Inverse Bestimmung elastischer Materialparameter aus dem Ausbreitungsverhalten geführter Ultraschallwellen mittels eines Convolutional-Neuronal-Networks

Mathias Held[1], Abdalla Rashwan[1], Jannis Bulling[1], Yevgeniya Lugovtsova[1], Maik Lauschkin[1], Jens Prager[1]

[1]*Bundesanstalt für Materialforschung und -prüfung, Unter den Eichen 87, 12205 Berlin*
Mathias.Held@bam.de, Jens.Prager@bam.de

Einleitung

Detailliertes Wissen über die charakteristischen Eigenschaften verwendeter Materialien ist Grundvoraussetzung für viele ingenieurtechnische Aufgaben. Zur Bestimmung der mechanischen Eigenschaften von Materialen gibt es verschiedene klassische, zerstörende Prüfverfahren. Eine Möglichkeit der zerstörungsfreien Bestimmung von elastischen Materialkonstanten in dünnen Körpern liegt in der Auswertung von Messergebnissen, die auf Basis des Ausbreitungsverhaltens geführter Ultraschallwellen gewonnen werden. Geführte Ultraschallwellen weisen ein multimodales und dispersives Verhalten auf. Dieses Verhalten kann mittels Dispersionsabbildungen dargestellt werden, vgl. Abb.1.

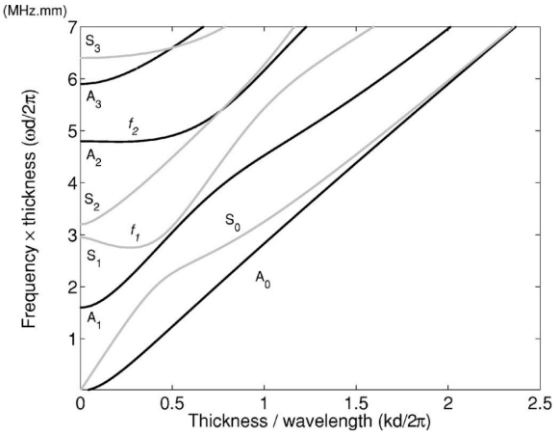

Abbildung 1: Dispersionsabbildung einer Stahlplatte mit den Moden A0 bis S3 und den Punkten der „Zero Group Velocity“ f_1 und f_2 , aus [1]

Um aus messtechnisch ermittelten Dispersionsabbildungen Rückschlüsse auf die Materialparameter zu ziehen, werden in der aktuellen Forschung verschiedene inverse Methoden diskutiert. In [2] wird ein Simulationsmodell basierend auf der SAFE-Methode an eine messtechnisch ermittelte Dispersionsabbildung angepasst. In [1] wird der Punkt der Zero Group Velocity (f1 und f2) zur Bestimmung der Materialparameter aus Dispersionsabbildungen verwendet, vgl. Abb.1.

Maschinelles Lernen und insbesondere Convolutional-Neural-Networks (CNNs) stellen eine Möglichkeit der automatisierten inversen Auswertung von Bilddaten dar. In diesem Beitrag werden CNNs verwendet, um die elastischen Konstanten von isotropen dünnen plattenförmigen Materialen bei verschiedenen Dicken aus Dispersionsabbildungen zu bestimmten.

Daten

Um ein überwachtes maschinelles Lernverfahren anzuwenden, ist ein vielfältiger Trainingsdatensatz zum Anlernen des Modells notwendig. Reale Messungen erfordern einen hohen zeitlichen Aufwand sowie eine Vielzahl verschiedener Materialproben. Daher wird für das Training ein synthetischer Datensatz verwendet. Hierfür wird eine semi-analytische Methode, die Scaled Boundary Finite Element Method (SBFEM), die in MATLAB programmiert wurde, verwendet [3], [4], [5]. Diese semianalytische Methode ermöglicht eine effiziente Simulation des Ausbreitungsverhalten geführter Wellen und eignet sich somit zur Generierung großer Datenmengen. Beispielsweise konnte in ca. 40h mit Hilfe eines modernen Computers ein Datensatz bestehend aus 4000 Dispersionsabbildungen erstellt werden. Das beispielhafte Vorgehen ist in Abb.2 dargestellt. Um eine realitätsnahe Anregung zu simulieren, wird ein Gaußscher Impuls mit begrenzter Bandbreite verwendet. Die simulierte Wellenausbreitung wird anhand der Auslenkung entlang einer Linie auf der Oberfläche ausgewertet. Mittels einer 2D-Fourier-Transformation (2D FFT) werden die gewonnen Daten aus dem Zeit-/Ortsbereich in den Frequenz-/Wellenzahlbereich transformiert und Dispersionsabbildungen erstellt.

