Neuronale Netze

Funktionen

Definition: Gegeben seien die Mengen

$$D \text{ (Definitionsbereich) und } W \text{ (Wertebereich).}$$

Eine *F*unktion $f : D \to W$ ordnet jedem $x \in D$ genau ein $y = f(x) \in W$ zu:

$$D \ni x \quad \longrightarrow \quad y = f(x) \in W.$$

Satz: Eine Funktion $f : D \to W$ ist äquivalent charakterisiert durch

a) die Auswertungsvorschrift von f oder

b) den Graph $G = \{(x, f(x)) \in D \times W \mid x \in D\}$ von f.

Einfache und „komplizierte“ Funktionen

$$\begin{array}{ccccccccc} \mathbb{R} & \ni & x & \longrightarrow & y & = & x^2 & \in & \mathbb{R} \\ \mathbb{R} & \ni & x & \longrightarrow & y & = & e^{-x^2} & \in & \mathbb{R} \end{array}$$

Einfache und „komplizierte“ Funktionen

$$
\begin{array}{rcccccll}
\mathbb{R} & \ni & x & \longrightarrow & y & = & x^2 & \in \mathbb{R} \\
\mathbb{R} & \ni & x & \longrightarrow & y & = & e^{-x^2} & \in \mathbb{R} \\
\text{Restaurant-Gäste} & \ni & x & \longrightarrow & y & = & \text{will } x \text{ bestellen} & \in \text{Speisekarte} \\
\text{Urlaubs-Fotos} & \ni & x & \longrightarrow & y & = & x \text{ zeigt ein Pferd} & \in \{\text{ja, nein}\} \\
\text{Proteomics-Daten} & \ni & x & \longrightarrow & y & = & x \text{ hat Krebs} & \in \{\text{ja, nein}\} \\
\text{Instagram-Daten} & \ni & x & \longrightarrow & y & = & x \text{ ist depressiv} & \in \{\text{ja, nein}\}
\end{array}
$$

Einfache und „komplizierte“ Funktionen

$\mathbb{R}$	$\ni$	x	$\longrightarrow$	y	$=$	x^2	$\in$	$\mathbb{R}$
$\mathbb{R}$	$\ni$	x	$\longrightarrow$	y	$=$	e^{-x^2}	$\in$	$\mathbb{R}$
Restaurant-Gäste	$\ni$	x	$\longrightarrow$	y	$=$	will x bestellen	$\in$	Speisekarte
Urlaubs-Fotos	$\ni$	x	$\longrightarrow$	y	$=$	x zeigt ein Pferd	$\in$	{ja, nein}
Proteomics-Daten	$\ni$	x	$\longrightarrow$	y	$=$	x hat Krebs	$\in$	{ja, nein}
Instagram-Daten	$\ni$	x	$\longrightarrow$	y	$=$	x ist depressiv	$\in$	{ja, nein}

Darstellung: $\mathbb{R}^n \ni x \ \to y = f(x) \in \mathbb{R}^m$ (Zahlen statt Objekte)

a) Auswertungsvorschrift? b) Graph $G = \{(x, f(x)) \mid x \in D\}$?

Einfache und „komplizierte“ Funktionen

$$
\begin{array}{rcccccll}
\mathbb{R} & \ni & x & \longrightarrow & y & = & x^2 & \in \ \mathbb{R} \\
\mathbb{R} & \ni & x & \longrightarrow & y & = & e^{-x^2} & \in \ \mathbb{R} \\
\text{Restaurant-Gäste} & \ni & x & \longrightarrow & y & = & \text{will } x \text{ bestellen} & \in \ \text{Speisekarte} \\
\text{Urlaubs-Fotos} & \ni & x & \longrightarrow & y & = & x \text{ zeigt ein Pferd} & \in \ \{\text{ja, nein}\} \\
\text{Proteomics-Daten} & \ni & x & \longrightarrow & y & = & x \text{ hat Krebs} & \in \ \{\text{ja, nein}\} \\
\text{Instagram-Daten} & \ni & x & \longrightarrow & y & = & x \text{ ist depressiv} & \in \ \{\text{ja, nein}\}
\end{array}
$$

Darstellung: $\mathbb{R}^n \ni x \to y = f(x) \in \mathbb{R}^m$ (Zahlen statt Objekte)

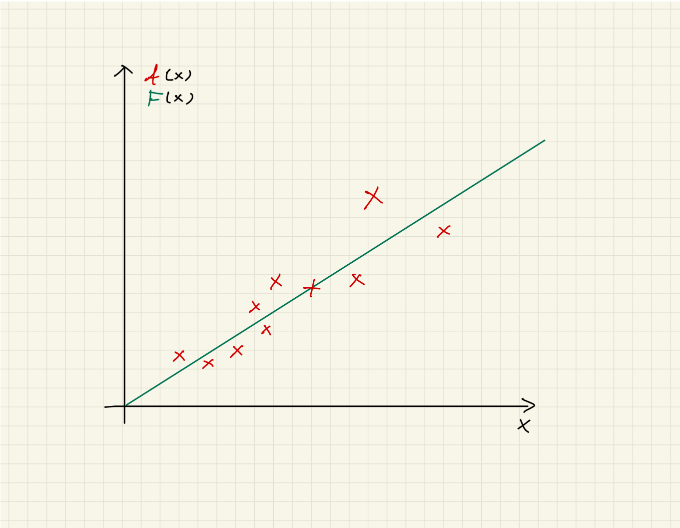
a) Auswertungsvorschrift? b) Graph $G = \{(x, f(x)) \mid x \in D\}$?

