

---

# Prozessakustik zur quantitativen Formstoffaufbereitung

---

Von der Fakultät für Ingenieurwissenschaften,  
Abteilung Maschinenbau und Verfahrenstechnik der

Universität Duisburg-Essen

zur Erlangung des akademischen Grades

einer

Doktorin der Ingenieurwissenschaften

(Dr.-Ing.)

genehmigte Dissertation

von

Philine Kerst

aus

Herdecke

**Tag der Disputation:** 29.03.2023

**Erstgutachter:** Prof. Dr. rer. nat. Johannes Gottschling

**Zweitgutachter:** Prof. Dr. rer. nat. habil. Robert Martin

# DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT  
DUISBURG  
ESSEN

*Offen im Denken*

ub | universitäts  
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

**DOI:** 10.17185/duepublico/78194

**URN:** urn:nbn:de:hbz:465-20230428-074338-6



Dieses Werk kann unter einer Creative Commons Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen 4.0 Lizenz (CC BY-NC-SA 4.0) genutzt werden.

*Wenn du das Geheimnis des  
Universums entschlüsseln  
möchtest, musst du anfangen  
in Energie, Frequenz und  
Vibration zu denken.*

– Nikola Tesla [1] –



## Abstract

As a natural product with limited resources, sand is of existential importance for various branches of industry. Foundry technology in particular requires considerable quantities for sand moulds and cores worldwide. However, not all sand is the same. Desert sand, for example, cannot be used for such moulds or cores, as the surface of the grains is too smooth for most applications. Depending on the casting process, the high-quality silica sands are therefore already recycled and thus reused. However, especially in the case of inorganically bonded moulding sand, it is still frequently not recycled but added to the waste system under environmentally relevant conditions. Existing regeneration processes often reach their analytical optimisation limits because the processes are not transparent enough to render inline analyses of the moulding sand possible.

The challenge at hand is that the moulding material quality and the regeneration progress cannot be determined in real time during the process under consideration, so that appropriate optimisations or process adjustments can be carried out during the regeneration process. The running process and the associated current preparation status cannot be measured directly. In the future, it should be possible to determine the moulding sand quality in real time during the process in order to actively control the preparation status directly with corresponding parameter settings. The feasibility of quantitative moulding sand determination during the running process offers the possibility to take into account environmentally relevant goals, such as reduced energy consumption, reduction of CO<sub>2</sub> emissions or resource conservation through more efficient control. Optimised plant control in terms of time, cost and energy efficiency can also be derived from this.

This research presents a methodology using Convolutional Neural Networks (CNN) for machine processing of sound and image data to map and evaluate the non-measurable changes of the moulding material in the running process via the process acoustics in real time. Assuming that the regeneration of the moulding material during the process are acoustically detectable, each process is recorded via a condenser microphone. Changes that occur during the process are extracted and learned using the methods mentioned below.

Audio signal processing offers a variety of analysis methods for feature extraction. The spectrogram is at the forefront of the analyses, because spectrograms visualise the time-frequency content of a signal with the help of the Short-Time Fourier Transform (STFT). The recorded audio data is thus transformed into image data and then analysed with CNNs, in the sense of image recognition. CNNs form a basic element of the methods of so-called deep learning (DL), a subfield of machine learning, and are already being used successfully in image recognition. This results in a combination of signal processing methods and machine learning methods for image processing to detect and classify process changes.



## Kurzfassung

Als Naturprodukt mit begrenzten Ressourcen ist Sand für diverse Industriezweige von existenzieller Bedeutung. Insbesondere die Gießereitechnik benötigt weltweit für Sandformen sowie -kerne beträchtliche Mengen. Dabei ist Sand nicht gleich Sand, sodass beispielsweise Wüstensande nicht für solche Gussformen bzw. -kerne eingesetzt werden kann, da die Oberfläche der Körner für die meisten Anwendungen zu glatt ist. Je nach Gussprozess werden die hochwertigen Quarzsande daher bereits im Kreislauf gefahren und wiederverwendet. Allerdings wird immer noch, insbesondere bei anorganisch gebundenem Formstoff häufig nicht rezykliert, sondern unter umweltrelevanten Auflagen dem Abfallsystem zugeführt. Bestehende dieser Regenerationsverfahren stoßen analytisch häufig an ihre Optimierungsgrenzen, da die Prozesse nicht transparent genug sind, um entsprechende Parameter zur Optimierung analytisch zu beschreiben. Somit sind Inlineanalysen der Formstoffqualität während des Prozesses unmöglich.

Die bestehende Herausforderung ist, dass die Formstoffqualität und der Regenerationsfortschritt nicht während des betrachteten Prozesses bestimmt werden können, um im Laufe des Regenerationsprozesses entsprechende Optimierungen oder Prozessanpassungen durchzuführen. Der aktuelle Aufbereitungszustand ist während der Regeneration prozessbedingt nicht direkt messbar. Die Formstoffqualität soll zukünftig während des Prozesses in Echtzeit bestimmbar sein, um so aktiv den Aufbereitungszustand mit entsprechenden Parametereinstellungen zu steuern. Die quantitative Formstoffbestimmung im laufenden Prozess bietet die Möglichkeit, umweltrelevante Ziele, wie einen verringerten Energiebedarf, Reduzierung des CO<sub>2</sub>-Ausstoßes oder Ressourcenschonung durch eine effizientere Steuerung zu berücksichtigen. Darüber hinaus kann eine optimierte Anlagensteuerung zu verbesserten Prozessfahrweisen hinsichtlich Zeit-, Kosten- sowie Energieeffizienz führen.

Im Rahmen dieser Forschungsarbeit wird eine Methodik mit Hilfe von Convolutional Neural Networks (CNN) zur maschinellen Verarbeitung von Ton- und Bilddaten vorgestellt, die die nicht messbaren Veränderungen des Formstoffes im laufenden Prozess über die Prozessakustik in Echtzeit abbilden und bewerten soll. Unter der Annahme, dass die Veränderungen des Formstoffes während des Prozesses akustisch erfassbar sind, wird jede Regeneration über ein Kondensatormikrofon aufgenommen. Die Veränderungen, die während des Prozesses auftreten, wurden mit den im Folgenden genannten Methoden extrahiert und erlernt.

Die akustische Signalverarbeitung (Audio Signal Processing) bietet eine Vielzahl an Analysemethoden zur Merkmalsextraktion (Feature extraction). Das Spektrogramm steht dabei im Vordergrund der Analysen, da Spektrogramme den Zeit-Frequenz-Inhalt eines Signals mithilfe der Short-Time-Fourier-Transformation (STFT) visualisieren. Die aufgenommenen Audiodaten werden damit zu Bilddaten transformiert und anschließend mit CNNs, im Sinne der Bilderkennung, analysiert. CNNs bilden einen Grundstein der Methoden des sogenannten Deep Learnings (DL), ein Teilgebiet des maschinellen Lernens, und werden bereits erfolgreich in der Bilderkennung eingesetzt. Damit ergibt sich eine Kombination aus Methoden der Signalverarbeitung und maschinellen Lernverfahren zur Bildverarbeitung, um Qualitätsveränderungen des regenerierten Sandes zu erkennen und zu klassifizieren.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Problemstellung und Ziel der Arbeit . . . . .	3
1.3	Aufbau der Arbeit . . . . .	3
<b>2</b>	<b>Theoretische Grundlagen</b>	<b>5</b>
2.1	Gießtechnische Grundlagen . . . . .	5
2.1.1	Formstoff . . . . .	6
2.1.2	Formstoffaufbereitung . . . . .	10
2.1.3	Formstoffqualität . . . . .	15
2.2	Akustische Signalverarbeitung . . . . .	21
2.2.1	Diskrete und Schnelle Fourier-Transformation . . . . .	24
2.2.2	Kurzzeit-Fourier-Transformation . . . . .	25
2.3	Künstliche Intelligenz und Maschinelles Lernen . . . . .	27
2.3.1	Maschinelles Lernen . . . . .	29
2.3.2	Neuronale Netze . . . . .	31
2.3.3	Deep Learning . . . . .	34
2.3.4	Convolutional Neural Networks . . . . .	35
2.3.5	Modellbewertung und - optimierung . . . . .	38
<b>3</b>	<b>Stand der Technik</b>	<b>47</b>
3.1	Regeneration anorganisch gebundener Formstoffe . . . . .	47
3.2	Audioklassifikation mittels CNN . . . . .	51
<b>4</b>	<b>Konzept und Versuchsstand</b>	<b>55</b>
4.1	Regenerationsprozess mittels Reibregenerator . . . . .	55
4.2	Problemstellung und Zielsetzung . . . . .	57
4.3	Lösungskonzept . . . . .	58
4.3.1	Versuchsstand . . . . .	60
4.3.2	Versuchs- und Datenvorbereitung . . . . .	61
4.4	Vorversuche/Vorbereitung . . . . .	63
4.4.1	Reproduzierbarkeit . . . . .	63
4.4.2	Bewertung des Regenerationsprozesses . . . . .	68
4.4.3	Konklusion . . . . .	77

<b>5</b>	<b>Machbarkeitsstudie</b>	<b>79</b>
5.1	Allgemeines Vorgehen . . . . .	80
5.1.1	Generierte Datenbasis . . . . .	80
5.1.2	Erstellung der Spektrogramme . . . . .	82
5.1.3	Augmentationsmethoden zur Datenerweiterung . . . . .	84
5.1.4	Convolutional Neural Network . . . . .	88
5.1.5	Auswertung . . . . .	91
5.2	Neusand vs. Regenerat . . . . .	92
5.2.1	Datenbasis . . . . .	92
5.2.2	Ergebnisse und Diskussion . . . . .	94
5.3	Erkennung von Störgeräuschen . . . . .	103
5.3.1	Datenbasis . . . . .	103
5.3.2	Ergebnisse und Diskussion . . . . .	104
5.4	Verschiedene Sandsorten . . . . .	107
5.4.1	Datenbasis . . . . .	107
5.4.2	Ergebnisse und Diskussion . . . . .	109
<b>6</b>	<b>Fazit und Ausblick</b>	<b>117</b>
6.1	Fazit . . . . .	117
6.2	Ausblick . . . . .	118

# Abbildungsverzeichnis

2.1	Form- und Gießverfahren . . . . .	6
2.2	Wichtige Bindersysteme zur Formherstellung . . . . .	9
2.3	Bindersysteme zur Kernherstellung . . . . .	10
2.4	Prozessverlauf von der Gussform bis zum Regenerat . . . . .	12
2.5	Der Formstoffkreislauf . . . . .	14
2.6	Abstufung der Kornformen . . . . .	16
2.7	Siebturm . . . . .	17
2.8	Siebe des Siebturms . . . . .	17
2.9	Histogramm und Verteilungskurve . . . . .	18
2.10	Siebdurchgang und Summenkurve . . . . .	18
2.11	Formkästen mit Prüfstäben . . . . .	21
2.12	Biegeversuch . . . . .	21
2.13	Kennwerte einer harmonischen Schwingung . . . . .	22
2.14	Analog-Digital-Wandler (i) Abtastung (ii) Quantisierung . . . . .	24
2.15	Fensterfunktionen Hann . . . . .	26
2.16	(i) Zeitbereich (ii) Frequenzbereich (iii) Zeit-Frequenz-Bereich . . . . .	27
2.17	Klassische Programmierung vs. Maschinelles Lernen . . . . .	29
2.18	Künstliches Neuron $j$ . . . . .	32
2.19	Mehrlagiges Perzeptron . . . . .	33
2.20	Teilgebiete der Künstlichen Intelligenz . . . . .	34
2.21	Faltungsschicht . . . . .	36
2.22	Pooling-Operationen . . . . .	37
2.23	Möglicher Aufbau eines CNN . . . . .	38
2.24	Datenaufteilung in Lern- und Testdaten . . . . .	40
2.25	Aktivierungsfunktionen . . . . .	42
2.26	Interpretation und Handlungsempfehlungen für das Modelltraining . . . . .	43
2.27	Augmentation von Audiodaten . . . . .	44
4.1	Der Reibregenerator und seine schematische Darstellung . . . . .	55
4.2	Darstellung des Materialflusses durch den Reibregenerator . . . . .	56
4.3	Reibstäbe . . . . .	56
4.4	Schlagplatten . . . . .	56
4.5	Lösungskonzept . . . . .	59

4.6	Verwendete Sensorik . . . . .	60
4.7	Korrelationsmatrix der Prozessdaten . . . . .	65
4.8	Vorgehensweise Versuchsreihe . . . . .	69
4.9	Korngrößenverteilung des Neusandes (Quarzsand: F34) . . . . .	71
4.10	Korngrößenverteilung der drei Versuchsreihen . . . . .	73
4.11	Festigkeitsanalyse der einzelnen Umläufe pro Regenerationszeit . . . . .	75
4.12	Festigkeitsanalyse pro Regenerationszeit im Vergleich zu Neusand . . . . .	76
4.13	Quarzsand F34 - Skala: 1 mm . . . . .	78
4.14	Quarzsand F34 - Skala: 250 $\mu\text{m}$ . . . . .	78
5.1	Beispiel für die erstellten Spektrogramme . . . . .	83
5.2	Verschiedene Augmentationsmethoden . . . . .	85
5.3	Convolutional Basis . . . . .	90
5.4	Konfusionsmatrix . . . . .	91
5.5	Abweichungen der Spektrogramme . . . . .	93
5.6	Analyse 1: Konfusionsmatrizen <i>Neusand</i> vs. <i>Nicht-Neusand</i> . . . . .	96
5.7	Analyse 2: Konfusionsmatrizen <i>Neusand</i> vs. <i>Nicht-Neusand</i> . . . . .	96
5.8	Auswertung der Test- und Verifizierungsdaten anhand der jeweiligen Konfusionsmatrix . . . . .	98
5.9	Modell 3: Darstellung der Prognose von Modell 3 . . . . .	100
5.10	Prognose von F34 Gebunden V2 . . . . .	102
5.11	Darstellung von Störgeräuschen . . . . .	103
5.12	Konfusionsmatrizen der drei DL-Modelle . . . . .	105
5.13	Auswertung der Prognosegüte . . . . .	107
5.14	Verteilungskurven der verschiedenen Sandsorten . . . . .	109
5.15	Korngrößen im Vergleich (grob, mittel, fein) . . . . .	109
5.16	Konfusionsmatrizen: Auswertung der Testdaten . . . . .	111
5.17	Konfusionsmatrizen: Auswertung der Verifizierungsdaten . . . . .	111
5.18	Konfusionsmatrizen der zwei Modelle . . . . .	113
5.19	Prognosen auf Basis der Verifizierungsdaten . . . . .	113
5.20	Korngrößenverteilung der Sandsorten S01 und S03 . . . . .	114
5.21	Prognosen der Sandsorten S01 und S03 . . . . .	115

# Tabellenverzeichnis

2.1	Eigenschaftsmerkmale zur gießereitechnologischen Bewertung von Formgrundstoffen . . . . .	8
2.2	Multiplikatoren . . . . .	19
3.1	Qualitätskriterien zur Beurteilung des Regenerationseffektes . . . . .	49
3.2	Anwendung von Augmentationsmethoden und Transfer Learning . . . . .	53
4.1	Verwendete Audiodatenlänge in der Forschung . . . . .	63
4.2	Dateiübersicht der Versuchsreihe zur Prozessstabilität . . . . .	67
4.3	Ergebnisse der Korrelationsuntersuchungen . . . . .	67
4.4	Kennwerte der Siebanalyse . . . . .	72
5.1	Vorliegende Datenbasis für Neusand . . . . .	80
5.2	Vorliegende Datenbasis binderbelasteter Sande . . . . .	81
5.3	Vorliegende Datenbasis - Regenerationszeit 5 Minuten . . . . .	81
5.4	Vorliegende Datenbasis - Regenerationszeit 10 Minuten . . . . .	82
5.5	Vorliegende Datenbasis - Regenerationszeit 15 Minuten . . . . .	82
5.6	Aufbau der Convolutional-Basis des CNNs . . . . .	89
5.7	Datenbasis zur Analyse <i>Neusand</i> vs. <i>Regenerat</i> . . . . .	92
5.8	Datensätze zur Verifizierung des Modells . . . . .	94
5.9	Datenbasis für Analyse 1 und 2 . . . . .	95
5.10	Modellkonfigurationen für die Analysen <i>Neusand</i> vs. <i>Nicht-Neusand</i> . . . . .	95
5.11	Auswertung der Prognosen auf den Verifizierungsdaten . . . . .	97
5.12	Angepasste Datenbasis zur Unterscheidung von <i>Neusand</i> und <i>Nicht-Neusand</i> . . . . .	98
5.13	Auswertung der Prognose der Langzeitaufnahmen . . . . .	99
5.14	Angepasste Datenbasis zur Erkennung des Qualitätsstatus des Regenerates . . . . .	100
5.15	Auswertung der Prognose . . . . .	101
5.16	Datenbasis des Modells zur Störgeräuschdetektion . . . . .	104
5.17	Trainingsparameter des Modells zur Störgeräuschdetektion . . . . .	105
5.18	Datenbasis für die Modelle zur Prognose der Sandsorte . . . . .	108
5.19	Modellkonfigurationen für die Analyse verschiedener Sandsorten . . . . .	110
5.20	Parametereinstellungen für die Modelle zur Analyse verschiedener Sandsorten . . . . .	112
5.21	Angepasste Datenbasis zur Unterscheidung von S01 und S03 . . . . .	114



# Abkürzungsverzeichnis

ADW	Analog-Digital Wandler
AFS	American Foundry Society
AI	Artificial Intelligence
AIF	Arbeitsgemeinschaft industrieller Forschungsvereinigungen
BN	Batch Normalization
CNN	Convolutional Neural Network
CV	Cross Validation
DCASE	Detection and Classification of Acoustic Scenes and Events
DFT	Diskrete Fourier-Transformation
DL	Deep Learning
DNN	Deep Neural Network
DOE	Design of Experiments
DSP	Digital Signal Processing
FFT	Fast Fourier-Transformation
FoMaSys	Formstoffmanagementsystem
FT	Fourier-Transformation
GPS	Global Positioning System
GPU	Graphics Processing Unit
KI	Künstliche Intelligenz
KKF	Kreuzkorrelationsfunktion
KMU	Kleine und mittlere Unternehmen
KNN	Künstliches Neuronales Netz

MFCC	Mel Frequency Cepstral Coefficient
ML	Maschinelles Lernen
MLP	Multilayer Perceptron
NN	Neuronales Netz
PCA	Principle Component Analysis
RAG	Rauheitsgrad
ReLU	Rectified Linear Unit
REM	Rasterelektronenmikroskop
RGB	Rot, Grün, Blau
RNN	Recurrent Neural Network
ROI	Return On Investment
SPS	Speicherprogrammierbare Steuerungen
STFT	Short Time Fourier Transformation
VDG	Verein Deutscher Gießereifachleute
VDI	Verband Deutscher Ingenieure
ZIM	Zentrales Innovationsprogramm Mittelstand
ZRE	Zentrum Ressourceneffizienz GmbH

# Symbolverzeichnis

## Kapitel 2:

$a_j$	Aktivierung von Neuron $j$
$A$	Amplitude
$AFS$	Kornfeinheitsnummer
$d_o$	Obere Kornklassengrenze
$d_u$	Untere Kornklassengrenze
$\delta$	Einheitsimpulsfunktion
$f$	Frequenz [Hz]
$f_A$	Abtastfrequenz
$f_{act}$	Aktivierungsfunktion
$f_{out}$	Ausgabefunktion
$g$	Gesamtkornmasse
$g_i$	Massenanteil der Kornklasse $i$
$GG$	Gleichmäßigkeitsgrad
$j$	Imaginäre Einheit
$k$	Anzahl Folds (Kreuzvalidierung)
$\max$	Maximum
$M1_i$	Multiplikator $S_{th}$ für Kornklasse $i$
$M2_i$	Multiplikator $MK_{clac}$ für Kornklasse $i$
$M3_i$	Multiplikator $ASF$ für Kornklasse $i$
$MK_{clac}$	Kalkulierte mittlere Korngröße
$n$	Zahlenindex

$N$	Anzahl der betrachteten Abtastwerte
$net_j$	Netzeingabe von Neuron $j$
$o_k$	Eingabeinformation von Neuron $k$
$o_j$	Eingabeinformation von Neuron $j$
$\varphi$	Phasenwinkel
$\varphi_0$	Nullphasenwinkel
$\pi$	Pi - Kreiszahl
$Q_1$	1. Quartil
$Q_3$	3. Quartil
$s$	Variable
$S$	Sigmoidfunktion
$So$	Sortierung
$S_{th}$	Theoretische spezifische Oberfläche
$t$	Zeitvariable
$t_A$	Abtastzeit
$T$	Periodendauer $\left[\frac{1}{\text{Hz}}\right]$
$T_0$	Beobachtungszeit
$\tanh$	Tangens hyperbolicus
$\theta_j$	Schwellenwert von Neuron $j$
$w_{kj}$	Gewichtete Verbindung von Neuron $k$ zu $j$
$\omega$	Kreisfrequenz
$x(t)$	Zeitkontinuierliches Signal
$x_n$	Zeitdiskretes Signal
$y(t)$	Zeitkontinuierliches Signal
$y_n$	Zeitdiskretes Signal
$Y_*(t)$	Spektrum des zeitdiskreten Signals $y_n$

### Kapitel 3:

$h$	Mel Frequenz
-----	--------------

**Kapitel 4:**

$K_n$	Konzentration des $n$ -ten Umlaufes
$\kappa$	Diskrete Laufvariable
$r$	Korrelationskoeffizient
$\tau$	Kontinuierliche Laufvariable
$x_i$	Stichprobenwert Zufallsvariable X
$\bar{x}$	Mittelwert einer Stichprobe zur Zufallsvariable X
$y_i$	Stichprobenwert Zufallsvariable Y
$\bar{y}$	Mittelwert einer Stichprobe zur Zufallsvariable Y



# Kapitel 1

## Einleitung

### 1.1 Motivation

Die „Potenziale der schwachen künstlichen Intelligenz für die betriebliche Ressourceneffizienz“ wurden 2021 im Rahmen einer Studie der *VDI Zentrum Ressourceneffizienz GmbH* (VDI ZRE) untersucht. Gerade kleine und mittlere Unternehmen (KMU), zu denen viele deutsche Gießereibetriebe zählen, stehen vor der großen Herausforderung neue Technologien zu adaptieren, denn die verfügbaren personellen, wirtschaftlichen, zeitlichen und auch natürlichen Ressourcen sind begrenzt und der Spielraum zum Experimentieren klein. Insbesondere die Einführung von künstlicher Intelligenz (KI) scheitert oft an der Verfügbarkeit der notwendigen digitalen Datenbasis. Dennoch ist die Steigerung der Ressourceneffizienz ein zentraler Hebel, um einen positiven Einfluss auf die Umwelt zu generieren und gleichzeitig Kosten zu senken. Die Politik, Kunden oder Investoren legen hierauf bereits vermehrt den Fokus. Zur Verbesserung der unternehmerischen Nachhaltigkeitsperformance werden effiziente Strategien, hauptsächlich im Bereich der Klimaneutralität gefordert. [27] Im Rahmen des *European Green Deals* Ende 2019 wurde als Ziel für die gesamte EU die Klimaneutralität bis 2050 verkündet, welche im selben Jahr auch vom deutschen Bundestag im Klimaschutzgesetz verankert wurde [104]. Die Notwendigkeit der Weiterentwicklungen der Fertigungsindustrie in Bezug auf Nachhaltigkeitsthemen wird auch durch die 2021 erstmals prämierte Kategorie „Sustainability“ des *Microsoft Intelligent Manufacturing Awards*<sup>1</sup> deutlich. Gewinner dieser Kategorie im Jahr 2021 war die Wienerberger AG, die sich mit der datengetriebenen Prozessoptimierung der Ziegelproduktion beworben hatte und, bei gleichbleibender Qualität, Energie- und Emissionsreduktionen erzielen konnte [131].

Gerade Gießereien gehören im industriellen Sektor zu den Branchen, in denen große Mengen an Ressourcen, wie natürliche Rohstoffe, Wasser und Energie, verbraucht werden. Zu diesen Rohstoffen zählen Form- und Kernsande sowie entsprechende Zusatzstoffe, die massenhaft für die Herstellung von Gussstücken notwendig sind [118]. Durch den hohen Einsatzgrad dieser natürlichen Ressourcen hat die Gießereipraxis erhebliche Auswirkungen auf die Umwelt. Bei Gussprozessen fallen zudem beträchtliche Mengen an festen und flüssigen Abfällen sowie giftige Prozessgase an, die in die Umwelt gelangen und diese negativ belasten [20]. Bezogen auf umwelttechnische Aspekte und die oben genannten Potenziale der betrieblichen Ressourceneffizienz durch künstliche Intelligenz sind zwei Komponenten für die vorliegende Arbeit entscheidend: Der Form- bzw. Kernsand und die Art des verwendeten Bindersystems. Die Wahl des Bindersystems ist

---

<sup>1</sup><https://www.rolandberger.com/de/Insights/Publications/Die-Gewinner-des-Microsoft-Intelligent-Manufacturing-Award-2021.html> (besucht am 14.06.2022)

eine entscheidende Variable für die Verringerung der Abgasemissionen [51]. Anorganische Systeme setzten im Gegensatz zu organischen Bindern bei der Oxidation durch den Abguss keine giftigen Gase frei, sondern lediglich Wasserdampf [101]. Sie sind daher weniger schädlich für die Umwelt und die menschliche Gesundheit. Des Weiteren zeichnen sich diese Systeme durch ihre niedrigen Kosten aus [118]. Das steigende Interesse an anorganischen Alternativen liegt zusätzlich im Anstieg der Deponiekosten für organische kunstharzgebundene Gießerei-Altsande. Dieser Anstieg ist auf die immer strenger werdenden Umweltvorschriften zurückzuführen. Anorganische Systeme sind dennoch nicht immer eine optimale Lösung, denn abhängig vom Gusswerkstoff sind sie nur bedingt an Stelle organischer Binder einsetzbar. [39, S. 58]

Sandgießen ist eines der am häufigsten verwendeten Verfahren in der Fertigungsindustrie, da es Vorteile in Bezug auf Kosten, Leistung und Qualität bietet [63]. Für die Fertigung einer Tonne Gussteile werden durchschnittlich etwa 8 bis 15 Tonnen Sand benötigt; während des Formprozesses und des Reinigungsprozesses des gegossenen Werkstücks werden bis zu 1,5 Tonnen Abfallsand produziert [113]. Dabei sind die qualitativen Unterschiede verschiedener Sandarten zu berücksichtigen. Wüstensand beispielsweise hat für die meisten Anwendungen eine zu glatte Oberfläche und kann daher schlecht für Gussformen bzw. -kerne eingesetzt werden. Zudem können vorhandene Salze und organische Verbindungen beim Einsatz solcher Sande zu unkontrollierten Reaktionen führen. Daher sind qualitativ hochwertige Sande, die für den industriellen Einsatz geeignet sind, zu einer wertvollen Ware geworden. [6] Durch diesen hohen Stellenwert des Rohstoffes Sand und des Fakts, dass durch Form- und Gussprozesse belastete Altsande einen Großteil der Gießereiabfälle ausmachen, ist die Wiederverwendung bzw. vollständige Rückgewinnung von Sanden einer der wichtigsten Entwicklungstrends der Gießereiindustrie [113]. Dabei dient wiedergewonnener Sand als Kreislaufprodukt, welches nach einer Aufbereitung erneut zur Herstellung von Kernen bzw. Formen für die Einzelfertigung sowie für Großserien eingesetzt werden kann [39, S. 12]. Im Hinblick auf den Anstieg der Abfallentsorgungskosten, die Schonung natürlicher Ressourcen sowie die Einhaltung der Umweltgesetze und -vorschriften stehen Gießereien aktuell vor vielfältigen zentralen Herausforderungen [63].

Neben metallurgischen Weiterentwicklungen sowie Prozess- und Produktoptimierungen liegt der Fokus kontinuierlicher Innovationen der Gießereibranche aktuell auf den stetig wachsenden Möglichkeiten der Digitalisierung [12, S. 7]. Im Rahmen der Digitalisierung sowie letztendlich auch durch die pandemische Lage wurden mehr und mehr Daten produziert. Dabei ist die gesamte Gesellschaft, aber auch die Wirtschaft, nicht nur Produzent, sondern gleichzeitig Verbraucher dieser Daten. [2, S. 1] Aufgrund der Existenz großer Datenmengen werden im Bereich des maschinellen Lernens, insbesondere durch künstliche neuronale Netze, derzeit erhebliche Fortschritte erzielt. Die Fortschritte und Neuentwicklungen zeigen sich beispielsweise in der Qualität automatischer Bilderkennungssysteme, die wiederum Auswirkungen auf eine Reihe weiterer Anwendungsfelder haben, etwa für selbstfahrende Autos und autonom agierende Drohnen oder im medizinischen Bereich für automatische Diagnosen von Krankheiten anhand bildgebender Verfahren wie Computertomografie, Ultraschall oder Magnetresonanztomografie. [133, S. 2] Künstliche neuronale Netze, die für Bild- und Audioklassifizierung sowie die Objekterkennung eingesetzt werden, fallen in den Teilbereich *Deep Learning* (DL) des maschinellen Lernens. Große Datenmengen sind hier notwendig, um leistungsfähige Netze erzeugen zu können [80]. Dabei wird zumeist versucht, den Erkenntnisgewinn des Menschen nachzuahmen. Insbesondere die Sinne des menschlichen Sehens und Hörens werden adaptiert. Im Fall der Audioklassifizierung wird sich der Adaption beider Sinne bedient, indem aus den Rohaudiodaten audiobasierte Bilder erstellt werden, die wiederum als Bildklassifikation analysiert werden. Die Relevanz von Geräuschemission ist nicht zu unterschätzen und wird bereits zur Leistungsbeurteilung von Maschinen eingesetzt. Das Hören ist für den Menschen einer der relevantesten Sinne: Neben einem hohen Einfluss auf

die Persönlichkeitsbildung dient es dem Menschen auch als empfindliches Warninstrument und ermöglicht die frühzeitige Erkennung von Gefahrensituationen. Dafür kann es sogar die Schlafphase unterbrechen, da Menschen auch im Schlaf hören. [116, S. 204]

## 1.2 Problemstellung und Ziel der Arbeit

Eines der bereits genannten und größten Probleme von Sandgießereien sind die hohen Abfallmengen, die durch die Sandformen entstehen. Durch die verschärften Umweltvorschriften konzentrieren sich viele der betroffenen Unternehmen daher auf einen nachhaltigen Gießprozess. [91] Die vorliegende Dissertation geht aus dem Forschungsprojekt SandREG (ZIM-Projekt, Förderkennzeichen: ZF4047818CM9) hervor, welches sich mit einer Lösung zu der genannten Problematik befasst: Der Fokus liegt auf einem mechanischen Regenerationsprozess von anorganisch gebundenem Formstoff mit einem Fließbettreiniger. Während des laufenden Prozesses können die Formstoffqualität und der Regenerationsfortschritt nicht bestimmt werden, wodurch im Laufe des Regenerationsprozesses keine entsprechenden Optimierungen oder Prozessanpassungen durchführbar sind. Der untersuchte Regenerationsprozess findet in einem geschlossenen System statt, wodurch der Qualitätszustand des Regenerates während des Prozesses mit dieser Anlage bisher nicht erfassbar ist. Aus Sicherheitsgründen kann die Anlage nur mit verschlossener Tür gestartet werden. Darüber hinaus ist es nicht möglich, ausreichende Probenmengen während des Betriebs, durch eine kleine Öffnung im Anlagenkorpus zu entnehmen, um signifikante Qualitätsanalysen durchzuführen. Der laufende Prozess und der damit verbundene aktuelle Aufbereitungszustand sind daher prozessbedingt nicht direkt messbar. Der Zweck eines jeden Regenerationsverfahrens ist, gebundene Sande wieder in ein Material umzuwandeln, dessen Eigenschaften denen von Neusand ähnlich sind.

Das übergeordnete Ziel dieser Dissertation ist, die oben genannten Potenziale der KI zur Förderung der Ressourceneffizienz im gießereitechnischen Kontext zu nutzen, um eine Lösung für die Prozessoptimierung hinsichtlich eines erfassbaren Formstoffqualitätsstatus zu erreichen. Wie bereits erwähnt, ist das Ziel des Regenerationsprozesses, einen belasteten Sand so aufzubereiten, dass neusandähnliche Eigenschaften zu erwarten sind. Übertragen auf die vorliegende Problemstellung soll im Verlauf der Arbeit das Prinzip der Audioklassifikation mit Hilfe KI-gestützter Modelle angewendet werden, um die Prozessfortschritte der Regeneration beurteilen zu können. Es wird geprüft, ob mit Hilfe des akustischen Ansatzes die Möglichkeit besteht, diesen intransparenten Prozess mittels seiner Prozessakustik analysierbar zu machen. Anhand des aufgestellten Konzeptes soll analysiert werden, inwieweit sich Formstoff durch den Prozessverlauf den akustischen Eigenschaften von Neusand wieder annähert. Dafür wird unbehandelter Sand, also Neusand, regeneriert und das zugehörige akustische Signal aufgenommen, um die Audiodaten des Neusandes als Lernbasis für ein Convolutional Neural Network (CNN) zu verwenden. Das Ergebnis dieser Analysen soll ermöglichen, den Regenerationsprozess von binderbehaftetem Sand so lange akustisch auszuwerten, bis die Prozessakustik dem gelernten Akustikprofil von Neusand ähnlich genug ist, um den Prozess zu stoppen.

## 1.3 Aufbau der Arbeit

Nachdem die Motivation sowie die Problemstellung inklusive Zielsetzung dieser Arbeit erläutert wurden, folgen in Kapitel 2 die zugehörigen theoretischen Grundlagen, die bereits aus dem vorangegangenen Abschnitt hervorgehen. Die Grundlagen sind in drei Themenbereiche aufgeteilt: Zunächst wird ein Überblick über die notwendigen gießereitechnischen Grundlagen gegeben, mit besonderem Fokus auf die Regeneration von

Formstoffen. Darauf folgt die Theorie der Signalanalyse, die sich im speziellen auf die Hintergründe und die Erstellung von Spektrogrammen fokussiert. Der abschließende Themenbereich befasst sich mit einer Einführung in die Thematik der künstlichen Intelligenz, wobei sich das Hauptaugenmerk auf den Aufbau und die Funktionsweise der Convolutional Neural Networks im Rahmen des Deep Learnings richtet.

Das 3. Kapitel befasst sich mit dem aktuellen Stand der Technik. Hier wird insbesondere Bezug auf den aktuellen Status der umwelttechnischen Relevanz der Regenerationstechnik eingegangen. Darauf aufbauend liegt der Fokus auf verschiedenen Forschungen zur Regeneration anorganisch gebundener Sandssysteme. Damit wird verstärkt auf den sinnvollen Einsatz von KI im Bereich der Formstoffsteuerung eingegangen und die vorliegende Arbeit entsprechend zum aktuellen Stand der Technik abgrenzt. Des Weiteren wird neben dem industriellen auch ein allgemeiner Fokus auf den Einsatz von Audioklassifizierungen gelegt.

Im 4. Kapitel „Konzept und Versuchsstand“ werden der verwendete Reibregenerator, die Problemstellung und die Zielsetzung sowie das entsprechende Lösungskonzept erläutert. Zusätzlich fließen wichtige Voruntersuchungen mit entsprechender Auswertung in dieses Kapitel ein, die eine Basis für das darauf folgende 5. Kapitel bilden.

Zur Durchführung des aufgestellten Lösungskonzeptes wird in Kapitel 5 die Umsetzung der zugehörigen Machbarkeitsstudie beschrieben. Insgesamt liegen drei Aspekte vor, die im Hinblick einer kontinuierlichen und ganzheitlichen Prozessüberwachung relevant sind und die differenziert untersucht werden. Die zugrundeliegende Datenbasis, die durchgeführten Analysen sowie eine entsprechende Auswertung inklusive Diskussion werden in diesem Abschnitt dargelegt.

Abschließend werden im 6. Kapitel die wesentlichen Ergebnisse hinsichtlich der aufgeführten Problemstellung sowie Zielsetzung zusammenfassend dargestellt. Mit einem Ausblick auf anknüpfende Forschungsthemen wird die vorliegende Dissertation beendet.

# Kapitel 2

## Theoretische Grundlagen

Der theoretische Hintergrund des betrachteten Ansatzes wird vor allem durch drei Themenfelder gebildet: Die gießereitechnischen Grundlagen, die akustische Signalanalyse und die theoretischen Grundlagen der künstlichen Intelligenz mit Fokus auf die verwendete maschinelle Lernmethode der Convolutional Neural Networks (CNN).

### 2.1 Gießertechnische Grundlagen

Die Gießertechnik umfasst unter anderem eine Vielzahl komplexer Prozesse sowie einzelner Prozessschritte, Materialien bzw. Materialzusammensetzungen und entsprechende Wirkzusammenhänge. Gussstücke werden als kleinste oder größte Bauteile realisiert und in unzähligen Bereichen, von der Luftfahrt über die Architektur und Bauindustrie bis zu medizinischen Bereichen, eingesetzt. Mitunter handelt es sich um eine enorm große und komplexe Branche. Um einen allgemeinen Überblick über die Bandbreite der Gießereiindustrie zu erhalten, der im Rahmen dieser Arbeit nicht behandelt werden kann, ist der Verein Deutscher Gießereifachleute (VDG) e.V.<sup>2</sup> eine gute Anlaufstelle.

Der Gießprozess ist ein sehr altes Verfahren, das bis 3500 v. Chr. nachgewiesen werden kann [109, S. 12]. Im Wesentlichen entstehen beim Gießen aus flüssigen, breiigen oder pastösen Metallen, Keramik- oder Kunststoffmassen sowie anderen Werkstoffen geometrisch definierte Gussteile mit bestimmten Eigenschaften. Gießen gehört zu den Fertigungsverfahren, welche in der Norm DIN ISO 8580 definiert werden. Fertigungsverfahren dienen im allgemeinen zur Herstellung von geometrisch bestimmten festen Körpern, einschließlich der Verfahren zur Gewinnung erster Formen aus formlosem Stoff, Verfahren zur Veränderung der Form sowie zur Veränderung der Stoffeigenschaften. Untergliedert werden sechs Verfahrenshauptgruppen: *Urformen*, *Umformen*, *Trennen*, *Fügen*, *Beschichten* und *Stoffeigenschaften ändern*. [39, 12, 109] Der Gießprozess wird der Verfahrenshauptgruppe *Urformen* zugeordnet. *Urformen* ist als formschaffendes Fertigungsverfahren eines festen Körpers aus formlosem Stoff durch Schaffen des Zusammenhaltes definiert [12, S. 10].

Die klassische Aufteilung von Form- und Gießverfahren wird zunächst in die Herstellung von Formteilen und Halbzeugen wie Bleche, oder Rohre unterschieden, wie das Schaubild 2.1 darstellt. Die bei einem Gießprozess entstehenden Formteile sind Rohteile und können durch Weiterverarbeitungen, wie spanende oder umformende Fertigungsverfahren, zu Fertigteilen mit hohen Genauigkeiten produziert werden. Es ist

---

<sup>2</sup><https://www.vdg.de/home> (besucht am 23.03.2022)

ebenfalls möglich, Werkstücke direkt in ihre endgültige Form zu gießen und sie damit verlustarm herzustellen. Eine effiziente Teilefertigung sollte daher so ausgelegt werden, dass die Rohteile unter Reduzierung der Anzahl notwendiger Fertigungsschritte konstruiert und möglichst nur an den Funktionsflächen spanlos oder spanend bearbeitet werden. [39, S. 8]

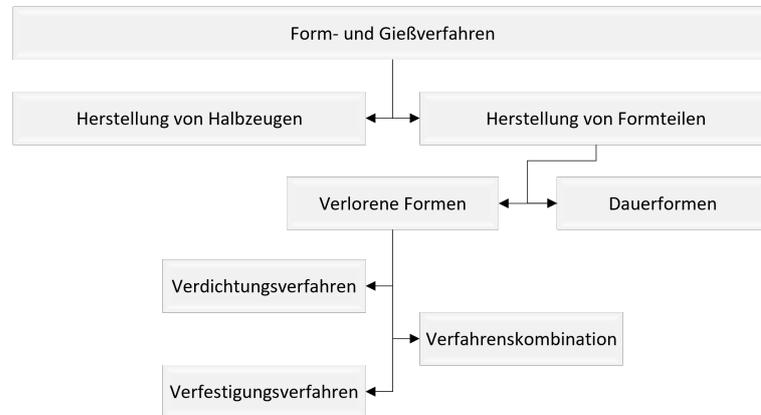


Abbildung 2.1: Form- und Gießverfahren (vgl. [12, S. 181])

Die Herstellung von Formteilen wird durch Verfahren mit verlorenen Formen oder mit Dauerformen realisiert. Dauerformen sind Formen, die für den dauerhaften Gießprozess ausgelegt und genutzt werden. Verlorene Formen hingegen entstehen durch Verdichtungs-, Verfestigungs- oder kombinierte Verfahren für jeden Abguss neu. Der direkte Wiedereinsatz ist im Gegensatz zu den Dauerformen nicht möglich, denn zum Entpacken des gegossenen Gussteils müssen die Gussformen zerstört werden, gehen also verloren. Verlorene Formen bestehen aus Formstoff, welcher sich aus mineralischen, feuerfesten, körnigen Grundstoffen, Bindemitteln sowie weiteren möglichen Zusatzstoffen zusammensetzt. Die Produktion von solchen Sandformen für den Gießprozess gehört ebenfalls zu der Verfahrenshauptgruppe *Urformen*, nämlich zur Untergruppe *Urformen aus dem körnigen oder pulverförmigen Zustand*. [12, S. 10], [39, S. 57] Sandformen bzw. Formstoff sind Hauptbestandteil des Sandgusses und der vorliegenden Arbeit, weshalb im Folgenden näher auf die genannten Bestandteile von verlorenen Formen eingegangen wird.

### 2.1.1 Formstoff

*Formstoff*, als Hauptbestandteil verllorener Formen, steht als Sammelbegriff für alle trockenen, feuchten, pastösen oder flüssigen Stoffe, die zur Herstellung von Gießformen verwendet werden können. Neben dem Formprozess von verlorenen Gussformen wird er auch für die Produktion von Kernen eingesetzt, die z.B. für Hohlräume in Gussteilen genutzt werden [10, S. 366].

Formstoffe unterscheiden sich hinsichtlich ihres Aufbaus, bei dem der Anteil des Formgrundstoffes als Hauptbestandteil zwischen 85 % und 98 % variieren kann, Formstoffbinder sowie mögliche Zusatzstoffe machen den restlichen Anteil von 2 % bis 15 % aus [35, S. 37]. Bezogen auf die eingesetzte Bindermenge bewegt sich der notwendige Formstoffhärteranteil zwischen 10 % und 50 %. Mögliche Zusatzstoffe liegen bei einem Anteil zwischen 0 % und 7 % [12, S. 193]. Als Formgrundstoff können beispielsweise Quarzsand, Korund, Zirkonoxid- oder Zirkonsilikatsand, als Bindemittel Wasserglas, Kieselsäure, oder Äthylsilikat eingesetzt werden [109, S. 34]. Durch die vielen bestehenden Möglichkeiten Formstoff herzustellen, werden Formstoffe abhängig von dem gewählten Bindersystem eingeteilt in [36, S. 246–247]:

- bentonitgebundene Formstoffe (auch Benton-Öl-gebundene Formstoffe),
- anorganische Formstoffbindersysteme mit z.B. Silikatbinder (Wasserglas), Zement, Kieselolen, Gips oder Salzbindersystemen und
- organische Formstoffbindersysteme mit z.B. Furanharzen, Phenolharzen oder Phenol-Resolharzen.

### Formgrundstoffe

Unter *Formgrundstoffe* werden Sande verstanden. Dabei handelt es sich sowohl um natürliche, also mineralische Sande, als auch um synthetische Sande bzw. Granulate. Bei einer Ansammlung vieler Partikel wird häufig auch von einem Kollektiv gesprochen. Bezogen auf unterschiedliche Anwendungsbereiche treten diese Kollektive auch unter den Synonymen Mischkomponenten, Haufwerke, Schüttgüter, Staubwolken, Suspensionen oder Schlämme auf. Dabei liegen Schüttgüter wie Sand, Zucker, Kohle, Kunststoffgranulat oder Tabletten in einem Partikelgrößenbereich von  $10 \cdot 10^{-6}$  bis  $10 \cdot 10^{-2}$  m [121, S. 3]. Mineralische Sande definieren sich über den Korndurchmesser des Kornhaufwerkes, welcher in einem Größenbereich von 0,02 und 2 mm liegt [12, S. 193], [35, S. 37], [47, S. 31]. Besteht ein Anteil eines Kornhaufwerkes aus Partikeln mit einem Korndurchmesser  $\leq 0,02$  mm, wird dieser Anteil als *Schlammstoffanteil* bezeichnet. Mit diesen Grenzwerten lässt sich Formgrundstoff massenanteilig in zwei Bereiche einteilen [129, S. 2]. In der Gießereiindustrie weisen Quarzrohstoffe einen hohen Einsatzgrad auf und stellen damit ein wichtiges Ausgangsmaterial dar [28, S. 11]. Allen durchgeführten Versuchen dieser Arbeit liegt Quarzsand als Basismaterial zu Grunde, sodass nachfolgend besonders auf Quarz als eingesetzter Formgrundstoff eingegangen wird.

Quarz ist ein mineralischer Rohstoff, gehört in Form von Sanden und Kies zu den Lockersedimenten und wird den Eigenschaftsrohstoffen zugeordnet [47, S. 31]. *Kristallines Siliziumoxid* ( $\text{SiO}_2$ ) ist der mineralogische Name für Quarz [29, S. 4], wobei Sandkörner, die mindestens einen  $\text{SiO}_2$ -Gehalt von 85-Massen-% aufweisen, zu den monomineralischen Sanden zählen [47, S. 32].

In Deutschlands Rohstoffvorkommen treten verbreitet Lagerstätten mit hochwertigen Quarzsanden auf. Sie liefern den Ausgangsstoff für eine Vielzahl an Baustoffen sowie für die Produktion von Industriegütern. Für den Einsatz in industriellen Bereichen steigt im Allgemeinen bei steigendem  $\text{SiO}_2$ -Gehalt auch die Qualität des Rohstoffes. Aufgrund seiner hohen Härte und weitestgehenden chemischen Beständigkeit ist Quarz ein sehr attraktiver und hochwertiger Rohstoff, an den hohe Anforderungen gestellt werden. [47, S. 33] Bei *Industriesanden* wird von Sanden mit einem  $\text{SiO}_2$ -Gehalt von  $>98$ -Massenprozent ausgegangen. Die Mineralzusammensetzung und damit die Qualität ist von der Lagerstätte abhängig. Umwelttechnische Einflüsse haben diese Mineralzusammensetzung über Jahrhunderte hinweg beeinflusst. [47, S. 32] Der sogenannte *Stoffbestand* liefert diesbezüglich einen wichtigen Kennwert, denn er gibt Informationen über die mineralogische und chemische Zusammensetzung von Sanden und Kies preis, sodass industrielle Anwender die Einsatzfähigkeit prüfen und Rückschlüsse zu den Lagergebieten ziehen können [47, S. 33]. Eine Auflistung von deutschen Lagerstätten mit entsprechenden Zusatzinformationen kann in „Quarzrohstoffe in Deutschland“ [29] nachgeschlagen werden.

Die Eignung von Quarzsanden in der Gießereiindustrie wird, wie bereits erwähnt, über den  $\text{SiO}_2$ -Gehalt definiert und je höher dieser Gehalt ist, desto weniger Beimengungen sind notwendig. Bei steigender Reinheit des Quarzsandes (Richtwerte liegen zwischen 94 % und  $\geq 99$  %), verhält sich das Material günstiger gegenüber thermischen Belastungen. [29, S. 15], [35, S. 42] Zur Aufbereitung sollte der Grubensand gewaschen und klassiert werden. Dabei liegt die Körnung in einem Bereich von 0,1 bis 0,5 mm, wobei diese

in vier Stufen (grobe, mittlere, feinere und feinste Körnung) unterteilt wird. Mit einer glatten Kornoberfläche, einer weiß bis gelblich-grauen Farbe und einem runden Quarzkorn entspricht der Sand dem gewünschten optischen Qualitätskriterium. Quarzsand zeichnet sich zudem durch einen nicht vorhandenen Kalkgehalt und möglichst geringem Schlammstoffanteil aus. [35, S. 42] Die häufig zitierte Aufstellung von Flemming und Tilch (1993) über die Eigenschaftsmerkmale zur gießereitechnologischen Bewertung von Formgrundstoffen ist auch für die vorliegende Arbeit von Bedeutung und wird in Tabelle 2.1 wiedergegeben. Quarzsande bieten den Vorteil, dass sie ohne große Einschränkungen für fast alle Bindersysteme genutzt werden können. Wenn es sich hingegen um thermisch hochbeanspruchte Formen bzw. Kerne handelt, werden bevorzugt synthetische Formgrundstoffe eingesetzt. [35, S. 37–38]

Tabelle 2.1: Eigenschaftsmerkmale zur gießereitechnologischen Bewertung von Formgrundstoffen [35, S. 38]

<b>Eigenschaftskomplex</b>	<b>Einzelmerkmal</b>
Chemische und mineralogische Zusammensetzung	Chemische Zusammensetzung z.B. SiO <sub>2</sub> -Gehalt, Anteil der Basiskomponente(n), weitere Bestandteile
	Realstruktur der Körner
	Chemismus der Kornsubstanz bzw. Kornoberfläche
Granulometrische und morphologische Eigenschaften	Korngrößenanteile, mittlere Korngröße, Feinkornanteil und Korngrößenverteilung des Haufwerkes
	Kornform
	Struktur der Kornoberfläche repräsentativer Körner
	„Aktivität“ der Kornoberfläche repräsentativer Körner
Physikalisch-technologische Eigenschaften	Zerkleinerungsverhalten des Kornhaufwerkes – Stabilitätsverhalten (Neigung zum Zerspringen einzelner Körner bei mechanischer und thermischer Beanspruchung)
	Härte repräsentativer Körner
	Dichte repräsentativer Körner
	Ausdehnung des Kornhaufwerkes
	Sinterverhalten des Kornhaufwerkes
	Thermische Kennwerte, z.B. Wärmeleitfähigkeit

### Formstoffbinder und -härter

Formstoffbinder wird zur Herstellung von Formstoffmischungen für Sandformen und Kerne verwendet. Im Allgemeinen verleiht das Bindersystem einem Kornhaufwerk durch physikalisch-chemische oder chemische Vorgänge die gewünschte Festigkeit und damit auch eine ausreichende Formstabilität für den Gießprozess [39, S. 58]. Mit einem prozentualen Anteil zwischen 2 % und 10 % des Formstoffsystems ist Formstoffbinder ein notwendiger Bestandteil zur Verfestigung. Das Bindersystem wird in einem Mischprozess gleichmäßig auf der Kornoberfläche aufgetragen, sodass eine optimale *Binderhülle* um das Korn entstehen kann. Wenn der angemischte Formstoff verdichtet wird, sorgt die Binderhülle an den vielen entstehenden Kontaktstellen der Körner dafür, dass die Körner aneinander haften. Diese Kontaktstellen werden als *Binderbrücken* bezeichnet und sind für den Zusammenhalt der Formstoffmischung verantwortlich. [36, S. 75–76], [35, S. 58–59]

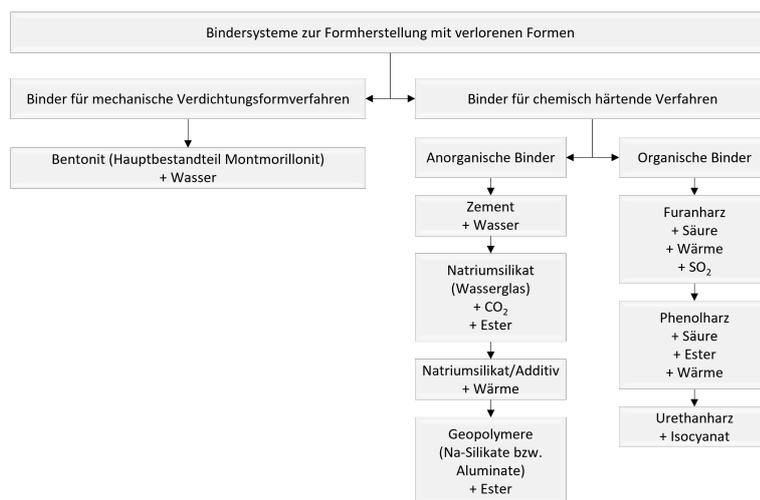


Abbildung 2.2: Wichtige Bindersysteme zur Formherstellung (vgl. [36, S. 251], [12, S. 197])

Abbildung 2.2 zeigt die wesentliche Unterteilung der Bindersysteme zur Formherstellung verllorener Formen. Wie bereits erläutert (Abbildung 2.1), wird bei der Formherstellung zwischen Verdichtungs- und Verfestigungsverfahren unterschieden; das gilt ebenfalls für die zu wählenden Bindersysteme. Für mechanische Verdichtungsverfahren wird hier der Formgrundstoff Bentonit mit Wasser vermischt. Der natürliche Rohstoff Bentonit enthält das notwendige Bindemittel bereits in Form vorhandener Tonminerale [35, S. 38], sodass in diesem Fall die Anreicherung mit Wasser ausreichend ist, um erneut eine stabile Form generieren zu können. Die chemisch härtenden Verfahren sind hingegen in unterschiedliche anorganische sowie organische Bindersysteme unterteilt. [36, S. 251] In dieser Arbeit wird mit einem anorganischen Wasserglasbinder (Natriumsilikat + Ester) gearbeitet, daher wird auf die anderen Bindersysteme nicht vertiefend eingegangen.

Neben den Verfahren für die Produktion von verlorenen Gussformen gibt es auch für die Kernherstellung verschiedene Bindeverfahren. Diese teilen sich in kaltselfthärtende, begasungshärtende und warm- oder heißhärtende Bindersysteme auf, die in Abbildung 2.3 dargestellt werden. Abgesehen von dem kaltselfthärtenden Wasserglas-Ester Verfahren wird auch hier nicht näher auf die anderen Systeme eingegangen.

Das selfthärtende Wasserglas-Ester Verfahren gliedert sich zu den anorganischen Bindersystemen, da Wasserglas ein anorganischer Binder ist. Mit einer Esterverbindung wird zwar ein organischer Härter in das System gegeben, dennoch gehören Esterverbindungen zu den unkritischen Verbindungen, da sie während der Pyrolyse beim Abguss nur sehr geringe schädliche Emissionen abgeben. Zudem nimmt die Härterzugabe mit 8 % bis 12 % bezogen auf die Bindermenge nur einen sehr geringen Anteil des gesamten Bindersystems ein. [101, S. 95] Bei einem kaltselfthärtenden Verfahren wird der Aushärtprozess bei Raumtemperatur und ohne zusätzliche Stoffe durchgeführt [101, S. 42]. Generell bestehen selfthärtende Bindersysteme aus einem Teil meist flüssige oder festen Binders und einem Teil flüssigen Härter. Der Formgrundstoff wird im Mischprozess zunächst mit dem Härter vermengt bevor der Binder hinzugegeben wird. Die Aushärtreaktion in Form einer Polymerisation beginnt, wenn Binder und Härter im Mischer aufeinandertreffen. Für die danach folgende Formgebung des angemischten Sandes wird dieser in das entsprechende Formmodell bzw. in die Kernvorlage gefüllt und verdichtet.

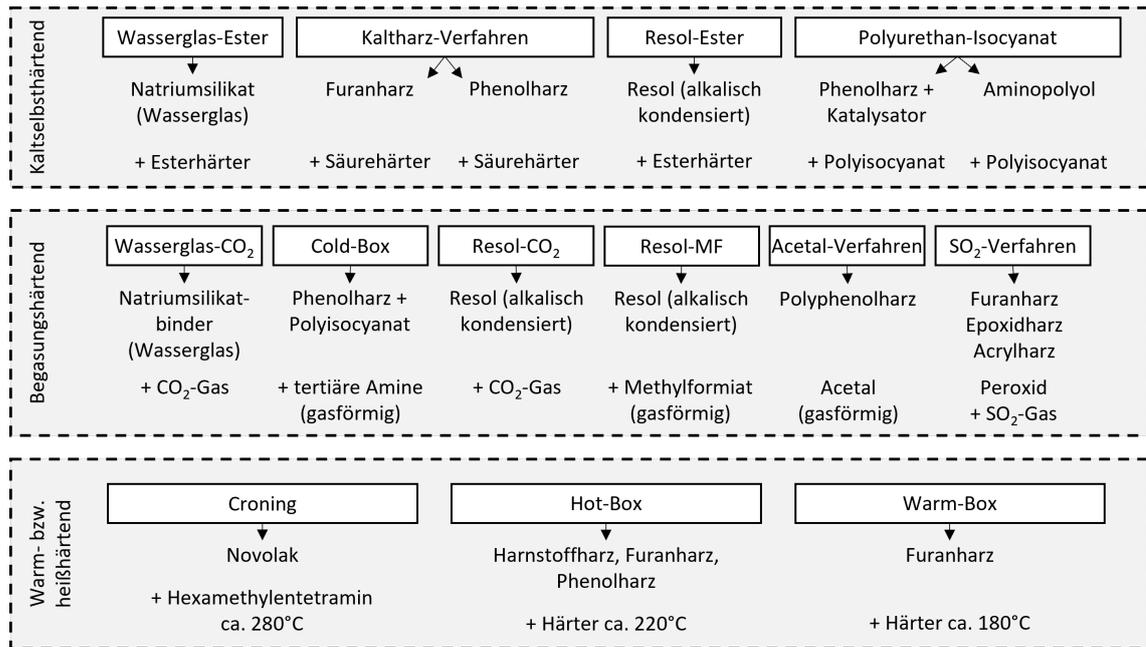


Abbildung 2.3: Bindersysteme zur Kernherstellung (vgl. [37, S. 314])

Für eine festgelegte Aushärtezeit verweilt das Sand-Binder-Gemisch in der Formgebung, bis die Festigkeit der entstandenen Form bzw. des Kernes ausreichend ist, um sich nicht mehr zu verändern. Die so produzierte Form wird aus der Formvorlage genommen, um in der verbleibenden Aushärtezeit offen weiter zu härten, denn durch das Entnehmen sind die Oberflächen zugänglich für die Atmosphäre. Nach einer Aushärtezeit von 24 Stunden bei Raumtemperatur wird keine signifikante Steigerung der Festigkeit mehr erwartet, sodass eine End- bzw. Maximalfestigkeit erreicht ist. Kriterien, die den Aushärteprozess beeinflussen, sind der Schlammstoffgehalt, die Temperatur und die Feuchtigkeit bezogen auf die Luft, den Formstoff sowie das Modell. [101, S. 73–75] Generell sind chemische Bindersysteme nur einmalig zu gebrauchen. Das bedeutet, dass für jede neue Form bzw. jeden neuen Kern ein neues Bindersystem mit dem Formgrundstoff vermischt werden muss [12, S. 248].

