

Woher kommt die Bezeichnung „Convolutional“?

Erarbeitet von
Dr. Ludmila Himmelspach

Lernziele	1
Inhalt	2
Einstieg	2
Woher kommt die Bezeichnung „Convolutional“?	2
Filterung von Bildern	3
Take-Home Message	7
Quellen	7
Weiterführendes Material	7
Disclaimer	7

Lernziele

- Du lernst, warum Convolutional Neural Networks so bezeichnet werden
- Du kannst erklären, was Zero Padding ist
- Du lernst, wie Filter auf Bilder angewendet werden

Inhalt

Einstieg

Hast Du Dich schon mal gefragt, warum künstliche neuronale Netzwerke für die Bildverarbeitung eigentlich Convolutional Neural Networks heißen? In diesem Video bekommst Du die Antwort auf diese Frage. Außerdem lernst Du, was Filter sind und wie diese auf Bilder angewendet werden.

Woher kommt die Bezeichnung „Convolutional“?

Erinnerst Du Dich noch daran, wie Bilder im Computer gespeichert werden? Ein digitales Bild besteht aus einzelnen, farbigen Pixeln. Diese werden sichtbar, wenn wir in die Bilder reinzoomen. Die Pixel innerhalb der Bilder sind rasterförmig angeordnet und enthalten Informationen über ihre Position und Farbe. Die Farbwerte werden durch ganze Zahlen repräsentiert und gespeichert, wobei 0 der schwarzen und 255 der weißen Farbe entsprechen. Einfachheit halber konzentrieren wir uns nur auf die Graustufenbilder.

Schauen wir uns ein einfaches 5 x 5-großes Beispielbild an. Es stellt ein weißes Kreuz auf schwarzem Hintergrund dar. Wenn wir die Farbwerte einzelner Pixel durch ganze Zahlen ersetzen, können wir unser Beispielsbild als eine 5 x 5-Matrix mit 25 Einträgen darstellen. Diese Darstellung hilft uns, zu verstehen, wie Bilder digital verarbeitet werden.

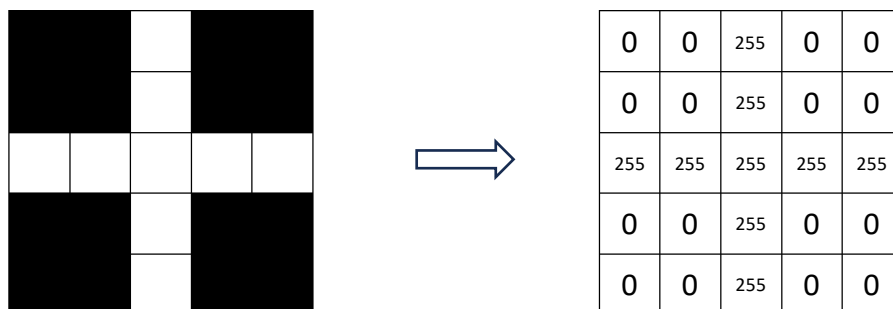


Abbildung 1: Beispielbild in Graustufen (links) und Ganzzahldarstellung (rechts).

Noch vor der Entwicklung der Convolutional Neuronal Networks wurden in der digitalen Bildverarbeitung Bilder mit Hilfe von unterschiedlichen *Filtern* bearbeitet und analysiert. Bei Anwendung eines Filters auf ein Bild werden alle Bildpunkte bzw. Pixel unter Einbezug ihrer Nachbapixel manipuliert. Die Anzahl der Nachbapixel wird durch die Größe der *Filtermaske* bestimmt. Bei einer Filtermaske handelt es sich um eine Matrix, deren Einträge, auch *Gewichte* genannt, die Art der Manipulation des Eingabebildes bestimmen.

Quelle [1]

Dir ist vielleicht der *Blur*- bzw. *Mittelwertfilter* bekannt, der einen verwischenden Effekt auf das Bild hat. Dieser Filter und seine Varianten werden oft benutzt, um zum Beispiel auf den Fotos Fältchen zu glätten oder den Hautton auszugleichen.

$$\frac{1}{9} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Die 3 x 3-Maske des Mittelwertfilters sieht wie oben angegeben aus.

