

## 47 Mathematische Architekturen für Neuronale Netze

Ohne Mathematik geht es auch bei der KI nicht. Optimale Netzwerkarchitekturen für komplexe Anwendungen sind am besten, wenn sie die zugrundeliegende mathematische Theorie abbilden. Dies gilt insbesondere für bildgebende Verfahren der Medizin.

KI, das ist doch Informatik, oder? So denken wohl die meisten, wobei es ohne Mathematik auch in diesem neuen, faszinierenden Forschungsfeld nicht geht. Ein besonders anspruchsvolles Feld der KI-Anwendungen sind sogenannte inverse Probleme, hierzu zählen die meisten bildgebenden Verfahren der Medizin ebenso wie die Parametrisierung von mathematischen Finanzmodellen oder Techniken der Überwachung und Steuerung komplexer Produktionsprozesse. Mit derartigen Problemen beschäftigten sich Prof. Dr. Peter Maass und seine Mitarbeiter Sören Dittmer, Tobias Kluth, Johannes Leuschner und Maximilian Schmidt an der Universität Bremen.

Aber was genau sind nun inverse Probleme? Das sind Probleme bei denen die gesuchte physikalische oder technische Größe nicht direkt gemessen, sondern indirekt bestimmt wird. Ein einfaches Beispiel ist die Messung der Höhe einer Quecksilbersäule (Messung) aus der dann die aktuelle Raumtemperatur (gesuchter Parameter) 'berechnet' wird. Dieses Beispiel ist allerdings doch zu einfach, im engeren Sinn spricht man von inversen Problemen erst, wenn zwei Schwierigkeiten hinzukommen. Als erstes werden nur Prozesse betrachtet, bei der beliebig kleine Änderungen der Messung zu beliebig großen Abweichungen des gesuchten Parameters führen können. Lange Zeit wurden derartige Probleme als „schlecht gestellt“ bezeichnet, in der Annahme, dass hier die physikalische Modellierung unvollständig oder ungenau sei. Erst Mitte der 1960er Jahre konnte mit Methoden der Funktionalanalysis gezeigt werden, dass diese Instabilität leider bei vielen Problemen inhärent ist und nicht durch bessere Messtechniken oder exaktere Modellierung behoben werden kann. Diese Instabilität wäre an sich kein Problem, wenn exakte Messdaten zur Verfügung stehen würden. Hier kommt allerdings die zweite Schwierigkeit ins Spiel: Messdaten sind nie exakt, sondern bestenfalls im Rahmen einer vorgegebenen Messgenauigkeit bekannt.

In Abbildung 66 sind drei Rekonstruktion aus CT-Daten abgebildet, die alle die vorgegeben Messdaten im Rahmen der Messgenauigkeit des Tomographen wiedergeben. Welche ist nun die richtige Rekonstruktion? Ein der grundlegenden Erkenntnisse der Theorie inverse Probleme besagt, dass dies bei derartig instabilen Problemen auf die jeweilige Anwendung ankommt. In dem vorliegenden Beispiel würde ein Arzt, der Diagnosen an Hand der Gewebestruktur stellt, die mittlere bevorzugen, ein Arzt, der rein an der Form der Organe interessiert ist, wohl eher die rechte auswählen.

Es gibt also keine „richtigen“ Lösungen gibt, sondern bestenfalls „gute“ Lösungen, die über ein Kriterium, das von Experten des Anwendungsgebiets individuell formuliert wird, ausgewählt werden.