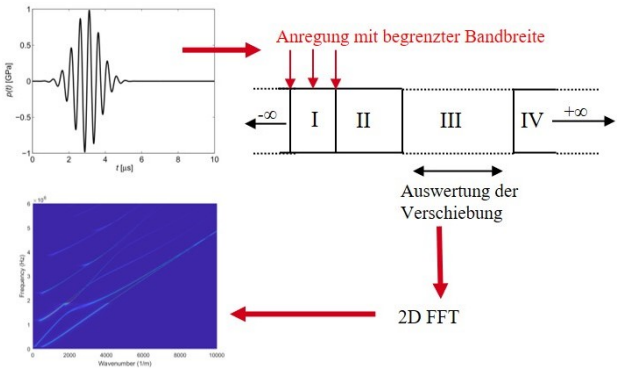

Abbildung 2: Schematische Darstellung des Vorgehens zur Erzeugung von Dispersionsabbildungen aus Simulationsdaten, nach [6]

Zur Reduzierung der variierbaren Eingabeparameter kann das E-Modul auf die Dichte normiert werden. Dies sieht man beispielsweise an der Formel zum Körperschall, vgl. (1).

$$c_{long} = \sqrt{\frac{E(1-\nu)}{\rho\,(1-\nu-2\nu^2)}} = \sqrt{\frac{E_{norm}(1-\nu)}{\rho_{norm}\,(1-\nu-2\nu^2)}} \quad (1)$$

$$mit\ E_{norm} = \frac{E}{\rho}\ und\ \rho_{norm} = 1$$

Für gängige isotrope metallische Materialen, wurde für E_{norm} und die Poissonzahl ν der in Tabelle 1 dargestellte Wertebereich gewählt. Die Variation der Dicke findet im Bereich 1-3 mm statt. Alle erzeugten Datenpunkte und ihre zugehörigen Materialparameter sind innerhalb des gewählten Bereichs zufällig verteilt. Der Testdatensatz beinhaltet die Daten, welche die Extremwerte des gewählten Bereichs abbilden.

Tabelle 1: Beispiele für Materialparameter verschiedener metallischer isotroper Materealien und des für die Simulation gewählten Wertebereichs

Material	Parameter		
	E-Modul [N/mm²]	Dichte [kg/m³]	E_{norm} [m²/s²]
Alu-minium	60 000 bis 72 000	2700	~ $22{,}2\cdot10^6$ bis ~ $26{,}7\cdot10^6$
Eisen	170 000 bis 196 000	7874	~ $21{,}6\cdot10^6$ bis ~ $24{,}9\cdot10^6$
Stahl	190 000 bis 210 000	7670 - 8360	~ $22{,}7\cdot10^6$ bis ~ $27{,}4\cdot0^6$
Magne-sium	42 000	1738	~ $24{,}2\cdot10^6$
Gewählt:	E-Modul / Dichte [m²/s²]		$21\cdot10^6$ bis $28\cdot10^6$
	Poissonzahl		0,2 bis 0,45

Dispersionsabbildungen

Der dargestellte Wertebereich in den Dispersionsabbildungen ist auf die jeweilige Dicke des simulierten Datenpunktes skaliert, vgl. Abb.1. Auf der y-Achse ist die Frequenz · Dicke, im Bereich 0 bis 5000 [Hz · m] und auf der x-Achse die Wellenzahl · Dicke im Bereich von 0 bis 1,5 [(1/m) · m] aufgetragen.

