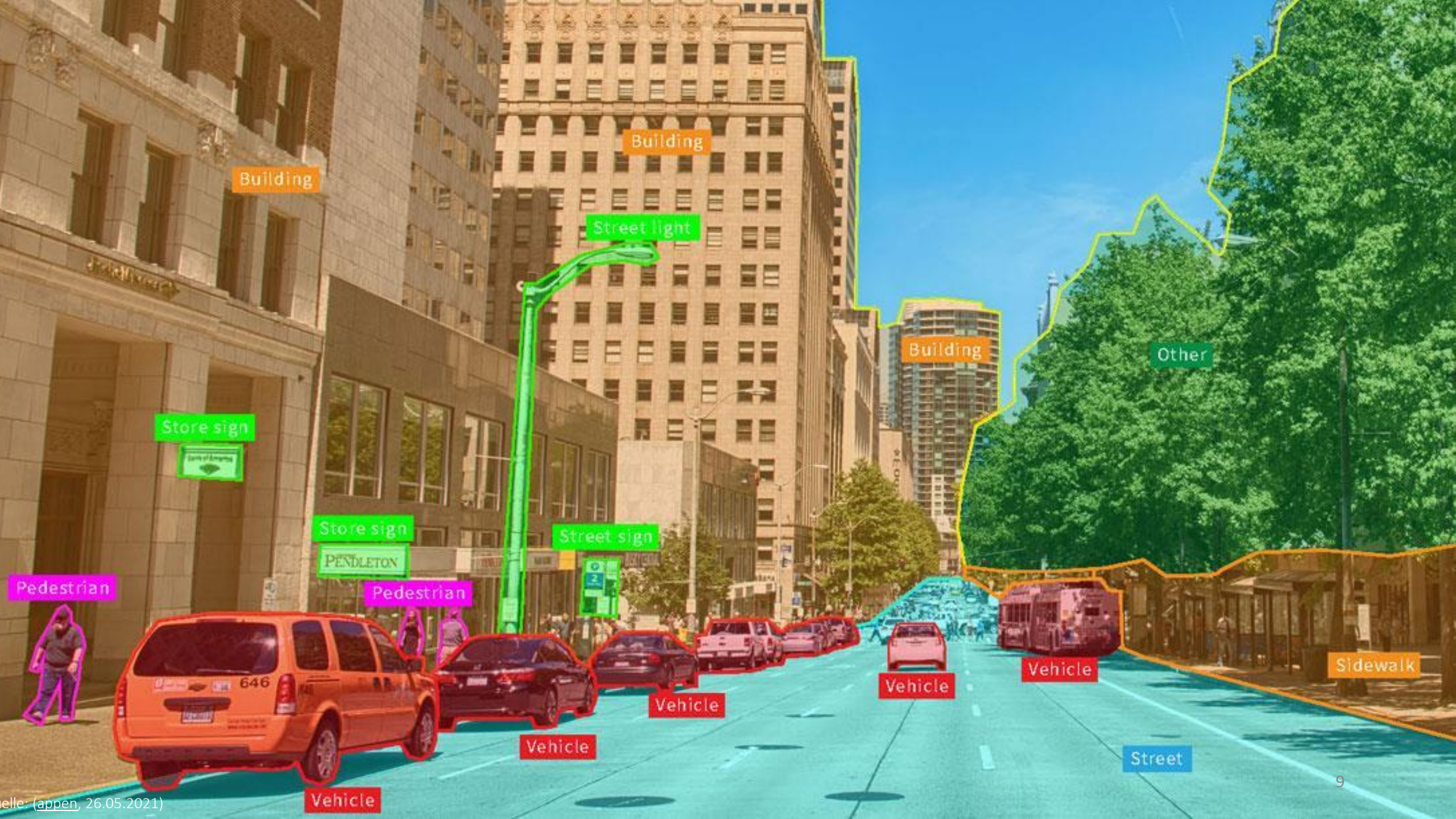


Quelle: (mathworks, 26.05.2021)

„Computer Vision (dt.: computerbasiertes Sehen) bezeichnet Systeme, die Objekte in digitalem Stand- und Bewegtbildmaterial erkennen und entsprechend verarbeiten.„

Zitat: Heller Martin

Quelle: Computer Vision (2021, Heller)

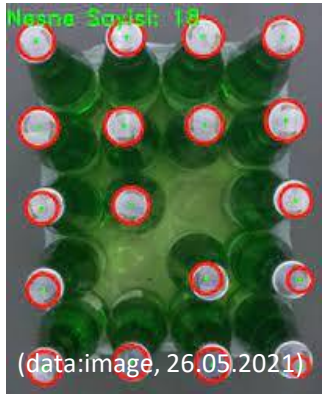


elle: (appen, 26.05.2021)

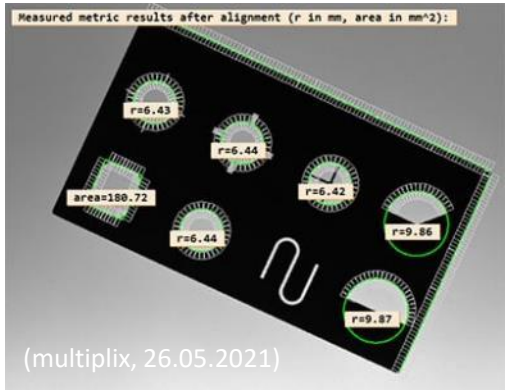
Machine Vision
(Industrie)

kontrolliert

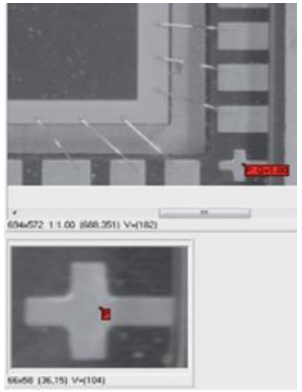
einfache geometrische Formen



(data:image, 26.05.2021)

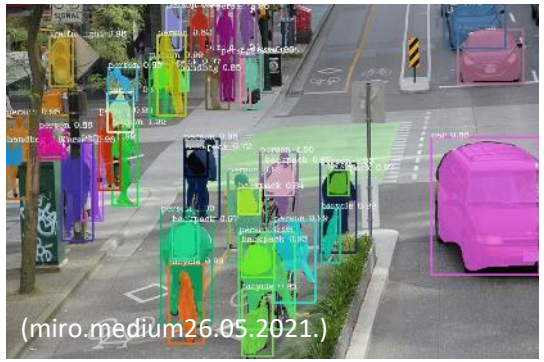


(multiplix, 26.05.2021)

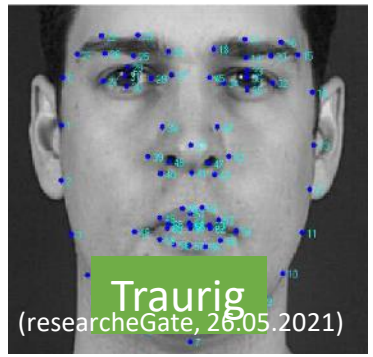


unkontrolliert

komplexe geometrische Formen


(miro.medium26.05.2021.)
Detektion

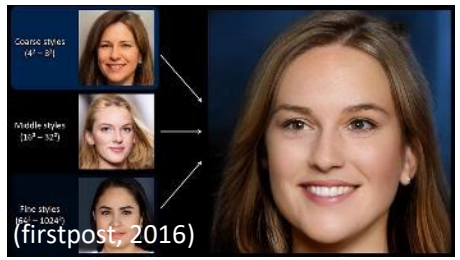

(pinterest, 26.05.2021.)
Segmentierung

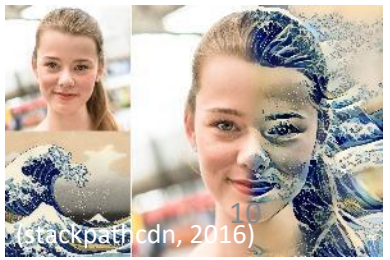


(researcheGate, 26.05.2021)
Analyse

Bildmanipulation

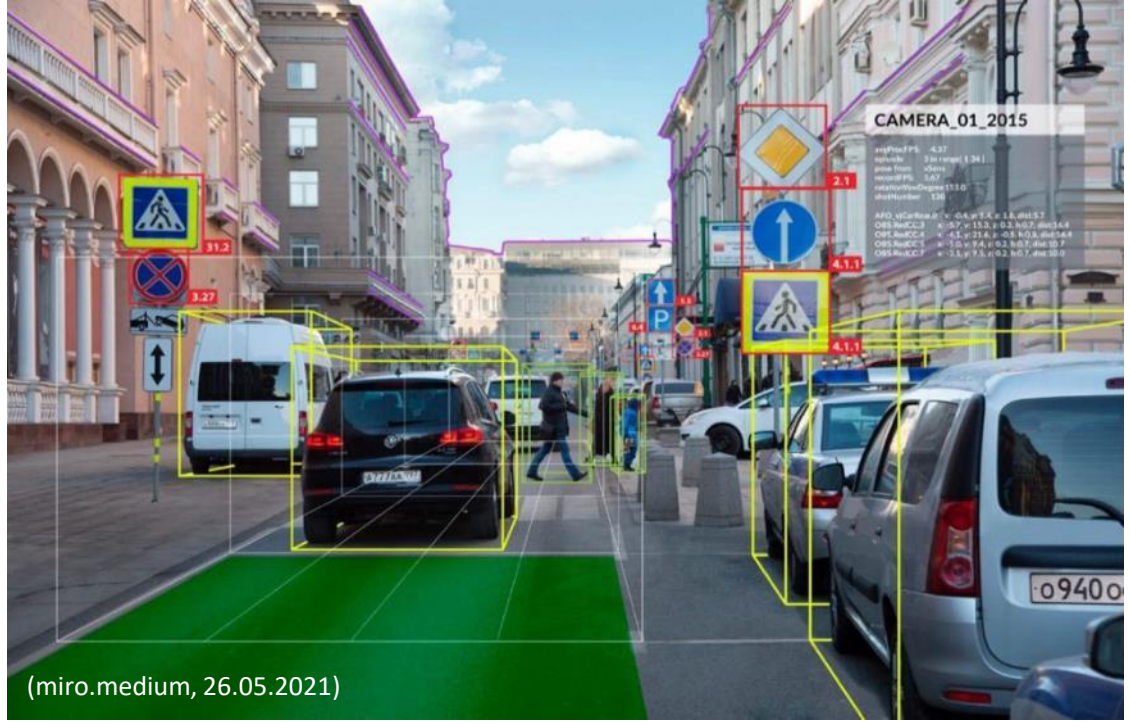

(Nvidia, 2016)


(firstpost, 2016)


(stackpath.cdn, 2016)