Derartige inverse Probleme zählen zu den komplexesten Herausforderungen der angewandten Mathematik und ihre Lösung ist zudem äußerst rechenintensiv. Erste Versu-

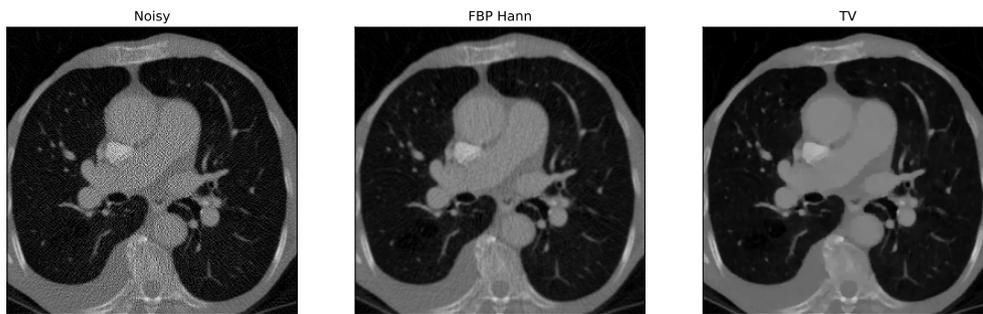


Abbildung 66: Drei CT-Rekonstruktionen, die aus demselben Datensatz erzeugt wurden. Die linke Rekonstruktion ist offensichtlich wertlos (noisy), obwohl sie die Messdaten genauso gut wiedergibt wie die beiden anderen. Ob die mittlere oder rechte Rekonstruktion bevorzugt wird, hängt davon ab, was der medizinische Experte damit machen will. Eine Diagnose, die auf der Gewebetextur der Organe beruht würde die mittlere bevorzugen, eine Diagnose, die rein auf Form und Morphologie der Organe abzielt die rechte.

che, dies mit Hilfe neuronaler Netze oder anderer Verfahren der KI effizienter zu lösen, sind allerdings gescheitert oder führten zu suboptimalen Ergebnissen. Eine direkte Anwendung der KI-Verfahren, die im Bereich der Computer Vision und Bildverarbeitung so erfolgreich sind, scheitert an der unvermeidbaren Instabilität derartiger inverser Probleme.

Bevor wir über konkrete Anwendungen neuronaler Netze für inverse Probleme reden, lassen Sie uns erst einmal klären, was neuronale Netze überhaupt sind und wie diese mathematisch beschrieben werden. Neuronale Netze sind dem menschlichen Nervensystem nachempfunden. Beim Menschen werden Reize über Nervenbahnen weitergeleitet und dabei entweder verstärkt oder reduziert. In den sogenannten Neuronen werden dann die ankommenden Reize verschiedener Nervenbahnen gebündelt und über einen bio-chemischen Prozess in einen ausgehenden Reiz zur Weiterleitung transformiert. Anschaulich kann man sich das Nervensystem als ein Gewirr von Nervenbahnen vorstellen, die die eingehenden Reize an eine erste Schicht von Neuronen weiterleiten. Von dort aus werden dann analog die ausgehenden Reize dieser ersten Neuronenschicht gewichtet und an eine zweite Schicht weitergeleitet usw.

Mathematisch gesprochen sind die Reize einfach Zahlen, die mit Gewichten multipliziert (Nervenbahn) und danach zusammengezählt werden. Die Aktion der Neuronen entspricht der Auswertung einer typischerweise nicht-linearen Funktion, diese wird Aktivierungsfunktion genannt. Das Ergebnis dieser Auswertung wird dann an die nachfolgenden Neuronenschichten weitergegeben.

Letztendlich entsteht eine Rechenvorschrift, die eine Menge von Zahlen als Eingabe bekommt und über die iterative Anwendung von Gewichtung (lineare Transformation,

entspricht Weiterleitung von Reizen entlang von Nervenbahnen) und Aktivierungsfunktion (entspricht der Aktionen, die in einem Neuron von statten gehen) eine festgelegte Anzahl von Ausgabewerten erzeugt. Ein einfaches neuronales Netz ist in Abbildung 67 dargestellt. Die Festlegung der Anzahl der Schichten, der Anzahl der Neuronen je Schicht sowie die Auswahl der Aktivierungsfunktion bestimmen die Architektur des Netzwerks.