Da die Frequenz der ausbreitungsfähigen Moden mit der Dicke skaliert und die Achsen der Dispersionsabbildungen auf die Dicke skaliert sind, ist der Verlauf der Moden für verschiedene Dicken mit den gleichen elastischen Konstanten identisch. Aufgrund der in der Simulation, vgl. Abb.2, begrenzten Bandbreite der Anregung, verschiebt sich jedoch der sichtbare Bereich der Moden bei verschiedenen Dicken und gleichen elastischen Konstanten in den Dispersionsabbildungen.

Bei Verwendung dieses inversen Modells wird somit Wissen über die Dichte, vgl. (1), und über die Dicke vorausgesetzt.

Vorverarbeitung

Zur weiteren Vereinfachung der Bilddaten werden die Achsenbeschriftungen aus den Dispersionsabbildungen gelöscht da diese keine für das CNN verwertbare Information darstellen. Um die in den Bildern vorhandene Information für das CNN in Relation zu setzen, ist es notwendig, den Wertebereich der Achsen für alle Abbildungen konstant zu halten. Zusätzlich ist in den generierten Daten die Anzahl der Pixel und das Seitenverhältnis der Abbildungen konstant.

Die Farbkodierung der mittels der SBFEM und der 2D FFT erzeugten Dispersionsabbildungen ist indirekt von der Amplitude des Signals abhängig. Die Höhe der Amplitude ist abhängig von der simulierten Anregung und stellt für diesen Anwendungsfall keine relevante Information dar. Die Reduzierung der Farbkanäle von RGB zu Grauwert-Abbildungen stellen somit eine Konzentration der gewünschten Information dar.

Zur Konvertierung der RGB Abbildungen in Grauwerte wird die Python Bibliothek OpenCV (cv2 4.5.0) verwendet. Um im Rahmen der Konvertierung digitale Artefakte zu entfernen wird ein Algorithmus mit Blurring, Dilatation, Erosion und einem adaptiven Threshold verwendet.

Zur Anwendung des maschinellen Lernverfahrens werden zusätzlich alle Ein- und Ausgabewerte bzw. Pixel in den Wertebereich null bis eins normiert. Die Normierung der Zielparameter erfolgt hierbei mittels der MinMax-Normierung und ist im Weiteren mit # gekennzeichnet.

Modell

Zur Erstellung des Modells und der Anwendung des überwachten maschinellen Lernverfahrens wird Python 3.8.5 mit den Bibliotheken Tensorflow (2.3) und Keras (2.4.3) verwendet. Sofern nicht weiter angegeben wurden die Standartparameter der entsprechenden Versionen gewählt. Alle Startparameter werden zufällig initialisiert und zur besseren Nachvollziehbarkeit mit entsprechenden Seeds versehen. Hinweise zur Problematik der Wiederholbarkeit von maschinellen Lernverfahren mit NVIDIA Grafikkarten in Tensorflow ist in [7] zu finden.

Die Grundarchitektur des gewählten Modells ist die eines Standard-Feed-Forward-Modells zur Klassifikation, welches an das Regressionsproblem angepasst wird. Für die Parameter der Faltungsschicht (Convolutional Layer) wurde eine Filtergröße von 3x3xTiefe mit Bias gewählt. Die Tiefe der jeweiligen Filter ist abhängig von der Tiefe der vorherigen Schicht. Die Anzahl der trainierbaren Parameter ergibt sich somit beispielsweise für die zweite Schicht zu 3x3x16x32+32 (Filtergröße * Tiefe_vorher * Anzahl_Filter + Anzahl_Bias), vgl. Tabelle 2. Die Schrittweite der Filter beträgt eins und der Bias wird mit null initialisiert.

