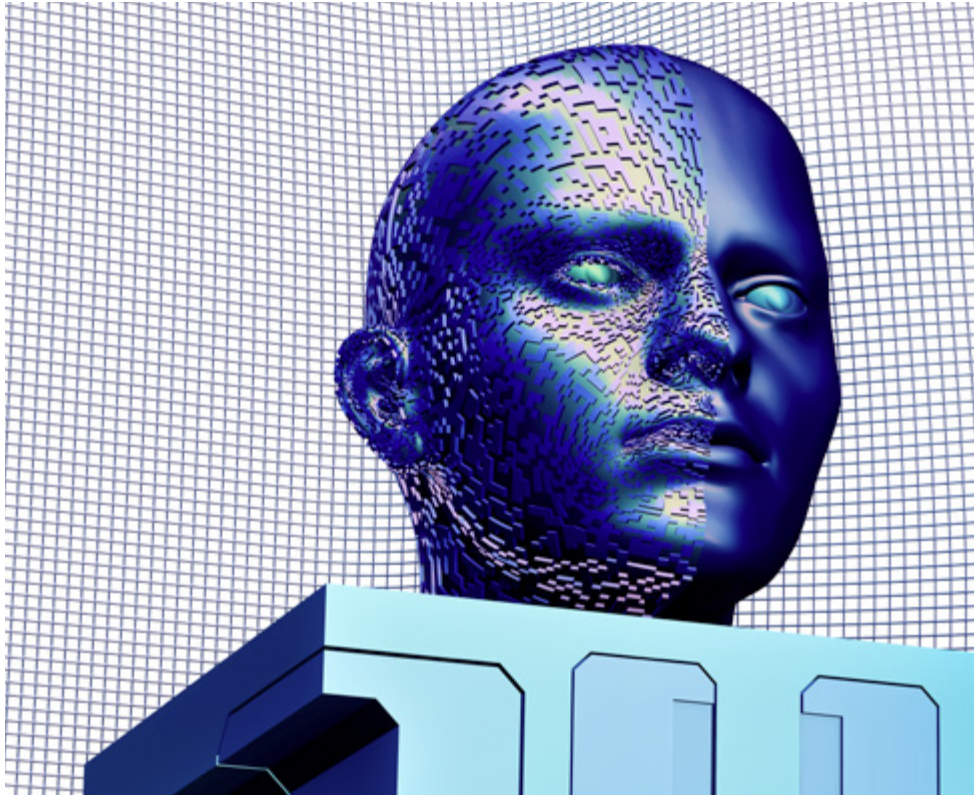




Mittelstand-Digital
**Zentrum
Hamburg**

LEITFADEN



ZUKUNFTSVISION ODER MEHR?

KI-GESTÜTZTE BILDVERARBEITUNG
FÜR DEN MITTELSTAND

Mittelstand-
Digital 

Gefördert durch:



Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

ZUKUNFTSVISION ODER MEHR?

KI-GESTÜTZTE BILDVERARBEITUNG FÜR DEN MITTELSTAND

Liebe Leserinnen und Leser,

künstliche Intelligenz ist längst nicht mehr nur bei großen Konzernen im Einsatz. Auch in kleinen und mittelständischen Unternehmen gibt es verschiedenste Anwendungsfelder für diese spannende Technologie.

Möchten Sie Inspektionsaufgaben oder andere optische Prozesse in Ihrem Unternehmen mithilfe von künstlicher Intelligenz automatisieren und so gleichbleibend hohe Qualität erzielen? Dieser Leitfaden verrät Ihnen, wie Sie dabei vorgehen können!

Erfahren Sie, wie Bildverarbeitung mit künstlicher Intelligenz funktioniert, welche Vor- und Nachteile es gibt und für welche Aufgaben KI grundsätzlich infrage kommt. Schritt für Schritt wird Ihnen gezeigt, welche Prozesse wann in einem Maschine-Learning-Projekt anstehen und was Sie beachten müssen. Außerdem erhalten Sie Hilfestellung bei der Entscheidung, wann Sie ein KI-Projekt inhouse machen oder auf die Unterstützung von Dienstleistern setzen sollten.

Für welchen Weg Sie sich auch entscheiden: Mit der richtigen Strategie wird Ihr Projekt zum Erfolg – und zu einem wirtschaftlichen und technologischen Wettbewerbsvorteil.

Prof. Dr.-Ing. Alexander Fay und Dr. Jakob Krooß

INHALTSVERZEICHNIS

EINS	Einleitung	04
ZWEI	Grundlagen: Maschinelles Lernen und neuronale Netze.....	05
	2.1 Klassische Bildbearbeitungsverfahren	07
	2.2 Lernende Bildverarbeitungsverfahren	08
DREI	Klassische vs. lernende Verfahren: Machine Learning in KMU	16
	3.1 Mögliche ML-Anwendungen für Ihr Unternehmen	19
	3.2 Verwendbare Technologien	20
	3.3 Herausforderungen bei der Anwendung in der Praxis	23
VIER	Ablauf eines ML-Projektes	25
FÜNF	Hardware und Software für ML-Projekte.....	31
SECHS	Eigenentwicklung oder Dienstleister?.....	35
SIEBEN	Fazit	40
ACHT	Literatur und weitere Informationen	44
NEUN	Mittelstand-Digital Zentrum Hamburg	46
ZEHN	Impressum	47

EINS

EINLEITUNG

In den letzten Jahren haben Methoden der künstlichen Intelligenz (KI) Aufgaben gelöst, die vorher unlösbar schienen. Vor allem die neuesten Entwicklungen in der Bildverarbeitung beeindrucken: Mit Kameras bestückte Autos fahren ohne Fahrer und krankhafte Veränderungen in medizinischen Ganzkörper-Scans werden vollautomatisch detektiert und klassifiziert. Bisher werden diese Technologien meist in großen Unternehmen eingesetzt. Doch auch für kleine und mittelständische Unternehmen (KMU) bietet bildverarbeitende KI enormes Potenzial. Verschiedenste Oberflächendefekte, schadhafte Bauteile, fehlende Komponenten oder auch andere Qualitätsmerkmale können so schon mit vergleichsweise geringem Einführungsaufwand zuverlässig und in Echtzeit erkannt werden. Während internationale Großkonzerne die finanziellen Mittel für individuelle und umfassende Komplettlösungen von Experten aufwenden können, ist der Einstieg für viele KMU jedoch eher schwierig.

Wo in Ihrem Unternehmen ist der Einsatz von KI in der Bildverarbeitung vielversprechend? Welche Lösung ist die richtige für Ihren Anwendungsfall und warum? Wie stellen Sie sicher, dass Ihre Lösung funktioniert? Wer kann Sie unterstützen? Und was ist sonst noch zu bedenken? Antworten auf Ihre Fragen erhalten Sie in diesem Leitfaden.

Der Leitfaden konzentriert sich auf die Verwendung von KI in der Bildverarbeitung und ist mit vielen Beispielen illustriert. Im Mittelpunkt stehen Verfahren des maschinellen Lernens (ML), insbesondere die sogenannten Convolutional Neural Networks. [Kapitel 2](#) liefert einen Schnelleinstieg in die begrifflichen Grundlagen und die Funktionsweise von Neuronen und Netzen. Keine Angst – es geht nicht um ein exaktes mathematisches Verständnis der Details. Vielmehr betrachten Sie das Ganze qualitativ, sodass Sie Vor- und Nachteile sowie Chancen und Einschränkungen nachvollziehen können. In [Kapitel 3](#) erhalten Sie einen Überblick, welche Aufgaben maschinelles Lernen übernehmen kann und auf welche Herausforderungen Sie dabei stoßen können. [Kapitel 4](#) zeigt die verschiedenen Projektphasen zur Einführung von ML in Unternehmen. Ein besonderer Fokus liegt dabei auf der Ressourcen- und Personalplanung, für Ihr KMU sicher ein wichtiger Aspekt. Schließlich erfahren Sie in [Kapitel 5 und 6](#) Wissenswertes rund um Hardware, Software und Dienstleister.

Mit diesem Know-how sollten Sie eigene ML-Projekte erfolgreich angehen und umsetzen können.

ZWEI

GRUNDLAGEN: MASCHINELLES LERNEN UND NEURONALE NETZE

Machen Sie sich zuerst mit ein paar Begriffen vertraut, bevor Sie mehr über das Prinzip von neuronalen Netzen bei der Bildgebung erfahren.

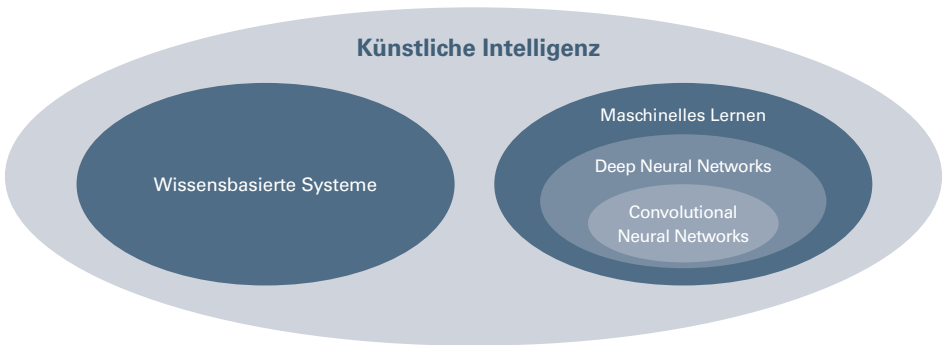


Abbildung 1:
Übersicht verschiedener Technologien
künstlicher Intelligenz

Der Begriff **künstliche Intelligenz (KI)** ist vielschichtig. Grundsätzlich lassen sich KI-Systeme in zwei Gruppen einteilen: Wissensbasierte Systeme sowie das maschinelle Lernen. **Wissensbasierte Expertensysteme** treffen anhand definierter Regeln, Zusammenhänge und Logiken nachvollziehbare Entscheidungen. Beispielhaft sei hier eine KI genannt, die Vorgänge im Rechtswesen automatisiert bearbeitet, etwa Anträge auf Entschädigung bei Verspätungen von Flug- oder Bahnreisen. Diese Expertensysteme sind in der Bildverarbeitung

KÜNSTLICHE INTELLIGENZ (KI)

„Künstliche Intelligenz ist der Überbegriff für Anwendungen, bei denen Maschinen menschenähnliche Intelligenzleistungen erbringen. Darunter fallen das maschinelle Lernen oder Machine Learning, das Verarbeiten natürlicher Sprache [...] und Deep Learning. Die Grundidee besteht darin, durch Maschinen eine Annäherung an wichtige Funktionen des menschlichen Gehirns zu schaffen – Lernen – Urteilen und Problemlösen.“ [1]

MASCHINELLES LERNEN (ML)

Das maschinelle Lernen fasst alle Methoden zusammen, die aus Beispieldaten Zusammenhänge ableiten, die anschließend auf weitere Daten übertragen werden können. Typischster Vertreter in der Bildverarbeitung sind die Convolutional Neural Networks (CNNs).

eher zu vernachlässigen. Das **maschinelle Lernen** (ML) hingegen setzt nicht auf vorgegebene Regeln, sondern definiert diese Regeln und Entscheidungsgrundlagen eigenständig anhand von Beispieldaten, für die eine gesuchte Aussage (z. B. „Das Bauteil ist (nicht) an der richtigen Position“) bereits bekannt ist. Anhand der gelernten Regeln kann für weitere Daten ebenfalls eine entsprechende Aussage getroffen werden.

In der Bildverarbeitung wird maschinelles Lernen immer häufiger eingesetzt. Die Gründe: Klassische, nicht auf KI basierende Bildverarbeitung ist aufwendig, erfordert spezialisierte Fachkräfte mit jahrelanger Erfahrung und scheitert noch immer an Problemstellungen, die für den menschlichen Blick keine Herausforderung darstellen. Für viele Anwendungen gibt es eine große Zahl an Daten in Form von Bildern – beste Voraussetzung also für das Anlernen der KI. Die jüngeren ML-Methoden zeigen Erfolge, die vor einigen Jahren undenkbar schienen und die klassische Bildverarbeitung an vielen Stellen überflügeln.

KLASSIFIKATION

Einteilung von Bildern anhand von Merkmalen in bestimmte Gruppen, z. B. unversehrtes vs. defektes Bauteil.