So wie der Mensch aus Erfahrung lernt kann nun ein neuronales Netz trainiert werden, in dem ihm Trainingsdaten mit bekanntem Ergebnis vorgelegt werden. Die Gewichte des Netzwerks werden dann so berechnet, dass das Netzwerk dieses Ergebnis bestmöglich reproduziert.

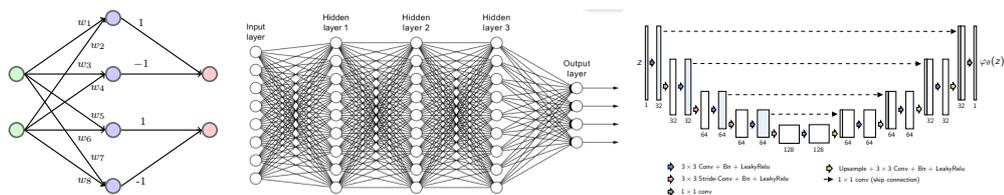


Abbildung 67: Links ist ein einfaches neuronales Netz zu sehen, das als Eingabe zwei Zahlen erhält und ebenso zwei Werte ausgibt, das mittlere Bild ist ein „fully connected feedforward“ Netz und das rechte Bild stellt das häufig verwendete „U-Net“ mit ca. 500.000 freien Parametern dar.

Um zu verstehen, was dies mit KI zu tun hat, muss man sich erst vergegenwärtigen, dass neuronale Netze, genauer gesagt „fully connected feed forward“ Netzwerke mit  $L$  Schichten als Realisierung von  $L$  Schritten eines Iterationsverfahrens interpretiert werden können. Ein derartiges neuronales Netz hat eine Vektor  $x_0$  als Eingabegröße. Die Werte des Vektors werden von links nach rechts, Schicht für Schicht weitertransportiert. Zwischen den Schichten, werden die einzelnen Werte mit Gewichten  $w$  multipliziert, das entspricht der Weiterleitung inkl. Dämpfung oder Verstärkung von Reizen entlang der Nervenbahnen. In den „Neuronen“ der einzelnen Schichten werden dann die einzelnen ankommenden Werte zusammengefasst und komponentenweise nicht-linearen Transformation unterworfen, das entspricht den biochemischen Reaktionen, die in den einzelnen Neuronen des Körpers z.B. entscheiden, ob der Reiz überhaupt weitergeleitet werden soll.

## Mathematische Grundlagen

Der Weg zu einem erfolgreichen Einsatz neuronaler Netze bei Inversen Problemen führt über Netzwerkarchitekturen, die aus tiefliegenden Forschungsergebnissen zur Analysis Inverser Probleme abgeleitet sind.

Hierzu braucht man Einblicke in die mathematischen Grundlagen, sowohl von inversen Problemen als auch von neuronalen Netzen.

Bei einem inversen Problem wird – in einer allgemeinen Formulierung – ein unbekannter Parameter  $x$ , für den lediglich Messwerte  $y^\delta$  bekannt sind, die über einen Messprozess  $F$  mit einer Messgenauigkeit  $\delta$  gemessenen Daten wurden. In dieser Formulierung sind alle Parameter  $x$ , die

$$\|F(x) - y^\delta\| \leq \delta$$

erfüllen, potentielle Lösungen des Problems. Unter diesen wird nun derjenige Parameter ausgewählt, der ein zusätzliches Kriterium  $R_{\text{expert}}$  am besten erfüllt. Die genaue Festlegung von  $R$  kann recht allgemein sein, hier können die Anwendungsexperten Eigenschaften kodieren, wie z.B. dass die Kanteninformation in den Rekonstruktionen besonders wichtig ist. Mit Hilfe der Lagrange-Multiplikatoren führt dies zu einem Minimierungsproblem

$$\min_x \|F(x) - y^\delta\|^2 + \alpha R_{\text{expert}}(x).$$

Derartige Minimierungsaufgaben können für ein nicht-lineares  $F$  und nicht-differenzierbare Funktionale  $R_{\text{expert}}$  nicht direkt gelöst werden. Ableiten nach  $x$  und Nullsetzen erlaubt aber die Konstruktion von Iterationsverfahren, die ausgehend von einem geeigneten gewählten  $x_0$  Schritt für Schritt die Lösung annähern. Derartige Gradientenabstiegsverfahren führen z.B. auf

$$x_{k+1} = \text{Prox}_R \left( x_k - \lambda F'(x_k)^* (F(x_k) - y^\delta) \right).$$