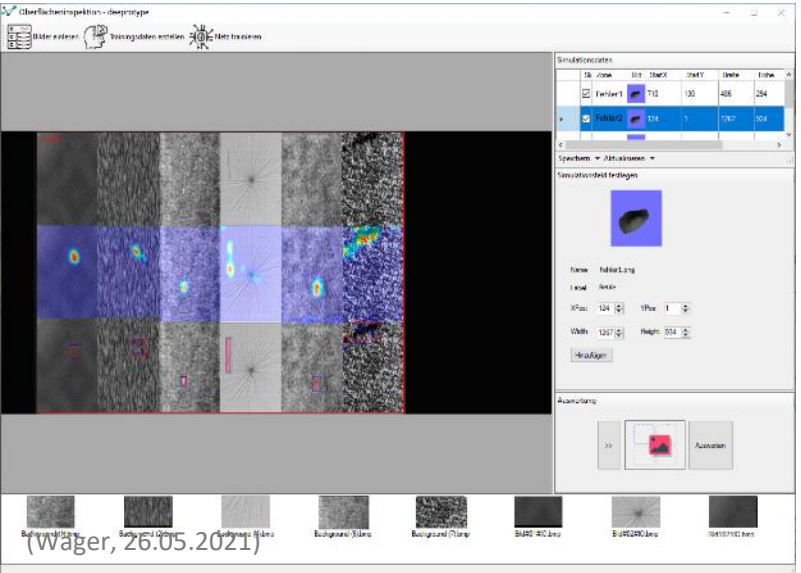
(Karras et. al., 2019)

(Saharan, 26.05.2021)



(miro.medium, 26.05.2021)

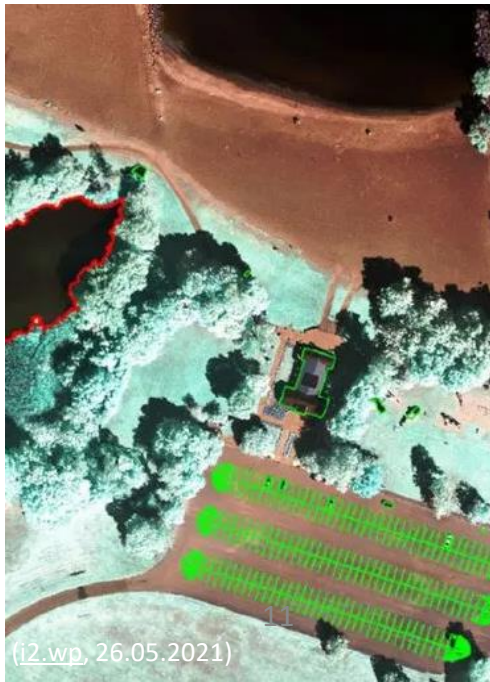
(techgenyz, 26.05.2021)



(Wager, 26.05.2021)

(Pubs.Rsna, 26.05.2021)

(pcwelt , 26.05.2021)



(i2.wp, 26.05.2021)

= KATZE

Quelle: (hslu, 26.05.2021)

Traditionelle Vorgehensweise CV

Qualitative Merkmale finden!

Merkmale finden

- Farbe
- Anzahl Extremitäten
- Anzahl Ohren
- Anzahl Schwanz
- Größe
- Augenfarbe

Merkmale extrahieren

Filter

- Kanten
- Ecken
- Linien
- Kreise

Klassifizieren

Regeln definieren

- If-Abfrage
- Statistik

KATZE?

Quelle: (MVTEC, 2021)

Quelle: (Lex Friidman, 2020)

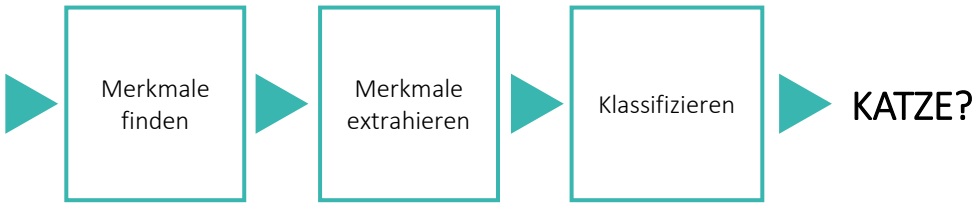
Quelle: (Lex Friidman, 2020)

Quelle: (katzen-fieber, 2021)

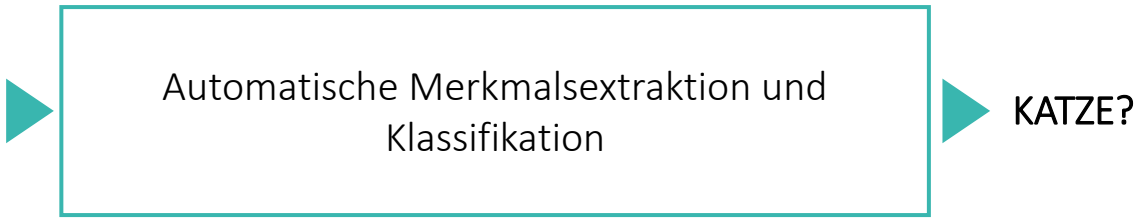
Quelle: (katzen.de, n. D.)

CV: Traditionell vs. Deep Learning

Traditionell



Deep Learning



(Tensorflow., 2019)
(MVTEC, 2021)

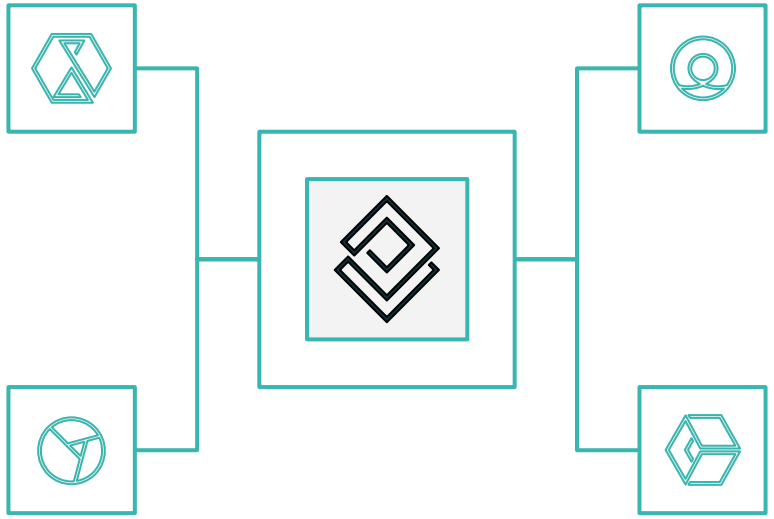
Wahl ist abhängig von der Aufgabe

Aufwand

Wenn die Ermittlung aller nötigen Parameter zu aufwendig ist.

Unbekannte Parameter

Wenn Parameter gänzlich bzw. teilweise fehlen.



Komplexität

Wenn die Erstellung eines Modells für die Abdeckung aller Möglichkeiten zu komplex wird.

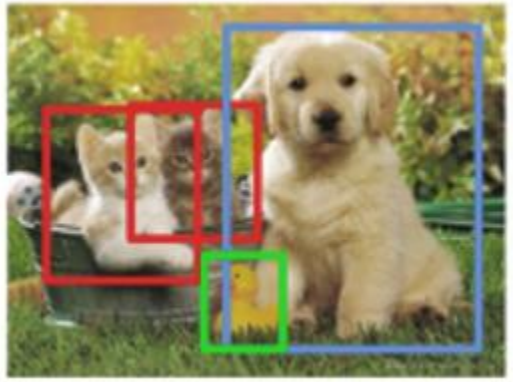
Flexibilität

Wenn das System auf unterschiedliche Objektvarianten gleichermaßen funktionieren muss

Computer Vision

Deep Learning

Klassifikationsvarianten

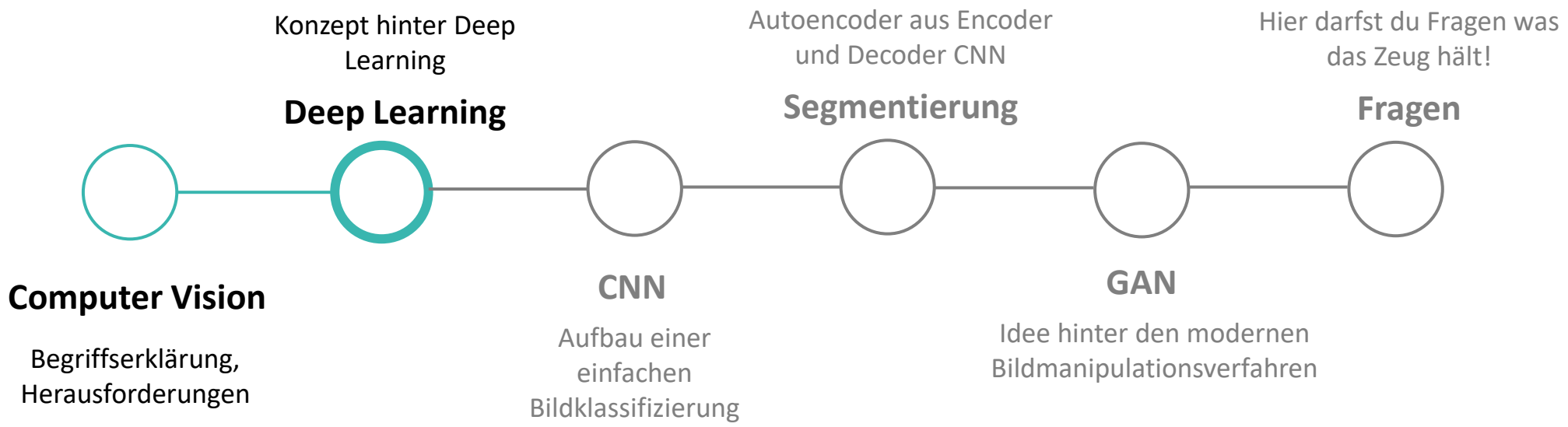
Bildklassifikation	Objektdetektion	Segmentierung
		
Katze	Katze, Hund, ...	Katze, Hund, ...
Um welches Objekt handelt es sich?	Wo und welche Objekte?	Welche Pixel gehören zu der Klasse?

einfach

mehrfach

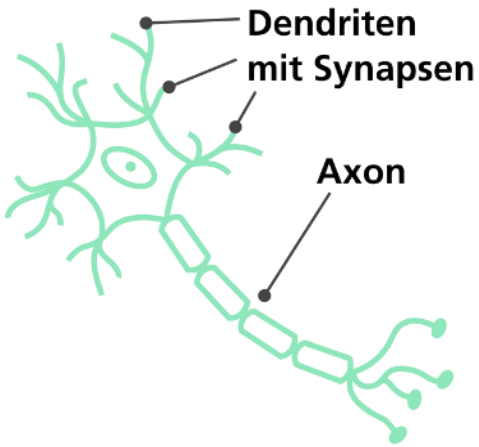
Bildquelle: (Tamir M., o. D.)