Bei der Filterung eines Bildes wird die Filtermaske sukzessiv über alle Bildpunkte geschoben. Deren Gewichte werden mit den darunterliegenden Pixelwerten des ursprünglichen Bildes multipliziert und aufaddiert. Die resultierenden Werte ergeben die Farbwerte der Pixel im Ausgabebild.

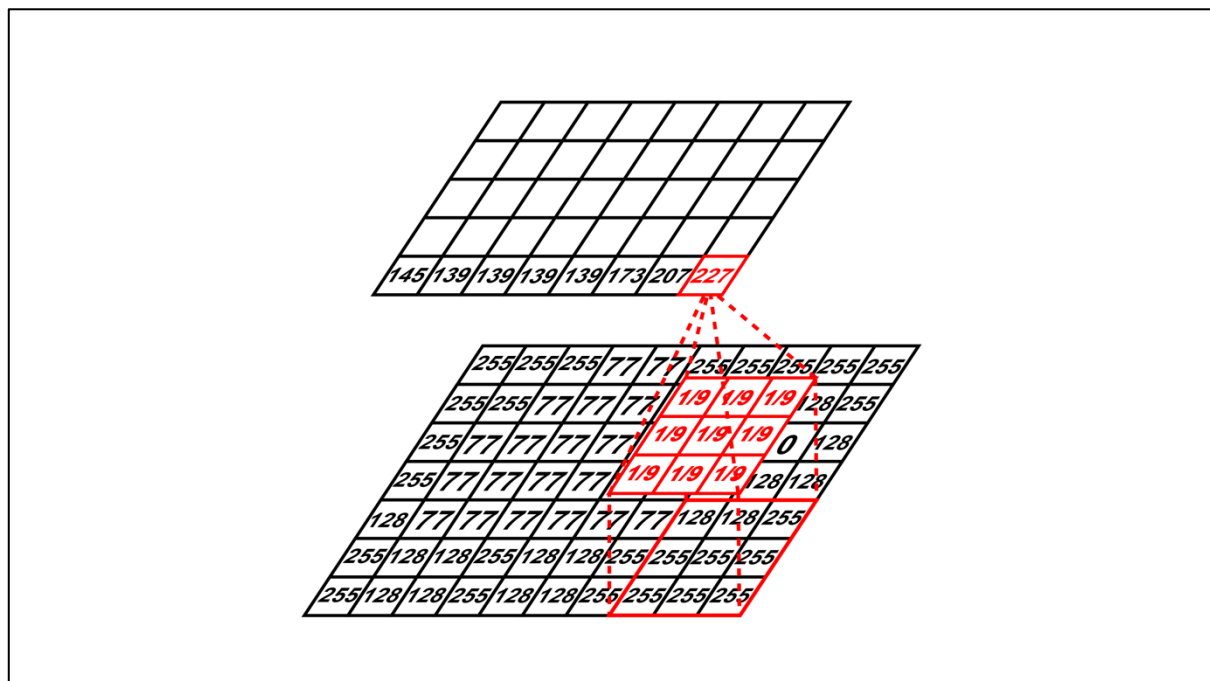


Abbildung 2: Filterungsvorgang in digitaler Bildverarbeitung.

Diesen Vorgang nennt man auch *Faltung* oder auf Englisch *Convolution*. Du kannst Dir wahrscheinlich schon denken, warum die neuronalen Netzwerke für die Bildverarbeitung *Convolutional Neural Networks* heißen. Richtig, weil in diesen Bildern gefiltert werden. Bevor wir uns anschauen, an welchen Stellen Filter in CNNs angewendet werden, sehen wir uns doch an, welche Filter es gibt und welche Veränderungen sie in Bildern bewirken.

