

# „Maschinelles Lernen in der physikalischen Akustik“

## Programm und Kurzfassungen der Vorträge

28. DEGA-Workshop „Physikalische Akustik“ im Physikzentrum Bad Honnef, gemeinsam veranstaltet vom Fachausschuss Physikalische Akustik der Deutschen Gesellschaft für Akustik (DEGA) und vom Fachverband Akustik der Deutschen Physikalischen Gesellschaft (DPG)

**Donnerstag, 19.10.2023**

Mittlerer Hörsaal (im rechten Seitenflügel, 1. OG)	
9:30	<b>Willkommen</b> Jens Prager, Christian Adams
9:40	<b>1. Domain Shift: Probleme und Herausforderungen beim Einsatz von ML in der Akustik</b> Olena Prianychnykova [a], Sascha Grollmisch [a], Saichand Gourishetti [a], Jakob Abeßer [a], Franca Bittner [a], Joachim Bös [a,b] <i>[a] Fraunhofer-Institut für Digitale Medientechnologie IDMT, [b] Fachgebiet Industrielle Anwendungen von Medientechnologien IAM, Technische Universität Ilmenau</i>
10:20	<b>2. Erhöhung der Simulationsgenauigkeit mittels KI</b> Amir Charmi, Lukas Beinlich, Mohsen Barzegar, Jannis Bulling, Yevgeniya Lugovtsova, Bengisu Yilmaz <i>Bundesanstalt für Materialforschung und -prüfung</i>
11:00	Kaffeepause
11:20	<b>3. Herausforderungen und Lösungen für industrielle KI in Microcontrollern und FPGAs</b> Christian Fuchs [a], Lukas Krupp [b], Julian Schauer [a], Tizian Schneider [a] <i>[a] Universität des Saarlandes, [b] Fraunhofer IMS, Duisburg</i>
12:00	<b>4. Design to Acoustics Through Deep Learning</b> Julius Schultz, Jan van Delden, Christopher Blech, Sabine C. Langer, Timo Lüddecke <i>TU Braunschweig</i>
12:40	Mittagessen
13:40	<b>gemeinsame Sitzung</b> des Fachausschusses Physikalische Akustik der Deutschen Gesellschaft für Akustik (DEGA) und des Fachverbands Akustik der Deutschen Physikalischen Gesellschaft (DPG) – Gäste sind willkommen!
14:20	<b>5. Verantwortungsethik beim maschinellen Lernen</b> Monika Gatt, Marcus Mäder, Steffen Marburg <i>TU München</i>
15:00	<b>6. Deep learning for parameterized spatio-temporal oscillations</b> Merten Stender [a], Jakob Ohlsen [b] <i>[a] TU Berlin, [b] TU Hamburg</i>
15:40	<b>Kaffeepause</b>

16:00	<b>7. Gitterfreie Schallquellencharakterisierung aus Mikrofonarraydaten mittels Transformer Architektur</b> Adam Kujawski, Ennes Sarradj <i>TU Berlin</i>
16:40	<b>8. Gridless beamforming with permutation invariant loss</b> Armin Goudarzi <i>DLR Göttingen</i>
17:20	<b>9. Neural Grid Compression: using neural networks to speed up beamforming deconvolution algorithms</b> Thiago Lobato <i>Head acoustics</i>
ca. 18:00	Ende des Vortragsprogramms am Donnerstag
ab 19:00	gemeinsames Abendessen im Lichtenberg-Keller