Fortschritt



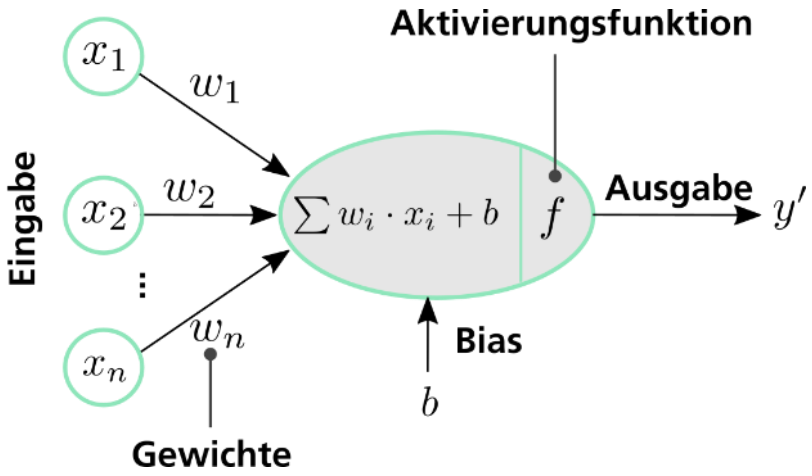
Biologische Inspiration für den Computer

Biologisches Neuron



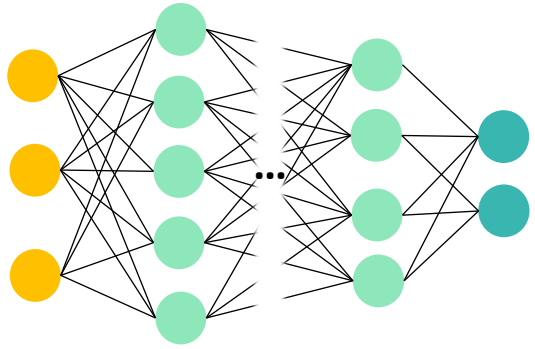
Berechnungsbaustein für das Gehirn

(Künstliches) Neuron



Rechenbaustein für ein "Neuronales Netz“

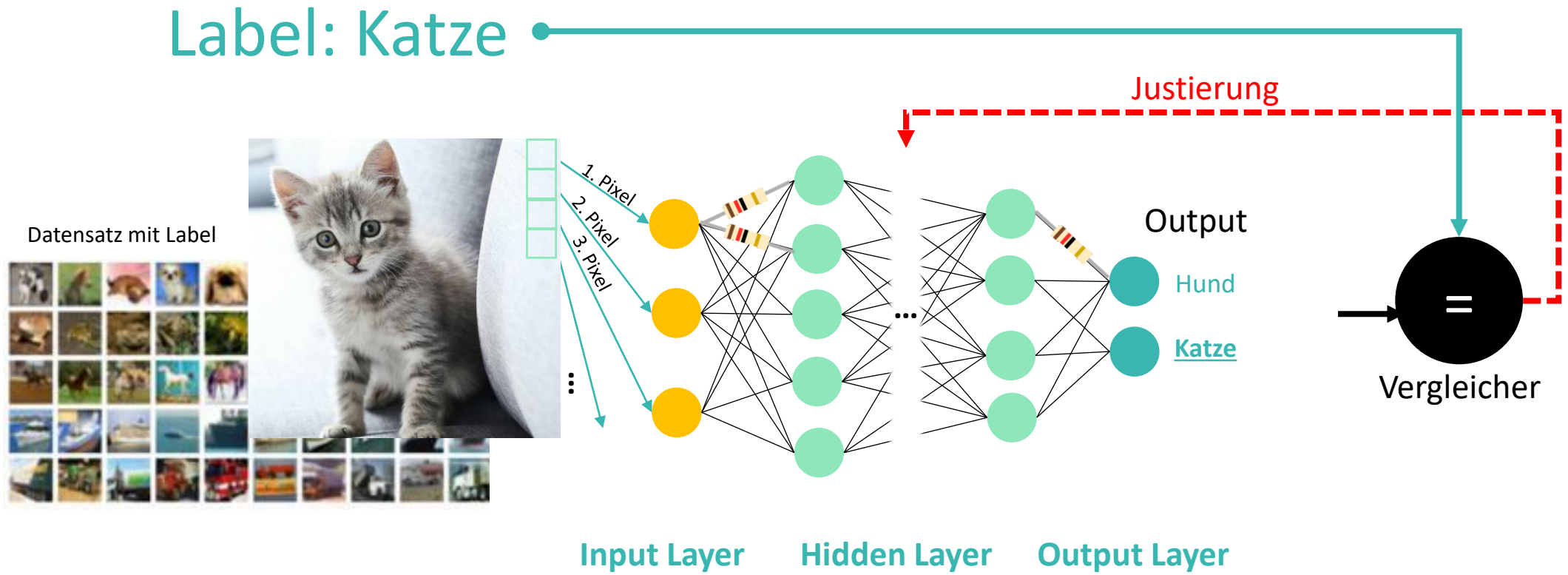
Neuronales Netzwerk



Verbund aus mehreren Neuronen

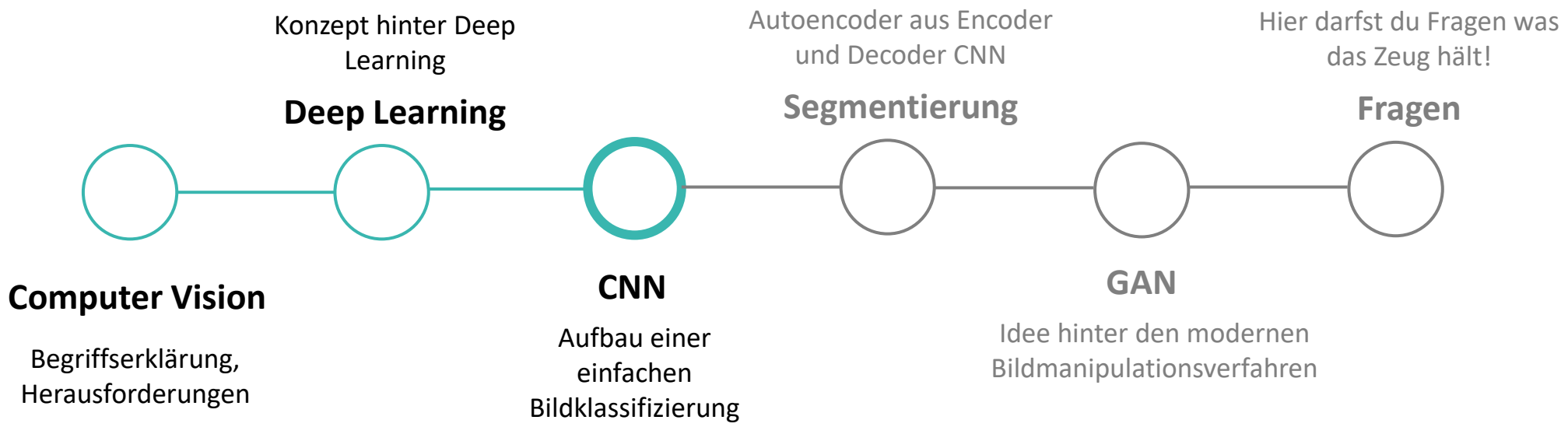
Quelle: (Rashid, T. & Langenau, F., 2017)
Quelle: (informatik-aktuell, n. D.)

Neuronales Netz (Training und Vorhersage)

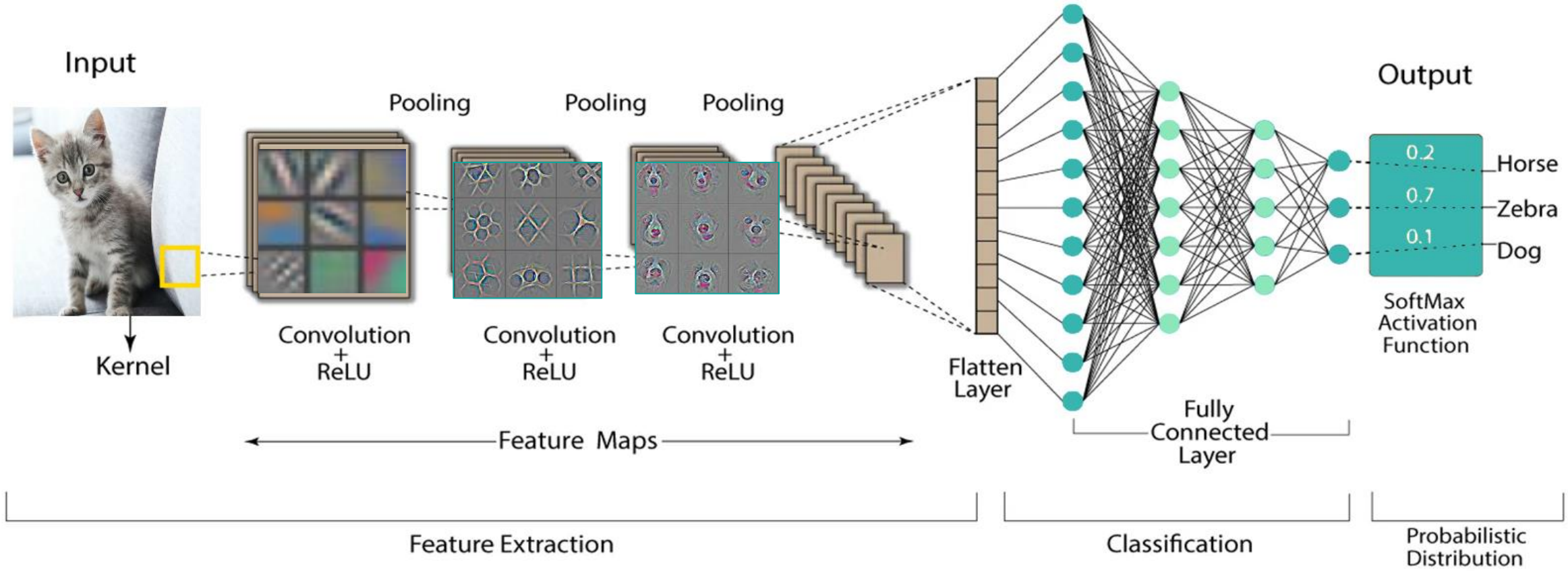


Quelle: (Rashid, T. & Langenau, F., 2017)
Quelle (CIFAR-10, 2014)

