

## Luftschallbasierte Rissdetektion von Metallteilen

Judith Liebetrau<sup>1</sup>, Sascha Grollmisch<sup>1</sup>, Johannes Nowak<sup>1</sup>

<sup>1</sup> Fraunhofer IDMT, 98693 Ilmenau, Deutschland, Email: {ltu, goh, noa}@idmt.fraunhofer.de

### Einleitung

Qualitätskontrolle als Teil des Produktionsprozesses dient zur Einhaltung festgelegter Qualitätsanforderungen. Mittels Zerstörungsfreier Prüfverfahren (ZfP) sollen verborgene Fehler in Bauteilen gefunden werden, ohne diese zu beschädigen. Röntgen- und Ultraschalluntersuchungen oder Computertomographie sind sehr zuverlässig in der Erkennungsrate von Fehlern, gleichzeitig aber kostenintensiv und daher nicht immer geeignet. Insbesondere für kleinere Bauteile (Umfang < 5 cm) stehen der Aufwand und Nutzen solcher Verfahren selten in Relation und es wird auf kostengünstigere Verfahren wie beispielsweise Sichtkontrolle durch die Mitarbeiter zurückgegriffen. Subjektiven Faktoren wie Erfahrungsschatz oder Ermüdung des Prüfers beeinflussen jedoch die Genauigkeit der Erkennung. Durch eine Automatisierung bzw. automatische Unterstützung der Prüfer könnten subjektive Einflussfaktoren verringert und somit der Qualitätssicherungsprozess verbessert werden.

Die automatische Rissdetektion auf Basis von Luftschallanalyse ist der Ausgangspunkt für die nachfolgend präsentierten Experimente. Es wird der Frage nachgegangen, ob sich Defekte an kleinen Drehteilen akustisch detektieren lassen und Luftschallanalyse als kostengünstige und automatisierte Vorstufe zur Sichtkontrolle dienen kann. Dazu müssen die Drehteile zum „Klingen“ angeregt, dieser Klang aufgenommen und anschließend analysiert werden, um die Drehteile automatisch den Klassen „in Ordnung“ (i.O.) bzw. „nicht in Ordnung“ (n.i.O.) zuzuordnen. Folgende Forschungsfragen leiten sich für die Experimente ab: Wie kann eine effiziente Klanganregung der Bauteile erfolgen? Unterscheidet sich der Anregungsklang von i.O. zu n.i.O. Teilen? Welche Merkmale können zur Unterscheidung genutzt werden? Welche Parameter beeinflussen die Erkennungsgenauigkeit?

### Vorüberlegung

Die Klopf- oder Klangprüfung als ZfP gilt als kostengünstig und leicht in die Serienfertigung integrierbar. Das Grundprinzip beruht darauf, dass ein Körper in seinen Eigenfrequenzen schwingt, sobald er angeschlagen wird. Diese Resonanzschwingungen lassen sich über Luft- oder Körperschallsensoren aufnehmen und mit eingelernten Mustern vergleichen. Dazu werden Standardmerkmale wie Frequenz, Amplitude oder Dämpfung genutzt und daraus Merkmale abgeleitet, die zur Charakterisierung der Bauteile genutzt werden. Diese klassische Herangehensweise eignet sich für die beschriebene Aufgabenstellung nur bedingt. Die zu untersuchenden Bauteile sind sehr klein, eine mechanische Anregung durch

einen Hammer erscheint als nicht zweckmäßig. Des Weiteren darf es während der Prüfung zu keiner Verunreinigung durch Fremdmetalle kommen. Die Anregung durch einen Gummi- oder Plastikhammer koppelt gegebenenfalls nicht genug Energie ein, um das Bauteil zum Schwingen zu bringen. Neben der Anregung der Bauteile gestalten sich die Merkmalsextraktion und vor allem die Adaptierbarkeit an veränderte Rahmenbedingungen eher schwierig. Um herkömmliche Verfahren nutzen zu können, sind Kenntnisse zum Schwingungsverhalten der Prüflinge erforderlich. Die Auswahl von allen möglicherweise relevanten Güteigenschaften benötigt Expertenwissen.