### **Freitag, 20.10.2023**

<b>Mittlerer Hörsaal (im rechten Seitenflügel, 1. OG)</b>	
9:00	<b>10. Entwicklung von Data-Augmentation-Strategien für Wälzlagerschäden</b> Katharina Marburg, Oliver Gnepper, André Schneider, Olaf Enge-Rosenblatt <i>Fraunhofer-Institut für Integrierte Schaltungen, Institutsteil Entwicklung Adaptiver Systeme, Dresden</i>
9:40	<b>11. Machine Learning-Driven Acoustic Mode Decomposition in Flow Ducts</b> Stefan Jacob <i>PTB/TU Braunschweig</i>
10:20	<b>Kaffeepause</b>
10:40	<b>12. Maschinelles Lernen zum Entrauschen hydroakustischer Daten</b> Christoph Pilger, Andreas Steinberg, Peter Gaebler und Johanna Lehr <i>Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Fachbereich B4.3: Erbebedienst des Bundes, Kernwaffenteststopp</i>
11:20	<b>13. Auswertung von EEG-Daten mit Maschinellern Lernen</b> Christian Koch <i>PTB</i>
12:00	<b>14. Machine-Learning zur Modellierung des akustischen Verhaltens von eBike-Antriebseinheiten</b> Annika Hayn <i>Robert Bosch GmbH</i>
12:40	<b>Zusammenfassung, Abschlussdiskussion und Verabschiedung</b> Jens Prager, Christian Adams
13:00	gemeinsames Mittagessen (Ende des Workshops)

# Kurzfassungen/Inhalte der Vorträge

Donnerstag, 19.10.2023

## 1. Domain Shift: Probleme und Herausforderungen beim Einsatz von ML in der Akustik

Olena Prianychnykova [a], Sascha Grollmisch [a], Saichand Gourishetti [a], Jakob Abeßer [a], Franca Bittner [a], Joachim Bös [a,b]

[a] Fraunhofer-Institut für Digitale Medientechnologie IDMT, [b] Fachgebiet Industrielle Anwendungen von Medientechnologien IAM, Technische Universität Ilmenau

Maschinelles Lernen, insbesondere in Form von künstlichen Neuronalen Netzen (Deep Learning), hat in den letzten Jahren sehr gute Ergebnisse in vielen Klassifikationsaufgaben erzielt und sich auch für die Analyse von Audiodaten als Stand der Technik etabliert. Trotz hoher Genauigkeit auf den vorab erstellten Testdaten ergeben sich jedoch oftmals Probleme im praktischen Einsatz, vor allem wenn sich die Umgebungsbedingungen anders darstellen als erwartet, beispielsweise durch den Austausch der Mikrofone oder andere akustische Bedingungen am Einsatzort. Diese Verschiebung zwischen den bestehenden Trainings- oder Testdaten und dem eigentlichen Anwendungsfall wird als Domain Shift bezeichnet und kann unter Umständen dazu führen, dass die erstellten Modelle keine verlässlichen Vorhersagen erzeugen. In diesem Vortrag werden bekannte Ursachen von Domain Shift für die Audioanalyse sowie die Implikationen für das maschinelle Lernen beleuchtet und aktuelle Lösungsansätze vorgestellt.

## 2. Erhöhung der Simulationsgenauigkeit mittels KI

Amir Charmi, Lukas Beinlich, Mohsen Barzegar, Jannis Bulling, Yevgeniya Lugovtsova, Bengisu Yilmaz

Bundesanstalt für Materialforschung und -prüfung

Structural Health Monitoring (SHM) mittels geführter Ultraschallwellen ermöglicht eine kontinuierliche Überwachung komplexer Bauteile und liefert umfangreiche Informationen über deren strukturelle Integrität. Die aufgezeichneten Signale sind jedoch komplex und werden daher zunehmend mittels KI verarbeitet, um Anomalien in der Systemintegrität automatisch zu erkennen. Eine Herausforderung dabei ist jedoch das Trainieren solcher KI-Modelle, da dafür große Mengen an Trainingsdaten inklusive diverser Schadensfälle benötigt werden, um eine hohe Genauigkeit und Zuverlässigkeit zu gewährleisten. Experimentelle Generierung dieser Daten ist wegen des enormen Zeit- und Kostenaufwands nicht möglich.