Fortschritt



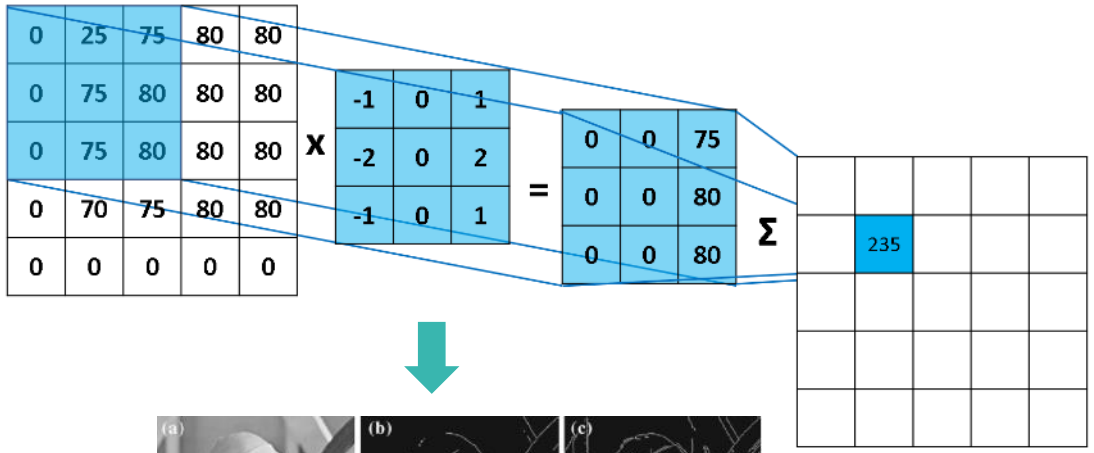
CNN - Convolutional Neural Network

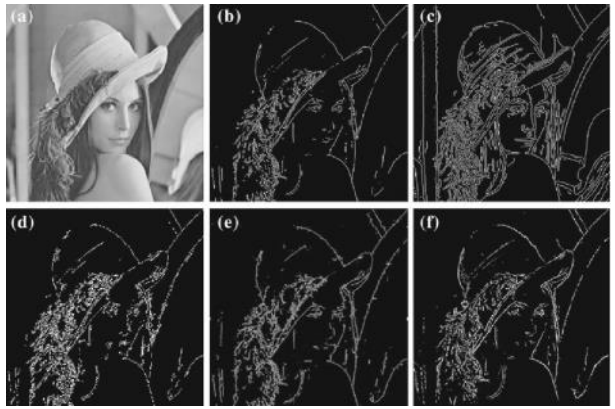


Quelle: (Swapna, 2021)
Quelle: (Zeiler A. & Fergus R., 2013)

Merkmalsextraktion

Convolutional

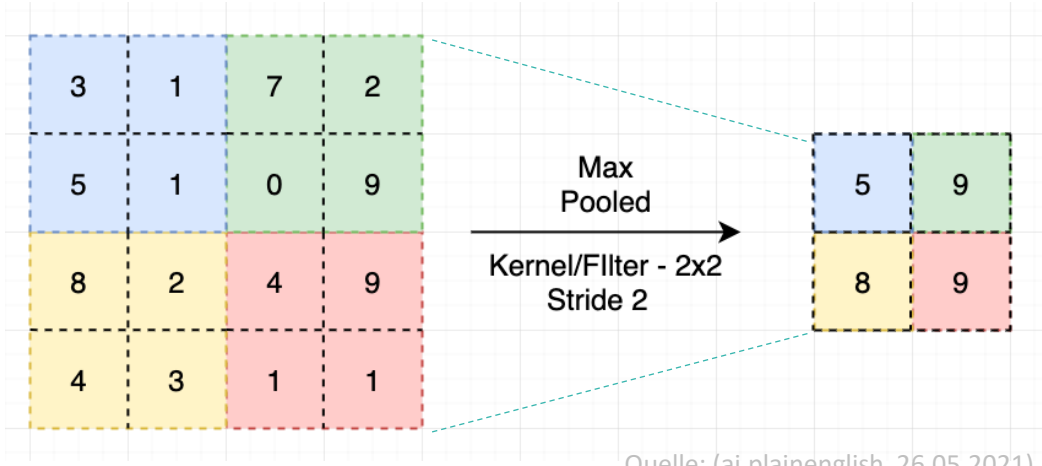




Quelle: (Verma, O., et al., 2013)

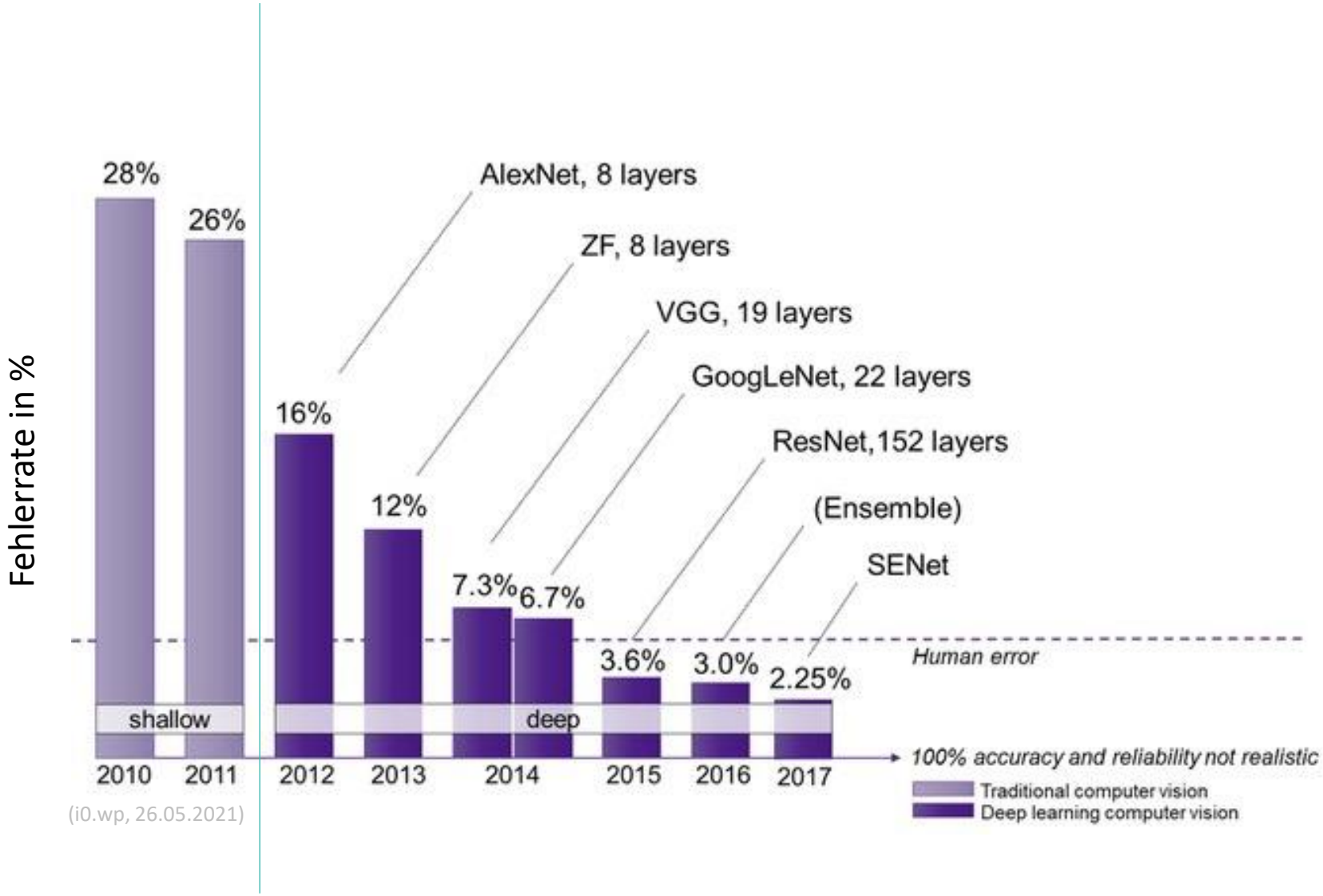
Quelle: (mlnotebook, 26.05.2021)

Pooling



Quelle: (ai.plainenglish, 26.05.2021)

IMAGENET

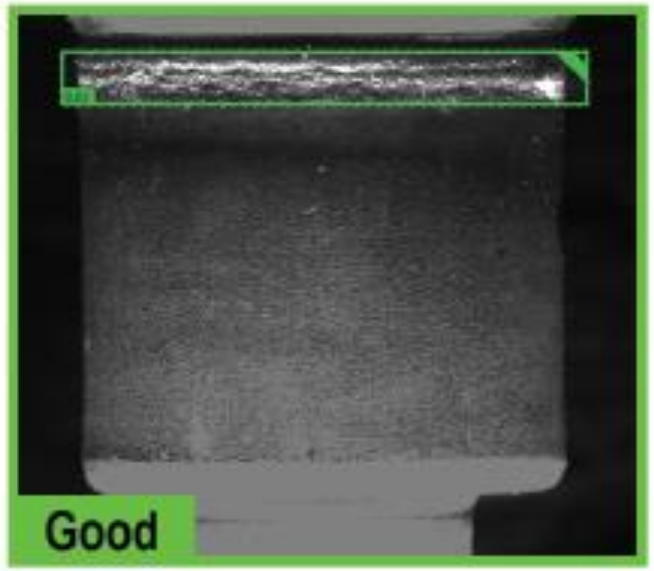


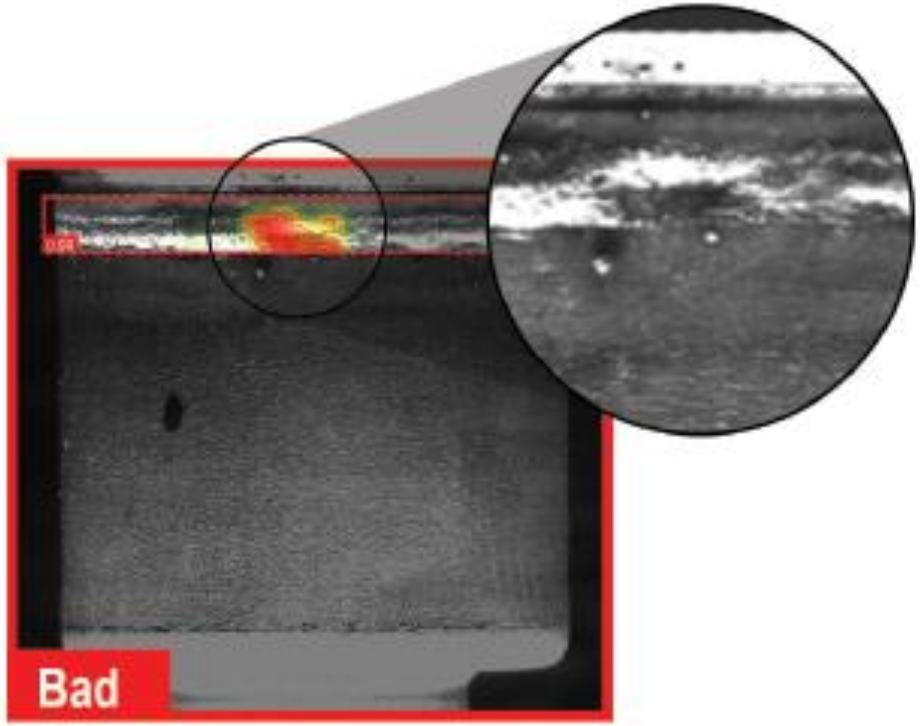
(i0.wp, 26.05.2021)