Ausweg: Approximation basierend auf Daten:

Wertetabelle: $G_N = \{(x, f(x)) \in D \times W \mid x \in D_N \subset D\}$

Ausgleichsproblem

Daten: $(x_i, y_i), \quad y_i = f(x_i). \qquad i = 1, \ldots, N$



parametrisierte Ansatzfunktion: $F_w(x) = wx, \quad w \in \mathbb{R}$

Bestimme den Parameter w so, dass F_w die Funktion f „gut" approximiert!

Ausgleichsproblem

Fehlerfunktional:

$$\mathcal{J}(w) = \sum_{i=1}^{N} (F_w(x_i) - y_i)^2 = \sum_{i=1}^{N} (w x_i - y_i)^2$$

Minimierungsproblem:

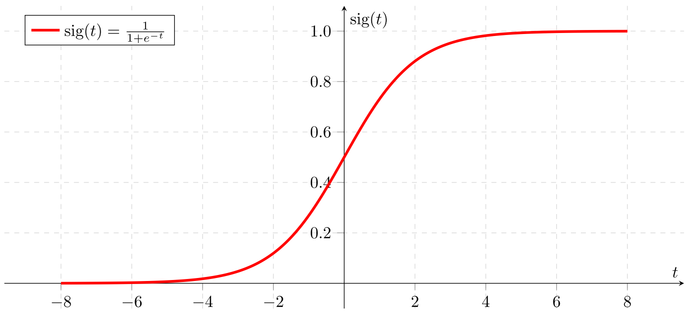
$$w \in \mathbb{R} : \qquad \mathcal{J}(w) \leq \mathcal{J}(v) \qquad \forall v \in \mathbb{R}$$

Lösung:

$$\mathcal{J}'(w) = 2 \sum_{i=1}^{N} (w x_i - y_i) x_i \stackrel{!}{=} 0 \qquad w = \left(\sum_{i=1}^{N} y_i x_i \right) \Big/ \left(\sum_{i=1}^{N} x_i^2 \right)$$

Neuronales Netz (1 Neuron)

Aktivierungsfunktion: $\sigma = 1/(1 + \kappa e^{-x})$



1 Schicht: $F_{w,b}(x) = \sigma(wx + b)$, Parameter $w,\ b \in \mathbb{R}$

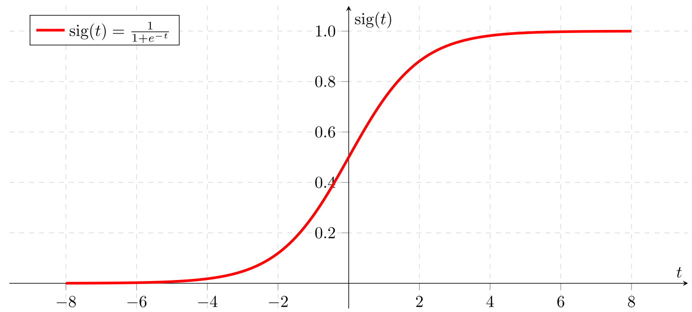
2 Schichten:

$$F_{W,b}(x) = \sigma(w_2\sigma(w_1x + b_1) + b_2) = g_2 \circ g_1(x), \quad g_i(z) = \sigma(w_iz + b_i)$$

Parameter $w_1,\ w_2,\ b_1,\ b_2 \in \mathbb{R}$

Tiefes Neuronales Netz (Deep Neural Network, 1 Neuron)

Aktivierungsfunktion: $\sigma(x) = 1/(1 + \kappa e^{-x})$



L Schichten:

$$F_{W,b}(x) = g_L \circ g_{L-1} \circ \cdots \circ g_1(x), \qquad g_i(z) = \sigma(w_i z + b_i)$$

Parameter $w_i,\ b_i \in \mathbb{R}, \quad i = 1, \ldots, L$

Neuronales Netz

L Schichten (Deep Neural Network):

$$F_{W,b}(x) = g_L \circ g_{L-1} \circ \cdots \circ g_1(x), \qquad g_i(z) = \sigma(W_i z + b_i), \quad W_i \in \mathbb{R}^{n_i, n_{i-1}},\ b_i \in \mathbb{R}^{n_i}$$

von n_{i-1} nach n_i Neuronen:

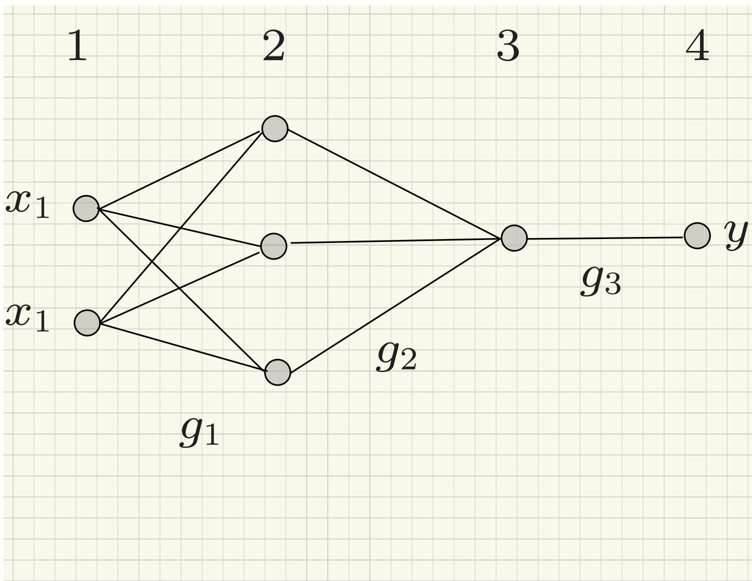
$$\mathbb{R}^{n_{i-1}} \ni z \mapsto g_i(z) = \sigma(W_i z + b_i) = \begin{pmatrix} \sigma((W_i z)_1 + b_{i,1}) \\ \vdots \\ \sigma((W_i z)_{n_i} + b_{i,n_i}) \end{pmatrix} \in \mathbb{R}^{n_i}$$

Parameter: $\qquad W = (W_1, \dots, W_L),\ b = (b_1, \dots, b_L)$

Neuronales Netz: Schematische Darstellung

4 Schichten (Deep Neural Network):

$$F_{W,b}(x) = g_3 \circ g_2 \circ g_1(x), \qquad g_i(z) = \sigma(W_i z + b_i),\ i = 1, 2, 3$$



Neuronales Netz: Parameter

4 Schichten (Deep Neural Network):

$$F_{W,b}(x) = g_3 \circ g_2 \circ g_1(x), \qquad g_i(z) = \sigma(W_i z + b_i),\ i = 1, 2, 3$$

Parameter: $W = (W_1, W_2, W_3)),\ b = (b_1, b_2, b_3)$

$$W_1 = \begin{pmatrix} W_{1,11} & W_{1,12} \\ W_{1,21} & W_{1,22} \\ W_{1,31} & W_{1,32} \end{pmatrix} \in \mathbb{R}^{3,2}, \quad b_1 = \begin{pmatrix} b_{1,1} \\ b_{1,2} \\ b_{1,3} \end{pmatrix} \in \mathbb{R}^3,$$

$$W_2 = \begin{pmatrix} W_{2,11} & W_{2,12} & W_{2,13} \end{pmatrix} \in \mathbb{R}^{1,3}, \quad b_2 = (b_{2,1}) \in \mathbb{R}$$

$$W_3 = (W_{3,11}) \in \mathbb{R}\ , \quad b_3 = (b_{3,1}) \in \mathbb{R}$$

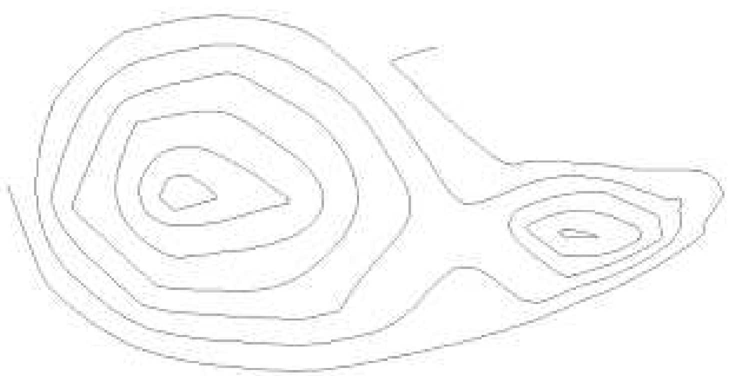
Bestimmung der Parameter: Ausgleichsproblem

Daten: (x_i, y_i), $\quad y_i = f(x_i), \quad i = 1, \ldots, N$

Fehlerfunktional:

$$\mathbb{R}^{n_1(n_0+1)+\cdots+n_L(n_{L-1}+1)} \ni \mathcal{J}(W,b) = \sum_{i=1}^{N} (F_{W,b}(x_i) - y_i)^2 \to \mathbb{R}$$

Methode des steilsten Abstiegs: Achtung: Lokale Minima!!