Filterung von Bildern

Wenden wir doch zuerst den Mittelwertfilter auf unser Beispielbild an. Damit Du siehst, wie die Pixelwerte des resultierenden Bildes berechnet werden, wenden wir den Filter auf die Matrixdarstellung unseres Bildes an. Wie der Filtername schon sagt, berechnet sich der

Wert eines jeden Pixels im Ausgabebild aus dem Mittelwert des Pixels und acht seiner Nachbarn im Ursprungsbild.

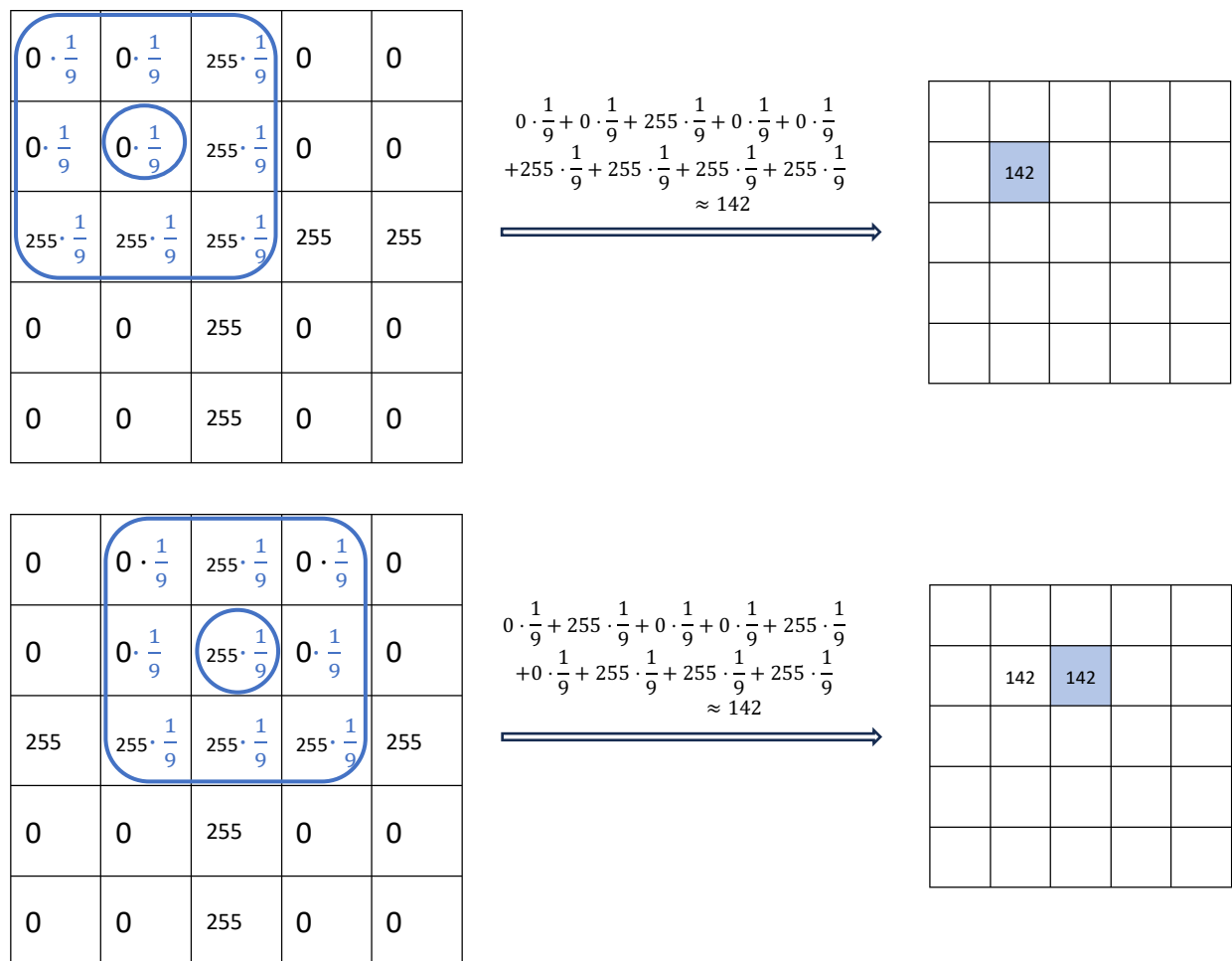


Abbildung 3: Berechnung der Pixelwerte des Ausgabebildes bei Anwendung des Mittelwertfilters.