Das hier vorgestellte System soll den voran genannten Limitierungen der klassischen Klangprüfung begegnen. Die Anregung der Drehteile erfolgt dabei passiv, indem die Bauteile auf eine Platte fallen gelassen werden. Das entstehende Aufprallgeräusch wird zur Analyse verwendet. Die Definition und Extraktion der relevanten Merkmale wird durch den Einsatz von maschinellen Lernverfahren realisiert. Ein explizites Expertenwissen bezüglich des Schwingungsverhaltens der Bauteile ist damit nicht notwendig.

### Grundlagen Maschinelles Lernen

Drei grundsätzliche Herangehensweisen können im Bereich des maschinellen Lernens (ML) unterschieden werden [1]: „Supervised“, „Semi-Supervised“ und „Unsupervised“. Beim Supervised-Ansatz liegen alle herangezogenen Daten vollständig annotiert vor. Das Semi-supervised-Verfahren arbeitet mit teilweise annotierten Datensätzen. Kommen Unsupervised-Ansätze zur Anwendung, liegen keinerlei Datenannotationen vor.

In dem vorliegenden Fall wurden Drehteile mit zugehöriger Kennzeichnung (i.O. und n.i.O.) zur Verfügung gestellt, um eine grundlegende automatische Unterscheidbarkeit zu testen. Deshalb wird Supervised-Ansatz genutzt und die Aufnahmen mit den Drehteilen entsprechend annotiert. Für diesen Ansatz existieren verschiedene ML-Algorithmen, welche je nach Anwendungsfall Vor- und Nachteile mit sich bringen. Insbesondere tiefe neuronale Netze (DNN – deep neural networks) zeichnen sich in vielen Bereichen gegenüber bisherigen Ansätzen aus [9], weshalb diese auch für den vorliegenden Artikel verwendet werden. Die Entwicklung künstlicher neuronaler Netze begann bereits in den 40er und 50er Jahren, etablierten sich erst in den letzten Jahren durch schnellere Hardware und größere Datenmengen im praktischen Einsatz. Inspiriert durch Abläufe im menschlichen Gehirn bestehen DNNs aus mehreren Schichten von Neuronen, welche durch gewichtete Verbindungen zu den vorherigen Schichten aktiviert werden und Signale an die nächste

Schicht weiterleiten. Als Eingangssignal der ersten Schicht werden die Merkmale (für Audiodaten beispielsweise ein Spektrogramm) verwendet und die Ausgabe der letzten Schicht entspricht den vorhergesagten bzw. erlernten Ergebnissen. Durch eine Trainingsphase werden die Gewichte in den Schichten angepasst und schrittweise optimiert. Die Parameterwahl, die Architektur des Netzwerkes sowie die vorhandene Datenmenge und -qualität bestimmt dabei maßgeblich die Klassifikationsgüte [4].

## Experimente

Zwei Experimente wurden durchgeführt. Ziel des ersten Experimentes (Exp1) war es, die prinzipielle Eignung des Testaufbaus für die Erzeugung von Luftschallaufnahmen zur Detektion von Rissen in kleinen Drehteilen zu untersuchen. Dabei stand die Frage im Vordergrund, ob sich die Aufprallgeräusche von i.O. bzw. n.i.O. Bauteilen unterscheiden, so dass diese mittels maschinellen Lernverfahren klassifiziert werden können. Des Weiteren sollte ermittelt werden, ob es Unterschiede in der Klassifikationsgenauigkeit in Abhängigkeit der Mikrofonanordnung und des Klangplattenmaterials gibt. In einem zweiten Experiment (Exp2) wurde die Klassifikationsgenauigkeit der Methode für ein Klangplattenmaterial und Mikrofonanordnung mit mehr Prüflingen untersucht.