Eine Alternative hierfür bieten die numerischen Simulationen, da dadurch Trainingsdaten für beliebige Anzahl an Schadenskonfigurationen erzeugt werden können. Das Hauptproblem dabei ist jedoch die Abweichung zwischen den experimentellen und simulierten Daten (durch z.B. Abweichungen bei der Geometrie, Randbedingungen oder Materialparameter). Um diese Abweichungen zu reduzieren bzw. zu eliminieren, wird eines KI-Modells mit einer Autoencoder-Decoder Architektur vorgestellt, die mittels eines Experiments kalibriert werden kann und anschließend in der Lage ist, die Genauigkeit der Simulationsergebnisse nachträglich zu erhöhen.

## 3. Herausforderungen und Lösungen für industrielle KI in Microcontrollern und FPGAs

Christian Fuchs [a], Lukas Krupp [b], Julian Schauer [a], Tizian Schneider [a]

[a] Universität des Saarlandes, [b] Fraunhofer IMS, Duisburg

Algorithmen des maschinellen Lernens nehmen zunehmenden Platz in unserem beruflichen und privaten Alltag ein. In der Akustik finden sie langsam Zugang, nachdem sie zur Spracherkennung bereits in großem Umfang verwendet werden. Im industriellen Kontext bedarf es zur Datenerfassung Messelektronik. Diese Elektronik, die häufig sowieso einen Microcontroller enthält, kann auch genutzt werden, um aufgenommene Daten direkt mit der vorhandenen Rechenpower auszuwerten. Auf diesem Gedanken bauen aktuell mehrere Forschungsprojekte, wie z.B. KI-MUSIK4.0, auf. Im

Rahmen dieses Projektes haben wir Möglichkeiten und Wege erforscht, KI direkt auf Hardwareebene auszuführen oder Modelle anzuwenden. Dazu gibt es verschiedene Möglichkeiten und vor allem verschiedene Hardwareplattformen, Programmiersprachen und Modellformate.

Dieser Vortrag soll einen Einblick liefern, welche Plattformen zur Inferenz von neuronalen Netzen oder auch klassischen maschinellen Lernalgorithmen auf Hardwareplattformen ermöglichen. Ein Fokus liegt darauf, unter welchen Umständen dies möglich ist und ein weiterer darauf, wo die Unterschiede in den verschiedenen Wegen liegen.

Es werden vor allem zwei Alternativen diskutiert, die in KI-MUSIK4.0 zum Einsatz kommen. Zum einen die Inferenz per Darstellung im ONNX-Format, also das Speichern und Lesen von Netzstrukturen, und zum anderen die Inferenz über das direkte Ausführen von Python Code auf der Hardwareplattform.

#### **4. Deep learning for frequency response prediction of dynamic benchmark models**

Julius Schultz, Jan van Delden, Christopher Blech, Sabine C. Langer, Timo Lüddecke

*TU Braunschweig*

Noise prevention is an important aspect of product design nowadays, to ensure health and comfort restrictions by the users. Enabling such a *design to acoustics* requires the acoustical engineers to understand wave propagation in technical systems already in an early design stage. To this end, a detailed understanding of the dynamical response of the system in the frequency domain is required. Such frequency response functions are usually computed using costly numerical simulations, such as the finite element method. Data-driven surrogate models rely on the offline – online paradigm, shifting the computational cost to the offline training phase, while providing a fast-to-evaluate surrogate model for the online phase. Having such a surrogate model at hand allows performing important tasks like design space exploration, optimization or uncertainty quantification.

In this work, we investigate the potential of machine learning based surrogates for the prediction of frequency response functions. A multi-mass oscillator serves as a starting point to investigate different network architectures, such as a multi-layer perceptron (MLP) or a transformer architecture. In a next step, we consider a vibrating plate solved via the finite element method (FEM) as a more sophisticated benchmark for surrogate modeling. This model allows defining complex input spaces, such as bead pattern variations, as well as high dimensional target quantities, such as the velocity distributions over the plate. Again, we investigate different network architectures like ResNet18/ViT or Fourier Neural Operator and compare them against a k-nearest-neighbor baseline.