Hier müssen wir uns allerdings um die Randpixel besonders kümmern, weil diese weniger als acht Nachbarpixel haben. Eine gängige Vorgehensweise ist Nullen um die Randpixel des Eingabebildes zu legen und mit diesen zu rechnen. Das wird übrigens *Zero Padding* genannt und findet in Convolutional Neural Networks auch ihre Anwendung.

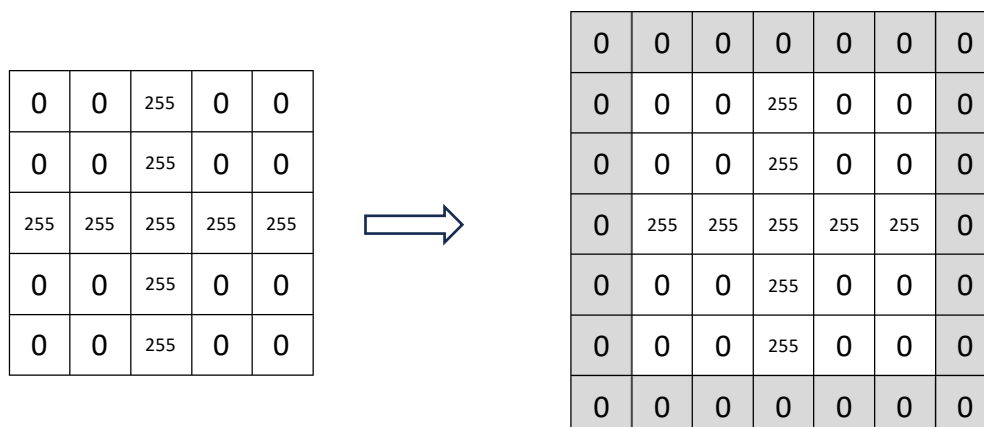


Abbildung 4: Zero Padding.

Wenn wir die Pixelwerte des resultierenden Bildes durch die entsprechenden Farben darstellen, sehen wir, dass der Mittelwertfilter alle Kanten im Bild verwischt hat, sodass man das Kreuz in der Mitte des Bildes kaum erkennen kann.