1.2M Bilder
1000 Klassen

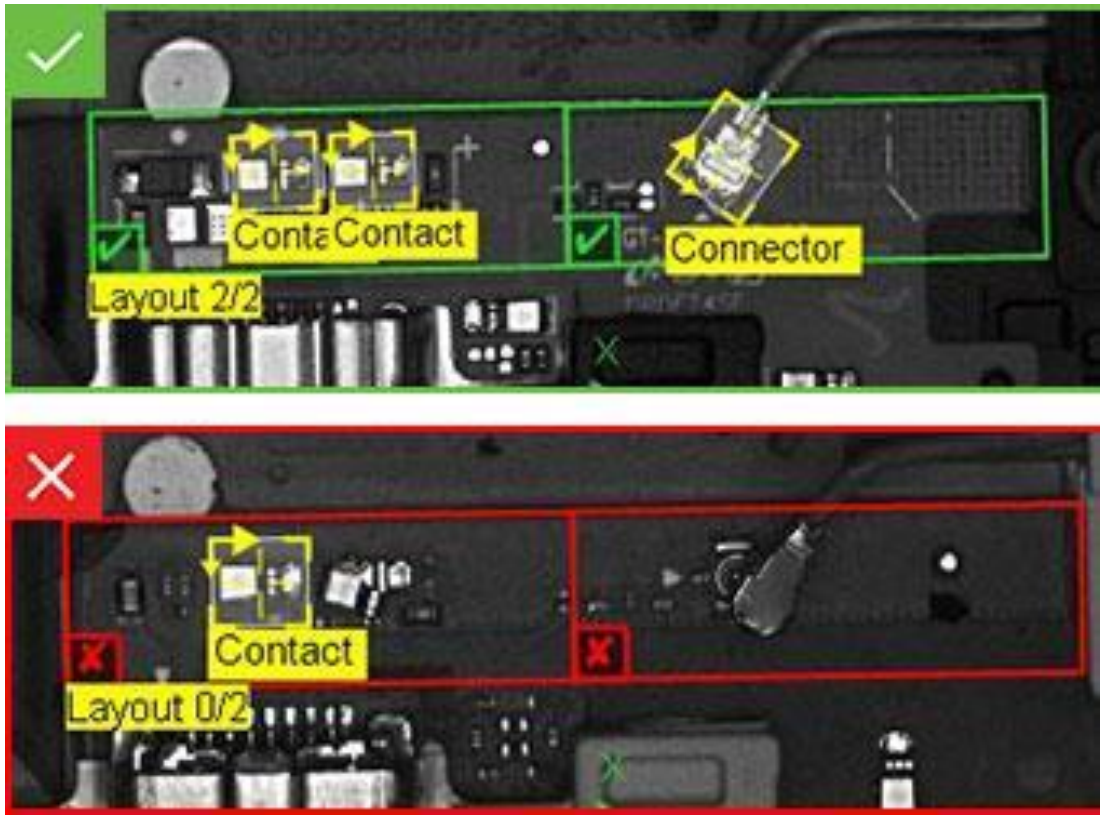
(community.cadence, 26.05.2021)