Tabelle 2: Darstellung der gewählten Architektur

Schichten	Ausgabe (h,b,t)	Parameter
Eingabe	656, 875, 1	
1.0 Faltung 1.1 Aktiv.funktion	654, 873, 16	160
2.0 Faltung 2.1 Aktiv.funktion 2.2 Pooling	326, 435, 32	4640
… Wdh. Schicht 2 mit Verdopplung der Filteranzahl …		
7.0 Faltung 7.1 Aktiv.funktion	16, 23, 1024	4719616
… Flattening …		
8.0 Voll-Verknüpft 8.1 Aktiv.funktion	1	376832

Weiterhin wird jeder Faltungsschicht eine ReLU-Aktivierungsfunktion hinzugefügt (vgl. Tabelle 2). Hierdurch werden alle negativen Werte zu null gesetzt. Als Pooling Funktion wird MaxPooling mit einer Schrittweite von zwei und einer Filtergröße von 2x2 verwendet. Hierdurch halbiert sich die Anzahl der Pixel der einzelnen Abbildungen in der Merkmalskarte (Feature Map). Die Tiefe bzw. Anzahl der Abbildungen in der Merkmalskarte bleibt konstant.

In der vollverknüpften Entscheidungsschicht ist kein Bias enthalten. Weiterhin wird ein separates Modell für jeden vorherzusagenden Parameter trainiert. Die Intention hierfür ist, die Einflüsse der Bildbereiche auf die verschiedenen Parameter eindeutiger zuordnen zu können.

Training

Das Training wird mittels einer Nvidia Quattro P4000 8GB ausgeführt und dauert ca. eine Stunde. Die gewählte Batch Size von 25 ist die größtmögliche bei 8 GB verfügbaren RAM.

Für die Anwendung des überwachten maschinellen Lernverfahrens wird der Datensatz aus 4000 Bildern in 3040 Trainings- und 760 Validations- sowie 200 Testdaten eingeteilt. Eine Durchmischung der Datensätze findet hierbei nicht statt. Für das Training wird der Adam Optimizer mit der mittleren quadratischen Abweichung (MSE) und sinkender Lernrate zur Skalierung der adaptiven Updates sowie Begrenzung der maximalen Updategröße verwendet, vgl. [8]. Zu Beginn wird eine Lernrate von 0,001 gewählt. Diese halbiert sich, sollte sich innerhalb der letzten 5 Epochen keine Verbesserung des MSE des Validationsdatensatzes von mindestens $5 \cdot 10^6$ ergeben haben. Das Training wird über 100 Epochen ausgeführt und nach jeder Epoche das beste Modell abgespeichert. Aufgrund der Trägheitsparameter des Adam Optimizers erfahren während des Trainings auch die mit null initialisierten Gewichte des Bias ein Update.

Auswertung

In einer vorrangegangen Veröffentlichung [9] wurde ein kleinerer Wertebereich sowie keine Dickenvariation untersucht. In dem dabei verwendeten Modell wird in den Faltungsschichten Zero Padding verwendet, um einen Informationsverlust in den Randbereichen zu vermeiden. Eine tiefergehende Evaluation dieser Methode ergab, dass Padding in Kombination mit den Filtern der Faltungsschicht, insbesondere durch Verwendung eines Bias, zu Streifen in den Randbereichen der Einträge in der Merkmalsschicht führt. Mit zunehmender Tiefe des CNN pflanzen sich diese zunehmend in die Bildmitte fort. Generell sollte daher bei der Verwendung von CNNs auf Padding verzichtet werden. Diese Annahme wurde kürzlich in [10] bestätigt.

Das zusammengefasste Ergebnis dieser Untersuchung ist in Tabelle 3 zu finden. Neben dem MSE# bezogen auf den MinMax normierten Wertebereich ist der reale mittlere absolute Fehler (MAE) in Prozent angegeben.