Anwendung: Diagnose von Krankheiten

Funktion: $\{\text{Proteom}\} \ni x \mapsto f(x) = y \in \{\text{Klassen}\}$

Deep Learning for Proteomics Data for Feature Selection and Classification

Sahar Iravani[1][0000–0002–8923–326X] and Tim O.F. Conrad[2,3][0000–0002–5590–5726]

[1] Zuse institute Berlin, Takustrasse 7, 14195 Berlin, Germany
iravani@zib.de
[2] Zuse institute Berlin, Takustrasse 7, 14195 Berlin, Germany
[3] Free University of Berlin, Arnimallee 6, 14195 Berlin, Germany
conrad@math.fu-berlin.de

Iravani S., Conrad T.O.F. (2019) Deep Learning for Proteomics Data for Feature Selection and Classification. In: Holzinger A., Kieseberg P., Tjoa A., Weippl E. (eds) Machine Learning and Knowledge Extraction. CD-MAKE 2019. Lecture Notes in Computer Science, vol 11713. Springer, Cham.

Anwendung: Diagnose von Depressionen

Funktion: $\{\text{Instagram-Daten}\} \ni x \mapsto f(x) = y \in \{\text{depressiv, nicht depressiv}\}$

Daten: bekannte Diagnosen $f(x_i) = y_i,\ i = 1, \ldots, N$

This AI System Can Diagnose Depression From Instagram Photos

Researchers created a machine-learning algorithm that was able to identify clinical depression with 70 percent accuracy, while human general practitioners achieved a 42 percent success rate.

By Glenn McDonald Published on 8/9/2017 at 2:25 PM

Heads up Instagram users — the pictures you post may be more revealing than you think.

This week, researchers at Harvard University and the University of Vermont released a study that suggests artificial intelligence systems can identify a depressed individual simply by looking at their Instagram photos.

According to the researchers, their algorithm has a 70 percent success rate when determining which Instagram users have been diagnosed as clinically depressed within the last three years. By comparison, general practitioners have about a 42 percent succes rate when diagnosing depression through in-person evaluations. While those figures aren't particularly rigorous from a statistical point of view — more on that in a bit — they do suggest that AI could be useful in screening for clinical depression.

See<er (2017)

Regular article | Open Access | Published: 08 August 2017

Instagram photos reveal predictive markers of depression

Andrew G Reece & Christopher M Danforth

EPJ Data Science 6, Article number: 15 (2017) | Cite this article

204k Accesses | 83 Citations | 3411 Altmetric | Metrics

This article has been updated

The Erratum to this article has been published in EPJ Data Science 2017 6:21

Abstract

Using Instagram data from 166 individuals, we applied machine learning tools to successfully identify markers of depression. Statistical features were computationally extracted from 43,950 participant Instagram photos, using color analysis, metadata components, and algorithmic face detection. Resulting models outperformed general practitioners' average unassisted diagnostic success rate for depression. These results held even when the analysis was restricted to posts made before depressed individuals were first diagnosed. Human ratings of photo attributes (happy, sad, etc.) were weaker predictors of depression, and were uncorrelated with computationally-generated features. These results suggest new avenues for early screening and detection of mental illness.

A.G. Reece., C.M. Danforth: Instagram photos reveal predictive markers of depression. EPJ Data Sci. 6, 15 (2017).

Anwendung: Diagnose von Depressionen

A.G. Reece., C.M. Danforth: Instagram photos reveal predictive markers of depression. EPJ Data Sci. 6, 15 (2017).

Abstract Using Instagram data from 166 individuals, we applied machine learning tools to successfully identify markers of depression. Statistical features were computationally extracted from 43,950 participant Instagram photos, using color analysis, metadata components, and algorithmic face detection. Resulting models outperformed general practitioners's average unassisted diagnostic success rate for depression. These results held even when the analysis was restricted to posts made before depressed individuals were first diagnosed. Human ratings of photo attributes (happy, sad, etc.) were weaker predictors of depression, and were uncorrelated with computationally-generated features. These results suggest new avenues for early screening and detection of mental illness.

Anwendung: Urlaubs-Fotos

Klassifizierung von Urlaubs-Fotos: Bild zeigt einen: Panda, Stopschild,...

Funktion: $\{\text{Pixel-Matrizen}\} \ni x \mapsto f(x) = y \in \{\text{Klassen}\}$

Daten: klassifizierte Urlaubs-Fotos $f(x_i) = y_i,\ i = 1, \ldots, N$

Anwendung: Urlaubs-Fotos

Klassifizierung von Urlaubs-Fotos: Bild zeigt einen: Panda, Stopschild,...