Eine der häufigsten Aufgaben in der Bildverarbeitung ist die **Klassifikation**, also die Einteilung von Bildern in Klassen. Ein einfaches Beispiel ist die Unterscheidung in „Bauteil in Ordnung“ bzw. „Bauteil defekt“. Es kann aber auch sehr viele Klassen geben („Bauteil hat Kratzer“, „Bauteil hat Bruchstelle“ etc.), im Extremfall sind das mehrere Tausend Klassen. Für diese Klassifikationsaufgaben sind sogenannte **neuronale Netze** hervorragend geeignet, also miteinander verknüpfte, künstliche Neuronen. Fast immer werden dabei sogenannte **Deep Neural Networks** (DNN) verwendet, bei Bildern insbesondere **Convolutional Neural Networks** (CNN).

2.1 Klassische Bildbearbeitungsverfahren

Klassische, also nicht auf Lernen basierende **Bildverarbeitungsalgorithmen** können je nach Anwendungsfall sehr unterschiedlich aussehen. Allerdings lässt sich ein Algorithmus für die Klassifikation fast immer in zwei Schritte aufteilen: die Merkmalsextraktion sowie die eigentliche Klassifikation anhand der **Merkmale**. Bei der Merkmalsextraktion werden bestimmte, mathematisch greifbare Eigenschaften des Bildes bzw. eines Objekts im Bild erfasst. Solche Merkmale können etwa die Größe, Farbe, Oberflächenbeschaffenheit, die Ausrichtung oder die Bewegungsgeschwindigkeit sein. Sie werden für gewöhnlich über mathematische Methoden aus den **Pixelwerten** gewonnen. Bei der eigentlichen Klassifikation wird dann anhand der Merkmale entschieden, zu welcher Klasse das Bild bzw. ein Objekt im Bild gehört. Dies kann zum Beispiel durch die Definition von Grenzwerten („Es handelt sich um ein Feuer, wenn sich mehr als 40 Pixel mit rötlicher Farbe im Bild befinden.“) geschehen.

Die gesuchten Merkmale sowie die Grenzwerte oder Regeln werden durch Fachkräfte mit Bildverarbeitungsexpertise festgelegt und als Algorithmus umgesetzt. Hierbei stehen zumeist auch Menschen mit speziellem Know-how für das Klassifikationsproblem zur Seite (z. B. ärztliches Fachpersonal bei der medizinischen Bildverarbeitung), insbesondere wenn die Unterschiede zwischen den Klassen für Fachfremde nicht direkt ersichtlich sind.

ALGORITHMUS

Feste Abfolge von vom Computer durchgeführten Arbeitsschritten. Die Abfolge wird mithilfe eines Quellcodes in einer Programmiersprache definiert. Bei gewünschten Anpassungen des Algorithmus muss zwingend auch der Quellcode angepasst werden. Je nach Komplexität sind hierfür i. d. R. ausgeprägte Programmierkenntnisse und Zeitaufwände erforderlich.

MERKMAL

Mathematisch beschreibbare Eigenschaft eines Bildes oder Objekts im Bild.

PIXEL

Von engl. Picture Element. Ein Bild setzt sich aus vielen einzelnen Bildpunkten, den sogenannten Pixeln, zusammen. Die Information der Pixel wird mit Zahlenwerten beschrieben. Hohe Zahlen entsprechen hellen, niedrige Zahlen dunklen Pixeln.

KÜNSTLICHES NEURON

Mathematische Abbildung einer menschlichen Nervenzelle, die allein sehr einfache, und im Verbund mit weiteren Neuronen komplexe Aufgaben lösen kann.

EINGANGSGEWICHTE

Eigenschaften eines Neurons, die bestimmen, wie sich das Neuron bei eingehenden Signalen verhält.

2.2 Lernende Bildverarbeitungsverfahren

DAS KÜNSTLICHE NEURON

Das **künstliche Neuron** bildet die Funktionsweise einer echten Nervenzelle, etwa im menschlichen Gehirn, mathematisch und somit von Computern berechenbar ab. Dabei ist der Mechanismus eines einzelnen Neurons alles andere als beeindruckend: Mehrere Eingangssignale (beim echten Neuron elektronische Impulse, beim künstlichen Neuron einfache Zahlenwerte), beispielsweise von anderen Nervenzellen, werden addiert. Falls in Summe ein Schwellwert überschritten wird, so wird ein Signal (ebenfalls elektrisch bzw. als Zahlenwert) ausgegeben und zum Beispiel an andere Neuronen weitergeleitet. Dabei haben die Eingangssignale unterschiedlich starken Einfluss darauf, ob eine Ausgabe gemacht wird; sie sind unterschiedlich stark gewichtet. So hat ein Eingangssignal, das fast immer aktiv ist, einen viel geringeren Einfluss als Signale, die nur in bestimmten Situationen eingehen. Diese **Eingangsgewichte**, die mit jedem Eingangssignal multipliziert werden, reichen aus, um das Verhalten eines Neurons vollständig zu beschreiben.

Entsprechend übersichtlich ist auch die Mathematik hinter dem künstlichen Neuron. Für den gesamten Prozess sind lediglich Additionen und Multiplikationen erforderlich. *Abbildung 2* zeigt die Elemente eines künstlichen Neurons, das seinem biologischen Gegenstück ähnelt.

Bereits ein einzelnes Neuron kann einfachste Klassifikationsaufgaben übernehmen. Der Vorteil dabei: Die Gewichte, also das Verhalten und somit die Klassifikationsgrundlage des Neurons, müssen nicht vorweg durch Fachleute festgelegt werden. Viel-

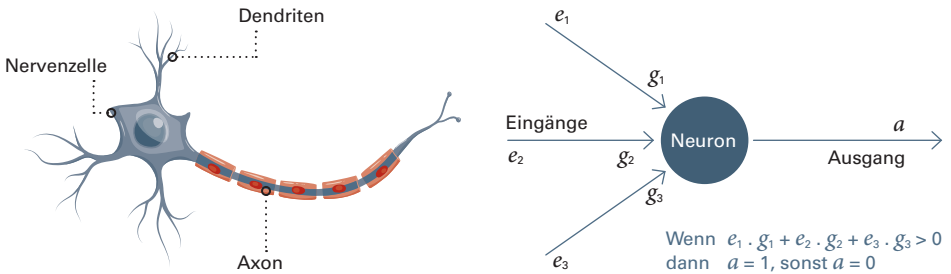


Abbildung 2: Vergleich zwischen Nervenzelle und künstlichem Neuron

mehr genügt es, geeignete Merkmale auszuwählen und eine ausreichende Menge an Beispieldaten mit bekannter Klasse bereitzustellen. Diese Daten werden auch als **Lerndaten** und die bekannte Zuordnung zu einer Klasse als **Label** bezeichnet.

Das künstliche Neuron kann dann anhand mathematischer Regeln eigenständig lernen, welche Gewichte zu einer optimalen Unterscheidung der Klassen führen: Zunächst werden die Startgewichte zufällig gewählt und alle verfügbaren Beispieldaten nacheinander mit den aktuellen Gewichten klassifiziert. Stimmt das Ergebnis mit dem Label überein, ist keine Änderung erforderlich. Wenn nicht, ist die Klassifikation noch fehlerhaft und die Gewichte werden leicht angepasst. Wurden alle Lerndaten mehrfach durchlaufen, ist davon auszugehen, dass die optimalen Gewichte gefunden wurden.

Beispiel 1 auf der folgenden Seite zeigt Ihnen anhand einer einfachen Illustration die Funktionsweise eines künstlichen Neurons. Hier lassen sich die beiden Klassen problemlos trennen. Ein Blick auf das Bild auf Seite 12 in Beispiel 2 verdeutlicht hingegen, dass bei komplizierteren Klassifikationsaufgaben mehr als ein Neuron erforderlich ist.

LERNDATEN

Daten (hier: Bilder), deren Klasse(n) bereits bekannt sind. Ein Neuron oder neuronales Netz kann anhand dieser Daten Klassifikationsregeln erlernen.

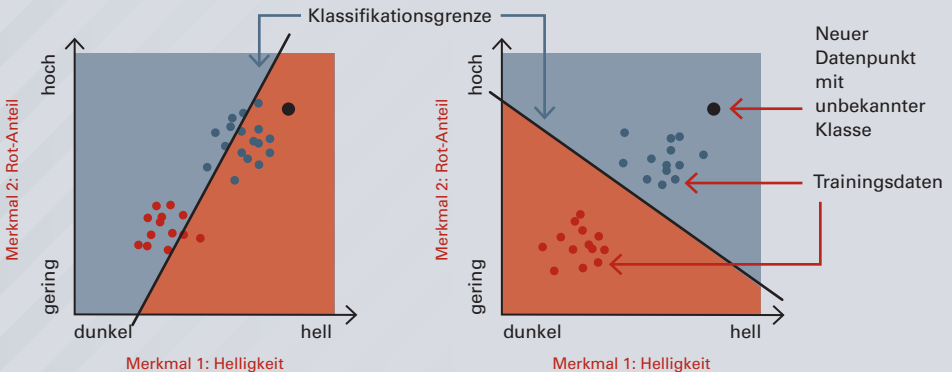
LABEL

Die manuelle Zuordnung, zu welche(r) Klasse(n) ein Bild gehört, wird auch als Label bezeichnet. Sie ist klar zu trennen von der nach dem Training durchgeführten, automatischen Klassifikation durch das Netz.

FUNKTIONSWEISE EINES KÜNSTLICHEN NEURONS

FEUER ERKENNEN (BEISPIEL 1)

Um zu erkennen, ob ein gegebenes Bild Feuer enthält, werden die einzelnen Bildpunkte betrachtet und – vereinfacht – in zwei Klassen eingeteilt. Wählt man einige dieser Pixel aus Testbildern aus und trägt ihre Helligkeit sowie den Rot-Anteil gegeneinander auf, so ist festzustellen, dass sich sowohl die Pixel ohne Feuer (orangene Punkte, eher dunkel, eher geringer Rot-Anteil) als auch die Pixel mit Feuer (blaue Punkte, eher hell, eher hoher Rot-Anteil) an einer bestimmten Stelle häufen.



Verlauf der Klassifikationsgrenze vor dem Lernen

Verlauf der Klassifikationsgrenze nach dem Lernen

Außerdem fällt auf, dass sich beide Klassen durch eine Gerade trennen lassen (rechtes Bild). Möchte man nun für ein weiteres (bzw. sehr viele) Pixel bestimmen, ob es zu einem Feuer gehört, gilt lediglich zu kontrollieren, auf welcher Seite der Gerade sich das Pixel einordnet (so liegt der schwarze Punkt oberhalb der Klassifikationsgrenze und gehört somit offensichtlich zu einem Feuer). Nichts anderes macht das Neuron: Die Lage der Klassifikationsgrenzen ist mit den sogenannten Gewichten im Neuron gespeichert. Die Auswertung (auf welcher Seite der Grenze befindet sich ein Eingabedatenpunkt?) kann mit einfachen Berechnungen sehr schnell durchgeführt werden.

Im Beispiel lässt sich die Gerade auch manuell leicht bestimmen. **In den allermeisten Fällen werden allerdings deutlich mehr als nur zwei Merkmale benötigt, um zwei Klassen eindeutig voneinander zu trennen.** Das Problem lässt sich dann grafisch nicht mehr darstellen und eine manuelle Trennung der Klassen ist nicht mehr möglich.

Die große Stärke des künstlichen Neurons liegt darin, dass es auch in diesen Fällen eine Trennung vornehmen kann, indem es die Position der Klassifikationsgrenze anhand von Beispieldaten mit bekannter Klasse erlernt. Zunächst wird die Grenze zufällig platziert (linkes Bild); anschließend werden nacheinander alle verfügbaren Beispieldatenpunkte mit der verfügbaren Klassifikationsgrenze ausgewertet. Stimmt das Ergebnis der Klassifikation mit der manuell bestimmten Klasse überein, ändert sich nichts; ist dies nicht der Fall, wird die Grenze ein Stück weit verschoben bzw. gedreht, sodass sie die Klassen besser trennt. Nachdem alle Punkte mehrfach auf diese Weise durchlaufen werden, liegt die Grenze so, dass beide Klassen korrekt getrennt wurden (rechtes Bild). Wertet man nun einen weiteren Datenpunkt/Pixel mit unbekannter Klasse aus (schwarzer Punkt), so wird durch das Neuron automatisch die richtige Klasse bestimmt.