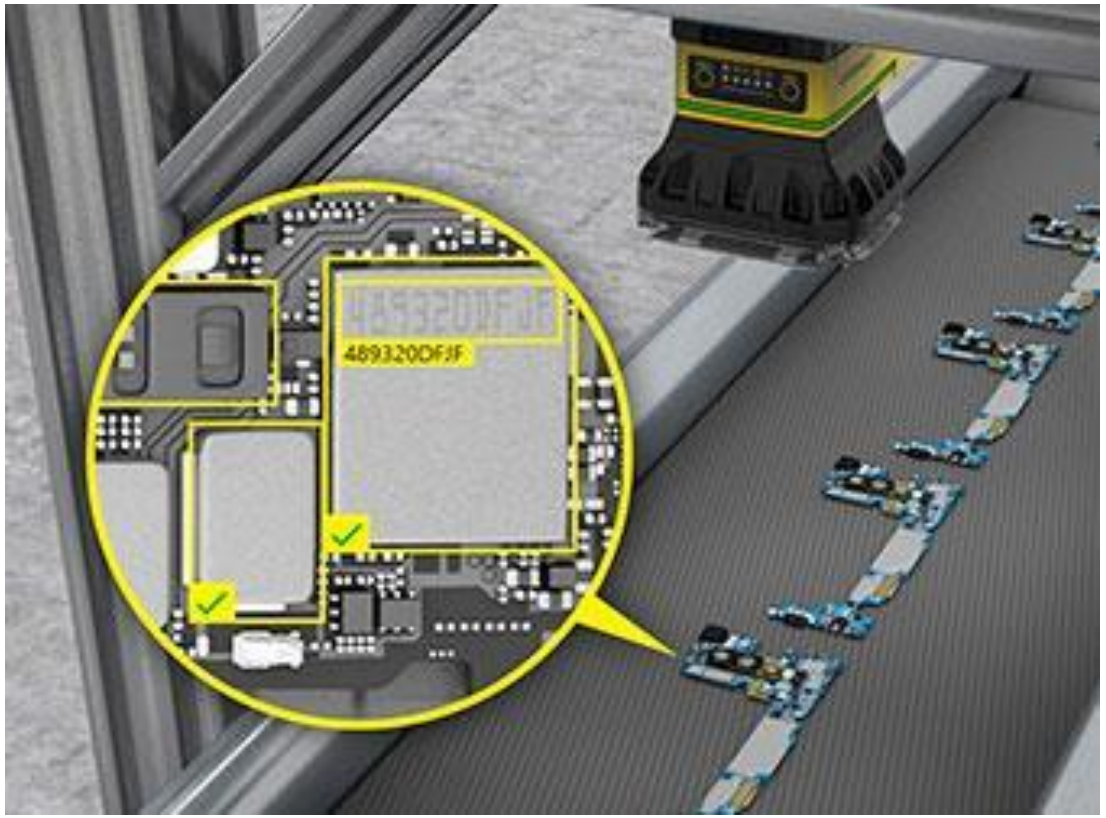
Praxisbeispiel in der Industrie



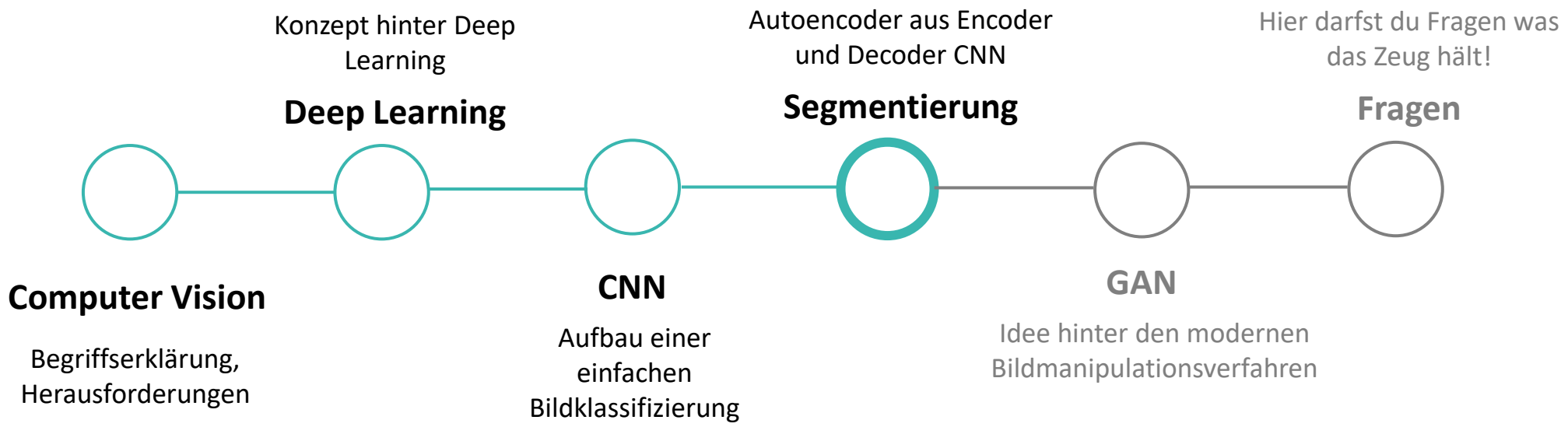


Praxisbeispiel in der Industrie

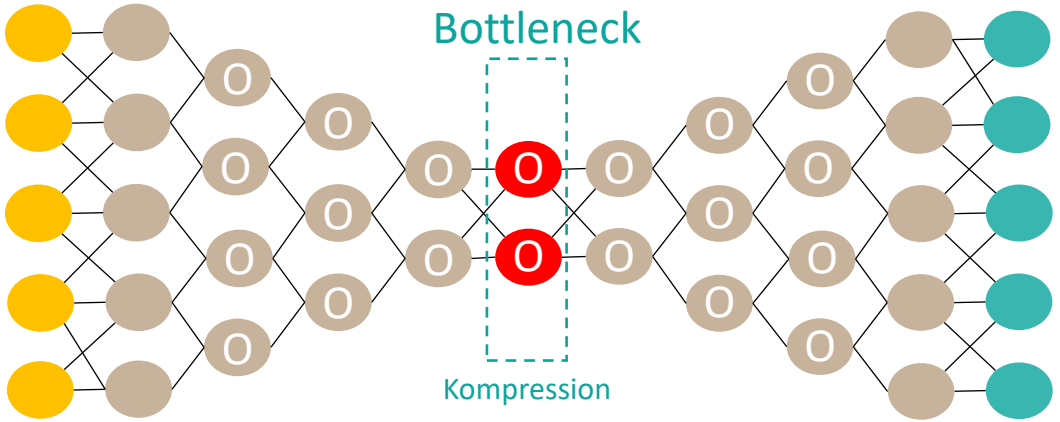


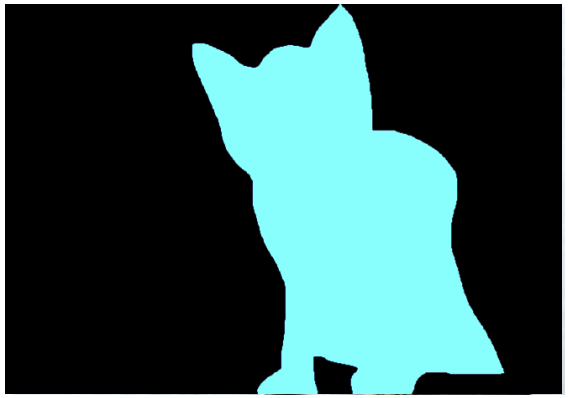


Fortschritt



Gedankenspiel



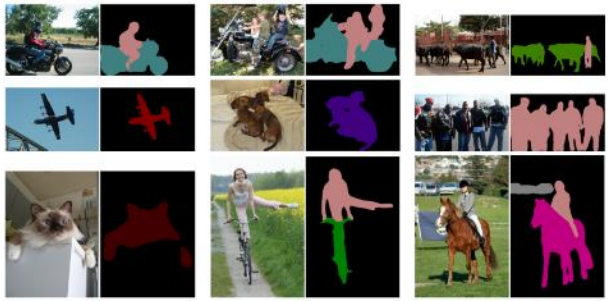


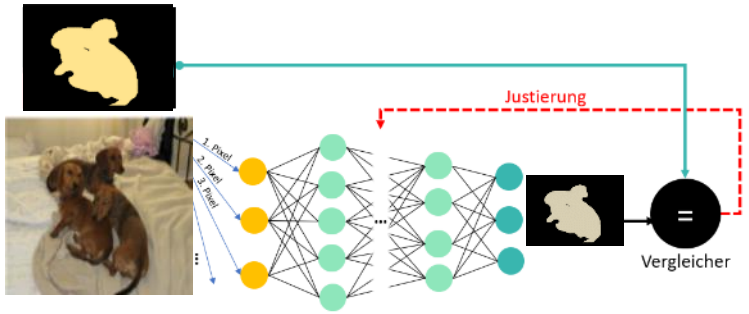
Eingabe

CNN Encoder

CNN^-1 Decoder

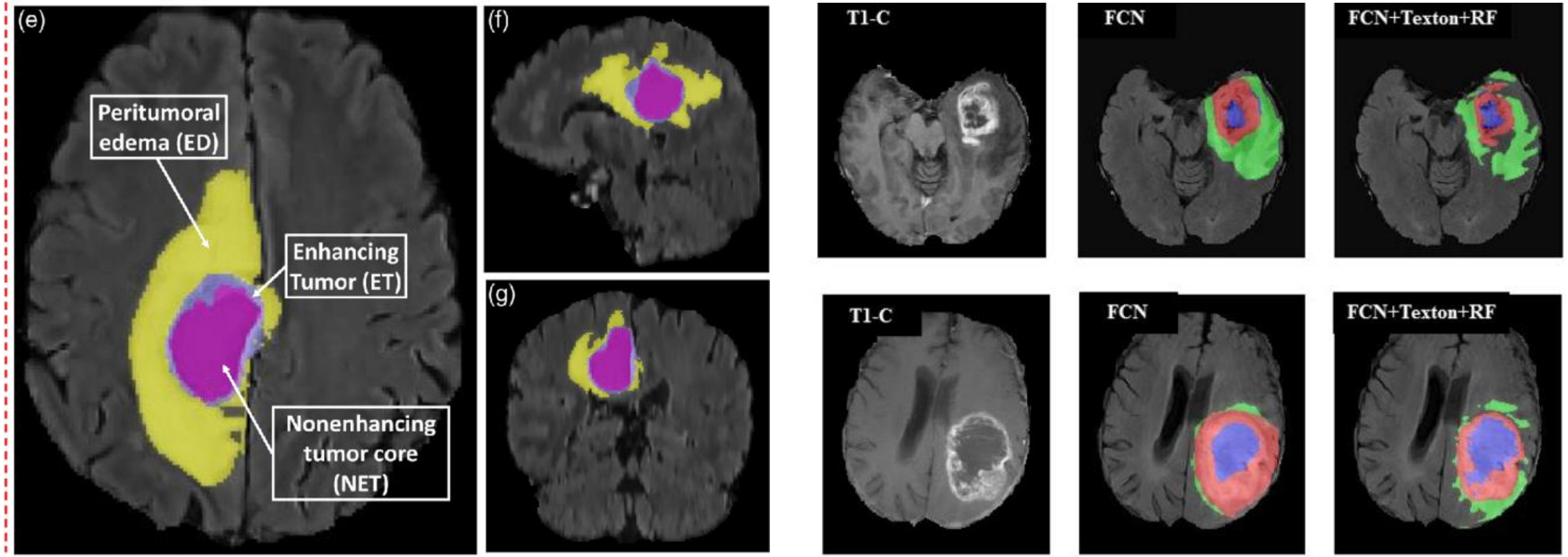
Ausgabe

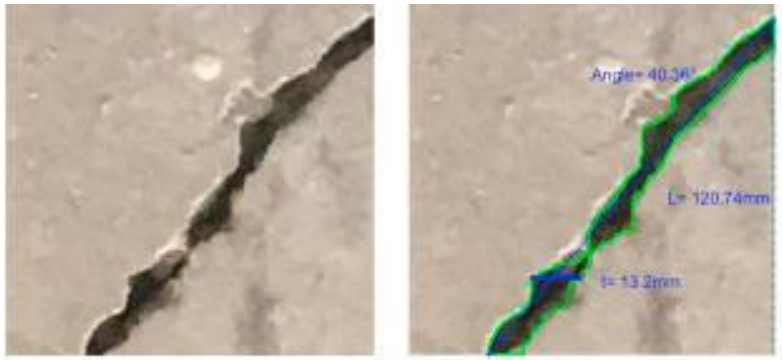




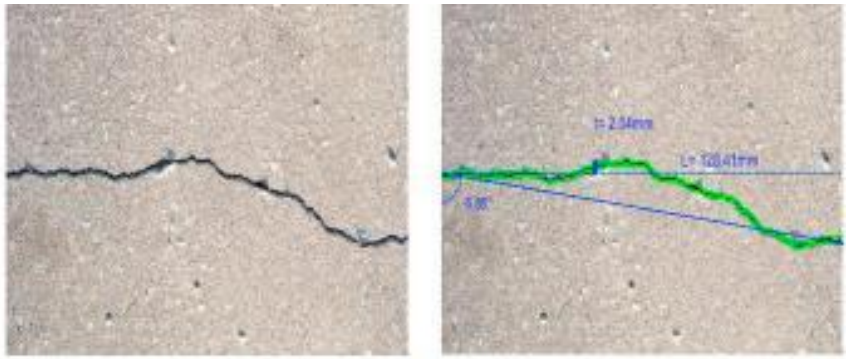
(Bank D., et. Al, 2021)

(Oles., A., 2018)

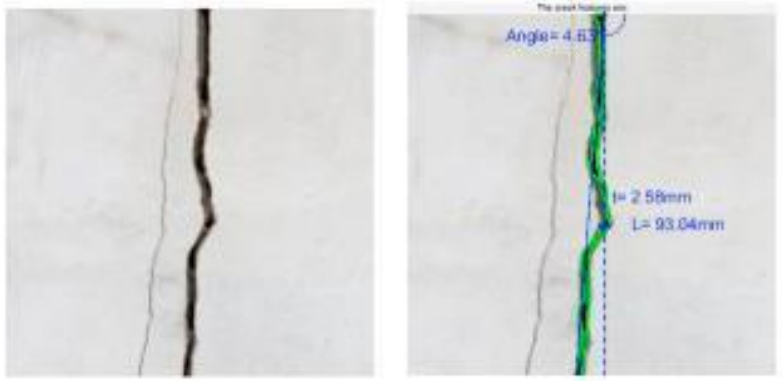

(e)
Peritumoral edema (ED)
Enhancing Tumor (ET)
Nonenhancing tumor core (NET)
(f)
(g)
T1-C
FCN
FCN+Texton+RF
T1-C
FCN
FCN+Texton+RF



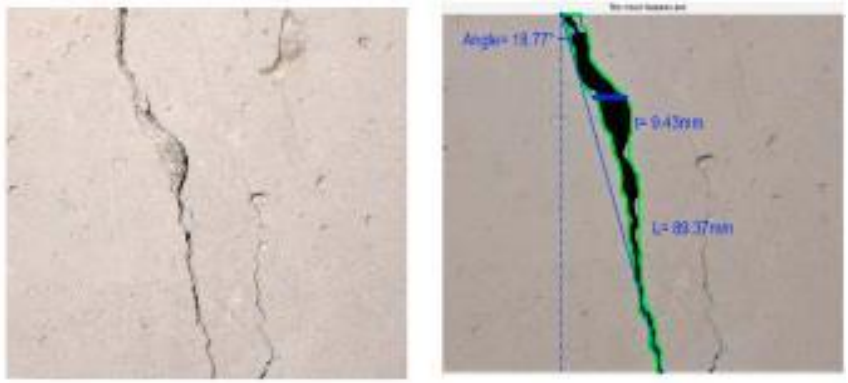
(a) Accurate quantification for a single VR crack



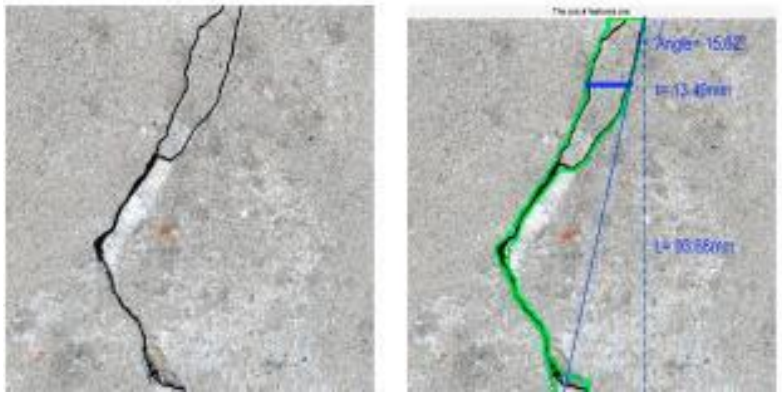
(b) Accurate quantification for a single HL crack

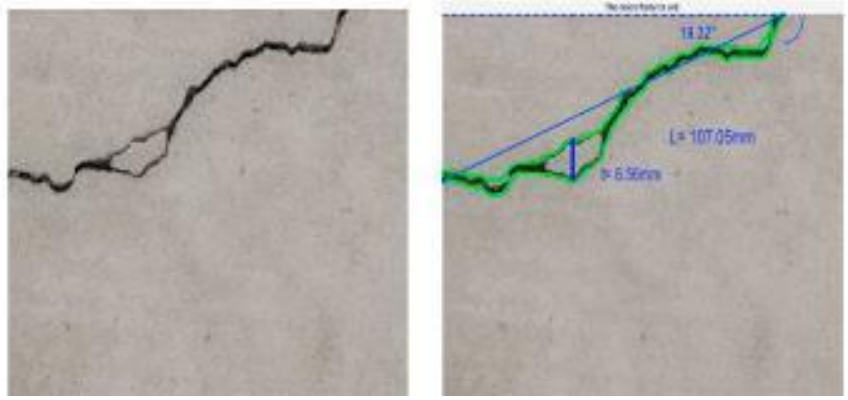


(c) Quantification of the widest crack for the case of non-intersecting VR cracks