Funktion: $\{\text{Pixel-Matrizen}\} \ni x \mapsto f(x) = y \in \{\text{Klassen}\}$

Daten: klassifizierte Urlaubs-Fotos $f(x_i) = y_i, \; i = 1, \ldots, N$

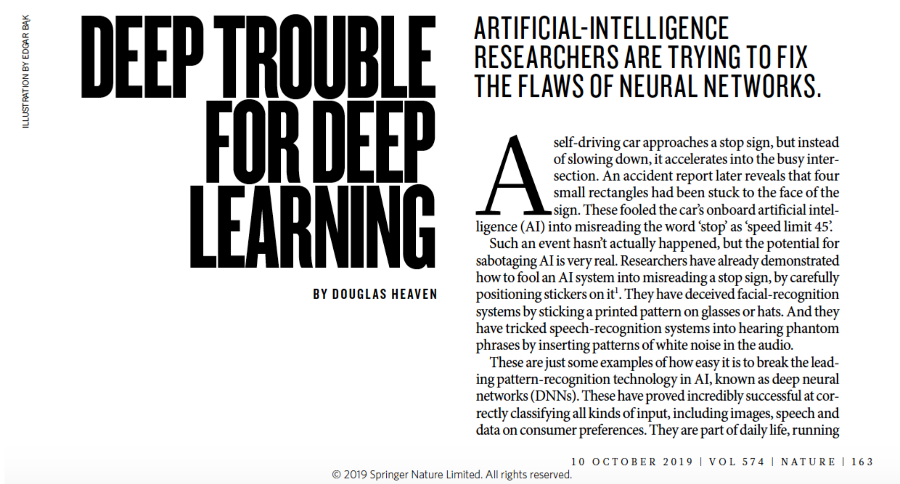


ILLUSTRATION BY EDGAR BĄK

DEEP TROUBLE FOR DEEP LEARNING

BY DOUGLAS HEAVEN

ARTIFICIAL-INTELLIGENCE RESEARCHERS ARE TRYING TO FIX THE FLAWS OF NEURAL NETWORKS.

A self-driving car approaches a stop sign, but instead of slowing down, it accelerates into the busy intersection. An accident report later reveals that four small rectangles had been stuck to the face of the sign. These fooled the car's onboard artificial intelligence (AI) into misreading the word 'stop' as 'speed limit 45'.

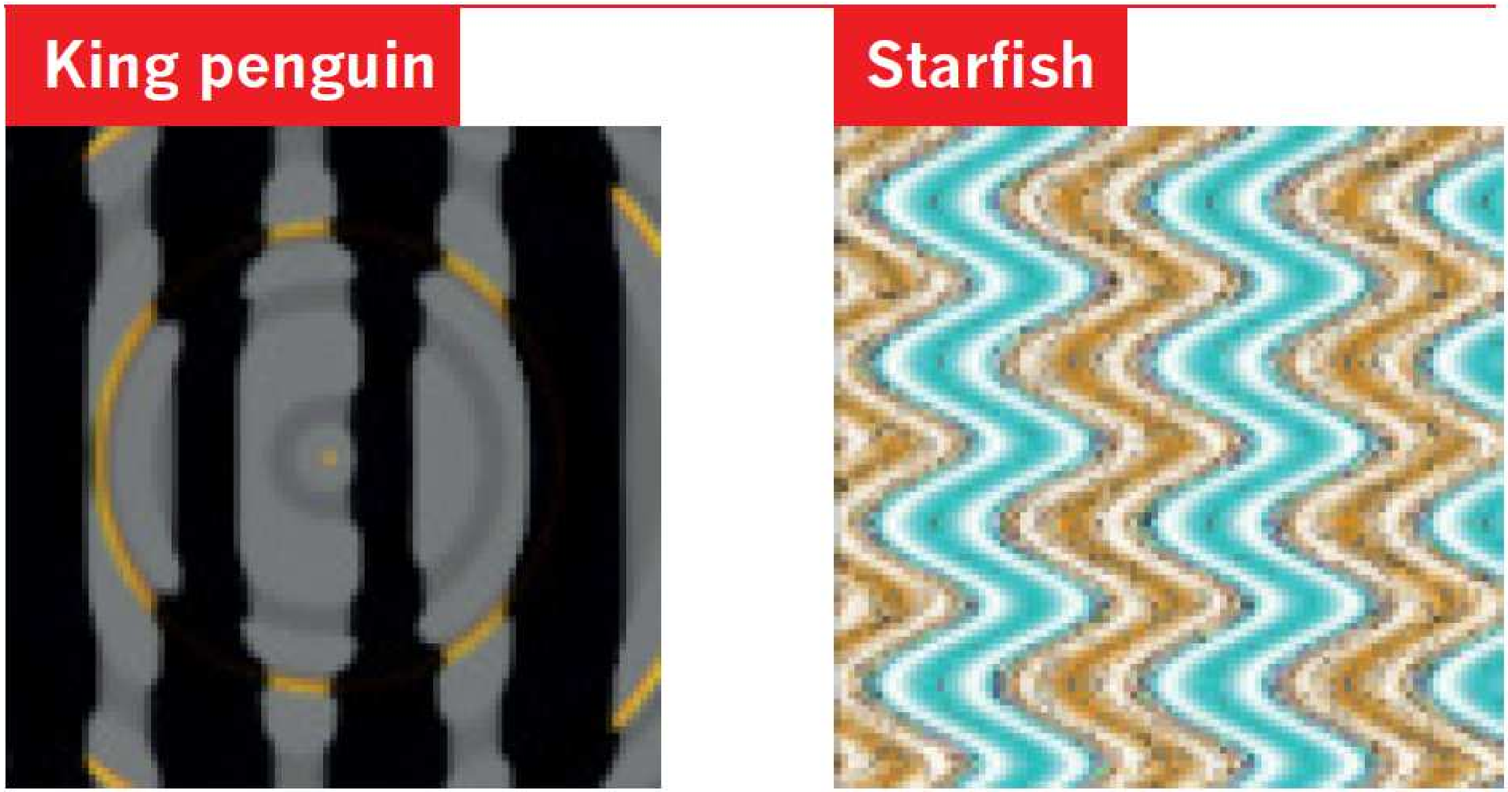
Such an event hasn't actually happened, but the potential for sabotaging AI is very real. Researchers have already demonstrated how to fool an AI system into misreading a stop sign, by carefully positioning stickers on it[1]. They have deceived facial-recognition systems by sticking a printed pattern on glasses or hats. And they have tricked speech-recognition systems into hearing phantom phrases by inserting patterns of white noise in the audio.