### Experimentalaufbau

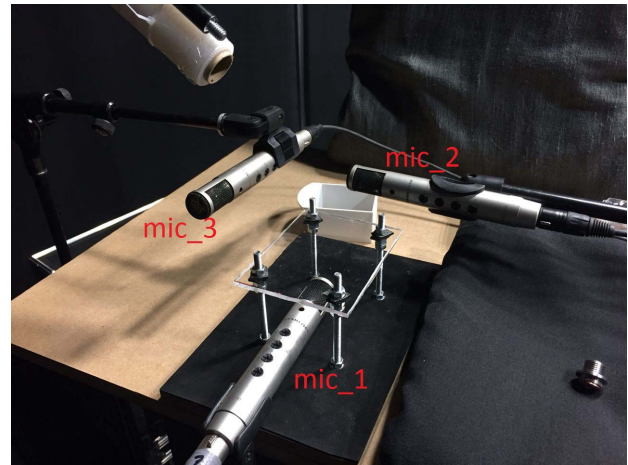
Wie in Abbildung 1 zu erkennen ist, besteht der Aufbau aus einer Fallröhre durch die die Drehteile aus einer definierten Höhe einzeln auf eine Klangplatte fallen. Das Aufprallgeräusch wird durch mehrere Mikrofone gleichzeitig aufgenommen. Dabei befindet sich mic\_1 unter der Klangplatte, mic\_2 seitlich neben der Klangplatte und mic\_3 leicht erhöht hinter der Klangplatte. Als Mikrofone werden 800 P48 mit einer Nierencharakteristik eingesetzt. Die Klangplatte ist austauschbar gestaltet, so dass verschiedene Materialien untersucht werden können. Zum Einsatz kommen mitteldichte Faserplatten (MDF) der Dicken 19 & 4,1 mm, sowie Plexiglas mit Dicken von 3,9 bzw. 1,7 mm.

### Datensatz

Da Exp1 die prinzipielle Machbarkeit der Luftschallanalyse mit dem gewählten Setup überprüfen soll, wurde zunächst ein kleiner Testdatensatz erzeugt. Jeweils 30 Aufnahmen des Aufprallgeräusches von n.i.O. und i.O. Teilen pro Klangplattenmaterial wurden mit einer Abstrategie von 48kHz angefertigt. In Exp2 wurden dann mehr Aufnahmen unter Verwendung der Mikrofon-Material-Kombination angefertigt, die die beste Erkennungsgenauigkeit in Exp1 geliefert hatte. Dabei wurden die Aufprallgeräusche von 223 i.O. Teilen und 166 n.i.O. aufgenommen und analysiert. Die einzelnen Aufprallgeräusche wurden manuell auf eine Länge von 50ms geschnitten und annotiert.

### Klassifikation

Für die Anwendung von DNNs werden die gesamten Daten meist in drei unabhängige Datensätze aufgeteilt:



**Abbildung 1:** Experimentalaufbau zur akustischen Rissdetektion. Die Werkstücke fallen aus einer definierten Höhe auf eine austauschbare Klangplatte. Das Aufprallgeräusch wird von drei Mikrofonen aufgenommen.

Trainings-, Validierungs- und Testdaten. Trainingsdaten dienen zum Lernen und Optimieren der Parameter des DNN. Anhand der ungesehenen Validierungsdaten wird die beste Parameterkombination ausgewählt, was ein Übertrainieren bzw. Auswendiglernen verhindern soll. Mittels der ebenfalls ungesehenen Testdaten wird die finale Klassifikationsgenauigkeit bestimmt. Für das beschriebene System wurden jeweils 60% von i.O. und n.i.O. als Trainings-, 20% als Validierungs- und 20% als Testdaten zufällig ausgewählt. Jede Untersuchung wurde 30 mal wiederholt (30-fache Kreuzvalidierung) und die Daten hierfür immer neu, nach dem Zufallsprinzip, verteilt.