### 2.1.2 Formstoffaufbereitung

Unter der *Formstoff- bzw. Formsandaufbereitung* werden Verfahren verstanden, deren Ertrag ein formgerechter, verwendungsfähiger Formstoff ist, welcher wieder in der Kernmacherei oder Formerei eingesetzt werden kann. Die Auslegung einer neuen bzw. unternehmensspezifischen Aufbereitungsanlage steht immer in Abhängigkeit zum eingesetzten Formstoffsystem. Beispielsweise ist bei bentonitgebundenen Formstoffsystemen bis zu 90 % des eingesetzten Materials wieder aufbereitbar, denn die Bindefähigkeit von Bentonit ist reversibel. [36, S. 246–247] Der Formstoff von bentonitgebundenen Modellen kann demnach durch die Zugabe von verbrauchten Bestandteilen wie Wasser, Binder oder anderen Zusatzstoffen aufbereitet werden [101, S. 172]. Chemische Bindersysteme werden mit dem Merkmal der „Regenerierfähigkeit“ gekennzeichnet, das bedeutet, dass die Möglichkeit besteht, Binderrückstände für den Wiedereinsatz des Formgrundstoffes von diesem zu entfernen [12, S. 249]. Der Regenerationsprozess spielt demnach für alle chemisch gebundenen Formstoffsysteme eine wichtige Rolle [101, S. 172]. Dabei können bis zu 95 % der

kaltharzgebundenen Altformstoffe regeneriert und dem Prozess wieder zugeführt werden [36, S. 249]. Für die Aufbereitung von Wasserglas-Ester-Altsand hat sich die mechanische Regenerierung als notwendig erwiesen, um die Binderrückstände von den Sandkörnern zu trennen [101, S. 115]. Laut Definition entspricht die Regenerierung von Altsanden der „Überführung eines Gießereialtsandes in einen Formgrundstoff mit neusandähnlichen Eigenschaften durch Entfernen verbrauchter Binderreste“ [12, S. 248]. Der Prozess der Formstoffaufbereitung beinhaltet folgende 5 Schritte [12, S. 249–250], [35, S. 368], [101, S. 173–175]:

1. Die *Knollenzerkleinerung* erfolgt nach der Entfernung der Gussteile sowie weiterer möglicher Metallteile. Der verbleibende gebundene Formstoff wird über einen Auspackrost gesiebt und vorzerkleinert. Die Siebdauer ist abhängig von der Maschenweite. Bei feinen Sieben zieht sich der Prozess in die Länge, benötigt dann aber keinen weiteren Zerkleinerungsprozess. Große Maschenweiten haben häufig ein weiteres Zerkleinerungsaggregat nachgeschaltet.
2. Zum Entfernen weiterer bestehender Metallteile im Formstoff, wie zum Beispiel Kühlkokillen, Spritzmetalle, kleinere Armierungsteile oder Grate, wird die *Metallabscheidung* durchgeführt. Abhängig vom vergossenen Metall werden Siebeinrichtungen, Wirbelstromtrennanlagen oder Magnetabscheider zur Entfernung eingesetzt.
3. Der verbleibende Formstoff wird im dritten Schritt *regeneriert*. In diesem entscheidenden Schritt werden die Binderhüllen von den einzelnen Körnern des Formgrundstoffes entfernt. Je nach Einsatzgebiet wird zwischen Kaltregenerierungen (mechanischer bzw. pneumatischer Regenerierung), thermischer oder Nassregenerierung unterschieden.
4. Das Regenerat muss nach dem Regenerationsprozess entstaubt werden. Das bedeutet, dass verbleibende gelöste Binderreste, Quarzabrieb oder weitere staubförmige Produkte aus dem Kornhaufwerk zu entfernen sind. Dieser Prozess wird auch als *Sichten* bezeichnet. Verbleibende staubförmige Produkte können beim Wiedereinsatz des Formgrundstoffes zu Gussfehlern im Gussteil führen. Im Gegensatz dazu bedingt ein gut regenerierter sowie entstaubter Formstoff eine Reduzierung der Binder- und Härtermenge, ohne signifikante Qualitätsverluste zu verursachen.
5. Zuletzt besteht die Möglichkeit der benötigten *Temperierung* des Formstoffes. Abhängig vom verwendeten Bindersystem entscheiden vorgegebene Temperaturbereiche über die Qualität des verarbeiteten Formstoffes. Die kritische Maximaltemperatur liegt bei etwa 30 °C.

Eine detailliertere grafische Darstellung des beschriebenen Prozessverlaufes einer Gussform bis zum Regenerat wird in Abbildung 2.4 dargestellt. Dort werden zudem die vier Möglichkeiten der Regeneration aufgelistet, die im Folgenden erläutert werden.

### **Thermische Regeneration**

Die thermische Regeneration ist ein Verfahren zur Aufbereitung organisch gebundener Formstoffe. Dabei werden organische Formstoffsysteme in einem Temperaturbereich zwischen 600 °C und 900 °C regeneriert. Dabei verbrennen die oxidierbaren Bestandteile und der Formgrundstoff bleibt in seiner Ursprungsform bestehen. Abhängig von der Behandlungstemperatur, der Verweilzeit, dem Sauerstoffgehalt der Verbrennungsgase und der Intensität des Wärmetauschers können Regenerierungsgrade von bis zu 100 % erreicht werden. Im Gegensatz dazu kann bei anorganischen Formstoffsystemen das Aufheizen des Formstoffsystems lediglich die Versprödung des Binders bewirken, jedoch keinen Abtrag. Im Nachgang zur thermischen

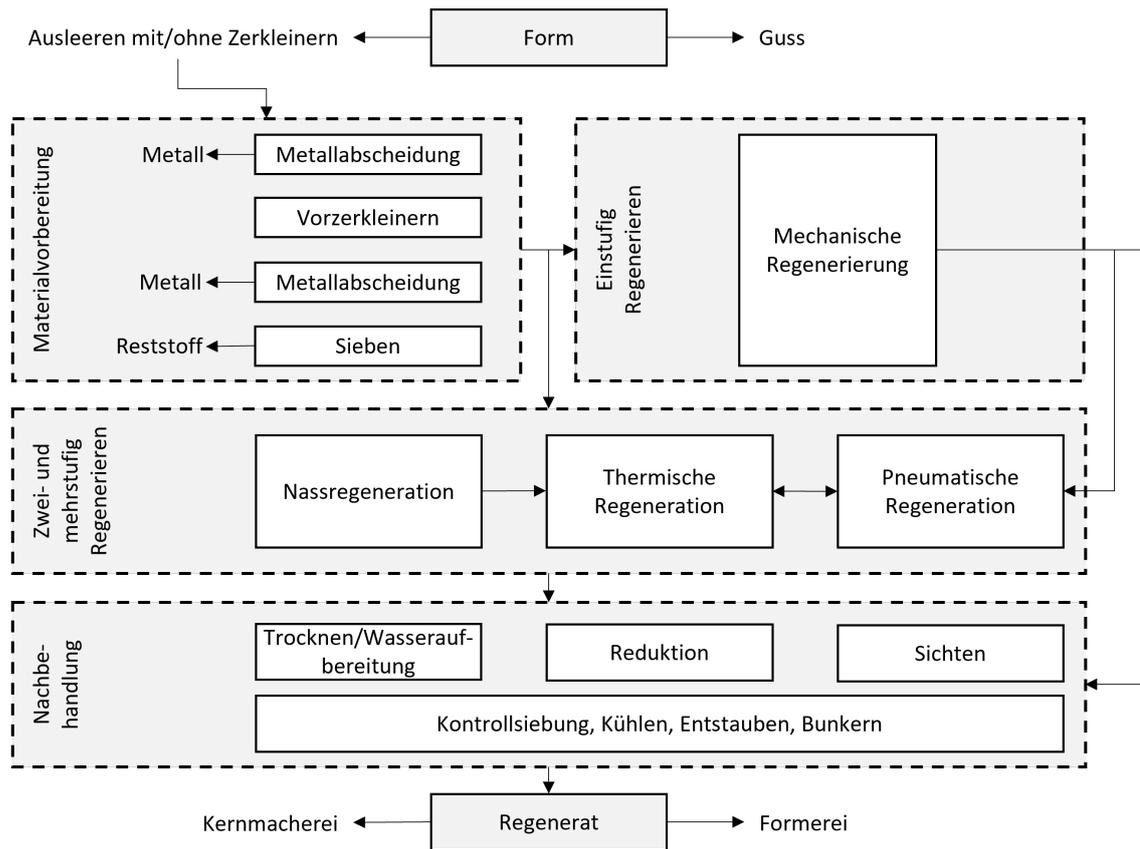


Abbildung 2.4: Prozessverlauf von der Gussform bis zum Regenerat (vgl. [12, S. 250] [35, S. 369])

Regeneration von anorganisch gebundenen Formstoffen ist noch eine mechanische Regeneration notwendig. [35, S. 372–373] Generell ist es bei dieser Art des Regenerierens wichtig, dass die gesamte Kornoberfläche frei zugänglich ist, sodass die umhüllende Bindehülle ganzheitlich oxidieren kann. Durch den hohen Energieaufwand und die notwendige Anlagengröße lohnt sich die thermische Regeneration nur bei einer entsprechend hohen Durchsatzmenge. [12, S. 251], [36, S. 266]

### Mechanische Regeneration

Die mechanische sowie die nachfolgend beschriebene pneumatische Regeneration sind in der Literatur auch unter dem Begriff der *Kaltregenerierung* zu finden. Das Regenerat wird in diesen Prozessen nicht zusätzlich beheizt, da die entstehende Wärme durch die vorherrschende Reibenergie auftritt. Beide Verfahren sind auf das Entfernen der Bindehüllen fokussiert. Beim mechanischen Regenerieren werden Schlag-, Reib- und Prallbeanspruchungen des Regenerates genutzt, um das Abtrennen des verbleibenden Binders zu provozieren. Dazu gehört auch die Reibung der Körner untereinander, zwischen Korn und Reibwerkzeug sowie zwischen Korn und Behälterwand. Die Intensität der mechanischen Regeneration bestimmt den erreichbaren Regenerierungsgrad. Der Vorteil den dieses Verfahren bietet, liegt in der Anwendbarkeit für organische als auch anorganische Formstoffsysteme. Zudem zählen diese Anlagen als „preislich attraktiv“ und weisen günstige Betriebskosten auf. Allerdings tritt bei steigender Intensität auch eine zunehmende Kornzerkleinerung ein, welche berücksichtigt werden muss. [12, S. 250], [35, S. 369]

### Pneumatische Regeneration

Unter der pneumatischen Regeneration wird die mehrstufige mechanische Regeneration verstanden. Dabei befindet sich der Altformstoff in einer sogenannten Prallrohranlage. Der darin befindliche Formstoff wird mit Hilfe eines Luftstroms durch ein Blasrohr gegen eine Prallplatte geblasen. Die entstehenden Reibbeanspruchungen im Blasrohr und der folgende Aufprall bewirken das Abtrennen spröder Binderhüllen, welche mit dem entstehenden Feinstaub über den Staubaustritt bereits teilweise abgesaugt werden können. Der abgeprallte Formstoff fällt wieder in das Auffangbecken oder in eine weitere Regenerierungszelle, um den Prozess erneut zu durchlaufen. Der erreichte Regenerierungsgrad ist abhängig von der Geschwindigkeit des Luftstroms, der Verweilzeit in einer Regenerierungszelle sowie der Anzahl der durchlaufenen Regenerierungszellen. Durch den Aufprall des Sandes auf das Hindernis wird jedoch nicht nur die Binderhülle zerschlagen, es führt auch zu erhöhtem Kornverschleiß, weshalb das Verfahren dahingehend optimiert wurde, dass das Korn vor dem Auftreffen auf das Hindernis abgebremst wird. Die Energie, die der Sand durch den Luftstrom gewinnt, wird beim Abbremsen in Regenerierungsarbeit umgewandelt. Dabei werden die Binderhüllen von den Sandkörnern abgetrennt, ohne dass das Korn selbst beschädigt wird. Diese Art der Regeneration ist aufgrund des hohen Energiebedarfs und des erhöhten Verschleißes sehr kostspielig. Dennoch lassen sich mit der pneumatischen Regeneration im Gegensatz zur mechanischen Regeneration für anorganisch gebundene Formstoffe deutlich höhere Regenerierungsgrade erreichen. [35, S. 370–371], [12, S. 251]

### Nasse Regeneration

Die nasse Regeneration basiert auf dem Prinzip der Attrition<sup>3</sup>. Durch autogenes Aneinanderreiben und -scheuern der Quarzkörner in einem Wasserbad werden Schadstoffe sowie Binderhüllen abgetrennt und ausgeschwemmt. Dieses Vorgehen wird auch als *nasse Sichtung* bezeichnet. Der Regenerierungsgrad ist bei diesem Verfahren von der Intensität des Attritionsprozesses, der Verweildauer in der Attritionszelle und vom Flüssigkeitsgrad des Altsand-Wasser-Gemisches abhängig. Ein großer Nachteil dieser Regeneration liegt in dem hohen Wasserverbrauch von etwa 4-5 m<sup>3</sup> Wasser pro Tonne Altsand. Die Nassregeneration kann nur erfolgreich mit wasserlöslichen Formstoffsystemen durchgeführt werden. Weitere Nachteile dieses Regenerationsverfahrens sind die Anlagengröße, die Aufbereitung bzw. die umweltkonforme Entsorgung des verunreinigten Wassers und des Klärschlammes sowie die Trocknung des Regenerates. Es bietet sich an, dieses Verfahren in Verbindung mit dem Betreiben von Druckwasserauspack- und Nassputzanlagen zu kombinieren. Vorteilhaft ist die sehr hohe Qualität des erzeugten Regenerates. Hier können Neusandqualitäten bzw. sogar bessere Formstoffeigenschaften erreicht werden. [12, S. 251], [35, S. 371–372]

Die Regenerierung von Formstoff gehört als Teilprozess zu der Formstoffaufbereitung. Dabei können nicht nur Altsande, sondern auch Schuttsande mit stark schwankenden Zusammensetzungen durch den Regenerationsvorgang wieder zu brauchbarem Formgrundstoff aufbereitet werden [35, S. 365]. Bereits im Jahr 1993 begründeten Flemming und Tilch die Notwendigkeit der Formstoffregeneration mit folgenden Aspekten [35, S. 365]:

1. Die Verfügbarkeit geeigneter Quarzsande verringert sich, der Aufwand für die Gewinnung und Aufbereitung nimmt zu,

---

<sup>3</sup> „Reinigung der Kornoberfläche von Mineralien durch gegenseitiges Abscheuern unter Zusatz einer Flüssigkeit, meist Wasser.“ [10, S. 68]

2. steigende Kosten für den Abtransport und die Deponie von Altsand bzw. Schutt,
3. zunehmende Deponieprobleme aufgrund steigender Anfallmengen und gesetzlicher Auflagen,
4. teilweise Nutzung technologischer Vorteile bei Verwendung von Regenerat (z.B. geringere Neigung zu Ausdehnungsfehlern, Einsparung von Bindern),
5. Wiederverwendung von Altsanden ist ein wirksamer Schutz der Umwelt und ihren Rohstoffressourcen.

Laut Bührig-Polaczek, Michaeli et al. ist es auch aus ökologischen und ökonomischen Gründen zwingend notwendig Altformstoffe zu regenerieren, um damit den Wiedereinsatz zur Herstellung von Kernen bzw. Formen zu gewährleisten [12, S. 248]. Dabei ist jeder Prozess auf das verwendete Formstoffsystem in Gießereien individuell anzupassen, denn unterschiedliche Formstoffsysteme beinhalten verschiedene (Schad-)Stoffe, die aus dem System entfernt werden müssen [35, S. 366].

### Wichtige Stationen der Kreislaufwirtschaft von Formstoff

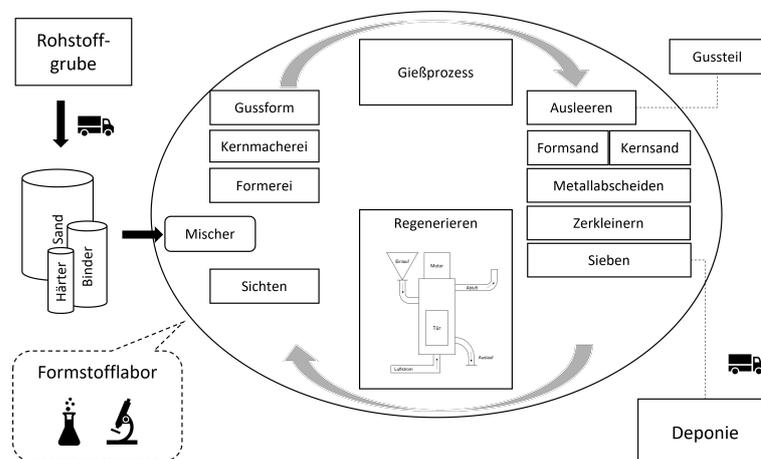


Abbildung 2.5: Der Formstoffkreislauf

Abbildung 2.5 stellt einen möglichen Formstoffkreislauf vom Rohstoff (Formgrundstoff) bis hin zum Abfallprodukt dar. Generell liegt hier ein physikalisch vernetztes System vor, das verschiedene Aggregate wie Kühler, Bunker, Mischer und viele weitere Komponenten und Faktoren wie Zyklonen, Schleudern, Absaugungen, Transportstreckenlängen oder Übergabestellen aufweist und u.a. durch das Gussprogramm und Umweltbedingungen beeinflusst wird [78].

Zu Beginn wird das Rohmaterial der Rohstoffgrube entnommen, gewaschen und klassiert. Gießereien besitzen vor Ort häufig eigene Silos oder Sandlager, damit ein gewisses Materialkontingent vorrätig deponiert werden kann. Von dort wird das Rohmaterial durch einen Mischprozess mit einem Bindersystem zu Formstoff angemischt, wobei das Mischverhältnis sowie die einzelnen Zusatzkomponenten gießereispezifisch und gussteilabhängig zu bestimmen sind. Je nach Rohstoff sowie Mischverhältnis wird der Formstoff für das Produzieren von Kernen oder Formteilen eingesetzt. Die Gesamteinheit aus beiden Komponenten (Gussform und -kern) sowie den Kernkästen und möglichen Hilfseinrichtungen (z.B. Ausstampfböden) bildet die so genannte *Modelleinrichtung*, die im anschließenden Gießprozess mit flüssigem Metall befüllt

wird [12, S. 181]. Nach der Aushärtung des Gussteils wird die Modelleinrichtung wieder entleert und das Gussteil verlässt den Formstoffkreislauf. Der verbleibende Formstoff setzt sich nach wie vor aus Kern- und Formsand zusammen. Metallreste oder andere Verunreinigungen wie Speiser, Filter etc. werden in den Schritten der Metallabscheidung, dem Kornvereinzeln und dem Sieben, aus dem Formstoffsystem separiert und entfernt.

Innerhalb dieses Kreislaufes wird der Formstoff durch unterschiedliche Faktoren beansprucht. Zum einen erfolgt diese durch die erhebliche thermische Belastung des Gießprozesses, die den Formstoff noch weiter Aushärten lassen und damit die Austrocknung des Formstoffes bedingen, wodurch direkt beteiligte Körner irreversibel beschädigt und zum Verschleißprodukt werden können. Des Weiteren führen die mechanischen Beanspruchungen innerhalb des Kreislaufes zur Beschädigung des Formstoffes. Beispielsweise beim Mischen, beim Entpacken der Gussstücke oder bei der Regeneration kommen mechanische Kräfte zum Einsatz, die zu Veränderungen der Kornform sowie der Korngröße führen können: Der dritte belastungstechnische Faktor für den Formstoff sind die chemischen Komponenten. Bei zu starker Verschmutzung des Sandes, sodass dieser nicht wiederverwendbar bzw. regenerierbar ist, muss der Sand aus dem Kreislaufsystem entfernt und auf einer Deponie entsorgt werden. Im Sinne nachhaltiger Ressourcenverwendung sollte das Ziel sein, diesen Abfallanteil so gering wie möglich zu halten, damit der Formstoff durch ein Regenerationsverfahren wiederholt eingesetzt werden kann.

Nach solch einem Regenerationsprozess muss der Formstoff vor dem Wiedereinsatz gesichtet werden, um verbliebene Staub- bzw. Binderpartikel vom Formstoff zu separieren und abzusaugen. Die Formstoffqualität wird durch die gängigen Formstoffprüfungen kontrolliert [36, S. 247]. Abhängig von dem Gießprozess und der Art der Verarbeitung gibt es unterschiedliche Anforderungen an Formstoffe, wie die Fließfähigkeit, die Formbeständigkeit, die Gasdurchlässigkeit, die Hochtemperaturbeständigkeit, die Zerfallsbeständigkeit und die Wiederaufbereitbarkeit.

### 2.1.3 Formstoffqualität

Im gesamten Gussprozess sichern Qualitätsprüfungen einen reibungslosen Prozessablauf. Dazu zählen unter anderem auch regelmäßige Formstoffprüfungen, die sich in sechs Gruppen gliedern lassen [36, S. 258–259]:

- Überprüfung der Eigenschaften des aufbereiteten Formstoffes,
- Überprüfung von Altsandeigenschaften,
- Überprüfung anfallender Staubqualität,
- Überprüfung anfallender Regenrate,
- Überprüfung der Formgrundstoffe im Anlieferungszustand,
- Überprüfung der zu entsorgenden Altsandchargen.

Im Allgemeinen wird die Qualität eines Materials oder Produktes entsprechend der Fähigkeit dieses Produktes, geforderte Produkteigenschaften zu erfüllen, bewertet [100]. Bezogen auf die Formstoffqualität werden die Produkteigenschaften im Allgemeinen mit Kenngrößen bzw. Maßzahlen der *Größe* bzw. der *Form* einzelner Partikeln definiert. Dabei ist zu berücksichtigen, dass unregelmäßig geformte Partikel nicht exakt beschrieben werden können. Beispielsweise hat die Angabe eines Längenmaßes nur einen

beschränkten Informationsgehalt. Kennzahlen, die die *Größe* eines Partikels beschrieben, werden als *Feinheitsmerkmal* bezeichnet. Dabei sind direkte geometrische Abmessungen, statistische Längen sowie geometrische und physikalische Äquivalentdurchmesser und die spezifische Oberfläche als Partikelgrößenmaß gebräuchlich. [121, S. 12] Die Partikelgröße und -form gehören neben den folgenden Qualitätskriterien zu den *granulometrischen* Eigenschaften [35, S. 44], [36, S. 259]:

- mittlere Korngröße, Korngrößenverteilung und AFS Kornfeinheitsnummer,
- Kornform und Eckigkeitsgrad,
- theoretische spezifische Oberfläche des Kornhaufwerkes,
- Schlämstoff und Feinanteile (Kornfraktionen im Bereich 0,1 bis > 0,02 mm),
- „Aktivität“ der Kornoberfläche hinsichtlich des Haftvermögens der Binderhülle (Bindeeigenschaften).

Neben den genannten granulometrischen Eigenschaften lassen sich weitere Kennwerte, nämlich *morphometrische* Eigenschaften, mit rasterelektronenmikroskopischen Untersuchungen (REM) ermitteln. [28, S. 66]:

- Körnungsaufbau (Korngefüge),
- Kornform und Kornrundung,
- Oberflächenbeschaffenheit.

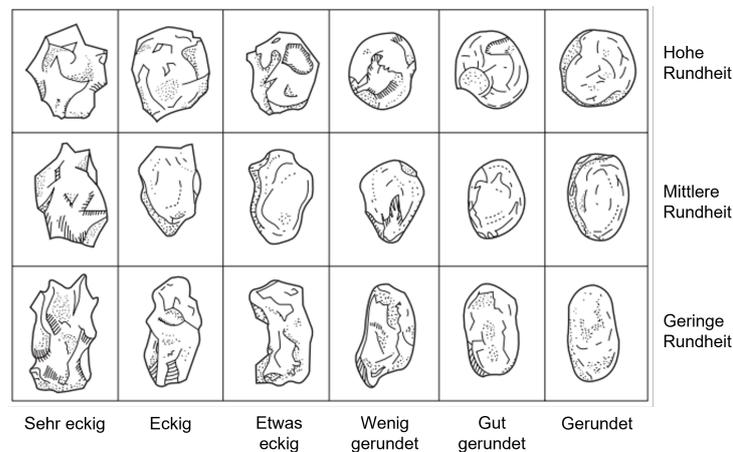


Abbildung 2.6: Abstufung der Kornformen (vgl. [9, S. 147])

Anhand abgestufter Merkmale können die Kornoberflächen mit einem Rauheitsgrad (RAG-Wert) bewertet und von glatt bis rau bzw. zerklüftet beschrieben werden [35, S. 47–49]. Bei der Kornform wird unterschieden zwischen rund bis gerundet, kantengerundet, eckig bis kantig oder splittrig. Dabei definiert der Eckigkeitskoeffizient (EK) die Kornform. Liegt dieser Kennwert bei 1, liegen ideal runde Körner vor. Reale Gießereisande liegen zwischen einem Eckigkeitskoeffizient von 1,1 und 1,3 (kantengerundet), 1,3 bis 1,7 (eckige Kornform) und 1,7 (splittrige Kornform). [35, S. 46] Das Korngefüge, also der kristalline Aufbau des Korns, kann entweder monokristallin oder polykristallin sein [129, S. 5]. Regelmäßige

Qualitätsuntersuchungen des Formstoffes sind wichtig, da auch innerhalb einer Lagerstätte Körner mit unterschiedlichen Rundungsgraden auftreten. Dabei ist der Rundungsgrad korngößenabhängig und feinere Körner weisen generell einen geringeren Rundungsgrad auf. [28, S. 69–70] Die Erfassung und Analyse der vorliegenden Kornform ist sehr aufwendig. Eine Annäherung beschreibt das Verhältnis der spezifischen Oberfläche des realen Kornhaufwerkes zur theoretisch berechneten spezifischen Oberfläche der Körner. [35, S. 46] Abbildung 2.6 verdeutlicht die beschriebenen Abstufungen der Kornform sowie der Kornrundungen grafisch.

### Korngrößenverteilung mittels Siebanalyse

Der Einsatz von Kiessanden in verschiedenen Industriezweigen ist von der Korngrößenverteilung und Form der Körner abhängig. Die Korngröße ist das wichtigste petrografische Merkmal für die Klassifizierung von Kiesen und Sanden und wird normalerweise durch Siebanalysen ermittelt [47, S. 34].

Die Siebanalyse ist ein Verfahren aus der Partikelmesstechnik, bei der Sande oder Kiese in mehrere so genannte *Fraktionen* oder *Körnungen* klassiert werden [121, S. 10, 285]. Mit Hilfe der Siebanalyse lassen sich die granulometrischen Kennwerte sowohl rechnerisch als auch zum Teil graphisch ermitteln [35, S. 46]. Grundlage für die Durchführung und Auswertung des Siebverfahrens bietet das VDG-Merkblatt P 27 *Prüfung von Formgrundstoffen. Bestimmung der granulometrischen Kennwerte* [129].



Abbildung 2.7: Siebturm



Abbildung 2.8: Siebe des Siebturms

Das Siebverfahren wird mit Hilfe eines Siebturmes durchgeführt, welcher beispielhaft in Abbildung 2.7 dargestellt ist. Der Turm setzt sich aus einer variablen Anzahl von Prüfsieben mit im Turmverlauf kleiner werdenden Maschenweiten zusammen. Abbildung 2.8 zeigt stellvertretend drei Siebe zur Verdeutlichung der abnehmenden Maschenweite. Die Maschenweiten der Siebe orientieren sich an einheitlichen Siebsatzintervallen, die nach ISO 3310, Teil 1 genormt sind. Der Siebturm wird mit einem Deckel durch das Festschrauben auf eine Vibrationsvorrichtung gespannt, welche mit einer einstellbarer Frequenz vibriert. Diese Vibration bringt die vorher eingefüllte Sandmenge in Bewegung, sodass sich die Körner eigenständig klassieren, indem sie durch die passenden Maschen der Siebe fallen. Die Rütteldauer ist ebenfalls einstell-

bar, um untereinander vergleichbare Analyseergebnisse erzielen zu können. Durch die unterschiedlichen Kornformen sowie möglichen Ecken und Kanten am Korn ist die Position des Kornes für den Siebdurchgang entscheidend. Ein längliches Korn mit geringem Durchmesser, wie zum Beispiel ein Reiskorn, würde liegend nur durch ein sehr grobmaschiges Sieb fallen, aufgerichtet passt es jedoch durch viel feinere Siebmaschen.

Zur Auswertung einer Siebanalyse werden die verwendeten Siebe vor dem eigentlichen Siebverfahren gewogen, die Messwerte erfasst und gespeichert. Nach dem Siebvorgang sind wieder alle Siebe inklusive der verbleibenden Sandmenge pro Sieb zu wiegen. Die Differenz beider Messwerte liefert das Sandgewicht, welches sich pro Sieb angesammelt hat. Dabei sind die Ergebnisse einer Siebanalyse immer abhängig von der Beschaffenheit der Analysesiebe sowie vom Verfahren selbst. [129, S. 1] Mit den erfassten Analysewerten lassen sich sowohl grafische als auch rechnerische Kennwerte ablesen bzw. berechnen. Die Gewichtswerte können als Häufigkeitsverteilung in einem Histogramm bzw. einer Verteilungskurve, Abbildung 2.9, oder grafisch als Summenkurve bzw. Siebdurchgangskurve, Abbildung 2.10, dargestellt werden. Aus den Diagrammen sind verschiedene Parameter wie z.B. der mittlere Korndurchmesser ablesbar. Die Sortierung ist für eine weitere Charakterisierung rechnerisch zu ermitteln. [47, S. 34]

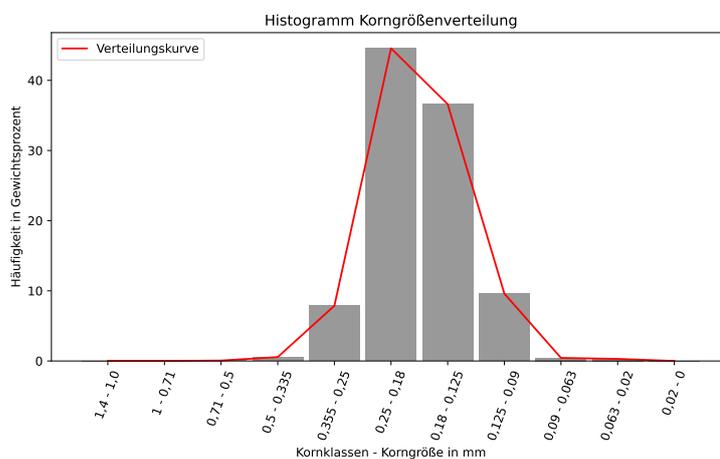


Abbildung 2.9: Histogramm und Verteilungskurve

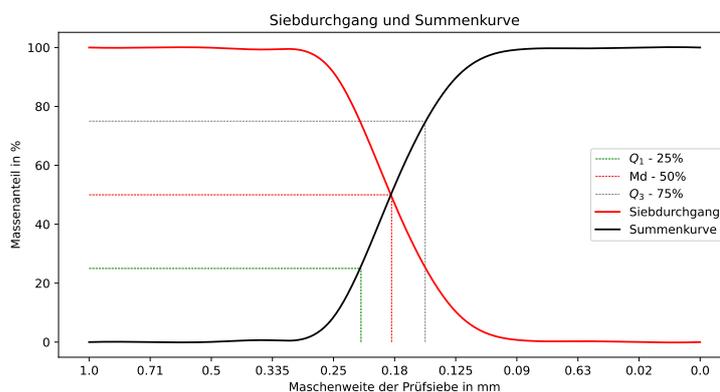


Abbildung 2.10: Siebdurchgang und Summenkurve

Aus der Summenkurve (Abbildung 2.10) lassen sich einige wichtige Kennwerte ablesen. Über den 50%-Wert der prozentualen Häufigkeit auf der Ordinate kann der Wert  $Md$  (*median diameter*), also die mittlere Korngröße, abgelesen werden und liefert den Wert für die Korngröße mit 50% Korndurchgang. Auf die selbe Art können die Werte für den Durchgang der Korngrößen bei 25% und 75% abgelesen werden. Hier wird das erste Quartil  $Q_1$  bei einem Prozentwert von 25 und das dritte Quartil  $Q_3$  bei 75% abgelesen. In Abbildung 2.10 sind alle drei Werte mit gestrichelten Linien farblich gekennzeichnet.

Mit beiden Messwerten lässt sich die *Sortierung*

$$S_o = \sqrt{\frac{Q_3}{Q_1}} \quad (2.1.1)$$

nach TRASK [28, S. 66] und der Gleichmäßigkeitsgrad

$$GG = Q_3 - Q_1 \quad (2.1.2)$$

[129, S. 5] berechnen. Der Gleichmäßigkeitsgrad ist ein Maß zur Beschreibung der Breite der Kornverteilung. Er berechnet sich aus den prozentualen Massenanteilen zwischen 25% und 75% des Siebdurchgangs. [10, S. 428]

Tabelle 2.2: Multiplikatoren [129, S. 4]

Kornklasse [mm]		Multiplikator $M1_i$ [ $\frac{\text{cm}^2}{\text{g}}$ ]	Multiplikator $M2_i$ [mm]	Multiplikator $M3_i$
$d_o$	$d_u$	$S_{th}$	$MK_{\text{calc}}$	$AFS$
5,6	4,0	5	4,733	2
4,0	2,8	7	3,347	3
2,8	2,0	9	2,366	4
2,0	1,4	13	1,673	6
1,4	1,0	19	1,183	9
1,0	0,71	26	0,843	15
0,71	0,5	37	0,596	25
0,5	0,335	53	0,421	35
0,355	0,25	75	0,298	45
0,25	0,18	105	0,212	60
0,18	0,125	148	0,150	81
0,125	0,09	211	0,106	118
0,09	0,063	296	0,075	164
0,063	0,02	546	0,035	275

Im Vergleich liefern die rechnerischen Kennwerte genauere Ergebnisse, denn graphische Auswertverfahren verlieren durch Subjektivität an Exaktheit [35, S. 46]. Dabei kommen drei kornklassenbezogene Multiplikatoren ins Spiel, die in Tabelle 2.2 aufgelistet werden. Dabei gibt  $d_o$  die obere und  $d_u$  die untere Grenze des Wertebereichs bzw. der Intervallgrenzen an, die eine Korn- bzw. Partikelklasse beschreibt [121, S. 25]. Mit den gegebenen Multiplikatoren und dem Massenanteil  $g_i$ , der sich beim Siebvorgang innerhalb der Grenzen  $d_o$  und  $d_u$  sammelt, lassen sich die Kennzahlen zur qualitativen Bestimmung eines

Kornhaufwerkes wie folgt berechnen.

**Mittlere Korngröße (Kalkuliert):**

Neben der aus dem Diagramm ablesbaren mittleren Korngröße, liefert die Kennzahl

$$MK_{calc} = \frac{\sum g_i \cdot M2_i}{g} \text{ [mm]} \quad (2.1.3)$$

die kalkulierte mittlere Korngröße. [129, S. 4]

**AFS Kornfeinheitsnummer:**

AFS steht für den Verband amerikanischer Gießereifachleute - **American Foundrymen's Society**. Die AFS Kornfeinheitsnummer

$$AFS = \frac{\sum g_i \cdot M3_i}{g} \quad (2.1.4)$$

ist eine dimensionslose Bewertungseinheit für Form- und Kernsande und gibt näherungsweise die Anzahl der Maschen eines Siebes in Zoll an. [36, S. 13], [129, S. 4]

**Spezifische Oberfläche:**

Die theoretische spezifische Oberfläche

$$S_{th} = \frac{\sum g_i \cdot M1_i}{g} \left[ \frac{\text{cm}^2}{\text{g}} \right] \quad (2.1.5)$$

wird unter der Annahme berechnet, dass die vorliegenden Körner ideale Kugelgestalt haben. Wie der Name der Kennzahl bereits beschreibt, handelt es sich nicht um die Berechnung der wahren Kornoberfläche. [28, S. 70] Folgende hier relevante Tendenzen sind für abnehmende Partikelgrößen zu nennen: Zunächst gilt, dass sich das Streuvermögen für Licht (Beugung, Reflexion, Absorption) verändert, wodurch sich die wahrgenommene Farbe der Körner verändert. Des Weiteren nimmt die spezifische Oberfläche bei abnehmender Partikelgröße zu. Eine Zunahme der spezifischen Oberfläche hat zur Folge, dass die Schüttdichte abnimmt, die Neigung zu Agglomeration und Wandhaftung wächst und die Festigkeit von Agglomeraten und Pellets größer wird. [121, S. 11], [129, S. 4]

**Biegefestigkeit**

Über die Festigkeit lässt sich generell die Beanspruchbarkeit eines betrachteten Werkstoffes erfassen. In diesem Fall wird die Festigkeit des Formstoffes durch einen Biegeversuch ermittelt, um die mechanischen Eigenschaften des Werkstoffes bestimmen zu können [10, S. 304].

Zur Durchführung dieses Versuches liegt ebenfalls ein VDG-Merkblatt (P 72 - *Bindemittelprüfung*) vor [130]. Zur Durchführung von Biegeversuchen sind zunächst die Art des Bindemittels, die Sandsorte, das Mischverhältnis sowie der Mischertyp zu erfassen. Eine Formstoffmischung beträgt optimalerweise zwischen 3-5 kg und wird mittels Labormischer aus den Komponenten, Formgrundstoff, Formstoffbinder und Formstoffhärter angemischt. Das Mischgut wird unmittelbar nach der Anfertigung in Formkästen gefüllt, verdichtet und das überschüssige Material bündig abgestreift, wie es in Abbildung 2.11 zu sehen ist. Nach einer festgelegten Entformungszeit von 30 Minuten werden die Prüfstäbe aus den Kästen entnommen. Danach gibt es fünf einzuhaltende Lagerzeiten, zu denen jeweils ein Teil der Biegestäbe geprüft wird. Diese



Abbildung 2.11: Formkästen mit Prüfstäben



Abbildung 2.12: Biegeversuch

liegen bei 30, 60, 90, 150, 270 und 1470<sup>4</sup> Minuten nach Herstellung der Sandmischung. Vor jeder Prüfung werden die Gewichte der Prüfstäbe gemessen. Die Biegefestigkeit wird ermittelt, indem ein zugespitzter Kolben so lange gegen den Formstoffstab schiebt, bis dieser der Kraft nicht mehr standhalten kann und bricht, welches in Abbildung 2.12 dargestellt wird. Der maximal aufgebrauchte Druck, gemessen in  $\frac{\text{N}}{\text{cm}^2}$ , wird erfasst und entspricht der Biegefestigkeit.

## 2.2 Akustische Signalverarbeitung

Für die Prozessüberwachung des untersuchten Regenerationsprozesses liegt der Fokus auf der Auswertung akustischer Prozessemissionen mit Maschinellen Lernmethoden. Zum besseren Verständnis des Vorgehens werden im Folgenden die notwendigen theoretischen Grundlagen der digitalen Signalverarbeitung, auch unter dem englischen Begriff *Digital Signal Processing* (DSP) bekannt, und die verwendeten Methoden erläutert.

Im Allgemeinen beschreibt ein *Signal* den zeitlichen Verlauf einer betrachteten physikalischen Größe, welche ein Träger von Information und Energie ist [103, S. 4], [85, S. 1]. Mit der Signalverarbeitung werden Methoden zur Extraktion dieser Informationen zur Verfügung gestellt. Die Auswahl passender Analysemethoden hängt von der Art des Signals und der Art der im Signal enthaltenen Informationen ab. Die digitale Signalverarbeitung wird bereits in vielen Bereichen der modernen Elektronik eingesetzt, in denen Informationen in einem digitalen Format verarbeitet werden, wie z.B. in den folgenden Bereichen: Sprachverarbeitung, Kommunikation, Biomedizin, Unterhaltungselektronik (Musiksynthese und digitales Audio und Video), Seismologie, Bildverarbeitung, Navigation oder genomische Signalverarbeitung. Generell geht es darum, das Signal mathematisch darzustellen und einen geeigneten Algorithmus zur Extraktion relevanter Informationen zu verwenden. Die meisten Signale der Natur liegen in analoger Form vor und sind kontinuierlich im Zeitbereich mit kontinuierlicher Amplitude. [105, S. 1–6]

<sup>4</sup>1470 Minuten entsprechen einer Aushärtzeit von 24h plus die 30 Minuten, die die Riegel in der Form verweilen.

Töne bzw. Geräusche im Sinne eines akustischen Ereignisses entsprechen akustischen Wellen, die als Druckschwingungen durch Luft, Gase oder Flüssigkeiten übertragen werden können [90, S. 19], [86, S. 1]. Ein akustisches Ereignis  $x(t)$  entspricht also der physikalischen Größe des zeitabhängigen Schalldrucks oder der Mikrofonspannung, wobei  $t$  als zeitliche Variable definiert ist [87, S. 3]. Grafisch lässt sich die Wellenform des Schalls, also die Veränderung des Luftdrucks, in einem Diagramm darstellen, wobei die Wellenform die Abweichung des Luftdrucks vom durchschnittlichen Luftdruck aufzeigt [90, S. 19]. Dabei setzt sich ein akustisches Signal meist durch die Summe unterschiedlicher reiner Töne zusammen, die wiederum als sinusförmiger Verlauf vorliegen<sup>5</sup> [86, S. 2–3].

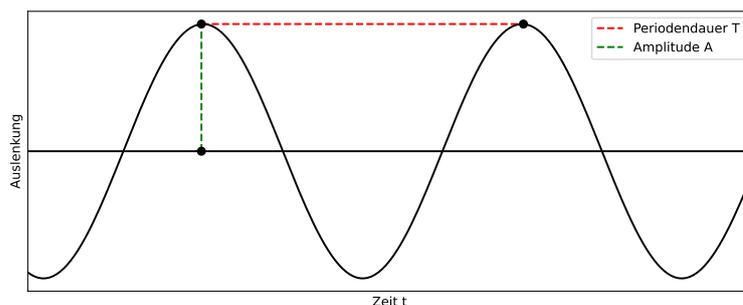


Abbildung 2.13: Kennwerte einer harmonischen Schwingung (vgl. [72, S. 16], [90, S. 21])

Es gibt drei Kennwerte, die zur grundlegenden Beschreibung einer Schwingung wichtig sind: die Periode, die Frequenz und die Phase. Die Periode bzw. Periodendauer

$$T = \frac{1}{f} \quad (2.2.1)$$

beschreibt die benötigte Zeit, um einen Zyklus der Schwingung zu durchlaufen. Abbildung 2.13 stellt die Periodendauer einer harmonischen Schwingung mit der rot gestrichelten Distanz dar. Die Frequenz der Schwingung berechnet sich entsprechend aus dem Kehrwert der Periodendauer.

### Harmonische Schwingungen

$$x(t) = A \cdot \cos(\omega t + \varphi_0) \quad (2.2.2)$$

können durch die Sinus- bzw. Kosinusfunktion vollständig beschrieben werden. Dabei wird die zeitliche Änderung der horizontalen bzw. der vertikalen Position eines Punktes auf einer Kreisbahn als harmonische Schwingung bezeichnet. Die Amplitude  $A$  ist der Betrag der maximalen Auslenkung der Schwingung und wird durch die grün gestrichelte Linie aus Darstellung 2.13 repräsentiert, während

$$\omega = 2\pi f = \frac{2\pi}{T} \quad (2.2.3)$$

die auf den Einheitskreis bezogene Kreisfrequenz der Schwingung darstellt. Bei einer harmonischen Schwingung wird die Phase durch den Phasenwinkel  $\varphi(t)$  beschrieben.  $\varphi_0$  ist der Nullphasenwinkel, also der Winkel zum Zeitpunkt  $t = 0$ . Er gibt an, wo sich ein Signal zum Zeitpunkt 0 befindet. [90, S. 21], [72, S. 15–17]

<sup>5</sup>Zur Visualisierung der beschriebenen Addition einzelner Signale zu einem gesamten Signal ist die folgende Website empfehlenswert: <https://teropa.info/harmonics-explorer/> (besucht am 21.03.2022)

Unter dem Begriff der *Signalverarbeitung* wird die Bereitstellung von *Systemen*, die zur kontrollierten Beeinflussung von Signalen genutzt werden, verstanden [103, S. 5]. Dabei kann das System ein Eingangssignal empfangen oder verändern und ein entsprechendes Ausgangssignal erzeugen. [85, S. 5], [87, S. 3] Akustische Signale sind in analoge und digitale Signale bzw. Systeme zu unterteilen. Ein analoges System wird dabei zur Verarbeitung und Übertragung zeitkontinuierlicher Signale genutzt, wohingegen ein digitales System zeit- und wertdiskrete Signale bearbeitet. [4, S. 24] Wertdiskret bedeutet, dass die digitalen Werte nur eine endliche, nach oben begrenzte Anzahl von Zuständen annehmen können [120, S. 80–81].

Die Zeitdiskretisierung wird durch die *Abtastrate* festgelegt, welche zur Digitalisierung eines analogen Signals entscheidend ist. Dafür wird das analoge Signal in einem genau festgelegten, gleichen Abstand abgetastet und die gemessenen Spannungswerte werden zu einem digitalen Wert konvertiert. [120, S. 80–81] Der Abstand zwischen den einzelnen Abtastwerten ist als *Abtastzeit*  $t_A$  definiert und ergibt in seinem Kehrwert die *Abtastfrequenz*  $f_A = \frac{1}{t_A}$ . Allerdings bestehen für minimale Abtastzeiten sowie maximale Abtastfrequenzen technische Grenzen, denn die Verarbeitung von Abtastwerten geschieht durch Signalprozessoren, welche durch ihre Verarbeitungsgeschwindigkeit reguliert werden und damit auch die Abtastfrequenz begrenzen. Bei dieser Zeitdiskretisierung eines kontinuierlichen Signals gehen die Signalwerte zwischen den Abtastzeitpunkten jedoch bei der Umwandlung in das digitale Signal verloren. Sind die Veränderungen zwischen den Abtastzeiten gering, kann das gesamte Signal vollständig rekonstruiert werden. Die Bedingungen für diese Rekonstruktion sind im sogenannten *Nyquist-Shannon Abtasttheorem* festgelegt. [103, S. 234–235]

Für die Umwandlung analoger in digitale Signale werden so genannte *Analog-Digital-Wandler (ADW)* eingesetzt. Neben der zeitlichen Abtastung wird das Signal auch amplitudenquantisiert. Das bedeutet, dass die Amplituden, genauso wie die Zeit, gerastert wird. [120, S. 81–82], [85, S. 150] Dabei hängt die Genauigkeit der Quantisierung von der verwendeten Amplitudenaufösung ab [87, S. 6]. Abbildung 2.14 zeigt beispielhaft die Unterschiede der Abtastung sowie der Quantisierung eines Signals. In blau sind jeweils die Abtastabstände bzw. die Quantisierungsschritte dargestellt. Heutzutage ist bereits der Großteil akustischer Messgeräte computergestützt und besitzt ADW zur Wandlung analoger in digitale Signale. Die erzeugten Computerdaten liegen normalerweise in binärer Darstellung vor. Durch die Abtastung des Schalldrucks über die Zeit sowie die Quantisierung der Amplitude werden aus dem analogen Signal Computerdaten generiert. [135, S. 22], [87, S. 6], [4, S. 212]

Im Allgemeinen kodiert ein digitales Signal alle Informationen, die notwendig sind, um ein Audiosignal (innerhalb der Grenzen des Nyquist-Shannon-Theorems) akustisch zu reproduzieren. Dabei ist zu ergänzen, dass Audiosignale akustische Signale sind, die in den Frequenzgrenzen des menschlichen Gehörs (16 Hz bis 16.000 Hz) liegen [90, S. 19], welche logarithmisch aufgeteilt werden können [116, S. 208].

Ein Schallereignis besitzt im Wesentlichen zwei wichtige Merkmale; nämlich die Klangfarbe und die Lautstärke [86, S. 1]. Ergänzende Charakteristika von Audiodarstellungen sind die Frequenz, Tonhöhe, Dynamik, und Intensität. Die Frequenz und die Tonhöhe (engl. Pitch) von reinen Tönen stehen dabei in einem Zusammenhang, eine Frequenz von 440 Hz entspricht der Tonhöhe A4 (Kammerton A). Bei Klangmischungen, wie zum Beispiel das Zusammenwirken mehrerer Instrumente in einem Musikstück, sind diese Zusammenhänge allerdings nicht mehr eindeutig. [90, S. 19–22] Auch die Klangfarbe kann nicht eindeutig extrahiert werden und ist eher eine Wahrnehmungseigenschaft des Klangs. Zur klangfarbenbasierten Strukturanalyse wird hier eine ursprünglich entwickelte Spracherkennungsmethode verwendet. Mit den *mel-frequency cepstral coefficients* (MFCCs) wird die grobe Form der spektralen Hüllkurve parametrisiert und die MFCC-basierenden Merkmale erfassen die klanglichen Eigenschaften des Signals. [90, S. 180–181]

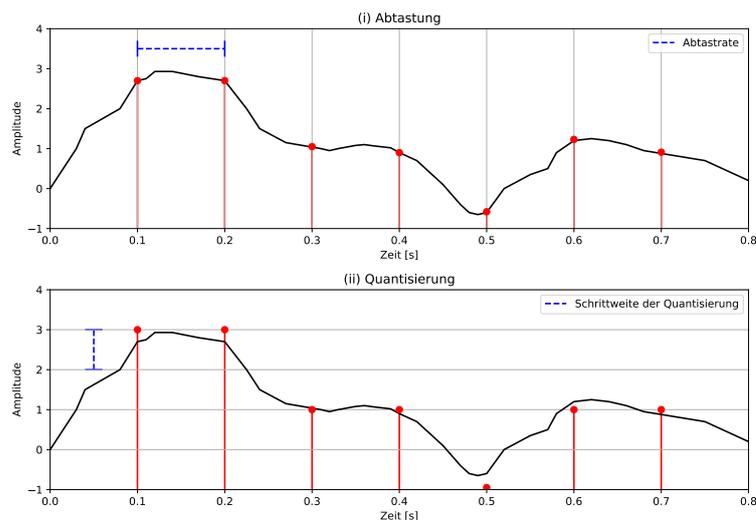


Abbildung 2.14: Analog-Digital-Wandler (i) Abtastung (ii) Quantisierung (vgl. [90, S. 61])

Um das Spektrum eines Signals abbilden zu können, wird im folgenden Kapitel die diskrete sowie die schnelle Fourier-Transformation erläutert. Diese Transformation ermöglicht eine Informationsextraktion in der ursprünglichen Domäne des Signals oder in einer transformierten Domäne [105, S. 1]. Für akustische Signale bedeutet dies, dass die Möglichkeit der Analyse sowohl im Zeit- als auch im Frequenzbereich besteht.

## 2.2.1 Diskrete und Schnelle Fourier-Transformation

Bei komplexen Schwingungsvorgängen wird häufig auf die Darstellung des zeitlichen Schwingungsverlaufes des Signals verzichtet. Stattdessen wird mit Hilfe von Fourier-Methoden das sogenannte *Spektrum* dargestellt [85, S. 7]. Dafür wird der zeitliche Verlauf in die einzelnen vorliegenden Frequenzen überführt. Die *Fourier-Transformation* (FT) stellt eine Integraltransformation dar und wird eingesetzt, um Signale aus dem Zeitbereich in den Frequenzbereich zu transformieren. Diese Umwandlung liefert die Basis, um bestimmte Signaleigenschaften zu extrahieren. [65] Durch diese Transformation gehen allerdings die Informationen der Phase verloren, wodurch das Spektrum weniger informativ als die vollständige Darstellung des zeitlichen Schwingungsverlaufes ist [99, S. 262]. Über die inverse Fourier-Transformation lässt sich die Zeitfunktion wieder aus dem Spektrum zurück gewinnen [92, S. 194].

Die Fourier-Transformation wurde um 1822 von Jean Baptiste Joseph Fourier (1768 - 1830) entwickelt [74, S. 1] und beruht auf der Transformation analoger, also kontinuierlicher Signale. Sie ist ein nützliches Werkzeug zur Bestimmung der Frequenzanteile periodischer Funktionen [103, S. 94]. Für die Umwandlung eines digitalen Signals vom Zeit- in den Frequenzraum nutzen Digitalrechner diskrete Signale bzw. generell diskrete Werte. Aus diesem Grund wird in dieser Arbeit die *diskrete Fourier-Transformation* (DFT) verwendet, bei der zusätzlich zum Zeitbereich auch der Frequenzbereich diskretisiert wird. [103, S. 272] Die schnelle Fourier-Transformation, auch bekannt als *Fast Fourier-Transformation* (FFT), ist eine effiziente und schnelle Umsetzung der diskreten Fourier-Transformation [136, S. 114]. Sie wird verwendet, um die abgetasteten diskreten akustischen Daten in eine diskrete Frequenzdarstellung zu transformieren [77, S. 10].

Wird ein zeitkontinuierliches Signal  $y(t)$  mit einer Abtastzeit  $t_A = \frac{1}{f_A}$  abgetastet, ergibt sich das abgetastete Signal  $y_n = y[n] = y(nt_A)$ ,  $n \in \mathbb{Z}$ . Insbesondere kann die Folge

$$y_n = y[n] = y(s) \cdot \sum_{i \in \mathbb{Z}} \delta(s - it_A), \quad s = nt_A, \quad n \in \mathbb{Z} \quad (2.2.4)$$

auch mit der Einheitsimpulsfunktion

$$\delta(t) = \begin{cases} 1, & t = 0 \\ 0, & \text{sonst} \end{cases} \quad (2.2.5)$$

erzeugt werden.

Somit ergibt sich eine Folge  $Y$  von Zahlen, für die die diskrete Fourier-Transformierte

$$Y^*(f) = \sum_{n=-\infty}^{\infty} y_n e^{-j2\pi f n t_A} \quad (2.2.6)$$

definiert ist [103, S. 250]. Wird der Beobachtungszeitraum auf ein endliches Intervall  $T_0 = [0; (N-1)t_A]$  eingeschränkt, ergibt sich eine Folge  $(y_0, y_1, \dots, y_{N-1})$  von  $N$  Abtastwerten [103, S. 272–273].

Der Index  $n$  des Zeitparameters läuft aufgrund der  $N$  endlichen Abtastwerte von 0 bis  $N-1$ . Für ein zeitdiskretes Signal ist das Spektrum mit der Periode  $f_A$  periodisch, sodass eine einzige Periode ausreichend für die Beschreibung des Spektrums ist. Für die Frequenzdiskretisierung wird die Periode des Spektrums durch endlich viele Spektrallinien beschrieben. Die Anzahl dieser Linien ist zwar frei wählbar, jedoch ist es sinnvoll sie gleich der Anzahl der Abtastwerte ( $N$ ) zu setzen. Als rechneraugliche Näherung ergibt sich für die DFT [103, S. 273]:

$$y_n^* = \sum_{i=0}^{N-1} y_i e^{-\frac{j2\pi n i}{N}}, \quad n = 0, \dots, N-1 \quad (2.2.7)$$

Die Schnelle Fourier-Transformation ist eine Verbesserung der klassischen DFT, die Symmetrie-Eigenschaften ausnutzt und die Berechnung dadurch beschleunigt. Dadurch verringert sich die Anzahl der Rechenoperationen, obwohl die FFT dieselben Eigenschaften wie die DFT besitzt [103, S. 279].