#### **5. Verantwortungsethik beim maschinellen Lernen**

Monika Gatt, Marcus Mäder, Steffen Marburg

*TU München*

Die zunehmende Technisierung und Digitalisierung der globalisierten Gesellschaft hat zu weitreichenden ethischen Problemen geführt. Klimatische Veränderungen einerseits und der Wunsch nach neuen, (teil-) automatisierten Lebensformen andererseits fordern ein neues, verantwortungsethisches Denken und Forschen. Der „Technische Imperativ“, der „Ökonomische Imperativ“ und der „Ethische (Kategorische) Imperativ“ stehen einander konkurrierend gegenüber. Das oftmals den Endverbraucher:innen und Kund:innen nicht ausreichend bewusste jedoch mittels lernender Maschinen umgesetzte unternehmerische Sammeln und Weitergeben von privaten Daten erschüttert aktuell das gesellschaftliche Vertrauen in Mensch und Technik zugleich.

In unserem Vorhaben vertreten wir einen verantwortungsethischen Forschungsansatz. Es wird ermittelt, welche vorhersehbaren ethischen Technikfolgen beim maschinellen Lernen perspektivisch erkannt und in den laufenden Prozess integriert werden können. Ziel ist es, negative Technikfolgen frühzeitig zu erkennen und zu beheben. Zwei Axiome sind bindend, erstens das moralische Axiom der Unantastbarkeit der menschlichen Würde und zweitens das kartesische Forschungsaxiom, das wissenschaftliche Forschung nicht auf subjektiver Gesinnung beruht, sondern wissenschaftliche Objektivität erzeugt.

## 6. Deep learning for parameterized spatio-temporal oscillations

Merten Stender [a], Jakob Ohlsen [b]

[a] TU Berlin, [b] TU Hamburg

Data-driven methods have become omnipresent across many fields of research, also in the field of structural vibrations and acoustics. This contribution introduces a very recent framework denoted Up-Net that allows for data fusion and time stepping of complex spatio-temporal dynamics. Using multiple encoding paths, the framework allows for fusing boundary conditions, spatial parameter distributions and field variables. Embedding the encoder-decoder structure into an auto-regressive feedback loop creates an autonomous and fully data-driven time stepping method. A range of successful applications are presented, e.g., for offshore pile driving noise prediction.

## 7. Gitterfreie Schallquellencharakterisierung aus Mikrofonarraydaten mittels Transformer Architektur

Adam Kujawski, Ennes Sarradj

TU Berlin

Mikrofonarrayverfahren leisten einen wesentlichen Beitrag zur Entwicklung geräuscharmer Maschinen und Fahrzeuge, indem sie es ermöglichen, Geräuschquellen zu identifizieren. Viele modellbasierte Verfahren zur Charakterisierung von Schallquellen mit Mikrofonarrays stützen sich auf physikalische Modelle, um Quellcharakteristiken wie Position und Stärke zu ermitteln. Meist erfolgt eine räumliche Diskretisierung des Quellbereichs mittels Berechnungsgitter und eine inverse Berechnung der Quellstärken aus den gemessenen Signalen. Abhängig von der Gitterauflösung sind diese Verfahren oft rechenintensiv oder in ihrer Rekonstruktionsgenauigkeit besonders bei großen Wellenlängen begrenzt.

Als vielversprechende Alternative gelten seit einigen Jahren datenbasierte Methoden aus dem Deep Learning Bereich, die eine gitterfreie Schätzung ermöglichen. Vorgestellt wird ein Modellansatz, mit dem die Quellcharakteristiken unter Verwendung einer Transformer Architektur und auf Grundlage der gemessenen Leistungsspektren geschätzt werden. Mit simulierten und auch experimentell ermittelten Daten wird demonstriert, dass das Verfahren nach erfolgtem Training zu schnelleren und bei großen Wellenlängen auch zu genaueren Schätzungen der Quellcharakteristiken im Vergleich zu modellbasierten Verfahren führt.