Tabelle 3: Ergebnis nach Training der CNNs

	E_{norm}	Poissonzahl
Epoche	100	59
Validationsdaten MSE#	$6.60*10^6$	$4{,}53*10^5$
Testdaten MSE#	$1{,}73*10^5$	$1{,}96*10^4$
Testdaten MAE in %	0,069 %	0,581 %

#Bezogen auf den MinMax normierten Wertebereich

Zur weiteren Evaluation des CNNs wird die Modellantwort auf ein leeres Bild, eine Matrix bestehend aus Nullen und ein volles Bild, eine Matrix bestehend aus Einsen, untersucht. Die Ergebnisse sind in Tabelle 4 dargestellt. Für ein leeres Bild zeigt sich, dass das Modell einen Wert nahe dem Mittelwert (0,5) des MinMax normierten zu vorhersagenden Wertebereichs und für ein volles Bild einen Wert deutlich über dem maximalen zu vorhersagenden Wert (1) annimmt. Mittels dieser Methode wird gezeigt, dass der ursprünglich mit null initialisierte Bias des Modells durch das Training den Mittelwert des zu vorhersagenden Wertebereichs der Trainingsdaten annimmt, vgl. Tabelle 4, „leeres Bild“.

Tabelle 4: Vorhersage des Trainierten Modells anhand eines leeren und eines vollen Bildes für den normierten Wertebereich (#)

	E_{norm} #	Poissonzahl #
Leeres Bild	0,595	0,58
Volles Bild	5,847	10,198

Zur weiteren Evaluation wird die Grad-CAM Class Activation Mapping) Methode und die Guided-Backpropagation, welche ursprünglich zur Evaluation von Klassifikationsproblemen entwickelt wurde, auf das Regressionsproblem angewandt. Beide Methoden haben zum Ziel, den Einfluss verschiedener Bildbereiche auf die jeweilige Vorhersage darzustellen und so das Ergebnis zu validieren und erklärbar zu machen. Da beide Methoden nicht das komplette Modell abbilden, werden die Darstellungen in Kombination bewertet.

In Abb. 3 ist beispielhaft das Ergebnis der beiden Methoden für eine spezifische Eingabe dargestellt. In der Eingabe sind deutlich der Verlauf der Moden und die Begrenzung des dargestellten Bereichs, aufgrund der begrenzten Bandbreite der simulierten Anregung, zu sehen. Weiterhin zeigen die beiden Analysemethoden, dass die Entscheidung des CNNs vom Verlauf der Kurven und nicht von anderen Merkmalen bzw. digitalen Artefakten abhängig ist. Bei Betrachtung des gesamten Testdatensatzes wird zudem deutlich, dass Bereiche hoher Frequenzen tendenziell einen stärkeren Einfluss auf das Ergebnis haben und eine Übereinstimmung mit der in [11] gezeigten Sensitivitätsanalyse auftritt. Die Vorhersage des Modells erfolgt somit ausgehend vom ungefähren Mittelwert, vgl. Tabelle 4, anhand der Krümmung der Kurven in den sensitiven Bereichen.

Eine stichprobenartige Analyse der Merkmalskarte der letzten Faltungsschicht nach Anwendung der Aktivierungsfunktion zeigt, dass nur wenige der 1024 Einträge während einer Vorhersage aktiv sind. Zwar unterscheidet sich die Aktivität der Felder der

Merkmalskarte entsprechend der jeweiligen Eingabe, jedoch sind abhängig vom zu vorhersagenden Parameter bzw. verwendetem Modell meist die gleichen Felder aktiv. Eine deutliche Reduzierung der Anzahl der Filter scheint somit möglich.

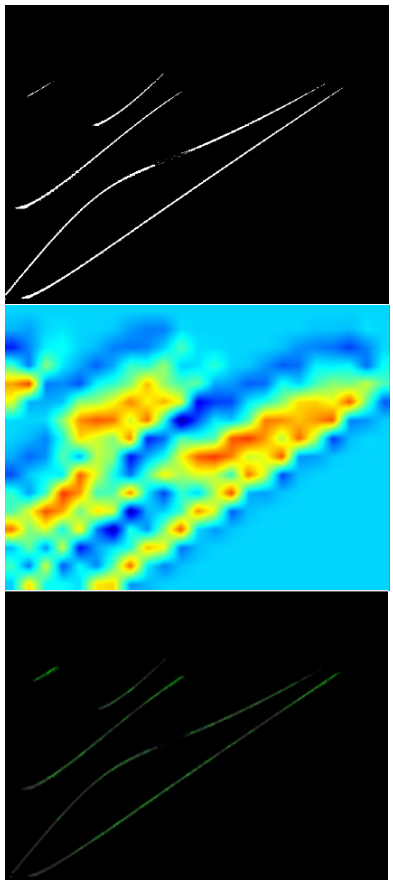

Abbildung 3: Beispielhafte Darstellung eines Ergebnisses der Grad-CAM und Guided-Backpropagation Analyse zur Vorhersage der Poissonzahl