### 2.2.2 Kurzzeit-Fourier-Transformation

Die Fourier-Transformation liefert detaillierte Informationen über den Frequenzgehalt eines Signals. Allerdings kann sie keine Auskunft darüber geben, wann diese Frequenzen innerhalb des Signals zeitlich auftreten. Für viele Analysen akustischer Daten ist es aber wichtig, gleichzeitig den Frequenzinhalt und seine zeitliche Entwicklung zu charakterisieren. Die Kurzzeit-Fourier-Transformation, auch bekannt unter dem englischen Ausdruck *Short Time Fourier-Transformation* (STFT), liefert die Lösung für diese Problematik [11, S. 70]. Die STFT zerlegt das vorliegende Signal in ausreichend kleine Teile, sodass diese als stationär angenommen werden können [84, S. 215]. Dadurch kann wiederum die Fourier-Transformation auf die einzelnen erzeugten Abschnitte angewendet werden, da diese in der Lage ist, nur periodische und stationäre Signale zu charakterisieren [11, S. 70]. Die Signalausschnitte werden mit sogenannten Fenstern bzw. deren entsprechenden Fensterfunktionen festgelegt, sodass eine gefensterte FFT berechnet wird. Jedem berechneten Einzelergebnis kann also eine bestimmte Position in der Zeit-Frequenz-Ebene zugeordnet

werden. [84, S. 215]

Die STFT wird im Kontext dieser Arbeit ebenfalls nur im diskreten Bereich betrachtet. Zusätzlich zum Spektrum  $F_x^\gamma[m, k]$  wird das Signal  $x(t)$  auch im Zeitbereich abgetastet. Der Beobachtungszeitraum entspricht wieder dem Intervall  $[0; (N - 1)t_A]$ , wobei  $t_A$  weiterhin die Abtastzeit beschreibt. Die diskrete STFT ist definiert durch [103, S. 492]

$$F_{x*}^\gamma[m, n] = \sum_{i=0}^{N-1} x_i \gamma_{mn}^*[i]. \quad (2.2.8)$$

Das zu untersuchende Signal wird damit durch ein sich verschiebendes Analysefenster betrachtet. Dabei wandert das Fenster das gesamte Signal entlang und generiert viele einzelne gefensterte Signale, die wiederum mit der klassischen FT transformiert werden [65]. Dabei beschreibt

$$\gamma_{km}^*[i] := \gamma_{i-m\Delta M} e^{-\frac{j2\pi ni}{N}} \quad (2.2.9)$$

das zeit- und frequenzverschobene Analysefenster, wobei  $\Delta M$  die Fensterbreite<sup>6</sup> angibt,  $x_i$  dem zeitdiskreten Signal entspricht und  $\gamma_i$  den Wert der Fensterfunktion an der Stelle  $i$  repräsentiert. Für die Fensterung eines Signals liegen unterschiedliche Fensterfunktionen vor. Eine der bekanntesten und häufig verwendeten Fensterfunktionen ist das Hann-Fenster, welches in Abbildung 2.15 links im Zeit- und rechts im Frequenzbereich für ein kontinuierliches Signal dargestellt ist. Mit der Form des Fensters, welche als Kosinusglocke bezeichnet wird, dämpft es das Signal an den Grenzen des Beobachtungsintervalls ab. Allerdings wird durch die Abblendung an den Fenstergrenzen auch die Amplitude und Leistung des Signals verringert, weshalb häufig mit überlappenden Fenstern gearbeitet wird. [72, S. 453], [84, S. 204]

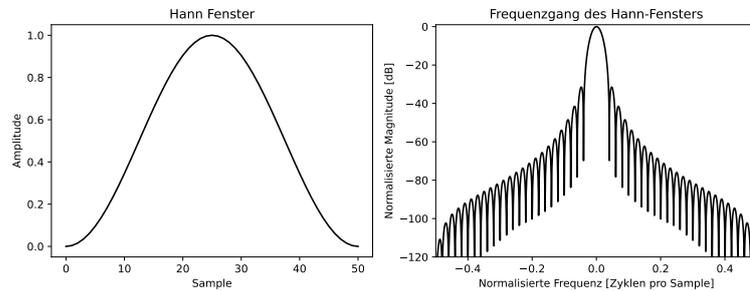


Abbildung 2.15: Fensterfunktionen Hann (vgl. [90, S. 97])

Mit dem Betragsquadrat

$$S_x^\gamma(\tau, f) = |F_x^\gamma(\tau, f)|^2 = \left| \int_{-\infty}^{\infty} x(t) \gamma^*(t - \tau) e^{-j2\pi ft} dt \right|^2 \quad (2.2.10)$$

der STFT ist die Lokalisierung des Frequenzinhalts in der Zeit möglich. Allerdings gehen bei diesem Übergang die Phaseninformation verloren. Das Diagramm von Frequenz und Zeit wird als Spektrogramm bezeichnet [11, S. 70], [103, S. 466].

Analysen im Zeitbereich können anhand der Abtastwerte des Signals durchgeführt werden, dabei handelt es sich durch die Eindimensionalität um eine einfache Analyseoption. Besonders bei akustischen Si-

<sup>6</sup>Im Folgenden wird angenommen, dass die Fensterbreite der Schrittweite der FT entspricht.

gnalen ist die Wellenform häufig das bevorzugte Format, um die Eigenschaften bzw. Charakteristika des Signals zu visualisieren [77, S. 10–11]. Die Wellenform, also die Darstellung im Zeitbereich, liefert jedoch keine Informationen über den Frequenzbereich, sodass eine Transformation notwendig wird, um diese Frequenzinformationen aus dem Signal zu extrahieren. Mit der DFT können die spektralen, also frequenzbezogene Merkmale untersucht werden, wobei hier die Zeitbezüge außer Acht gelassen werden. [43] Die Schwierigkeit in der Kombination beider Bereiche liegt in dem Bestreben, gleichzeitig hohe Auflösung im Zeit- als auch im Frequenzbereich zu erreichen. Mit dem Spektrogramm wird ein guter Mittelweg gegeben, der die Informationen der beiden Bereiche mit geringerer Auflösung darstellt. [11, S. 74] Abbildung 2.16 illustriert in der oberen Reihe die beschriebenen Auflösungsgrenzen des Zeit-, des Frequenz- und des Zeit-Frequenz Bereiches. Darunter befindet sich die Visualisierung eines aufgenommenen Sprachsignals „Hallo, das ist ein Sprachsignal.“ als Wellenform, also als Darstellung des Zeitbereiches, das zugehörige Spektrum, also die Darstellung des Frequenzbereiches und das Spektrogramm, die Kombination aus Zeit- und Frequenzbereich.

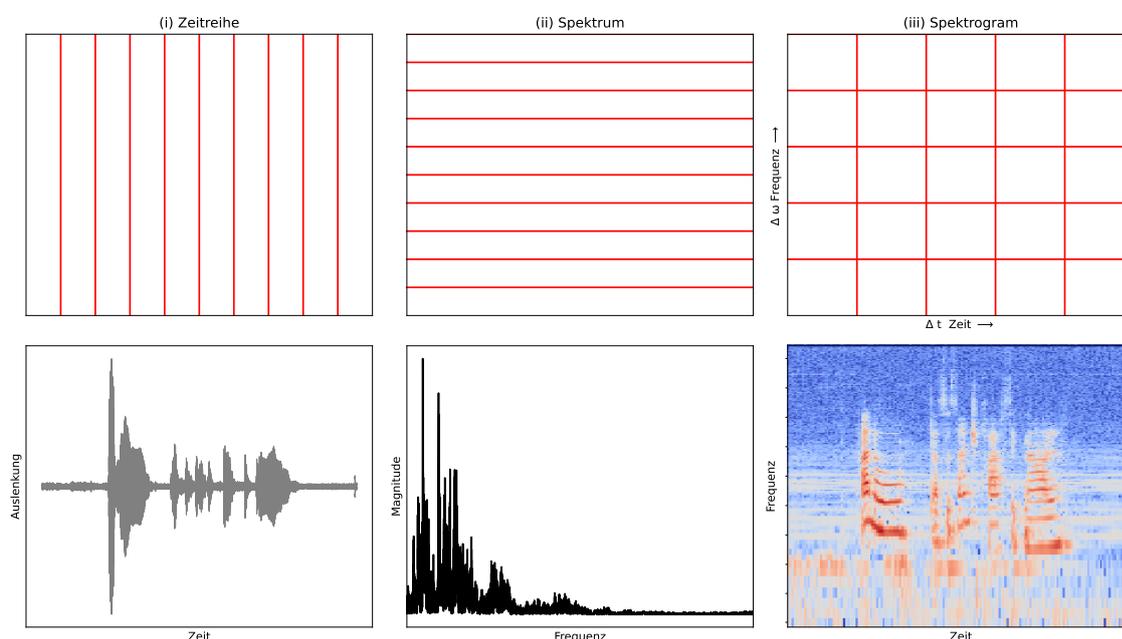


Abbildung 2.16: (i) Zeitbereich (ii) Frequenzbereich (iii) Zeit-Frequenz-Bereich (vgl. [11, S. 74])

## 2.3 Künstliche Intelligenz und Maschinelles Lernen

*Künstliche Intelligenz (KI)* ist die deutsche Übersetzung für den englischen Begriff *Artificial Intelligence (AI)*. Das Ziel von KI ist, die komplexen kognitiven Fähigkeiten des Menschen, wie zum Beispiel das Wahrnehmen, die Objekterkennung, das Lernen, das Sprachverstehen, das Schlussfolgern oder das Problemlösen zu untersuchen und bestmöglich maschinell umzusetzen [79, S. 2]. Im Wesentlichen soll ein Computer in die Lage versetzt werden, ähnliche intellektuelle Tätigkeiten wie ein Mensch ausführen zu können. Im Rahmen der KI-Forschung zielt ein intelligentes System auf die Erfüllung der folgenden fünf Aufgaben ab [40, S.15]:

1. Logisches Denken,
2. Treffen von Entscheidungen bei Unsicherheit,
3. Planen,
4. Lernen,
5. Kommunikation in natürlicher Sprache.

Technische Systeme oder Maschinen, die alle fünf Aspekte beherrschen würden, entsprächen einer universalen Intelligenz und fielen in den Bereich einer *starken Künstlichen Intelligenz*. Wird nur eine Teilmenge der oben genannten Bereiche von einem intelligenten System bedient, wird von einer *schwachen Künstlichen Intelligenz* gesprochen. Gerade diese spezialisierten Systeme einer schwachen KI werden aktuell rasant weiterentwickelt und finden Anwendung in Bereichen von Produktion, Planung [40, S. 16] oder, wie bereits erwähnt, der effizienten Nutzung von Ressourcen [27]. Hinzu kommt die Tatsache, dass von vielen technischen Geräten vermehrt intelligentes Verhalten erwartet wird. Diese intelligenten technischen Systeme erkennen eigenständig ihre Ziele und können diese aufgrund ihrer KI-basierten Programmierung auch erreichen. [79, S. 1–2] Lämmel (2020) kritisiert die Übersetzung, die die Assoziationen eines künstlichen Wesens wecken könnte, welches die Menschheit beherrscht (und nicht andersherum). Die Inhalte des Fachgebietes könnten seiner Meinung nach besser mit *gekünstelte Intelligenz*, *maschinelle Intelligenz* oder *synthetische Intelligenz* übersetzt werden [73, S. 1]. Generell sind die Entwicklung und die Nutzung von Methoden der KI mit vielen ethischen Fragen verbunden, die für die gesamte Gesellschaft relevant sind. Ein verantwortungsvoller Umgang mit Künstlicher Intelligenz setzt eine ausführliche systematische Darstellung sowie Diskussion voraus. [46, S. 22] Künstliche Intelligenz teilt sich dabei in verschiedene Teilgebiete auf; acht dieser Gebiete werden im Folgenden gelistet [79], [33]:

1. Problemlösen und maschinelles Beweisen (*problem solving and theorem proving*),
2. Spracherkennung und Verstehen natürlicher Sprache (*natural language understanding*),
3. Bildverarbeitung und Bilderkennung (*image processing and vision*),
4. Lernen (*learning*),
5. Expertensysteme (*expert systems*),
6. Qualitatives Schließen (*qualitative reasoning*),
7. Intelligente Roboter (*robotics*),
8. KI-Hardware und KI-Software.

Neben diesen Teilgebieten gibt es auch Fachgebiete, die durch ihre starke KI-Prägung direkt zur Künstlichen Intelligenz gezählt werden, wie zum Beispiel Klassifikationen und Mustererkennungen, automatisches Programmieren, kognitive Psychologie oder Computeralgebra (Formelmanipulation) [79, S. 5]. Aus den oben genannten (Teil-)Bereichen der Künstlichen Intelligenz sind die folgenden drei relevant für diese Dissertation: *Maschinelles Lernen (ML)*, *Klassifikation bzw. Mustererkennung* und *Bildanalyse in Kombination mit Ansätzen aus dem Bereich der (Audio-)Signalverarbeitung*

### 2.3.1 Maschinelles Lernen

Künstliche Intelligenz und Maschinelles Lernen haben große Überschneidungen, denn viele integrierte Algorithmen des ML werden im Bereich der KI eingesetzt [40, S. 16], [11, S. 154]. Die Algorithmen des Maschinellen Lernens werden als interdisziplinäre Disziplin aus Mathematik und Informatik verstanden und basieren auf Verfahren der Optimierungsrechnung [11, S. 154], [40, S. 329]. Dabei sind Verfahren des Maschinellen Lernens die Grundlage von Programmsystemen, die aus vorgegebenen Daten „lernen“ sollen. Ausgehend von der Datenlage können Muster und Zusammenhänge „erkannt“, Tatsachen- und Regelwissen „erlernt“, oder Priorisierungen adaptiert werden [46, S. 20]. Im Vergleich zur klassischen Programmierung, bei der der Mensch die Regeln implementiert, die wiederum auf die Eingabedaten angewendet werden, um damit die Ausgabe zu generieren, werden beim Maschinellen Lernen Daten zur Verfügung gestellt und die Regeln sind das Ergebnis. Mit diesen vom Maschinellen Lernen erstellten Regeln können unbekannte Daten untersucht und entsprechende Antworten eigenständig ermittelt werden, siehe Abbildung 2.17. [16, S. 23]

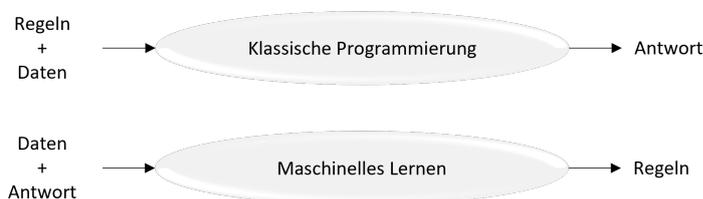


Abbildung 2.17: Klassische Programmierung vs. Maschinelles Lernen (vgl. [16, S. 23])

Mit den Algorithmen des Maschinellen Lernens können verschiedene Ziele verfolgt werden, wodurch auch die Regeln für jede Problemstellung individuell ausfallen. Im Bereich des Data Minings werden die Maschinellen Lernmethoden zur Extraktion aussagekräftiger Merkmale genutzt, es wird vom *Feature Engineering* gesprochen [134, S. 43]. Des Weiteren können Daten generierten Gruppen zugeordnet werden, wenn sie aussagekräftigen und eindeutig erkennbaren Mustern entsprechen. Die Mustererkennung ist dem Fachbegriff *Clustering* zugeordnet. Solche Methoden können beispielsweise bei wichtigen Entscheidungsprozessen unterstützen. [11, S. 154] In industriellen bzw. wirtschaftlichen Bereichen können Maschinelle Lernmethoden auf Prozessdaten angewendet werden, um Merkmale von Prozessen digital zu simulieren oder unbekannte Muster bzw. Regelmäßigkeiten zu identifizieren [2, S. 2]. Ausgehend von der Problemstellung können ML-Methoden in drei Lernkategorien unterschieden werden:

1. Überwachtes Lernen (*Supervised Learning*),
2. Bestärkendes Lernen (*Reinforcement Learning*),
3. Unüberwachtes Lernen (*Unsupervised Learning*).

#### Überwachtes Lernen

Beim *überwachten Lernen* benötigt das Maschinelle Lernmodell sowohl die Eingabe- als auch die Ausgabewerte. Die Daten werden in *funktional unabhängige (Feature)* und *funktional abhängige (Label)* Daten unterteilt. Das funktional abhängige Merkmal entspricht den Ausgabewerten, die von den funktional unabhängigen Variablen abhängen. Einem Modell des überwachten Lernens ist damit das Ergebnis bereits

bekannt. Ziel des Modells ist es, die vorliegenden Zusammenhänge, die zwischen Ein- und Ausgabewerten bestehen, zu erfassen und zu erlernen. Daten, die bereits mit ihrem Ergebnis vorliegen, werden auch als *gelabelte Daten* bezeichnet [16, S. 50]. Hat ein Modell die Zusammenhänge erlernt und ist generalisierungsfähig, kann es für Eingabedaten, die ohne Ausgabe vorliegen, entsprechende Ausgabewerte prognostizieren. [133, S. 23] Beim überwachten Lernen wird zwischen *Klassifikation* und *Regression* unterschieden. Regressionsprobleme haben eine kontinuierliche Größe, also eine reelle Zahl, als Output [89, S. 28]. Klassifikationen hingegen ordnen Daten in vorgegebene Kategorien. Diese Kategorien einer Klassifizierungsaufgabe werden auch als *Klassen* bezeichnet [16, S. 50]. Liegen nur zwei Klassen vor, wird von einer *binären Klassifikation* gesprochen [89, S. 27].

### Bestärkendes Lernen

Bei Problemstellungen, bei denen ein Algorithmus lernen muss, wie er sich zum Lösen eines Problems „verhalten“ muss, kommt das *bestärkende Lernen* (engl. *Reinforcement Learning*) zum Einsatz [141, S. 6]. In vielen Fällen wird dabei auch auf Methoden des überwachten Maschinellen Lernens zurückgegriffen. Methoden des bestärkenden Lernens basieren meist auf einem Softwareagenten, der anhand von positiven bzw. negativen Rückmeldungen des Systems Schlüsse für das Lernverhalten zieht. Der Softwareagent kann aus den Prognosen des überwachten Lernens des Systems eine Nutzenfunktion ableiten, die ihm wiederum prognostiziert, welche Aktion voraussichtlich den höchsten Nutzen erzielt. [40, S. 24–25] Dabei geht es nicht um eine einzelne Entscheidung, sondern um die Folge von richtig getroffenen Entscheidungen, denn auch komplexe Ziele sollen möglichst optimal erreicht werden können. Ein gutes Beispiel ist das Spiel Schach: Viele Regeln gibt es nicht, doch entscheidend für den Spielerfolg ist die Aufeinanderfolge von richtig gewählten Zügen. [2, S. 14]

### Unüberwachtes Lernen

Liegt einem Maschinellen Lernmodell eine Datenbasis weder mit zugehörigen Zielwerten (Labeln) vor, noch besteht die Möglichkeit, die Ergebnisse der Prognose positiv bzw. negativ zu bewerten, kommen Methoden des *unüberwachten Lernens* ins Spiel. Diese sind darauf ausgelegt, eine gegebene Datenmenge auf unbekannte Strukturen und Muster zu untersuchen bzw. diese zu finden. [40, S. 25] Bekannte unüberwachte Lernmethoden sind z.B. Cluster- und Faktorenanalysen oder die Hauptkomponentenanalyse (engl. *Principle Component Analysis* (PCA)) [48, S. 15].

### Daten

Die Grundlage für jedes datengetriebene Modell sind Daten. Dabei ist der Kontext zu den vorliegenden Daten wichtig. Fehlt der Bezug bzw. der Kontext, unter dem vorliegende Daten analysiert werden sollen, können die Ergebnisse zumeist nicht zielführend interpretiert werden. Erst vor dem Hintergrund einer klar formulierten Zielsetzung ist die Extraktion entsprechender Information aus vorliegenden Daten möglich. Gleiches gilt für die Datenqualität, die ebenfalls im Kontext einer Problem- oder Aufgabenstellung beurteilt werden muss. [56, S. 87–88] Getreu dem Motto *Garbage in, Garbage out* spielt die Datenqualität eine entscheidende Rolle hinsichtlich der Prognoseergebnisse [56, S. 93]. Die Datenvorbereitung ist demnach ein wesentlicher Schritt, der durchgeführt werden sollte, bevor die Daten als Eingabe für ein Maschinelles Lernmodell genutzt werden. Zwei generelle Punkte, die zu berücksichtigen sind, sind die *Selektion* und die *Integration* der Daten. Unter der Selektion wird das Zusammenstellen genau der Datenmenge verstan-

den, die für die Aufgabenstellung von Nutzen ist. Die Integration beschreibt die Zusammenstellung aus unterschiedlichen Quellen. Weiter können vorliegende Daten auf ihre *Vollständigkeit* bzw. *fehlende Werte* untersucht werden. Es gibt viele Möglichkeiten mit Unvollständigkeiten umzugehen, um die Datenbasis qualitativ zu verbessern, denn diese hat erheblichen Einfluss auf die letztendliche Prognose des Modells. [73, S. 233]

Daten können in strukturierte und unstrukturierte Daten unterschieden werden. *Strukturierte Daten* liegen meist tabellarisch vor, denen strukturierte Informationen zugrunde liegen. Diese Struktur wird implizit durch die Daten vorgegeben [56, S. 146]. Die Spalten einer solchen Datenmenge beschreiben jeweils ein Merkmal, welches häufig auch als *Feature*, *Variable* oder *Einflussgröße* bezeichnet wird. Die zugehörigen Zeilen enthalten zu den entsprechenden Merkmalen numerische Werte oder Informationen. Eine Zeile, die die Werte aller Features enthält, wird als vollständiger Datensatz bezeichnet. [40, S. 18] Unter *unstrukturierten Daten* werden solche Daten verstanden, aus denen zunächst keine eindeutigen Informationen hervorgehen. Für die Interpretation der Daten ist hier die Interpretation eines Informationsempfängers notwendig. [56, S. 146–147] Beispielsweise bestehen Bilder aus Pixeln, die zwar strukturiert beschrieben werden können, denn jedes Pixel hat eine feste Position und enthält einen Farbwert; allerdings wird nicht übermittelt, was das Bild zeigt. Gleiches gilt für freie Texte oder Emails: Auch hier liegen Informationen über Buchstaben und Wörter vor, der Inhalt ist maschinell jedoch nicht extrahierbar und damit unstrukturiert. [40, S. 18–19]

### 2.3.2 Neuronale Netze

Die Anfänge *Neuronaler Netze* (NN) liegen in den 1940er Jahren und basieren auf den Forschungen von Warren McCulloch und Walter Pitts [34]. Ziel dieser Forschung war die Entwicklung eines einfachen mathematischen Modells, welches die Funktionsweise und das Verhalten eines menschlichen Neurons im Gehirn erklären sollte [40, S. 168]. Das menschliche Gehirn ist in der Lage, Millionen von zusammengeschalteten Nervenzellen zu trainieren und Steuerungsaufgaben zu übernehmen [73, S. 190]. Gerade in diesem Bereich sind die komplexesten biologischen Strukturen auch heute noch nicht vollständig erforscht und nachvollziehbar [40, S. 168]. Der Vergleich eines Künstlichen Neuronalen Netzes mit den Funktionsweisen eines menschlichen Gehirns ist daher nur bedingt richtig. Dennoch wird mit Maschinellen Lernmodellen versucht, Computerprogramme bzw. Algorithmen zu entwickeln, die in analoger Weise lernen können und Gelerntes auch anwenden sollen [73, S. 190]. Als Teil der KI können Neuronale Netze sowohl den Ingenieurwissenschaften als auch den Erkenntniswissenschaften zugeordnet werden, mit den jeweiligen Zielen der kognitiven Modellierung, also der Simulation kognitiver Prozesse, die anhand ihrer Informationsverarbeitung nachvollzogen werden, als auch zur Konstruktion intelligenter Systeme, die sich den menschlichen Wahrnehmungs- und Verstandsleistungen maschinell nähern. [46, S. 2]

Generell können Künstliche Neuronale Netze (KNN) auf ein verallgemeinerndes Verhalten trainiert werden und sind damit nicht nur in der Lage die Trainingssituation, sondern auch darüber hinaus unbekannte Daten zu prognostizieren. Erreicht wird diese Fähigkeit durch das Zusammenschalten einer großen Anzahl künstlicher Neuronen. [73, S. 191] Wie sich ein solches künstliches Neuron aufbaut wird in Abbildung 2.18 dargestellt.

Ein künstliches Neuron ist auch unter dem Begriff des *einlagigen Perzeptrons* bekannt. Es wurde in den 1950er Jahren als erste Idee eines Künstlichen Neuronalen Netzes entwickelt [73, S. 190]. Die Funktionsweise lehnt sich an den Informationsaufnahmeprozess zum Beispiel des menschlichen Sehens an. Im ersten Schritt wird etwas gesehen und die Information wird zum Gehirn weitergegeben, dort analysiert

und als das entsprechende Objekt erkannt.

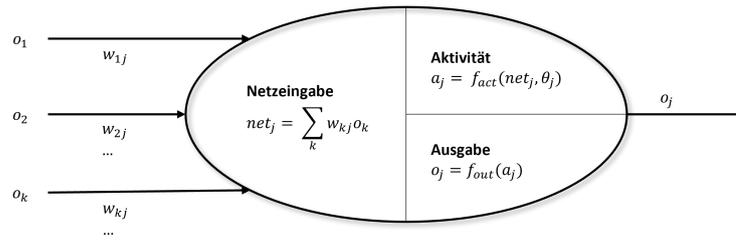


Abbildung 2.18: Künstliches Neuron  $j$  (vgl. [73, S. 192])

Das einlagige Perzeptron (Single Layer Perceptron) adaptiert dieses Vorgehen. Es gibt also einen Input, eine verarbeitende Position und einen entsprechenden Output. Mit dieser einen Eingangs- und Ausgangsschicht entspricht das einlagige Perzeptron einem einfachen vorwärts gerichteten NN. [73, S. 204] Der Input des Neurons sind die Eingabeinformationen  $o_k$ , die über gewichtete Verbindungen  $w_{kj}$  in das Neuron geleitet werden. Aus diesen beiden Werten, die pro Verbindung in das Neuron eingehen, setzt sich die Netzeingabe

$$net_j = \sum_k w_{kj} \cdot o_k \quad (2.3.1)$$

additiv zusammen. Es wird von der *Propagierungsfunktion* gesprochen, die die Summe über alle eingehenden gewichteten Netzverbindungen bildet. Im nächsten Schritt wird eine *Aktivierungsfunktion* auf die errechnete Propagierungsfunktion angewendet. Unter Berücksichtigung eines festzulegenden Schwellenwertes  $\theta_j$  wird eine neue Aktivität

$$a_j = f_{act}(net_j, \theta_j) \quad (2.3.2)$$

bestimmt, die festlegt, wie stark das künstliche Neuron aktiviert wird. Mit dieser Aktivität kann die finale Ausgabe des Neurons

$$o_j = f_{out}(a_j) \quad (2.3.3)$$

bestimmt werden, welche wiederum als Eingabeinformation für das nächste Neuron innerhalb eines Netzes dient.

Die Wahl der Aktivierungsfunktion ist abhängig von der Problemstellung. Im Allgemeinen unterstützt sie die Schaltfunktion eines Neurons und hat einen mehr oder weniger sigmoiden Charakter<sup>7</sup>. Folgende Aktivierungsfunktionen werden häufig verwendet [73, S. 192–193], [133, S. 12]:

1. Lineare Funktionen,
2. Gleichrichter-Funktion,
3. Logistische Funktion,
4. Tangens Hyperbolicus,

<sup>7</sup>Wenn eine Funktionskurve S-förmig verläuft, wird sie als sigmoid bezeichnet.

## 5. Schwellwertfunktion.

Dieser spezielle und sehr einfache Aufbau des einlagigen Perzeptrons hat nur eine trainierbare Verbindungsschicht [73, S. 204] und das Training erfolgt über eine festzulegende Anzahl von Lerniterationen [133, S. 14–15].

In den 1980er Jahren wurde das einstufige Perzeptron zu einem mehrschichtigen bzw. *mehrlagigen Perzeptron* (Multilayer Perceptron (MLP)) weiterentwickelt. MLPs zählen zu den Künstlichen Feedforward-Netzen, die mindestens aus drei Schichten bestehen. Ergänzend zu der Ein- und Ausgabeschicht gibt es demnach mindestens eine weitere, die so genannte *verborgene Schicht* (engl. Hidden Layer). Jede Schicht ist nur mit der jeweils nächsten Schicht verbunden. Durch diesen Aufbau kommen MLPs anhand mehrerer Verarbeitungsschritte zu einer Entscheidung. Das gilt sowohl für Regressions- als auch für Klassifikationsprobleme. [40, S. 181], [133, S. 16–17], [89, S. 99] Der Aufbau eines mehrschichtigen Perzeptrons wird in Abbildung 2.19 verdeutlicht, wobei die Neuronen im Netzwerk als Knoten und die Verbindungen als gerichtete Kanten dargestellt werden. Der Aufbau des Netzwerkes und die Struktur der Verarbeitungseinheiten und ihrer Verbindungen lässt sich durch den Begriff der *Netzwerktopologie* beschreiben. Aus der Netzwerktopologie können z.T. auch Rückschlüsse auf den Netztyp oder auf die Fähigkeiten zur Informationsverarbeitung gewonnen werden [17, S. 179].

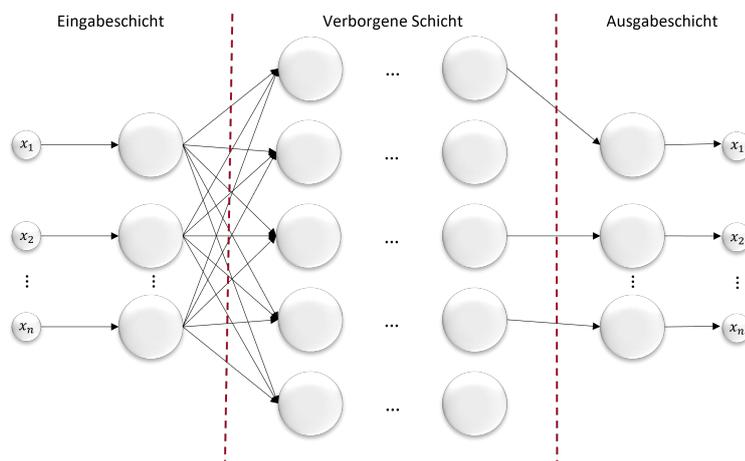


Abbildung 2.19: Mehrlagiges Perzeptron (vgl. [133, S. 16])

Der wesentliche Vorteil eines MLPs gegenüber dem einlagigen System liegt in der Fähigkeit jede stetige Funktion zu approximieren (Universal Approximation Theorem). Dafür sind MLPs prinzipiell in der Lage, jede berechenbare Funktion zu realisieren [73, S. 204], was sie für Deep Learning-Anwendungen, auf die im nächsten Kapitel eingegangen wird, attraktiv macht [73, S. 222].

Um die Verbindungsgewichte eines vorwärts gerichteten Netzes mit mindestens einer verdeckten Schicht anzupassen und dafür zu sorgen, dass das Netz aus den Inputdaten lernt, wird der Backpropagation-Algorithmus, ein Verfahren zur Fehlerrückführung, genutzt. Gegenüber einem einlagigen Perzeptron liegt das Problem bei MLPs in den Fehlerwerten der verdeckten Schicht(en), da sich diese nicht bestimmen lassen. Bekannt sind nur die vorliegenden Ausgabewerte, die als Referenz für den Fehler der prognostizierten Ausgabe des Netzes dienen. Ausgehend von der auftretenden Abweichung werden bei der Backpropagation die Gewichte rückwärts von der Ausgabe- in Richtung Eingabeschicht angepasst. [73, S. 211, 213] Der Algorithmus zum Trainieren des Modells umfasst die folgenden drei Schritte [133, S. 17]:

1. Die Gewichtsmatrizen der Schichten werden zunächst mit Zufallswerten initialisiert, um die Eingabedaten durch das Netz zu leiten.
2. Die berechneten Ausgabewerte werden mit den gegebenen Werten verglichen und die Abweichung als Fehler gespeichert. Wird ein festgelegter Schwellenwert überschritten, wird der nächste Schritt durchgeführt. Bleibt die generelle Abweichung unter diesem Wert ist das Training beendet.
3. Ausgehend von der Ausgabeschicht läuft die Fehlerrückmeldung bis zur Eingabeschicht zurück. Durch die Modifikation der Gewichtsmatrizen wird so der Fehlerwert sukzessive minimiert.

Für diese Minimierung nutzt der Backpropagation-Algorithmus das Verfahren des Gradientenabstiegs. Mit diesem Verfahren wird versucht, ein (lokales) Optimum zu finden, indem der Abstieg der Fehlerfunktion in abfallender Richtung bestimmt wird [133, S. 18].

### 2.3.3 Deep Learning

*Deep Learning* (DL) beschreibt Lösungen, die mehrschichtige Künstliche Neuronale Netze implementieren und damit in der Lage sind, schrittweise Informationen aus Eingabedaten zu extrahieren, um eine Ausgabe zu prognostizieren. Die Eingabedaten durchlaufen dabei unterschiedliche Schichten des Netzes, in denen entsprechende Merkmale extrahiert und gelernt werden. [133, S. 22] Die Darstellung der Schichten ist auch unter dem Begriff *Kompositionalität* bekannt. Das bedeutet, dass komplexe Daten als Kombination von elementaren Komponenten dargestellt werden. Diese Komponenten werden immer weiter in noch einfachere Komponenten zerlegt, bis eine atomare Einheit als Ergebnis vorliegt. Gerade diese Fähigkeit, Komplexität mit kompositorischen Darstellungen zu handhaben, macht Deep Learning so wirkungsvoll [141, S. 5]. DL wird häufig als eigenes Gebiet der KI ausgewiesen, es stellt allerdings nur ein Teilgebiet dar, wie in Darstellung 2.20 zu sehen ist. Deep Learning wird demnach als Teilgebiet den Künstlichen Neuronalen Netzen zugeordnet, die wiederum einen Teil des Maschinellen Lernens darstellen, welcher zur KI gehört [133, S. 23].

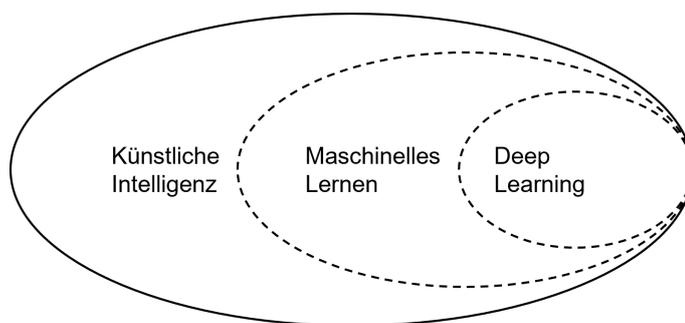


Abbildung 2.20: Teilgebiete der Künstlichen Intelligenz (vgl. [16, S. 22])

Mit DL können viele komplexe Problemstellungen angegangen werden, die mit flachen Künstlichen Neuronalen Netzen (engl. shallow neural networks) nicht gelöst werden können. Darunter fallen unter anderem Bildklassifikations- und Vorhersageaufgaben [141, S. 6], die mit einem Deep Learning-Modell, nämlich den Convolutional Neural Networks (CNN), gelöst werden können. Die verwendeten Algorithmen sind zielgerichtet auf den Anwendungsfall zuzuschneiden und bieten vielversprechende Lösungsansätze [89, S. 99]. Das Feature Engineering, also das Extrahieren von Merkmalen aus gegebenen Datenmengen,

welches bei klassischen Machine-Learning-Algorithmen spezielles Domänenwissen voraussetzt, ist im Deep Learning nicht notwendig. [133, S. 2].

### 2.3.4 Convolutional Neural Networks

Der Ursprung des *Convolutional Neural Networks* (CNN) geht auf die Forschung von Yann LeCun [75] aus dem Jahr 1989 zurück, in der eine Netzwerktopologie beschrieben wurde, die aus einer großen Anzahl hintereinandergeschalteter Schichten (engl. Layer) besteht [133, S. 23]. CNN sind vorwärts gerichtete Netze, die auch häufig unter der Abkürzung *ConvNet* in der Literatur zu finden sind. Diese Art von Netz wurde speziell für die Bildverarbeitung entwickelt, da einfache vorwärts gerichtete Netze bei der Mustererkennung schnell an ihre Grenzen stoßen. Gerade im Bezug auf die räumliche Position eines zu erkennenden Objektes können flache NN gegenüber CNN nicht mithalten. Neben der Bildklassifikation werden CNNs auch für Audioklassifizierungen oder die Analyse von Zeitreihen (engl. time series) eingesetzt. [73, S. 252], [40, S. 352], [133, S. 26], [57, S. 51]

Diese Art von Netz verdankt den Namen „Convolutioal Neural Network“ seiner zugrundeliegenden Faltungsoperation, wobei Faltungen (engl. convolution) in der Mathematik, speziell in der Funktionalanalysis, eingesetzt werden [73, S. 252]. Neben der Faltungsoperation kommt es bei den CNNs auch auf die gesamte Netzarchitektur an. So wird die Abfolge von aufeinander folgenden Faltungs- und Poolingschichten als eine spezielle Form der Bildaufbereitung oder Datenvorbereitung betrachtet und genutzt, um automatisch Merkmale zu generieren, die den räumlichen Bezug eines Objektes berücksichtigen. Diesem komplexen Aufbau folgt eine einfache klassische Netztopologie, die für die finale Klassifikation verantwortlich ist. [40, S. 352], [73, S. 258] Zusammenfassend besteht ein CNN grundlegend aus folgenden Schichten [73, S. 253]:

1. Faltungsschicht (Convolutional Layer),
2. Pooling Layer (Zusammenfassung),
3. Voll vernetzte Schichten (Klassifikation).

#### Faltungsschicht (Convolutional Layer)

Die Funktionsweise der *Faltungsschicht* lässt sich gut an dem Beispiel einer Bilddatei erklären. Computer erfassen Bilder anhand von Pixeln, die durch eine  $n \times m$ -Matrix sowohl die Höhe als auch die Breite des Bildes beschreiben. Jedem Pixel kann ein spezieller Farbwert zugeordnet werden, der sich z.B. aus den Farbkanälen Rot, Grün und Blau (RGB) in Form von drei Farbtintensitäten zusammensetzt. [40, S. 19]

Die Faltungsschicht im CNN, also der Convolutional Layer, wird zur Erkennung bzw. zum Lernen der vorhandenen lokalen Muster der Bilder verwendet [16, S. 164]. Die zu lernenden Bilddaten werden der Faltungsschicht in Form von Zahlenwerten als Eingangsdaten übergeben. Mit einem Filter bzw. einer Faltungsmatrix mit deutlich geringerer Größe als die Eingabematrix (z.B.  $3 \times 3$  oder  $5 \times 5$ ) werden die Muster der Eingabedaten gefiltert. Das Vorgehen ist in etwa vergleichbar mit einer Schablone, die mit einem kleinen Sichtfenster über ein Bild bewegt wird, um nur den zu sehenden Bildausschnitt zu betrachten. Die Faltungsmatrix bzw. die Schablone einer jeweiligen Schicht wird von links oben bis nach rechts unten über die jeweiligen Bilder geführt. Mit diesem Vorgehen wird immer nur ein bestimmter Ausschnitt des Bildes verarbeitet. Für jede Position der Faltungsmatrix wird ein Ergebniswert für die Ergebnismatrix berechnet. Dazu werden die Werte der Bilddatei mit den Zahlenwerten der Faltungsmatrix multipliziert und zu einem Ergebniswert addiert, der an die Ergebnismatrix übergeben wird. Die Zahlenwerte der

Faltungsmatrix werden dabei initial zufällig gewählt und während des Trainings fortwährend angepasst. Der Mittelpunkt einer Faltungsmatrix wird auch als *Kern* bzw. im englischen als *Kernel* bezeichnet. [73, S. 254], [133, S. 24] Die Ergebnismatrix der gesamten Operation wird auch als *Feature Map* bezeichnet [40, S. 354]. Abbildung 2.21 verdeutlicht die Vorgehensweise der Faltungsschicht, wobei rot umrandet die Faltungsmatrix dargestellt wird, die im ersten Schritt links oben auf die Ecke des Bildes gesetzt wird und sich in den folgenden Schritten immer weiter über das Bild bewegt. Im ersten Schritt wird für das gezeigte Beispiel ein Wert von  $5 \cdot 92 + (-1) \cdot 4 + (-1) \cdot 16 + 0 \cdot 21 = 440$  in die resultierende Feature-Map eingetragen.

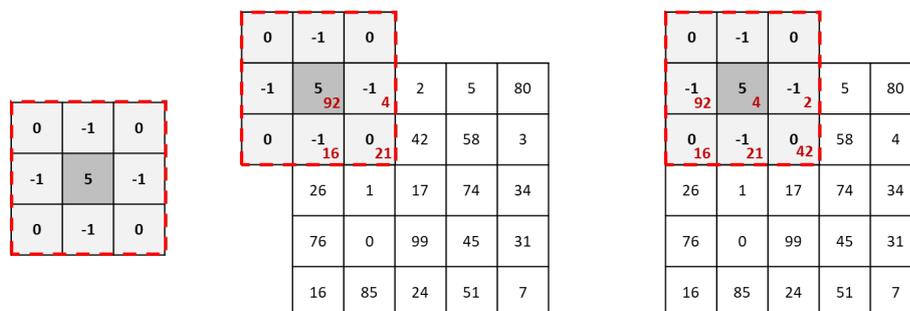


Abbildung 2.21: Faltungsschicht (vgl. [133, S. 24])

Die ursprünglichen Eingabedaten werden durch die Faltungsschicht und die verwendeten Faltungsmatrizen so verändert, dass in der Ergebnismatrix unterschiedliche Eigenschaften (Features) der Eingabedaten betont werden. Wichtige Merkmale bzw. Features eines Bildes, die für eine spätere Klassifikation von zentraler Bedeutung sein könnten, können so in den Faltungsschichten herausgearbeitet werden. Diese Merkmale alleine reichen jedoch für eine finale Klassifikation noch nicht aus. Es fehlt der Schritt der Informationsverdichtung und Verallgemeinerung. Dieser Verarbeitungsschritt wird mit dem Pooling erreicht. [40, S. 356]

### Pooling Layer (Zusammenfassung)

Innerhalb eines CNNs folgt der Faltungsschicht in der Regel eine *Pooling-Schicht*. In dieser Schicht werden die Ergebnisse der Faltung durch das so genannte Pooling reduziert und überflüssige Informationen verworfen [133, S. 25]. Dafür wird zunächst die Ergebnismatrix bzw. die Feature Map der Faltung an die Pooling-Schicht übergeben. Die Verdichtung der Informationen, auch *Downsampling* genannt [16, S. 170], kann mit den folgenden zwei Operationen durchgeführt werden:

1. Average Pooling,
2. Max Pooling.

Ein z.B.  $2 \times 2$  großer Pooling-Filter wird für die Pooling-Operation über die Feature Map gelegt und reduziert die Matrix damit um 75 % ihrer ursprünglichen Dimension [133, S. 121]. Abbildung 2.22 stellt einen  $2 \times 2$  Pooling-Filter auf einer  $4 \times 4$  Pixel großen Feature Map dar und verdeutlicht die beiden möglichen Pooling-Optionen. Beim Max Pooling wird jeweils der Maximalwert innerhalb des Pooling-Filters gesucht und an die Ausgabematrix der Pooling-Schicht weitergeben. In Abbildung 2.22 beträgt der Maximalwert 40. Beim Average Pooling wird hingegen der Durchschnitt aller Werte innerhalb des  $2 \times 2$ -Filters zur Übergabe an die Feature Map berechnet:  $\frac{1}{4} \cdot (21 + 40 + 6 + 13) = 20$ . [73, S. 256–257]

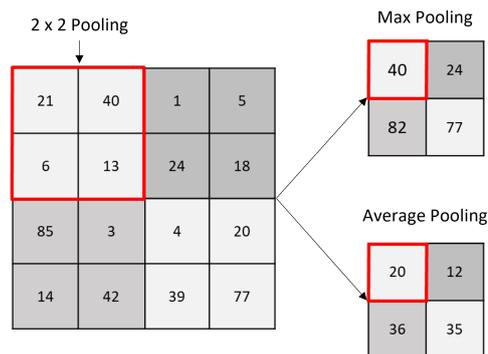


Abbildung 2.22: Pooling-Operationen (vgl. [133, S. 121])

Die Feature Map, die dem Pooling übergeben wird, enthält alle Eigenschaften der Ausgangsdaten. Zu diesem Zeitpunkt ist diese Feature Map noch empfindlich gegenüber der Position eines zu erkennenden Objektes in den Bilddaten. Durch das Pooling werden die Eigenschaften des Bildes summiert und komprimiert weitergegeben. Die Verdichtung unter gleichzeitiger Beibehaltung wesentlicher Informationen der Ausgangsdaten wirkt gegen diese räumliche Empfindlichkeit. Des Weiteren hat die Dimensionsreduktion einen positiven Effekt auf die Performance des Netzes beim Training. [73, S. 256–257] Im Allgemeinen wird die Abfolge von einer Convolutional- und einer Pooling-Schicht in der Architektur eines CNN mehrfach wiederholt. Dabei gilt es, die genaue Anzahl der Wiederholungen stets individuell auf die vorliegende Problemstellung anzupassen. [40, S. 357]

### Voll vernetzte Schicht

Den letzten Teil eines CNNs bildet ein klassisches einfaches Neuronales Netz, also ein MLP. Dieses besteht aus einer oder mehr voll vernetzten Zwischenschichten (hidden layer) und einer voll vernetzten Ausgabeschicht [73, S. 257]. In der Literatur wird dieser voll vernetzte Teil des CNNs auch *Fully-connected-Layer*, *Dense-Layer* oder *Perzeptron-Schicht* genannt. Diese Schicht nimmt am Ende des Faltungs- und Pooling-blocks die Feature Map der letzten Pooling-Schicht für die finale Klassifikation entgegen. Dies stellt eine gute Möglichkeit dar, nichtlineare Kombinationen der Merkmale des CNNs zu lernen. [133, S. 25] Die Größe des Fully-connected-Layers bzw. die feste Anzahl von Neuronen richtet sich nach der Anzahl der Klassen, die das Netz prognostizieren soll und die bereits zu Beginn des Trainings feststeht [73, S. 257], [133, S. 121–122]. Zur Berechnung der prozentualen Wahrscheinlichkeit der Zugehörigkeit zu einer Klasse wird die Softmaxfunktion häufig als letzte Schicht gewählt [133, S. 25], [73, S. 257]. Softmax verstärkt dabei den Maximalwert gegenüber den anderen prognostizierten Werten, wodurch das Neuronale Netz gezwungen wird, eine klare Aussage hinsichtlich der Prognose zu treffen [134, S. 102].

*Recurrent Neural Networks* (RNN), eine andere Architektur von Neuronalen Netzen, die insbesondere für Aufgaben der Sequenzverarbeitung geeignet sind, bilden in dem hier vorliegenden Fall eine mögliche Alternative zu den CNNs [16, S. 288]. Sie können unter anderem für die Regression von Zeitreihen, zur Klassifikation von Zeitreihen, zum Aufspüren von Anomalien in Zeitreihen und zur Kennzeichnung von Sequenzen eingesetzt werden [16, S. 298]. Im Rahmen dieser Dissertation wurden die Prozessgeräusche von Regenerationen aufgenommen, womit die Audiodaten einer zeitlichen Sequenzen entsprechen. Die Untersuchungen hinsichtlich des Qualitätsstatus fokussierten sich ebenfalls auf eine zeitliche Prognose;

nämlich, wann sich der Status des Regenerates im laufenden Prozess zu Neusand wandelt. Damit stellt sich die Frage, warum CNNs anstelle von ebenso möglichen RNNs verwendet wurden. Ein RNN kann anhand einer Schleife beschrieben werden, die die beim vorangegangenen Durchlauf der Schleife berechnete Größen wiederverwendet [16, S. 254]. Im Bezug auf Bilddaten bedeutet das, dass sich ein RNN beispielsweise die erste Pixelreihe des Bildes vornimmt und sich relevante Aspekte in dieser Reihe „merkt“. Mit diesen Informationen nimmt sich das Modell jede einzelne Pixelreihe vor und speichert pro Pixelreihe weitere relevante Aspekte. Dieses Vorgehen ist sehr rechenintensiv und es ist bekannt, dass CNNs eine schnellere Alternative bieten. Zudem wird ein Bild durch ein CNN immer wieder als Ganzes betrachtet. Nach jeder Faltung wird nach neuen Informationen innerhalb des Bildes gesucht, sodass es immer wieder als Ganzes betrachtet wird und nicht nur einmalig reihenweise nach Informationen gescannt wird. Aus den genannten Gründen wurden für die durchgeführten Analysen CNNs verwendet.

Ein beispielhafter Aufbau eines solchen CNNs stellt Abbildung 2.23 dar. Hier werden zwei Blöcke der Convolution- und Pooling-Kombination hintereinander geschaltet, bevor die Daten der voll vernetzte Schicht übergeben werden.

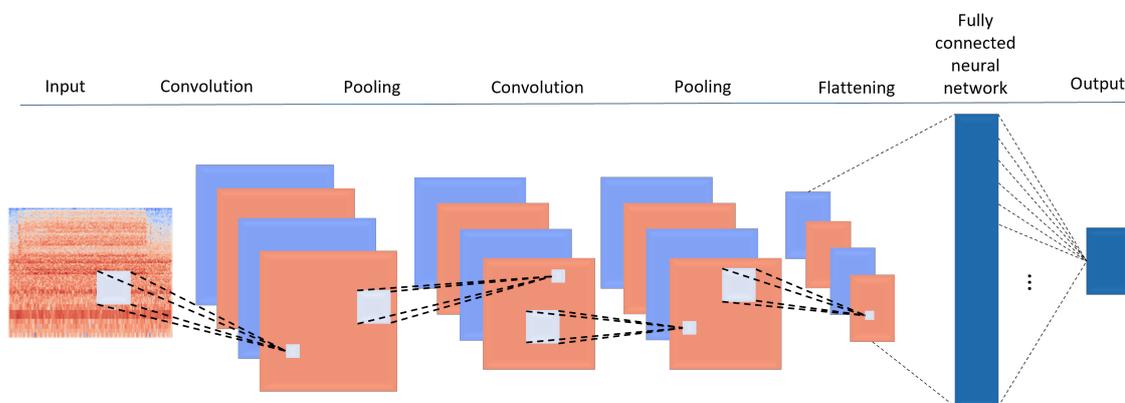


Abbildung 2.23: Möglicher Aufbau eines CNN

### 2.3.5 Modellbewertung und -optimierung

Das Ziel beim Einsatz neuronaler Netze, bezogen auf die Bildklassifikation, besteht darin, Bilder oder Objekte in Bildern korrekt zu klassifizieren, die nicht Bestandteil der Trainingsmenge des Modells waren, also zu generalisieren [73, S. 227]. Die Herausforderung bei der Modellauswahl und -bewertung besteht darin, dass die Minimierung des empirischen Risikos für eine Verallgemeinerung nicht gewährleistet werden kann. [98, S. 200]. Unter der Verallgemeinerung eines Modells wird die gleichbleibende Leistung dieses Modells auf unbekanntem Daten verstanden [16, S. 133], [89, S. 235]. Dieses Verhalten ist nicht zwangsläufig gegeben, sondern hängt von der Netzarchitektur, insbesondere von der Anzahl der Verbindungen zwischen den einzelnen Neuronen ab. Mit einer zu groß gewählten verdeckten Schicht kann es vorkommen, dass sich das Netz jedes einzelne Trainingsmuster merkt. Es lernt die Muster demnach auswendig und verallgemeinert diese nicht [73, S. 241]. Würde ein Modell auf der gesamten vorliegenden Datenbasis trainieren, wäre es damit in der Lage, für jeden Datensatz in dieser Datenmenge die korrekte Antwort zu prognostizieren; aber eine Aussage über die Generalisierungsfähigkeit kann so nicht getroffen werden [89, S. 17]. Daher ist die Auslegung und Optimierung eines neuronalen Netzes ein wesentlicher Faktor für die Performance.

Einflussfaktoren des Generalisierungsfehlers können folgende sein [73, S. 240–241], [89, S. 111], [40, S. 179]:

1. Die **Netzgröße**, die insbesondere durch die Zahl der Neuronen in den inneren Schichten in Abhängigkeit von der Zahl der zu lernenden unterschiedlichen Kategorien bestimmt wird.
2. Die **Anzahl und Art der Verbindungen** und deren zugehörigen Gewichte.
3. Das **Training** des Netzes, bei dem die Lernverfahren, die Lernparameter und die Zahl der Lernschritte das Ergebnis beeinflussen.

Zunächst wird das Training eines Modells hinsichtlich der Relevanz der Datenaufteilung in unterschiedliche Datenmengen betrachtet. Danach wird der Fokus auf eine dieser Teilmengen, nämlich die Validierungsmenge, gelegt, für die es verschiedene Auswahlmethoden gibt, die sich ebenfalls auf die Modellqualität auswirken können. Ergänzend wird zur Modelloptimierung auf zwei weitere mögliche Modell-Schichten eines CNNs eingegangen; zum einen auf die Dropout-Funktion und zum anderen auf verschiedene Aktivierungsfunktionen. Das für DL-Modelle wichtige Datenvolumen wird abschließend behandelt. Es gibt zwei Möglichkeiten, eine nicht ausreichende Datenbasis zu kompensieren: die Datenaugmentation und das Transfer-Learning.

### **Trainings-, Validierungs- und Testdaten**

Ziel beim Training eines datengetriebenen Modells ist die Maximierung der Generalisierungsfähigkeit bzw. die Minimierung des Generalisierungsfehlers. Um das Modell bei der Verallgemeinerung des Gelernten zu unterstützen, sollte die vorliegende Datenmenge in drei unterschiedliche Datenmengen, die Trainings-, Test- und Validierungsdaten aufgeteilt werden. Mit den drei Teilmengen wird dem Netz in den zugehörigen Trainings-, Validierungs- oder Testphasen jeweils das bis dahin noch unbekannte Datenpaket mit ähnlichen Mustern präsentiert. Warum dieses Vorgehen einen essentiellen Einfluss auf die Modellqualität hat, ergibt sich durch die Erläuterung der genannten Teildatenmengen [133, S. 192], [73, S. 240–241]:

#### **1. Testdatenmenge:**

Um die Qualität des Netzwerkes final zu beurteilen sind die Testdaten wichtig. Entscheidend ist, dass diese Daten dem Modell im Vorfeld niemals präsentiert wurden, also weder bereits in den Trainings- noch in den Validierungsdaten auftauchen. Daher bietet es sich an, die Testdaten als erstes aus einer vorliegenden Datenbasis zu entnehmen. Nach dem Modelltraining sowie der Modelloptimierung werden die Testdaten verwendet, um die finale Güte der Prognose bzw. des Klassifikators oder Regressionsmodells zu bewerten; sie werden dem Modell also bis zur Testphase vorenthalten. [40, S. 209], [16, S. 134], [98, S. 200]

#### **2. Trainingsdatenmenge:**

Nach Entfernung der Testdaten kann die Trainingsdatenmenge zufällig aus der verbleibenden Gesamtdatenmenge ausgewählt und zum Training des Maschinellen Lernmodells verwendet werden. Mit dem prozentual größten Anteil der Ausgangsdatenbasis dienen die Trainingsdaten als Input für das Netz. Die dann noch verbleibenden Daten entsprechen den Validierungsdaten. [16, S. 133], [98, S. 200], [40, S. 208]

### 3. Validierungsdatenmenge:

Die Validierungsdaten dienen der Leistungsbeurteilung des Modells. Dabei gibt es verschiedene Methoden, mit denen die Validierungsdaten ausgewählt und von den Trainingsdaten separiert werden können. [16, S. 133]

Mit dieser Datenmenge wird das trainierte Modell zunächst getestet, indem der Testfehler ausgehend von den Validierungsdaten geschätzt wird. Mit den gewonnenen Informationen können die *Hyperparameter* des trainierten Modells optimiert werden. Im Falle von Neuronalen Netzen beschreiben die Hyperparameter beispielsweise die Anzahl oder Größe der einzelnen Layer. Mit der Validierungsmenge wird also weitergelernt, indem das Ergebnis als Feedbacksignal zu Auswahl geeigneterer Parameter genutzt wird. Dieser Optimierungsschritt kann beliebig oft wiederholt werden, um schlussendlich eine optimierte Parameterkonstellation zu ermitteln. Durch diesen Prozess lernt das System immer weiter, weshalb es zur finalen Bewertung erneut mit den verbliebenden unbekanntenen Testdaten auf die erreichte Güte getestet werden sollte. [16, S. 133–134], [98, S. 200]

In der Literatur wird häufig nur die Aufteilung in Test- und Trainingsdaten genau beschrieben, ohne noch zusätzlich eine Validierungsdatenmenge abzugrenzen. Daher wird für die vorliegende Arbeit auf die Verwendung von Trainings-, Validierungs- und Testdaten hingewiesen. Im Folgenden werden die Trainings- und Validierungsdatenmengen unter dem Begriff **Lerndaten** zusammengefasst, da die aufgestellten Modelle mit dieser Menge optimal auf die finale Bewertung mit den Testdaten vorbereitet werden. Gerade für Deep-Learning-Verfahren wird diese Dreiteilung der Datenmenge genutzt [133, S. 68]. Abbildung 2.24 verdeutlicht die beschriebene Zugehörigkeit und Aufteilung einer gesamten Datenmenge. Generell gilt, dass es in allen drei Datenmengen keine Überschneidungen geben darf, um die Modellqualität nicht negativ zu beeinflussen.