DEEP NEURAL NETWORKS

Die hohe Leistungsfähigkeit des menschlichen Gehirns kann erst durch die Verknüpfung mehrerer Neuronen erreicht werden. Dies lässt sich auch auf künstliche Neuronen übertragen: So schließen sich beliebige Verschaltungen mehrerer Neuronen zu einem **neuronalen Netzwerk** (NN, engl. Neural Network) zusammen. Für die Klassifikation bietet sich ein Untertyp der NN an, das **Deep Neural Network** (DNN). In diesem sind die Neuronen nicht beliebig positioniert, sondern in mehrere Schichten eingeteilt. Die verfügbaren Merkmale bilden die Eingänge der ersten Schicht, die Ausgänge einer jeden Schicht die Eingänge der jeweils nächsten Schicht. Die Neuronen der letzten Schicht geben das Klassifikationsergebnis aus. Mit steigender Anzahl an Schichten können auch komplexere Aufgaben gelöst werden; das Netz wird tiefer, daher der Name **Deep** Neural Network. Die meisten Schichten sind voll vernetzt – jedes Neuron einer Schicht ist mit jedem Neuron der nächsten Schicht verknüpft.

NEURONALES NETZWERK (NN)

Beliebige Verknüpfung mehrerer künstlicher Neuronen.

DEEP NEURAL NETWORK (DNN)

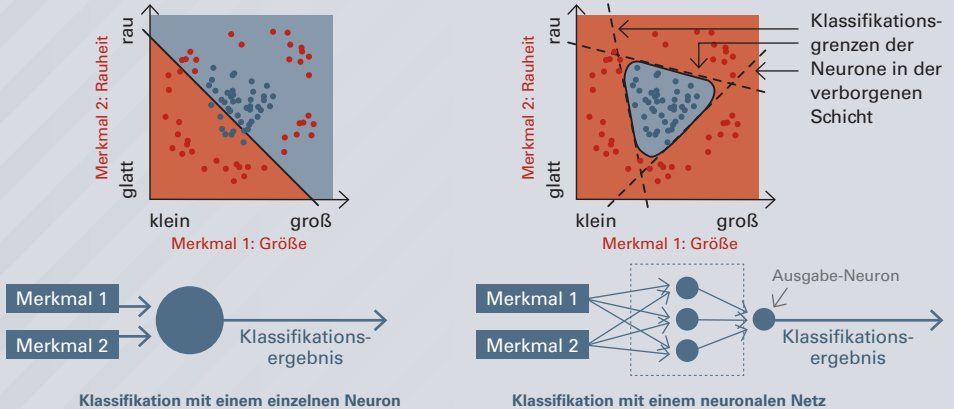
Untertyp des neuronalen Netzwerks, in dem die Neuronen in Schichten aufgebaut sind.

Hinweis:
Das Beispiel dient der Anschaulichkeit und hat keinen Bezug zur realen Geologie.

FUNKTIONSWEISE EINES KÜNSTLICHEN NEURONALEN NETZWERKS

EDELSTEINE BESTIMMEN (BEISPIEL 2)

In diesem Beispiel sollen in einer Edelsteinmine Steine für die manuelle Untersuchung automatisiert vorsortiert werden. Jeder gewonnene Stein wird fotografiert und die beiden Merkmale „Größe“ und „Rauheit“ mit klassischen Bildverarbeitungsmethoden bestimmt.



Trägt man die beiden Merkmale der Beispieldaten mit bekannter Klasse gegeneinander auf, so zeigt sich, dass insbesondere Steine mittlerer Größe und Rauheit die gesuchten Edelsteine enthalten (blaue Punkte im Zentrum). Alle weiteren Steine (dargestellt als gelbe Punkte) gehören zur Klasse der Nichtedelsteine. Offensichtlich lassen sich die beiden Klassen nicht durch eine einzige Klassifikationsgrenze trennen (linkes Bild). Verknüpft man jedoch mehrere Neuronen miteinander, so können auch unregelmäßig „geformte“ Klassen beschrieben werden (rechtes Bild). In diesem Beispiel muss der untersuchte Datenpunkt jeweils auf einer bestimmten Seite jeder Klassifikationsgrenze der drei Neuronen in der verdeckten Schicht liegen, damit das Ausgabe-Neuron den Datenpunkt als potenziellen Edelstein klassifiziert.

Die abgerundeten Ecken ergeben sich aus der Aktivierungsfunktion, ein mathematisches Element, das für die korrekte Funktion des neuronalen Netzes unerlässlich, hier jedoch nicht von weiterer Bedeutung ist.

Mit steigender Anzahl an Neuronen können auch komplizierte Klassifikationen vorgenommen werden. Insbesondere kann ein Netz mehrere verborgene Schichten enthalten; die Ausgänge der ersten Schicht bilden dann die Eingänge der zweiten Schicht und so weiter. Weiterhin können mehr als ein Ausgabe-Neuron verwendet und so zeitgleich mehrere Klassen betrachtet werden.

Der Lernprozess ist etwas komplizierter als beim einzelnen Neuron. Die Korrektur der Klassifikationsgrenzen muss nicht nur für die Ausgabe-Neuronen, sondern auch für alle vorhergehenden Schichten durchgeführt werden. Der „Fehler“, also der Unterschied zwischen der Klassifikation mit dem aktuellen, noch fehlerhaften Netz und der manuellen, also korrekten Klassifikation wird ausgehend vom letzten Neuron bis zur ersten verborgenen Schicht „durchgereicht“ (engl. Backpropagation). Dieser Prozess ist im Vergleich zur Klassifikation selbst sehr aufwendig und kann einen wesentlichen Zeitaufwand bedeuten.

CONVOLUTIONAL NEURAL NETWORKS

Deep Neural Networks konnten in der Vergangenheit für viele Anwendungsgebiete sehr gute Ergebnisse erzielen. Auf Bilder waren sie aber zunächst nicht anwendbar, da die Netze für die extrem vielen Pixel innerhalb eines Bildes einfach zu groß gewesen wären. Erst mit der Erfindung der Convolutional Neural Networks (CNN) wurde dieses Problem gelöst. Hier werden dem eigentlichen Netz mehrere sogenannte Faltungsschichten vorgeschaltet. Als **Faltung** (engl. convolution) bezeichnet man eine mathematische Operation, mit der einfache Strukturelemente wie Ecken, Kanten oder schlichte Muster in einem Bild gefunden werden können. Der Ausgang einer jeden Faltungsschicht bildet den Eingang der nächsten Faltungsschicht, wodurch in jeder Schicht komplexere Informationen gewonnen werden können. So werden beispielsweise in den ersten Schichten gerade Linien im Bild erkannt. In den folgenden Schichten könnten diese Linien als Fell eines Tieres identifiziert werden.

Der Clou ist, dass die genaue Art der Faltung und somit der gesuchten Merkmale nicht wie in der klassischen Bildverarbeitung fest vorgegeben ist, sondern ebenfalls anhand der Trainingsdaten gelernt wird!

Das Netz entscheidet also selbst anhand der Daten, ob nach Fell oder doch nach Federn gesucht werden muss. Da in jeder Schicht mehrere Faltungen unabhängig voneinander durchgeführt werden, lassen sich mehrere Merkmale parallel ermitteln. Diese fließen dann in die im vorherigen Abschnitt beschriebenen Schichten aus Neuronen ein. Nun entscheidet sich, ob die zuvor ermittelten Merkmale

FALTUNG

(engl. convolution)

Einfache, sich aus Multiplikation und Addition benachbarter Pixel zusammensetzende mathematische Operation, mit der bestimmte Merkmale, etwa einfache Strukturelemente wie Ecken, Kanten oder schlichte Muster in Bildern gefunden werden können.

„Fell“, „Augen“, „Schlappohren“ und „langes Gebiss“ zu einem Hund oder einer Katze gehören. CNN weisen also eine ähnliche Struktur auf wie die Klassifikation mit klassischer Bildverarbeitung: Zunächst werden Merkmale bestimmt, anhand derer anschließend die Zuteilung zu Klassen durchgeführt wird.



Wichtig für die faltenden Schichten: Zwischen zwei Faltungen wird das Zwischenergebnis, das dieselbe Größe hat wie das Eingangsbild, jeweils verkleinert („Pooling“). Da der Abstraktionsgrad mit jeder Faltung steigt, geht dabei kaum Information verloren. Die jeweils nächste Schicht sowie insbesondere die Neuronen im letzten Schritt können jedoch deutlich kleiner entworfen werden, wodurch sich die Berechnungsgeschwindigkeit stark erhöht. Ein weiterer Vorteil der Faltungen ist die vorgegebene Struktur des Netzes: Im Gegensatz zu den DNN sind nur benachbarte Pixel im Bild miteinander verknüpft. Auch hierdurch kann das Netz weiter verkleinert werden. [Abbildung 3](#) zeigt die Struktur eines typischen CNN.

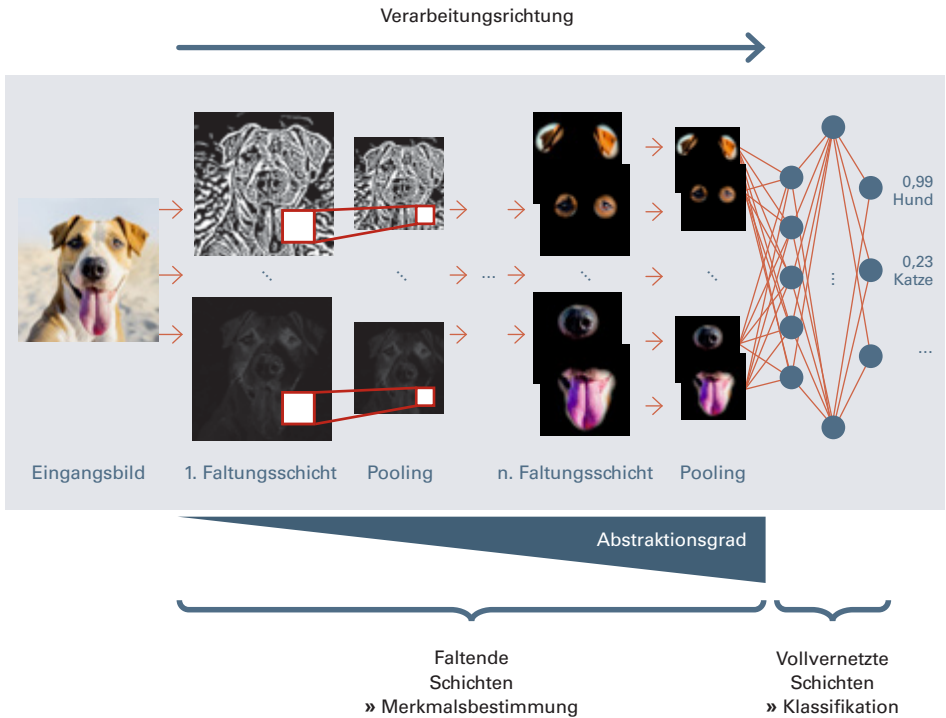


Abbildung 3: Struktur eines typischen Convolutional Neural Network (CNN)

Durch die Faltungen und die dazwischenliegenden Verkleinerungen entsteht also ein Netz, das bei der Bildverarbeitung um Größenordnungen kleiner, schneller und erfolgreicher ist als ein entsprechendes DNN.

Convolutional Neural Networks haben die Bildverarbeitung revolutioniert. Bereits bei ihrer Einführung konnten sie viele Aufgaben besser lösen als über Jahre entwickelte klassische Ansätze. Inzwischen wird die Leistungsfähigkeit von CNN ständig erhöht, unter anderem auch durch speziell zugeschnittene Hardware. Aber Achtung: Methoden des maschinellen Lernens und damit auch die CNN sind keine Wundermittel, die alle bisherigen Ansätze ablösen. Einige Aufgaben können nach wie vor nur oder zumindest besser durch klassische Methoden gelöst werden.

DREI

KLASSISCHE VS. LERNENDE VERFAHREN: MACHINE LEARNING IN KMU

Wie können Sie entscheiden, ob sich für eine Aufgabenstellung in Ihrem Unternehmen ein neuronales Netz, insbesondere ein Convolutional Neural Network, anbietet? Halten Sie sich den wesentlichen Unterschied zwischen der klassischen und der auf maschinellem Lernen basierenden Bildverarbeitung noch einmal vor Augen: In der klassischen Bildverarbeitung werden die Mechanismen, nach denen eine Entscheidung getroffen wird, durch menschliche Expertise manuell hergeleitet und entwickelt. Bei lernenden Verfahren hingegen geben nicht Sie diese Mechanismen vor, sondern diese werden automatisiert aus Beispieldaten abgeleitet, also gelernt.