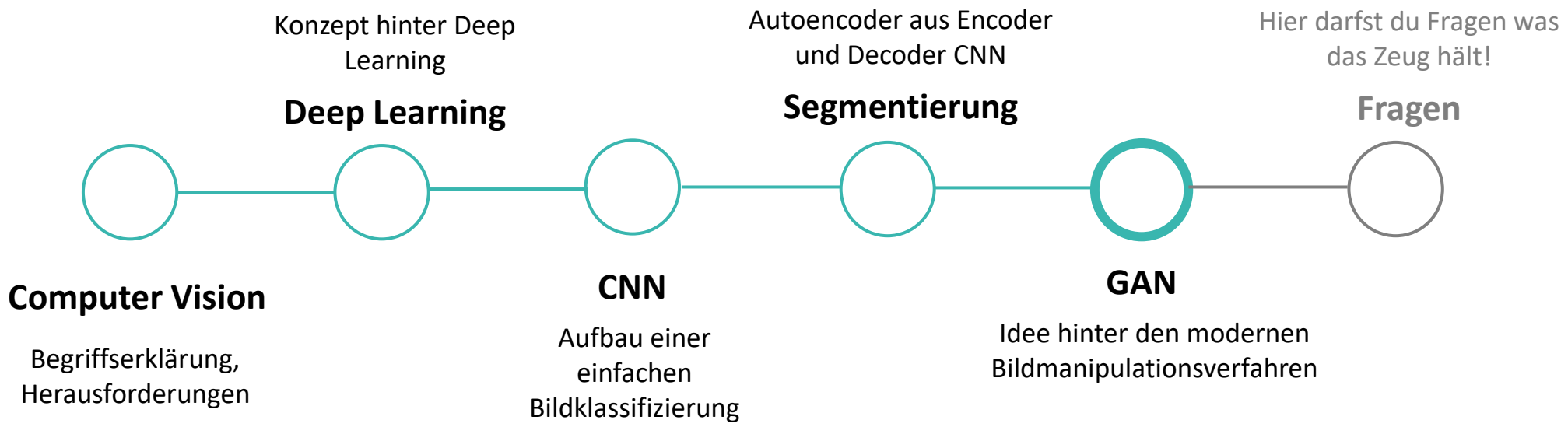


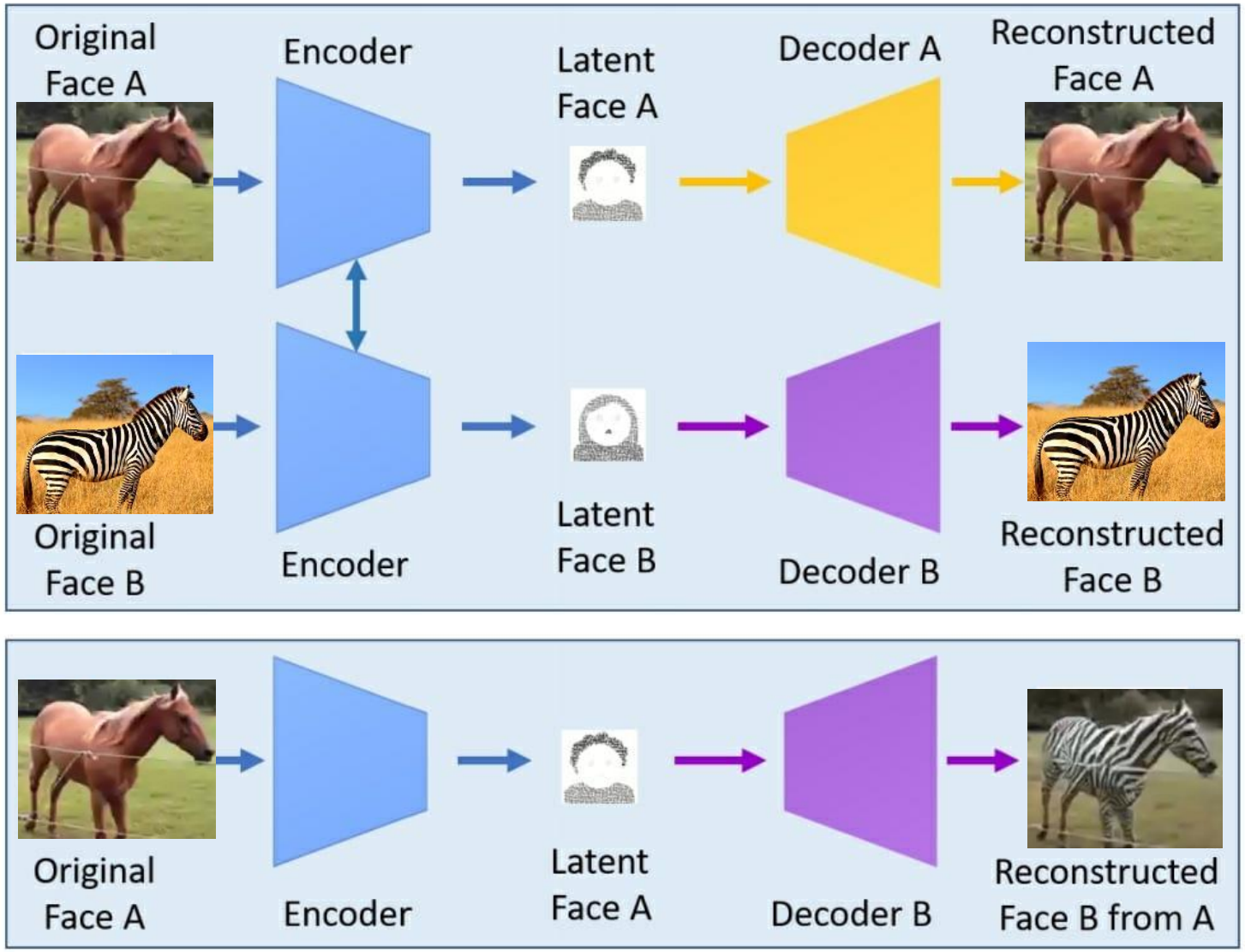
(d) Quantification of the widest crack for the case of non-intersecting VL cracks





Fortschritt




Original
Face A
Encoder
Latent
Face A
Decoder A
Reconstructed
Face A
Original
Face B
Encoder
Latent
Face B
Decoder B
Reconstructed
Face B
Original
Face A
Encoder
Latent
Face A
Decoder B
Reconstructed
Face B from A

GAN - Generative Adversarial Networks

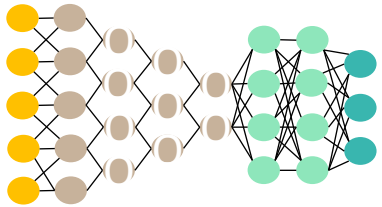
Generierendes Gegenläufiges Netzwerk

$$\min_G \max_D V(\mathrm{D},\mathrm{G}) = E_x\left[\log\left(D\left(x\right)\right)\right] + E_z\left[\log(1 - D\left(G(z)\right)\right]$$

Grundwahrheit

Diskriminator

Deep Convolutional Network (DCN)



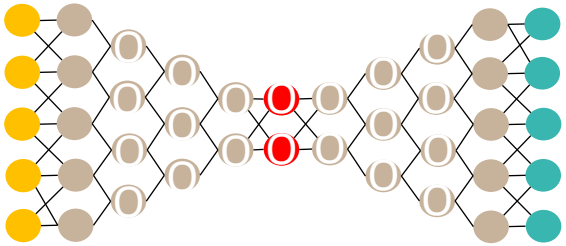
Real

FAKE

Eingabe

Generator

(Autoencoder)



Generiert

Quelle: Generative Adversarial Networks (I. Goodfellow et al., 2014)
Quelle: (M. Wäger ,2020)

CYCLE-GAN

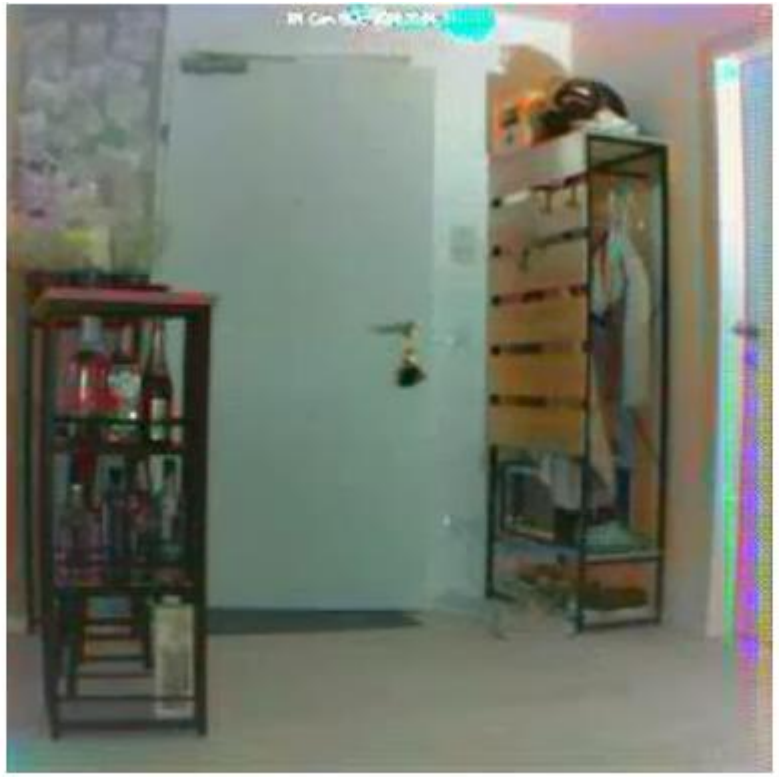

00:11,04

(Liu, et al., 2021)

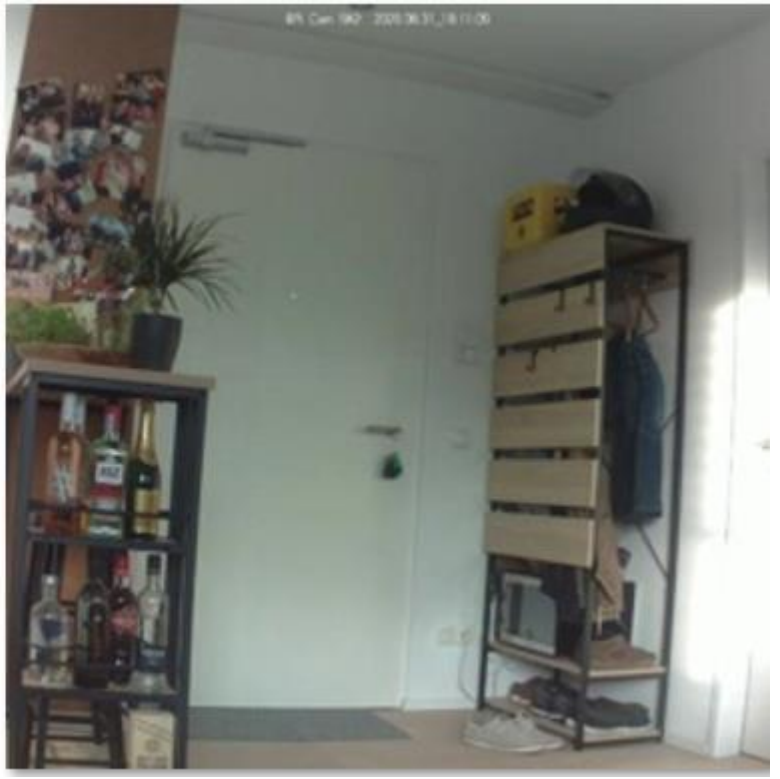
Video: https://www.youtube.com/watch?v=gg0F5JjKmhA