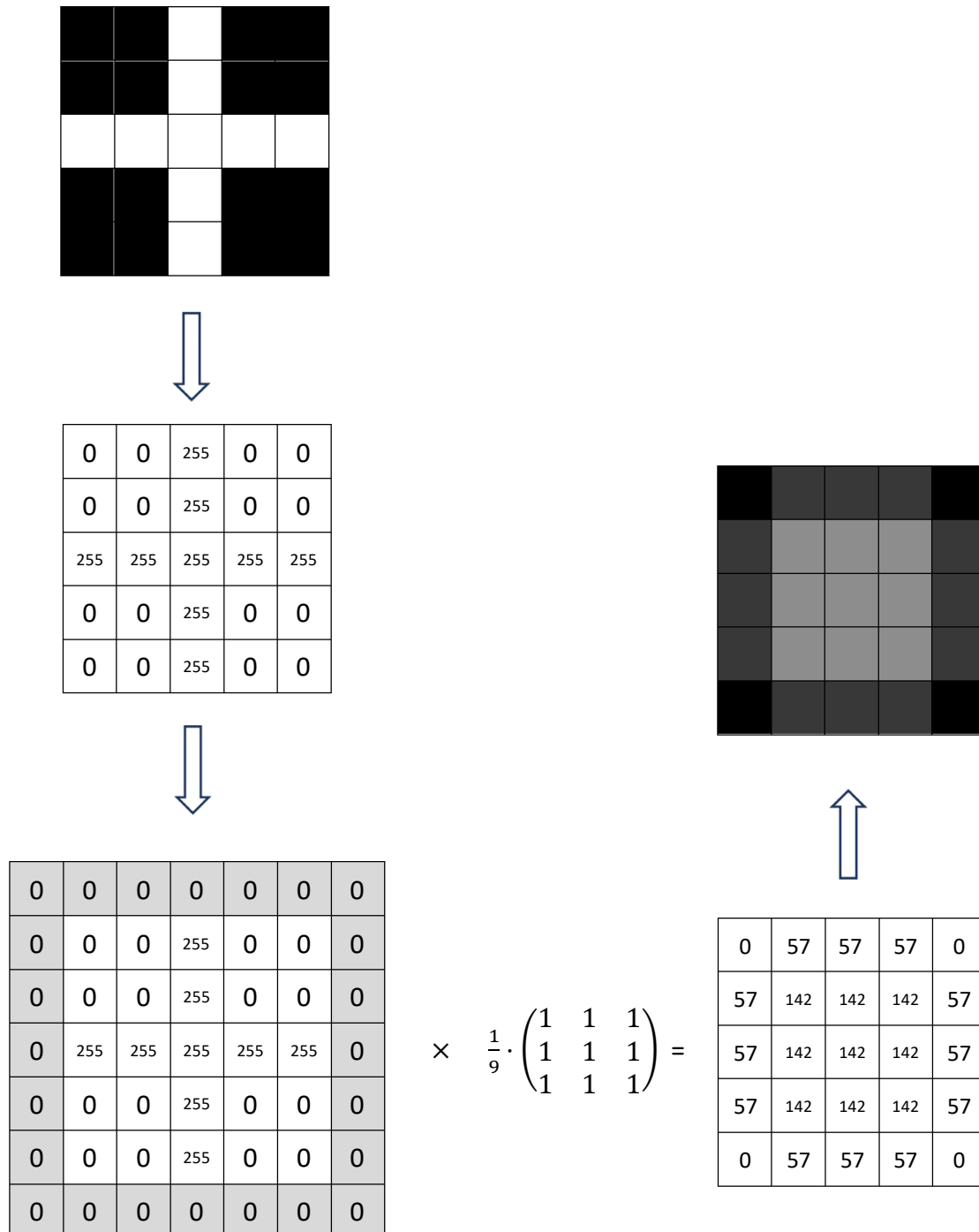


Abbildung 5: Anwendung des Mittelwertfilters auf das Beispielbild.

Lassen wir die Filtermaske, die in der mittleren Spalte Einsen und sonst Nullen enthält, über unser Beispielbild schieben, so wird im Ausgabebild der senkrechte Balken hervorgehoben, während die waagerechte Linie verwischt wird. Diesen Filter bezeichnet man auch als *vertikalen Kantenfilter*, weil er eben die vertikalen Kanten im Bild hervorhebt.