Als Eingangsparameter für das 2-Klassenproblem wurde das gemittelte Magnitudenspektrogramm des Aufprallklanges mit einer Fenstergröße von 512 Samples und einer Sprungweite von 64 Samples verwendet. Andere Fenstergrößen sowie das Aneinanderreihen mehrerer spektraler Frames statt der Mittlung wurden ebenfalls überprüft. Die beschriebene Variante erzeugte jedoch die besten Ergebnisse. Das verwendete DNN umfasst zwei versteckte Schichten (Feed-Forward) mit 128 und 64 Neuronen. Zusätzliche Schichten bzw. Neuronen führten zu keiner Verbesserung der Klassifikationsgenauigkeit. Nach jeder versteckten Schichten wurde Dropout-Schichten mit einer Dropout-Wahrscheinlichkeit von 0.5 eingefügt, welche Übertraining verhindern sollen [8]. Die zufällige Initialisierung der Gewichte erfolgte durch Glorot Initialisierung [3]. Für die versteckten Schichten wurde als Aktivierungsfunktion ReLU (Rectified Linear Unit) [7] genutzt. Zur Beschleunigung der Trainingsphasen kam als Optimierungsalgorithmus Adam [5] mit einer Lernrate von 0.001 zum Einsatz. Categorical Cross Entropy wurde als Kostenfunktion verwendet [6]. Die Implementierung des Systems erfolgte in Python3 unter Nutzung von Keras<sup>1</sup> und Tensorflow<sup>2</sup>.

<sup>1</sup><https://keras.io>

<sup>2</sup><https://www.tensorflow.org>

## Ergebnisse

Als Ergebnis der Klassifikation werden verschiedene Maße ausgegeben. Die mittlere Erkennungsgenauigkeit (acc mean) mit entsprechender Standardabweichung (acc stdev) gibt einen gemittelten Wert über die Ergebnisse der Kreuzvalidierungsläufe an. Ebenfalls wird die höchste und niedrigste Erkennungsrate der Durchläufe angegeben. Eine Angabe der Erkennungsrate alleine reicht für eine Bewertung der Praxistauglichkeit nicht aus, sofern die Erkennungsrate kleiner als 100% ist. Je nach Anwendungsfall ist die Anzahl der nicht erkannten Schlechteile (false positive - FP) oder die Anzahl der detektierten Pseudofehler (false negative - FN) als relevanter Parameter zu untersuchen, wie in Tabelle 1 dargestellt.

**Tabelle 1:** Genereller Aufbau einer Konfusionsmatrix

		Predicted	
		i.O.	n.i.O.
True	i.O.	<i>TP</i>	<i>FN</i>
	n.i.O.	<i>FP</i>	<i>TN</i>

## Experiment 1

In Tabellen 2-4 sind die Ergebnisse des DNN jeweils für vier Klangplatten und die drei Mikrofonpositionen aufgezogen. Die höchsten Erkennungsraten liefern die Audio-

**Tabelle 2:** Erkennungsraten Exp1 für mic\_1

Material	Mikrofon 1 (mic_1)			
	acc mean	acc stdev	min	max
MDF 19 mm	84.4%	12.7%	50%	100%
MDF 4,1 mm	65.6%	15.4%	16.7%	91.7%
Plexiglas 3,9 mm	56.4%	16.1%	25%	91.7%
Plexiglas 1,7 mm	56.7%	13%	33.3%	83.3%

**Tabelle 3:** Erkennungsraten Exp1 für mic\_2

Material	Mikrofon 2 (mic_2)			
	acc mean	acc stdev	min	max
MDF 19 mm	68.6%	15.3%	41.7%	100%
MDF 4,1 mm	65%	14.6%	41.7%	91.7%
Plexiglas 3,9 mm	61.7%	12.9%	33.3%	83.3%
Plexiglas 1,7 mm	54.4%	12.7%	25%	75%

aufnahmen in der Kombination MDF mit 19 mm Dicke (MDF19) und mic\_3 bzw. mic\_1. Gleichzeitig ist zu beobachten, dass Aufnahmen von mic\_3 über alle Materialien