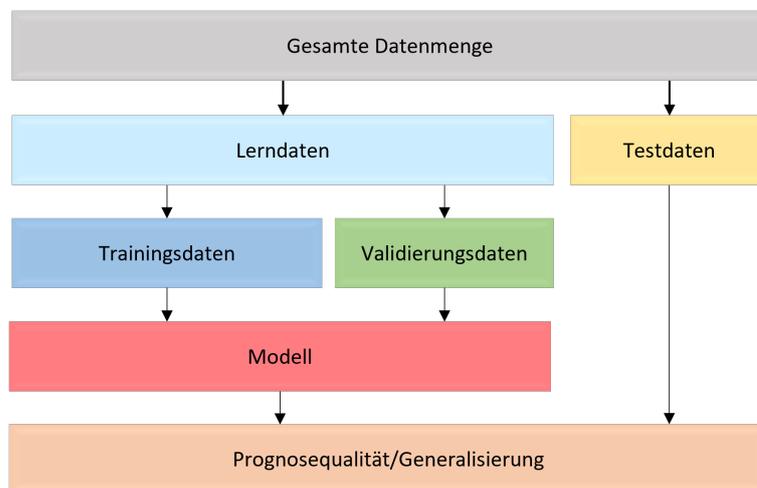


Abbildung 2.24: Datenaufteilung in Lern- und Testdaten (vgl. [133, S. 68])

### Validierungsmethoden

Validierungsdaten werden verwendet, um die Prognosequalität maschineller Lernmodelle während des Trainings zu prüfen bzw. zu überwachen. Dabei beschreibt die Prognosequalität den durchschnittlichen

Generalisierungsfehler des aufgestellten maschinellen Lernmodells auf die bisher unbekannte Validierungsdatenmenge [52, S. 241]. Abhängig von dem verfügbaren Volumen sowie der Relevanz einer Datenmenge kann zur Modellbewertung zwischen einer *einfachen Validierung* und der *Kreuzvalidierung* gewählt werden. Für die *einfache Validierung* (engl. *Hold-Out Validation*) wird bei einem hohen Datenvolumen und einer gleich verteilten Relevanz bezogen auf das zu untersuchende Problem davon ausgegangen, dass die Lerndatenmenge zu einem wählbaren prozentualen Anteil in Trainings- und Validierungsdaten geteilt werden kann, ohne dem Modell dadurch wichtige implizite Informationen vorzuenthalten. [16, S. 121] Ein häufig verwendetes Verhältnis entspricht 70 % Trainings- sowie 30 % Validierungsdaten [98, S. 200].

Ist die verfügbare Datenmenge nicht ausreichend oder haben die Daten unterschiedlich hohe Wertigkeiten im Sinne der Analyseaufgabe, reicht die einfache Validierung häufig nicht aus, um die Komplexität des gesamten Modells mit den Trainingsdaten zu erfassen. In diesen Fällen wird die so genannte Kreuzvalidierung gewählt. In der Literatur tritt diese auch unter dem englischen Begriff *Cross Validation (CV)* bzw. der entsprechenden Abkürzung auf. Es gibt drei Variationen der Kreuzvalidierung:

#### 1. Zufällige Untergruppen-Validierung:

Dieses Verfahren ergänzt die Methodik der *einfachen Validierung*, indem diese mehrfach durchgeführt wird, um letztendlich den Mittelwert aller Qualitätswerte der Metamodelle über den Mittelwert zu einer Qualitätsaussage zu kombinieren. Dafür wird bei jedem Durchgang der selbe prozentuale Anteil zufällig gewählter Validierungsdaten genutzt. Durch diese zufällige Auswahl kann es jedoch dazu kommen, dass es Datenpunkte gibt, die entweder niemals oder auch mehrfach in der Validierungsmenge auftreten, sodass sich die Bedeutung dieser Daten auf das Modell auswirken kann. [114, S. 316]

#### 2. $k$ -fold-Kreuzvalidierung:

Um die zuvor beschriebene Problematik der ungleichmäßigen Datenverteilung auf Trainings- und Validierungsmengen zu verhindern, wird die Lerndatenmenge in  $k$  gleichgroße und sich nicht überschneidende Mengen  $K_1, \dots, K_k$  eingeteilt. Das aufgestellte Lernmodell wird dann  $k$  mal ausgeführt und nach dem jeweiligen Validierungsscore bewertet. Für jeden Durchlauf wird eine andere der  $k$  Teilgruppen als Validierungsmenge gewählt, während die verbleibenden  $k - 1$  Gruppen die Trainingsdatenmenge bilden. Es ist nicht unüblich, dass sich die  $k$  Validierungsscores deutlich voneinander unterscheiden. Aus diesen jeweiligen Qualitätswerten wird zur finalen Modellbewertung der Mittelwert berechnet. Dieses Verfahren gewährleistet eine gleichmäßige Datenverteilung zur Auswertung der Modellqualität, denn es wird vermieden, dass die Hyperparameter durch eine zufällig ungünstig gewählte Validierungsdatenmenge bestimmt werden. Häufig gewählte Werte zur Gruppierung sind  $k \in \{4, 5, 10\}$ . [114, S. 316], [16, S. 121], [52, S. 242], [98, S. 201], [89, S. 236]

#### 3. Leave-One-Out Kreuzvalidierung:

Die Leave-One-Out-Kreuzvalidierung kann als besondere Methode der  $k$ -fachen Kreuzvalidierung betrachtet werden. Die Validierungsdatenmenge besteht dabei nur aus einem Datensatz. Diese Methode ist insbesondere für große Datenmengen sehr zeitaufwändig, kann jedoch bei einer kleineren Anzahl von Datensätzen zu besseren Prognosequalitäten führen. [52, S. 242] [114, S. 316], [89, S. 241]

### Dropout

Neben der klassischen Architektur eines CNNs kommen häufig auch Dropout-Schichten zum Einsatz. Das Dropout-Verfahren dient zur Regularisierung eines Modells und kann in der Trainingsphase eines Modells

beliebig eingesetzt werden. Der Eingabevektor einer Dropout-Schicht entspricht dem Ausgabevektor der vorherigen Schicht. Das Dropout-Verfahren schaltet dabei einen gewissen festzulegenden Anteil der Neuronen zufällig aus, indem diese Neuronen mit dem Wert 0 neu bewertet werden. [16, S. 149], [133, S. 26] Der neu erzeugte Ausgabevektor führt dazu, dass sich das Modell im nächsten Trainingsschritt nicht auf bestimmte Aktivierungen in einem gegebenen Feedforward-Pfad verlassen kann. Auf diese Weise werden redundante Wege im Netz ausgebildet, um mehr Neuronen für das gleiche Ziel zu verwenden. [40, S. 261] Eingesetzt wird dieses Verfahren, um der Gefahr der *Überanpassung (Overfitting)* entgegenzuwirken, das Trainingsergebnis des Modells zu verbessern und das Training ggf. zu beschleunigen [40, S. 261], [133, S. 26]. Die Dropout-Quote, also der stillgelegte Neuronenanteil, liegt meistens zwischen 0,2 und 0,5 [16, S. 149].

### Aktivierungsfunktion

Die Aktivierungsfunktion in Deep Learning-Modellen ist essentiell für die Leistungsfähigkeit und die Verallgemeinerungsfähigkeit des Modells. Ohne eine Aktivierungsfunktion werden nur lineare Transformationen der Daten erlernt [16, S. 102]. Gegenüber linearen Modellen können durch den Einsatz nicht linearer Aktivierungsfunktionen wesentlich kompliziertere Funktionen von einem Modell erfasst und erlernt werden [89, S. 104]. Häufig verwendete Aktivierungsfunktionen sind die logistische Funktion (Sigmoid), Tangens hyperbolicus und der lineare Gleichrichter (Rectified Linear Unit) [57, S. 55]. Die drei Funktionen sind in

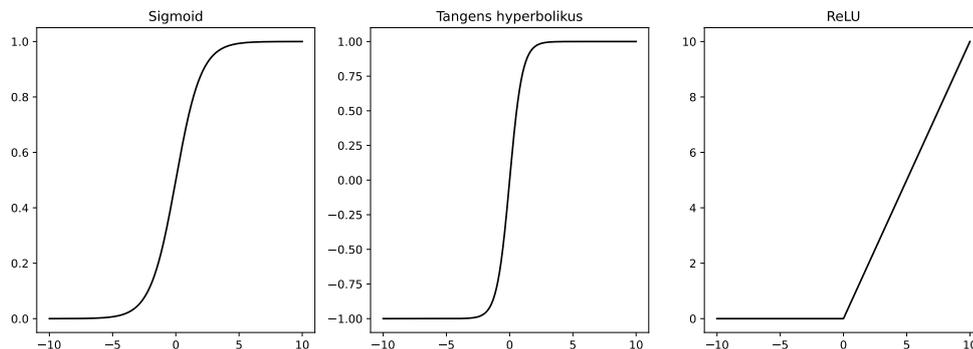


Abbildung 2.25: Aktivierungsfunktionen (vgl. [57, S. 55])

Abbildung 2.25 abgebildet und mit den Formeln des Tangens hyperbolicus

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.3.4)$$

und der logistische Funktion

$$S(x) = \frac{1}{1 + e^{-x}} \quad (2.3.5)$$

definiert. Insbesondere in CNNs findet die *Rectified Linear Unit (ReLU)*-Aktivierungsfunktion

$$f(x) = \max(0, x) \quad (2.3.6)$$

häufig Anwendung. Die ReLU-Schicht ersetzt dabei negative Werte durch 0 und markiert damit diesen bestimmten Wert, ohne die Dimension der Eingabe zu verändern. [133, S. 26]

### Datenvolumen

Das Datenvolumen ist gerade für DL-Modelle ein wichtiger Erfolgsfaktor. Bildverarbeitende CNNs zur Bildklassifikation oder zur Spracherkennung benötigen einige tausend, besser zehntausende Bilder, um sinnvoll trainieren zu können [73, S. 259]. Existierende DL-Modelle wie ResNet50 [53], VGG16 [115] oder MobileNet [58] berechnen eine halbe Million Gewichte und mehr im Training.

Die Komplexität der Modellarchitektur hängt dabei nicht von der verfügbaren Datenmenge ab, sondern von der Variabilität der Trainingsdaten. Wenn mit dem Sammeln zusätzlicher Daten eine höhere Anzahl Varianten abgedeckt werden kann, kann mit dieser erhöhten Datenmenge auch ein komplexeres Modell konstruiert werden. Dieses kann, ohne die Gefahr einer Überanpassung, gute Prognosequalitäten erzielen. [89, S. 31], [133, S. 192] Ein wesentlicher Nachteil dieser komplexen Modelle ist die aufwendige Lernphase, die jedoch durch die parallelisierte Verarbeitung mithilfe von GPUs und verteilten Systemen unterstützt werden kann [133, S. 26]. Andrew Ng, ein chinesisch-US-amerikanischer Informatiker, empfiehlt die Kennzahlen *Trainingsfehler*, *Testfehler*, *Validierungsfehler* und *Trainingsvalidierungsfehler* in Bezug auf die eingesetzten Datenmengen zu berücksichtigen<sup>8</sup> [133, S. 193]. Entsprechende Handlungsempfehlungen sind Abbildung 2.26 zu entnehmen.

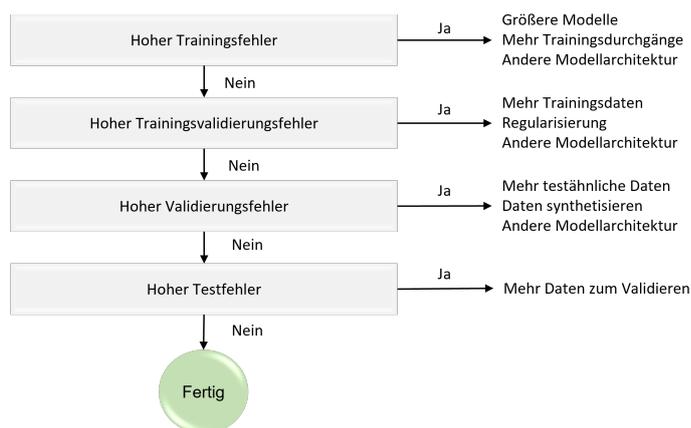


Abbildung 2.26: Interpretation und Handlungsempfehlungen für das Modelltraining (vgl. [133, S. 194])

Zusätzliche Daten können beim überwachten Lernen einen erheblichen Einfluss auf die Modellqualität haben. Allerdings ist das einfache Duplizieren keine sinnvolle Option zur Erhöhung des Datenvolumens. [89, S. 31], [133, S. 192] Eine bessere Möglichkeit zur synthetischen Vergrößerung der Datenmenge bieten Methoden der Dataaugmentation.

### Dataaugmentation

Methoden zur Dataaugmentation sind ein gängiges Hilfsmittel, um eine vorliegende Datenbasis so zu erweitern, dass eine Leistungsverbesserung des bestehenden Modells erzielt werden kann. Die Ausgangsdaten werden durch die Anwendung von Augmentationsmethoden modifiziert und neu gespeichert, wodurch sich

<sup>8</sup><https://www.youtube.com/watch?v=F1ka6a13S9I> (besucht am 26.07.2022)

die Ausgangsdatenmenge erweitert und einer Überanpassung des Modells entgegengewirkt wird. Es ist zu berücksichtigen, dass diese Erweiterung keinen Informationsgewinn für das Modell darstellt. Dem Netz werden lediglich die vorhandenen Informationen auf eine andere Art und Weise präsentiert. [80, 18, 106], [16, S. 184–189] Das Modell wird beim Training niemals zwei völlig identische Daten vorfinden. Dadurch kann das Modell weitere wichtige Merkmale der Daten erfassen, was zu einer Verbesserung der Generalisierungsfähigkeit führen kann [16], [73, S. 228]. Dabei ist entscheidend, dass die Datenerweiterung nur auf die Trainingsdaten angewendet wird, wie es auch in den Analysen von Takahashi, Gygli et al. [125] gemacht wird.

Die Auswahl passender Methoden zur Datenaugmentation ist von der Problemstellung abhängig. Verfahren, die auf Bilddaten angewendet werden, sind zum Beispiel das Zerschneiden von Bildern, Spiegelungen, Translation oder Rotation bzw. Verzerrungen, Änderungen des Kontrastes, der Helligkeit oder der Farbsättigung. [106], [73, S. 228] In Bezug auf die vorliegenden audiobasierten Bilddaten ist bei der Augmentierung Vorsicht geboten, denn die Bildaugmentierungsmethoden sind nicht ohne weiteres auf Spektrogramme anwendbar, auch wenn das Spektrogramm als Bilddatei an ein CNN übergeben wird. Diese Bilddaten basieren aber auf Audiodaten, sodass es sinnvoller ist, Augmentationsmethoden direkt auf die Rohaudiodaten anzuwenden und im Nachhinein die Bildtransformation durchzuführen. Es gibt allerdings auch Methoden, die auf Spektrogramme anzuwenden sind, wobei das Signal als solches durch die Bildveränderungen nicht verändert wird. Es macht beispielsweise keinen Unterschied, ob eine Audiodatei in drei separate Dateien geteilt wird, um dann drei Spektrogramme zu erzeugen, oder die Ausgangsdatei als Spektrogramm dargestellt und dieses im Zeitbereich in drei Teile zerschnitten wird. Abbildung 2.27 zeigt die beiden erklärten Optionen der Datenerweiterung für Audiodateien. Für die Erweiterung bzw. Veränderungsmöglichkeiten von Audiodaten orientiert sich die vorliegende Arbeit unter anderem an der MATLAB-Toolbox *AudioDegrader*. Aus der Veröffentlichung [80] von April 2021 geht hervor, dass es sich um die erste Toolbox handelt, die sich auf die Augmentierung von Audiodaten spezialisiert hat. Viele der dort vorgestellten und in MATLAB implementierten Methoden greifen auf die *Audio Degradation Toolbox* zurück [83].

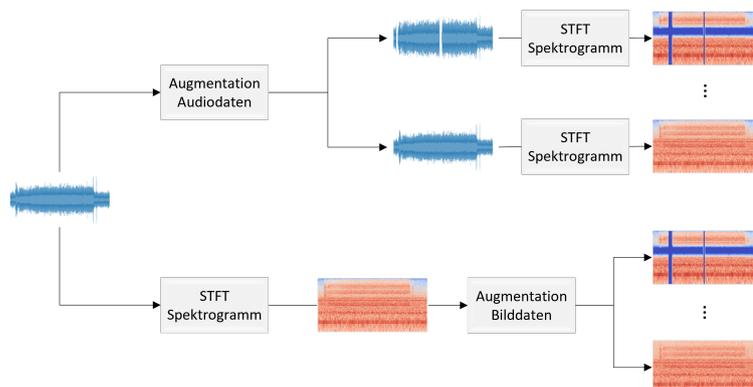


Abbildung 2.27: Augmentation von Audiodaten (vgl. [80])

## Transfer Learning

Ein anderer oder auch ergänzender Ansatz, um mit einer geringen Datenmenge erfolgreiche Prognosen zu erzielen, ist das *Transfer Learning*. Hier werden bereits bestehende und trainierte Netze für die eigene

Problemstellung verwendet. Es wird sich also an dem bereits gewonnen Wissen aus einem ähnlichen, bereits gelöstem Problem bedient und auf das eigene übertragen. Dafür wird der gelernte Block von Convolutional- und Pooling-Schichten wiederverwendet und nur das dichte Netz ab der Fully-connected-Schicht wird auf das eigene Anliegen spezifiziert. [40, S. 393–394] Viele der Netze, wie z.B. die VGG-Topologie [115], ResNet-50 [53] oder AlexNet [71], die für Transfer Learning genutzt werden, basieren jedoch auf Bilddaten und nicht auf Audiodaten bzw. Spektrogrammen. In der Forschung von Palanisamy, Singhanian et al. (2020) wird belegt, dass Transfer Learning auch dann für Audioklassifikationen möglich ist, wenn das Ausgangsnetz mit normalen Bildern trainiert wurde. Hierfür wurde systematisch untersucht, wie viele von den trainierten Gewichten für das Lernen von Spektrogrammen nützlich sind. Es konnten Validierungsgenauigkeiten durch Bildnetze erzielt werden, die für die Prognose der untersuchten Audiodatenmengen (UrbanSound8K<sup>9</sup>, ESC-50<sup>10</sup>) dem Stand der Technik entsprechen. Damit sind Bildnetzwerke in der Lage, exzellente Ergebnisse bei der Audioklassifikation zu erzielen. [95] Mittlerweile gibt es auch Soundnetzwerke, die nur auf Basis von Audiodaten trainiert wurden und z.B. in [128] bessere Klassifikationsergebnisse als Bildnetzwerke erreichen. Es wurde festgestellt, dass bei Transfer Learning die Nachtrainingskonfiguration erhebliche Auswirkungen auf die Klassifizierungsgenauigkeit und die erforderlichen Rechenressourcen haben kann [128].

---

<sup>9</sup><https://urbansounddataset.weebly.com/urbansound8k.html> (besucht am 13.05.2022)

<sup>10</sup><https://dagshub.com/kinkusuma/esc50-dataset> (besucht am 13.05.2022)



# Kapitel 3

## Stand der Technik

Der Hintergrund der vorliegenden Arbeit wird vor allem durch zwei Forschungsfelder gebildet: Die Regeneration anorganisch gebundener Formstoffe im gießereitechnischen Kontext und die Audioklassifikation mit Convolutional Neural Networks im allgemeinen und industriellen Rahmen.

### 3.1 Regeneration anorganisch gebundener Formstoffe

Die wesentlichen Herausforderungen, denen sich die Gießereiindustrie im Rahmen der Digitalisierung sowie immer strenger werdender Umweltvorschriften stellen muss, wurden bereits in der Einleitung dargelegt. Neben dem Forschungsziel dieser Dissertation gibt es bereits Forschungsfortschritte in ähnlichen sowie angrenzenden Bereichen. Die zahlreichen chemischen Bindemöglichkeiten, ebenso wie die Vielzahl der Regenerationsoptionen, bieten hohe Kombinationspotentiale und damit auch unterschiedliche Forschungsschwerpunkte. Die Abgrenzung dieser Dissertation liegt insbesondere in den gießereitechnischen Aspekten der Regenerationsmöglichkeiten von belasteten Sanden mit einem anorganischen Bindersystem sowie der Digitalisierung formstoffbezogener Überwachungssysteme.

Unumstritten ist, dass organische Bindemittel weltweit einen erheblich größeren Marktanteil als anorganische haben. Das liegt daran, dass anorganische Bindemittel nicht für jeden Einsatzzweck geeignet sind, da sie bei Temperaturen oberhalb von 900 °C ihre physikalischen Grenzen erreichen [88]. Dennoch gibt es nennenswerte Innovationen im Bereich der Regenerationstechnik, die im Sinne eines nachhaltigeren Gießprozesses stehen. Einen allgemeinen Überblick über die Fortschritte der Regenerationstechnik und dem Wiedereinsatz belasteter Altsande mit besonderem Fokus auf die polnische Gießereiindustrie liefern die Veröffentlichungen von Jozef Stanislaw Dańko [20, 21, 23, 22]. Des Weiteren versuchen Kayal und Chakrabarti (2008) allgemeingültige Aussagen hinsichtlich einer optimalen Regenerationsstrategie zu treffen, doch aufgrund der vielen variablen Faktoren (Art des Bindersystems, Art der Regeneration, Art des Formstoffes, usw.) kann kein Regenerationsverfahren hervorgehoben werden. Das Ergebnis dieser Forschungsarbeit liegt schlussendlich in der individuellen Auslegung von Regenerationsanlagen bzw. -verfahren, um anfallende Formstoffabfälle aufzubereiten und damit weiterhin sinnvoll nutzen zu können. [64] Auch das finnische Unternehmen *Finn Recycling*<sup>11</sup> folgt dem Trend der Ressourcenschonung und kämpft gegen die hohen Abfallmengen von Altsanden bei gleichzeitig steigenden Importen von Neusanden. Mit ihrem eigens entwickelten Regenerationsverfahren wirbt *Finn Recycling* mit hohen Einsparungspotenzialen von etwa 15 %

---

<sup>11</sup>[www.finnrecycling.fi](http://www.finnrecycling.fi) (besucht am 13.09.2022)

bis 30 % hinsichtlich der Sandkosten. Zusätzlich geht damit eine deutlich verbesserte CO<sub>2</sub>-Bilanz einher. Mit der Kombination aus thermischer sowie mechanischer Regeneration bietet *Finn Recycling* ein Regenerat mit neusandähnlichen und teils sogar besseren Eigenschaften als Neusand. Das produzierte Regenerat weist eine sehr hohe Feinheit und die Möglichkeit der einstellbaren Korngrößenverteilung sowie des Grades der Entstaubung auf. [91]

Bezüglich anorganisch gebundener Formstoffsysteme haben Haanappel, Linke et al. (2022) ein Verfahren zur Regenerierung von wasserglasgebundenen Gießereisanden entwickelt, welches als mechanisch adsorptiver Prozess beschrieben werden kann und unter der folgenden Bezeichnung *WO 2017/137113A1/EP3202927B1, Method and Device for regeneration Foundry Sand* patentiert ist. Die schonende Sandaufbereitung, die robuste und kostengünstige Anlagentechnik sowie die Einsatzmöglichkeit für viele verschiedene Bindemittelsysteme sind wesentliche Vorteile des entwickelten *Rotareg*, insbesondere für wasserglasgebundene Gießereisande. [51] Zum Vergleich wird der Schleifscheibenregenerator herangezogen, der in der Veröffentlichung von Sun, Jing et al. (2011) vorgestellt wird und zu den mechanischen Regenerationsverfahren zählt. Der Altsand wird hierbei durch die Einwirkung eines Luftstroms in einen fluidisierten Zustand versetzt und durch eine rotierende Schleifscheibe mit hoher Geschwindigkeit in Bewegung gebracht. Kollisionen zwischen dem Sand, der Schleifscheibe und dem Maschinenkörper führen zu Reibeffekten. Die gelösten Binderreste und Feinstaubpartikel werden durch ein Staubabscheidungssystem aus dem Prozess entfernt. Mit einer geeigneten Rotationsgeschwindigkeit kann die Regeneration von Harzsand beliebigen Anforderungen angepasst werden, sodass das Regenerat für Grauguss, Sphäroguss, Stahlguss, Kupferguss und Aluminiumguss eingesetzt werden kann. [123] Auch die *Klein Anlagenbau AG* beschäftigte sich mit der Entwicklung eines Regenerationsverfahrens und bietet mit ihrem eigenen Clustreg-Verfahren eine nahezu emissionsfreie Regenerierung von Gießereialtsanden. In einem ersten Schritt werden die Binderreste durch die mechanische Regenerierung vom Sand gelöst. Danach wird dieser vorbehandelte Sand mit einem Trägerstoff und einem Haftvermittler vermengt, sodass die gelösten Binderreste und Additive an dem Trägerstoff haften. Das Gemisch aus Trägerstoff und gesammelten Abfallpartikeln wird im letzten Schritt vom Korn entfernt. [108] Neben den genannten Forschungsarbeiten zur Anlagentechnik wird auch die gesamte Prozesskette der Wiederaufbereitung untersucht, wie z.B. durch Stachowicz und Granat (2014, 2016), die sich der Prozessoptimierung des gesamten Wiederaufbereitungsprozesses von wasserglasgebundenen Formsandabfällen widmeten. In ihrem Experiment geht es um die Verbesserung des klassischen Zyklus vom Anmischen über das Mikrowellenhärten, das thermische Belasten mit 800 °C, das Abkühlen, das Zerkleinern und zuletzt das mechanische Regenerieren. Die Innovation bezieht sich dabei auf die Reaktivierung verbleibender Binderreste im Sandsystem nach dem Regenerationsvorgang. So wird vor dem Anmischen ein zusätzlicher Aktivierungsschritt in den Zyklus der Wiederaufbereitung integriert. [118, 119] Ein anderer Ansatz zur Aufbereitung anorganisch gebundener Sande entstand unter der Annahme, dass durch eine deutliche Verringerung der Temperatur die auf der Oberfläche der Sandkörner befindliche Wasserglashülle brüchiger wird. Ziel dieses Experimentes von Dańko und Dańko (2005) war die Optimierung des Regenerationseffektes, welcher mit der Zugabe von flüssigem Stickstoff in das Regenerat erreicht werden sollte. Bei diesem Vorgehen wird der belastete Sand im Aufbereitungsprozess auf –70 °C gekühlt, wodurch sich das Volumen des Bindemittels sowie der Sandkörner verändert und erhöhte Spannungen an der Kornoberfläche entstehen. [21] Einen großen Bereich in der Aufbereitung wasserglasgebundener Formstoffe deckt die Nassregeneration ab. Kang, Kim et al. (2018) und Kim, Bae et al. (2021) untersuchten den Einsatz von nassregeneriertem anorganisch gebundenem Altsand in der Kernproduktion [63, 67]. Die Abgrenzung zwischen diesen Experimenten liegt in der Zugabe eines Antihafmittels in das wasserglasbasierende Bin-

demittel zur Erzielung besserer Ausklopfeigenschaft bzw. eines besseren Entkernverhaltens [67]. Mit dem selben Ziel entwickelten Izdebska-Szanda, Balinski et al. (2012) einen chemisch modifizierten anorganischen Binder, der dem sonst vorherrschenden Nachteil einer hohen Restfestigkeit entgegenwirkt. [61] Auch Etemayer, Erhard et al. (2021) haben das Entkernverhalten anorganisch gebundener Quarzsande untersucht, mit dem Ziel, dieses bereits nach dem Anmischen vorhersagen zu können [31]. Mit der Entwicklung eines Biobinders auf Proteinbasis, der sich ebenfalls zur Wiederaufbereitung eignen soll, beschäftigten sich Shi, Liu et al. (2012) [113]. Neben der Forschung zu Altsanden, die aus natürlichen Formgrundstoffen bestehen, gibt es auch Forschung, die sich mit der Regeneration von anorganisch gebundenen künstlichen Sanden beschäftigt [68].

Das Vermischen von Form- und Kernsand ist ein wichtiger zu berücksichtigender Faktor im Bereich eines Formstoffumlaufsystems, denn Form- und Kernsande werden nicht zwangsläufig mit identischen Bindersystemen angemischt. In Bezug auf das Mischen eines bentonitgebundenen Formstoffsystems mit einem zum Beispiel alkalischen anorganischen Kernbindersystem hat dies bei der Wiederverwendung negativen Einfluss auf die Formstoffeigenschaften. Bisher wird dieser Problematik mit einem erheblichen Austausch des Umlaufformstoffes durch Neusand entgegengewirkt. [140] Bifano, Fischerauer et al. (2020) stellen ein Verfahren vor, welches sich dieser Problemstellung annimmt. Dafür nutzen sie die Impedanzspektroskopie als Methode zur Charakterisierung verschiedener Arten von Quarzsand, Chromitsand und Sandmischungen. Zudem verfolgen sie das Ziel, den Regenerationsprozess von Sandmischungen aus verschiedenen Sanden und Bindemitteln so zu optimieren, dass das Regenerat für den Wiedereinsatz ausreichend gut gereinigt wurde. Um sicher zu stellen, dass sich die Spektren unterschiedlicher Sande signifikant unterscheiden, wurden die jeweiligen Sande zunächst einzeln vermessen. Mit dem Ergebnis, dass jede Sandart eine charakteristische Impedanzkurve mit unterschiedlichen Merkmalen im Nyquist-Diagramm aufweist, wurde ein parametrisiertes Modell entwickelt, welches die wichtigen charakteristischen Merkmale der Eigenschaften der Impedanzen erfasst. So kann eine Identifizierung und Klassifizierung von Sandgemischen durchgeführt werden. [6, 7]

Obwohl sich die Art der Regeneration und die verwendeten Bindersysteme in den gezeigten Forschungen unterscheiden, nutzen die meisten der vorgestellten Arbeiten identische Qualitätskriterien, um den Erfolg des Regenerationsverfahrens zu untersuchen, welche zusammenfassend in Tabelle 3.1 aufgelistet sind.

Tabelle 3.1: Qualitätskriterien zur Beurteilung des Regenerationseffektes

Methoden	Verwendet in:
Granulometrische Kennwerte	[91, 51, 64, 68, 21, 6, 31, 108]
Biegefestigkeit	[51, 91, 118, 119, 63, 64, 68, 21, 108]
Zugfestigkeit	[118, 119]
Gasdurchlässigkeit	[64, 21]
SiO <sub>2</sub> -Gehaltes im Waschwasser (Nassregeneration)	[67]

Neben der Anpassung bzw. Verbesserung der einzelnen Komponenten der Prozesskette des Wiederaufbereitungsprozesses, wie z.B. der verschiedenen Bindersysteme oder der Art der Regeneration, spielen auch indirekte Faktoren eine Rolle für die Effizienz eines Prozesses. Beispielsweise ist die Qualität der Herstellungsausrüstung entscheidend, wie Li et al. (2022) belegen; die Arbeit legte den Fokus auf die anorganische Kernherstellung, wobei herausgefunden wurde, dass eine Kernsandtemperatur von 35 °C nicht

überschritten und eine Standby-Zeit von idealerweise weniger als 10 Minuten eingehalten werden sollte. Neben diesen beiden Parametern hat auch die Temperatur des katalytischen Gases Einfluss auf die Fließfähigkeit des Kernsand und auf die anfängliche Festigkeit des Kerns. [76] Ein weiterer wichtiger Faktor ist der Gesamtenergieverbrauch von Gießereien, denn nach dem Schmelzbetrieb ist die Sandaufbereitung, besonders in Nassgussgießereien, der zweitgrößte Energieverbraucher. Für veraltete Maschinen, die für die genannten Prozesse genutzt werden, schlägt Wolfgang Ernst eine Umrüstung der Hardware vor. Es ist die Rede vom Tausch alter Motoren gegen Energiesparmotoren sowie von Modifikationen der Steuerung, die bei einem möglichen Leerlauf eine Sandaufbereitungsanlage beispielsweise zum Stillstand bringen kann. [30]

Gerade für solche Steuerungsoptionen ist der Einsatz Künstlicher Intelligenz von Vorteil. Die vorwiegend mittelständisch organisierte Gießereiindustrie in Deutschland steht dabei vor dem Problem, ausreichend ausgebildete Fachkräfte zu finden [78]. Gerade für KMU ist die Adaption neuer Technologien durch die begrenzt verfügbaren personellen, wirtschaftlichen, zeitlichen und auch natürlichen Ressourcen herausfordernd [27]. Im Hinblick auf die Analyse komplexer Fertigungsprozesse, wie die der Gießereitechnik, wird in der *Giesserei Special 2021* die Software *Eidominer* als entsprechende KI-Anwendungsumgebung für überwachtes maschinelles Lernen vorgestellt. Mit dem Ziel, auch ungeschultem Personal den Zugang zu KI-Modellen zu gewährleisten, bündelt die Software Daten-Exploration, Daten-Vorverarbeitung und Modell-Training inklusive Evaluation. [15] Weitere Herausforderungen und Chancen im gießereitechnologischen Kontext werden von Bührig-Polaczek, Rudack et al. 2021 im Artikel *Digitalisierung in der Gießereitechnik* erläutert. Bezogen auf die Formstoffsteuerung wird ein Acht-Schritte-Plan zur Implementierung eines digitalen Verfahrens für den Einsatz maschineller Lernmethoden vorgestellt, womit sich die Energie- und Rohstoffeffizienz positiv beeinflussen ließe. [13] Allerdings ist bereits die Erfassung prozessrelevanter Daten zur Qualitätssicherung und Prozesssteuerung in vielen Gießereien ungenügend, sodass wechselseitige Abhängigkeiten einzelner Prozessschritte häufig unerkannt bleiben. Um dem entgegenzuwirken, ist die Digitalisierung als wichtiger und fester Bestandteil in der Unternehmensstrategie zu verankern. [25, 78, 27]

Einen wesentlichen Ansatz zur Prozesssicherheit in der Formstoffaufbereitung lieferte die Dissertation von H. Gemming (2003). Mit Hilfe von Fuzzy-Regelalgorithmen wurde ein Steuerungskonzept entwickelt, welches durch den Einsatz moderner regelungstechnischer und informationstechnischer Verfahren eine Verbesserung der Regelung der Formstoffaufbereitung erlauben sollte. Dabei wurde darauf geachtet, dass auch das aus Erfahrungen der Anlagenbetreiber gewonnene Wissen in die Reglerdefinitionen einfließen konnte. [42] In Bezug auf die Bestimmung eines optimalen Arbeitspunktes einer Regenerierungsanlage – unter Berücksichtigung maximaler Regenerierqualität verschiedener Ausgangssande – stellt E. Westhoff seine Forschung zur *On-line-Analyse von Formsandsystemen* vor. Das Ziel der Entwicklung und Optimierung eines Verfahrens zur online-Bewertung der Regenerierqualität mit Hilfe einer (N)PAK-Analytik wird erreicht, allerdings konnten rein anorganische Bindersysteme mit der verwendeten Analytik nicht online charakterisiert werden. [137] Gerade Formsand ist ein sich zeitlich ständig veränderndes Gefüge: Die Zugabe von Wasser, Umgebungsänderungen der Temperatur, Verdunstung, Bewegung, Mischen oder Auflockern können die Eigenschaften von Formsand beeinflussen. Auch regelmäßige Untersuchungen von Proben bilden jeweils nur Momentaufnahmen ab. Falsche Formstoffeigenschaften wirken sich im Gießprozess auf die Qualität der Gießereierzeugnisse aus. Dennoch werden die Prozesse bzw. die gesamte relevante Prozesskette noch nicht kontinuierlich überwacht, was gerade für den Einsatz von KI-Lösungen wichtig wäre. [6, 78] In dem Artikel „Formsand 4.0“ wird das Formsandmanagementsystem *FoMaSys* der Michen-

felder Elektrotechnik GmbH & Co. KG. vorgestellt, welches die oben genannten Einflussfaktoren auf die Formstoffqualität in einem einheitlichen System bündeln soll. Es handelt sich um ein Cyber-physisches System für die bentonitgebundene Formsandaufbereitung und ergänzt die klassische zentrale Steuerung. Das System bietet neben prozesstechnischen sowie wirtschaftlichen Vorteilen eine integrierte Datenbank mit verfügbarem Prozesswissen sowie Kunden-Feedback aus über 40 Jahren praktischer Erfahrung im Bereich der Formsandaufbereitung. Damit wird dem Fachkräftemangel sowie dem Verlust der „Erfahrungswerte der Altvorderen“ zum Teil entgegengewirkt. Für eine erfolgreiche Implementierung des FoMaSys ist die umfangreiche Datenerfassung aller relevanten Prozessparameter erforderlich, um mit diesen ein skalierbares System anbieten zu können, welches sich auf die Darstellung und Funktionalität der Formsandaufbereitung, insbesondere auf die Überwachung der Formsandqualität, konzentriert. [78]

Neben der Überwachung des Formstoffes selbst ist auch die Anlagenüberwachung relevant für die zu erzeugende Gussqualität. Künstliche Intelligenz wird daher z.B. auch für die Vorhersage des Verschleißzustandes von Kernschießmaschinen genutzt, damit Wartungen genau zu dem Zeitpunkt durchgeführt werden, an dem es erforderlich ist (*Predictive Maintenance*). Zudem kann ein implementiertes KI-System dazu beitragen, komplexe Prozesse zu überwachen und zu optimieren. Daten aller Temperatur-, Luftdruck- und Luftfeuchtigkeitssensoren, die in der Regel in einer Gießerei oder speziell an und um einen Schmelzofen herum verteilt sind, können zusammengeführt werden und anhand von Referenzuntersuchungen die Güte der gefertigten Erzeugnisse vorhersagen. Für den Erfolg eines solchen Systems müssen fortwährend neue Daten aufgenommen werden, um ein stetiges Lernen zu ermöglichen. [69] Auch für die qualitative Prüfung von Formstoffen gibt es bereits intelligente Systeme. Die Bildanalyse-Software der Firma *Olympus*<sup>12</sup> bietet eine Analysesoftware zur Bildererkennung, -segmentierung sowie -klassifizierung. Die Software ist für den Bereich der Werkstoffprüfung ausgelegt und unterstützt mit Hilfe von Künstlicher Intelligenz, insbesondere Deep Learning, Analysen von Mikroskopiebildern. [94] Darüber hinaus gibt es weitere Bereiche der Gießereitechnik, die von KI-Systemen profitieren können. In dem Projekt „Ultradünnwand-Druckgießen“ des Fachgebietes Gießereitechnik der Universität Kassel werden Methoden der Bildererkennung beispielsweise eingesetzt, um Oberflächendefekte auf den untersuchten Bauteilen verlässlich zu identifizieren und zu klassifizieren [41].

## 3.2 Audioklassifikation mittels CNN

Für die Forschung im Fachbereich der Audioklassifikation steht häufig die Untersuchung und Identifizierung von Umgebungs-, Umwelt- oder Tiergeräuschen im Vordergrund. Dabei kann es um die Erkennung von Vögeln anhand ihres Gesanges [117], die Charakterisierung von Lautäußerungen von Katzen [93] oder um die Erkennung von Tieren in Tropenwäldern zum Schutz der biologischen Vielfalt [124] gehen. Die *Detection and Classification of Acoustic Scenes and Events (DCASE)* treibt durch regelmäßig neue Aufgaben und Herausforderungen die Forschung im Bereich der Klassifizierung und Detektion von (Umwelt-) Geräuschen zusätzlich an. Die dadurch produzierten Ergebnisse können in zugehörigen Foren diskutiert und veröffentlicht werden, wodurch die Organisatoren des DCASE Forscher aus dem akademischen und industriellen Umfeld zusammenbringen und damit einen Beitrag zu diesem dynamischen Forschungsfeld leisten<sup>13</sup>. Auch in medizinischen Fachbereichen wird die Audioklassifikation mit maschinellen Lernmethoden, insbesondere den CNN, zum Beispiel zur Schnarcherkennung [138] eingesetzt. Ein weiteres großes

---

<sup>12</sup><https://www.olympus-global.com/> (besucht am 06.10.2022)

<sup>13</sup><https://dcase.community/> (besucht am 29.06.2022)

Forschungsgebiet bietet die Musik-Audioklassifikation, die u.a. die Identifizierung von Instrumenten, Genre, Liedern oder Soundtracks aus Videos beinhaltet [5, 112, 55]. Neben den genannten akustischen Signalen werden aber auch andere Signalarten mit CNN untersucht, wie zum Beispiel seismische Aktivitäten zur Überwachung von Vulkanen, bei denen sich der Netzinput aus Signalen von z.B. magmatischen Fluiden zusammensetzt [19].

Auch im industriellen Umfeld wird der Einsatz von Audioklassifikationsmodellen erforscht, z.B für die vorausschauende Wartung von Industrieanlagen. Hierfür entwickelten Henze, Gorishti et al. (2019) einen audiobasierten Ansatz, in dem eine Kombination der zwei maschinellen Lerntechniken, Anomalieerkennung und Klassifizierung, zum Einsatz kommt. Die Analyse des Anlagenzustandes bietet die Chance auf ein maximales Zeitintervall zwischen Reparaturen sowie eine minimierte Anzahl außerplanmäßiger Ausfälle, welche wiederum eine kosteneffiziente Wartung bedingen. Die durchgeführte Machbarkeitsstudie zeigte, dass das Prozessmodell in einem industriellen Umfeld, mit einem Mikrofon als Sensordatenquelle, eingesetzt werden kann. [54] Im Gegensatz zur Betrachtung von nur einer industriellen Anlage widmeten sich Kim, Lee et al. (2021) einem schallbasierten System zur gleichzeitigen Überwachung mehrerer Anlagen, also eines gesamten Fertigungsprozesses. Das entwickelte System beinhaltet für jede Anlage ein eigenes CNN, welches nur den Betriebsstatus der eigenen Anlage klassifiziert und die Geräusche aller anderen Geräte als Rauschen betrachtet. Der Vorteil dieses Überwachungssystems ist, dass es selbst bei gleichzeitigem Betrieb mehrerer Geräte/Anlagen möglich ist, genau eines dieser Geräte und dessen Status zu identifizieren. Das weiterführende Ziel ist daher die Entwicklung eines gesamtheitlichen Systems, welches auf Grundlage des Betriebszustands jedes Geräts, das in einem Fertigungsprozess verwendet wird, den Fortschritt des gesamten Fertigungsprozesses analysiert. [66] Eine weitere Option des industriellen Einsatzes von Audioklassifikationsmodellen wird in der Forschung von Yang, Peng et al. (2021) für die Überwachung von Förderbändern in der Kohleindustrie entwickelt, die bisher von Mitarbeitern manuell inspiziert wurden. Während der manuellen Prüfungen wurde besonders auf visuelle Fehler und akustische Veränderungen geachtet, doch durch die Subjektivität dieser Analysen wurde der Betrieb in der Kohleindustrie erheblich durch Störungen beeinträchtigt. Das daraufhin entwickelte intelligente audiobasierte Fehlerklassifizierungssystem für Förderbandrollen erreicht eine Genauigkeit der Fehlerzustandsklassifizierung von 90 %. [139] Ferner können auch Oberflächenbeschaffenheiten von Metallkugeln oder verschiedenen anderen Arten von Schüttgut (Schrauben, Muttern, etc.), die durch ein Rohr gleiten, anhand von Geräuschanalysen mit hoher Genauigkeit klassifiziert werden. Johnson, Grollmisch (2021) haben in ihrer Veröffentlichung nach einer Methode gesucht, die die Robustheit gegenüber Domänenverschiebungen bestehender Erkennungssysteme erhöhen kann, idealerweise ohne dass neue Daten aufgezeichnet oder die Modelle neu trainiert werden müssen. Durch die Anwendung einer adaptiven Normalisierungsmethode und neun Augmentierungsmethoden zur Datenerweiterung konnte die Leistung von zwei Geräuscherkennungsmodellen für die IIMB- und IIT-Datensätze um 28 % bzw. 39 % gesteigert werden. [62] Ryan, Takafuji et al. (2020) veröffentlichten einen Ansatz basierend auf selbstüberwachtem Lernen zur Klassifizierung von Geräuschen. Sie vergleichen den selbstüberwachten Ansatz mit den Ergebnissen eines vortrainierten Modells zu Klassifikation von Motorgeräuschen. Mit unterschiedlichen Augmentationsmethoden erwies sich das selbstüberwachte Training als wirksam zur Verbesserung der Leistung des nachgelagerten Klassifikationsmodells der Motorgeräusche gegenüber dem vortrainierten Modell. [110]

Wie in Kapitel 2.2 erläutert, kann jedes akustische Signal, welches das Ergebnis einer Kombination verschiedener Frequenzen ist, auf seine akustischen und individuellen Merkmale untersucht werden [77, S. 5]. Im Sinne der Audioklassifikation mit maschinellen Lernmodellen hat sich das Spektrogramm, die

Darstellung des Signals im Zeit-Frequenz-Bereich (s. Kapitel 2.2.2), als sinnvolle und nutzbare Darstellung erwiesen [117, 5]. Für Analysen bzw. Geräuschidentifikationen, die der menschlichen subjektiven Wahrnehmung von Geräuschen nahe kommen soll, wurde in der psychologischen Akustik neben der Frequenz die subjektiv empfundene Tonhöhe, die *Tonheit*  $h$ , eingeführt. Die Tonheit verhält sich nicht linear zu der Frequenz des Signals, sondern folgt über einen weiten Frequenzbereich vielmehr einem logarithmischen Verlauf, sodass für die Tonheit eine neue Skala, die sogenannte Mel-Skala eingeführt wurde. Mit Hilfe der Gleichung

$$h(f) = 2595 \cdot \log_{10} \left( 1 + \frac{f}{700\text{Hz}} \right) \quad [\text{mel}] \quad (3.2.1)$$

kann eine Frequenz  $f$  [Hz] in die entsprechende Tonheit  $h(f)$  umgerechnet werden [97, S. 94–95], [49, S. 6–7]. Sollen nun Geräusche von einem Modell so erlernt werden, wie der Mensch diese wahrnimmt, bietet sich das mel-Spektrogramm an. Dabei wird zunächst die STFT auf das Signal angewendet und das entstehende Spektrogramm wird danach mit Formel (3.2.1) mel-skaliert [126, 142, 112, 66]. Unabhängig davon, welche Art des Spektrogramms für die Analysen gewählt wurde, kann es zusätzlich Sinn machen, die Frequenzachse in eine logarithmische Skala umzurechnen; sie kann Ergebnisse liefern, die der menschlichen Wahrnehmung noch näher kommen [66].

Methoden zur Datenaugmentierung und der Einsatz von Transfer Learning finden auch in der Literatur immer häufiger Anwendung. Tabelle 3.2 gibt einen Überblick über den erfolgreichen Einsatz beider Methoden.

Tabelle 3.2: Anwendung von Augmentationsmethoden und Transfer Learning

Datenaugmentierung		Transfer Learning	
Methode	Verwendet in:	Methode	Verwendet in:
Time Stretch	[81, 112, 111, 24, 124, 62, 110]	VGG19	[142, 55, 124]
Pitch Shift	[117, 5, 112, 126, 24, 111, 62, 110]	VGGNet	[125]
Add Noise	[117, 81, 112, 124, 126, 24, 111]	AlexNet	[55]
Verstärkung	[81, 126, 112]	ResNet-50	[55]
Time Shift	[117]		
Combining Audio	[117]	YAMNet	[81, 128, 93]
SpecAug	[62, 110]	VGGish	[81, 128, 126]

Die Augmentation wird dabei nicht nur zur Erweiterung der Datenmenge verwendet, sondern ist auch nützlich bei unausgewogenen Datenmengen [81]. Da die Methoden selbst variabel einstellbar sind, werden einige von ihnen auch mehrfach mit unterschiedlichen Parametereinstellungen zur Erweiterung der Datenmenge verwendet, wie beispielsweise die Tonhöhenverschiebung oder das Hinzufügen verschiedener Hintergrundgeräusche. In der rechten Hälfte von Tabelle 3.2 wird der Einsatz von Transfer Learning in der Literatur in Bezug auf die verschiedenen vortrainierten Netze gelistet. Dabei befinden sich im oberen Teil der Tabelle die bildbasierten Netze und unten, durch einen Absatz getrennt, die Netze, die auf Audiodaten basieren. Einige der angegebenen Quellen finden sich sowohl in der linken als auch in der rechten Hälfte der Tabelle 3.2, da es keine Seltenheit ist, dass durch Augmentierung erweiterte Datenmengen auch mit bereits trainierten Netzen klassifiziert werden.



# Kapitel 4

## Konzept und Versuchsstand

Um eine Möglichkeit zu finden, einen betrachteten Regenerationsprozess hinsichtlich der Regeneratqualität zu optimieren, wurde im Rahmen dieser Arbeit praktische Versuche an einem Reibregenerator vorgenommen. Die während des Regenerationsprozesses aufgezeichneten Akustikdaten dienen als Grundlage der anschließenden durchgeführten Untersuchung maschineller Lernmethoden.

### 4.1 Regenerationsprozess mittels Reibregenerator

Der verwendete Reibregenerator<sup>14</sup> wurde in Zusammenarbeit der Firmen *Domnick Verfahrens- und Anlagentechnik* und *Technofond Gießereihilfsmittel GmbH*<sup>15</sup> in den Jahren 2008 bis 2012 entwickelt und für das durch die AIF (Arbeitsgemeinschaft industrieller Forschungsvereinigungen) geförderte Forschungsprojekt SandREG mit dem Förderkennzeichen ZF4047818CM9 zur Verfügung gestellt. Abbildung 4.1 zeigt die Maschine, die zur mechanischen Regeneration von Gießereisanden eingesetzt wird, in ihrer originalen Form (links) sowie in der graphisch erstellten Darstellung (rechts).

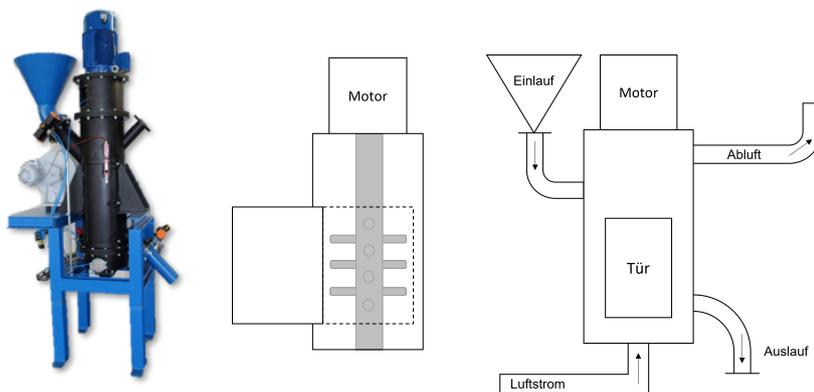


Abbildung 4.1: Der Reibregenerator und seine schematische Darstellung

<sup>14</sup><https://www.reibregenerator.de/> (besucht am 23.03.2022)

<sup>15</sup><https://www.technofond.de/> (besucht am 23.03.2022)

Die Funktionsweise und der Aufbau der Anlage lassen sich anhand des Materialflusses des gesamten Regenerationsprozesses, welcher mit Hilfe der fünf Teilbilder in Abbildung 4.2 dargestellt wird, anschaulich erläutern. Der oben links platzierte Einlauftrichter hält das Material vor Prozessbeginn bereit. Der kornvereinzelte Sand wird so lange durch eine manuell steuerbare Klappe im Trichter gehalten, bis diese zum Start der Regeneration über ein Bedienpaneel geöffnet wird, sodass das Material aufgrund der Schwerkraft in den Reibregenerator fließt. Die Prozessschritte des Einlaufs werden exemplarisch im 2. und 3. Bild von links der Abbildung 4.2 gezeigt. Der Regenerationsprozess selbst lässt sich bildlich nicht festhalten, da er in einem geschlossenen System bzw. hinter verschlossener Tür abläuft (Bild 4 v.l.). Zum Austragen des Materials wird eine weitere Klappe am unteren Teil des Korpus geöffnet. Durch die anhaltende Rotation der Welle wird der Sand sukzessive aus der Anlage getragen und kann über den Auslauf in einem Auffangbehälter gesammelt werden.

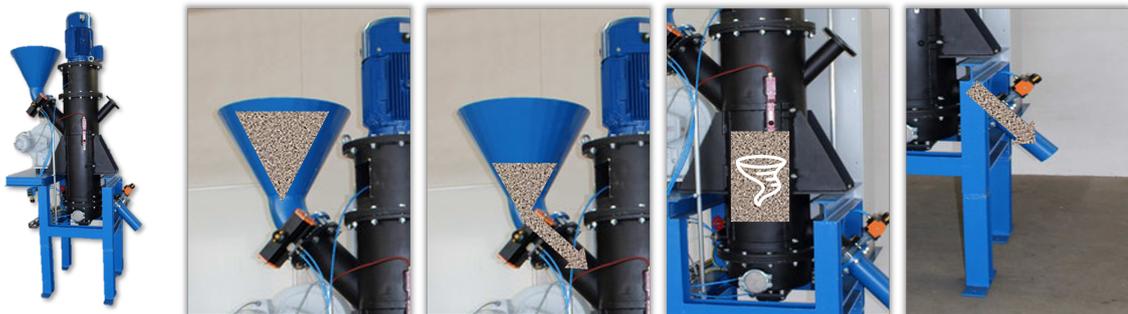


Abbildung 4.2: Darstellung des Materialflusses durch den Reibregenerator

Im Inneren der Anlage befindet sich eine Welle, die durch den aufgesetzten Motor angetrieben wird. An der Welle sind Reibstäbe befestigt. Grafisch ist dies in Abbildung 4.1 erfasst, eine reale Ansicht bieten die Bilder in Abbildung 4.3. An den unteren Stäben können zusätzliche Schlagplatten befestigt werden, um einem zu schnellen Verschleiß entgegenzuwirken (s. Bild 4.4).



Abbildung 4.3: Reibstäbe



Abbildung 4.4: Schlagplatten

Aus Sicherheitsgründen kann die Anlage nur mit verschlossener Tür gestartet werden. Sobald sich das binderbelastete Material innerhalb des Korpus befindet, wird dieses durch die Rotation der Reibstäbe und einen zusätzlichen vertikalen Luftstrom mechanisch regeneriert. Durch die Bewegung im Kornhaufwerk entsteht Reibenergie sowohl zwischen den einzelnen Körnern als auch zwischen dem Korn und dem Reibstab sowie zwischen dem Korn und der Außenwand des Anlagenkorpus. Die durch die Reibung entstehende Wärmeenergie trägt zur Versprödung der Binderhüllen bei, welche durch die Rotation wiederum besser abgeschlagen werden können. Der kontinuierliche Luftstrom trägt einen Teil dieser abgetragenen Binderhüllen sowie andere Feinpartikel aus dem gesamten System, dennoch sollte das Sandsystem nach der Regeneration zusätzlich gesichtet werden. Ein wesentlicher Nachteil der mechanischen Regeneration liegt im auftretenden Kornabrieb, der sich nicht verhindern lässt, also auch in den untersuchten Regenerationen im Rahmen dieser Dissertation auftritt.

## 4.2 Problemstellung und Zielsetzung

Die wesentlichen Aspekte des Prozesses lassen sich wie folgt zusammenzufassen:

1. Bei dem betrachteten Verfahren handelt es sich um ein mechanisches Regenerationsverfahren mit einem sogenannten Fließbettreiniger, der hier für die Regeneration von anorganisch gebundenen Sanden eingesetzt wird.
2. Das Ziel des Regenerationsverfahrens besteht darin, den gebundenen Sand wieder in ein Material umzuwandeln, das den Eigenschaften von Neusand ähnlich ist.
3. Da der Prozess in einem geschlossenen System stattfindet, ist er nicht transparent und der Qualitätszustand des Regenerates ist während des Prozesses mit dieser Anlage bisher nicht zu erfassen. Versuche, während des laufenden Betriebs signifikante Sandproben zu entnehmen, um die aktuelle Qualität des Sandes über die gängigen Formstoffqualitätskriterien zu bestimmen, blieben ohne Erfolg.

Die Zielsetzung dieser Forschungsarbeit liegt darin, eine Möglichkeit zu finden, den hier betrachteten Regenerationsprozess hinsichtlich der Regeneratqualität zu optimieren. Das bedeutet, dass der Prozess genau zu dem Zeitpunkt zu stoppen sein soll, an dem das sich im System befindende Regenerat wieder der gewünschten Qualität entspricht. Ein klassischer Ansatz liegt in der Aufstellung einer umfangreichen Versuchsplanung (Design of Experiments (DOE)), durch die genügend Erkenntnisse über den Fortschritt der Regeneration gesammelt wird.