Die folgenden Überlegungen helfen Ihnen für den konkreten Anwendungsfall:

- Voraussetzung für lernende Verfahren ist, dass ausreichend große Datensätze mit Bildern vorhanden sind. Jedes Bild innerhalb dieser Datensätze benötigt ein Label, zu welcher Klasse das Bild jeweils gehört. Da für ein gut funktionierendes Netz viele Tausend Trainingsbilder nötig sind, kann der Aufwand von Beschaffung und Labeln dieser Daten schnell höher sein als die Entwicklung eines klassischen Algorithmus. Womöglich stehen Ihnen ohnehin nicht genügend Daten zur Verfügung. Sofern diese Daten aber bereits verfügbar sind, sollten Sie eine mögliche ML-Umsetzung erwägen.

ML FÜR BRANDERKENNUNG IN UNTERNEHMEN

Die Erkennung von Bränden mit ML-Methoden ist ein großes Forschungsfeld und kommt auch in Betrieben häufig zum Einsatz. Um Feuer bzw. Flammen oder Rauch per ML zu bestimmen, können zahlreiche große Datenbanken mit frei zugänglichen Bildern genutzt werden. Bei Verpuffungen, also schlagartigen, ansatzweise explosionsartigen Verbrennungen, stehen hingegen kaum Daten zur Verfügung – nicht zuletzt, da es ein extrem hoher Aufwand ist, eine Explosion herbeizuführen, der auch mit Gefahren verbunden ist. Die wenigen verfügbaren Aufnahmen von Verpuffungen sind zudem meist nicht mit Hochgeschwindigkeitskameras entstanden und somit für die rechtzeitige Erkennung im Alltag unbrauchbar.

Bei einem klassischen Bildverarbeitungsalgorithmus ist nachvollziehbar, warum für ein Bild ein bestimmtes Resultat erzielt wird. Tauchen trotzdem unerwartete Ergebnisse auf, kann die Auswertung schrittweise rekonstruiert und der Algorithmus bei Bedarf nachjustiert werden, so dass entsprechende Unregelmäßigkeiten nicht erneut auftreten. Bei einem lernenden Algorithmus hingegen ist für den Menschen nicht offensichtlich, warum ein Netz zu einem bestimmten Ergebnis kommt. Beispielsweise ist es bei der bildlichen Unterscheidung zwischen Hunden und Katzen zwar möglich, dass das Netz die Struktur von Fell sowie die Form der Augen als wesentliche Merkmale erlernt hat. Es kann aber auch sein, dass das Netz andere, komplett abstrakte Merkmale betrachtet, die für den Menschen überhaupt nicht interpretierbar sind. Dies kann sowohl Vor- als auch Nachteil sein. Häufig finden neuronale Netze Merkmale, die bei der Klassifikation äußerst hilfreich sind, die der menschlichen Expertise aber zuvor entgangen sind, da sie entweder nicht naheliegend oder für Menschen nicht interpretierbar sind. Aber dieses Vorgehen birgt auch immer eine Gefahr. So können aufgrund der Zusammensetzung der Testdaten Merkmale betrachtet werden, die eigentlich überhaupt keine Rolle spielen.

Daraus ergibt sich eine weitere Herausforderung für lernende Systeme: Eine Zertifizierung für sicherheitskritische Aufgaben (z. B. autonomes Fahren) ist deutlich erschwert. So ist es im Gegensatz zu klassischen Bildverarbeitungssystemen nicht möglich, einzelne Schritte des Algorithmus zu prüfen, sondern nur den gesamten Algorithmus. Ein weiteres Risiko besteht

DAS PANZER-BEISPIEL

Im Internet kursiert folgende Anekdote: Ein neuronales Netz wurde trainiert, um Panzer in Bildern zu erkennen. Es wurden hervorragende Ergebnisse erzielt bei Versuchen mit den im Training verwendeten Daten. Ein Einsatz im Feld zeigte jedoch, dass das Netz völlig „zufällige“ Ergebnisse lieferte. Erst nach langen Untersuchungen kam man darauf, dass die Trainingsdaten mit Panzern bei Sonnenschein, die Bilder ohne Panzer bei dichter Wolkendecke aufgenommen wurden – das Netz suchte nicht nach Panzern, sondern nach Wolken und war dementsprechend unbrauchbar. Auch wenn es sich hier vermutlich um eine Urban Legend handelt [2] und sich solch extremen Auswirkungen durch eine durchdachte Auswahl der Trainingsdaten verhindern lassen, zeigt es anschaulich die Risiken von lernenden Verfahren.

dadurch, dass selbst minimale Änderungen am Bild eine Anpassung der Ausgabe, etwa des Klassifikationsergebnisses, zur Folge haben können. Ein gezieltes Ausnutzen dieses Umstands („single pixel attack“) ist für viele Anwendungen ein ernstzunehmendes Risiko.

- Natürlich können auch manuell erzeugte Modelle fehlerhaft sein. Werden solche Fehler bemerkt oder entsteht aus anderen Gründen ein Änderungsbedarf (etwa bei veränderten Anforderungen), so ist dies bei klassischen Algorithmen meist mit bedeutendem Aufwand verbunden. Zumindest in einigen Fällen können solche Änderungen bei lernenden Verfahren erheblich einfacher umgesetzt werden, indem das Training mit einigen zusätzlichen Trainingsbildern erneut durchgeführt wird.
- Bei der Klassifikation mit vielen verschiedenen Klassen muss im traditionellen Ansatz für jede Klasse ein eigener Algorithmus entwickelt werden. Ein neuronales Netz hingegen kann alle Klassen gleichzeitig auswerten.

Bei der **Bestimmung von qualitativen Informationen**, insbesondere komplexen Klassifikationsaufgaben sind lernende Verfahren erfahrungsgemäß den klassischen Verfahren überlegen. Dies gilt umso mehr, wenn Abweichungen zwischen Klassen (z. B. die genaue Ausprägung eines Oberflächendefekts) existieren und/oder die Umgebungsbedingungen (z. B. Beleuchtung oder Hintergrund) variabel sind.

Bei der **Bestimmung von quantitativen Informationen** wie Größen, Winkeln oder Geschwindigkeiten haben lernende Verfahren hingegen Schwächen. Hier sind meist klassische Methoden das Mittel der Wahl. Ein Beispiel: Das Ablesen der Ausrichtung eines Messzeigers auf einem runden, analogen Druckmessgerät erfolgt mit traditionellen Methoden zuverlässig und problemlos. Gleiches gilt für das Auslesen von Bar- oder Aztec-Codes.

In einigen Fällen bietet es sich an, einen hybriden Ansatz wählen, also bestimmte Teile des Algorithmus klassisch aufbauen (etwa bei der Vorbereitung der Daten oder der Bestimmung einiger Merkmale) und die lernenden Verfahren nur dort einsetzen, wo diese ihre Stärken haben.

3.1 Mögliche ML-Anwendungen für Ihr Unternehmen

Aus Unternehmenssicht lassen sich mögliche Anwendungsfälle von künstlicher Intelligenz und maschinellem Lernen grob in zwei Gruppen teilen:

- Interne Fälle: KI/ML verbessert einen Prozess (z. B. Qualitätskontrolle) bzw. macht ihn wirtschaftlicher
- Externe Fälle: KI/ML wird Teil eines vom Unternehmen verkauften Produkts bzw. Services

Die Überlegung, KI und ML ins eigene Produktportfolio einzubinden, scheint im Rahmen von sich ändernden Kundenbedürfnisse und dadurch neu öffnenden Märkten attraktiv. Aber seien Sie vorsichtig, gerade wenn Sie sich neu in das Gebiet KI wagen! Es können große Investitionen nötig sein sowie eine wesentliche Umgestaltung der Strukturen und Prozesse in Ihrem Unternehmen. Je nach Anwendung messen Sie sich zudem auf dem Markt mit Unternehmen, die diesen Wandel schon vollzogen und ihre Produkte oder Services bereits etabliert haben. Im Rahmen der Bildverarbeitung gehören dazu primär der Verkauf von Hardware, Software, Tools, Bibliotheken sowie Cloud-, Beratungs- und weiteren Services, die auf KI/ML basieren. Diese können sowohl darauf abzielen, die Kunden bei der Erstellung eigener Netze zu unterstützen, als auch darauf, der Kundschaft ein fertiges Endprodukt zu liefern, beispielsweise eine Software für die Erkennung von krebsverdächtigen Zellen in mikroskopischen Aufnahmen. Aber auch der Verkauf von durch KI gewonnenen Informationen (z. B. aus Satellitenaufnahmen) kann sich als profitables Geschäftsmodell erweisen.

Einen leichteren Einstieg in die Technologie KI/ML bieten zumeist Anwendungen, die in Ihrem Unternehmen selbst zu finden sind. Bezogen auf die Bildverarbeitung eignen sich insbesondere Aufgaben in der Qualitätskontrolle. Hier können Sie Prozesse automatisieren, die zuvor manuell durchgeführt wurden. Dazu gehört das Prüfen auf Schäden, Fehler sowie teilweise auch auf falsche Positionierung. Der Aufwand richtet sich nach der Komplexität der Aufgabe, mit vergleichsweise kleinem Aufwand lassen sich häufig schon gute Ergebnisse erzielen. Weitere Anwendungsgebiete sind Assistenzsysteme wie Apps für die optische Erkennung und genaue Benennung von vielen Tausend verschiedenen Bauteilen (z. B. in der Instandhaltung) oder Systeme, die die Logistik oder die Arbeit von Produktionsrobotern unterstützen (z. B. Pick-and-Place-Anwendungen).

3.2 Verwendbare Technologien

In ihrer Ursprungsform eignen sich Convolutional Neural Networks für die **Klassifikation**, also die Einordnung ganzer Bilder in bestimmte Klassen. Das Bild wird als Ganzes ausgewertet, allerdings ohne Ortsinformationen. So ist beispielsweise möglich, ein Bauteil auf Risse zu prüfen. Wo genau sich diese befinden, bleibt jedoch unklar. Dabei spielt es

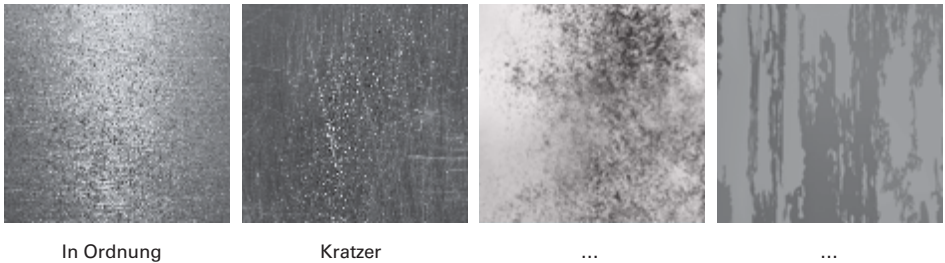


Abbildung 4: Beispiel für die Klassifikation

für ein gut trainiertes Netz keine Rolle, wie ein gesuchtes Objekt orientiert ist. Ist es aber nicht vollständig sichtbar oder ein nur sehr kleiner Teil des Gesamtbilds, kann dies die Performance des Netzes beeinflussen.

Die fehlende Ortsinformation bei der Klassifikation ist in einigen Anwendungen ein großer Nachteil. Glücklicherweise existieren Anpassungen für CNN, die auch eine ungefähre Lokalisierung der Objekte ermöglichen. Hierfür wird dem CNN (siehe [Abbildung 3, Seite 15](#)) ein weiterer, zumeist ebenfalls lernender, Mechanismus vorgeschaltet, der markante relevante Bereiche im Bild vorauswählt, die dann einzeln durch das CNN ausgewertet werden. Die rechteckige Form dieser Bereiche gibt entsprechend ausgewerteten Bildern ihr typisches Aussehen ([Abbildung 5](#)). Diese Herangehensweise wird als

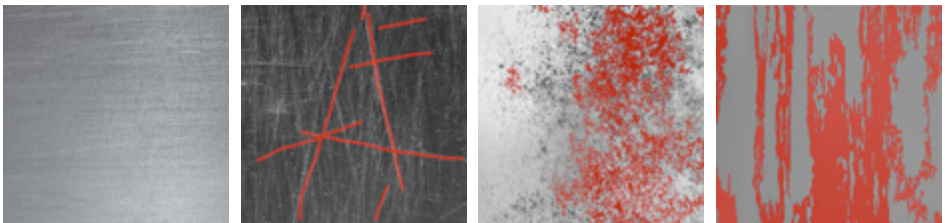


Abbildung 6: Beispiel für die semantische Segmentierung

Objekterkennung bezeichnet. Ein weiterer Vorteil dabei: Es können mehrere unterschiedliche Objekte in einem Bild gleichzeitig erkannt werden.