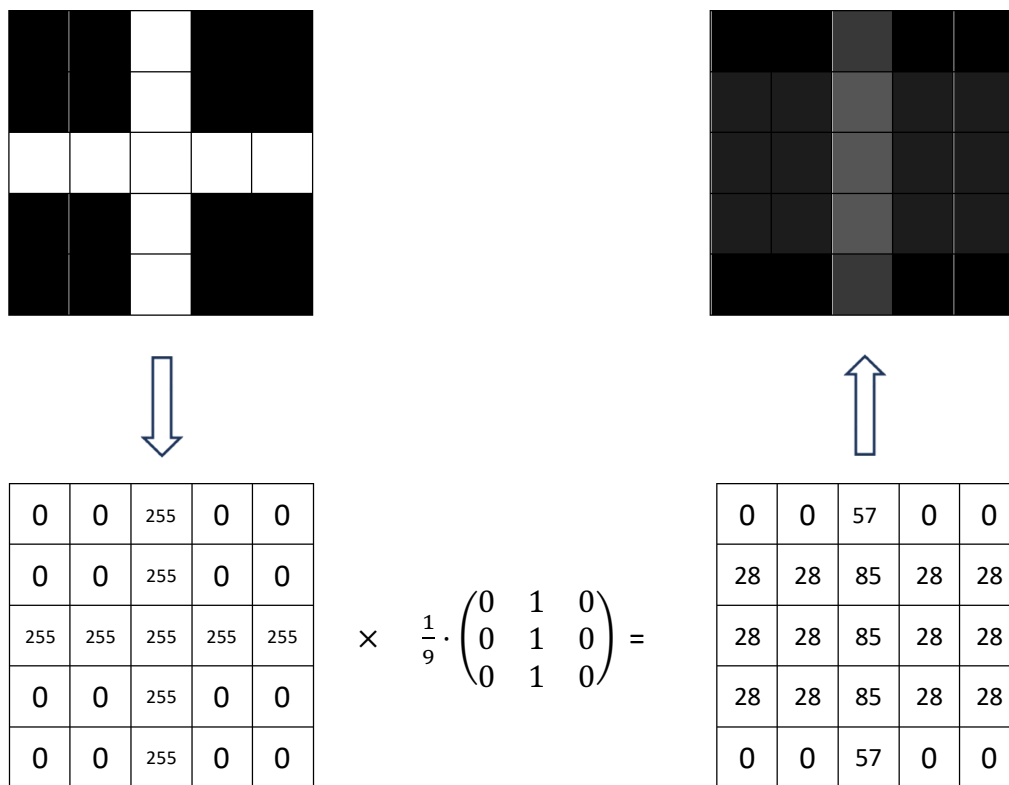


Abbildung 6: Anwendung des vertikalen Filters auf das Beispielbild.

Filtern wir unser Beispielbild mit der Filtermaske, die in der mittleren Zeile Einsen und sonst Nullen enthält, wird im Ausgabebild die waagerechte Linie hervorgehoben und der vertikale Balken verwischt. Deswegen bezeichnet man diesen Filter als *horizontalen Kantenfilter*.

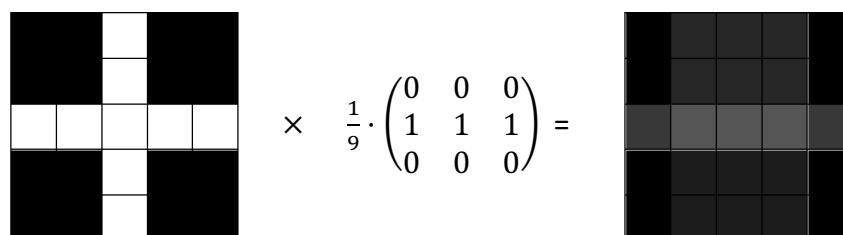


Abbildung 7: Anwendung des horizontalen Filters auf das Beispielbild.

Kannst Du Dich noch daran erinnern, was Hubel und Wiesel über die Neuronen im visuellen Cortex herausgefunden haben, die gleiche Wahrnehmungsfelder haben? Einige von ihnen nehmen nur Linien mit bestimmter Ausrichtung wahr.

Quelle [2]

Das, was uns angeboren zu sein scheint, wird in den Convolutional Neural Networks mit Hilfe von unterschiedlichen Filtern realisiert.

Take-Home Message

In diesem Video hast Du erfahren, wie Convolutional Neural Networks zu ihrem Namen kamen. Außerdem hast Du gelernt, wie Bilder mit Filtern verarbeitet werden.

Quellen

Quelle [1] Jähne, B. (2002). *Digitale Bildverarbeitung* (5., überarb. und erw. Aufl.). Springer.

Quelle [2] Bellebaum, C., Thoma, P., Daum, I. (2011). *Neuropsychologie*. Germany: VS Verlag für Sozialwissenschaften. (S 34)

Weiterführendes Material

Chollet, F. (2018). *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. MITP-Verlags GmbH & Co. KG.

Disclaimer

Transkript zu dem Video „Bildklassifikation und Bildsegmentierung: Woher kommt die Bezeichnung „Convolutional“?“, Dr. Ludmila Himmelsbach.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.