Aus Fachgesprächen des vorangegangenen Forschungsprojektes SandREG ging ein alternativer und interessanter Ansatz hervor, der sich auf Erfahrungen der Anlagenbediener mit langjähriger Prozess Erfahrung stützt und auf akustische Unterschiede während des Regenerationsprozesses verweist. Generell wird die Akustik bereits bei der Leistungsbeurteilung von Maschinen eingesetzt, indem die Geräuschemission als Qualitätsmerkmal, in Form von akustischer Wertigkeit, berücksichtigt wird. Dabei spielen auch die Sinne des Menschen in solchen Bereichen und vor allem in den täglichen Lebenslagen eine entscheidende Rolle [116, S. 204]. Im Alltag begegnen Menschen Situationen, in denen Sie aus Intuitionen und Erfahrungswerten Handlungen ableiten. Bezogen auf akustische Ereignisse eignet sich das Beispiel des Mahlens von Nüssen oder Kaffeebohnen: Unterschiedliche Nussorten erzeugen beim Mahlen unterschiedliche Geräuschemissionen aufgrund ihrer Größe, Härte, Form oder des Mahlgrades. Gleiches gilt etwa für die

Wahrnehmung von akustischen Veränderungen bei Kaffeevollautomaten, die nur noch eine Restmenge an Bohnen zur Verfügung haben und während des Vorgangs leer laufen. Wird dieser Status akustisch erkannt, kann der Vorratsbehälter für die nächste Nutzung wieder mit Kaffeebohnen befüllt werden. Um zurück auf den vorliegenden Regenerationsprozess zu kommen, leitet sich aus den Erfahrungswerten der Prozessmitarbeiter die folgende zu untersuchende Hypothese für diese Arbeit ab:

„Mittels der Prozessakustik des hier betrachteten Regenerationsprozesses kann eine Aussage über die Regeneratqualität innerhalb dieses geschlossenen Prozesssystems getroffen werden.“

### 4.3 Lösungskonzept

Ausgehend vom Stand der Technik werden akustische Analysen bereits in den unterschiedlichsten Branchen eingesetzt. Mit der zu untersuchenden Hypothese legt diese Arbeit den Fokus auf die Übertragung der menschlichen akustischen Wahrnehmung auf ein maschinelles System, wodurch die digitale Signalanalyse von Audiodateien einen wesentlichen Bestandteil ausmacht. Um klare Analyseergebnisse und eine daraus abgeleitete Beurteilung der Hypothese liefern zu können, gilt es, die folgenden drei Leitfragen zunächst theoretisch und im Anschluss praktisch zu beantworten:

1. Nach welchem Qualitätskriterium des Regenerates wird der Prozess ausgelegt?

Die zu Beginn entscheidende Frage ist die des Qualitätskriteriums, auf das das Regenrat untersucht werden soll. Generell ist der notwendige Aufbereitungsgrad von Formstoff gießerei-, prozess- und produktspezifisch zu bewerten. Einen allgemeingültigen Qualitätsstatus festzulegen ist daher nicht Ziel dieser Arbeit. Die Definition, dass eine Regeneration der *Überführung eines Gießereialtsandes in einen Formgrundstoff mit neusandähnlichen Eigenschaften durch Entfernen verbrauchter Binderreste* [12, S. 248] entspricht, ist hier von zentraler Bedeutung: Neusand ist das eingesetzte Rohmaterial, sodass das Ziel festgelegt wird, dessen Eigenschaften mit dem Regenerat wieder so nah wie möglich zu kommen.

2. Lässt sich das festgelegte Qualitätskriterium akustisch erkennen?

Das Qualitätskriterium bezieht sich, wie zuvor definiert, auf die *neusandähnlichen Eigenschaften*. Das bedeutet, dass das Regenerat nach Durchführung des beschriebenen Regenerationsprozesses ähnliche Analysewerte aufweisen sollte wie Neusand. Der gewählte Lösungsansatz bezieht sich somit auf die Untersuchung akustisch wahrgenommener Unterschiede während des Prozessverlaufes. Zur Simulation der Sinneswahrnehmung *Hören* gibt es drei zu beachtende Aspekte [116, S. 207]:

- (a) Übertragung des hörbaren Schallstimulus zu den Rezeptoren,
- (b) Umwandlung des Stimulus in elektrische Signale,
- (c) Auswertung der Signale nach Schallparametern.

Um herauszufinden, inwieweit sich das im Regenerator befindliche Material auf die akustischen Emissionen des Prozesses auswirkt, wird jeder Regenerationsprozess aufgenommen (a). Das Mikrophon ist dabei direkt mit einem Computer verbunden, sodass das analoge Signal in ein digitales umgewandelt und als Audiodatei gespeichert wird (b). Mit der Auswertung des Signals wird dann ein akustisches Profil des Prozesses erstellt (c). Unter der Annahme, dass sich die Prozessakustik während des Regenerationsprozesses verändert, weil sich das Material im Inneren verändert, sieht das Vorgehen zur

Beantwortung dieser Frage wie folgt aus: Für den Vergleich von Regenerat und Neusand in einem akustischen Rahmen wird zunächst ein akustisches Profil von Neusand erstellt, indem der Regenerationsprozess von Neusand aufgenommen wird. Diese Aufnahmen dienen später als Vergleichsbasis. Der gesamte Regenerationsprozess des binderbehafteten Sandes wird ebenfalls aufgenommen. Die Audiodaten des Regenerates sind dann mit dem Profil des Neusandes zu vergleichen. Das akustische Profil eines Neusandes kann dabei nur für den selben Sandtyp verwendet werden; für die Erstellung weiterer Profile muss jedes System individuell erfasst und kalibriert werden.

### 3. Lässt sich der Status des Aufbereitungsprozesses akustisch erfassen?

Durch das mechanische Regenerieren sollen die Binderhüllen des belasteten Sandes entfernt werden, um diesen erneut zur Produktion von Kernen bzw. Gussformen zu verwenden. Mit dem Regenerationseffekt geht auch ein gewisser Quarzabrieb der Körner einher, der in diesem Fall implizit berücksichtigt werden soll. Dafür wird Neusand mit einer unüblich langen Regenerationszeit von 60 Minuten regeneriert und der Prozess akustisch aufgenommen, wodurch zum einen der beschriebene Effekt des Abriebs erfasst und zum anderen das mehrfache kurze Regenerieren (von z.B.  $4 \times 15$  Minuten) simuliert wird. Zusätzlich kann mit den Aufnahmen analysiert werden, ob sich innerhalb des bestehenden akustischen Profils zeitliche Klassen bilden lassen, die gewissen Veränderungen entsprechen, also einen Regenerationsstatus repräsentieren können. Zum Beispiel wäre es möglich, dass bei einer 15 minütigen Regeneration erfassbare Unterschiede zwischen den Minutenblöcken 2-6 und 10-14 bestehen.

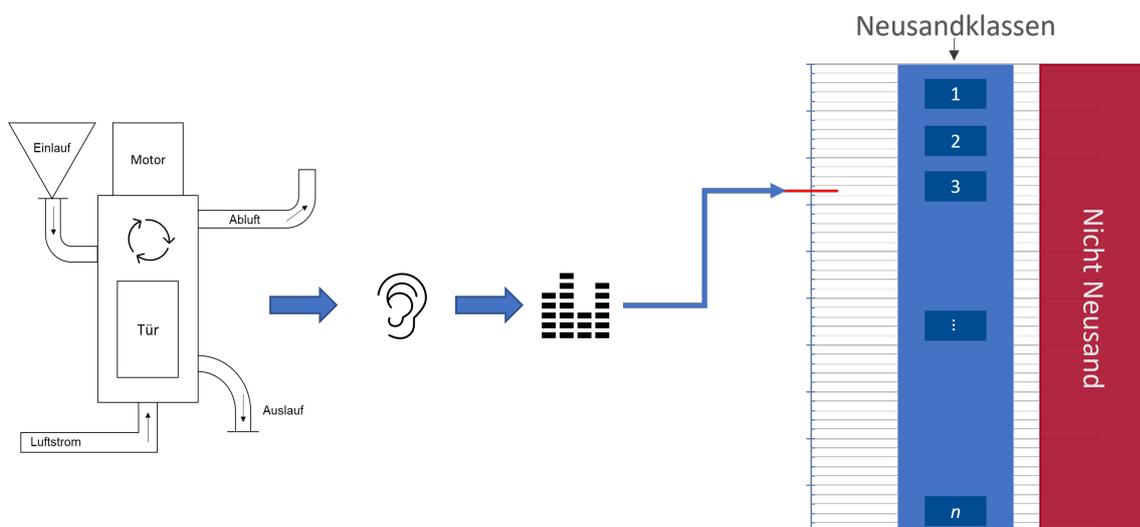


Abbildung 4.5: Lösungskonzept

Abbildung 4.5 beschreibt grafisch das beschriebene Vorgehen für die Beantwortung der Fragen 2 und 3. Der dargestellte Ablauf soll zunächst mit Neusand durchgeführt werden. Dem laufenden Prozess wird maschinell „zugehört“, die analogen Signale in digitale umgewandelt und falls möglich, in zeitliche Kategorien, die Neusandklassen, klassifiziert. Unter der Annahme, dass nun dieses akustische Profil des verwendeten Neusandtyps als Basis besteht, wird der Prozess mit binderbehaftetem Sand wiederholt. Auch diesem Prozess wird zugehört, die Signale umgewandelt und mit den Neusandklas-



schinen und Anlagen nach ISO 19816. Dieser Sensor misst die effektive Schwingungsgeschwindigkeit an einer nicht rotierenden Bauteiloberfläche, welche in ein Analogsignal am Stromausgang umgewandelt wird. Der Transmitter wurde am Träger des Korpus montiert, indem eine Gewindebohrung (M8) mit mindestens 10 mm Tiefe eingerichtet und der Sensor verschraubt wurde. Der Hersteller gibt für den verbauten Sensor eine Ansprechverzögerung von 0,5 Sekunden und einen Frequenzbereich zwischen 10 und 1000 Hz an [3].

Der *Strömungssensor SAxx04* der Firma *ifm*<sup>18</sup> ist ein Sensor zur Überwachung von flüssigen sowie gasförmigen Medien. Dabei werden die Prozessgrößen *Strömung* sowie *Medientemperatur* erfasst. Der Sensor erfasst die Strömung nach dem kalorimetrischen Messprinzip und ist am Zulaufrohr des Luftstromes verbaut.

Die Temperatursensoren sind möglichst ausgewogen über den Korpus verteilt, um das gesamte Temperaturspektrum erfassen zu können. Insgesamt sechs Temperaturwerte werden unmittelbar am bzw. im Korpus gemessen, wobei sich die empfindlichen Sensoren in einer Ummantelung befinden, um vor erhöhtem Verschleiß geschützt zu werden. Dadurch ist immer mit einer gewissen Trägheit der realen Informationsaufnahme zu rechnen. Ein Mantel-Thermoelement Typ K (-100 bis +1100 °C) wurde trotz des Verschleißrisikos durch eine Öffnung der Korpuswand direkt in die Anlage geleitet, um einen Referenzwert für die tatsächlich vorherrschende Sandtemperatur im Inneren der Anlage erfassen zu können. Zusätzlich wurden Sensoren im Auslass und im Abluftrohr verbaut. Die Medientemperatur der einströmenden Luft wird, wie bereits erwähnt, über den Strömungssensor erfasst.

Das Mikrofon ist ein USB-Kondensatormikrofon (Anua MIC-900) welches die Prozessgeräusche als 16-bit Stereosignal mit einer Abtastrate von 44100 Hz an einen angeschlossenen Laptop weitergibt. Dieser Mikrofontyp entspricht genormten Genauigkeitsanforderungen, die elementar für Absolutmessungen sind [86, S. 291]. Jeder Regenerationsprozess wird mit der frei verfügbaren Audioeditor und -rekorder Software *Audacity*<sup>19</sup> aufgenommen.

Die Steuerung des Reibregenerators erfolgt mit einem SIEMENS SIMATIC Automatisierungssystem (S7-1200)<sup>20</sup>. Dieses umfasst verschiedene speicherprogrammierbare Steuerungen (SPS), mit denen eine breite Palette von Automatisierungsaufgaben gelöst werden kann. Die Motorfrequenz sowie die Luftzufuhr lassen sich über die Steuerung manuell einstellen oder als Automatisierung programmieren.

### 4.3.2 Versuchs- und Datenvorbereitung

#### Prozessdaten

Im Allgemeinen haben Maschinen-, Apparat- und Prozessparameter einen Einfluss auf die charakteristischen Parameter eines Stoffsystems. Dazu können die Betriebsweise (kontinuierlich oder diskontinuierlich), die Betriebsbedingungen (Drehzahl, Temperatur, Druck, Dauer, Durchsatz, Geschwindigkeit), die Maschineneinstellungen, die Dosierungen oder Konzentrationen, die Produktzufuhren bzw. -entnahmen oder die Reaktionsführungen und die zeitliche Fahrweise zählen [121, S. 5]. Für die durchgeführten Versuche mit dem Reibregenerator ist daher eine Parametereinstellung gewählt worden, welche über alle Versuche hinweg nicht verändert wurde. Dadurch wurden mögliche anlagenspezifische Einflüsse konstant gehalten, um keine ungewollten Abweichungen in den Messwerten zu verursachen. Die in drei Dateien gespeicherten Prozessdaten jeder Regeneration wurden ebenfalls immer nach der selben Vorgehensweise zusammengeführt und aufbereitet. Alle erfassbaren Werte speichert die Anlage sekundlich und mit einem Zeitstempel inkl.

<sup>18</sup><https://www.ifm.com/de/de> (besucht am 28.03.2022)

<sup>19</sup><https://www.audacity.de/> (besucht am 10.06.2022)

<sup>20</sup><https://support.industry.siemens.com/cs/start?lc=de-WW> (besucht am 10.06.2022)

Datum. Parameter mit identischem Zeitstempel wurden mit den zugehörigen Messwerten untereinander, also blockweise, aufgelistet. Für die Auswertung sowie zum Vergleich der Regenerationsprozesse wurden auf Prozessdatenbasis folgende Schritte immer identisch durchgeführt:

**1. Daten zusammenführen:**

Die erfassten Datenblöcke der Prozess- und Temperaturdaten werden untereinander in eine neue Datei kopiert und nach dem Zeitstempel sortiert, um wieder eine korrekte zeitliche Reihenfolge der Parameter herzustellen.

**2. Zuordnung wichtiger Meldungen:**

Über das Meldearchiv können die genauen Zeitpunkte des hier entscheidenden Ein- und Auslaufes den gesamten Prozesswerten zugeordnet werden, womit sich einzelne Regenerationen aus einem Mehrfach- bzw. Dauerbetrieb leicht identifizieren lassen.

**3. Normierung der Regenerationen:**

Jeder Regenerationsvorgang wird von Beginn bis zum Ende zeitlich normiert. Da die Prozesswerte sekundlich abgespeichert werden, bietet es sich an, den Prozess sekundlich zu normieren, sodass der Start einer jeden Regeneration von null Sekunden bis zum Endzeitpunkt des Prozesses durchgezählt wird.

**4. Umstrukturierung der Daten:**

Die einzelnen Parameter werden tabellarisch untereinander erfasst. Mit Hilfe einer Pivot-Tabelle lassen sich die Daten so umstrukturieren, dass die Parameter als Variablenüberschrift in die oberste Zeile verschoben werden und die jeweiligen Messwerte dem Zeitverlauf entsprechend darunter zugeordnet werden.

Nach der Zusammenführung und Umstrukturierung werden die Rohdaten im nächsten Schritt der Aufbereitung von Ausreißern befreit, denn die Steuerung speichert während des Prozesses vereinzelt unbrauchbar hohe Werte. Dieses Phänomen tritt bei vereinzelt Datensätzen und nicht bei jeder Regeneration auf, wodurch diese Ausreißer leicht zu identifizieren sind. Dieser Fehler wird durch das Löschen des Datensatzes eliminiert und führt zu keinem signifikanten Informationsverlust. Eine zweite Eigenheit der Datenerfassung ist das regelmäßige Überspringen einer Sekunde in der Messdatenspeicherung. Das bedeutet, dass etwa alle 90 Sekunden eine zeitliche Lücke in den Rohdaten entsteht. Die Temperaturdaten des externen Temperatursensors erfassen jedoch diese Sekundenwerte, wodurch ein Datensatz für die gesamten Prozessdaten mit nur einem Messwert entsteht, welcher ebenfalls während der Aufbereitung entfernt wird. Aufgrund der vorliegenden hohen Datendichte wurde auf das synthetische Ersetzen dieser fehlenden Werte verzichtet.

## **Audiodaten**

Die Anlage befindet sich in einer der großen Forschungshallen der Universität Duisburg-Essen, in der einige wenige Büros, eine Vielzahl weiterer Klein- sowie Großanlagen und unterschiedlichste Analysevorrichtungen untergebracht sind. Die akustischen Aufnahmen der Regenerationsprozesse wurden immer von der gleichen Mikrofonposition durchgeführt und es wurde für alle Aufnahmen stets darauf geachtet, dass keine parallel laufenden Experimente anderer Forschungskollegen\*innen in der selben Halle durchgeführt wurden. Dennoch ließen sich vereinzelt Störgeräusche nicht vermeiden. Die Aufnahme jedes einzelnen

durchgeführten Regenerationsprozesses wurde bereits vor dem Einlauf des Regenerates gestartet und erst nach dem Auslauf gestoppt. Damit wurde sichergestellt, dass der gesamte Prozess akustisch erfasst wurde. Die Tonspuren liegen in dem softwarespezifischen *aup*-Format sowie als exportierte *wav*-Dateien vor. Für die Signalanalysen sind der Ein- und Auslauf sowie auffällige akustische Störungen nicht relevant. Im ersten Schritt zur Datenaufbereitung wurden daher alle aufgenommenen Rohdaten so zugeschnitten, dass nur die Aufnahme des reinen Regenerationsprozesses übrig blieb. Die Gesamtaufnahmen der einzelnen Regenerationen weisen Längen von insgesamt 5 bis 60 Minuten auf. Als Gesamtdatei ließen sich diese Daten nicht sinnvoll verarbeiten, sodass eine weitere Aufteilung notwendig war. Die gewählte Länge zur Aufteilung der Audiospuren orientierte sich an dem ungefähren Durchschnitt von 3 Sekunden, der sich aus der Literatur zu Audioklassifikationen und bestehenden, frei zugänglichen, akustischen Datenmengen wie UrbanSound8K und ESC-50 ergibt, vgl. Tabelle 4.1. Auftretende Störgeräusche wurden bereits während der Aufnahmen dokumentiert sowie im Nachhinein durch eine erneute akustische manuelle Prüfung der Tonspur identifiziert und für die Analysen eliminiert, sodass nur saubere und qualitativ hochwertige Audio-Samples verblieben.

Tabelle 4.1: Verwendete Audiodatenlänge in der Forschung

Quelle/Datensatz:	[66]	[111]	[24]	[5]	[142]	[117]	ESC-50	UrbanSound8K
Länge der Audiodateien:	1 s	4 s	3 s	2,56 s	2 s	3 s	5 s	≤ 4 s

## 4.4 Vorversuche/Vorbereitung

Den wesentlichen Analysen hinsichtlich der vorliegenden Problemstellung gingen einige notwendige Vorversuche voraus. Dazu zählten zum einen Versuche zur Reproduzierbarkeit der erzeugten Daten und zum anderen die Untersuchung der Regenerationsleistung der Anlage, um mit den klassischen Formstoffqualitätskriterien die erreichbare Regeneratqualität beurteilen zu können.

### 4.4.1 Reproduzierbarkeit

Unter der Reproduzierbarkeit wird die Vergleichbarkeit der Messung von Versuchen, die unter gleichen Bedingungen durchgeführt wurden, verstanden. Diese setzt voraus, dass der Arbeitsablauf sowie die Arbeitsbedingungen bekannt sind und die Erfassung der Daten mit der notwendigen Sorgfalt durchgeführt wird. Um die Reproduzierbarkeit der Prozess- sowie der Audiodaten für den hier vorliegenden Prozess nachzuweisen, wurde überprüft, ob Regenerationsprozesse mit gleichbleibenden Parametereinstellungen und identischem Chargengewicht einen Output produzieren, der sich im Vergleich untereinander sehr ähnelt bzw. nur eine geringe Schwankungsbreite aufweist.

#### Prozessdaten

Für die Prozessdaten wird die Korrelationsmatrix genutzt, welche alle wechselseitigen Korrelationen innerhalb einer Datenmenge wiedergibt. Die Korrelation beschreibt die *Stärke* des linearen Zusammenhanges zwischen zwei Stichproben X und Y. Für die Analyse werden jeweils Wertepaare  $(x_1, y_1)$ ,  $(x_2, y_2)$ , ... ,

$(x_n, y_n)$  gebildet und der Korrelationskoeffizient

$$r = r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4.4.1)$$

berechnet [26, S. 519], [32, S. 126].

Das Ergebnis liefert  $r$ -Werte in einem Bereich von  $-1 \leq r \leq 1$ . Bei  $r = 0$  liegt keine Korrelation vor: Es wird von einem unkorrelierten Wertepaar gesprochen, bei dem kein linearer Zusammenhang besteht. Eine perfekte lineare Korrelation liegt bei  $|r| = 1$  vor und je näher  $r$  dem Betragswert 1 kommt, desto stärker ist der vorliegende lineare Zusammenhang der untersuchten Variablen. Ein positives Vorzeichen ( $r > 0$ ) wird als positive Korrelation interpretiert, sodass von einer Tendenz ausgegangen wird, bei der die Werte  $x_i, y_i$  um eine Gerade mit positiver Steigung liegen.  $r < 0$  ist entsprechend die negative Korrelation, also die Tendenz, dass die Werte  $x_i, y_i$  um eine Gerade mit negativer Steigung liegen. [32, S. 131], [114, S. 381–382], [26, S. 521] Die symmetrische Korrelationsmatrix besitzt auf der Hauptdiagonalen ausschließlich den Wert 1, da der Korrelationskoeffizient einer Variable, die mit sich selbst korreliert wird, immer 1 ergibt [122, S. 160–161].

Zur Analyse des Anlagenverhaltens hinsichtlich der Reproduzierbarkeit dienen die gespeicherten Prozessdaten, wobei die verwendete Datenbasis aus 15 Parametern und 40495 Datenzeilen besteht, die über insgesamt 11 Regenerationsprozesse aufgenommen wurden. Für diese Abhängigkeitsuntersuchung wurde explizit darauf geachtet, die variablen Prozessparameter der Zeit, der einstellbaren Motorfrequenz sowie der durchströmenden Luftmenge für jede Regeneration zu verändern. Die Korrelationsmatrix in Abbildung 4.7 stellt die verfügbaren Variablen als Wertepaar mit sich selbst und mit jeweils allen weiteren Parametern dar. Die zugehörigen Korrelationskoeffizienten werden in den einzelnen Matrixfeldern wiedergegeben, welche zusätzlich farblich in ihrer Wertigkeit repräsentiert werden. Die erste Zeile der Matrix zeigt eine hohe Korrelation zwischen Laufzeit und Temperatur der Anlage, mit Werten zwischen 0,62 und 0,82. Neben der Zeit haben auch die Motordaten des Frequenzumrichters (Drehmoment, Frequenz, Strom und Wirkleistung) einen Einfluss auf die Temperatur. Ausgehend von der Einstellbarkeit der Frequenz ist klar, dass die Parameter des Frequenzumrichters sich gegenseitig beeinflussen und sehr hohe Korrelationen aufweisen. Durch die Korrelation zwischen Frequenz und Wirkleistung mit 0,91 haben entsprechend auch beide Parameter Einfluss auf die entstehenden Temperaturen, was in den Zeilen *FU\_Frequenz* und *FU\_Wirkleistung* der Abbildung 4.7 erkennbar ist. Ein ähnliches Einflussverhalten zeigen die Daten des Schwingungssensors, welcher zudem mit  $r = 0,54$  eine gewisse Abhängigkeit zur einstellbaren Frequenz aufweist. Die untere Hälfte der Korrelationsmatrix stellt alle erfassten Temperaturdaten gegenüber. Im vierten Quadranten, dem Temperatur-Block unten rechts der Matrix, ergeben sich sehr hohe Korrelationen der Temperaturdaten untereinander. Dies weist auf einen gleichmäßigen und stabilen Temperaturverlauf der Anlage hin. Mit diesem Erkenntnis könnte auf einige der Sensoren für die Temperaturerfassung verzichtet und das Sensorsystem verschlankt werden, ohne Informationen über den Prozess zu verlieren. Auch das Material sowie die eingesetzte Menge des Materials, welches in der Anlage regeneriert wird, spielen eine Rolle in Bezug auf das Anlagenverhalten. Diese Einflüsse werden jedoch nicht durch das Sensorsystem erfasst und fließen damit nicht in diese vorliegende Analyse ein.

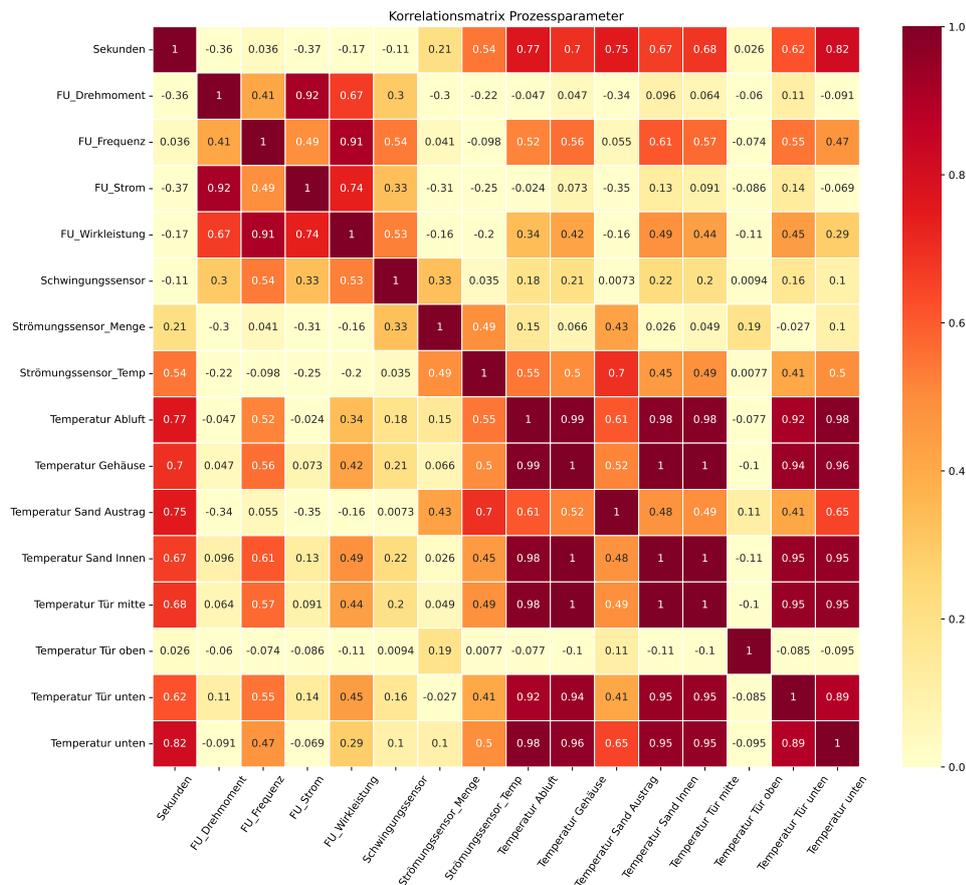


Abbildung 4.7: Korrelationsmatrix der Prozessdaten

### Audiodaten

Für den Vergleich von Audiodaten wird häufig die *Kreuzkorrelation* von zwei zeitabhängigen Signalen  $x(t)$  und  $y(t)$  verwendet. Im Allgemeinen stellt die Kreuzkorrelation ein Maß für die Ähnlichkeit zweier Signale dar [38, S. 27–28], das die gegenseitige lineare Abhängigkeit beschreibt [72, S. 414]. Zusätzlich berücksichtigt die Kreuzkorrelation die funktionale Abhängigkeit der Ähnlichkeit von einer zeitlichen oder örtlichen Verschiebung der beiden Signale. Es kann demnach das Maß der Übereinstimmung zweier Signale für jede Verschiebung bestimmt werden [84, S. 90], um genau die Verschiebung zu identifizieren, an der die Übereinstimmung am höchsten ist. Die Verwendung der Kreuzkorrelation und die Erfassung einer möglichen Signalverschiebung wird beispielsweise in der GPS-Anwendung (Global Positioning System) genutzt, welche zur Ermittlung der geografischen Position dient. Die Position wird aus der Entfernung zu mehreren Satelliten berechnet, indem eine Repliksequenz des Nutzers erzeugt wird. Diese Repliksequenz entspricht der verzögerten Version der vom GPS-Satelliten gesendeten Sequenz. Durch die Messung dieser Verzögerung kann die Entfernung zwischen dem GPS-Satelliten und dem Nutzer bestimmt werden, die dann zur Bestimmung der Nutzerposition verwendet wird [84, S. 90], [105, S. 95].

Die Kreuzkorrelation kann für den diskreten und kontinuierlichen Fall berechnet werden. Für zeitdis-

krete reelle stationäre Prozesse wird die Kreuzkorrelationsfunktion (KKF)

$$\varphi_{xy}(\kappa) = \sum_{n=-\infty}^{\infty} x[n] \cdot y[n + \kappa] \quad (\text{diskret}) \quad (4.4.2)$$

$$\varphi_{xy}(\tau) = \int_{-\infty}^{\infty} x(t) \cdot y(t + \tau) dt \quad (\text{kontinuierlich}) \quad (4.4.3)$$

aus den Signalen  $x(t)$  und  $y(t)$  berechnet, indem diese gegeneinander verschoben werden und für jede Verschiebung (engl. *lag*) die Übereinstimmung berechnet wird: Dafür werden beide Signale miteinander korreliert, also die Summe der Produkte korrespondierender Werte beider Signalfolgen gebildet [84, S. 91]. Wenn sich beide Funktionen ähnlich sind, ergibt sich ein Maximum in der KKF.

Für die Zeitdifferenz wird im kontinuierlichen Fall die Laufvariable  $\tau$  und im diskreten  $\kappa$  anstelle von  $t$  für die (absolute) Zeit verwendet [38, S. 27–28]. Aufgrund der Stationarität der Prozesse wird für die Zeitmittelwerte von einer unendlich ausgedehnten Musterfolge ausgegangen, sowohl in die negative als auch in die positive Richtung. Für die Verschiebung sind alle ganzen Zahlen zugelassen [136, S. 272]. Die Funktion liefert Betragswerte zwischen 0 und 1. Negative Werte bis  $-1$  sind bei gegenphasigen Signalen ebenfalls möglich. Entspricht das Ergebnis der normierten KKF dem Wert 1, sind die Signale identisch, bei einem Wert von 0 sind die Signale unkorreliert [14]. Je ähnlicher die Signale also zu einem gewählten Zeitpunkt sind, desto größer wird der Wert der berechneten Kreuzkorrelation [102, S. 46]. Die Kreuzkorrelationsfunktion ist weiterhin die Basis der spektralen Kreuzleistungsdichte  $S_{xy}(f)$ . Diese wird durch die Fourier-Transformation der KKF gebildet und ist für die Berechnung der Übertragungsfunktion und der Kohärenz wichtig [72, S. 414–415].

In dem hier vorliegenden Fall handelt es sich jedoch nicht um Audiosignale, die aufgrund verschiedener Sensoren zeitversetzt aufgenommen wurden, sondern um Signale aus jeweils eigenständigen Versuchen. Da also kein zeitlicher Versatz vorliegen kann, bietet eine Untersuchung im Zeitbereich keinen Mehrwert. Durch die Fourier-Transformation wird vom Zeit- in den Frequenzbereich des Signals gewechselt. Da die Kreuzkorrelation auf den Zeitbereich ausgelegt ist, ist die Aussagekraft bei Anwendungen auf den Frequenzbereich stark eingeschränkt. Aus diesem Grund wird zudem die klassische Korrelation auf die fouriertransformierten Signale angewendet.

Für die Analyse der Reproduzierbarkeit der Audiodaten wird der Prozess mit gleichbleibenden Parametereinstellungen und mit jeweils einer einheitlichen Menge Neusand wiederholt. Mit der verfügbaren Sandmenge konnten insgesamt 9 Versuche durchgeführt werden. Daraus ergeben sich entsprechend neun Audiodateien, die diese Prozesse akustisch wiedergeben. Die vorliegenden Aufnahmen sind ca. 30 Minuten lang und wurden zum besseren Handling in 10-Sekunden-Aufnahmen zerschnitten. Die verbleibenden Sekunden einer Gesamtdatei, mit denen keine komplette 10s-Datei erzeugt werden konnte, wurden hier vernachlässigt. Die erzeugten 10-Sekunden-Audiodateien wurden bei 0 beginnend durchnummeriert, um die zeitliche Abfolge nachvollziehen zu können. Einlauf-, Auslauf- sowie Störgeräusche wurden aus den Audiodaten entfernt. Für den Vergleich der Aufnahmen wurden zufällig 150 Signalpaare aus insgesamt 1523 Dateien gebildet, wobei die Auswahl unter den Restriktionen, dass die Dateien eines Paares aus zwei unterschiedlichen Prozessaufnahmen stammen und keine doppelten Paare bzw. Dateien auftreten, erfolgte. Die Paarwahl sowie die Berechnung der Korrelationskoeffizienten wurden in Python durchgeführt. Tabelle 4.2 listet die daraus resultierende Gesamtanzahl verfügbarer Daten für den Ähnlichkeitsvergleich in den ersten beiden Spalten sowie einige der 150 Vergleichspaare als Beispiel in der dritten Spalte auf. Die Dateinamen enthalten an zweiter Stelle die Zuordnung, aus welchem Regenerationsversuch sie stammen,

und an dritter Stelle die Nummerierung zur zeitlichen Zuordnung. Das bedeutet, dass das Beispieldatenpaar der ersten Tabellenzeile aus 4.2 „Sandsound V6 53 & Sandsound V2 101“ sich aus der 53. Datei der sechsten Regenerationsaufnahme und der 101. Datei des zweiten Regenerationsversuches zusammensetzt.

Tabelle 4.2: Dateiübersicht der Versuchsreihe zur Prozessstabilität

Versuche	Anzahl 10s-Datei	Beispielpaare	
Prozessstabilität V1	180	Sandsound V6 53	Sandsound V2 101
Prozessstabilität V2	181	Sandsound V7 7	Sandsound V9 40
Prozessstabilität V3	180	Sandsound V2 94	Sandsound V5 106
Prozessstabilität V4	183	Sandsound V3 98	Sandsound V6 143
Prozessstabilität V5	179	Sandsound V1 151	Sandsound V8 21
Prozessstabilität V6	180	Sandsound V4 11	Sandsound V1 12
Prozessstabilität V7	180	Sandsound V8 122	Sandsound V7 45
Prozessstabilität V8	180	Sandsound V5 31	Sandsound V3 57
Prozessstabilität V9	180	Sandsound V9 171	Sandsound V2 30
<b>Summe aller Dateien:</b>	<b>1523</b>	<b>Summe Versuchspaare: 150</b>	

Häufig wird auf die Logarithmierung von Signalen zurückgegriffen, um einen Wertebereich über mehrere Größenordnungen, wie z.B. sehr kleine und sehr große Werte, besser abbilden zu können, wie es in Kapitel 2.2 erläutert wurde. Durch die Logarithmierung wird der Wertebereich zu deutlich kleineren Intervallen zusammengefasst. Dadurch lassen sich die Unterschiede der Größenordnungen in Diagrammen übersichtlicher darstellen als in linearem Maßstab. Für eine bessere Visualisierung wird daher häufig auch die Diagrammchse logarithmiert, damit auch sehr kleine bzw. sehr große Werte sichtbar werden [44, S. 59], [70, S. 60, 167]. In diesem Fall wurde die Logarithmierung im Frequenzbereich auf die Absolutwerte der fouriertransformierten Signale angewendet.

Für die insgesamt 150 Dateipaare wurde jeweils die Kreuzkorrelation der logarithmierten fouriertransformierten Signale, die Korrelation der fouriertransformierten Signale und die Korrelation der logarithmierten fouriertransformierten Signale berechnet. Tabelle 4.3 fasst die Minimal- und Maximalwert der berechneten Ergebnisse zusammen, aus denen die Spannweite der Ergebnisse interpretiert werden kann. Des Weiteren sind der Mittelwert und der Median angegeben.

Tabelle 4.3: Ergebnisse der Korrelationsuntersuchungen

	Kreuzkorrelation (log)	Korrelation	Korrelation (log)
<b>Minimum</b>	0,84	0,56	0,84
<b>Maximum</b>	0,88	0,74	0,87
<b>Mittelwert</b>	0,86	0,64	0,86
<b>Median</b>	0,86	0,63	0,85

#### 1. Interpretation der Ergebniswerte der Kreuzkorrelation (log):

Die Kreuzkorrelation wurde im logarithmierten Frequenzbereich durchgeführt und weist hohe Korrelationsergebnisse zwischen 0,84 und 0,88 auf. Für ca. 80 % dieser Ergebnisse lag die höchste Korre-

lation im Fall keiner Verschiebung ( $lag = 0$ ) vor, für die restlichen 20 % trat eine Verschiebung auf, die jedoch keine Regelmäßigkeiten aufwies. Da es sich um Informationen zur Frequenzverschiebung handelt und diese für die Reproduzierbarkeit nicht sinngemäß interpretiert werden können, wurden für die selbe Datenbasis zwei weitere Berechnungen durchgeführt.

## 2. Interpretation der Ergebniswerte der Korrelation:

Die mittlere Spalte der Tabelle 4.3 gibt die Korrelationskoeffizienten der fouriertransformierten Signale wieder. Hier ist eine größere Spannweite der Ergebnisse zu erkennen, die Werte liegen in einem Bereich von 0,56 bis 0,74. Mit Werten von 0,64 bzw. 0,63 für Mittelwert und Median liegen jedoch keine sehr hohen Korrelationskoeffizienten für die Datenpaare vor. Durch die hohe Bandbreite an Frequenzwerten akustischer Signale, die sowohl sehr kleine als auch sehr große Werte umfasst, ist allerdings die Korrelation der logarithmierten Daten deutlich aussagekräftiger.

## 3. Interpretation der Ergebniswerte der Korrelation (log):

Die sich ergebenden Korrelationskoeffizienten zeigen deutlich höhere Werte, die den Ergebnissen der zuerst durchgeführten Kreuzkorrelationsberechnung sehr nahe kommen. Mit einem Mittelwert von 0,86 und dem Median von 0,85 kann den untersuchten Dateipaaren eine hohe lineare Abhängigkeit nachgewiesen werden. Damit ist die Reproduzierbarkeit der Akustikdaten belegt, sodass für die folgenden Analysen davon ausgegangen wurde, dass bei konstant gehaltenen Bedingungen auch vergleichbare akustische Signale aufgenommen werden.

### 4.4.2 Bewertung des Regenerationsprozesses

Obwohl der zentrale Fokus der vorliegenden Problemstellung auf der Verarbeitung und Analyse der akustischen Sandprofile liegt, ist die Quantifizierung der Wirkleistung der Anlage ebenfalls wichtig. Im Folgenden wird daher zunächst die Versuchsdurchführung und die Auswertung der Siebanalyse sowie der Biegebiegefestigkeit beschrieben.

#### Versuchsdurchführung

Zur Verifizierung der erzielbaren Sandqualität nach dem Regenerationsprozess ist eine umfangreiche Versuchsreihe durchgeführt worden. Dafür wurde Quarzsand F34 als Ausgangsmaterial verwendet und jeweils drei separate Chargen mit einer Sandmenge von je 30 kg vorbereitet. Jede Charge sollte den in Abbildung 4.8 dargestellten Kreislauf insgesamt 10 mal durchlaufen; dieses Vorgehen simuliert das übliche Umlaufsystem von Formstoff in Gießereien. Auch die *Klein Anlagebau AG* hat mit einem zehnfachen Umlaufsystem ihr entwickeltes Clustreg-Verfahren verifiziert [108].

Durch die aufgezeigte Vorgehensweise (Abbildung 4.8) wird deutlich, dass der Sand in diesem Kreislauf nicht zusätzlich thermisch belastet, also kein Abguss durchgeführt wird. Bei einem Gießvorgang kommt nur die äußere Schicht der gesamten Form direkt mit dem flüssigen Metall und damit mit hohen Temperaturen in Kontakt. Aufgrund der generell eher geringen thermischen Belastung (bis maximal 900 °C) anorganisch gebundener Systeme und der sofort einsetzenden Abkühlung, bleiben einige Bereiche im Form- sowie Kernpaket auf Raumtemperaturniveau [51]. Zusätzlich können gebrochene Kerne bzw. Formen als Grund für die Untersuchung nicht thermisch belasteter Systeme genannt werden. In diesem Fall wurde auf das Erhitzen des angemischten Sandes verzichtet, ebenso wie bei den Untersuchungen durch Kang, Kim et al.

(2018) und Sun, Jing et al. (2011) [63, 123]. Da es sich nur um binderbelasteten Sand handelt, der nicht thermisch belastet wurde, wird im Folgenden von *Regenerat* anstelle des Begriffes Formstoff gesprochen.

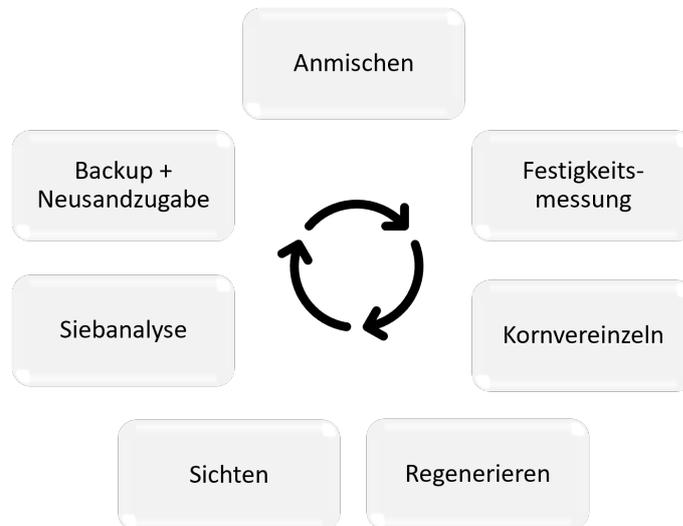


Abbildung 4.8: Vorgehensweise Versuchsreihe

Für das **Anmischen** wurde ein Flügelmischer mit 10kg Fassungsvermögen verwendet, sodass der Mischvorgang für jede Charge dreimal wiederholt wurde, um die gesamte Sandmenge mit dem Bindersystem zu benetzen. Dafür wird nach einer Homogenisierung von 60 Sekunden zuerst flüssiger Härter und nach weiteren 30 Sekunden ebenfalls flüssiger Binder der Firma *Technofond Gießereihilfsmittel GmbH* hinzugegeben. Die Mischzeiten sind in der Anlagensteuerung festgelegt und damit für jeden Mischprozess identisch. Als Binder wurde Tefosil 10 verwendet, ein flüssiger alkalischer Geopolymerbinder auf Basis anorganischer Polymere zur CO<sub>2</sub>- und Selbstaushärtung. Die Binderdosierung bei Selbstaushärtung, wie in diesem Fall, liegt zwischen 1,6–2,4 Gewichts-%. Der Härter (H-TC 50) ist für das anorganische Geopolymer-Bindersystem ausgelegt. Die empfohlene Dosierung liegt zwischen 10,0 bis 17,0 Gew.-% bezogen auf die Bindermenge. Die Binder- und Härtermenge blieb für die Vergleichbarkeit für jeden Umlauf gleich. Der angemischte Sand wird im zweiten Schritt zu Biegeriegeln geformt und händisch verdichtet; dieses Vorgehen wurde in Kapitel 2.1.3 beschrieben. Mit fünf Formkästen, die pro Charge jeweils zweimal mit Sand befüllt wurden, konnten insgesamt 84 Biegeriegel geformt und anteilig nach Aushärtzeiten von 30, 60, 90, 150, 270 und 1470 Minuten nach dem Mischprozess vermessen werden (**Festigkeitsmessungen**). Nachdem alle Messwerte erfasst wurden, mussten die gebundenen und verdichteten Körner wieder **vereinzelt** werden; dies erfolgte über einen mechanischen Rüttelrost. Durch die Rüttelbewegung reiben sich die Körner nach und nach von den Biegeriegeln ab und werden vereinzelt wieder aufgefangen. Der **Regenerierungsprozess** der drei Chargen läuft immer identisch ab. Die Anlage wird zuvor durch eine Spülcharge aufgewärmt, damit die Ausgangsbedingungen der Regeneration einen kontinuierlichen Prozess simulieren. Die drei Chargen haben jeweils unterschiedliche Regenerationszeiten, nämlich 5, 10 und 15 Minuten. Je nach Regenerationsverfahren können die Regenerationszeiten deutlich variieren. Für die hier betriebene mechanische Regeneration lassen sich vergleichbare Prozesszeiten in den Veröffentlichungen [21] und [118] finden. Für die Vergleichbarkeit der Ergebnisse sind die variablen Parameter *Motorfrequenz* und *Strömungsmenge* bei jedem Prozess gleichbleibend. Auch wenn der Reibregenerator bereits ein Abluftsystem hat, wurden die regenerierten und abgekühlten Sande zusätzlich **gesichtet**, um verbleibende Staub-

und Mikropartikel aus den jeweiligen Chargen zu entfernen. Das so entstandene Regenerat wurde auf seine granulometrischen Eigenschaften untersucht, wobei die Korngrößenverteilung das wichtigste petrografische Merkmal für die Klassifizierung von Kiesen und Sanden ist und mit der **Siebanalyse** ermittelt wird [47, S. 34]. Für jedes Regenerat und für jeden untersuchten Sand wurden insgesamt drei Siebanalysen mit jeweils 50g durchgeführt, wofür Siebe der Firma Retsch<sup>21</sup> verwendet wurden. Die Messwerte wurden durch die Angabe des Mittelwertes, des Medians sowie des minimalen und maximalen Wertes zusammengefasst bzw. ausgewertet. Bevor der Zyklus von vorne beginnt, wurden aus jeder Charge 5 kg Sand als Backup separiert und luftdicht verschlossen eingelagert. Die **Backups** haben zwei Funktionen: Zum einen können notwendige Analysen zu einem späteren Zeitpunkt erneut durchgeführt werden und zum anderen simuliert der Vorgang den auftretenden Sandverlust während des Prozesses [8, S. 73] bei gleichzeitiger **Zugabe von Neusand**. Für solche Umlaufsysteme kann in Abhängigkeit der Anzahl  $n$  Umläufe und des anteiligen Zusatzes bzw. der Abnahme  $z$  die Neusandkonzentration

$$K_n = (1 - (1 - z)^n) \cdot 100\% \quad (4.4.4)$$

berechnet werden [8, S. 68]. Für die vorliegenden Versuche befinden sich 30 kg Sand im Umlauf, wobei pro Umlauf durchschnittlich 6 kg Neusand hinzugegeben werden. Damit kann eine Neusandkonzentration von  $K_{10} = 89,26\%$  für den zehnten Umlauf berechnet werden.

Bevor die Auswertung der Ergebnisse erfolgt, ist darauf hinzuweisen, dass die gesamte Versuchsreihe nur einmal durchgeführt wurde. Daher können die Ergebnisse dieser Versuchsreihe selbst nicht als signifikant eingestuft werden. Es wird angenommen, dass bei der Wiederholung der Versuchsreihe unter identischen Bedingungen ähnliche, einer zufälligen Abweichung unterliegende, Ergebnisse erzielt werden können. Wie in der Versuchsdurchführung beschrieben, wurden jedoch bei der Durchführung der Festigkeitsmessungen sowie bei der Siebanalyse mehrere Messwerte erzeugt, um diese zu einem Wert zu mitteln.

Die erste Versuchslinie mit einer Regenerationszeit von 5 Minuten musste bereits nach dem dritten Umlauf abgebrochen werden, weil die Riegel bei der Entnahme aus den Formkästen zerbrachen. Zu diesem Zeitpunkt hatten sie bereits 30 Minuten in der Form ausgehärtet und hätten eine ausreichende Festigkeit für die Entnahme erreichen sollen. Ausgehend von diesem Ergebnis kann die Schlussfolgerung gezogen werden, dass eine Regenerationszeit von 5 Minuten nicht ausreichend ist, um das hier untersuchte Regenerat längerfristig einzusetzen.

### Auswertung der Siebanalyse

Der für die Versuchsreihe verwendete Quarzsand F34 weist folgende dargestellte Korngrößenverteilung auf: In den ersten drei Kornklassen (1,4–1,0 mm), (1,0–0,71 mm) und (0,71–0,5 mm) sowie in der letzten Kornklasse (0,02–0 mm) befinden sich keine bzw. vernachlässigbare Mengen Sand. Ab der vierten Kornklasse (0,5–0,335 mm) baut sich über die fünfte Klasse (0,335–0,35 mm) das Maximum der Verteilung in der sechsten Klasse (0,25–0,18 mm) auf. Mit einem Anteil von 51,06 % befindet sich hier etwa die Hälfte der vermessenen Probenmenge. Die anteilig zweitgrößte Menge befindet sich in der siebten Kornklasse (0,18–0,125 mm), wobei die Sandmengen in den Klassen 5 und 7 nicht gleich verteilt vorliegen. In den verbleibenden Klassen 8 (0,125–0,09 mm), 9 (0,09–0,063 mm) und 10 (0,063–0,02 mm) sammelt sich der restliche Anteil von insgesamt ca. 5,5 %, wengleich sich der Großteil dieses Anteils auf dem achten Sieb befindet. Die für den untersuchten Quarzsand F34 beschriebene Korngrößenverteilung, welche eine annähernd

<sup>21</sup><https://www.retsch.de/> (besucht am 13.01.2022)

glockenförmige Verteilungskurve aufweist, ist in Abbildung 4.9 dargestellt.

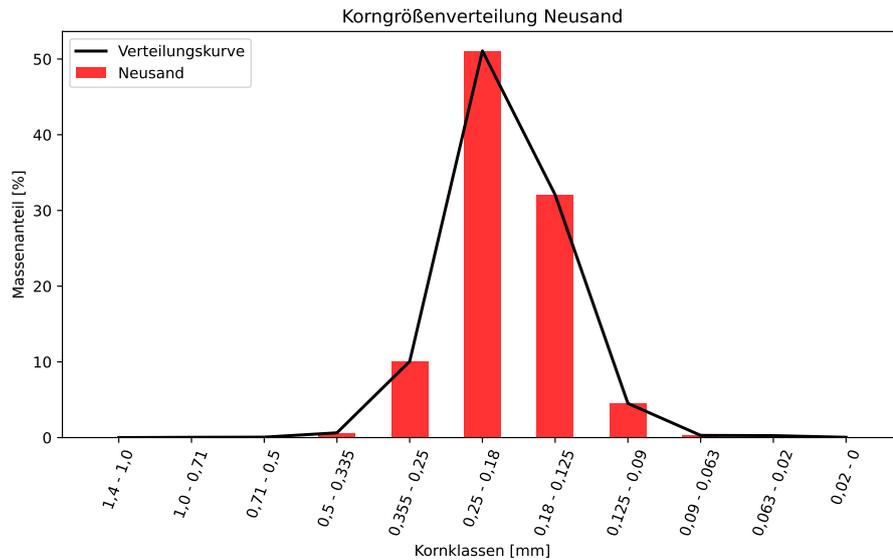


Abbildung 4.9: Korngrößenverteilung des Neusandes (Quarzsand: F34)

Neben dem Histogramm können aus den Daten der Siebanalyse zudem die Kennwerte der theoretischen spezifischen Oberfläche, die AFS-Feinheitsnummer, der Gleichmäßigkeitsgrad sowie die mittlere Korngröße berechnet werden. Die Formeln zur Berechnung dieser Kennwerte wurden in Kapitel 2.1.3 behandelt. Tabelle 4.4 enthält die Daten für den untersuchten Neusand in der obersten Zeile, die Werte dienen als Ausgangspunkt für den Vergleich mit den Daten der durchgeführten Untersuchungen der Regenerate.

Für die drei Regenerationszeiten 5, 10 und 15 Minuten werden die entsprechenden Kennwerte jeweils untereinander abgegrenzt und den entsprechenden Umläufen zugeordnet. Für die theoretische spezifische Oberfläche ergeben sich im ersten Umlauf für alle drei Regenerationszeiten im Vergleich zu Neusand einmalig höhere Werte. Für alle weiteren Umläufe stellt sich ein tendenziell sinkender Verlauf der berechneten Werte ein. Kleinere Abweichungen lassen sich durch Schwankungen aufgrund von Messfehler erklären. Die sinkenden Werte der theoretischen spezifischen Oberfläche weisen auf eine stetige Zunahme der Partikelgröße hin, welche sich auch in der AFS-Feinheitsnummer widerspiegelt, die als ein Maß für die relative Kornfeinheit interpretiert werden kann und in dem vorliegenden Fall über die Versuchsreihen hinweg immer weiter absinkt. Begründet werden kann dieser Effekt durch die Zugabe des Binders, der wiederum durch die Regeneration nicht vollständig abgeschlagen wird. Der Gleichmäßigkeitsgrad gibt den prozentualen Massenanteil zwischen 25 % und 75 % der gesamten Sandmenge an und ist ein Maß für die Breite der Kornverteilung [10]. Auch hier ist die Senkung des Massenanteils in den vorgegeben Grenzen zu beobachten, was auf eine generelle Veränderung der Partikel deutet. Die Werte der mittleren kalkulierten Korngröße weisen darauf hin, dass sich ein Großteil der Masse weiterhin im sechsten Sieb sammelt, nämlich bei einer Korngröße von 0,19 bis 0,21 mm.

Für weitere Informationen hinsichtlich der granulometrischen Kornveränderung werden in Abbildung 4.10 die Histogramme für die drei Regenerationszeiten untereinander dargestellt. Von oben nach unten zeigt die Darstellung die Analyseergebnisse der (i) 5-minütigen, der (ii) 10-minütigen und der (iii) 15-minütigen Regenerationen. Aus der untersuchten Korngrößenverteilung des Neusandes (Abbildung 4.9) ergab sich, dass 98,37 % der Sandmenge sich auf die mittleren Siebe 4, 5, 6, 7 und 8 verteilt, weshalb in Abbildung

Tabelle 4.4: Kennwerte der Siebanalyse

Reg.-Zeit	Umlauf	theoertische Spez. Oberfläche [cm <sup>2</sup> /g]	AFS	Gleichmäßigkeitsgrad [%]	Mittlere (calc) Korngröße [mm]
Neusand		121,61	68,36	82,31	0,20
5 Min	1	125,36	70,43	80,18	0,19
	2	117,88	66,58	77,17	0,21
	3	114,32	64,77	77,44	0,21
10 Min	1	125,04	70,25	81,21	0,19
	2	122,17	68,78	79,74	0,20
	3	119,08	67,19	78,91	0,20
	4	117,71	66,48	78,51	0,20
	5	114,47	64,82	76,96	0,21
	6	115,04	65,11	77,19	0,21
	7	114,48	64,83	77,61	0,21
	8	112,83	63,99	76,92	0,21
	9	111,85	63,49	76,57	0,21
	10	112,59	63,85	77,40	0,21
15 Min	1	123,06	69,23	80,46	0,20
	2	119,43	67,38	78,90	0,20
	3	119,75	67,50	80,08	0,20
	4	119,29	67,28	79,58	0,20
	5	117,32	66,25	78,95	0,20
	6	116,94	66,08	77,89	0,21
	7	118,32	66,78	79,49	0,20
	8	112,08	63,59	76,15	0,21
	9	113,68	64,40	77,14	0,21
	10	114,38	64,74	78,82	0,21

4.10 auf die Darstellung der verbleibenden Kornklassen (1, 2, 3 sowie 9, 10 und 11) verzichtet wird. Alle drei Teilgrafiken sind identisch aufgebaut: Auf der x-Achse, welche jeweils identisch skaliert ist, sind die Kornklassen mit den entsprechenden Maschenweiten der Siebe aufgetragen. Der Massenanteil pro Sieb bzw. pro Kornklasse wird prozentual auf der y-Achse abgebildet. Die farbliche Zuordnung der Umläufe ist ebenfalls für alle drei Histogramme identisch, daher wurde in Teildarstellung (iii) auf die erneute Wiedergabe der Legende verzichtet. Die Reihenfolge der dargestellten Balken pro Kornklasse folgt dem durchgeführten Umlaufsystem, das bedeutet, dass pro Kornklasse links als erstes in rot immer Neusand abgebildet wird und sich nach rechts folgend Umlauf 1, Umlauf 2 usw. einreihen. Über dem jeweiligen Balken einer Kornklasse gibt der zugehörig Prozentwert die durchschnittliche absolute Abweichung eines jeden Umlaufes zu Neusand an. Da die Versuchsreihe der Regeneration mit einer Prozesszeit von 5 Minuten nach dem dritten Umlauf abgebrochen wurde, sind nur die aufgenommenen Werte bis Umlauf 3 abbildbar.