In einigen Anwendungen reicht eine ungefähre Ortsangabe durch die Rechtecke nicht mehr aus. Hier hilft die letzte der drei typischen Technologien in der lernenden Bildverarbeitung, die sogenannte (semantische) **Segmentierung**. In dieser wird nicht dem gesamten Bild oder einem Bereich, sondern jedem einzelnen Pixel eine Klasse zugeordnet. Typischerweise werden die verschiedenen Klassen dann mit einer farblichen Zuordnung im Bild markiert (Abbildung 6). Vereinfacht lässt sich das für die Segmentierung benötigte Netz als CNN ohne die vollvernetzte Schicht beschreiben. Die pixelbasierte Information bleibt hier erhalten und die Verkleinerung des Bildes zwischen den Faltungsschichten wird wieder rückgängig gemacht.

OBJEKTERKENNUNG
 Klassifikation mit vorhergehender Regionenvorhersage. Jede der rechteckigen Regionen wird einzeln klassifiziert.

SEGMENTIERUNG
 Pixelweise Zuordnung einer Klasse.

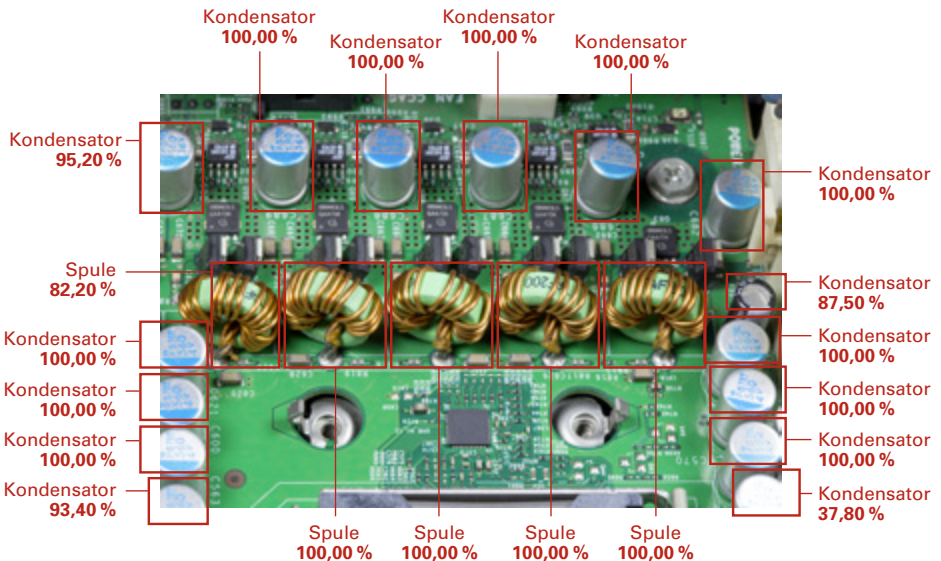


Abbildung 5: Beispiel für die Objekterkennung

RECURRENT NEURAL NETWORKS

Netze, die keine einzelnen Daten, sondern Datenreihen auswerten können. Informationen aus vorherigen Datenpunkten (z. B. Bildern) werden zwischengespeichert und können so in die Auswertung des aktuellen Bildes einfließen.

NOVELTY DETECTION

Neuronales Netz, das Auffälligkeiten wie Schäden nicht explizit klassifiziert, sondern lediglich „bemerkt“, dass Daten von den üblichen Messwerten abweichen.

Zunächst scheint insbesondere die Segmentierung dem klassischen CNN aufgrund der zusätzlichen Informationen weit überlegen. Allerdings müssen diese auch für die Beispieldaten bereitgestellt werden. Im Gegensatz zur Klassifikation müssen Klassen also nicht nur für jedes Bild, sondern für jedes Pixel gelabelt werden. Aufgrund der Vielzahl benötigter Bilder kann dies in einen mehrwöchigen Prozess ausarten, für den je nach Anwendungsfall Fachkräfte erforderlich sind. Dies sollten Sie bei der Auswahl der geeigneten Technologie berücksichtigen.

Mit der Klassifikation, der Objekterkennung und der Segmentierung sind die drei großen lernenden Technologien für die Bildverarbeitung abgedeckt. Es gibt zahlreiche weitere Entwicklungen für Spezialfälle und Anpassungen für hybride Ansätze. Hervorzuheben sind die sogenannten **rekurrenten Netze**, die auch Bildfolgen, also Videos, auswerten können. So können auch Veränderungen im Bild wahrgenommen werden, wie das Flackern von Flammen. Diese Erweiterung kann auf bei der Klassifikation ebenso wie bei der Objekterkennung und der Segmentierung genutzt werden. Außerdem interessant ist auch die **Novelty Detection**, bei der nur mit Bildern einer Klasse (z. B. „in Ordnung“) trainiert wird. Das Netz ist anschließend in der Lage, Abweichungen in anderen Bildern (z. B. Schäden) zu erkennen, allerdings können diese nicht klassifiziert werden („hier ist etwas anders, möglicherweise nicht in Ordnung“ statt „Kratzer“). Die Novelty Detection können Sie vor allem dann nutzen, wenn nur wenige Trainingsbilder mit Auffälligkeiten wie Schäden vorliegen und somit ein Training der Schadensklassen nicht möglich ist.

3.3 Herausforderungen bei der Anwendung in der Praxis

Warum ist es nun so schwierig, diese gut funktionierenden Technologien in einem Unternehmen, insbesondere in einem KMU zu integrieren? Neben der Datenlage und Aufbereitung gibt es weitere wichtige Punkte. So ist zwar die grundlegende Struktur eines jeden CNN identisch, im Detail unterscheiden sie sich jedoch deutlich voneinander. Die Anzahl an Schichten, die Größe und die pro Schicht durchgeführten Faltungen, die Anzahl der Neuronen in der vollvernetzten Schicht, die Anzahl an Trainingsdurchläufen und viele weitere mathematische Parameter beeinflussen die Performance eines Netzes stark. Natürlich gibt es gewisse Daumenregeln und erprobte Netze, die in anderen Anwendungen gut abgeschnitten haben. Die ideale Lösung kann jedoch je nach Anwendungsfall, auch abhängig von den verwendeten Lerndaten, anders aussehen. Der Leitsatz „Viel hilft viel“ gilt nicht beim Design von neuronalen Netzen! Durch immer mehr Neuronen und/oder Schichten verschlechtert sich die Performance eines Netzes ab einem gewissen Punkt deutlich: Das Netz ist zwar groß genug, die Lerndaten „auswendig zu lernen“. Im Feld zeigen sich aber schlechtere Ergebnisse, da sich das Netz kein „abstraktes Denken“ mehr zulegen musste (Abbildung 7).

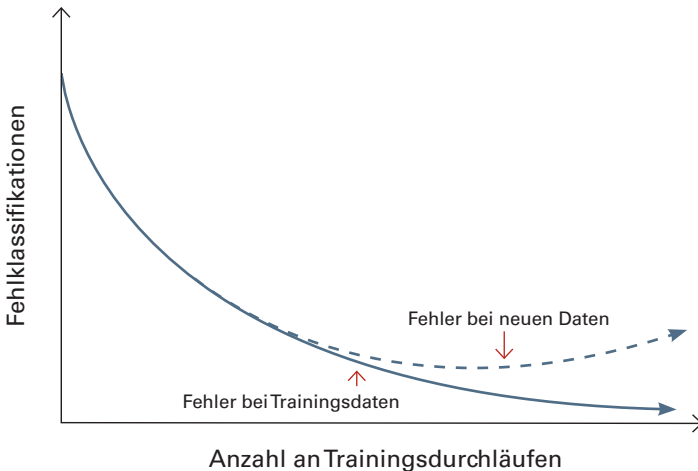


Abbildung 7: Die Anzahl der Fehlklassifikationen von Testdaten sinkt mit zunehmender Anzahl an Trainingsdurchläufen. Bei neuen Daten steigt die Anzahl jedoch ab einem gewissen Punkt wieder. Ein vergleichbarer Effekt ist auch für weitere Parameter wie die Größe des Netzes zu beobachten.

Die Erhöhung der Komplexität eines neuronalen Netzes führt nicht unbedingt auch zu einer Verbesserung der Ergebnisse. Vielmehr muss die Komplexität an die Aufgabenstellung angepasst werden.

ML-Methoden sind keine Plug-and-Play-Technologie, die Sie von der Stange kaufen können. Für optimale Ergebnisse müssen sich Fachleute in Ihrem Unternehmen zwingend damit auseinandersetzen und Anpassungen der KI an die den Anwendungsfälle vornehmen.

Inwiefern diese Optimierung getrieben werden muss, ist auch vom Anwendungsfall abhängig. Bei einer automatisierten Vorsortierung für ein effizienteres Recycling ist der ein oder andere zusätzliche Fehler durch fehlende Anpassungen des verwendeten Netzes vermutlich verzeihbar. Hier können Sie bezüglich des Aufwands auch Kompromisse eingehen und etwas schlechtere Ergebnisse in Kauf nehmen. Für entsprechende Anwendungen bietet sich möglicherweise sogar der Einsatz einer Smart Camera an (ein sehr wirtschaftlicher und unkomplizierter Ansatz, [siehe Kapitel 5, Seite 31](#)). Bei medizinischen Anwendungen wäre das Überspringen der Anpassungsphase hingegen fatal. Selbst wenn durch die Anpassung des Netzes die Fehlerquote nur um den Bruchteil eines Prozentes reduziert werden kann, ist dies zwingend erforderlich.

VIER

ABLAUF EINES ML-Projektes

Nun wird es konkret! In diesem Kapitel erfahren Sie, wie Sie Machine-Learning-Anwendungen und -Methoden step by step in Ihrem Unternehmen einführen. [Abbildung 8](#) zeigt die einzelnen Schritte. Auch wenn der Fokus auf Anwendungen mit Bildverarbeitung liegt, lassen sich die meisten Aussagen auf ML-Anwendungen allgemein beziehen.



Abbildung 8:
Ablauf eines ML-Projekts.
Modifiziert aus [4].

1. IDENTIFIKATION DES ANWENDUNGSFALLS UND DEFINITION VON ZIELEN UND KENNZAHLEN

Warum und wo wollen Sie ML in Ihrem Unternehmen integrieren? Die betriebliche Ausgangssituation lässt sich generell in drei Gruppen einteilen (modifiziert nach [3]):

Problemgetrieben

Wie kann ich mein operatives Problem mit Daten/KI lösen?

Bei der problemgetriebenen Ausgangslage besteht ein konkretes ungelöstes, operatives Problem, das durch eine neue Technologie gelöst werden soll.

Technologiegetrieben

Wie kann ich KI in meinem Unternehmen einsetzen?

Bei der technologiegetriebenen Ausgangslage ist zunächst keine spezifische Anwendung im Fokus. Vielmehr wird nach einem Anwendungsfall gesucht, um die Technologie im Unternehmen zu etablieren.

Datengetrieben

Wie kann ich vorliegende Daten wertschöpfend nutzen?

Bei der datengetriebenen Ausgangslage haben Sie den Wert von bereits vorliegenden Daten erkannt und möchte diese mit Methoden des ML gewinnbringend in Ihrem Unternehmen nutzen.