In einer weiteren Untersuchung wurde die Möglichkeit der Vorhersage der Dicke mit dem gezeigten Modell untersucht. Die Vorhersage erreichte mit einem MAE von 0,493 % eine sehr gute Genauigkeit, ähnlichen zu den in Tabelle 3 dargestellten Ergebnissen. Die Analyse mittels der gezeigten Methoden ergab jedoch, dass die Vorhersage nicht anhand des Verlaufs der Kurven, sondern anhand der Verschiebung des sichtbaren Bereichs der Kurven erfolgt und somit abhängig von der Bandbreite der simulierten Anregung ist, vgl. Abb. 2.

Zusammenfassung

Es wurde gezeigt, dass eine Vorhersage der isotropen elastischen Konstanten mittels CNNs anhand von Dispersionsabbildungen möglich ist. Die Verwendung von ursprünglich für Klassifikationsprobleme entwickelten Analysemethoden wie Grad-CAM und Guided Backpropagation kann auch zur Interpretation bei Regressionsproblemen angewandt werden. Mit den gezeigten Evaluationsmethoden kann validiert werden, dass die Vorhersage der Werte nicht anhand von digitalen Artefakten erfolgt. Die gezeigte Methode der Analyse mittels leerer und voller Bilder kann zudem zur weiteren Erklärbarkeit der Vorhersage von Neuronalen Netzen beitragen. Zur weiteren Validation der Modelle sind Untersuchungen mit realen, messtechnisch gewonnenen Dispersionsabbildungen notwendig, um abschließend eine Transformation der Methode auf anisotrope Materialien durchzuführen.

Literatur

[1] Clorennec, D., C. Prada, and D. Royer, Local and noncontact measurements of bulk acoustic wave velocities in thin isotropic plates and shells using zero group velocity Lamb modes. Journal of Applied Physics, 2007. 101(3).

[2] Webersen, M., et al., Guided ultrasonic waves for determining effective orthotropic material parameters of continuous-fiber reinforced thermoplastic plates. Ultrasonics, 2018. 84: p. 53-62.

[3] Gravenkamp, H., C. Song, and J. Prager, A numerical approach for the computation of dispersion relations for plate structures using the Scaled Boundary Finite Element Method. Journal of Sound and Vibration, 2012. 331(11): p. 2543-2557.

[4] Gravenkamp, H., et al., The computation of dispersion relations for three-dimensional elastic waveguides using the Scaled Boundary Finite Element Method. Journal of Sound and Vibration, 2013. 332(15): p. 3756-3771.

[5] Rose, J.L., Ultrasonic guided waves in solid media. 2014, New York NY: Cambridge University Press. xxii, 512 pages.

[6] Gravenkamp, H., C. Birk, and C. Song, Simulation of elastic guided waves interacting with defects in arbitrarily long structures using the Scaled Boundary Finite Element Method. Journal of Computational Physics, 2015. 295: p. 438-455.

[7] Access: 10.06.2021 https://github.com/NVIDIA/framework-determinism.

[8] Kingma, D.P. and J. Ba Adam: A Method for Stochastic Optimization. 2014. arXiv:1412.6980.

[9] Held, M., et al., Inverse Determination of Elastic Material Parameters from Ultrasonic Guided Waves Dispersion Measurements using Convolutional Neuronal Networks, in SMSI 2021 - Measurement Science. 2021. p. 239-240.

[10] Alsallakh, B., et al. Mind the Pad -- CNNs can Develop Blind Spots. 2020. arXiv:2010.02178.

[11] Bulling, J., et al., Sensitivity of Ultrasonic Guided Waves to Elastic Constants: A Numerical Study, in European Workshop on Structural Health Monitoring. 2021. p. 759-768.