Die Histogramme der Umläufe innerhalb der drei Regenerationszeiten bestätigen die zuvor erkannte

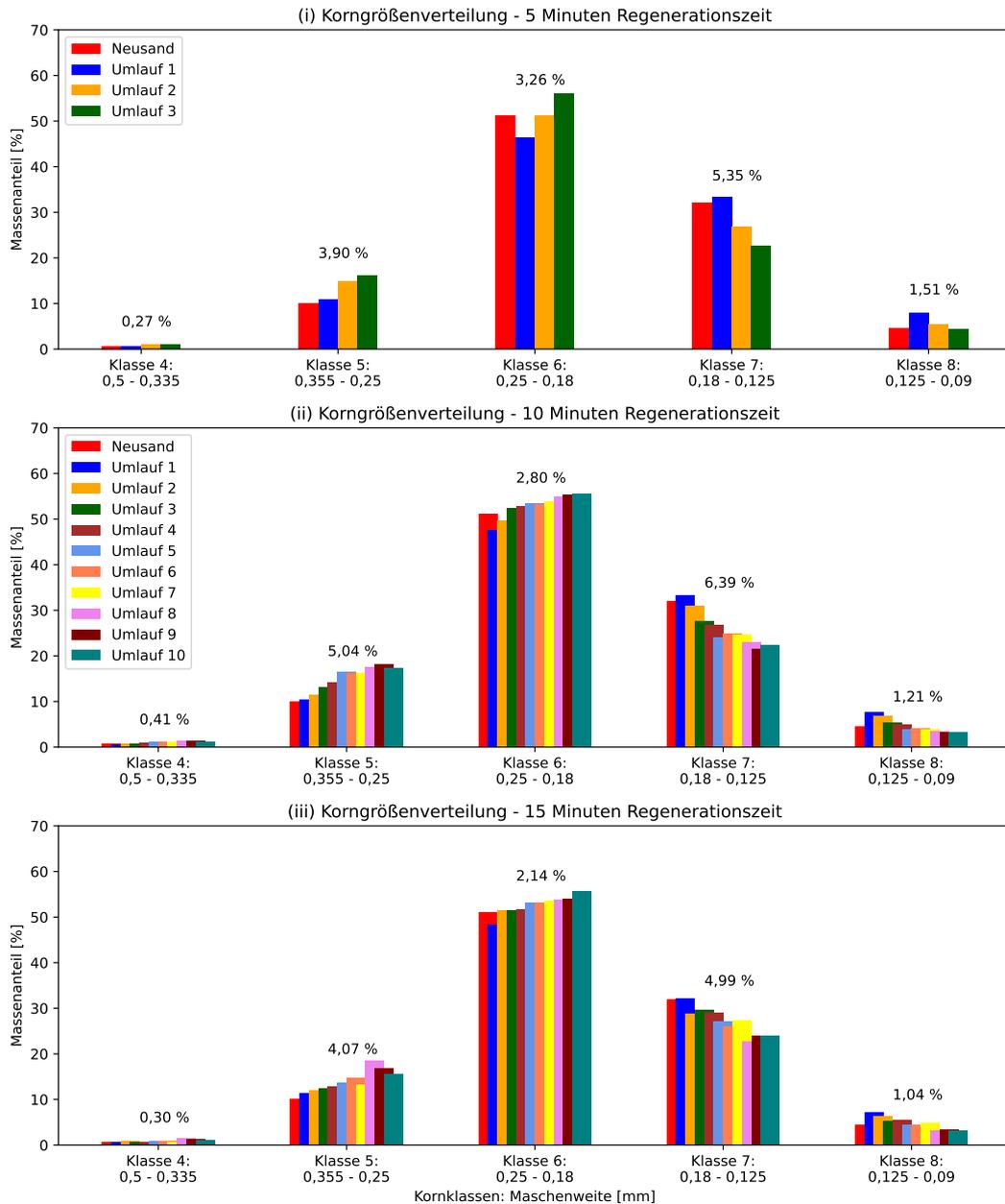


Abbildung 4.10: Korngrößenverteilung der drei Versuchsreihen

Veränderung der Partikelgröße. Für alle drei Darstellungen ergeben sich die geringsten Veränderungen im Vergleich zu Neusand in der vierten (0,5–0,335 mm) und achten (0,125–0,09 mm) Kornklasse. Die auftretenden Veränderungen liegen für alle drei Regenerationszeiten in Klasse 4 unter 0,5 % und in Klasse 8 unter 1,6 %. Das bedeutet, dass sich die gröberen sowie die feineren Körner der gesamten Sandmenge weniger stark verändern. Allerdings sind diese Randkornklassen für eine gesamtheitliche Interpretation vernachlässigbar, da der Anteil im Bezug auf die Gesamtmenge sehr klein und damit nicht signifikant ist. Die relevanten Veränderungen treten in den mittleren drei Kornklassen auf, wobei der prozentual größte Unterschied aller drei Regenerationszeiten im Vergleich zu Neusand in Kornklasse 7 (0,18–0,125 mm) auftritt, nämlich mit einer durchschnittlichen absoluten Abweichung von 6,93 % bei einer Regenerationszeit

von 10 Minuten. Die geringste Veränderung innerhalb der mittleren drei Kornklassen findet in der sechsten Klasse (0,25-0,18 mm) statt. Obwohl sich die kalkulierte mittlere Korngröße (s. Tabelle 4.4) kaum verändert, bzw. lediglich zwischen 0,19 und 0,21 mm schwankt, liegen die Abweichungen zu Neusand pro Umlauf bei einer Regenerationszeit von fünf Minuten (i) bei 3,26 %, bei 10 Minuten (ii) bei 2,80 % und bei 15 Minuten (iii) bei 2,14 %. Für die sechste Kornklasse kann demnach bei steigender Regenerationszeit eine sinkende Veränderung gegenüber Neusand festgestellt werden. Die verbleibende fünfte Kornklasse (0,335–0,25 mm) weist gemeinsam mit der siebten Kornklasse die größten Abweichungen zu Neusand auf, was bedeutet, dass der Regenerationsprozess, unabhängig von der Regenerationszeit, keine gleichmäßige Auswirkung auf die Veränderung der Partikelgröße hat.

Die Ergebnisse der Siebanalysen zeigen für alle drei Regenerationszeiten ein veränderndes Verhalten der Partikelgröße, wie es auch in der Literatur für mechanische Regenerationsprozesse beschrieben wird (vgl. Kapitel 2.1.2). Andere Forschungsarbeiten zur Regeneration von Gießereisanden weisen auf eine generelle Abnahme der Partikelgröße hin, wie zum Beispiel bei Sun, Jing et al. (2011) [123] und Danko, Danko (2005) [21]. Bezogen auf die vorliegenden Kennwerte der Siebanalysen wurde festgestellt, dass die Regeneration, unabhängig von der Regenerationszeit, die größte Veränderung des Korngefüges in der fünften und siebten Kornklasse bewirkt. Aus diesen Erkenntnissen folgt, dass die Veränderung der Sandkörner nicht gleichmäßig auftritt. Zur Erzielung der ursprünglichen Korngrößenverteilung des Neusandes könnten die entsprechend notwendigen Mengen von klassiertem Sand zum Regenerat hinzugefügt werden, um die produzierten Veränderungen auszugleichen. Wichtig anzumerken ist, dass eine allgemeingültige Interpretation der vorliegenden Ergebnisse nicht möglich ist, da die Ansprüche der Gießereien an die Formstoffqualität variieren. Aspekte wie die Korngrößenverteilung, die Geometrie oder die Oberfläche einzelner Körner können nicht allgemeingültig in die Definition von Sandqualität einbezogen werden [6]. Erwiesen ist allerdings die Auswirkung der veränderten Korngrößenverteilung auf die Festigkeit, die sich auch in den folgenden Ergebnissen widerspiegelt, denn bei veränderter Partikelgröße verändert sich auch die Festigkeit [10], [121, S. 11].

### Auswertung der Biegebiegefestigkeit

Neben der Siebanalyse wurden sowohl mit Neusand, als auch mit dem Regenerat eines jeden Umlaufes pro Zyklus, insgesamt 84 Biegebiegefestigkeitsersteller erstellt und anteilig zu sechs Messzeiten vermessen. Pro Messzeit entstanden dementsprechend 14 bzw. 15 Riegel, deren Festigkeiteswerte zu einem aussagekräftigen Messwert gemittelt wurden, um Messchwankungen entgegenzuwirken. Die Messungen für alle drei Regenerationszeiten, (i) 5 Minuten, (ii) 10 Minuten und (iii) 15 Minuten, sind untereinander mit den aufeinander folgenden Umläufen in Abbildung 4.11 dargestellt. Die x-Achse bildet in logarithmischer Darstellung die Aushärtezeit ab; so können der große Sprung zwischen den Messzeiten 270 und 1470 Minuten sowie die Festigkeiteswerte, insbesondere in den kleineren Bereichen (30, 60, 90 Minuten), deutlich besser abgebildet und analysiert werden. Die sechs Messzeiten sind durch die vertikal gestrichelten Linien in jedem Diagramm gekennzeichnet. Die maximal erreichte Festigkeit vor dem Bruch des Riegels wird auf der y-Achse in  $\frac{N}{cm^2}$  angegeben, wobei das obere Teildiagramm bei einer Skalierung von  $350 \frac{N}{cm^2}$  endet und die beiden weiteren bis  $400 \frac{N}{cm^2}$  skaliert sind, da insgesamt höhere Festigkeiteswerte erreicht wurden. Die farbliche Zuordnung ist für jedes der drei Teildiagramme identisch, sodass die Legende nur einmalig abgebildet wird, aber für alle Teilgrafiken gültig ist. Neusand ist, wie in den Abbildungen zuvor, in rot und mit einer dickeren Linie dargestellt.

Die Entscheidung für den Abbruch der Versuchsreihe bei 5-minütiger Regenerationszeit wird durch die

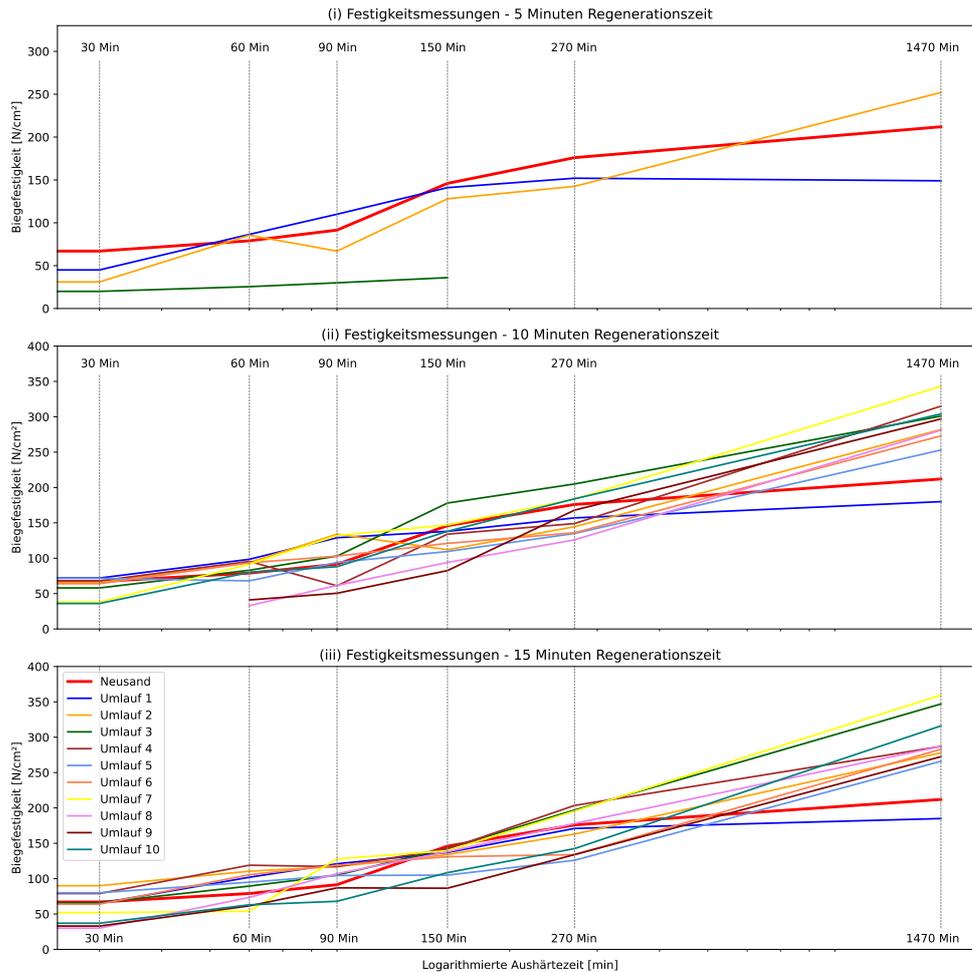


Abbildung 4.11: Festigkeitsanalyse der einzelnen Umläufe pro Regenerationszeit

Festigkeitsanalyse schnell verdeutlicht. Umlauf 3 wird in Abbildung 4.11 in dunkelgrün dargestellt und weist in Teildiagramm (i) über die gesamte Aushärtezeit bereits deutlich geringere Festigkeitswerte auf. Zudem ist nur eine sehr geringe Steigerung der Festigkeit über die Zeit erkennbar; anders als es bei den vorherigen Umläufen 1 und 2 sowie bei Neusand abzulesen ist. Aufgrund der geringen Festigkeit traten bereits bei der Entnahme der Riegel aus den Formkästen Brüche auf, sodass ein Großteil nicht vermessen werden konnte. Für den dritten Umlauf konnten dadurch insgesamt nur 36 von 84 Riegel vermessen werden, die sich auf die Messzeiten von 30 bis 150 Minuten aufteilen ließen, wodurch für die späten Messzeiten nach 270 und 1470 Minuten keine Riegel verblieben. Die Versuchsreihe wurde auf Grundlage der erläuterten Ergebnisse nach dem dritten Umlauf abgebrochen.

Für die Regenerationszeiten von 10 und 15 Minuten konnten die geplanten 10 Umläufe durchgeführt werden, wobei sich der Festigkeitsverlust über die Anzahl der Umläufe, zumindest für die kürzeste Aushärtezeit von 30 Minuten, auch bei der 10 minütigen Prozesszeit bemerkbar macht. So konnten für das Regenerat der Umläufe 8 und 9 durch zu geringe Festigkeiten der Riegel keine Messwerte zur Messzeit 30 min erfasst werden. Die mittlere Darstellung (ii) in Abbildung 4.11 weist dementsprechend eine Lücke bei 30 Minuten für die Umläufe 8 und 9 auf. Erst ab 60 Minuten Aushärtezeit nimmt die Festigkeit der Regenerate der Umläufe 8 und 9 wieder zu, wenngleich diese bis zur maximal erreichten Festigkeit nach 24 Stunden

unterhalb der erreichten Werte von Neusand liegen. In der untersten Abbildung (iii) der dreigeteilten Darstellung 4.11 sind die Festigkeitsverläufe der Versuchsreihe bei 15 minütiger Regenerationszeit abgebildet. Hier konnten alle Messwerte erfasst und abgebildet werden, wengleich zu den kurzen Aushärtezeiten ab Umlauf 7 geringere Festigkeiten gegenüber Neusand gemessen wurden. Sowohl bei 10 als auch bei 15 Minuten Regenerationszeit erreichen alle Regenerate, außer das des ersten Umlaufes, höhere maximale Festigkeiten als Neusand. Generell lässt sich sowohl graphisch als auch rechnerisch kein erklärbarer Zusammenhang in der Entwicklung der Festigkeiten in Abhängigkeiten des Umlaufes feststellen.

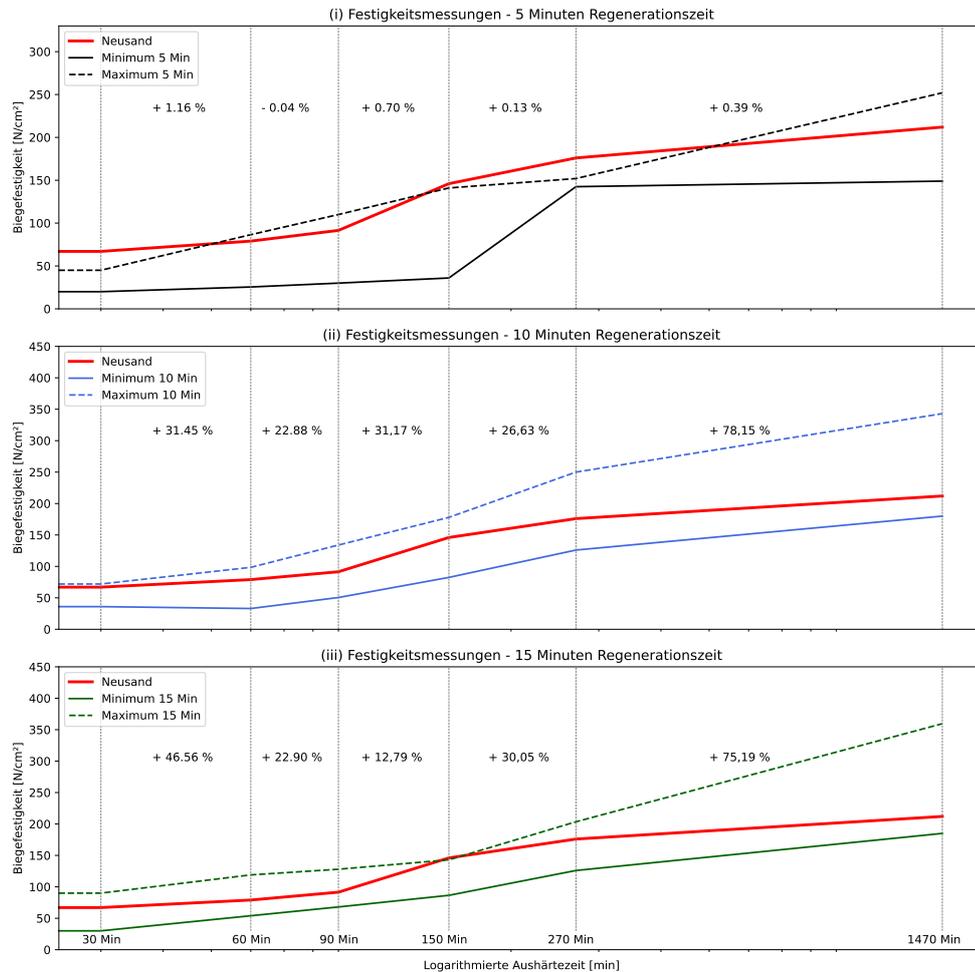


Abbildung 4.12: Festigkeitsanalyse pro Regenerationszeit im Vergleich zu Neusand

Zur weiteren Beurteilung der Ergebnisse wird Abbildung 4.12 herangezogen. In gleicher Reihenfolge sind hier von oben nach unten die Ergebnisse der Regenerationszeiten von (i) 5, (ii) 10 und (iii) 15 Minuten abgebildet, wobei jeweils nur der Bereich zwischen den maximal sowie den minimal erreichten Festigkeitswerten im Vergleich zu Neusand dargestellt wird. Zusätzlich wurden die Messwerte aller Regenerate pro Umlauf und Messzeit gemittelt und die prozentuale Abweichung zur gesamten gemittelten Festigkeit der nachfolgenden Messzeit berechnet und in den Diagrammen angegeben. Das bedeutet, dass für das Regenerat mit fünf minütiger Regenerationszeit die mittlere Festigkeit der drei einfließenden Umläufe von einer Aushärtezeit von 30 auf 60 Minuten um 1,16% steigt. Die Abweichungen fallen hier für den gesamten Verlauf so gering aus, da die geringen Festigkeitswerte des Regenerates des dritten Umlaufes die gesamte

Berechnung beeinflussen, da sie zu  $\frac{1}{3}$  in das Ergebnis einfließen. Dies ist auch durch den Verlauf bzw. den grau hinterlegten Bereich dieser Versuchsreihe erkennbar, der keiner gleichmäßigen Form folgt, wie bei den unteren beiden Darstellungen zu erkennen ist.

Für die Regenerate, die jeweils 10 bzw. 15 Minuten regeneriert wurden, ergeben sich im Vergleich Verlaufsgebiete, die die Messwerte des Neusandes umhüllen. Die Unterschiede zu Grafik (i) der Abbildung 4.12 werden auch durch die berechneten prozentualen Veränderungen deutlich. Insbesondere in der ersten Stunde unterscheiden sich die Festigkeitssteigerung um ca. 15%. Gleiches gilt für den Zeitsprung von 90 auf 150 Minuten; hier sind es sogar über 18% Abweichung. Den größten Sprung weisen beide Versuchsreihen beim Erreichen der jeweiligen Maximalfestigkeit bei einer Messzeit nach 24 Stunden auf. Die durchschnittliche Steigerung der Festigkeit liegt hier bei 78,15% bzw. 75,19%. Dieser Anstieg lässt sich durch die veränderte Partikelgröße während des Regenerationsprozesses bei gleichbleibendem Bindersystem erklären. Durch das mechanische Regenerieren erreichen die Körner generell eine höhere Rundheit und die beschriebene Partikelgrößenveränderung tritt ein. Für eine gleichbleibende Festigkeit wäre also ein angepasstes Bindersystem erforderlich. Aufgrund der gleichbleibenden Binder- bzw. Härtermenge ist die erläuterte Konsequenz der erzielten höheren Festigkeitswerte gegenüber Neusand erkenn- und erklärbar.

### 4.4.3 Konklusion

Anhand der Regenerationsversuche wurde beurteilt, inwiefern das Ziel, den gebundenen Sand wieder in Material umzuwandeln, dass den Eigenschaften von Neusand ähnlich ist, erreicht werden kann. Zunächst wurde festgestellt, dass eine festgelegte Regenerationszeit von 5 Minuten für das untersuchte Formstoffsystem aus Quarzsand F34 und einem anorganischen Wasserglas-Bindersystem nicht ausreicht, um das erzeugbare Regenerat langfristig in einem Umlaufsystem zu nutzen. Diese Erkenntnis ist für den weiteren Verlauf dieser Arbeit, insbesondere für das Labeling der aufgenommenen Akustikdaten, sehr wichtig. Damit konnte eine Prozesslaufzeit bestimmt werden, in der der regenerierte binderbelastete Sand nicht wieder Neusand bzw. einer neusandähnlichen Struktur entsprechen konnte, nämlich bis zu einem Prozesszeitpunkt von mindestens fünf Minuten.

Aus der Siebanalyse konnte die Erkenntnis gezogen werden, dass der Regenerationsprozess, unabhängig von der Regenerationszeit, keinen gleichmäßigen Einfluss auf die Veränderung der Partikelgröße hat. Insbesondere die fünfte (0,335 - 0,25) und siebte (0,18 - 0,125) Kornklasse erfuhren im Vergleich zu Neusand die durchschnittlich größte prozentuale Veränderung. Bei einem gleichbleibenden Bindersystem mit immer identischen Mengenzugaben bewirkt die Partikelgrößenveränderung den generellen Anstieg der Festigkeit gegenüber Neusand. Dieser Effekt bestätigt sich durch die Messergebnisse der Biegeriegel festigkeiten: Über 10 Umläufe erreichten die Biegeriegel tendenziell höhere Maximalfestigkeiten als Neusand, schwanken jedoch bei den frühen Messzeiten ungleichmäßig um die Festigkeitswerte von Neusand.

Ein weiteres Qualitätsmerkmal von Quarzsand ist die tatsächliche Struktur der Quarzkörner. Die morphologischen Merkmale von Sanden sind die Kornoberfläche und die Kornform, die wiederum durch den Kantigkeitsgrad bestimmt wird. Solche Analysen werden mit den erzeugbaren Bildern eines Rasterelektronenmikroskops (REM) durchgeführt, um die Oberflächenstruktur und das Bindungsverhalten von Sanden zu untersuchen. Abbildungen 4.13 und 4.14 zeigen beispielhaft den verwendeten Quarzsand F34 mit einer Maßstab von 1 mm in dem linken Bild (4.13) und in dem rechten (4.14) von 250  $\mu\text{m}$ . Neusand ist generell für eine eher eckige, kantige Struktur mit unebener bzw. rissiger oder zerklüfteter Oberfläche bekannt. Eine Regeneration führt dazu, dass das Korn durch den Prozess eine kantengerundete Struktur annimmt. Die hohe Auflösung, die Vergrößerung der Sandkörner und die Möglichkeit einer zusätzlichen

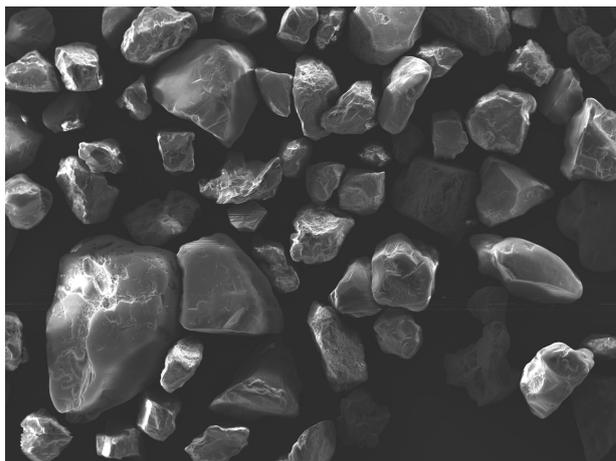


Abbildung 4.13: Quarzsand F34 - Skala: 1 mm

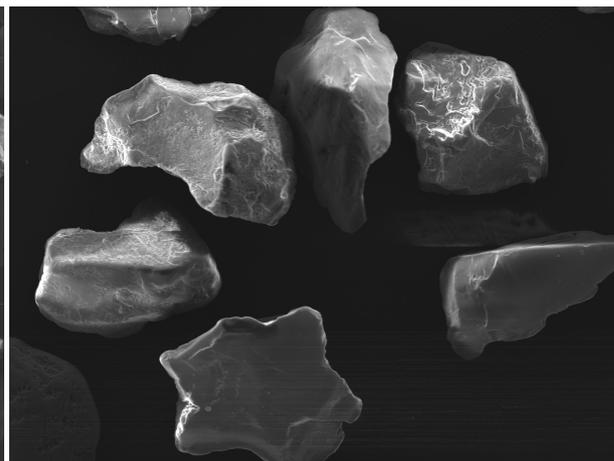


Abbildung 4.14: Quarzsand F34 - Skala: 250 µm

chemischen Analyse sind entscheidende Vorteile des REMs, welche in Bezug auf gezielte Forschungen zur Kornveränderungen unverzichtbar sind. REM-Analysen sind allerdings sehr zeitaufwendig und bilden nur einen Bruchteil eines zu untersuchenden Schüttguts ab; die dargestellten Sandkörner auf den gezeigten Bildern (Abbildungen 4.13 und 4.14) sind sogar zählbar. Es kann zwar ein Eindruck der Kornstruktur abgebildet und gewonnen werden, aber für das hier verwendete Chargengewicht von 30 kg und die Anzahl der Versuchsreihen würde sich ein erheblicher Mehraufwand für eine signifikante Untersuchung der morphologischen Veränderung ergeben. Die REM-Analysen sowie die klassischen und gängigen Formstoffqualitätskriterien gewährleisten mit einem gewissen Zeitaufwand allerdings die tatsächliche Bestimmung des Ist-Zustandes des Regenerates nach einem Regenerationsprozess. Wie Bifano, Fischerauer et al. (2020) jedoch bereits in ihren Untersuchungen bemängeln, ist mit diesen Kriterien keine Inline-Prozessüberwachung der Sandqualität während der Sandregenerierung möglich [6]. Aus diesem Grund wird auf den zusätzlichen Aufwand umfangreicher REM-Analysen sowie anderer Untersuchungen verzichtet, um im Folgenden den Fokus auf das Forschungsziel einer kontinuierlichen akustischen Beurteilung des Regenerationsprozesses, also einer Inline-Überwachung, zu legen.

# Kapitel 5

## Machbarkeitsstudie

Die im vierten Kapitel vorgestellte Problemstellung führte zu drei Leitfragen des Lösungskonzeptes. Die erste Frage, die sich auf das Qualitätskriterium des Regenerates bezieht, wurde bereits durch die Auswertung der Ergebnisse der in Kapitel 4.4.2 diskutierten Versuchsreihe beantwortet. Die zwei verbleibenden Leitfragen, inwieweit sich das festgelegte Qualitätskriterium der Neusandähnlichkeit akustisch erkennen und inwieweit sich der Status des Aufbereitungsprozesses akustisch erfassen lässt, werden im Folgenden anhand einer Machbarkeitsstudie beantwortet. Dabei wurde geprüft, ob mit der akustischen Prozesserfassung eine kontinuierliche Inline-Prozessüberwachung möglich ist. Drei Aspekte, die im Hinblick einer kontinuierlichen und ganzheitlichen Prozessüberwachung relevant sind, wurden untersucht:

### 1. Neusand vs. Regenerat

In Kapitel 4.3 stellt Abbildung 4.5 das Konzept für die akustische Erkennung des Regenerationsfortschrittes dar. Zwei wesentliche Fragen traten dabei auf, die mit dieser ersten Analyse untersucht wurden: Ist ein DL-Modell in der Lage, den akustischen Unterschied zwischen Neusand und unzureichend regeneriertem (Alt)sand zu erlernen und zu generalisieren? Wenn ja, inwieweit ist es sinnvoll und möglich, einen Status des Regenerationsprozesses anhand der akustischen Aufnahmen prognostizieren zu können? Dieser Status bezieht sich auf die Formstoffqualität, die von unzureichender Sandqualität bis zu hinreichend neusandähnlicher Sandqualität reicht.

### 2. Störgeräusche

Allgemein laufen Prozesse, insbesondere im realen Betrieb, nicht störungsfrei. Besonders im Hinblick auf Akustikaufnahmen sollten Umgebungs- und Hintergrundgeräusche von anderen Fertigungsprozessen, Anlagen, Mitarbeitern, benachbarten produzierenden Unternehmen etc. während der Audiodatenauswertung berücksichtigt werden. Wie auch in [54] wurde der hier betrachtete Prozess in einer Demo-Umgebung, in diesem Fall in einer universitären Forschungshalle, und nicht im realen Betrieb untersucht. Henze, Gotshti et al. (2019) lösten für ihre Forschung künstlich erzeugten Fehlerzustände aus, um diese prozesstechnisch erfassen zu können [54]. Für die geplanten Analysen dieser vorliegenden Forschung wurde ein Teil der aufgetretenen Störgeräusche genutzt, um zu prüfen, ob ein DL-Modell diese im laufenden Prozess von störfreien Aufnahmen unterscheiden kann, also Störungen als solche erkennt.

### 3. Verschiedene Sandsorten

Für die Analysen zur Unterscheidung von Regenerat und Neusand wurde nur eine Sandsorte, nämlich Quarzsand F34, verwendet. In der Praxis ist es üblich, dass unterschiedliche Sand-Binder-Systeme zur Herstellung von Sandformen bzw. -kernen verwendet werden. Zur Erkennung und Charakterisierung unterschiedlicher Sandsorten haben Bifano, Fischerauer et al. (2020) das Verfahren der Impedanzspektroskopie genutzt, um die Sandsorten anhand ihrer unterschiedlichen Impedanzkurven zu charakterisieren [6]. Angelehnt an dieses Vorgehen wird untersucht, ob es mit Hilfe eines CNNs möglich ist, unterschiedliche Sande anhand der aufgenommenen Prozessakustik des Regeneriervorgangs zu identifizieren.

Mit den Untersuchungen dieser drei Teilaspekte soll abschließend eine Aussage hinsichtlich der anfangs aufgestellten Hypothese getroffen werden. Bevor jedoch auf die drei Analysen eingegangen und das Ergebnis diskutiert wird, werden die vorliegende Datenbasis sowie das allgemeine Vorgehen zur Erstellung der Spektrogramme, der erweiterten Datenmenge durch das Augmentieren und der CNN erläutert.

## 5.1 Allgemeines Vorgehen

### 5.1.1 Generierte Datenbasis

Jede durchgeführte Regeneration wurde akustisch erfasst. Die Rohaudiodaten wurden zuerst von den Ein- und Auslaufgeräuschen bereinigt, sodass nur noch die reinen Prozessaufnahmen verblieben, die anschließend, wie in Kapitel 4.3.2 beschrieben, in 3s-Audiodateien geschnitten wurden, wobei unvollständige Restaufnahmen vernachlässigt wurden.

Tabelle 5.1: Vorliegende Datenbasis für Neusand

Sandsorte	Aufnahmen	Rohdatei in [hh:mm:ss]	Reine Prozesslänge in [hh:mm:ss]	Anzahl 3s-Dateien
H33	V1 H33 - 60min	01:01:23630	00:59:50	1196
	V2 H33 - 60min	01:00:47664	01:00:00	1200
Sand 01	V1 S01 - 60min	01:04:09958	01:00:10	1203
	V2 S01 - 60min	01:01:36722	01:00:25	1208
Sand 02	V1 S02 - 60min	01:01:40628	01:00:00	1200
	V2 S02 - 60min	01:01:48627	01:00:00	1200
Sand 03	V1 S03 - 60min	01:03:10432	01:01:15	1225
	V2 S03 - 60min	01:00:36228	00:59:55	1198
H33 T	V1 H33 T - 60 min	01:03:34865	01:00:05	1201
	V2 H33 T - 60 min	01:01:36503	01:00:53	1217
F34	V1 F34 - 60 min	01:04:54292	01:00:10	1203
	V2 F34 - 40min	00:43:13031	00:40:10	803

Aus den Tabellen 5.1 und 5.2 bis 5.5 gehen die in ihrer Gesamtheit vorliegenden akustischen Datenmengen für die Kategorien *Neusand* resp. *Nicht-Neusand* hervor. Die *Nicht-Neusand*-Daten werden

entsprechend den Regenerationszeiten in drei einzelnen Tabellen dargestellt.

Tabelle 5.1 beinhaltet die akustische Datenbasis der verfügbaren *Neusand*-Aufnahmen, die mit insgesamt 6 verschiedenen Neusanden aufgenommen wurden. Quarzsand H33 stand aus zwei Bezugsquellen zur Verfügung und wird daher zweimal in Tabelle 5.1 aufgeführt. Unter der Annahme, dass sich die Qualität des Regenerates bei einer Regenerationszeit von z.B. einmalig 20 Minuten oder zweimalig 10 Minuten nicht unterscheidet, wurden die *Neusand*-Regenerationen, bis auf eine Ausnahme, immer ca. 60 Minuten lang durchgeführt, um den Prozess von mehreren Regenerationsvorgängen zu simulieren. Mögliche auftretende Quarzkornveränderungen während des Prozesses wurden mit diesen Langzeitaufnahmen implizit mit aufgenommen und demnach berücksichtigt. Für jede Sandsorte wurden jeweils zwei Langzeitaufnahmen generiert. Die Tabelle gibt damit einen Gesamtüberblick über die jeweiligen Rohaudioaufnahmen der Regenerationsvorgänge in der dritten Spalte. Die reine Prozesslänge beschreibt die Länge der Aufnahme nach dem Zuschnitt, also dem Entfernen der Ein- und Auslaufgeräusche des Prozesses. In der letzten Spalte wird die Anzahl der generierten 3-Sekunden-Dateien angegeben, die sich aus dem reinen Prozess zuschneiden ließen. Auftretende Störgeräusche sind in den hier aufgelisteten Daten noch vorhanden.

Die zweite Kategorie vorliegender Audiodaten betrifft die Datenbasis der Prozessakustik von regenerierten binderbelasteten Sanden, also den Regeneraten. Die dazu durchgeführte Versuchsreihe wurde, wie bereits in Kapitel 4.4.2 beschrieben, mit der Sandsorte Quarzsand F34 durchgeführt. Diese Versuche lieferten, neben den Erkenntnissen zur Wirkleistung der Anlage, dementsprechend auch akustische Prozessdaten der Regenerationen eins belasteten Sandsystems. Neben diesem Umlaufsystem (Tabellen 5.3, 5.4 und 5.5) konnten mit den Sandsorten F34 und H33 noch drei resp. zwei Langzeitaufnahmen von jeweils etwa 60 Minuten erzeugt werden, die in Tabelle 5.2 aufgeführt sind. In allen vier Tabellen wird wieder die Länge der Rohdateien, die reinen Prozesslängen sowie die Anzahl sich daraus ergebender 3s-Dateien aufgelistet.

Tabelle 5.2: Vorliegende Datenbasis binderbelasteter Sande

Sandsorte	Aufnahmen	Rohdatei in [hh:mm:ss]	Reine Prozesslänge in [hh:mm:ss]	Anzahl 3s-Dateien
H33 T	H33 1x gebunden V1	01:01:19099	01:00:10	1203
	H33 1x gebunden V2	01:01:47252	01:00:36	1212
F34	F34 1x gebunden 60 min V1	01:03:38649	01:00:53	1217
	F34 1x gebunden 60 min V2	01:02:04126	01:01:14	1224
	F34 Umlaufsystem 60 min	01:02:28092	01:01:38	1232

Tabelle 5.3: Vorliegende Datenbasis - Regenerationszeit 5 Minuten

Sandsorte	Aufnahmen	Rohdatei in [hh:mm:ss]	Reine Prozesslänge in [hh:mm:ss]	Anzahl 3s-Dateien
F34	5min U1	00:04:40	00:04:40	88
	5min U2	00:05:06	00:05:06	102
	5min U3	00:04:25	00:04:25	88
	5min U4	00:05:51	00:05:51	117
	5min UR	00:13:25	00:13:25	268

Tabelle 5.4: Vorliegende Datenbasis - Regenerationszeit 10 Minuten

Sandsorte	Aufnahmen	Rohdatei in [hh:mm:ss]	Reine Prozesslänge in [hh:mm:ss]	Anzahl 3s-Dateien
F34	10min U1	00:08:21	00:08:21	161
	10min U2	00:09:30	00:09:30	190
	10min U3	00:10:00	00:10:00	200
	10min U4	00:09:11	00:09:11	183
	10min U5	00:11:13	00:11:13	224
	10min U6	00:08:10	00:08:10	163
	10min U7	00:10:00	00:10:00	200
	10min U8	00:08:22	00:08:22	167
	10min U9	00:08:10	00:08:10	163
	10min U10	00:08:49	00:08:49	176

Tabelle 5.5: Vorliegende Datenbasis - Regenerationszeit 15 Minuten

Sandsorte	Aufnahmen	Rohdatei in [hh:mm:ss]	Reine Prozesslänge in [hh:mm:ss]	Anzahl 3s-Dateien
F34	15min U1	00:15:15	00:15:15	305
	15min U2	00:22:30	00:15:00	300
	15min U3	00:14:25	00:14:25	288
	15min U5	00:16:50	00:15:00	300
	15min U6	00:13:10	00:13:10	263
	15min U7	00:15:10	00:15:10	303
	15min U8	00:14:20	00:14:20	286
	15min U9	00:08:10	00:08:10	163
	15min U10	00:11:15	00:11:15	225

### 5.1.2 Erstellung der Spektrogramme

Wie auch für die bereits vorangegangene Berechnung der Kreuzkorrelations- bzw. Korrelationskoeffizienten für die Analyse der Reproduzierbarkeit der Audiodaten, wird für die Erstellung der Spektrogramme die Programmiersprache Python<sup>22</sup> benutzt. Die Skripte für die Berechnungen wurden über Jupyter Notebooks ausgeführt. Jupyter<sup>23</sup> hat sich mit diesen Notebooks, die sowohl Texte beinhalten als auch Programme ausführen können, zu einem Universalwerkzeug entwickelt, welches neben Python auch die Ausführung von über 40 Programmiersprachen innerhalb einer einheitlichen Web-Oberfläche ermöglicht. [133, S. 63] Zum Laden der Audiodaten, zur Durchführung der STFT und zur Erstellung der resultierenden Spektrogramme wird die *librosa*-Bibliothek<sup>24</sup> verwendet, die wiederum auf mathematische Funktionen von *numpy*<sup>25</sup> zurückgreift.

Die Erstellung der Spektrogramme der 16-Bit Stereo Audiodaten mit einer Abtastrate von 44100 Hz

<sup>22</sup><https://www.python.org/> (besucht am 14.07.2022)

<sup>23</sup><https://jupyter.org/> (besucht am 30.06.2022)

<sup>24</sup><https://librosa.org/> (besucht am 30.06.2022)

<sup>25</sup><https://numpy.org/> (besucht am 30.06.2022)

erfolgte für jede Analyse identisch. Für die Ausführung der STFT gibt es einige variable Werte, die für die Berechnungen relevant sind und im Rahmen dieser Arbeit immer gleichbleibend in die Transformationen einfließen. Die STFT transformiert das ursprüngliche Signal in den Zeit-Frequenz-Bereich, indem die DFT auf die jeweiligen Fenster angewendet wird. Die Länge eines solchen gefensterten Signals wird in diesem Fall mit der *Frame Length* und der Überlappung<sup>26</sup> bestimmt. Mit welcher Fensterfunktion das Ausgangssignal transformiert werden soll, ist ebenfalls variabel und muss somit auch individuell angegeben werden. Weist dieses Fenster eine geringere Länge (*Win Length*) bezüglich der angegebenen *Frame Length* auf, wird die *Win Length* mit Nullwerten so weit aufgefüllt (*Zeropadding*) bis beide Fensterlängen zueinander passen.

Für die generierten Spektrogramme wurde, wie in den Forschungen von Sprengel, Jaggi et al. (2016) und Schlüter, Grill (2015), das Hanning-Fenster verwendet [117, 112]. Die Fenstergröße des Hanning-Fensters entspricht einer *Frame Length* von 2048 und einer *Hop Length* von 1024. Die x-Achse der Spektrogramme gibt den zeitlichen Verlauf des Signals in Sekunden an und die zugehörige Frequenz wird logarithmisch auf der y-Achse abgebildet. Die Farbwerte entsprechen der Intensitäten der jeweiligen Frequenzen und werden mit der integrierten Farbskalierung *coolwarm* mit Matplotlib<sup>27</sup> erzeugt. Abbildung 5.1 zeigt stellvertretend ein Beispiel für alle erzeugten Spektrogramme, wobei die Bilder dem CNN in reiner Form, also ohne Achsenbeschriftungen, ohne die Farbskala und ohne Rahmen als Input übergeben werden.

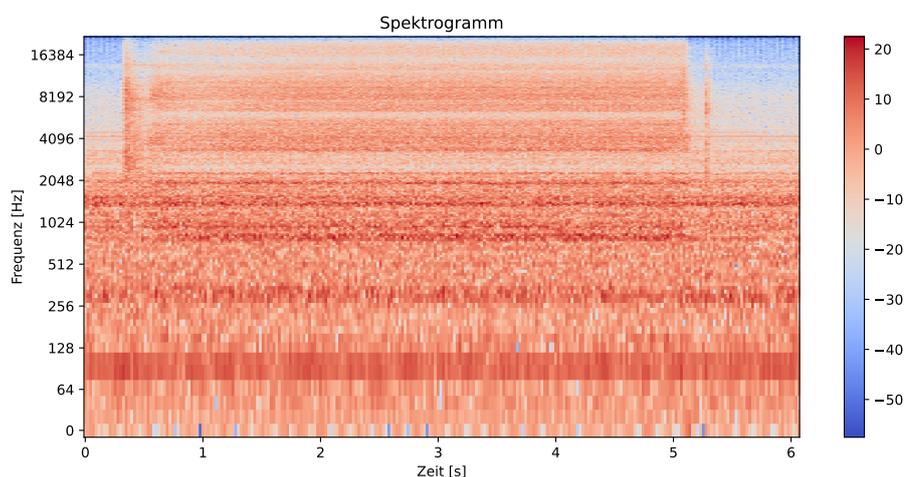


Abbildung 5.1: Beispiel für die erstellten Spektrogramme

Häufig werden auch grau-skalierte Spektrogramme als Eingabe für das CNN verwendet, wie in den Untersuchungen von Bian, Wang et al. (2019) [5]. Ein Grund dafür kann die Dimensionsreduktion der vorliegenden Audiodatenbasis sein. Für die Erläuterung ist ein kurzer Exkurs in die Datenverarbeitung von maschinellen Lernanwendungen notwendig: Die Eingabe für maschinelle Lernanwendungen entspricht einem mehrdimensionalen Array aus Daten, welches als *Tensor* bezeichnet [45, S. 35] und in der linearen Algebra zur Beschreibung eines Datenobjekts genutzt wird. Ein Skalar entspricht einem nulldimensionalen Tensor, ein Vektor einem eindimensionalen Tensor, eine Matrix einem zweidimensionalen Tensor und ein mehrdimensionales Feld entspricht einem N-dimensionalen Tensor [133, S. 99]. Bilder werden durch ihre Höhe, Breite und die Anzahl ihrer Farbkanäle dimensioniert, sodass diese Informationen in einen dreidimensionalen Tensor überführt werden [133, S. 110]. Ein farbiges Bild mit  $128 \times 128$  Pixeln wird als  $(128, 128, 3)$ -Tensor ausgedrückt, wobei die dritte Position innerhalb des Tensors die Anzahl der

<sup>26</sup>Die Überlappung ist durch  $\max\{0, \text{FrameLength} - \text{HopLength}\}$  gegeben

<sup>27</sup><https://matplotlib.org/> (besucht am 06.09.2022)

drei Farbkanäle Rot, Grün und Blau repräsentiert. Ein grau-skaliertes Bild der selben Größe würde aufgrund der eindimensionalen Farbauswahl durch einen Tensor der Größe  $(128, 128, 1)$  dargestellt werden. Im Rahmen der vorliegenden Dissertation wurde eine Bachelorarbeit betreut, die u.a. den Farbeinfluss der Spektrogramme auf die Leistungsfähigkeit eines CNNs untersuchte. Dafür wurden neben den aufgenommenen akustischen Prozessdaten der Regenerationsversuche auch Sounds aus den Datensätzen ESC50 und UrbanSound8k genutzt und jeweils Spektrogramme mit den Farbenskalen *magma*, *viridis* und *greys*, einer Graustufenskala, erzeugt, um damit die Leistungsfähigkeit anhand eines vortrainierten CNNs zu überprüfen. Die Analysen mit den drei verwendeten Datenmengen ergaben keinen signifikanten Einfluss auf die Leistung des neuronalen Netzes. Die Farben geben also weder mehr, noch weniger Informationen der Eingabedaten preis; eine mögliche Begründung dieser Erkenntnis wird aus dieser Bachelorarbeit übersetzt wiedergegeben: „Dies lässt sich durch die Tatsache erklären, dass die Bilder von Spektrogrammen auf Grundlage von Tonaufnahmen erstellt werden, die nicht durch Farben bestimmt sind. Bei einem Ton, z. B. einer Kirchenglocke, bleibt der Inhalt des Geräusches immer noch eine Kirchenglocke, unabhängig von der Farbe der Spektrogrammes, welches aus dem Klang des Kirchenglockentons erzeugt werden würde. Spektrogramme werden auf Grundlage der Frequenz des Klangs über die Zeit erstellt, wodurch ist die Frequenz unabhängig von der verwendeten Farbskala ist, die bei der Umwandlung in das Spektrogramm gewählt werden würde.“ [127]

### 5.1.3 Augmentationsmethoden zur Datenerweiterung

Aus der Literatur geht hervor, dass die Prognosequalität und die Qualität des Generalisierungsverhaltens von DL-Modellen bei einer erhöhten Ausgangsdatenmenge gesteigert werden kann. Die Anwendung von Augmentationsmethoden zur Erweiterung einer Datenmenge oder zum Ausgleich unausgewogener Datenmengen ist in der Forschung zu Audioklassifizierungsproblemen gängig, sodass auch für die vorliegenden Daten einige Methoden zur Erweiterung verwendet wurden. Wie bereits in Kapitel 2.3.5 und in [112] erläutert, ist zu beachten, dass einige Techniken zur Datenerweiterung nicht auf Spektrogramme angewendet werden sollte, da sie ein transformiertes Audiosignal auf visueller Ebene darstellen und wichtige Informationen verzerrt werden können. Daher wurden die Erweiterungsmethoden für die folgenden Analysen dieser Arbeit nur auf die Audiodaten angewendet, sodass das Signal selbst verändert und erst danach zu einem Spektrogramm transformiert wurde.

Die Librosa-Bibliothek bietet einige Methoden zur Datenerweiterung, von denen die Methoden zur Tonhöhenverschiebung und zur Zeitdehnung bzw. -stauchung für die Erweiterung der Datenmenge verwendet wurden. Neben Librosa bietet die Python-Bibliothek *Audiomentations*<sup>28</sup> eine Vielzahl von Methoden sowie die Möglichkeit, mehrere Erweiterungsmethoden kombiniert auf ein Signal anzuwenden. Die verwendeten Augmentationsmethoden werden im Folgenden aufgelistet und erläutert. Die englischen in der Literatur häufig verwendeten Begriffe sind jeweils in Klammern hinter den Methoden zu finden. Unter den Erläuterungen einer jeden Methode werden die expliziten Veränderungen angegeben, die im Rahmen der durchgeführten Augmentierung dieser Arbeit verwendet wurden.

Abbildung 5.2 bildet die im nachfolgend erläuterten Augmentationsmöglichkeiten ab, wobei links oben das Spektrogramm des ursprünglichen Rohsignals abgebildet wird. Es handelt sich hier um ein Störgeräusch, damit die Effekte der Augmentation besser zu erkennen sind. Die jeweiligen Veränderungen des Ausgangssignals sind den Bildüberschriften der einzelnen Spektrogramme zu entnehmen.

---

<sup>28</sup><https://github.com/iver56/audiomentations> (besucht am 14.07.2022)

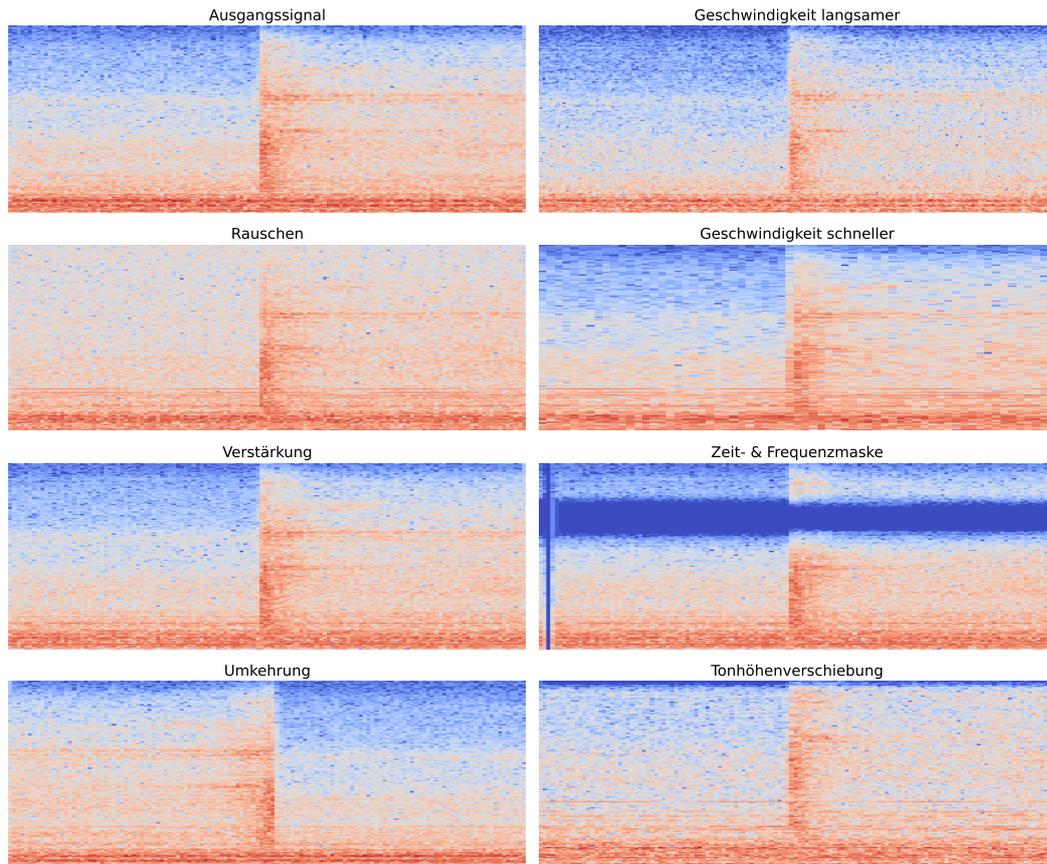


Abbildung 5.2: Verschiedene Augmentationsmethoden

### 1. Geschwindigkeit ändern (Time Stretch)

Diese Augmentationsmethode verändert die Geschwindigkeit des Signals um einen bestimmten Prozentsatz. Es wird empfohlen, die Geschwindigkeit nur dann zu verändern, wenn diese keiner wichtigen Eigenschaft des Signals entspricht. [80] Änderungsfaktoren kleiner 1 verlangsamen das Signal, Faktoren größer 1 beschleunigen es. Ryan, Takafuji et al. (2020) wählten insgesamt 4 Einstellungen, um die Signale zu verändern und ihre vorliegende Datenmenge zu vergrößern: 2, 5, 0,2, 0,5 [110]. Einen geringeren Wertebereich nutzten Johnson und Grollmisch (2021) [62] mit Faktoren zwischen 0,9 und 1,1.

Die Veränderung der Geschwindigkeit wird in Abbildung 5.2 zweifach dargestellt, da das Signal einmal gestaucht und einmal gestreckt wird. Konkret wird das Signal zum einen um den Faktor 0,5, also um die Hälfte verlangsamt, und zum anderen mit Faktor 2 gestreckt, also doppelt so schnell dargestellt.

### 2. Hintergrundgeräusche/Rauschen hinzufügen (Impulse Response/Add Noise)

Mit diesem Verfahren wird einer Audiodatei ein Rauschen oder andere Hintergrundgeräusche hinzugefügt. Diese Geräusche können Umgebungsgeräusche wie zum Beispiel Hundebellen, eine befahrene Straße oder ähnliches beinhalten. Mit den Rohdaten und den verrauschten Daten lässt sich ein Modell erzeugen, welches durch seine Generalisierung auch im späteren Einsatz robust gegen verunreinigte

Signale ist. Diese Methode ist zum Teil auch auf Spektrogramme anwendbar, indem den Bilddaten ein Rauschen hinzugefügt wird. Ein gewisser Prozentsatz der Pixel wird dafür mit einer Zufallszahl multipliziert, deren Mittelwert eins und deren Varianz ein wählbarer Parameter ist. [80]

Neben dem Hinzufügen von Stör- oder Hintergrundgeräuschen kann ein Audiosignal auch so verändert werden, als würde es von einer bestimmten Quelle erzeugt werden. So können zum Beispiel die Verzerrungen, die durch das Soundsystem eines Smartphones entstehen oder das Echo und die Hintergrundgeräusche in einer großen Halle, simuliert werden. Gerade, wenn ein Modell robust gegenüber Umgebungsgeräuschen sein soll, ist dieses Verfahren wichtig. [80] Die verwendeten Hintergrundgeräusche sind in der Literatur sehr unterschiedlich und durch die große Vielfalt ergeben sich erhebliche Möglichkeiten. Beispielsweise nutzten Sun, Maeda et al. (2021) die Geräusche von Regen, Donner, Flugzeugen, Kettensägen und Autos bzw. Trucks [124] als hinzugefügte Hintergrundgeräusche. Durch ihre Vielseitigkeit wird diese Augmentationsmethode häufig verwendet [117, 81, 112, 126, 24, 111].

Die in dieser Arbeit veränderten Signale fließen mit einem Rauschfaktor (White Noise) von 0,2 in die erweiterte Datenmenge ein.

### 3. Zeit- & Frequenzmaske (Time & Frequency Masking )

Die Zeit- bzw. Frequenzmaske kann als unabhängige, aber auch als kombinierte Methode genutzt werden. Das Prinzip ist für beide Methoden identisch: Eine Zeitmaske entspricht einem festzulegenden Zeitbereich des Signals oder des Spektrogramms, welcher mit dem Wert 0 belegt wird. Das bedeutet, dass für den festgelegten Bereich die Informationen herausgenommen und dem DL-Modell beim Training vorenthalten werden. Gleiches gilt für den Frequenzbereich. Es können auch mehrere Zeiten bzw. Frequenzen aus einem Signal entfernt werden, um den Informationsverlust zu erhöhen [80]. Der Ursprung dieser Methode geht auf eine Veröffentlichung von Park, Chen et al. aus dem Jahr 2019 zurück [96]. Die dort beschriebene Veränderung wird vor dem eigentlichen Training des Modells durchgeführt. Wang, Zou et al. haben ein Verfahren entwickelt, welches Augmentationsmethoden sowohl auf den Eingaberaum als auch auf den verborgenen Raum der tiefen neuronalen Netze anwendet, um die Eingabe und die dazwischenliegenden Merkmalsrepräsentationen zu verbessern [132].

In der in Abbildung 5.2 dargestellten Kombination aus Zeit- und Frequenzmaske wurden die Breite der Zeitmaske mit den Min- bzw. Max-Werten von 0,02 und 0,025 definiert und das Frequenzband mit den Werten 0,3 für das Min-Frequenzband und 0,4 für das Max-Frequenzband eingestellt.

### 4. Verstärkung (Gain)

Die Verstärkung eines Signals wird in Maguolo, Paci et al. grundsätzlich empfohlen [80]. Neben der ganzheitlichen Verstärkung kann auch nur eine zufällig gewählte Zeitspanne des Signals von einer Verstärkung beeinflusst werden. Bezieht sich diese Zeitspanne auf den Anfang bzw. das Ende des Signals führt der Effekt zum Ein- oder Ausblenden (Gain Transition). Mit der Verstärkungsmethode aus der *Audiomentations*-Bibliothek wird die Überblendung des Signals auf einer logarithmischen Skala durchgeführt, um menschlichen akustischen Wahrnehmung näher zu kommen<sup>29</sup>. Die Verstärkung des Signals wird als Augmentationsmethode in [81, 112, 126] ohne spezifische Angaben aufgeführt.

<sup>29</sup><https://github.com/iver56/audiomentations/blob/master/README.md> (besucht am 15.06.2022)

Um das Signal zu Verstärken bzw. die Lautstärke zu erhöhen, wird der Ton mit einem zufälligen Amplitudenfaktor zwischen 6 und 12 multipliziert.

## 5. Umkehrung (Reverse)

Bei der Umkehrung wird das Signal entlang der Zeitachse umgekehrt; der Effekt ist auch als Zeitinversion bekannt. Im visuellen Bereich ist die Spiegelung eines Bildes eine weit verbreitete Augmentierungstechnik, die auf das Audiosignal transformiert wurde. Dies kann im Zusammenhang mit Audioklassifizierungsaufgaben von Bedeutung sein. Die Umkehrung von Audiosignalen wurde erfolgreich von Guzhov, Raue et al. (2021) angewendet [50] und fließt auch in dieser Arbeit in die Erweiterung der Datenbasis ein.