Gerade wenn Sie technologiegetrieben in das Feld KI einsteigen wollen, ist es sinnvoll, sich bei mehreren infrage kommenden Ideen zunächst niedrige Ziele zu stecken und auf schnelle Erfolge abzielen. So können Sie eigene Erfahrungen machen und Know-how sammeln, das Sie dann in komplexeren Projekten nutzen können. Starten Sie nicht mit einem langwierigen Aufbau von Strukturen und Prozessen – setzen Sie diese etwas später um, wenn Sie bereits erste Erfolge mit der Technologie erzielen konnten. Die Suche nach möglichen Anwendungsfällen sollte dabei in interdisziplinären Teams stattfinden. Wählen Sie für die ersten Projekte Anwendungsfälle, in denen keine personenbezogenen oder sensiblen Daten verwendet werden.

Unabhängig von der Ausgangslage sollten Sie einen potenziellen Anwendungsfall zunächst auf die Umsetzbarkeit mit klassischen Methoden, mit einem ML-Verfahren und auf die Verfügbarkeit geeigneter Daten prüfen.

Entscheidend für den Erfolg eines Projekts ist die klare Definition von Zielen. Dazu ein Beispiel: Leiterplatten sollen durch Unterstützung von lernender Bildverarbeitung effizienter recycelt werden. Dafür sind die Ziele detailliert zu spezifizieren: Sollen lediglich Vorder- und Rückseiten unterschieden werden? Der Bestückungsgrad? Sollen einzelne Komponenten klassifiziert und/oder der Ertrag pro Platine abgeschätzt werden? Sollen spezielle, finanziell besonders ertragsreiche Bauteile erkannt werden? All diese Ansätze sind sinnvoll und können eine Ertragssteigerung erzielen. Sie erfordern jedoch völlig andere Herangehensweisen und unterschiedliche neuronale Netze. Sich hier nicht von Vornherein festzulegen, wird im besten Fall zu erheblichem Mehraufwand, schlimmstenfalls zum Scheitern des Projekts führen.

In jedem Fall, insbesondere aber wenn Sie technologiegetriebenen agieren, sollten Sie den Return of Investment (RoI) abschätzen. Wenn Sie eine Anwendung ohne effektiven Mehrwert einführen, wird diese kaum langfristig im Unternehmen etabliert werden. Gerade für kleine und mittlere Unternehmen ist ein möglichst geringer Investitionsbetrag und ein frühzeitiger RoI aufgrund beschränkter Ressourcen gefordert. Empfehlenswert ist in diesem Fall die Umsetzung von Projekten, die sich innerhalb von zwei Jahren amortisieren [6]. Der RoI sollte mit bestehenden oder auch neuen Kennzahlen (Key Performance Indicators, KPI) quantitativ gemessen werden. Wollen Sie beispielsweise lediglich die (durch KPI gemessene) Erkennungsrate bei fehlerhafter Ware erhöhen oder soll der Anteil fehlerhafter Ware selbst durch das KI-System (durch KPI messbar) reduziert werden? Kleine Details machen hier einen großen Unterschied!

Um den Erfolg eines Projekts sicherzustellen, müssen klare Ziele definiert werden, deren Erfolg anhand von entsprechenden Kennzahlen quantitativ gemessen werden kann.

2. ERHEBUNG UND KATEGORISIERUNG DER TRAININGSDATEN

Wenn Sie einen Anwendungsfall für eine ML-Lösung identifiziert haben, müssen Sie im nächsten Schritt die für das Training des Netzes benötigten Bilder sammeln und kategorisieren, also labeln. Je nach Anwendung liegen diese Bilder bereits vor oder müssen zunächst erzeugt werden.

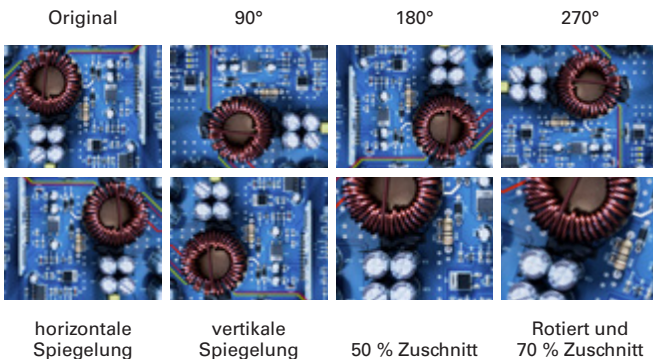


Abbildung 9:
Data Augmentation

DATA AUGMENTATION

Vergrößerung der Menge an verfügbaren Daten durch Drehen, Spiegeln, Vergrößern, Verzerren usw.

TRANSFERLERNEN

Verwendung eines vortrainierten Netzes.

Für komplexere Aufgaben sind Bilderzahlen im fünf- bis sechsstelligen Bereich erforderlich. Solche Zahlen zu erreichen, ist selbst für große Unternehmen herausfordernd. Glücklicherweise gibt es einige Tricks, um die benötigte Zahl an Bildern zu reduzieren. In fast allen Netzen wird die sogenannte **Data Augmentation** eingesetzt. Bei dieser wird jedes verfügbare Bild mehrfach gespiegelt, gedreht und verzerrt und in der Größe verändert. In Kombination ergeben sich so aus jedem Originalbild viele weitere Bilder, die ebenfalls für das Training eingesetzt werden können (siehe [Abbildung 9](#)). In einigen Fällen lassen sich Bilder auch mithilfe von Simulationen in großer Zahl künstlich erzeugen.

Eine weitere Möglichkeit, die benötigte Anzahl an Bildern zu reduzieren, ist das sogenannte **Transferlernen**. Über verschiedene, offen zugängliche Quellen sind zahlreiche vordefinierte Netze verfügbar, die bereits mit einer sehr großen Menge an Bildern verschiedenster Klassen vortrainiert wurden. Übertragen auf das Beispielnetz „Haustier bestimmen“, „kennen“ diese Netze zwar weder Hunde noch Katzen, aber bereits die Bildeigenschaften von Ecken, Kanten, Fell, Ohren etc. Diese Netze können nun mit einer vergleichsweise geringen Anzahl an Bildern (z. B. vierstellig) weitertrainiert werden, wobei die vorhandenen Merkmale an den Anwendungsfall angepasst und mit den richtigen Klassen verknüpft werden. Das resultierende Netz zeigt bei deutlich geringerer Anzahl an Trainingsbildern vergleichbare Ergebnisse wie ein nicht vortrainiertes Netz.

Das Potenzial des Transferlernens sollten Sie nicht unterschätzen! Auch größere oder auf ML spezialisierte Unternehmen setzen darauf.

Data Augmentation und Transferlernen sind mächtige Werkzeuge für die effiziente Nutzung von neuronalen Netzen.

3. AUSWAHL, IMPLEMENTIERUNG, TRAINING UND OPTIMIERUNG DES NETZES

Abhängig vom Anwendungsfall und den verfügbaren Daten wählen Sie als Nächstes eine geeignete ML-Architektur (z. B. eine der Unterformen des CNN). Dabei sollten Sie festlegen, inwiefern klassische Bildverarbeitung in die Auswertung mit einbezogen wird (hybrider Ansatz) und welche Softwarebibliotheken und Entwicklungsplattformen genutzt werden sollen.

Im nächsten Schritt implementieren Sie das gewählte Netz. Dabei sind auch die nicht gelernten Parameter wie die Anzahl an Schichten und der Neuronen pro Schicht vorläufig festzulegen.

In den meisten Fällen müssen Sie die zuvor gesammelten Daten noch für das Netz aufbereiten. So kann es z. B. sein, dass die Auflösung oder der Datentyp der Bilder noch nicht zum Netz passen.

Anschließend kann das Netzwerk trainiert werden. Typischerweise nutzen Sie dabei den größten Teil der verfügbaren Daten für das Training. Den verbleibenden Rest verwenden Sie zum Testen in zwei separaten Gruppen für die Parameteroptimierung sowie die spätere, abschließende Validierung. Die ideale Struktur Ihres Netzes finden Sie heraus, indem in einem iterativen Prozess verschiedene Kombinationen aus Anzahl der Schichten, Anzahl der Neuronen pro Schicht, Anzahl an Trainingsdurchläufen etc. getestet werden und die beste Kombination ausgewählt wird.

4. VALIDIERUNG, INTEGRATION UND INBETRIEBNAHME

Bei der abschließenden Validierung sollten Sie neben den Testdaten selbst auch erweiterte Stresstests durchführen, beispielsweise mit verrauschten oder verwackelten Bildern oder Gegenbeispielen. Sollten sich dabei noch Probleme offenbaren, müssen Sie die vorherigen Schritte teilweise wiederholen und gegebenenfalls weitere Lerndaten hinzufügen.

HIGH LEVEL VS. LOW LEVEL

Entwicklungsabteilungen arbeiten bei der Programmierung für gewöhnlich in sogenannten High-Level-Sprachen oder -umgebungen, die den Entwicklungsaufwand durch einen hohen Abstraktionsgrad reduzieren. Sie sind gegenüber hardwarenahen Low-Level-Sprachen jedoch recht langsam und teilweise schwerer zu integrieren.

Andernfalls können Sie das Netz in das Gesamtsystem, also z. B. die Produktionsanlage, integrieren. Hierfür muss das Netz zunächst aus der **High-Level-Programmierungsumgebung** in eine **Low-Level-Programmiersprache** überführt werden.

Nach der Inbetriebnahme Ihres neuronalen Netzes lässt sich anhand der gewählten Kennzahlen prüfen, ob die zuvor definierten Ziele erreicht wurden und Sie können gegebenenfalls nachsteuern.

5. ZYKLISCHE ÜBERPRÜFUNG

Auch nach der Inbetriebnahme ist das Projekt für Ihr Unternehmen nicht vollständig abgeschlossen. Vielmehr sollten Sie den Prozess weiterhin zyklisch untersuchen und im Bedarfsfall weitere Trainingsdaten hinzufügen, Kennzahlen ergänzen etc.

Achtung:

Dieser Leitfaden ersetzt keine Rechtsberatung. Während der gesamten Durchführung eines ML-Projekts sind gewisse rechtliche Aspekte zu beachten bezüglich Datenschutz, Haftung und Urheberrechten. Je nach Anwendungsfall kommen diese Felder zum Teil oder alle zum Tragen. Eine detaillierte Darstellung der aktuellen Rechtslage finden Sie im VDI-Statusreport Maschinelles Lernen in KMU [4] (ebenfalls nicht im Sinne einer Rechtsberatung).

FÜNF

HARDWARE UND SOFTWARE FÜR ML-PROJEKTE

Welche technische Ausstattung benötigen Sie, um ML-Projekte in Ihrem Unternehmen umzusetzen? Mit dieser Kurzbetrachtung von erforderlicher Soft- und Hardware erhalten Sie schnell einen ersten Überblick.

HARDWARE

Für die Auswertung eines Bildes mit einem neuronalen Netz sind mathematisch betrachtet lediglich Multiplikationen und Additionen erforderlich. Dieser Prozess ist also sehr unkompliziert und benötigt dementsprechend vergleichsweise wenig Rechenleistung. Je nach Netz und Anwendung ist eine Auswertung innerhalb weniger Millisekunden auch für gewöhnliche Hardwaresysteme durchaus möglich. Der eigentliche Rechenaufwand entsteht vielmehr im Lernprozess, da hier nicht nur ein Bild betrachtet wird, sondern alle für das Lernen bereitgestellten Bilder durchlaufen werden, in den meisten Fällen sogar mehrfach. Dieser Vorgang kann je nach Datenlage und Netz mehrere Tage dauern.

Da das Training in der Entwicklungsphase nicht nur einmal, sondern aufgrund der Optimierung vielfach durchgeführt wird, ist eine effiziente Hardware essenziell. Deshalb wird das Training aktuell nicht auf normalen Prozessoren, sondern nahezu ausschließlich auf Grafikkarten durchgeführt. Deren Aufbau erlaubt ein gleichzeitiges Ausführen mehrerer gleichartiger Berechnungsschritte, wodurch die Geschwindigkeit beim Training immens gesteigert wird. Längst entwickeln führende Hersteller auch speziell auf ML-Aufgaben zugeschnittene Grafikkarten.