## 8. Gridless beamforming with permutation invariant loss

Armin Goudarzi

DLR Göttingen

We present gridless beamforming using deep learning with a permutation invariant loss to solve the inherent vector-to-set problem. Due to the gridless architecture, true and estimated sources are matched based on the minimum error to compute their loss. The matching can be efficiently performed using the Hungarian matching algorithm. The proposed architecture allows the combination of regression and classification variables to estimate the number of sources and their location, power, and other properties simultaneously. The method can include broadband sources, multipoles, and parameterized distributed sources.

## 9. Neural Grid Compression: using neural networks to speed up beamforming deconvolution algorithms

Thiago Lobato

Head acoustics

Beamforming results are limited by the spatial resolution of the microphone array used, which may lead to sources close to each other being considered as one. Deconvolution methods, such as DAMAS and CMF, produce better results in these situations. However, they have a high computational cost and thus often lack sufficient speed to be used in real-time applications. A solution to this

is to use a hybrid method to perform deconvolution using a neural network to compress the source grid and then solve the inverse problem only at the compressed locations, which can speed up the deconvolution on high-resolution grids by orders of magnitude. In this contribution we briefly discuss the application of neural grid compression in typical deconvolution algorithms and also show how it can be combined with an extended formulation of DAMAS to achieve real-time deconvolution of sources with complex radiation patterns (e.g. dipoles and quadrupoles).

**Freitag, 20.10.2023**

## **10. Entwicklung von Data-Augmentation-Strategien für Wälzlagerschäden**

Katharina Marburg, Oliver Gnepper, André Schneider, Olaf Enge-Rosenblatt  
*Fraunhofer-Institut für Integrierte Schaltungen, Institutsteil Entwicklung Adaptiver Systeme, Dresden*

Bei der Zustandsüberwachung von Wälzlagern werden Algorithmen des Maschinellen Lernens eingesetzt, um Schäden anhand von Körperschallsignalen zu erkennen und zu klassifizieren. Eine Herausforderung besteht jedoch darin, an ausreichend Trainingsdaten jeder Klasse zu gelangen, um diese Algorithmen zu trainieren. Da es nicht wirtschaftlich ist, beschädigte Wälzlager zu sammeln oder sie künstlich zu beschädigen, enthalten reale Datensätze fast ausschließlich Daten gesunder Lager, während die Schadklassen selten vertreten sind. Dieses Ungleichgewicht führt in der Regel zu einer Überanpassung des Modells an die Klasse der gesunden Lager. Um diesem Problem zu begegnen, lassen sich Data-Augmentation-Methoden einsetzen, um ausgehend von vorhandenen Zeitreihen künstliche Daten zu erzeugen und die unterrepräsentierten Klassen mit ihnen zu ergänzen. In diesem Beitrag wird darauf eingegangen, welche grundlegenden Ansätze für Data Augmentation von Hüllkurven und ihren Spektren existieren. Es werden Aspekte diskutiert, die bei einer Anwendung auf Wälzlagerdaten berücksichtigt werden sollten. Darauf aufbauend werden Klassifikationsergebnisse eines Faltungsnetzes präsentiert, die mit dem Wälzlagerdatensatz der Universität Paderborn gewonnen wurden.

## **11. Machine Learning-Driven Acoustic Mode Decomposition in Flow Ducts**

Stefan Jacob  
*PTB/TU Braunschweig*

Commonly employed in the field of flow duct analysis, acoustic mode decomposition serves as a technique for separating the constituent pressure wave components. Traditional methodologies rely on mathematical descriptions of wave motion. However, in this study, we introduce a novel approach to mode decomposition that leverages machine learning for the separation of acoustic mode content. We train a neural network using data derived from a limited set of numerical solutions of the Linearized Navier-Stokes equation within the frequency domain, specifically in a straight duct. Subsequently, we assess the network's performance using experimental data.