Um zu verstehen, was dies mit KI zu tun hat, muss man sich erst vergegenwärtigen, dass neuronale Netze, genauer gesagt „fully connected feed forward“ Netzwerke mit  $L$  Schichten als Realisierung von  $L$  Schritten eines Iterationsverfahrens interpretiert werden können. Ein derartiges neuronales Netz hat einen Vektor  $x_0$  als Eingabegröße. Die Werte des Vektors werden von links nach rechts, Schicht für Schicht weitertransportiert. Zwischen den Schichten, werden die einzelnen Werte mit Gewichten  $w$  multipliziert, das entspricht der Weiterleitung inkl. Dämpfung oder Verstärkung von Reizen entlang der Nervenbahnen. In den „Neuronen“ der einzelnen Schichten werden dann die einzelnen ankommenden Werte zusammengefasst und komponentenweise weiterverarbeitet, das entspricht den biochemischen Reaktionen, die in den einzelnen Neuronen des Körpers z.B. entscheiden, ob der Reiz überhaupt und wenn ja wie stark weitergeleitet werden soll.

Mathematisch formuliert lässt sich ein fully connected feedforward Netzwerk mit  $L$  Schichten zur Lösung des inversen Problems also ebenfalls als iterative Rechenvorschrift schreiben. Zu gegebenem  $y^\delta$  und mit beliebig gewähltem  $x_0$  wird  $b = -\lambda F'(x_0)^* y^\delta$  gesetzt. Dann entspricht die Anwendung des neuronalen Netzes der Iterationsvorschrift

$$x_{k+1} = \varphi(Wx_k + b).$$

Für die spezielle Wahl  $W = I - \lambda F'(x_k)^* F(x_k)$ ,  $b = -\lambda F'(x_k)^* y^\delta$  und  $\varphi = \text{Prox}_R$  sieht man die Analogie zu dem analytischen Gradientenabstiegsverfahren. Entscheidend ist aber, dass die Aktivierungsfunktion  $\varphi$  des Netzwerks entsprechend der mathematischen Theorie als proximal Mapping  $\text{Prox}_R$  der Experteninformation  $R$  gewählt

wird. Darauf wäre man ohne die mathematische Theorie im Hinterkopf nicht gekommen. In diesem Fall entspricht jedenfalls der Output des Netzwerks, der mit  $\Phi(y^\delta, W)$  bezeichnet wird, exakt der  $L$ -ten Iterierten des analytischen Gradientenabstiegsverfahrens.

Der Vorteil der neuronalen Netze, ist jedoch, dass die Matrix  $W$  und der Bias-Vektor  $b$  an Hand von experimentell gemessenen Daten optimal an die jeweilige Anwendung angepasst werden: Hierzu werden zu mehreren bekannten Parametern  $x_i, i = 1, \dots, N$  die Messdaten  $y_i$  erhoben. Danach werden die Gewichte  $W$  über die Minimierung der sogenannten Loss-Funktion

$$\sum_{i=1}^N \|\Phi(y_i^\delta, W) - x_i(W)\|^2$$

bestimmt. So kann das Netzwerk optimal an eine gegebenen Datenstruktur angepasst werden und selbst feinste Modellierungsdetails des Messoperators sind in den Messdaten enthalten. Damit wird das Netzwerk dahin getrimmt, Outputs zu erzeugen, die ähnlich denen der Testdaten sind.