Die Phasen, in denen ein Netz trainiert wird, umfassen nur einen Bruchteil der Gesamtlaufzeit eines Projekts. Für ein einzelnes Projekt angeschaffte und aufgrund der benötigten Hardware eher teure Rechnersysteme werden also nur für kurze Zeit verwendet. Entsprechend haben Anbieter wie Amazon Web Services oder Microsoft Azure dieses Problem bereits früh erkannt und bieten speziell auf das Training zugeschnittene, cloudbasierte Großrechnersysteme zur Miete an. Für einzelne Projekte ist dies meist am kostengünstigsten. Schnittstellen zu typischen Entwicklungsumgebungen machen diesen Ansatz zudem sehr zugänglich. Werden hingegen viele Projekte durchgeführt oder

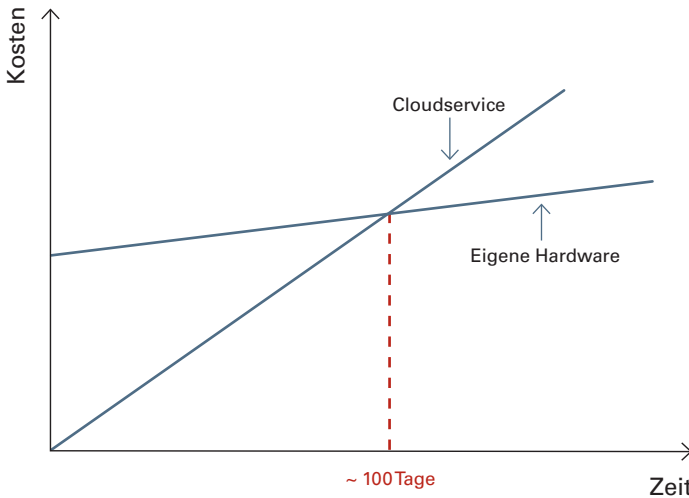


Abbildung 10: Kostenentwicklung eigene Hardware vs. Cloudservice [5].

die Rechnersysteme auch für andere Aufgaben verwendet, übersteigen die Mietkosten auf Dauer die Anschaffungskosten. Grob überschlagen lohnt sich die Anschaffung eigener Hardware ab etwa 100 Tagen Vollzeitraining [5] (vgl. Abbildung 10). Wenn Sie Cloud Services nutzen möchten, sollten Sie bedenken, dass die zum Lernen verwendeten Daten das Unternehmen verlassen. Vergessen Sie also nicht, bei besonders sensiblen Informationen die Datensicherheit zuvor zu prüfen.

Cloudbasierte Dienste können eine sinnvolle Option für das Training der Netze sein.

Für die im Vergleich zum Training einfach zu bewerkstellende Auswertung eines Bildes mit dem fertig trainierten Netz können Sie je nach benötigter Berechnungsgeschwindigkeit prozessor- oder grafikartenbasierte Systeme nutzen. Hersteller bieten zudem immer häufiger spezialisierte Hardwarebausteine für diese Auswertung an.

Eine weitere, recht junge Entwicklung ist die sogenannte **Smart Camera**, also eine Kombination aus Kamera und auswertender Hardware. Diese Geräte können mit einem zuvor erstellten Netz bespielt werden. Der Vorteil: Die Auswertung findet bereits in der Kamera statt! Es muss also nicht das ganze Bild übertragen werden, sondern das Ergebnis einer **Klassifikation** (z. B. defekt/in Ordnung) kann direkt an weitere am Prozess beteiligte Maschinen weitergegeben werden. Allerdings lässt die Hardware der Smart Camera nur bestimmte Netzstrukturen zu, beispielsweise die Anzahl an Schichten und Neuronen – die Anpassung der Netzstruktur an eine bestimmte Aufgabe ist nur eingeschränkt möglich. Für komplexere Anwendungen oder besonders hohe Erwartungen an die Fehlerrate sind Smart Cameras also aktuell noch nicht geeignet. Zudem lässt sich bisher nur die **Klassifikation** mit Smart Cameras umsetzen. Die **Objekterkennung** und **Segmentierung** haben sich hier noch nicht etabliert.

SMART CAMERA

Kombination aus Kamera und Hardware, die die aufgenommenen Daten direkt mit einem neuronalen Netz auswertet. Smart Cameras können aktuell nur für weniger komplexe Klassifikationsaufgaben eingesetzt werden.

SOFTWARE

Wenn Sie ein ML-Projekt umsetzen wollen, müssen Sie nicht alle benötigten Softwareelemente von Grund auf selbst entwickeln. Es existieren zahlreiche Bibliotheken und spezialisierte Software mit umfangreichen und flexiblen Methoden.

ML-Bibliotheken stellen Bausteine für den Aufbau von neuronalen Netzen in gängigen Programmiersprachen wie Python zur Verfügung. Sie sind in den meisten Fällen kostenlos und sehr flexibel einsetzbar. Allerdings erfordern sie nach wie vor ein umfangreiches Verständnis der verwendeten ML-Mechaniken und dementsprechend ausgebildetes Fachpersonal.

ML-BIBLIOTHEK

Sammlung von Funktionen und Softwarebausteinen, die frei verfügbar sind und in gängigen Programmiersprachen eingesetzt werden können.

INTEGRIERTE ML-SOFTWARE

Zumeist ostenpflichtige Entwicklungsumgebungen, die kein detailliertes ML-Wissen voraussetzen.

Integrierte ML-Software hingegen ist stärker gekapselt und somit auch für weniger spezialisiertes Personal zugänglich. Im Gegensatz zu ML-Bibliotheken wird integrierte ML-Software von den entwickelnden Firmen vermarktet und ist in der Regel nicht kostenfrei. Zum Service der Firmen gehören dafür auch Aktualisierungen inklusive Rückwärtskompatibilität. Durch die starke Kapselung handelt es sich bei auf kommerzieller ML-Software basierenden Entwicklungen meist um Blackbox-Systeme, die genaue Funktionsweise ist im Gegensatz zu ML-Bibliotheken nicht einsehbar. Im direkten Vergleich ist integrierte Software zudem weniger flexibel.

Zusätzlich ist Software für das effektive Sammeln, Labeln und Verwalten von Lerndaten sinnvoll.

Wenn Sie sich in Ihrem Unternehmen nicht mit der ML-Methodik auseinandersetzen, wohl aber die Vorteile nutzen wollen, können Sie auf Dienstleister zurückgreifen, die sowohl die Entwicklung des neuronalen Netzes als auch die Aufbereitung der Daten und das Training übernehmen.

SECHS

EIGENENTWICKLUNG ODER DIENSTLEISTER?

Welcher Ansatz ist der richtige für Ihr Unternehmen? Klären Sie dafür die Frage nach der langfristigen Strategie Ihres Betriebs bezüglich ML. Der Aufbau von Personal, Expertise und technologischer Infrastruktur für ein einzelnes, kleines Projekt wird ökonomisch kaum zielführend sein. Insbesondere werden die ML-spezifischen Kompetenzen bei einzelnen Projekten nur temporär benötigt, was dafür spricht, diese Expertise extern zuzukaufen. Möchten Sie jedoch längerfristig von ML profitieren, ist der Aufbau und die Pflege von **eigenem Know-how** unbedingt empfehlenswert!

NÖTIGE KOMPETENZEN UND PERSONALLAGE

Ob spezialisiertes Personal mit ML-Bibliotheken, ein Mittelweg über integrierte ML-Software oder die Anwerbung von Dienstleistungsagenturen den besten Weg darstellt, ist eine Frage der Geschäftsmodelle und der langfristigen Unternehmensstrategie.

Wenn Sie ML in Ihrem Unternehmen ohne entsprechendes Personal integrieren möchten, ist der Einsatz von **Smart Cameras** eine Option. Diese sind empfehlenswert, wenn die zu lösende Aufgabe (aus ML-Sicht) weniger komplex ist (siehe Kapitel 3.2, Seite 20). Beispiele sind die Unterscheidung „Bauteil/Produkt in Ordnung“, „Bauteil/Produkt defekt“ oder „Bauteil/Produkt fehlt“. Bei der Verwendung von Smart Cameras brauchen Sie (neben der Fleißarbeit des Labelns und dem relevanten Domänenwissen) lediglich Fachpersonal für die Anlagentechnik, das die Kamera integriert.

Möchten Sie hingegen komplexere ML-Projekte im Unternehmen selbst durchführen, benötigen Sie KI/ML-Expertise. Qualifizierte Fachkräfte mit Spezialwissen im maschinellen Lernen (z. B. Data Scientists) sind hochbegehrte und entsprechend rare. Erfahrungen zeigen jedoch, dass eine jahrelange Expertise aufgrund der zur Verfügung stehenden

Bibliotheken und Softwarepakete und tiefere Informatik- oder Programmierfähigkeiten nicht zwingend erforderlich sind. Vielmehr ist ein ausgeprägtes Interesse an der Technologie entscheidend.

Da ML eine junge Technologie ist, haben viele interessierte Studierende sowie Absolventinnen und Absolventen der MINT-Fächer (Mathematik, Ingenieurwesen, Naturwissenschaften, Technik) bereits Kontakt mit KI, ML und Softwarebibliotheken. An den Universitäten ist die Bereitschaft groß, sich mit dieser Schlüsseltechnologie zu beschäftigen.

Wenn Sie Studierende etwa über Praktika oder Abschlussarbeiten im Unternehmen mit ML-Projekten betrauen, dürfen Sie keine marktreife Lösung erwarten. Aber gerade wenn Sie noch nicht wissen, ob eine größere Investition in ML für Ihr Unternehmen sinnvoll ist, können studentische Hilfskräfte eine kostengünstige Machbarkeitsstudie erarbeiten, auf der weitere Aktivitäten aufbauen können. So können Sie auch begehrte Fachkräfte schon frühzeitig an Ihr Unternehmen binden. Besonders als KMU können Sie so mit geringem Investitionsvolumen erste Schritte in Richtung ML gehen.

Neben dem ML-Know-how ist die Verfügbarkeit und Qualität der Daten entscheidend. Je nach Datentyp liegen die Kompetenzen hier beispielsweise bei den Fachkräften für Messtechnik, Controlling oder Qualitätsmanagement. Unterschätzen Sie nicht den Aufwand beim Labeln der Daten und nehmen Sie dies unbedingt in die Personalplanung auf. In den meisten Fällen lässt sich hierfür bestehendes Personal abstellen. Weitere Kompetenzen sind je nach Projekt im Bereich Hardware und insbesondere bei Onlineanwendungen in der Anlagen- und Automatisierungstechnik erforderlich. Häufig sind diese bereits im Unternehmen vorhanden und lassen sich bei Bedarf durch Schulungen ergänzen.

Um den Erfolg mit ML sicherzustellen, muss die Belegschaft in Ihrem Unternehmen ML als Technologie akzeptieren. Stellen Sie vorab sicher, dass künstliche Intelligenz als Chance und nicht als Risiko wahrgenommen wird.

Die Akzeptanz der Technologie ML ist zwingende Voraussetzung für deren erfolgreichen Einsatz im Unternehmen.

Kommunizieren Sie mit allen Mitarbeitenden, insbesondere späteren Anwendenden, von Anfang an klar und binden Sie sie in das Projekt mit ein. Unternehmensweit muss die Bereitschaft bestehen, Kapazitäten für die Entwicklung und das Datenmanagement vorzuhalten.

Zusammenfassend zeigt sich, dass häufig bereits ein großer Teil der erforderlichen Kompetenzen im Unternehmen vorhanden ist und Kompetenzlücken gezielt durch Fortbildungen geschlossen werden können. Die Expertise für ML hingegen muss in den meisten Fällen zugeführt werden. Natürlich können Sie vorhandenes Personal weiterbilden, das Angebot an (teils sogar kostenfreien) Fortbildungen ist gut. Erfahrungen zeigen jedoch, dass solche Bestrebungen nur teilweise erfolgreich sind. Sollten Sie diesen Weg dennoch wählen, sollte Ihr Team uneingeschränkte Motivation und Lernbereitschaft mitbringen.

Wenn Sie schließlich die ML-Expertise im Unternehmen etabliert haben, sollte Sie diese durch regelmäßige Weiterbildung up to date halten, um auch weiterhin von Innovationen profitieren zu können.

DIENSTLEISTUNG GEZIELT EINBINDEN

Da Expertise im Bereich ML zumeist nicht im Unternehmen zu finden ist und nur temporär gebraucht wird, liegt es nahe, das benötigte Wissen extern zuzukaufen und die zugehörigen Prozesse an Dienstleistungsunternehmen zu übertragen.

Die Aufgaben, die Dienstleistende gut übernehmen können, umfassen insbesondere die Auswahl, die Implementierung, die Datenaufbereitung, das Training und die Parameteroptimierung. Je nach Projekt können auch die Erhebung und/oder die Kategorisierung der Trainingsdaten sowie die Validierung und die Integration delegiert werden. Die Identifikation geeigneter Aufgaben und die Integration in Ihre Geschäftsmodelle, Prinzipien, die Sicherstellung des Return on Investment etc. kann Ihnen jedoch niemand abnehmen!