These are just some examples of how easy it is to break the leading pattern-recognition technology in AI, known as deep neural networks (DNNs). These have proved incredibly successful at correctly classifying all kinds of input, including images, speech and data on consumer preferences. They are part of daily life, running

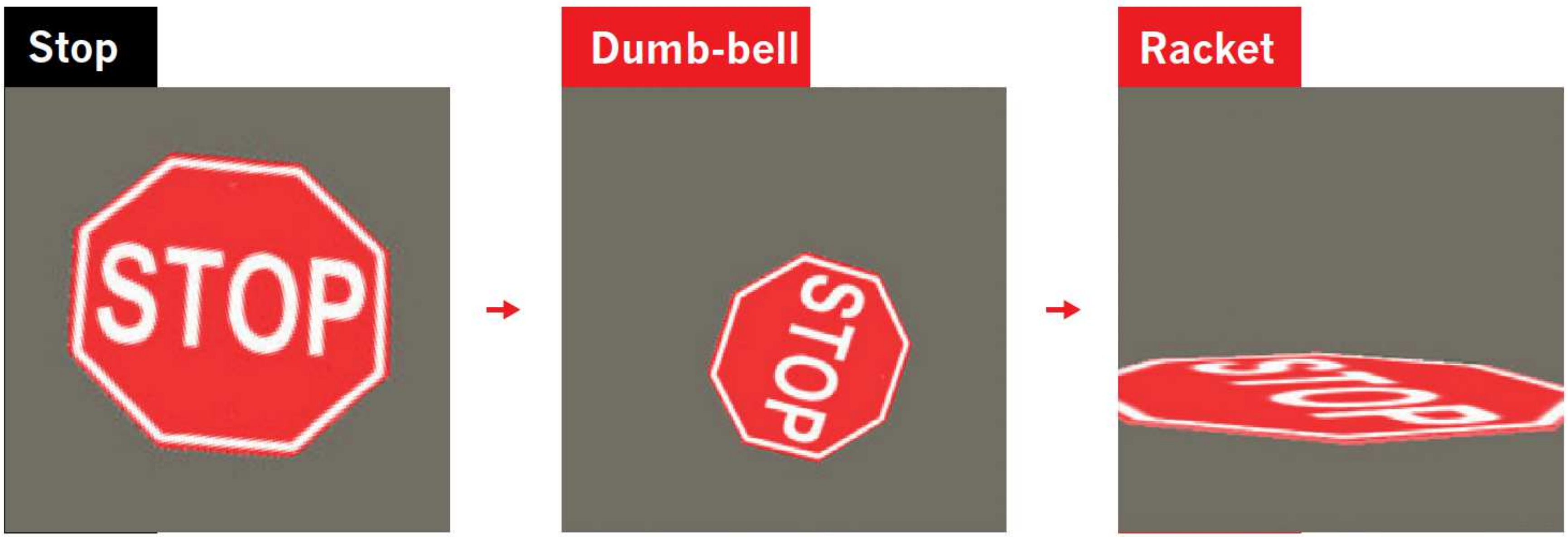
10 OCTOBER 2019 | VOL 574 | NATURE | 163