## Anwendung

Als Anwendungsbeispiel wollen wir ein besonders anspruchsvolles und relativ neues tomographisches Verfahren wählen. Bei dem sogenannten „magnetic particle imaging“ (MPI) werden kleinste Nanopartikel in die Blutbahn injiziert. Diese Nanopartikel haben einen magnetisierbaren metallischen Kern und eine bioaktive Hülle. Diese Partikel werden nun mit dem Blutfluss transportiert und erlauben – sofern man ihre aktuelle Position nachverfolgen kann – detaillierten Aufschluss über Blockaden in Blutgefäßen und über die Herzfunktion. Prinzipiell kann die bioaktive Hülle auch so konstruiert werden, dass sich die Nanopartikel an bestimmten metabolisch charakteristischen Regionen ansiedeln und so detailliert funktionale Erkenntnisse liefern. Diese Nanopartikel werden danach über den natürlichen Kreislauf wieder ausgeschieden und sind so potentiell eine harmlose Alternative zu radioaktiven Kontrastmitteln.

Dabei entsteht allerdings die Frage, wie kann die Position der Nanopartikel nachverfolgen nachdem sie injiziert wurden? Hierzu wird ein dynamisches Magnetfeld außerhalb des Körpers angelegt, um gezielt Nanopartikel in bestimmten Bereichen zu magnetisieren. Die dynamische Änderung dieses äußeren Magnetfelds führt nun dazu, dass Nanopartikel anfangen zu schwingen und ein eigenes elektromagnetisches Feld zu erzeugen. Dieses wird wiederum durch Spulen außerhalb des Körpers gemessen. In einem einfachen linearen Modell wird das elektromagnetische Feld, das von einem einzelnen Nanopartikel in Position  $x$  erzeugt wird, mit  $s(x, t)$  bezeichnet. Ist nun  $c(x)$  die Anzahl oder Konzentration der Nanopartikel an der Position  $x$  so entsteht das Gesamtsignal durch Summation (Integration) der Beiträge an allen Positionen im Körper  $\Omega$

$$y(t) = \int_{\Omega} c(x)s(x, t) dx .$$

Das ist allerdings ein stark vereinfachtes Modell das z.B. sowohl die komplexe Mescharakteristik der messenden Spulen und die Partikel-Partikel Interaktionen ebenso wie den Einfluss der Größenverteilung der Nanopartikel außer Acht lässt. Dies lässt sich alles nur sehr schwer modellieren. Aber, wie gesagt, alle diese Modellierungseinheiten sind in hinreichend guten Testdatensätzen enthalten. Dies macht den Einsatz daten-getriebener Verfahren in diesem Bereich so attraktiv.

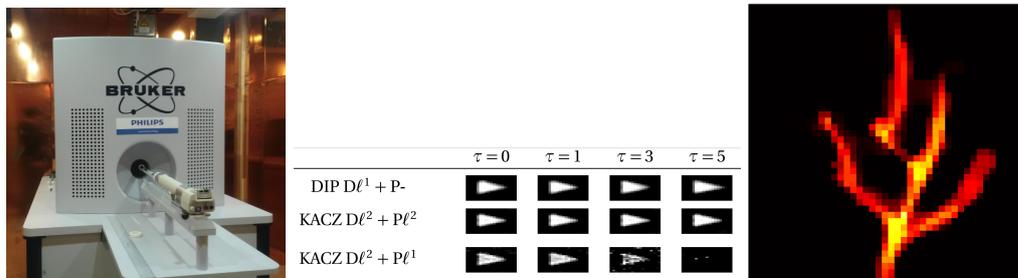


Abbildung 68: Links ist ein MPI-Gerät des UKE (Universitätsklinikum Hamburg Eppendorf) zu sehen, die mittleren Grafiken sind Rekonstruktionen eines real gemessenen Phantoms, das einmal herkömmlich (KACZ) und einmal über neuronale Netze (DIP) rekonstruiert wurde, das rechte Bild zeigt eine KI-Rekonstruktion eines Adernbaums.