## **12. Maschinelles Lernen zum Enttauschen hydroakustischer Daten**

Christoph Pilger, Andreas Steinberg, Peter Gaebler und Johanna Lehr  
*Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Fachbereich B4.3: Erbebendienst des Bundes, Kernwaffenteststopp*

Die Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) betreibt als deutscher Repräsentant und nationales Datenzentrum für den Kernwaffenteststopp (CTBT) systematische Studien und langjährige Auswertungen von Seismometer-, Infraschall- und Hydrophon-Daten des Internationalen Monitoring Systems (IMS) für den CTBT.

Dabei werden auch Wasserschall-Daten von Hydroakustikstationen des IMS betrachtet, die zur Registrierung von (nuklearen) Explosionen in Indischem, Pazifischem und Atlantischem Ozean installiert sind. Diese Daten sind, insbesondere im tieffrequenten Bereich um/unter 1Hz, oft stark veräusert durch Brandungs-, Dünungs-, Wind- und Wellenphänomene und erschweren so die Registrierung anthropogener Einzelquellen wie z.B. Explosionen.

Ein Ansatz auf der Basis maschinellen Lernens zum Entrauschen (Denoisen) dieser Hydroakustik-Daten wird vorgestellt, der auf einer Autoencoder-Decoder Architektur und einem Trainingsdatensatz von natürlichen Erdbeben basiert, aber ebenso transferierbar und generalisierbar ist für die Anwendung der Explosionsüberwachung. Insbesondere ergibt sich eine klare Verbesserung von quell-spezifischen Signal-zu-Rausch-Verhältnissen.

### **13. Auswertung von EEG-Daten mit Maschinellern Lernen**

Christian Koch

*PTB*

Daten, die mit Elektroenzephalographiemethoden (EEG) gewonnen werden stellen aus drei Gründen große Herausforderungen an eine Auswertung: Sie verfügen über ein sehr geringes Signal-zu-Rausch-Verhältnis, sind immer nicht-stationär und weisen eine hohe interindividuelle Streuung auf. Im Vortrag wird gezeigt, wie maschinelles Lernen unter Verwendung eines Autoencoders zur Verminderung des Rauschens eingesetzt werden kann. Weiterhin lassen sich neuronale Netze zur Unterstützung der Erkennung von Signalen anwenden. In einem Beispiel sollen akustisch evozierte Potentiale mit Infraschall als Stimulus nachgewiesen werden. In einem zweiten Beispiel geht es um die Erkennung von Lautheitsunterschieden, um die Erkennung von Hörschwellen in Zukunft verbessern zu können.

### **14. Machine-Learning zur Modellierung des akustischen Verhaltens von eBike-Antriebseinheiten**

Annika Hayn

*Robert Bosch GmbH*

Die Nachfrage an eBikes ist in den letzten Jahren stark angestiegen und somit auch die Ansprüche der Nutzer an die Qualität. Dabei zählt das Geräuschverhalten des eBikes als wichtiger Qualitätsparameter. Um den Marktansprüchen gerecht zu werden, befasst sich dieses Forschungsvorhaben mit der Optimierung des akustischen Verhaltens von eBike-Antriebseinheiten aus der Serienproduktion. Ziel des Forschungsvorhabens ist es schon während des Produktionsprozesses eine Aussage über die Wirkung unterschiedlicher Fertigungsqualitäten der Einzelteile hinsichtlich des NVH-Verhaltens im Endprodukt treffen zu können.

Hierbei soll der Einfluss verschiedener Parameter in der Einzelteilerfertigung von Zahnrädern in Bezug auf das NVH-Verhalten im zusammengebauten Endprodukt, der eBike-Antriebseinheit, untersucht werden. Es sollen anschließend Machine-Learning-Ansätze eingesetzt werden, um das NVH-Verhalten auf Grundlage der Fertigungsparametern aus dem Zahnradschleifprozess vorherzusagen.

Dieser Vortrag soll einen Überblick über das geplante Forschungsvorhaben geben. Es werden die Rahmenbedingungen vorgestellt und auf erste Ansätze und Anforderungen an das Vorhaben dargestellt.