Eingabe

Generiert

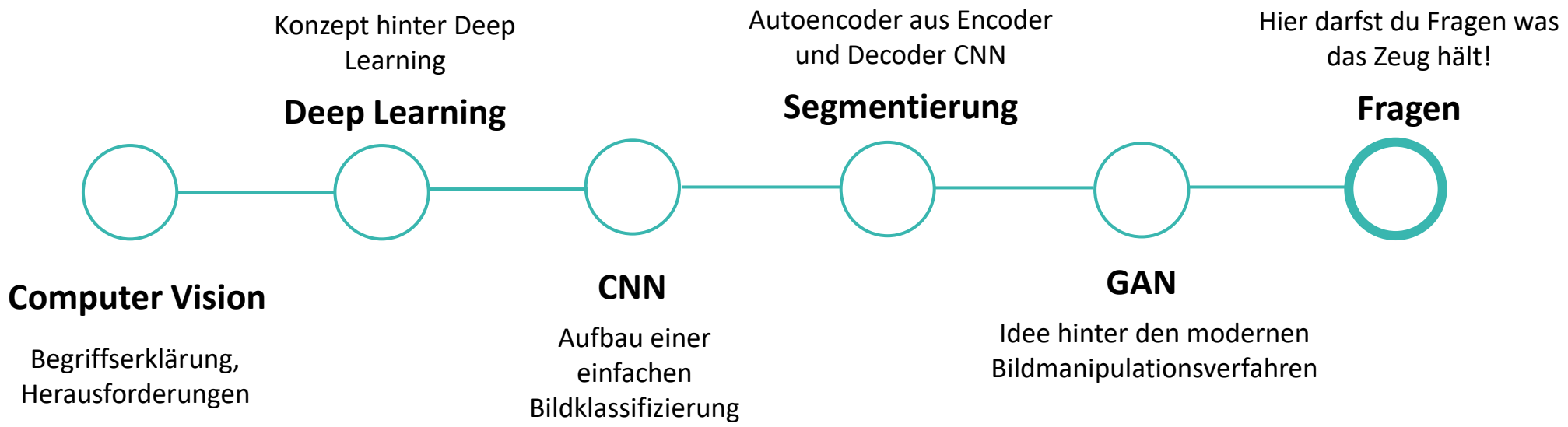


Grundwahrheit



Quelle: (M. Wäger ,2020)

Fortschritt

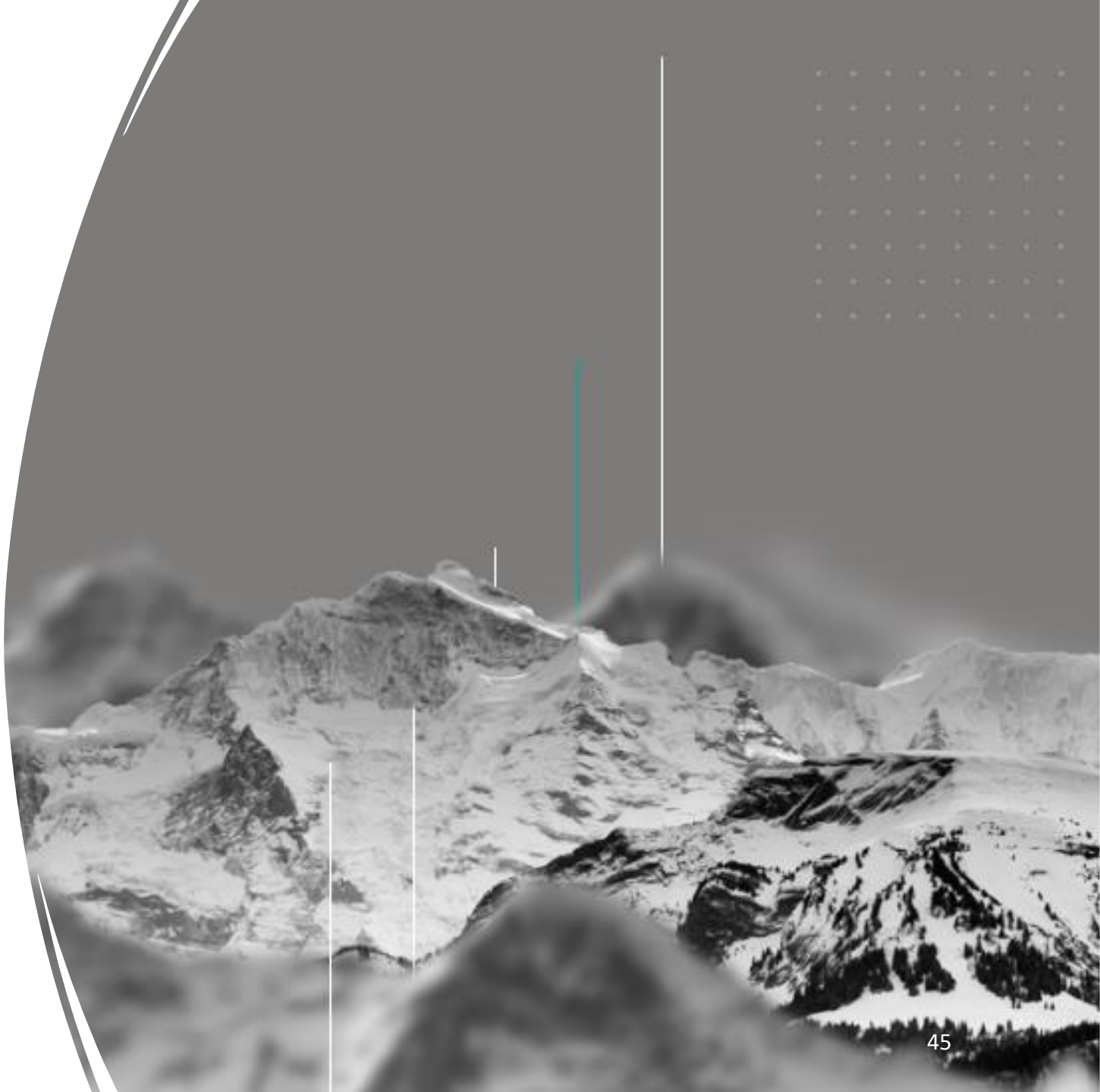


Fragen?

contact@deeprotype.com
+43/670 5550903

www.deeprotype.com

deeprotype



Referenzen

J. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2242-2251, doi: 10.1109/ICCV.2017.244.

Petrov I., Gao, D., Chervoniy, Nikolay, Liu, Kunlin, Marangonda, Sugasa, Umé, Chris, Jiang, Jian, RP, Luis, Zhang, Sheng, Wu, Pingyu, Zhang, Weiming. (2020). DeepFaceLab: A simple, flexible and extensible face swapping framework.

Wagner, F., A. Sánchez, Y. Tarabalka, R. G. Lotte, Matheus Pinheiro Ferreira, M. P. Aidar, E. Gloor, O. Phillips and L. Aragão. "Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images." Remote Sensing in Ecology and Conservation 5 (2019): 360-375.

*Liu, G., F. Reda, K. Shih, T. Wang, A. Tao, B. Catanzaro. "*Image Inpainting for Irregular Holes UsingPartial Convolutions*" ECCV 2018*, 2018

*Bank D., Koenigstein N., Giryes, R. "*Autoencoders*" Eprint 2003.05991, 202*

Programmersought.com. 2021. *Deep learning to change faces: the faceswap project practice - Programmer Sought*. [online] Available at: <https://www.programmersought.com/article/5207825274/> [Accessed 27 May 2021].

Sugahara, M., 2021. *misu007/einstein-vision-object-detection-annotation-tool-sample*. [online] GitHub. Available at: <https://github.com/misu007/einstein-vision-object-detection-annotation-tool-sample> [Accessed 26 May 2021].

Lin TY. et al. (2014) Microsoft COCO: Common Objects in Context. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham. https://doi.org/10.1007/978-3-319-10602-1_48

Cooper, G., 2019. *New Vision Technologies For Real-World Applications*. [online] Semiconductor Engineering. Available at: <https://semiengineering.com/new-vision-technologies-for-real-world-applications/> [Accessed 26 May 2021].

Oles., A., „Fast Semantic Segmentation“. [online] Available at https://github.com/oandrienko/fast-semantic-segmentation, 2018

Deng, J. et al., 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. pp. 248–255.

LeCun, Y. & Cortes, C. (2010), 'MNIST handwritten digit database', .

Verma, O., Hanmandlu, Madasu, Sultania, Ashish, Parihar, Anil. (2013). [Image] A novel fuzzy system for edge detection in noisy image using bacterial foraging. Multidimensional Systems and Signal Processing. 24. 10.1007/s11045-011-0164-1.

Tensorflow. 2021. *Machine Learning Zero to Hero (Google I/O'19)*. [Image] Available at: <https://www.youtube.com/watch?v=VwVg9jCtqaU> [Accessed 26 May 2021].

Rashid, T. & Langenau, F., 2017. Neuronale Netze selbst programmieren: Ein verständlicher Einstieg mit Python: Dpunkt.Verlag GmbH. Verfügbar unter: <https://books.google.at/books?id=_vwTMQAACAAJ> [Zugang am 2019-03-07].

Mvtec.com. 2021. *Klassische Bildverarbeitung vs. Deep Learning*. [online] Available at: <https://www.mvtec.com/de/technologien/deep-learning/klassische-bildverarbeitung-vs-deep-learning> [Accessed 26 May 2021].

CIFAR-10, A Krizhevsky, V Nair, G Hinton online: http://www. cs. toronto. edu/kriz/cifar. 2014

Wäger, M., "Maschinelle Bildkolorierung von IR-Nachtaufnahmen mithilfe zusätzlicher Farbinformationen", 2020

Zeiler, M. & Fergus R., " Visualizing and Understanding Convolutional“ Networks, 2013

Saharan, A. (2020). Creating a Human Pose Estimation Application with NVIDIA DeepStream [Human Pose Estimation]. https://developer.nvidia.com/blog/creating-a-human-pose-estimation-application-with-deepstream-sdk/

Karras, T. Laine, S. & Aila, T., A Style-Based Generator Architecture for Generative Adversarial Networks, 2019

Abbildungsverzeichnis

hamburger-tierschutzverein, n.d. [image] Available at: <: https://www.hamburger-tierschutzverein.de/images/images/0983_2020-Q3/Katze_Odessa-4zu3.jpg> [Accessed 26 May 2021].

appen, “Segmentierung Von Objekten.” Https://Appen.com/, appen.com/wp-content/uploads/2019/04/SLIDER-Appen_image_annotation_05.jpg.

mathworks, n.d. [image] Available at: <https://it.mathworks.com/help/matlab/matlab_prog/ch_data_types8.gif> [Accessed 26 May 2021].

hslu, (2021) Katzenbild. [image]. Available at: https://blog.hslu.ch/majorobm/files/2019/02/Titelbild.jpg [Accessed 26 May 2021].

katzen.de, 2021. [image] Available at: <https://www.katzen.de/magazin/wp-content/uploads/2015/02/Fotolia_57635875_S-b.jpg> [Accessed 26 May 2021].

iformatik-aktuell, " Vergleich biologisches Neuron mit künstlichem Neuron" , n. D. [image]. https://www.informatik-aktuell.de/fileadmin/templates/wr/pics/Artikel/03_Betrieb/abb1_neuron_schaaf.png, [Accessed 26 May 2021].

*miro.medium.com. (n.d.). [Image] https://miro.medium.com/max/2000/1*M-r4QQDrWfwGzLS8YSc0ng.jpeg*

Pubs.Rsna [Computer Vision in der Medizin]. (2019). Pubs.Rsna.Org. [Image] https://pubs.rsna.org/cms/10.1148/ryai.2019180001/asset/images/medium/ryai.2019180001.fig4.gif

Techgenyz. (n.d.). Techgenyz. https://www.techgenyz.com/wp-content/uploads/2018/06/pokemon-go-augmented-reality.jpg