## 6. Tonhöhenverschiebung (Pitch Shifting)

Das Signal wird bei diesem Verfahren, dass von Maguolo, Paci et al. als eine der häufigsten Augmentierungsverfahren deklariert wurde, durch eine Erhöhung oder Senkung der Frequenz verändert. [80] Die Tonhöhe (engl. Pitch) bezieht sich dabei auf die Wahrnehmung der Grundfrequenz eines Geräusches. Diese reicht von niedrigen oder tiefen bis zu hohen oder spitzen Tönen [77, S. 6]. Auf Spektrogramme kann die Ton- und auch die Zeitverschiebung angewendet werden, da das Spektrogramm sowohl Informationen des Zeit- als auch des Frequenzbereichs abdeckt [80]. Sprengel, Jaggi et al. (2016) nutzten eine Tonhöhenverschiebung von etwa 5%, wobei größere Verschiebungen zu schlechteren Ergebnissen führten [117]. Schlüter und Grill (2015) erzielten mit dieser Augmentationsmethode in ihren Untersuchungen im Vergleich zu anderen die besten Ergebnisse [112]. Konkrete Verschiebungen werden in dem Paper von Salamon und Bello 2017 [111] mit den Werten  $-2$ ,  $-1$ ,  $1$ ,  $2$  genannt, wobei sich Ryan, Takafuji (2020) [110] an diesen Werten orientiert und sie entsprechend identisch übernommen hat. Auch Johnson, Grollmisch (2021) nutzten die Grenzen von  $-2$  bis  $2$  und verwendeten zwischen diesen Werten Halbtonsprünge für die Datenerweiterung [62]. Keine spezifischen Angaben zu den verwendeten Parametern der Verschiebung wurden in [126, 5, 24] genannt.

Die Tonhöhenverschiebung wird für die Erweiterung der hier vorliegenden Datenmenge mit einem Faktor von 8 durchgeführt.

## 7. Kombination

Eine weitere Möglichkeit der Datenaugmentierung ist die gleichzeitige Anwendung verschiedener Methoden auf das Ausgangssignal, die beispielsweise durch die Kombination der Zeit- und Frequenzmaske in einem der Spektrogramme der Abbildung 5.2 zu sehen ist. Zusätzlich besteht die Option, die jeweiligen Methoden, die kombiniert werden sollen, mit einer wählbaren Wahrscheinlichkeit einfließen zu lassen. Sofern die Wahrscheinlichkeiten nicht bei 100% liegt, wird der Kombinationsfunktion die Entscheidung über die Auswahl der Methoden überlassen. Das bedeutet, dass für jedes neu generierte Spektrogramm jeweils eine neue Kombination von Methoden gewählt und angewendet wird. Das erschwert allerdings die Vergleichbarkeit der Auswirkung einzelner Methoden auf die Modellgüte im Nachhinein. Dennoch erzielte die zufällige Kombination von Augmentierungen gegenüber den einzeln angewendeten Augmentierungsmethoden laut Johnson und Grollmisch (2021) eine Erhöhung der Robustheit<sup>30</sup> des trainierten Modells [62].

---

<sup>30</sup>Robustheit beschreibt das Potential eines Systems, Veränderungen standzuhalten, ohne die ursprünglichen Strukturen anzupassen.

### 5.1.4 Convolutional Neural Network

Die CNNs werden mit dem Open-Source-Framework Keras<sup>31</sup>/TensorFlow<sup>32</sup> in Python aufgebaut und ausgeführt. Keras bietet für den schichtweisen Aufbau eines CNNs das *Sequential Model*, mit dem das Netz vom Input-Layer sequenziell bis zum Output-Layer manuell angelegt werden kann [40, S. 220]. Üblicherweise werden die Inputdaten zuerst an eine Convolutional-Schicht übergeben, welche in Keras mit mehreren Funktionen zur Erzeugung der nötigen Datenstruktur zur Verfügung steht. Diese Funktion existiert 1-, 2- oder 3-dimensional (1D, 2D, 3D) und in Form einer transponierten Version, da die Datenstruktur abhängig von der Dimensionalität der Eingabedaten bzw. des Tensors ist. Da die Spektrogramme für die durchgeführten Analysen immer identisch transformiert werden und damit eine Bildgröße von  $271 \times 446$  Pixeln aufweisen, fließen sie in jedes Netz mit dem Tensor (271, 446, 3) als Input ein. Die Convolutional-Schicht erzeugt einen individuellen Faltungs-Kernel, mit dem die Schicht-Eingabe gefaltet wird. Unter anderem können die Filterdimension (*kernel\_size*), die Aktivierungsfunktion sowie der Parameter *Padding* manuell ausgewählt und definiert werden. [133, S. 120] Mit Keras kann auf einige Aktivierungsfunktionen zurückgegriffen werden, wie zum Beispiel die bereits erläuterten Funktionen Sigmoid, Tangens hyperbolicus und ReLU (Kapitel 2.3.5) [40, S. 226]. Beim Pooling können, ähnlich wie beim Convolutional-Layer, Parameter, wie die *Dimension* des Pooling-Filters und der Parameter *Padding* individuell ausgewählt werden. Durch dieses Bausatzverfahren liefert Keras die Möglichkeit, für jede Problemstellung individuell angepasste Netze zu erstellen. Wenn alle Schichten hinzugefügt wurden, wird das Netz unter Angabe einer Fehlerfunktion (*loss*) sowie eines wählbaren Optimierungsalgorithmus kompiliert. [40, S. 222], [133, S. 120] Durch das Zusammenschalten der einzelnen Schichten, die jeweils durch einfache Verarbeitungsschritte funktionieren, erreichen CNNs ihre Leistungsfähigkeit [73, S. 192].

Im Rahmen der bereits erwähnten Bachelorarbeit [127] wurde eine Netztopologie entwickelt, die für die folgenden Analysen übernommen wird, da bereits gute Ergebnisse damit erzielt werden konnten. Das CNN baut sich über vier Faltungsblöcke auf, wobei sich jeder Block aus zwei Faltungsschichten mit einem  $3 \times 3$ -Kernel, gefolgt von einer Batch-Normalisierung und einer ReLU-Aktivierungsschicht zusammensetzt. Zwischen den Faltungsblöcken befinden sich  $2 \times 2$  Max-Pooling-Schichten, wohingegen die letzte Schicht dieser gesamten Faltungsbasis durch eine globale durchschnittliche Pooling-Schicht abgeschlossen wird. Der Aufbau dieser Faltungsbasis wird in Tabelle 5.6 dargestellt. Für das Modelltraining bietet Keras eine „Callback“-Funktion an, mit der die Gewichte des Netzes bzw. das gesamte Modell zu einem bestimmten Zeitpunkt gespeichert werden kann. In den für diese Arbeit trainierten Netzen wurde immer das Modell, welches die bisher beste Leistung, gemessen an den Werten der Lossfunktion und der Modellgüte (Accuracy), sowohl auf den Trainings- als auch auf den Validierungsdaten erzielt hat, gespeichert. So gibt es zur finalen Auswertung die Möglichkeit, entweder das Modell des vollendeten Trainings für weitere Prognosen zu nutzen oder jenes, welches während des Trainingsprozesses als das Beste abgespeichert wurde.

Neben den Basis-Schichten (Convolutional- und Pooling-Schichten) fließen weitere Schichten bzw. Parameter in die Architektur des hier verwendeten CNNs ein. Die *Batch Normalization (BN)* wird nach jeder Faltungsschicht durchgeführt, was durch die dargestellte Topologie des Netzes in Tabelle 5.6 nachvollzogen werden kann. Die *Batch Normalization*-Schicht wurde 2015 von Sergey Ioffe und Christian Szegedy entwickelt, um der Problematik, dass ein normierter Schichtinput nicht einem normierten Output entspricht, innerhalb eines maschinellen Lernmodells entgegen zu wirken [60]. Die BN-Schicht sorgt dafür, dass die gewählten Stapel (engl. Batches) während des Trainings normiert werden, indem intern ein gleitender

<sup>31</sup><https://keras.io/> (besucht am 30.06.2022)

<sup>32</sup><https://www.tensorflow.org/> (besucht am 30.06.2022)

Tabelle 5.6: Aufbau der Convolutional-Basis des CNNs

Layer (type)	Output Shape	Anzahl Param
conv2d (Conv2D)	(None, 271, 446, 64)	1792
batch_normalization ()	(None, 271, 446, 64)	256
activation (ReLU)	(None, 271, 446, 64)	0
conv2d (Conv2D)	(None, 271, 446, 64)	36928
batch_normalization ()	(None, 271, 446, 64)	256
activation (ReLU)	(None, 271, 446, 64)	0
maxpooling2d (MaxPooling2D)	(None, 135, 223, 128)	0
conv2d (Conv2D)	(None, 135, 223, 128)	73856
batch_normalization ()	(None, 135, 223, 128)	512
activation (ReLU)	(None, 135, 223, 128)	0
conv2d (Conv2D)	(None, 135, 223, 128)	147584
batch_normalization ()	(None, 135, 223, 128)	512
activation (ReLU)	(None, 135, 223, 128)	0
maxpooling2d (MaxPooling2D)	(None, 67, 111, 265)	0
conv2d (Conv2D)	(None, 67, 111, 265)	295168
batch_normalization ()	(None, 67, 111, 265)	1024
activation (ReLU)	(None, 67, 111, 265)	0
conv2d (Conv2D)	(None, 67, 111, 265)	590080
batch_normalization ()	(None, 67, 111, 265)	1024
activation (ReLU)	(None, 67, 111, 265)	0
maxpooling2d (MaxPooling2D)	(None, 33, 55, 512)	0
conv2d (Conv2D)	(None, 33, 55, 512))	1180160
batch_normalization ()	(None, 33, 55, 512)	2048
activation (ReLU)	(None, 33, 55, 512))	0
conv2d (Conv2D)	(None, 33, 55, 512)	2359808
batch_normalization ()	(None, 33, 55, 512)	2048
activation (ReLU)	(None, 33, 55, 512)	0
global_averagage_pooling2d ()	(None, 512)	0

Durchschnittswert und die Varianz des betrachteten Stapels berechnet wird. Bekannte Netz-Architekturen, wie z.B. die des ResNet50-Modells, weisen die Batch Normalisation mehrfach auf. Dabei wird diese Schicht typischerweise sowohl nach einem Convolutional Layer als auch nach einem Fully-connected-Layer eingesetzt. [16, S. 333–334]

Weitere variable Parameter, die das Training des Netzes beeinflussen können, sind die so genannten Hyperparameter, wie u.a. die *Batch Size* und die Anzahl der *Epochen* (engl. Epoch). Die Trainingsphase eines Modells wird in Epochen unterteilt, wobei die festzulegende Anzahl angibt, wie häufig das Training mit den Trainingsdaten wiederholt werden soll [107, S. 155]. Während solch einer Iteration werden die Trainingsdaten als gebündelte Datenstapel (Batches) dem ML-Modell übergeben. Mit der Batch Size wird die Stapelgröße festgelegt, wobei TensorFlow eine standardisierte Größe von 32 vorschlägt. [48, 320]

Auf die Faltungsbasis der verwendeten Netz-Architektur folgt das voll vernetzte NN mit drei Dense-Layern, wobei den ersten beiden Schichten jeweils eine Dropout- und eine Batch Normalization-Schicht folgt. Beide Dropoutschichten fließen mit einem Wert von 0,4 in das Modell ein. Alle drei Dense-Layer nutzen den *l1.l2-Regularizer*, der sowohl *L1*- als auch *L2*-Sanktionen anwendet. Die *L1*-Regularisierung entspricht dabei der Straffunktion, die proportional zum Absolutwert der Gewichtskoeffizienten ist, also der *L1*-Norm der Gewichtungen. Hingegen ist die Straffunktion der *L2*-Regularisierung proportional zum Quadrat der Werte der Gewichtskoeffizienten, also der *L2*-Norm der Gewichtungen. [16, S. 147] Während des Kompilervorgangs des Modells fließen in Keras der Optimierer *Adam* und die Loss-Funktion *categorical\_crossentropy* mit in das Modell ein. Die Adam-Optimierung ist ein stochastisches Gradientenabstiegsverfahren, das auf der adaptiven Schätzung von Momenten erster und zweiter Ordnung beruht<sup>33</sup>. Die Verlustfunktion beschreibt, wie ein neuronales Netz seine Leistung für die Trainingsdaten beurteilen kann. [16] Die hier verwendete Kreuzentropie-Verlustfunktion kann eingesetzt werden, wenn zwei oder mehr Klassen vorliegen.

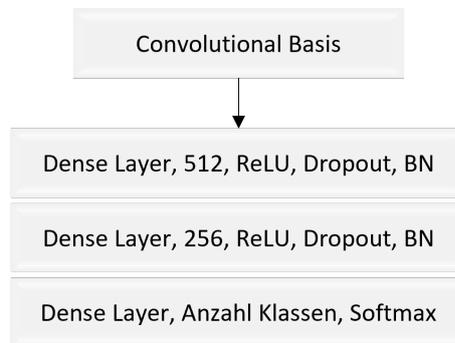


Abbildung 5.3: Voll vernetzter Block des CNNs (vgl. [127])

Die verwendeten Hyperparameter sind von den Daten und der Problemstellung sowie der aufgestellten Netz-Architektur abhängig, sodass keine allgemeingültigen Werte zur Verfügung stehen. Aus der Literatur gehen zum Teil Richtwerte hervor, so liegt z.B. der Dropout meistens zwischen 0,2 und 0,5 [16, S. 149]. Keras bzw. TensorFlow geben auch Standardwerte vor, die genutzt werden können, aber auch individuell variiert werden dürfen. Wie unterschiedlich eingesetzte Netztopologien aussehen können, zeigen folgende Beispiele: In der Veröffentlichung von Sprengel, Jaggi et al. (2016) wird ein Netz aus 5 Faltungsschichten, denen jeweils eine Max-Pooling-Schicht folgt, aufgebaut. Der Netzaufbau wird mit einem Dense-Layer vor der letzten Softmax-Schicht beendet. Vor jeder Faltungs- sowie vor der Dense-Schicht wird eine Batch-Normalisierung durchgeführt. Zusätzliche Dropout-Schichten gehen mit dem Wert 0,2 in die Eingabeschicht, mit 0,4 in die voll vernetzte Schicht und mit 0,4 in die Softmax-Schicht des Modells ein. Trainiert wird mit einer Batch-Size von 8 oder 16 und der Lossfunktion *categorical\_crossentropy*. [117] Inoue, Vinayavekhin et al. (2019) arbeiten hingegen mit nur zwei Faltungsblöcken. Auf die erste Convolutional-Schicht mit Batch-Normalisierung und ReLU-Aktivierungsfunktion folgt eine Max-Pooling- sowie eine Dropout-Schicht mit dem Wert 0,2. Der zweite Faltungsblock besteht aus zwei Faltungsschichten mit Batch-Normalisierung und ReLU-Aktivierungsfunktion, auf die eine globale Max Pooling sowie wieder eine Dropout-Schicht folgt, wobei nun ein Dropout von 0,5 verwendet wird. Abschließend wird das Netz mit einem Dense-Layer und nachfolgendem Softmax beendet. Die Lerndaten werden mit einer 4-fachen

<sup>33</sup>[https://www.tensorflow.org/api\\_docs/python/tf/keras/optimizers/Adam](https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam) (besucht am 02.08.2022)

Kreuzvalidierung validiert und die Trainingsdaten mit einer Batch Size von 256 in 500 Epochen für das Training verwendet. Kompiliert wird mit einer Lernrate von 0,0001 und dem Optimizer Adam. [59]

### 5.1.5 Auswertung

Zur Bewertung der Prognosegüte eines voll trainierten Modells für Klassifizierungsprobleme dient die *Konfusionsmatrix*. Die Matrix bildet in ihren Zeilen die tatsächlichen Labels ab und in den Spalten die prognostizierten Werte. Es lässt also sich ablesen, wie häufig jede der möglichen Kombinationen aus der ermittelten Klasse und der tatsächlichen Klasse vorgekommen ist, also welcher Anteil von dem Modell korrekt klassifiziert wurde [133, S. 194], [89, S. 282]. In Abbildung 5.4 wird die Konfusionsmatrix einer binären Klassifikation dargestellt. Auf den Achsen sind jeweils die Klasse 0 bzw. 1 markiert. Die Werte auf der Diagonalen der Matrix geben die korrekt prognostizierte Anzahl oder, wie in Abbildung 5.4, den prozentual richtig prognostizierten Anteil der Klasse an. In Abbildung 5.4 sind demnach 92,96 % der Prognosen für Klasse 1 und 100 % für Klasse 0 korrekt zugeordnet worden, wohingegen 7,04 % der zu Klasse 1 gehörenden Eingabe fälschlicherweise als Klasse 0 prognostiziert wurden.

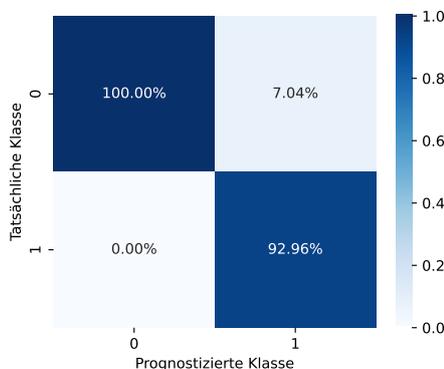


Abbildung 5.4: Konfusionsmatrix

Entsprechend der Anzahl der Klassen wird die Prognose der vorgestellten CNNs als ein Vektor ausgegeben, der die berechnete Wahrscheinlichkeit der Klassenzugehörigkeit angibt. Ein Modell, welches zwischen zwei Klassen unterscheidet und eine unbekannte Eingabe prognostizieren soll, gibt beispielsweise den Ausgabevektor  $[0,920\ 280\ 9, 0,072\ 717\ 61]$  aus. Die Anzahl der Koordinaten des Vektors entspricht dabei der Anzahl der Klassen, wobei der Index des Vektors auf die jeweilige Klasse verweist. In dem gezeigten Beispiel würde das Modell mit einer 92%igen Konfidenz die zu prognostizierende Eingabe der ersten Klasse zuordnen und mit ca. 7% der zweiten. Eine weitere Möglichkeit der Darstellung ist die 1-aus-n-Codierung (One-Hot-Codierung, engl. One-Hot-Encoding) [134, S. 110]. Es handelt sich ebenfalls um eine Vektordarstellung, deren Koordinaten wieder der Anzahl vorliegender Klassen entspricht. Anstelle der Wahrscheinlichkeiten erfolgt die Ausgabe binär, also mit 0 oder 1, indem der Maximalwert der prognostizierten Wahrscheinlichkeiten mit 1 und die verbleibenden Koordinaten mit 0 besetzt werden. Wenn das trainierte Modell eine Prognose ausgibt, wird an der Stelle des Index der am wahrscheinlichsten prognostizierten Klasse der Wert 1 angegeben, ansonsten 0. Für das oben genannte Beispiel entspricht die Prognose dem Vektor  $[1, 0]$ , weil 0,920 280 9 das Maximum der Ausgabewerte darstellt. Diese Darstellung benötigt nur einen geringen Aufwand bei der Datenvorbereitung und erzielt robuste Modellergebnisse [48, S. 25]. Im Fall der in dieser Arbeit vorgestellten Prognosen werden beide Vektoren erzeugt und abgespeichert. Zur

Berechnung der Prognosegüte wird die *Accuracy* berechnet, welche den Anteil der richtig vorhergesagten Beobachtungen im Verhältnis zu allen Beobachtungen angibt [48, S. 21].

## 5.2 Neusand vs. Regenerat

Die erste Analyse zielt auf die Unterscheidung zwischen neuem und gebundenem Sand ab. Ausgehend von diesem Ziel ergeben sich die zwei Klassen *Neusand* und *Nicht-Neusand*.

### 5.2.1 Datenbasis

Tabelle 5.7: Datenbasis zur Analyse *Neusand* vs. *Regenerat*

<i>Nicht-Neusand</i>		<i>Neusand</i>	
Versuch	Anzahl Datensätze	Versuch	Anzahl Datensätze
U1 5 min	87	F34 Umlaufsystem	100
U2 5 min	88	F34 Gebunden V1	101
U4 5 min	70	F34 Gebunden V2	95
UR 5 min	85	V1 F34 - 60 min	768 / 360
U1 10 min	72	V2 F34 - 40 min	759
U2 10 min	91		
U3 10 min	74		
U4 10 min	46		
U5 10 min	68		
U7 10 min	78		
U8 10 min	58		
U9 10 min	87		
U10 10 min	60		
U1 15 min	39		
U3 15 min	76		
U5 15 min	77		
U6 15 min	87		
U7 15 min	74		
U8 15 min	57		
U9 15 min	88		
U10 15 min	76		
F34 Gebunden V1	94		
F34 Gebunden V1	95		
F34 Umlaufsystem	96		
<b>Summe:</b>	1823 / 1415		1823 / 1415

Anhand der Ergebnisse der vorangegangenen Regenerationsversuche (Kapitel 4.4.2) wurde festgestellt, dass bis zu einer Regenerationszeit von 5 Minuten keine neusandähnliche Qualität des Regenerates erreicht

werden kann. Diese Information wird nun genutzt, um alle akustischen Aufnahmen der Regenerationsprozesse, die mit binderbehaftetem F34 durchgeführt wurden, bis zu einem Zeitpunkt von 5 Minuten als *Nicht-Neusand* zu labeln, also Sand, der noch nicht ausreichend neusandähnlich ist. Damit ergeben sich 1823 Audiodateien aus insgesamt 24 Regenerationsvorgängen zur Klasse *Nicht-Neusand*, die sich jeweils aus den Dateien der ersten fünf Prozessminuten, exklusive derer mit Störungen, zusammensetzt. Eine Auflistung der jeweiligen Anzahl einfließender Datensätze wird in Tabelle 5.7 dargestellt. Die unterschiedlichen Datenmengen der einzelnen Regenerationsversuche gehen aus den unterschiedlich häufig auftretenden Störgeräuschen hervor, die aus den jeweiligen 5 Minuten-Paketen entfernt wurden. Ausgehend von dem feststehenden Datenvolumen des *Nicht-Neusandes* wird die identische Anzahl Dateien auch für die *Neusand*-Klasse zusammengestellt, um keinen unausgewogenen Dateninput für das Netz zu erzeugen. Die 1823 Datensätze der Klasse *Neusand* setzen sich zum einen anteilig aus den zwei Neusandaufnahmen mit Quarzsand F34 zusammen und zum anderen aus den jeweilig letzten 5 Minuten der Aufnahmen der Langzeitregenerationen von binderbelastetem Sand. Diese Chargen wurden jeweils 60 Minuten in der Anlage bearbeitet, sodass davon ausgegangen wird, dass die Sandqualität in diesen letzten 5 Minuten hinreichend neusandähnlich ist. Die Klasse *Neusand* ist tabellarisch in der rechten Hälfte der Tabelle 5.7 dargestellt.

In dieser Tabelle 5.7 befinden sich in der *Nicht-Neusand*-Spalte sechs rot markierte Versuche mit der entsprechenden Anzahl einfließender Datensätze. Während des Erstellens der Spektrogramme ist aufgefallen, dass, trotz identischer Bedingungen während der Prozessaufnahmen sowie der identischen Vorgehensweise zur Erzeugung der Spektrogramme, die sechs markierten Dateien ein anderes Erscheinungsbild der Spektrogramme aufweisen. Abbildung 5.5 verdeutlicht diese Unterschiede durch den Vergleich der Spektrogramme je einer Datei aus den Umläufen U8 und U9 der Regenerationsprozesse mit 15 Minuten Laufzeit. Die Transformationen der zwei 3-Sekunden-Dateien werden in Abbildung 5.5 nebeneinander dargestellt, um die Abweichungen deutlich zu machen. Es ist gut erkennbar, dass sowohl im sehr niedrigen als auch im sehr hohen Frequenzbereich Unterschiede vorliegen. Alle rot markierten Versuche weisen diese abweichende Struktur des in 5.5 links dargestellten Spektrogramms auf.

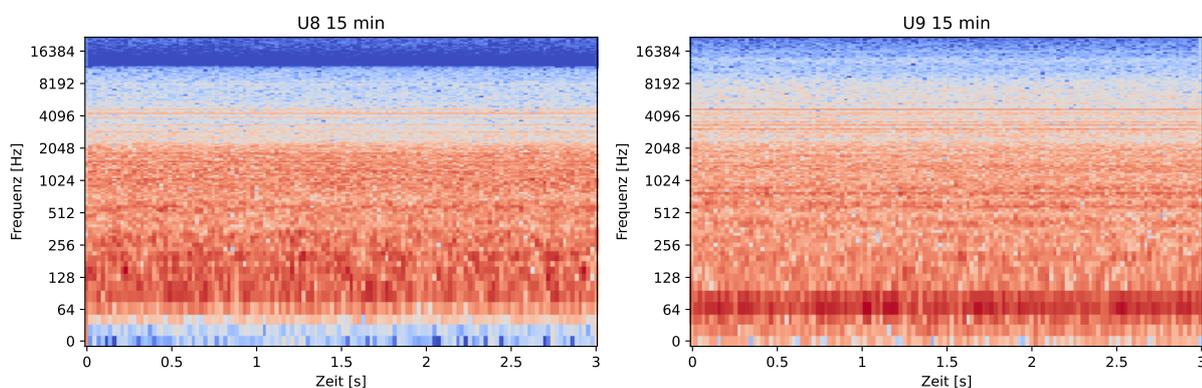


Abbildung 5.5: Abweichungen der Spektrogramme

Die auftretenden Veränderungen lassen sich nicht nachvollziehen oder anhand veränderter Bedingungen erklären. Aus diesem Grund wird eine zweite Datenbasis erstellt, in der die rot markierten Aufnahmen entfernt wurden. Anhand der Prognoseergebnisse kann möglicherweise überprüft werden, ob die auftretenden Unterschiede in den Spektrogrammen Auswirkungen haben oder nicht. Damit liegen zwei Datenmengen vor, die im ersten Fall aus 1823 Dateien aus 24 Regenerationen besteht und im Zweiten aus der reduzierten

Menge von 1415 Dateien aus insgesamt 18 Regenerationen. Die *Neusand*-Klasse wird ebenfalls auf die neue Anzahl von 1415 Datensätzen reduziert. Die Veränderung ist in der rechten Spalte durch die rot markierte Datensatzmenge nachvollziehbar.

## 5.2.2 Ergebnisse und Diskussion

Im ersten Schritt werden die Auswirkungen der erwähnten Unterscheide in den Spektrogrammen untersucht. Bevor die dafür aufgestellten Modelle näher beschrieben und die Ergebnisse diskutiert werden, ist zu berücksichtigen, dass die separierten Testdaten den Trainings- bzw. Validierungsdaten vermeintlich sehr ähnlich sind. Beispielsweise könnte aus der Prozessaufnahme zu Versuch „U1 5 min“ die Teildatei Nr. 50 (Minute 02:30 - 02:33) als Testdatei ausgewählt und dem Training des Netzes damit vorenthalten werden. Die Nachbardateien Nr. 49 und 51, also drei Sekunden vorher bzw. nachher, werden hingegen möglicherweise in der Trainings- oder Validierungsphase eingesetzt und sind dem trainierten Netz somit bekannt. Die Testdatei Nr. 50 wird ihren beiden Nachbarn sehr ähnlich sein und damit für das Modell vermutlich leicht zu klassifizieren.

Tabelle 5.8: Datensätze zur Verifizierung des Modells

Vorenthaltende Datensätze		Verbleibende Datensätze	
Versuch	Datensätze	Versuch	Datensätze
U3 5 min	61	U3 10 min	74
U6 10 min	150	U2 10 min	70
U2 15 min	286	U5 10 min	103
		U7 10 min	74
		UR 5 min	134
		U9 10 min	41
		U10 15 min	84
		U3 15 min	163
		U5 15 min	179
		U6 15 min	136
		U7 15 min	172
		U9 15 min	41
		V1 F34 - 60 min	403
		F34 Gebunden V1	1036
		F34 Gebunden V2	1102
		F34 Umlaufsystem	1081

Für eine aussagekräftigere Beurteilung der CNNs werden den Modellen daher zusätzlich *unbekannte Daten* aus separaten Versuchen zur Prognose präsentiert. Damit soll nicht nur die Güte der Prognose überprüft, sondern auch eine Verifizierung des Modells anhand eines realen Prozesses simuliert werden. Tabelle 5.8 listet diese unbekanntens Datensätze in der linken Spalten auf. Die drei Aufnahmen der Regenerationsversuche wurden bewusst von der Modelldatenbasis ausgeschlossen, fließen also nicht in die Ausgangsdatenmenge mit ein und sind entsprechend nicht in Tabelle 5.7 zu finden. In der rechten Spalte

führt Tabelle 5.8 die bisher nicht verwendeten Audiodateien der Aufnahmen aus dem Umlaufsystem auf. Da nur die ersten 5 Minuten der Regenerationsprozessaufnahmen in die Klasse *Nicht-Neusand* einfließen, werden hier – ab Minute 6 – alle verbleibenden Aufnahmen zur Verifizierung verwendet. Allerdings müssen die Ähnlichkeiten dieser Daten, die die Dateien einer Regenerationsprozessaufnahme implizit verbindet, bei der Auswertung der Prognosen berücksichtigt werden.

Unter der Hypothese, dass die unterschiedlich ausfallenden Spektrogramme, die in Kapitel 5.2.1 vorgestellt wurden, Einfluss auf die Modellgüte und damit die Prognosequalität haben, wurden für die zwei aufgestellten Ausgangsdatenmenge jeweils drei CNNs trainiert. Den trainierten Modellen wurden anschließend die Verifizierungsdaten zur prognose präsentiert, um zu überprüfen, ob die Unterschiede in den Spektrogrammen und damit in den vorliegenden Ausgangsdatenmengen einen signifikanten Einfluss auf die Prognosegüte ausüben. Tabelle 5.9 fasst die Anzahl der Ausgangsdatensätze noch einmal zusammen und enthält weiterführend die Aufteilung in Test-, Validierungs- und Trainingsdaten. Dazu wurden zunächst 10 % der Datenbasis als Testdaten separiert und die verbleibende Datenmenge mit der Hold-Out-Methode zufällig in 70 % Trainings- sowie 30 % Validierungsdaten aufgeteilt. Die beiden Klassen sind für die Analyse und für eine vereinfachte Auswertung mit den Ziffern 0 und 1 kodiert worden; die Zuordnung kann aus den Angaben der Spalte „Klasse“ in Tabelle 5.9 entnommen werden. **Analyse 1** beschreibt die vollständige Datenbasis mit 1823 Ausgangsdateien und **Analyse 2** die Untersuchungen mit reduzierter Datenbasis.

Tabelle 5.9: Datenbasis für Analyse 1 und 2

Analyse	Klasse	Ausgangsdatensätze	Trainingsdatensätze	Validierungsdatensätze	Testdatensätze
<b>Analyse 1</b>	Nicht-Neusand (0)	1823	1149	492	182
	Neusand (1)	1823	1149	492	182
<b>Analyse 2</b>	Nicht-Neusand (0)	1415	891	382	142
	Neusand (1)	1415	891	382	142
	Aug-Nicht-Neusand (0)	9434	8910	382	142
	Aug-Neusand (1)	9434	8910	382	142

Tabelle 5.10: Modellkonfigurationen für die Analysen *Neusand* vs. *Nicht-Neusand*

Analyse	Modell	Batch Size	Epochen	Optimierung & Lernrate	Lossfunktion
<b>Analyse 1</b>	1	32	50	Adam 0,001	categorical crossentropy
	2	16	80		
	3	32	30		
<b>Analyse 2</b>	1	64	50	Adam 0,001	categorical crossentropy
	2 (Aug)	128	50		
	3	32	30		

Für beide Analysen 1 & 2 wurden drei Modelle (Modell 1 - Modell 3) mit den Parametereinstellungen trainiert, die in Tabelle 5.10 aufgeführt sind. Alle Modelle basieren auf den genannten Ausgangsdaten; nur das zweite Modell der Analysereihe 2 verwendete eine erweiterte Datenbasis für das Modelltraining. Diese

Abweichung ist mit der Abkürzung „Aug“ in den Tabellen 5.9 und 5.10 entsprechend gekennzeichnet.

Jedes fertig trainierte und generalisierte Modell wird anhand der Testdaten überprüft und über die zugehörige Konfusionsmatrix ausgewertet. Die Prognosen der Modelle von Analyse 1 sind in den Matrizen der Abbildung 5.6 abgebildet. Im zweiten Modell wird eine Prognosegenauigkeit von 100% erreicht, die Testdaten werden alle korrekt klassifiziert. Für die anderen beiden Modelle ergeben sich ähnlich gute Ergebnisse mit nur wenigen Falschprognosen.

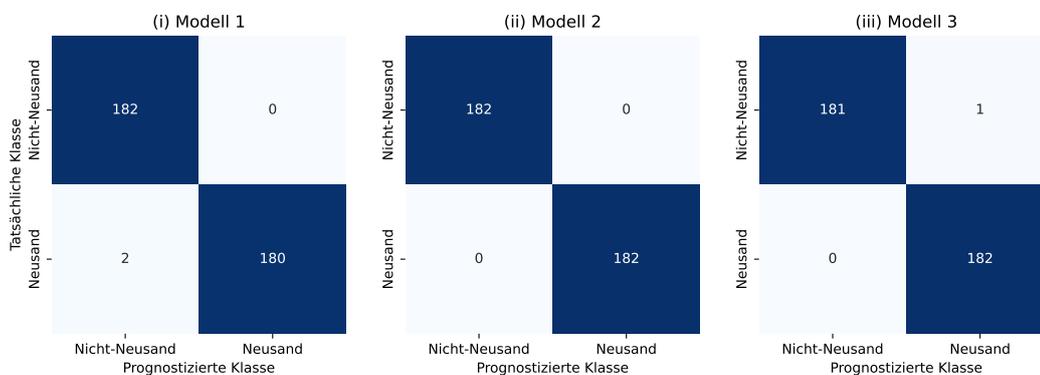


Abbildung 5.6: Analyse 1: Konfusionsmatrizen *Neusand* vs. *Nicht-Neusand*

Auch die Modelle, die im Zuge der zweiten Analyse mit reduzierter Datenbasis durchgeführt wurden, erzielten sehr gute Prognoseergebnisse auf den Testdaten. Abbildung 5.7 reiht die zugehörigen Auswertungen nebeneinander auf.

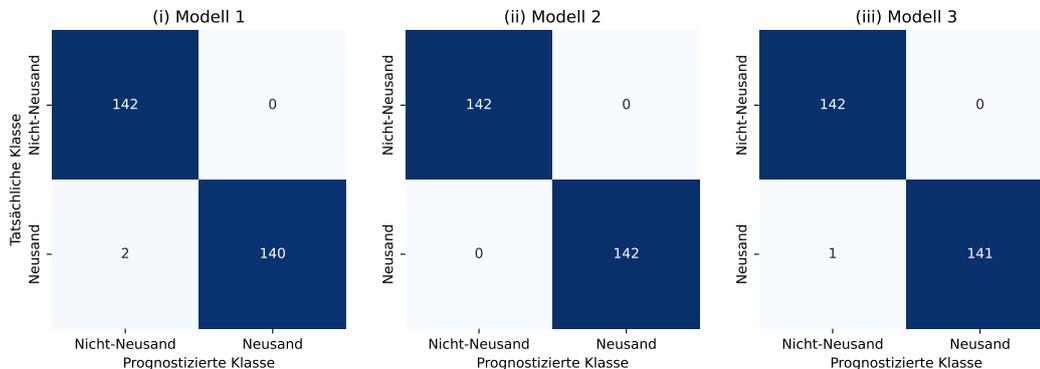


Abbildung 5.7: Analyse 2: Konfusionsmatrizen *Neusand* vs. *Nicht-Neusand*

Wie zu Beginn dieses Kapitels beschrieben sind die Prognoseergebnisse auf den Testdaten kritisch zu beurteilen. Daher werden im nächsten Schritt die in Tabelle 5.8 aufgelisteten Verifizierungsdaten allen Modellen zur Prognose vorgelegt. Der Fokus wird hierbei auf die drei Prozessaufnahmen der linken Tabellenspalte gelegt, da diese den Modellen bisher gänzlich unbekannt sind. Für die Auswertung der Ergebnisse liegt der Schwerpunkt zunächst auf den Unterschieden der Prognosen der Analysen 1 & 2. Mit dem Vergleich soll eine der beiden Datenmengen festgelegt werden, mit der die wesentlichen Leitfragen dieser Dissertation nachfolgend beantwortet werden sollen. Die Prognosen der Verifizierungsdaten werden in Tabelle 5.11 präsentiert. Dabei wird nur die Summe der pro Klasse korrekt klassifizierten Daten ausge-

geben und nicht, wie in einer Konfusionsmatrix, die zusätzliche Information der Falschzuordnungen. Da es nur um den Vergleich der Resultate geht, soll nicht weiter auf die Klassifizierung selbst eingegangen werden. Im Vergleich fällt auf, dass nur die Prognose der Eingaben des „U3 5 Min“-Versuches eine minimale Abweichung aufweisen: 2 von 183 Eingaben wurden in Analyse 2 anders prognostiziert. Alle anderen Prognosen sind im Vergleich identisch. Aufgrund dieser Ergebnisse kann davon ausgegangen werden, dass die unterschiedlichen Spektrogramme keine negativen Auswirkungen auf die Modellgüte haben, sodass die ursprünglich gewählte Datenmenge von 1823 Datensätzen für weitere Analysen vollständig genutzt wird.

Tabelle 5.11: Auswertung der Prognosen auf den Verifizierungsdaten

Versuch	Modell	Analyse 1		Analyse 2	
		Nicht-Neusand (0)	Neusand (1)	Nicht-Neusand (0)	Neusand (1)
U3 5 Min	1	61	0	60	1
	2	61	0	60	1
	3	61	0	61	0
U6 10 Min	1	150	0	150	0
	2	150	0	150	0
	3	150	0	150	0
U2 15 Min	1	286	0	286	0
	2	286	0	286	0
	3	286	0	286	0

Nachdem die Datenbasis zur weiteren Verwendung identifiziert wurde, geht es nun um die erste der ursprünglichen zwei Fragestellungen, nämlich inwieweit sich das festgelegte Qualitätskriterium der Neusandähnlichkeit akustisch erkennen lässt. Übertragen auf die vorliegenden Daten ist die Frage, ob es mit Hilfe eines DL-Modells möglich ist, unzureichend regenerierten Sand als *Nicht-Neusand* und hinreichend regenerierten Sand, der also ausreichend neusandähnlich ist, als *Neusand* zu klassifizieren. Da aus den vorherigen Untersuchungen des Regenerationsverhaltens der Anlage hervorging, dass eine Prozesszeit von 5 Minuten nicht ausreichend für eine brauchbare Sandqualität ist, liegen mit den drei unbekanntem Prozessaufnahmen (U3 5 min, U6 10 min und U2 15 min) Audiodaten vor, die mindestens für die ersten 5 Minuten als *Nicht-Neusand* prognostiziert werden müssten. Für das Modelltraining werden die zusammengestellten Daten aus Tabelle 5.7 verwendet, allerdings wird die gesamte Prozessaufnahme *V2 F34 - 40 min*, die mit Neusand aufgenommen wurde, aus der Datenbasis für das Modelltraining entfernt. Der Versuch (*V2 F34 - 40 min*) wird also mit 759 Datensätzen aus der ursprünglich zusammengestellten Datenmenge der *Neusand*-Klasse ausgeschlossen. Die Audiodaten dieses Versuches sind dem trainierten Modell damit unbekannt und können nach dem Training ebenfalls zur Verifizierung verwendet werden, um zu prüfen, ob sie auch als Neusand klassifiziert werden. Die angepasste Datenmenge sowie die Aufteilung in Test- (10 % der Gesamtdaten), Trainings- (70 % der Lerndaten) und Validierungsdaten (30 % der Lerndaten) befinden sich in Tabelle 5.12.

Das CNN wurde mit einer Batch-Size von 32 und 80 Epochen trainiert, wobei der Optimierer *Adam* mit einer Lernrate von 0,001 und *categorical\_crossentropy* als Verlustfunktion eingesetzt wurde. Auf Basis der Testdaten liegt nach dem Training eine Modellgüte von 100 % vor, die die blau dargestellte Konfusionsmatrix aus Abbildung 5.8 widerspiegelt. Wie bereits erwähnt, sind die Testdaten den Lerndaten zu ähnlich, um sich auf diese Auswertung verlassen zu können. Aus diesem Grund werden dem Modell

Tabelle 5.12: Angepasste Datenbasis zur Unterscheidung von *Neusand* und *Nicht-Neusand*

Klasse	Ausgangsdatensätze	Trainingsdatensätze	Validierungsdatensätze	Testdatensätze
Nicht-Neusand (0)	1823	1149	492	182
Neusand (1)	1064	671	287	106

zusätzlich die Verifizierungsdaten, also 61 Datensätze für U3 5 Minuten, 150 Datensätze für U6 10 Minuten, 286 Datensätze für U2 15 Minuten und 795 Datensätze für F34 V2, zur Prognose präsentiert. Abbildung 5.8 stellt in der grünen Matrix (ii) die Prognose auf den Verifizierungsdaten dar, wobei die beiden Klassen *Neusand* und *Nicht-Neusand* auf der x-Achse abgebildet werden und die vier Vorhersagen jeweils untereinander auf der y-Achse aufgetragen werden. Es ist deutlich zu erkennen, dass die vorenthaltenen Regenerationsaufnahmen von dem trainierten Modell korrekt prognostiziert werden können. Die Daten der Neusandaufnahme (F34 V2) werden zu ca. 90 % (686 von 795 Datensätzen) als *Neusand* klassifiziert wohingegen alle drei *Nicht-Neusand*-Versuche zu nahezu 100 % als *Nicht-Neusand* erkannt werden. Die Frage, ob sich die Prozessakustik des hier verwendeten Regenerationsverfahrens mit dem Reibregenerator nutzen lässt, um Regenerat und Neusand akustisch differenzieren zu können, kann mit den vorgestellten Ergebnissen mit „Ja“ beantwortet werden.

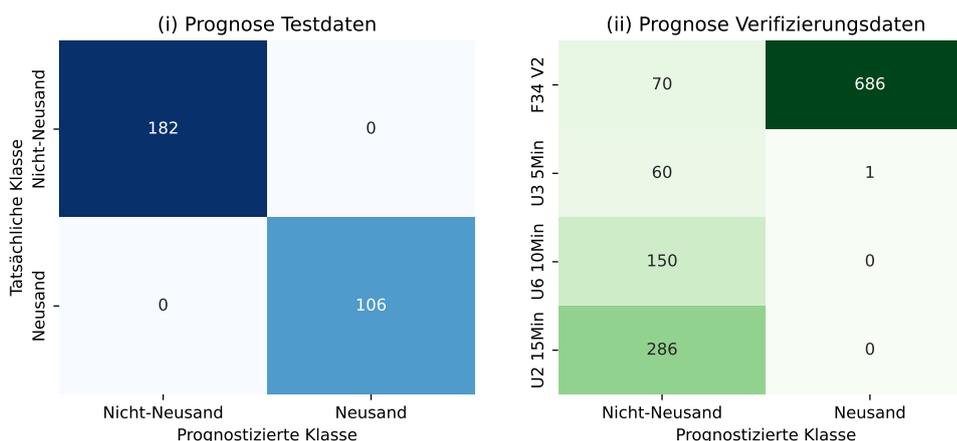


Abbildung 5.8: Auswertung der Test- und Verifizierungsdaten anhand der jeweiligen Konfusionsmatrix

Die zweite Leitfrage, die innerhalb dieses Kapitels beantwortet werden soll, bezieht sich auf den Status des Aufbereitungsprozesses und inwieweit sich dieser akustisch erfassen lässt. Während der bereits durchgeführten Analysen zur Identifizierung der passenden Datenmenge sowie zur Vorhersage von Regenerat bzw. Neusand ist während ergänzender Analysen aufgefallen, dass die Prognosen der verbliebenen Datensätze, also die aufgeführten Daten der rechten Spalte von Tabelle 5.7, für das Umlaufsystem immer *Nicht-Neusand* vorhersagten. Für alle diese Prognosen waren den trainierten Netzen aber jeweils die ersten 5 Minuten der Prozessaufnahmen bereits bekannt, sodass die Klassifizierung zu eben dieser Klasse aufgrund ihrer Ähnlichkeit nicht unerwartet ist. Interessanter gestalteten sich die Prognoseergebnisse der Langzeitaufnahmen „F34 Gebunden V1“, „F34 Gebunden V2“ und „F34 Umlaufsystem“, von denen jeweils die ersten 5 Minuten zur Klasse *Nicht-Neusand* und die letzten 5 Minuten zur *Neusand*-Klasse zugeordnet wurden. Basierend auf den Modellen der ersten Analyse, die in Tabelle 5.10 vorgestellt und ei-

gentlich für die Untersuchung der Datenbasis genutzt wurden, ergaben sich die in Tabelle 5.13 aufgelisteten Vorhersagen für die Langzeitaufnahmen.

Tabelle 5.13: Auswertung der Prognose der Langzeitaufnahmen

Versuch	Modell	Analyse 1		Prozentuale Zuordnung	
		Nicht-Neusand	Neusand	Nicht-Neusand	Neusand
F34 Gebunden V1	1	619	417	60%	40%
	2	509	527	49%	51%
	3	503	533	49%	51%
F34 Gebunden V2	1	870	232	79%	21%
	2	918	184	83%	17%
	3	728	374	66%	34%
F34 Umlaufsystem	1	672	409	62%	38%
	2	674	407	62%	38%
	3	573	508	53%	47%

Die Prognosen fallen sehr unterschiedlich aus, wie die prozentualen Angaben aus Tabelle 5.13 zeigen. Größtenteils wird etwas mehr als die Hälfte der Daten zur Klasse *Nicht-Neusand* zugeordnet. Dieses Prognoseverhalten ist logisch nachvollziehbar, da dem Modell die Aufnahmen bereits zum Teil bekannt sind, nämlich anhand der ersten und letzten 5 Minuten, die in das Modelltraining einfließen. Interessanter als die Klassifizierung selbst ist der Prognoseverlauf, der nachfolgend genauer betrachtet wird, um herausfinden, wie sich die Prognose über den zeitlichen Verlauf verhält und ob bzw. zu welchem Zeitpunkt sich das Modell „umentscheidet“. Zur Visualisierung wurden die Vorhersagen des dritten Modells verwendet, weshalb diese Daten in Tabelle 5.13 zur besseren Zuordnung rot markiert wurden. Abbildung 5.9 zeigt die Verläufe der Prognosen für alle drei Regenerationsaufnahmen. Dafür sind jeweils auf den y-Achsen die Klassen 0-*Nicht-Neusand* und 1-*Neusand* aufgetragen. Die x-Achse beschreibt den kodierten Zeitverlauf der Regenerationsaufnahmen<sup>34</sup>, wodurch die Graphen somit den zeitlichen Verlauf der Klassenzuordnung darstellen. Das bedeutet für (i), dass die ersten ca. 540 Audioaufnahmen als *Nicht-Neusand* klassifiziert werden, während die letzten ca. 340 Aufnahmen (ab 860), als ausreichend neusandähnlich erkannt werden, um der Klasse 1 zugeordnet zu werden. Zwischen diesen Bereichen treten deutliche Schwankungen in der Prognose auf, die durch die Sprünge zwischen den Klassen 0 und 1 in dem Diagramm erkennbar sind.

Sowohl die Anfangs- als auch der Endbereiche aller Aufnahmen sind klar einer der beiden Klassen zugeordnet. Die Unterschiede innerhalb dieser drei Darstellungen liegen in den Start- bzw. Endzeitpunkten der Schwankungen und den auftretenden Häufigkeiten der Sprünge zwischen den Klassen. Die Vorhersage für F34 Gebunden V1 der Abbildung 5.9 (i) beginnt bei 540 zu springen; die Schwankung endet bei ca. 860. In diesem Zeitraum sind die Sprünge unterschiedlich dicht, so dass die Abstände ab dem x-Achsenabschnitt von 700 immer breiter werden. Anders verhält sich der Klassifikationsverlauf von F34 Gebunden V2, dargestellt in Abbildung 5.9 (ii). Hier überwiegt im Bereich zwischen ca. 580 und 800 der x-Achse eher die Tendenz zu Klasse 0. Dieses Verhalten ändert sich im Bereich von 800 bis etwa 930; dort tauchen die Schwankungen zwischen 0 und 1 häufiger auf und werden damit im Diagramm dichter abgebildet.

<sup>34</sup>Der Zeitverlauf wird durch die Nummerierung der Eingabe angegeben. Durch die Multiplikation mit  $\frac{3s}{60}$  kann der exakte Aufnahmezeitpunkt ermittelt werden.

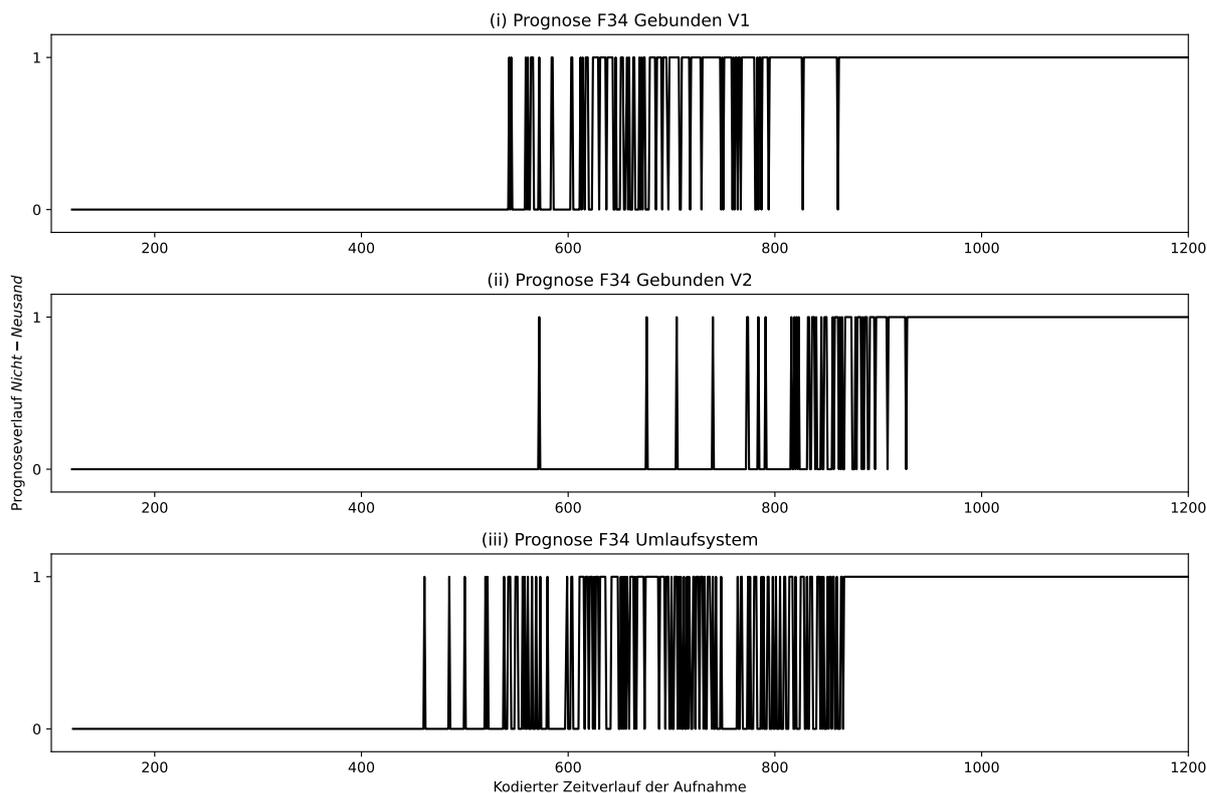


Abbildung 5.9: Modell 3: Darstellung der Prognose von Modell 3

Der dritte Prognoseverlauf (iii) in Abbildung 5.9 beginnt im Vergleich zu den anderen beiden etwas früher zu Springen; in etwa bei Datei 450 (Minute 22:30). Zu Beginn ist ebenfalls ein eher größerer Abstand zwischen den Prognosesprüngen zu erkennen, der auch im weiteren Verlauf der Prognose nach recht dichten Abschnitten mehrfach erneut auftaucht.

Da jedoch von allen drei gezeigten Verifizierungsaufnahmen Teile dieser Aufnahmen den trainierten Modellen bereits bekannt waren, wird im nächsten Schritt geprüft, ob sich eine vergleichbare Prognose auch auf unbekanntem Daten erzielen lässt. Ausgangsbasis bilden wieder die Dateien, die in Tabelle 5.7 aufgeführt sind, allerdings wird aus beiden Klassen der Datenanteil der Regenerationsaufnahme „F34 Gebunden V2“ entfernt. Damit wird „F34 Gebunden V2“ aus dem Modelltraining ausgeschlossen und es liegt eine Langzeitaufnahme vor, die zur Verifizierung des Modells verwendet werden kann.

Tabelle 5.14: Angepasste Datenbasis zur Erkennung des Qualitätsstatus des Regenerates

Klasse	Ausgangs- datensätze	Trainings- datensätze	Validierungs- datensätze	Test- datensätze
Nicht-Neusand (0)	1728	1089	466	173
Neusand (1)	1722	1085	465	172

Die Aufteilung der Datenbasis erfolgt wieder durch das Separieren von 10 % Testdaten mit anschließender Teilung in 70 % in Trainings- und 30 % Validierungsdaten. Die Anzahl der so entstandenen Datenmen-

gen sind in Tabelle 5.14 angegeben. Das CNN wurde mit einer Batch-Size von 32 und 80 Epochen trainiert. Als Optimizer wurde wieder *Adam* mit einer Lernrate von 0,001 verwendet und *categorical\_crossentropy* als Verlustfunktion eingesetzt. Das beste gespeicherte Modell liefert auf den Testdaten ein Accuracy von 100 % mit einem Verlustwert von 0,0178. Ausschlaggebender als die Auswertung der Testdaten sind jedoch die Prognoseergebnisse der Verifizierungsdaten.

Neben F34 Gebunden V2 wurden erneut drei Aufnahmen U3 5 Minuten, U6 10 Minuten und U2 15 Minuten für die Verifizierung des Modells aus den Lern- bzw. Testdaten ausgeschlossen. Die Prognosen für alle vier Prozessaufnahmen finden sich in Tabelle 5.15. Im Vergleich zu den vorherigen Auswertungen lassen sich für die drei Versuche des Umlaufsystems keine Veränderungen feststellen: Nach wie vor werden diese zu 100 % der Klasse *Nicht-Neusand* zugeordnet. Für die Prognose auf den Verifizierungsdaten der Regeneration F34 Gebunden V2 ergibt sich eine Verteilung von 25 % *Nicht-Neusand* und 75 % *Neusand*. Im Vergleich zu den Prognosen aus Tabelle 5.13 fällt auf, dass anteilig mehr Daten als *Neusand* klassifiziert werden.

Tabelle 5.15: Auswertung der Prognose

Versuch	Klassenzuordnung		Prozentuale Zuordnung	
	Nicht-Neusand	Neusand	Nicht-Neusand	Neusand
U3 5 Min	61	0	100%	0%
U6 10 Min	150	0	100%	0%
U2 15 Min	286	0	100%	0%
F34 Gebunden V2	308	909	25%	75%

Unabhängig von der anteiligen Zuordnung der Daten in die beiden Klassen ist der visualisierte Verlauf der Prognose zur Beurteilung der Ergebnisse entscheidend. Aus diesem Grund wird die Vorhersage von F34 Gebunden V2 ebenfalls grafisch dargestellt. Abbildung 5.10 zeigt für die Auswertung der Ergebnisse drei verschiedene Grafiken. In der oberen Grafik (i) wird die 1-aus-n kodierte Prognose wiedergegeben, darunter (ii) befindet sich ein Teilausschnitt von (i), der die Sprünge zwischen den Prognosen auf der angepassten Skala von 150 bis 450 der y-Achse detaillierter aufzeigt. Die unterste Darstellung (iii) stellt die prognostizierten Konfidenz dar, wodurch die Klassifizierung des Modells noch besser nachvollzogen werden kann. Um ein generelles Trendverhalten der Prognose erkennbar zu machen, wurden für die Prognosedaten sowohl die binäre Prognose als auch die berechneten Wahrscheinlichkeiten, mit dem gleitenden Mittelwert geglättet. Dafür wurden jeweils das arithmetische Mittel 30 aufeinanderfolgender Werte der Prognosedaten als roter Verlauf über die einzelnen Diagramme gelegt.