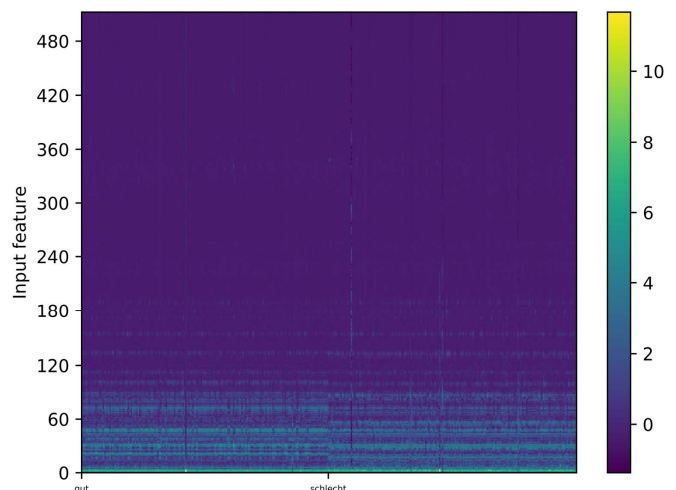
**Tabelle 4:** Erkennungsraten Exp1 für mic\_3

Material	Mikrofon 3 (mic_3)			
	acc mean	acc stdev	min	max
MDF 19 mm	84.2%	13.7%	41.7%	100%
MDF 4,1 mm	73.3%	11.2%	41.7%	91.7%
Plexiglas 3,9 mm	60.6%	13%	33.3%	83.3%
Plexiglas 1,7 mm	53.1%	14%	25%	83.3%

besser klassifiziert werden als die von mic\_1. Die Aufprallgeräusche auf Plexiglas werden durch das DNN deutlich schlechter kategorisiert. Auch die Aufnahmen mit Mikrofon mic\_2 führen zu schlechteren Erkennungsraten. Auf eine Aufstellung der Konfusionsmatrix wurde zu diesem Zeitpunkt verzichtet, da Exp1 zunächst nur die prinzipielle Machbarkeit der Luftschallanalyse zeigen soll. Die automatische Erkennung anhand des Aufprallklanges erscheint möglich, jedoch schwanken die Ergebnisse zwischen den Durchläufen stark. Deshalb wird in Exp2 untersucht, ob sich die Schwankungen zwischen den Kreuzvalidierungsläufen ausgleichen lassen und somit eine robustere Erkennung möglich wird.

## Experiment 2

Für das zweite Experiment wurden neu aufgenommene Aufprallgeräusche auf MDF19 untersucht. In Abbildung 2 ist für 512 Frequenzbänder (Bereich 0 - 24 kHz) die Energieverteilung pro Audioaufnahme aufgetragen. Auf der linken Seite sind die Energiespektren der 233 i.O. Teile dargestellt, auf der rechten Seite die der 166 n.i.O. Teile. Es ist deutlich zu sehen, dass es einen Unterschied in den tiefen und mittleren Frequenzbereichen beider Klassen gibt, jedoch auch innerhalb einer Klasse. Dies bedeutet, dass der Klang zwischen n.i.O. und i.O. Teilen charakteristische Unterschiede aufweist, es sich aber nicht um einen eindeutigen Unterschied handelt. Die Variation



**Abbildung 2:** Energieverteilung i.O. und n.i.O. Teile

innerhalb einer Klasse liegt vermutlich darin begründet, dass durch den Aufbau nur die Fallhöhe vorgegeben ist, die Geschwindigkeit oder Technik mit der die Drehteile in die Röhre geworfen werden jedoch nicht. Dementsprechend trafen die Drehteile nicht immer exakt gleich auf der Klangplatte auf.

**Tabelle 5:** Erkennungsraten Exp2 für mic\_03 MDF 19 mm

Mikrofon	Material MDF 19			
	acc mean	acc stdev	min	max
mic_1	99.6%	0.001%	97.2%	100%
mic_3	98.6%	0.01%	97.2%	100%

Dennoch fallen Erkennungsgenauigkeiten in Tabelle 5 für beide Mikrofone sehr hoch aus. Es zeigt sich, dass die größere Datenmenge zu einer deutlich erhöhten Genauigkeit führt und einzelne Variationen in den Mustern auch erlernt werden. Beide Mikrofonpositionen liefern vergleichbare Ergebnisse. Im besten Durchgang wurde eine Genauigkeit von 100% erzielt, so dass hier nur die Konfusionsmatrix des schlechtesten Durchgangs von mic\_1 dargestellt wird (Tabelle 6). Hierbei wird sicht-