D. Heaven: Deep Trouble for Deep learning. Nature 163 (2019)

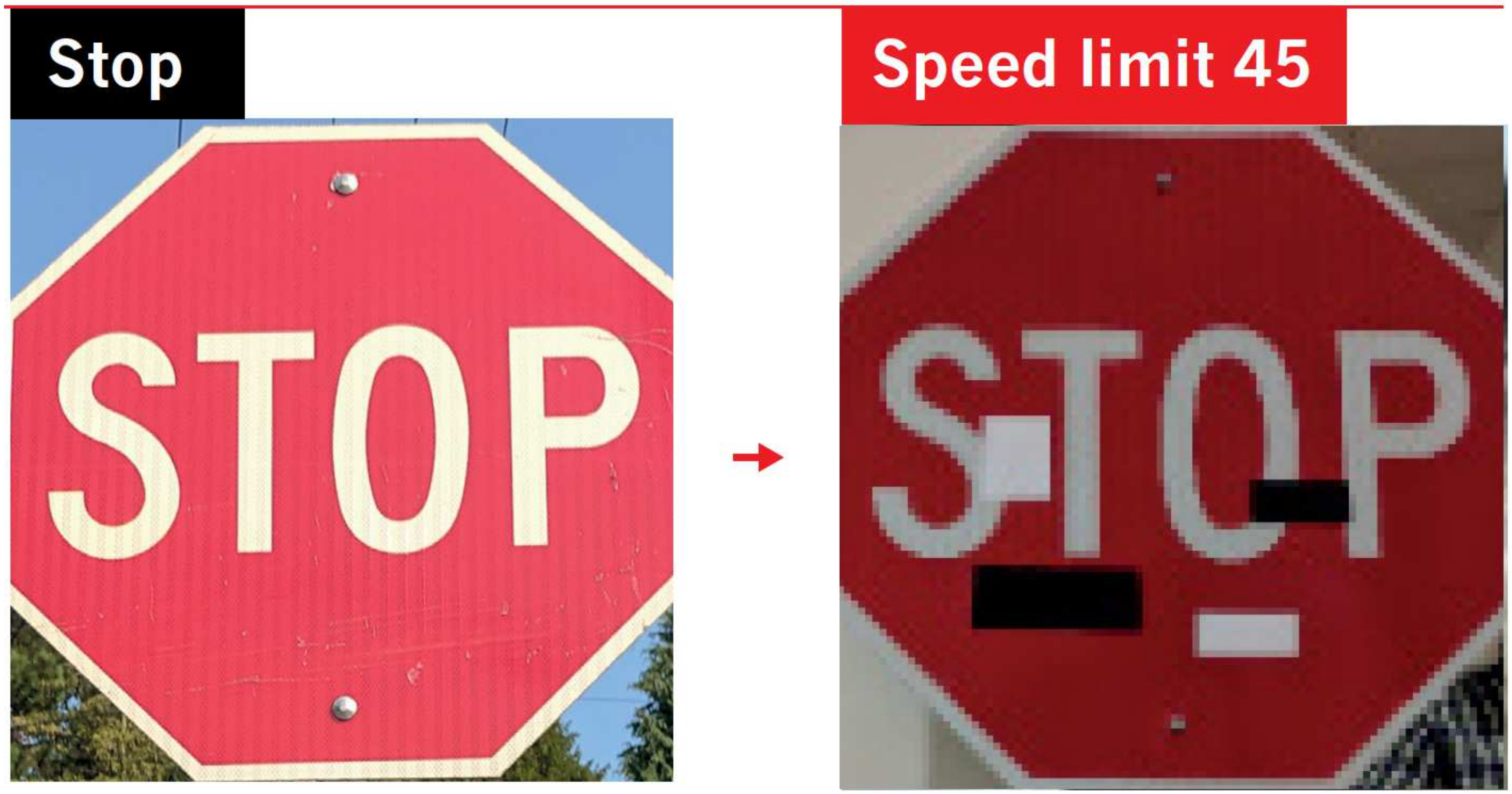
Tiefe Fantasie



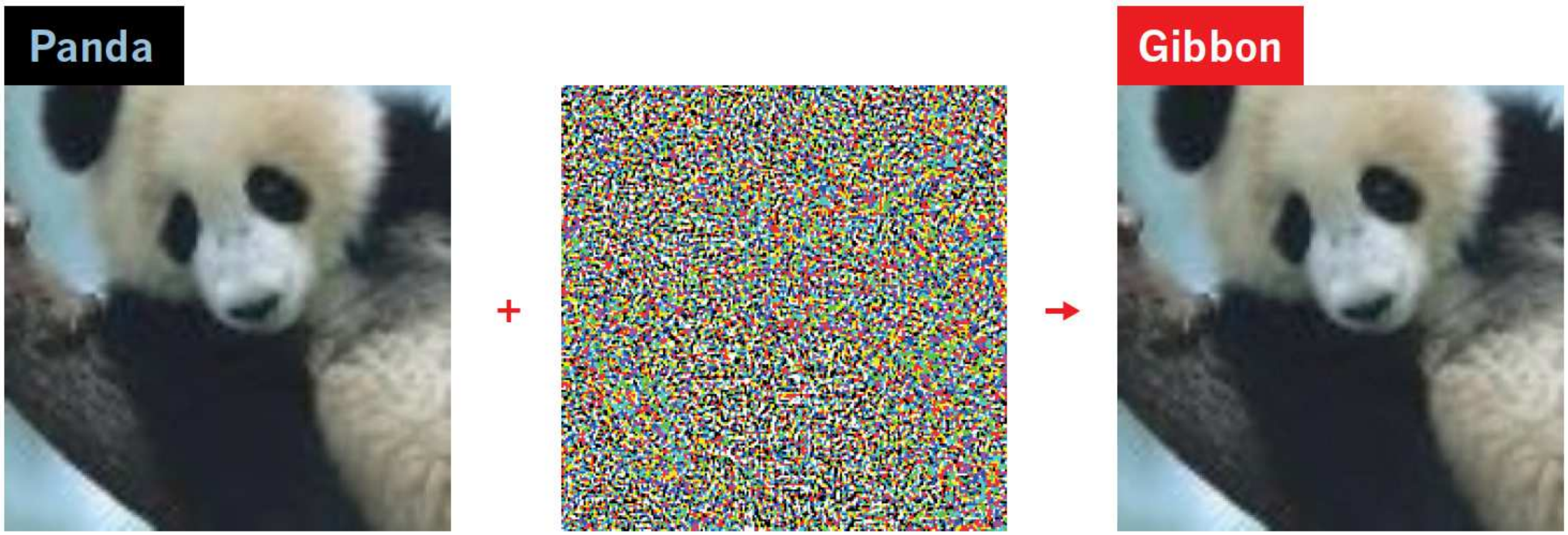
Falsche Perspektive



Gemeine Pflaster