Einige der Prozesse eines ML-Projektes können nur vom Unternehmen selbst bearbeitet werden.

Natürlich können Sie sich (z. B. im Rahmen von interdisziplinären Workshops) auch hier begleitend unterstützen lassen, das nötige Know-how kann jedoch nur das Unternehmen selbst einbringen.

Welche der Teilschritte eines ML-Projekts an Externe abgegeben werden können, zeigt Abbildung 11.

Befassen Sie sich bitte auch mit den Nachteilen einer externen Lösung! Am schwersten wiegt, dass Sie im eigenen Unternehmen auf diesem Wege kaum Expertise im Bereich ML sammeln können. In jedem weiteren Projekt, in dem ML eingesetzt werden soll, müssen Sie erneut jemanden beauftragen. Zudem können nachträgliche Änderungen oder Anpassungen oft (trotz vom Dienstleistungsunternehmen offen gestalteter Prozesse) nicht selbstständig umgesetzt werden und sind somit immer mit erheblichem finanziellen Mehraufwand verbunden.

Abbildung 11: Die blau eingefärbten Prozesse können von Dienstleistenden übernommen werden. Je nach Projekt gilt dies auch für die schraffierten Punkte. Die rot eingefärbten Prozesse hingegen erfordern die Expertise des Unternehmens selbst und können nicht abgegeben werden. Modifiziert aus [4].



FINANZIELLE FÖRDERUNG FÜR ML-PROJEKTE NUTZEN?

Neben der Beauftragung von Dienstleistungsunternehmen gibt es eine weitere Möglichkeit, externes Know-how in Ihrem Unternehmen zu nutzen. Bei der sogenannten Forschungs- und Entwicklungsförderung werden durch Ministerien auf Länder-, Bundes- oder EU-Ebene anteilig Mittel für innovative Projekte bereitgestellt. Dabei ist vorgegeben,

dass die Technologieentwicklung im Verbund mit weiteren Partnern, insbesondere auch aus dem universitären Bereich, durchgeführt werden muss. Auch ML-Start-ups bieten sich als Partner an. Diese sind manchmal sogar in direkter Nachbarschaft zu finden.

Die Förderung durch öffentliche Gelder erwartet ein gewisses Maß an Innovation, also eine Mindest-Sprunghöhe. Insbesondere die umfangreiche EU-Förderung ist für die meisten Unternehmen nicht erreichbar. Zudem müssen die Ergebnisse offengelegt werden. Eine Förderung kommt also nicht für jedes Projekt infrage.

Die Politik hat die Bedeutung von KI und ML, insbesondere auch für KMU, jedoch erkannt und das Förderangebot ist groß. Immer wieder ergeben sich (teilweise abhängig vom Bundesland) Möglichkeiten, auch kleine, weniger komplexe Projekte teilfinanzieren zu lassen. Einige nützliche Links für weitere Informationen finden Sie auf [Seite 44](#) dieses Leitfadens.

Unabhängig davon, ob Partnerschaften als Dienstleistung, durch Förderung oder anderweitig entstehen, sollte Sie sorgfältig auswählen. Das Feld KI ist riesig, und niemand kann alle Methoden kennen, geschweige denn anwenden. Wenn Sie ein ML-Projekt durchführen wollen, suchen Sie also keine Fachkräfte für KI, sondern speziell für ML, idealerweise mit Vorerfahrungen in der Bildverarbeitung.

Das Forschungsfeld KI ist riesig. Bei der Partnerwahl sind Vorerfahrungen mit den relevanten Methoden entscheidend.

SIEBEN

FAZIT

Maschinelles Lernen in der Bildverarbeitung kann in Unternehmen unterschiedlichster Größe Mehrwert generieren. Offensichtlichstes Anwendungsgebiet ist das Qualitätsmanagement, in dem Sie Arbeitsabläufe schneller und/oder sicherer gestalten können. Die Anwendungsgebiete sind jedoch vielfältig und von Unternehmen zu Unternehmen und Geschäftsmodell zu Geschäftsmodell unterschiedlich. Es ist Ihre Aufgabe, diese Anwendungsgebiete in Ihrem eigenen Unternehmen zu identifizieren.

Geben Sie sich Zeit, Sie können bei maschinellem Lernen nicht auf eine fertige, voll funktionsfähige Lösung zurückgreifen. Vielmehr müssen Sie bereit sein, neben finanziellen Ressourcen auch Zeit, Personal und Engagement zu investieren. Für viele KMU ist dies zunächst ein Hindernis. Allerdings lohnt sich je nach Intensität der Nutzung und Kürze des erwarteten Return on Investment der Aufwand. Während einige Prozesse von ML-Projekten durchaus komplett ausgelagert werden können, setzen andere die nur im Unternehmen vorhandene Expertise voraus. Langfristig ist die Verwendung von ML besonders fruchtbar, wenn die Expertise im Unternehmen selbst entsteht. So kann ML Teil der Unternehmensstruktur und -identität werden.

Lassen Sie sich nicht von der hohen Komplexität eines ML-Projekts abschrecken! Die Einstiegshürde ist nicht hoch. So kann die Bearbeitung einer ersten, kleinen Projektidee durch interessierte Studierende oder ein Start-up ein guter Einstieg mit vielen neuen Erkenntnissen für Ihr Unternehmen sein. Einfache Anwendungen lassen sich womöglich bereits mit einer Smart Camera umsetzen, und somit ohne ML-Spezialkenntnisse. Sollte die Nutzung von ML langfristig vielversprechend sein, können in Zukunft weitere Investitionen folgen. Gutes Gelingen!

IHRE NOTIZEN

ACHT

LITERATUR UND WEITERE INFORMATIONEN

Informieren Sie sich über diesen Leitfaden hinaus über zahlreiche frei zugängliche Quellen weiter und vertiefen Ihr Wissen in ausgewählten Teilbereichen. Im Folgenden finden Sie einige hilfreiche Studien, Berichte aus Gremien und weitere Onlinequellen mit nützlichen Informationen.

ALLGEMEINE INFORMATIONEN

- VDI-Statusreport „Maschinelles Lernen – Künstliche Intelligenz mit neuronalen Netzen in optischen Mess- und Prüfsystemen“, November 2019.
- VDI-Statusreport „Maschinelles Lernen in KMU – Künstliche Intelligenz im eigenen Unternehmen nutzen“, November 2020.
- An executive's guide to AI – McKinsey,
<https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/An%20executives%20guide%20to%20AI/An-executives-guide-to-AI.ashx>,
 abgerufen am 07.07.2021.
- KI im Mittelstand – Potenziale erkennen, Voraussetzungen schaffen, Transformation meistern. Plattform Künstliche Intelligenz, Juni 2021,
<https://www.acatech.de/publikation/ki-im-mittelstand-potenziale-erkennen-voraussetzungen-schaffen-transformation-meistern/>,
 abgerufen am 17.02.2022.

F&E-FÖRDERUNG

- Forschungs- und Innovationsförderung – Ein Wegweiser für kleine und mittlere Unternehmen. Förderberatung des Bundes,
https://www.bmbf.de/upload_filestore/pub/Forschungs_und_Innovationsfoerderung.pdf,
 abgerufen am 07.07.2021.
- Förderdatenbank, www.foerderdatenbank.de, abgerufen am 07.07.2021.

FREI ZUGÄNGLICHE BROWSERBASIERTE DEMOS ZUR FUNKTIONSWEISE NEURONALER NETZE

- <http://playground.tensorflow.org>
 - <https://www.tensorflow.org/js/demos>
-

QUELLEN

- [1] „Was ist künstliche Intelligenz?“
<https://news.sap.com/germany/2018/03/was-ist-kuenstliche-intelligenz/>,
abgerufen am 01.07.2021.
- [2] „The Neural Net Tank Urban Legend“: <https://www.gwern.net/Tanks>,
abgerufen am 01.07.2021.
- [3] Prof. Wolfgang Kersten, Sebastian Lodemann: „Welchen Einfluss hat KI auf KMU-Prozesse und Mitarbeiter? Ein Einblick in den Bereich Supply Chain Management“ (Vortrag).
https://www.hamburg-logistik.net/fileadmin/user_upload/blog/bilder/Veranstaltungen/VA_LIHH/do.innovation/2019/E_-_TUHH_doinnovation.pdf,
abgerufen am 30.08.2021.
- [4] VDI-Statusreport „Maschinelles Lernen in KMU – Künstliche Intelligenz im eigenen Unternehmen nutzen“, November 2020.
<https://www.vdi.de/ueber-uns/presse/publikationen/details/vdi-statusreport-maschinelles-lernen-in-kmu>,
abgerufen am 17.02.2022.
- [5] KI in der Bildverarbeitung – Vortrag der IDS Imaging Development Systems GmbH beim KI Lab Kurpfalz, <https://www.youtube.com/watch?v=VZ0YDjXnDw>,
abgerufen am 17.02.2022.
- [6] KI im Mittelstand – Potenziale erkennen, Voraussetzungen schaffen, Transformation meistern. Plattform Künstliche Intelligenz, Juni 2021.

NEUN

MITTELSTAND-DIGITAL ZENTRUM HAMBURG

Für kleine und mittlere Unternehmen bei Fragen und Herausforderungen der digitalen Transformation.

KONTAKT:

Mittelstand-Digital
Zentrum Hamburg
Handelskammer Hamburg

Adolphsplatz 1
20457 Hamburg
Tel.: +49 40 36138-443
kompetenzzentrum@hk24.de

PROJEKTPARTNER:

Konsortialführer des Mittelstand-Digital Zentrum Hamburg und zentraler Ansprechpartner für Unternehmen ist die Handelskammer Hamburg.



WEITERE LEITFÄDEN FINDEN SIE HIER:



QR-Code mit dem
Smartphone
abscannen



...und viele mehr!

www.digitalzentrum-hamburg.de
www.twitter.com/DigiZentrum_HH

Weitere Projektpartner im Mittelstand-Digital Zentrum Hamburg sind:

Technische
Universität Hamburg



Hochschule für
Angewandte
Wissenschaften
Hamburg



Handwerks-
kammer
Hamburg



KOOPERATIONSPARTNER:

Helmut-Schmidt-
Universität Hamburg



ZEHN

IMPRESSUM

HERAUSGEBER:

Prof. Dr.-Ing. Alexander Fay
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg
Kooperationspartner des Mittelstand-Digital Zentrum Hamburg

AUTOREN:

Dr. Jakob Krooß
Für das Mittelstand-Digital Zentrum Hamburg

LEKTORAT:

Katrin Meyer

GESTALTUNG:

www.lockvogel-hamburg.de

BILDNACHWEIS:

ludariimago/stock.adobe.com (1), L.Darin/stock.adobe.com (9),
AndSus/stock.adobe.com (11), jagodka/stock.adobe.com (14),
grafikplusfoto/stock.adobe.com (14), Gelpi/stock.adobe.com (14),
Happy monkey/stock.adobe.com (14), Photoboyko/stock.adobe.com (15),
Ольга Васильева/stock.adobe.com (20), kichigin19/stock.adobe.com (20),
MdLothfor/stock.adobe.com (20), enterphoto/stock.adobe.com (20),
Sattaya/stock.adobe.com (21), Suraphol/stock.adobe.com (27)

AUFLAGE:

1. Auflage, 02/2023

Was ist Mittelstand-Digital?

Mittelstand-Digital informiert kleine und mittlere Unternehmen über die Chancen und Herausforderungen der Digitalisierung. Regionale Kompetenzzentren helfen vor Ort dem kleinen Einzelhändler genauso wie dem größeren Produktionsbetrieb mit Expertenwissen, Demonstrationszentren, Netzwerken zum Erfahrungsaustausch und praktischen Beispielen. Das Bundesministerium für Wirtschaft und Klimaschutz ermöglicht die kostenlose Nutzung aller Angebote von Mittelstand-Digital. Weitere Informationen finden Sie unter www.mittelstand-digital.de

www.digitalzentrum-hamburg.de



MITTELSTAND-DIGITAL ZENTRUM HAMBURG

Adolphsplatz 1, 20457 Hamburg
Tel.: +49 40 36138-443, kompetenzzentrum@hk24.de