**Tabelle 6:** Konfusionsmatrix für Exp2 mic\_1

		Predicted	
		i.O.	n.i.O.
True	i.O.	94.44%	5.56%
	n.i.O.	0.0%	100.0%

bar, dass einige i.O. Teile als n.i.O. eingeordnet wurden. Dies bedeutet, dass alle defekten Teile im Qualitätssicherungsprozess identifiziert worden wären, aber auch einige Pseudofehler auftreten. Für das finale Ergebnis der Erkennung wurde der höchste Wert der Ausgangsschicht des DNN gewählt, was einem Schwellwert von 0.5 bei zwei Klassen entspricht. Durch eine Verschiebung des Schwellwertes kann die Häufigkeit von FN bzw. FP entsprechend der Anforderungen des Produktionsprozesses angepasst werden.

## Zusammenfassung

Eine luftschallbasierte Erkennung von Rissen in Metallteilen ist mit der verwendeten Apparatur prinzipiell möglich. Eine Abhängigkeit der Erkennungsrate von dem Klangplattenmaterial und der Mikrofonierung wurde gezeigt. Die besten Erkennungsraten lieferten die Aufprallgeräusche auf einer 19 mm dicke MDF-Platte, die durch Mikrofone in leicht erhöhter Position hinter sowie unter Klangplatte aufgenommen worden sind. Für kleine Datenmengen zeigte das verwendete DNN eine Erkennungsrate von durchschnittlich 84.4%. Je nach Aufteilung der Daten wurden sogar alle Drehteile richtig klassifiziert. Es kam zu deutlichen Schwankungen in der Erkennungsrate je nach untersuchter Aufteilung Training und Testdaten. Durch eine erhöhte Datenmenge konnten diese Schwankungen reduziert und eine minimale Genauigkeit von 97.2% erreicht werden.

Der starken Variabilität in den Aufprallgeräuschen kann auf zweierlei Weise begegnet werden. Zum einen könnte ein weitaus größerer Datensatz akquiriert werden, so dass die Vielfalt der Aufprallgeräusche durch das DNN gelernt werden kann. Dies setzt jedoch voraus, dass man dem DNN idealerweise alle möglichen Arten von Aufprallgeräuschen für den Lernprozess zur Verfügung stellt. Dies wiederum bedingt einen sehr gut annotierten Datensatz, der zeitaufwendig erstellt werden muss. Ein praktikablerer Weg ist jedoch die Anpassung des Aufbaues, um noch kontrollierbarere Fallbedingungen zu schaffen. Der subjektive Faktor würde dadurch verringert und die Vorhersagegenauigkeit deutlich stabilisiert werden.

In den Experimenten wurde aktuell nicht untersucht, wie sich Störgeräusche auf die Erkennungsrate auswirken. Es ist zu vermuten, dass übermäßige Hintergrundgeräusche zu einer Verringerung der Erkennungsgenauigkeit führen werden. Diesem kann wiederum durch konstruktive Veränderung (Kapselung) der Fallapparatur begegnet werden oder aber durch verbesserte Signalverarbeitung wie beispielsweise Quellentrennung [2].

## Literatur

- [1] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer-Verlag New York Inc, 1 edition, 2006.
- [2] Estefanía Cano, Johannes Nowak, and Sascha Grollmisch. Exploring sound source separation for acoustic condition monitoring in industrial scenarios. In *2017 24th European Signal Processing Conference (EUSIPCO)*, Kos Island, Greece, 2017.
- [3] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 249–256, 2010.
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- [5] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.
- [6] James D. McCaffrey. Why you should use cross-entropy error instead of classification error or mean squared error for neural network classifier training, November 2013.
- [7] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.
- [8] Nitish Srivastava, Geoffrey E Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1):1929–1958, 2014.
- [9] Y. Lecun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.