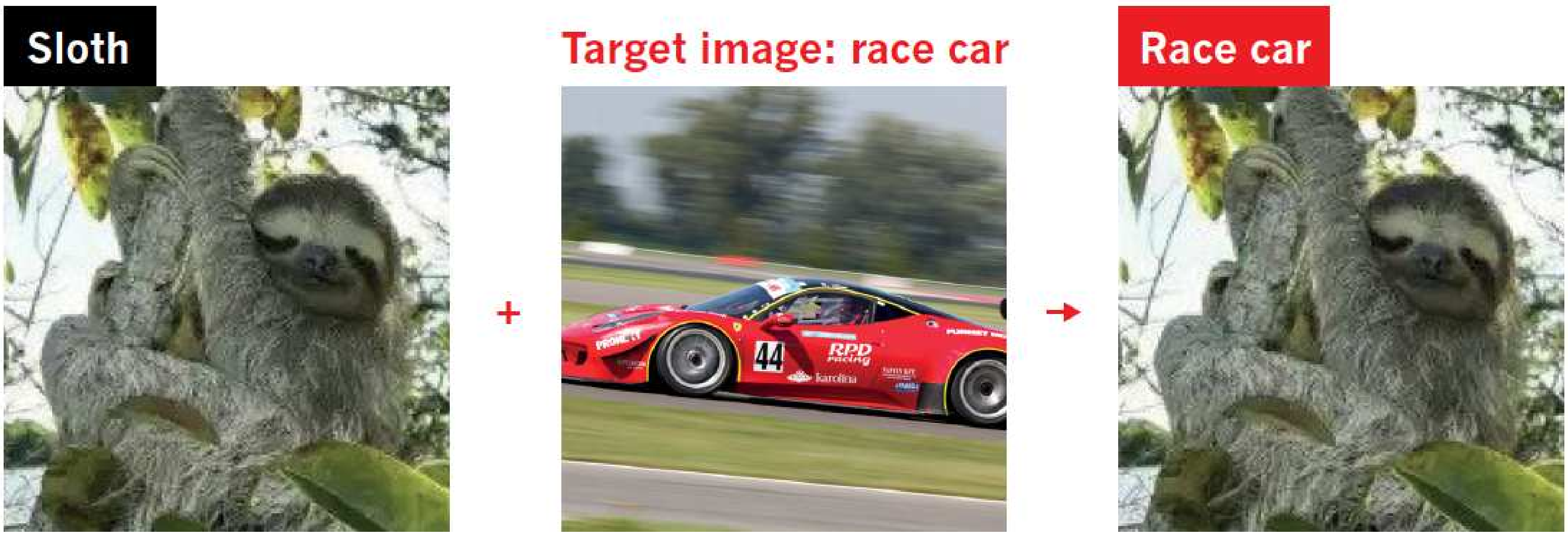


Gemeines Rauschen



Schlecht konditioniertes Problem? Instabiler Algorithmus?

Zielsicheres Rauschen Rauschen



Schlecht konditioniertes Problem? Instabiler Algorithmus?