Der Prognoseverlauf weist im Vergleich zu den bereits analysierten Langzeitprognosen eine ähnliche Form auf. Die in (i) dargestellte gesamte Prognose zeigt bis zu einem Zeitpunkt von ca. 10 Minuten (der erste Sprung liegt auf der x-Achse bei 188) die Klassifizierung von *Nicht-Neusand*, was das Regenerat aufgrund der vorangegangenen Analysen bis zu einem Zeitpunkt von 5 Minuten auch ist. Danach folgt die, in (ii) etwas deutlicher dargestellte, Übergangsphase der Prognose, die sich über den x-Achsenabschnitt von 200 bis etwa 400 und demnach über einen Zeitraum von etwa 10 Minuten zieht. Diese Übergangsphase wird auch durch den geglätteten roten Verlauf in (iii) gut sichtbar, der bei etwa 200 einen deutlich erkennbaren Sprung auf das Niveau zwischen 0,4 und 0,6 macht, welcher sich ca. bei 400 wiederholt. Ab 400 nähert sich die Kurve stetig dem Wert 1, also neusandähnlicher Qualität an und wird damit der *Neusand*-Klasse zugeordnet. Die Prognosesprünge zwischen den Klassen 0 und 1, die in der oberen Darstellung

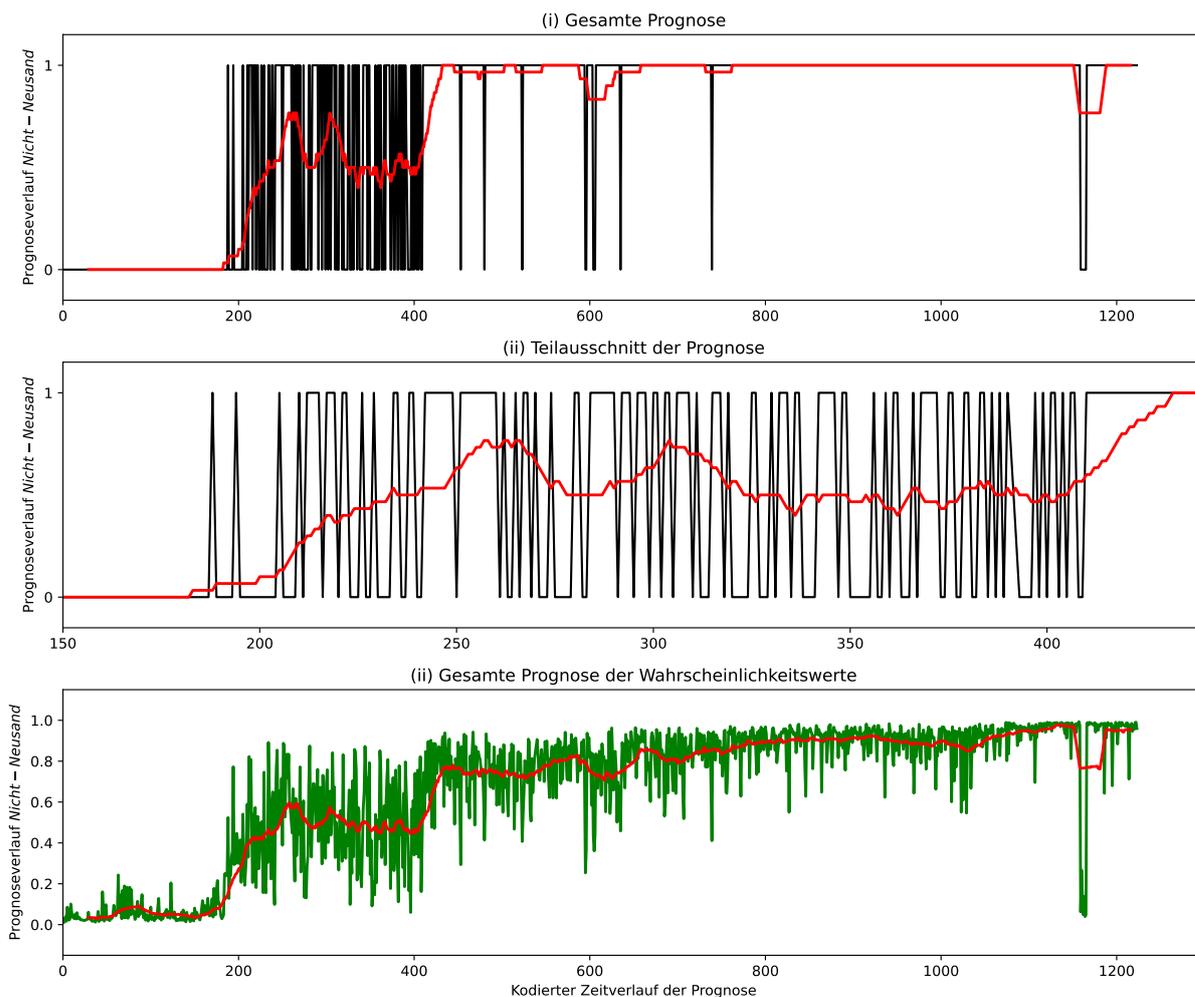


Abbildung 5.10: Prognose von F34 Gebunden V2

ab 400 dargestellt werden, können als Ausreißer interpretiert werden, da sie Anteilig einen Prozentsatz von 2% ausmachen. Die geglätteten Kurven, sowohl in Abbildung 5.10 (i) als auch (iii), übergehen dieses Ausreißer mit nur kleinen Schwankungen. Die Übergangsphase zwischen 200 und 400, welche in 5.10 (ii) vergrößert dargestellt wird, zeigt kein eindeutiges Trendverhalten der Prognose. Diese Phase könnte in weiterführender Forschung näher untersucht werden, indem ergänzende Untersuchungen gezielt auf die Übergangsphase ausgelegt werden, da aktuell nur die Label für die „Extremfälle“ bekannt sind. Insgesamt belegt die Auswertung, dass das trainierte Modell in der Lage ist, eine bislang unbekannte Regenerationsprozessaufnahme zu verarbeiten und eine plausible interpretierbare Prognose zu liefern.

Zusammenfassend haben die Analysen ergeben, dass es auf Basis der im Rahmen dieser Arbeit erfassten Datenlage möglich ist, zum einen Neusand und Regenerat, also binderbehaftete Altsande, akustisch zu differenzieren und des Weiteren auch einen Zeitbereich zu identifizieren, in dem die Qualität des Sandes immer neusandähnlicher wird. Für einen praktischen Anwendungsfall innerhalb eines Gießereibetriebes sind solche Modelle immer individuell zu kalibrieren und bestenfalls immer weiter zu optimieren. Je mehr Informationen den Modellen dort zur Verfügung gestellt werden können, desto genauer kann der prognos-

tizierte Zeitraum eingegrenzt werden. Insbesondere hier ist das Fachwissen der Mitarbeiter entscheidend, die mögliche Grenzwerte bereits aus ihren Erfahrungen kennen und die dafür notwendigen Daten entsprechend zu einer Trainingsbasis zusammenstellen könnten. Aus dem Konzept ging die erweiternde Idee hervor, zusätzliche Neusandklassen zu bilden. Diese weitere Unterteilung würde in Bezug auf den praktischen Einsatz nur Sinn ergeben, wenn unterschiedliche Regeneratqualitäten innerhalb einer Gießerei produziert werden müssen. Auf Basis der durchgeführten Versuche und den daraus gesammelten Informationen wäre eine weitere Klassenbildung im Sinne dieser Arbeit nicht zielführend. Auf diesem Grund wurde die Umsetzung dieser genannten Konzeptidee nicht weiter verfolgt und somit verworfen.

## 5.3 Erkennung von Störgeräuschen

Im Zuge der Audiodatenvorbereitung wurden akustische Störungen identifiziert und aus den Rohaudiodaten entfernt, damit die Folgeanalysen auf reinen und qualitativ hochwertigen Ausgangsdaten basieren können. Da nicht nur in dem hier vorliegenden Prozess akustische Störungen auftreten, sondern es insbesondere in hoch frequentierten Gießereibetrieben zu Störgeräuschen des Prozesses oder zu Störungen durch Hintergrundgeräusche kommen kann, wurden die eliminierten Aufnahmen zur Detektion von Störungen mittels Deep Learning genutzt.

### 5.3.1 Datenbasis

Insgesamt 600 Audiodateien, die aus insgesamt 15 verschiedenen Regenerationsprozessen gefiltert wurden, bilden die Datenbasis für die Klasse *Störgeräusche*. Ausgewählt wurden Geräusche wie zum Beispiel der Sound der Ein- und Auslaufklappen oder Ventilbewegungen der Anlage oder, Gespräche sowie metallische oder dumpfe Geräusche, die durch hinunter- bzw. umgefallene Gegenstände verursacht wurden. Für eine bessere Vergleichbarkeit wurden die Störungen ebenfalls in 3-Sekunden Aufnahmen gespeichert. Abbildung 5.11 zeigt stellvertretend drei Spektrogramme aufgetretener Störgeräusche.

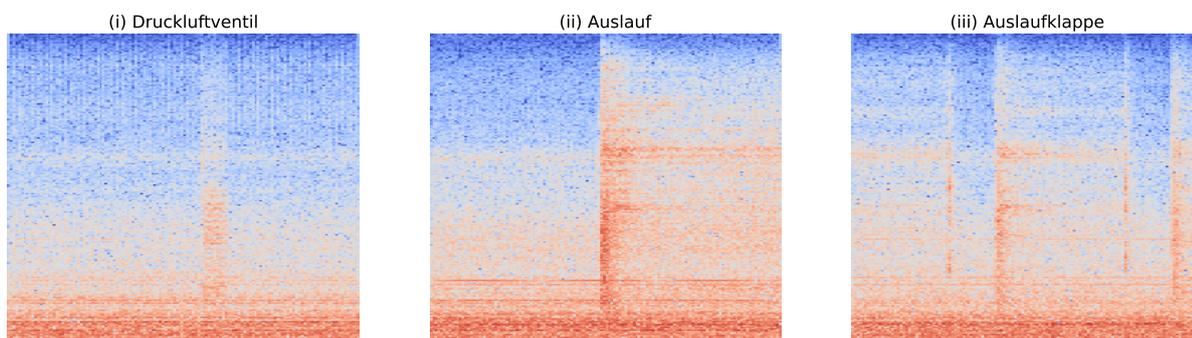


Abbildung 5.11: Darstellung von Störgeräuschen

Im linken Spektrogramm (i) ist ein relativ regelmäßig auftretendes Störgeräusch zu erkennen: Der Luftausstoß eines Druckluftventils. Die weiteren beiden Spektrogramme repräsentieren Geräusche des Austragprozesses zum Ende der Regeneration. Durch das Öffnen der Auslaufklappe entsteht ein Rauschen; zum einen durch den Austrag des Sandes aus der Anlage und zum anderen durch die lautere Prozessakustik aufgrund des Öffnens des geschlossenen Systems. Die mittlere Abbildung (ii) stellt diesen beschriebenen Vor-

gang dar. Eine vollständige Entleerung der Anlage ist aufgrund ihrer Konstruktion kaum möglich, sodass immer eine geringe Restmenge Regenerat im Reibregenerator verbleibt. Innerhalb eines Umlaufsystems kommt es dadurch zu einem erhöhten Sandverlust, der wiederum mit Neusand ausgeglichen werden muss. Des Weiteren erreicht die Anlage beim stetigen Einfüllen einer festgelegten Chargenmenge irgendwann ihre Leistungsgrenze, da die Sandmenge innerhalb der Anlage zu groß und damit nicht mehr handhabbar wird. Um beiden Szenarien entgegen zu wirken, wurde der Auslaufvorgang in die Länge gezogen, indem die Auslaufklappe immer wieder geöffnet und nach Austritt der Restsandmenge geschlossen wurde. Das wiederholte Öffnen bzw. Schließen zeigt Abbildung (iii) in 5.11.

Um dem CNN eine ausgewogenen Datenmenge als Input zu liefern, wurden für die Klasse *Nicht-Störgeräusch* 600 Audiodateien verwendet, die ebenfalls der einheitliche Dateilänge von 3 s entsprachen. Bei der Zusammenstellung der *Nicht-Störgeräusch*-Klasse wurden aus möglichst vielen unterschiedlichen Regenerationsprozessen störfreie Teilaufnahmen verwendet, um eine hohe Bandbreite verschiedener Prozessaufnahmen zu gewährleisten. Zusätzlich wurde bei der Zusammenstellung der Klasse *Störgeräusche* darauf geachtet, nicht alle 1499 detektierten Störungen mit in das Modell einfließen zu lassen, damit auch in diesem Fall Störgeräusche zur Verifizierung verbleiben konnten. Zur Leistungsbeurteilung können dem trainierten Modell demnach alle Rohaufnahmen präsentiert werden, um genau die Audioeingaben zu identifizieren, die nicht störfrei sind. Ein solches DL-Modell könnte der eigentlichen Statusprognose der Sandqualität vorgeschaltet werden. Damit bestünde die Möglichkeit, in einem ersten Schritt die Prozessaufnahme auf Reinheit zu prüfen und nur dann in das folgende Modell weiterzuleiten, wenn sie störfrei ist.

### 5.3.2 Ergebnisse und Diskussion

Für die erste Analyse der zusammengestellten Datenmengen wurden zunächst 10 % Testdaten aus der Ausgangsdatenbasis entnommen und die verbleibende Menge in 70 % Trainings- und 30 % Validierungsdaten geteilt. Auf Basis dieser Datenaufteilung wurden mehrere CNNs trainiert. Für die jeweiligen Modelle sind die Mengen der Datensätze in Tabelle 5.16 angegeben.

Tabelle 5.16: Datenbasis des Modells zur Störgeräuschdetektion

Analyse	Klasse	Ausgangsdatensätze	Trainingsdatensätze	Validierungsdatensätze	Testdatensätze
<b>Normal</b>	Störgeräusch (0)	600	378	162	60
	Nicht-Störgeräusch (1)	600	378	162	60
<b>Augmentiert</b>	Störgeräusch (0)	4002	3780	162	60
	Nicht-Störgeräusch (1)	4002	3780	162	60
<b>k-Fold</b>	Störgeräusch (0)	600	Fold 5: 430	110	60
	Nicht-Störgeräusch (1)	600	Fold 5: 430	110	60

Die Spaltenüberschriften geben jeweils Auskunft über die Anzahl der Ausgangs-, Trainings-, Validierungs- und Testdatensätze. In der ersten Spalte **Analyse** wird auf das Vorgehen der jeweiligen Analyse hingewiesen. Im ersten Fall wird die Datenbasis „normal“ verwendet. Für das zweite Modelle wird die Trainingsdatenbasis erweitert und besteht aus 3780 Trainingsdatensätzen. Dafür wurden die in Abschnitt 5.1.3 erläuterten Augmentationsmethoden auf die Rohaudiodaten angewendet, um die Trainingsdatenmenge um

ein 10-faches auszudehnen. Das dritte Modell basiert wiederum auf der ursprünglichen Datenbasis, welche diesmal durch die Kreuzvalidierung mit  $k = 5$  in Trainings- und Validierungsdaten aufgeteilt wurde. Tabelle 5.16 listet die gewählte Aufteilung des fünften Folds, da diese die beste Prognose lieferte.

Mit unterschiedlichen Hyperparametern wurden mehrere Modelle trainiert und drei repräsentative Ergebnisse ausgewählt, die im Folgenden ausgewertet werden. Das CNN für die „normale“ Analyse zwischen *Störgeräuschen* und *Nicht-Störgeräuschen* wurde mit einer Batch-Size von 32, 100 Epochen, der Adam-Optimierung mit einer Lernrate von 0,001 und der „categorical crossentropy“ als Lossfunktion trainiert. Das Training des DL-Modells, für das die augmentierte Datenmenge als Basis diente, wurde mit den folgenden Parametern gestartet: Batch-Size = 32, 100 Epochen, Adam-Optimierung mit einer Lernrate von 0,001 und „categorical crossentropy“-Lossfunktion. Das dritte Modell weist ebenfalls eine Batch-Size von 32 auf, wurde jedoch nur mit 30 Epochen trainiert. Tabelle 5.17 listet für alle drei Analysen die Werte der Verlustfunktion sowie die Accuracy der leistungsstärksten Netze auf.

Tabelle 5.17: Trainingsparameter des Modells zur Störgeräuschdetektion

Analyse	Auswertung	Trainingsphase	Validierungsphase	Testphase
<b>Normal</b>	Loss	0,1013	0,3	0,093
	Accuracy	0,9931	0,9594	1,0
<b>Augmentiert</b>	Loss	0,0227	0,7815	1,1858
	Accuracy	0,9959	0,8344	0,7083
<b>k-Fold</b>	Loss	0,0691	0,0666	0,0691
	Accuracy	0,9976	1,0	1,0

Für die Modelle mit einfacher Validierung sowie mit Kreuzvalidierung werden auf den Testdaten sehr gute Ergebnisse erzielt. Die Werte der Verlustfunktion sind mit 0,093 und 0,0691 sehr niedrig und die Accuracy weist eine 100%ige Güte auf. Diese lässt sich entsprechend auch in den zugehörigen Konfusionsmatrizen (i) und (iii) in Abbildung 5.12 ablesen, die die Zuordnung der Testdaten in die jeweiligen Klassen angeben.

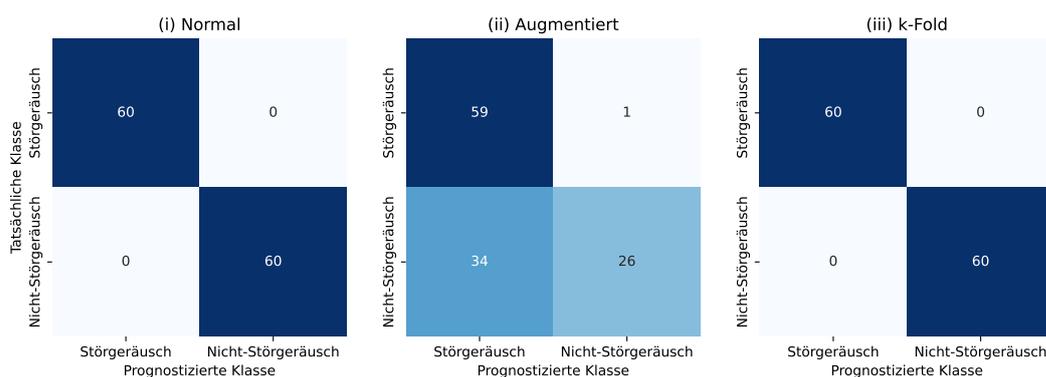


Abbildung 5.12: Konfusionsmatrizen der drei DL-Modelle

Eher ungewöhnlich ist das Ergebnis des CNNs basierend auf der augmentierten und damit deutlich größeren Datenbasis. Häufig führt ein erhöhtes Datenvolumen auch zu einer Leistungssteigerung, wie in

Kapitel 3.2 verdeutlicht wurde. In dem hier vorliegenden Fall sind jedoch nur die Vergleichswerte aus der Trainingsphase (Tabelle 5.17) in einem ähnlichen Bereich wie die der anderen beiden Modelle. In der Validierungs- und auch in der Testphase kristallisiert sich kein gutes Verallgemeinerungsverhalten des Netzes heraus. Dies lässt sich auch aus der zugehörigen Konfusionsmatrix (Abbildung 5.12 (ii)), die die Prognoseergebnisse der Testdaten zusammenfasst, erkennen. Die Zuordnung in Klasse 0, also die Erkennung der Störgeräusche, weist noch eine hohe Genauigkeit von 59 (von 60) richtig klassifizierten Dateien auf, allerdings kann das Modell die Daten der Klasse 1 (*Nicht-Störgeräusch*) nicht gut differenzieren: Nur 26 Dateien werden richtig prognostiziert, die verbleibenden 34 wurden fälschlicherweise als Störgeräusch identifiziert und Klasse 0 zugewiesen.

Zur Verifizierung der Modelle stehen, wie bereits erwähnt, 23 ganzheitliche Aufnahmen verschiedener Regenerationsprozesse zur Verfügung. Den trainierten Modellen werden nacheinander alle Prozessaufnahmen zur Prognose vorgelegt. Die Dateien sind bereits im Vorfeld mit 0 (*Störgeräusch*) bzw. 1 (*Nicht-Störgeräusch*) manuell gelabelt worden. Die Prognose wird One-Hot-codiert und im Falle einer Störung als  $[1, 0]$  ausgegeben. Die Prognose für jede Datei wird dann mit dem zuvor z.B. zugeordneten Label der Eingabe verglichen, um eine Aussage über die Richtigkeit treffen zu können. Um die Ergebnisse von 3 mal 23 Auswertungen besser zusammenfassen zu können, wird für jede Aufnahme der korrekt prognostizierte Anteil prozentual bestimmt und pro Analyse zusammenfassend in einem Boxplot-Diagramm dargestellt.

Ein Boxplot besteht dabei aus drei wesentlichen Komponenten: Der Box, den Antennen und dem eingezeichneten Median. Zusätzlich können auch auftretende Ausreißer mit in das Diagramm integriert werden. Die Box umspannt den Interquartilsabstand zwischen dem ersten und dem dritten Quartil der vorliegenden Datenmenge. Die aus der Box reichenden Antennen bilden den Minimal- bzw. Maximalwert der vorliegenden Datenreihe ab. Im Falle von erkannten Ausreißern werden diese aus dem Boxplot ausgeschlossen und separat im Diagramm abgebildet. Die Antennengrenzen werden dann korrigiert und entsprechen den Minimal- bzw. Maximalwerten der verbleibenden Datenmenge ohne Ausreißer. Der Median, der als Strich durch die Box zu erkennen ist, ist ein Lageparameter der Statistik und teilt die Datenmenge in zwei gleich große Hälften. Er gibt genau den Wert an, für den 50 % aller Messwerte kleiner oder gleich und 50 % der Messwerte größer oder gleich sind. Dadurch ist der Median dadurch im Gegensatz zum Mittelwert (arithmetisches Mittel) nicht empfindlich gegenüber Ausreißern. [44, S. 405] Abbildung 5.13 zeigt die Boxplots zur Auswertung der drei vorgestellten und trainierten Modelle anhand der Verifizierungsdaten.

Die Zugehörigkeit der Plots zu den Modellen kann über die y-Achse abgelesen werden. Auf der x-Achse ist die Skalierung der Prognosequalität bis 100 % angegeben. Ausgehend von den vorgestellten Ergebnissen des Modelltrainings ist es nicht verwunderlich, dass das augmentierte Modell die schlechteste Prognosegüte aufweist. Zwar liegt das erzielte Maximum der 23 Prozessaufnahmen bei 84 %, das Minimum allerdings bei nur 11 %. Die hohe Spannweite spiegelt sich auch in der Box wider, die von 28 % bis 61 % reicht. Der Median der Datenmenge liegt bei etwa 50 %, womit die Generalisierungsunfähigkeit des Modells ebenfalls belegt wird. Die Ergebnisse der anderen beiden Modelle weisen hingegen sehr gute Prognosen anhand der Verifizierungsdaten auf. Für die Auswertung beider trainierter CNNs ergibt sich ein Median von 91 % und Maximalwerte von 98 % (k-Fold) bzw. 97 % (normal). Die Spannweite hinsichtlich der Minimalwerte geht jedoch etwas weiter auseinander. Die Analyseergebnisse, basierend auf dem kreuzvalidierten Modell, weisen das Minimum bei 69 % auf. Es ist zu berücksichtigen, dass zwei ausreißende Prognosen bei 45 % bzw. 57 % nicht in den Boxplot integriert werden. Die in violett dargestellte Box spannt den Interquartilsabstand zwischen 85 % und 94 % auf. Die Verifizierung des *normalen* Modells liefert die orangene Box, die mit ihren Grenzen bei 86 % und 94 % etwas kompakter ist. Das angepasste Minimum liegt bei 68 %, denn bei

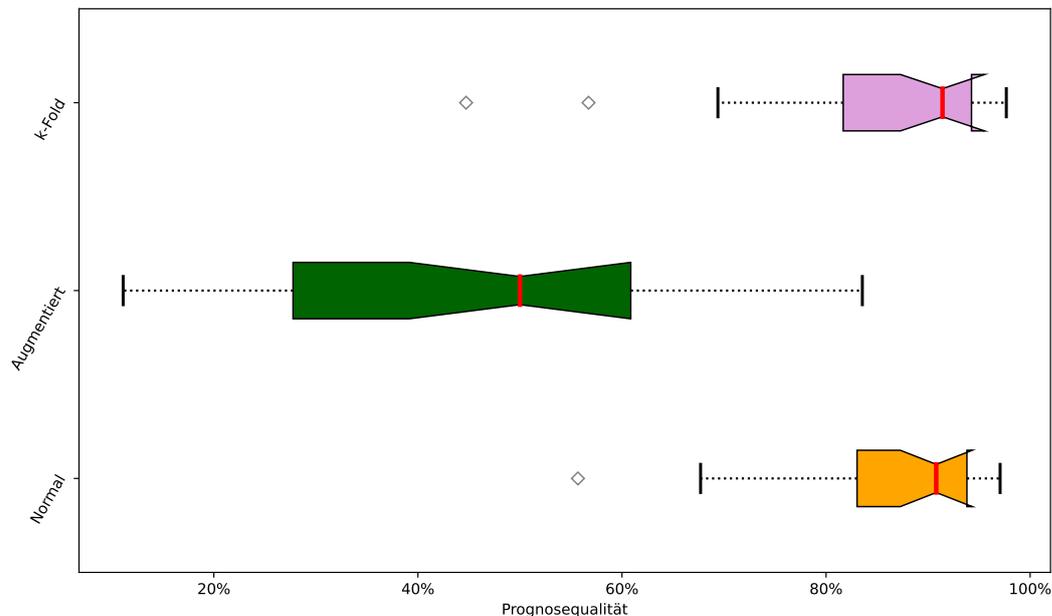


Abbildung 5.13: Auswertung der Prognosegüte

56% wurde die Prognose als Ausreißer erkannt, welcher nicht mit in den Boxplot einfließt. Inwieweit hier ein aussagekräftiger Vergleich zwischen den beiden Modellgüten gezogen werden kann bzw. sollte ist schwer zu beurteilen. Fakt ist jedoch, dass, im Gegensatz zum Modell basierend auf der augmentierten Datenmenge, deutlich bessere Prognosegüten erzielt werden können. Aufgrund der Rohdatenbasis der Modelle, die ohne Datenerweiterungen zu guten Prognoseergebnissen führten, können zudem Speicherkapazitäten und Rechenzeit gespart werden.

Durch die vorgestellten Ergebnisse wurde gezeigt, dass eine Unterscheidung zwischen einer störfreien Audioaufnahme eines Prozesses und möglichen Störungen mit Deep Learning-Modellen möglich ist. Der Einsatz eines solchen Modells im laufenden Betrieb bietet die weiterführende Option, die prognostizierten Störungen zu speichern, um die Datenbasis kontinuierlich mit neuen prozessspezifischen Szenarien zu erweitern. Dadurch besteht die Möglichkeit, das bestehende CNN kontinuierlich neu bzw. weiter zu trainieren, um noch genauere Prognosen erzielen zu können. Zum anderen besteht die Chance, bei einer ausreichend verfügbaren Datenlage, zwischen den auftretenden Störgeräuschen zu unterscheiden, also das Klassifizierungsproblem auszuweiten. Möglicherweise gibt es kritische Störgeräusche, z. B. im Bereich vorausschauender Wartung (Predictive Maintenance), die bei einer frühzeitigen Erkennung einen langen Stillstand der Anlage verhindern könnten.

## 5.4 Verschiedene Sandsorten

### 5.4.1 Datenbasis

Wie bereits zu Beginn des 5. Kapitels erwähnt, wurden auch Regenerationsprozesse mit unterschiedlichen Sandsorten durchgeführt. Inwieweit diese Sandsorten von einem DL-Modell unterschieden werden können, wird anhand der folgenden Datenbasis analysiert: Insgesamt fünf Sandtypen fließen in die Ausgangsda-

tenmenge ein, für die jeweils zwei Chargen Neusand á 30 kg regeneriert und dabei die Prozessakustik aufgenommen wurde (Tabelle 5.1). Die Regenerationsprozesse wurden für ca. 60 Minuten durchgeführt. Mit diesen Langzeitaufnahmen bietet es sich an, zwei Modelle zu erstellen, um die Unterscheidbarkeit der Sande zu untersuchen. Da jeweils zwei Aufnahmen pro Sandsorte vorliegen, können für das erste Modell jeweils die ersten Versuchsaufnahmen als Lernbasis für das DL-Modell verwendet werden. Die verbleibenden zweiten Prozessaufnahmen dienen dann als Datenbasis zur Verifizierung. Für das zweite Modell wird diese Konstellation der Datenzusammenstellung getauscht, sodass diesmal die zweiten Prozessaufnahmen die Lerndatenbasis bilden und die darauf trainierten Modelle mit den Audioaufnahmen der ersten Versuche verifiziert werden können. Beide Varianten sind mit der Anzahl entsprechender Datensätze in Tabelle 5.18 dargestellt, wobei Störgeräusche bereits aus den Datenmengen entfernt wurden. Durch die nicht exakt identische Prozesslaufzeit und die individuell auftauchenden Störgeräusche bilden sich minimale Mengenunterschiede der Ausgangsdaten, die jedoch nicht signifikant genug sind, um in diesem Fall von einer unausgewogenen Datenmenge zu sprechen. Aufgrund dessen, dass je Sandsorte nur zwei Aufnahmen existieren, muss beachtet werden, dass keine dichte Datenbasis der Geräuschaufnahmen pro Sandsorte vorliegt. Übertragen auf die bekannte MNIST-Datenmenge [75], die handgeschriebene Ziffern von 0 bis 9 beinhaltet, könnte das hier vorliegende Szenario wie folgt heruntergebrochen werden: Zur Identifizierung der Ziffern von 0 bis 9 liegen 10 handschriftlich verfasste Notizen dieser Ziffern von nur zwei Personen vor. Die geschriebenen Ziffern der ersten Person soll dem Modell nun dazu dienen, die Ziffern der zweiten Person zu identifizieren. Da jede Handschrift einzigartig ist, wird es einem Modell basierend auf dieser Zweierkonstellation schwer fallen, ein verallgemeinerndes Muster zu erkennen. Daher umfasst die eigentliche MNIST-Datenbank auch eine Trainingsdatenmenge von 60 000 anstelle von 20 (2 mal 10) Ziffernbeispielen und 10 000 Testdaten [75].

Tabelle 5.18: Datenbasis für die Modelle zur Prognose der Sandsorte

Klasse	Aufnahme	Ausgangsdatensätze	Trainingsdatensätze	Validierungsdatensätze	Testdatensätze	Daten zur Verifizierung
F34 (0)	V1	1198	755	323	120	V2: 759
H33 (1)	V1	1140	718	308	114	V2: 1179
S01 (2)	V1	1168	739	316	117	V2: 1145
S02 (3)	V1	1167	735	315	113	V2: 1161
S03 (4)	V1	1160	731	313	116	V2: 1138
F34 (0)	V2	759	478	205	76	V1: 1198
H33 (1)	V2	1179	743	318	118	V1: 1140
S01 (2)	V2	1145	722	309	114	V1: 1168
S02 (3)	V2	1161	731	314	116	V1: 1167
S03 (4)	V2	1138	717	307	114	V1: 1160

Um ein noch besseres Verständnis für die Unterschiede der Sande zu bekommen, werden erneut die granulometrischen Eigenschaften als Qualitätskriterium herangezogen. Abbildung 5.14 beinhaltet die Verteilungskurven der fünf Sandarten. Zum einen fällt auf, dass sich die Korngrößen der Sande S01, S02 und S03 über weniger Klassen verteilen, wodurch sich die prozentualen Spitzen dieser Verteilungskurven

erhöhen. Die drei Sandsorten sind speziell zusammengestellte Sande der Firma Quarzwerke<sup>35</sup> und sind in Abbildung 5.15 dargestellt. Es liegt ein sehr grober und ein sehr feiner Sand sowie ein Sand mit mittlerer Korngrößenverteilung vor, wie es auch aus Diagramm 5.14 ablesbar ist. Die Sorten H33 (blaue Verteilungskurve) und F34 (rote Verteilungskurve) erstrecken sich über mehr Kornklassen als die anderen drei, was sich in den abgeflachten Kurven dieser Sande widerspiegelt. Generell sind sich die beiden Sorten in ihrer Verteilung ziemlich ähnlich, denn die Kurven liegen zum größten Teil übereinander.

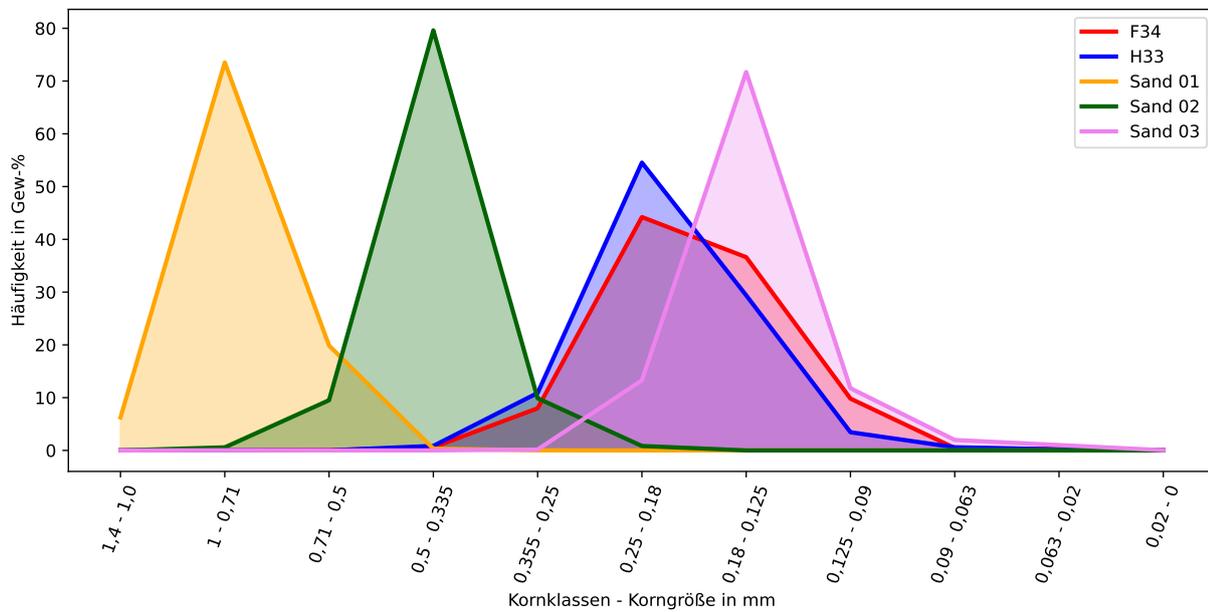


Abbildung 5.14: Verteilungskurven der verschiedenen Sandsorten



Abbildung 5.15: Korngrößen im Vergleich (grob, mittel, fein)

## 5.4.2 Ergebnisse und Diskussion

Unter Berücksichtigung der Tatsache, dass die Datenlage für diese Untersuchung nicht optimal ist, wurden einige Modelle mit den beiden zur Verfügung stehenden Datenmengen trainiert. Tabelle 5.18 beinhaltet In-

<sup>35</sup><https://www.quarzwerke.com/> (besucht am 19.08.2022)

formationen über die Anzahl der Ausgangs-, Trainings-, Validierungs-, Test- und Verifizierungsdatensätze. Die Aufteilung beschreibt die Separierung von zunächst 10 % Testdaten und anschließender einfacher Validierung mit 70 % Trainings- und 30 % Validierungsdaten. Die Verifizierungsdaten liegen als vollständige Prozessaufnahme der Sande vor und sind damit eindeutig gelabelt. Mit dem in Kapitel 5.1.4 beschriebenen CNN und den in Tabelle 5.19 gelisteten Modellkonfigurationen wurden drei DL-Modelle erstellt und trainiert. Dabei basiert die Aufteilung der Inputdaten der ersten beiden Modelle auf der beschriebenen einfachen Validierung. Das dritte Modell nutzt zur Validierung der Lerndaten die Kreuzvalidierung mit einem Fold von  $k = 5$ .

Tabelle 5.19: Modellkonfigurationen für die Analyse verschiedener Sandsorten

Modell	Batch Size	Epochen	Optimierung & Lernrate	Lossfunktion
1	32	30	Adam	categorical
2	16	30	0,001	crossentropy
3 ( $k$ -Fold)	64	50		

Die drei Modelle liefern auf den Testdaten zum Teil unterschiedliche, aber generell gute Ergebnisse. Modell 1 weist bei der Evaluation der Testdaten einen Wert von 0,3277 der Verlustfunktion und eine Accuracy von 0,9207 auf. Die zugehörige Konfusionsmatrix (i) wird in Abbildung 5.16 dargestellt. Dabei fällt auf, dass bei der Prognose zwei Klassen mit Falschprognosen auftreten, nämlich zum einen für die Sandsorte H33, in der zwei Spektrogramme fälschlicherweise der Klasse S02 zugeordnet wurden. Zum anderen können die Spektrogramme der Sandsorte S03 nicht gut klassifiziert werden. Das Modell scheint nicht zwischen S02 und S03 differenzieren zu können und ordnet die Bilder somit beiden Klassen zu. Während des Trainings des zweiten Modells, wurde das Modell der besten Trainingsepoche mit einem Losswert von 0,0438 und einer Accuracy von 1.0 während der Validierungsphase abgespeichert. Die Auswertung der Testdaten lieferte daraufhin Werte von 1,545 und 0,8052 für die Verlustfunktion und die Accuracy. Hier wird bereits deutlich, dass das Modell an Prognosegenauigkeit einbüßt, was sich auch in der Konfusionsmatrix (ii) in Abbildung 5.16 widerspiegelt. In diesem Fall hat das Modell 2 Probleme mit den Spektrogrammen des Sandes S02, denn die 113 Spektrogramme des Sandes S02 werden Anteilig den Klassen F34, S01 und S03 zugeordnet.

Das dritte Modell wurde als das Beste im zweiten Fold des Kreuzvalidierten Trainings erfasst und gespeichert. Im Gegensatz zum ersten Modell sind die Werte der Verlustfunktion (1,7095) sowie der Accuracy (0,8) der Testphase etwas schlechter, aber ähnlich den Werten des zweiten Modells. Die Prognose der Testdaten zeigt Matrix (iii) der Abbildung 5.16. Die hier auftretende Problematik betrifft die Sandsorte H33, die anstelle von H33 als S01 prognostiziert wird. Diese Fehlinterpretation der Spektrogramme lässt sich auch nicht auf eine ähnliche Kornklassenverteilung zurückführen, denn die Verteilungskurven dieser Sande liegen erkennbar auseinander, wie aus Abbildung 5.14 hervorgeht.

Auch in diesem Fall sind die drei Prognoseergebnisse der Testdaten kritisch zu beurteilen. Wie bereits in der Auswertung der Prognosen der *Neusand vs. Regenerat*-Analysen in Kapitel 5.2.2, stammen die separierten Testdaten aus den identischen Gesamtaufnahmen, die für die Trainings- und Validierungsphasen des Modelltrainings verwendet werden. Dadurch ist eine gewisse Ähnlichkeit der Daten implizit vorhanden. Tabelle 5.18 gibt daher direkt die Verifizierungsdaten mit an, um die Modellprognosen anhand unbekannter Daten zu testen. Die Verifizierungsdaten werden den trainierten Modellen nun jeweils in ihrer Gesamtheit

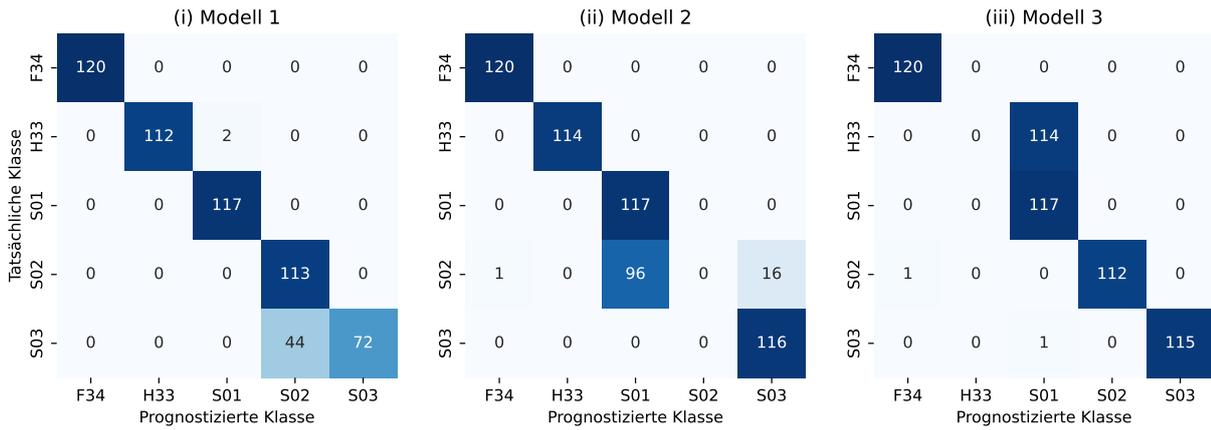


Abbildung 5.16: Konfusionsmatrizen: Auswertung der Testdaten

zur Prognose übergeben und die Matrizen der Abbildung 5.17 repräsentieren die erzeugten Ergebnisse. Zum besseren Vergleich ist die ideale Vorhersage oben links in Abbildung 5.17 als „Ideal“ angegeben. Die Klassenzuordnung der Sandsorten ist der Achsenbeschriftung zu entnehmen.

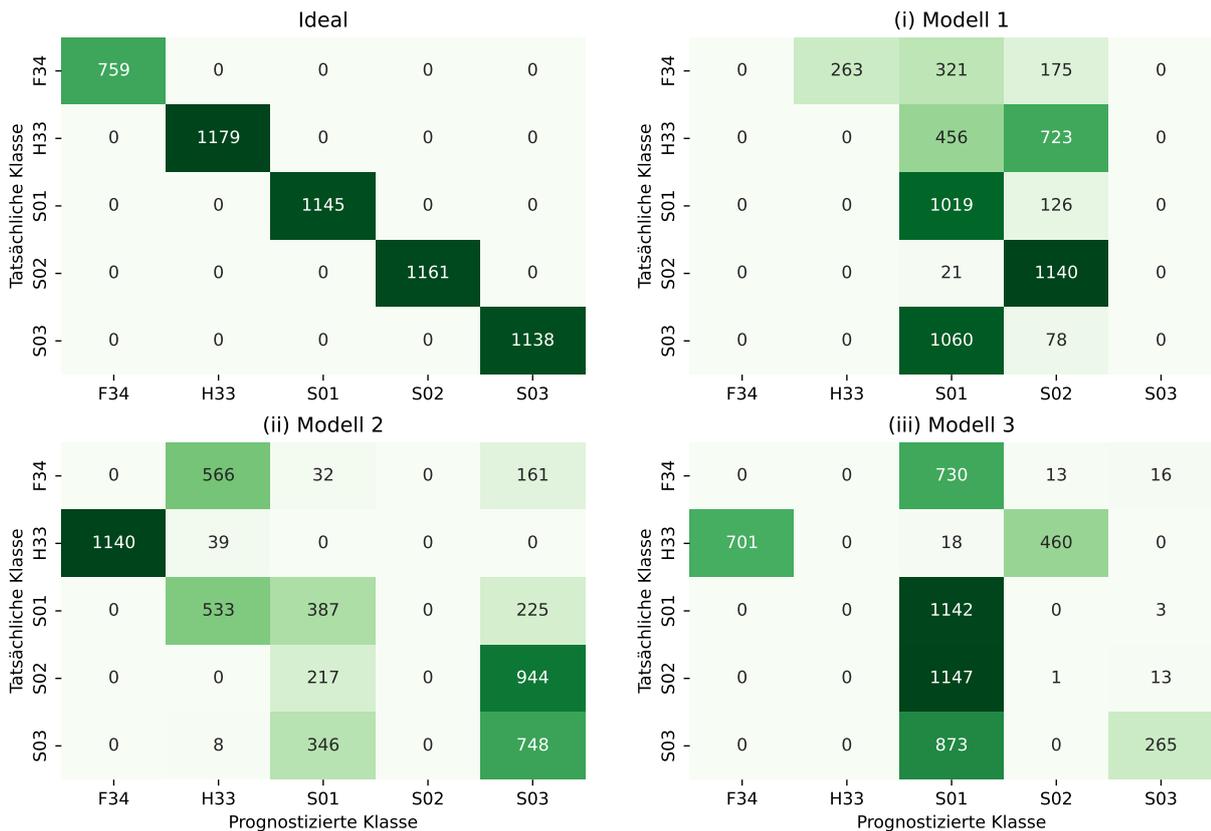


Abbildung 5.17: Konfusionsmatrizen: Auswertung der Verifizierungsdaten

Auf den ersten Blick ist in keiner der vorliegenden Prognosematrizen eine Diagonale zu erkennen, die dem Ideal ähnlich ist, was auf keine gute Prognosequalität hinweist. Das erste Modell prognostiziert einen Großteil der Verifizierungsdaten als den Klassen S01 und S02 zugehörig. Die Audioaufnahmen des Sandes F34 werden weder als dieser prognostiziert, noch werden andere Sandsorten als F34 erkannt; gleiches gilt für den Sandtyp S03, wodurch in den beiden Spalten der Klassen F34 und S03 vertikale 0-Reihen entstehen. Lediglich die Daten der Klassen S01 und S02 wurden zu ca. 89 % (1019 von 1145) bzw. 98 % (1140 von 1161) korrekt vorhergesagt. Somit wurden nur 2 von 5 Klassen richtig prognostiziert, was in einer Accuracy von 40 % resultiert und für keine gute Verallgemeinerungsfähigkeit des Modells spricht. Die Prognose des zweiten Modells (ii) liefert ähnlich schlechte Ergebnisse. Für die Betrachtung der Diagonalen ergeben sich für die Klassen F34 0 %, für H33 3 %, für S01 34 %, für S02 0 % und für Klasse S03 66 % richtig klassifizierte Verifizierungsdaten. Des Weiteren fällt auf, dass in Klasse H33 1140 der 1179 verfügbaren Verifizierungsdaten als Sandsorte F34 anstatt als H33 erkannt wurden. Die sehr ähnliche Korngrößenverteilung beider Sande könnte ausschlaggebend für diese Fehlinterpretation sein, siehe dazu Abbildung 5.14. Auch für die Prognosen des dritten Modells kann keine gute Bilanz gezogen werden. Zum Vergleich wird wieder die Diagonale der Matrix (iii) betrachtet, wobei auffällt, dass nur die Sandsorte S01 nahezu vollständig korrekt klassifiziert wurde. S03 weist immerhin 265/1138 richtig zugeordnete Spektrogramme auf, was einem Prozentsatz von etwa 23 % entspricht. Für alle verbleibenden Sandsorten ist eine Tendenz der Prognosen zu Klasse S01 zu erkennen. So werden 96 % des F34, 98 % des S02 und 77 % des S03 fälschlicherweise als S01 prognostiziert.

Aus allen drei Modellen ist keine Tendenz in den Prognosen zu erkennen. Mit zwei richtig prognostizierten Klassen weist das erste Modell mit einer Accuracy von ca. 40 % im Vergleich zu den anderen beiden das „beste“ Ergebnis auf. Die anderen beiden Modelle erreichen im Verhältnis der richtig prognostizierten zu allen prognostizierten Verifizierungsdaten 22 % (Modell 2) und 26 % (Modell 3). Für alle drei Modelle kann auf Grundlage der verwendeten Daten keine Verallgemeinerungsfähigkeit nachgewiesen werden. Um zu Prüfen, inwieweit die verwendete Datenbasis die Ergebnisse beeinflusst („Garbage in, Garbage out“), werden für eine zweiten Versuchsreihe die Ausgangsdaten mit den Verifizierungsdaten vertauscht. Tabelle 5.18 stellt diesen Tausch und die neue Aufteilung der Daten im unteren Bereich dar. Mit der getauschten Datenzusammenstellung wird erneut untersucht, inwieweit die im Folgenden vorgestellten Modelle Sandsorten unterscheiden können. Die Spezifikationen zweier Modelle sind Tabelle 5.20 zu entnehmen, wobei für das erste Modell auf die einfache Validierung zur Aufteilung der Datenmenge zurückgegriffen wird. Im zweiten Fall wird die Ausgangsdatenmenge mit einem Fold von  $k = 5$  Kreuzvalidiert.

Tabelle 5.20: Parametereinstellungen für die Modelle zur Analyse verschiedener Sandsorten

Modell	Batch Size	Epochen	Optimierung & Lernrate	Lossfunktion
1	128	200	Adam	categorical
2 ( $k$ -Fold)	64	50	0,001	crossentropy

Beide Modelle erzielten im Training gute Ergebnisse. Modell 1 wurde während des Trainings mit Werten der Verkostfunktion von 0,0079 auf den Trainings-, 0,0078 auf den Validierungs- und 0,0078 auf den Testdaten gespeichert. Die Werte der berechneten Accuracy lagen in allen drei Phasen bei 1, also 100 %. Abbildung 5.18 bildet auf der linken Seite die zugehörige Konfusionsmatrix (i) ab, die die perfekte Zuordnung der Testdaten bestätigt. Das zweite Modell lieferte im vierten Fold mit einer Aufteilung von

3874 Trainings- und 970 Validierungsdatensätzen das beste gespeicherte Modell, welches ebenfalls eine Prognosegenauigkeit von 100% auf den Testdaten erreicht, siehe (ii) in der Abbildung 5.18. Auch hier ist es wichtig, die Ergebnisse unter Berücksichtigung der Ähnlichkeit der Testdaten zu den Lerndaten zu beurteilen, weshalb der Fokus erneut auf die Auswertung der Verifizierungsdaten gelegt wird.



Abbildung 5.18: Konfusionsmatrizen der zwei Modelle



Abbildung 5.19: Prognosen auf Basis der Verifizierungsdaten

Die Prognoseergebnisse der Verifizierungsdaten sind in den Matrizen der Abbildung 5.19 abgebildet. Die erste Matrix stellt abermals das Ideal, also eine 100%ige Prognosegenauigkeit dar, um die Modellprognosen besser vergleichen zu können. Ohne erneut im Detail auf die Auswertungen der Verifizierungsdaten einzugehen, sprechen die berechneten Prognosegüten der Matrizen für sich. Die Matrix des ersten Modells (i) weist einen Wert von 9,5% und Modell 2 (ii) von 6% korrekt klassifizierter Daten auf. Der Tausch der Datenmengen hat also zu keiner Verbesserung der Prognose geführt. Auch in diesem Fall können die Modelle auf Basis der vorliegenden Daten die Sandsorten anhand ihrer Prozessakustik also nicht differenzieren.

Generell wird aus den Analysen deutlich, dass die Datenlage, wie bereits zu Beginn vermutet, nicht aussagekräftig genug ist. Es liegen zu wenig, insbesondere unterschiedliche, Aufnahmen der Regenerati-

onsprozesse der Sande vor, um einem DL-Modell genug Informationen zur Generalisierung zu liefern. Um die Machbarkeit einer Unterscheidung von Sandsorten auf Basis der verfügbaren Datenlage ggf. doch noch nachweisen zu können, wird zusätzlich eine dritte Datenbasis erstellt, die sich nur aus den Daten der zwei Sandsorte S01 und S03 zusammensetzt. Aufgrund der binären Zusammenstellung besteht die Möglichkeit, dass es einem DL-Modell hier leichter fallen könnte, zwischen den Sandsorten zu differenzieren. Die reduzierte Ausgangsdatenmenge wird noch einmal in Tabelle 5.21 aufgeführt. Die ursprüngliche Aufteilung in Test- und Lerndaten bleibt bestehen und kann ebenfalls dieser Tabelle entnommen werden.

Tabelle 5.21: Angepasste Datenbasis zur Unterscheidung von S01 und S03

Klasse	Aufnahme	Ausgangsdatensätze	Trainingsdatensätze	Validierungsdatensätze	Testdatensätze	Daten zur Verifizierung
S01 (2)	V1	1168	739	316	117	V2: 1145
S03 (4)	V1	1160	731	313	116	V2: 1138

Zusätzlich wurde darauf geachtet, dass ein deutlicher Unterschied der Korngrößenverteilung der Sandsorten vorliegt, um dem Modell die Differenzierung ebenfalls zu erleichtern. Die verwendeten Sande mit entsprechender Korngrößenverteilung sind in Abbildung 5.20 dargestellt.

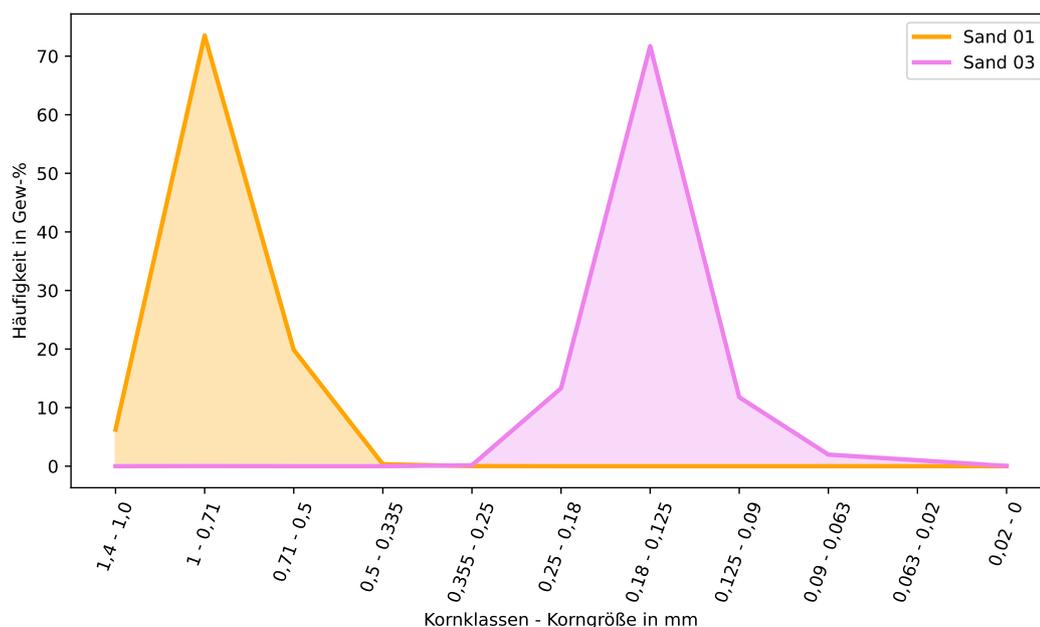


Abbildung 5.20: Korngrößenverteilung der Sandsorten S01 und S03

Das Modelltraining auf Basis der reduzierten Sandsorten wird mit einer Batch-Size von 32 und 80 Epochen durchgeführt. Wieder kommen der Optimierer *Adam* (0,001) und die Verlustfunktion *categorical\_crossentropy* zum Einsatz. Wie bereits in allen vorherigen Analysen, werden auch in diesem Fall die Testdaten perfekt prognostiziert, wie es die linke Matrix in Abbildung 5.21 zeigt. Da die Ausgangsdatenmenge aus den ersten Versuchsaufnahmen der Sande besteht, können die zweiten Audioaufnahmen der Regenerationsprozesse wieder zur Verifizierung verwendet werden. Für S01 stehen 1145 und für S03 1138 Verifizierungsdatensätze zur Verfügung, die dem trainierten Modell für die Auswertung übergeben wur-

den. Die Prognose des Modells ist in der rechten Matrix in Abbildung 5.21 (ii) dargestellt. Anhand dieser Klassifizierung lässt sich eine Genauigkeit von ca. 70 % berechnen, was bedeutet, dass das Modell gewisse Unterschiede anhand der gegebenen Datenbasis erfassen kann. Die Prognose ist zwar noch nicht sehr gut, kann aber womöglich mit einer erweiterten und dichten Datenmenge verbessert werden.

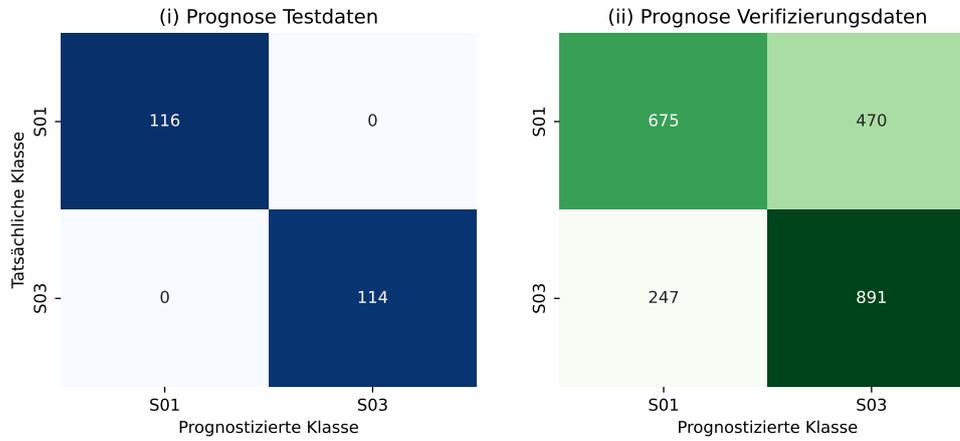


Abbildung 5.21: Prognosen der Sandsorten S01 und S03



# Kapitel 6

## Fazit und Ausblick

Im Rahmen dieser Dissertation wurde ein Regenerationsverfahren zur Aufbereitung anorganisch gebundener Formstoffe mithilfe des so genannten Reibregenerators anhand auftretender Prozessgeräusche untersucht. Ziel der Arbeit war die Beurteilung des Prozessfortschrittes mit KI-gestützten Modellen, insbesondere Verfahren zur Audioklassifikation. Unbehandelter Sand, also Neusand, wurde regeneriert und das zugehörige akustische Signal aufgenommen, um die Audiodaten des Neusandes als Lernbasis für CNNs zu verwenden. Ausgehend von dieser Idee wurde im Sinne einer Machbarkeitsstudie die folgende Hypothese zur Überprüfung aufgestellt:

Mittels der Prozessakustik des hier betrachteten Regenerationsprozesses kann eine Aussage über die Regeneratqualität innerhalb dieses geschlossenen Prozesssystems getroffen werden.

Diese Hypothese wurde anhand von drei Leitfragen untersucht: (1) Nach welchem Qualitätskriterium des Regenerates wird der Prozess ausgelegt? (2) Lässt sich das festgelegte Qualitätskriterium akustisch erkennen? (3) Lässt sich der Status des Aufbereitungsprozesses akustisch erfassen? Im Verlauf der schriftlichen Ausführung wurden alle drei Fragen durch Untersuchungen mit entsprechenden Auswertungen beantwortet.

### 6.1 Fazit

Das Qualitätskriterium wurde als die Reproduktion der „Eigenschaften von Neusand“ festgelegt. Diese wurden in einem ersten Schritt, durch klassische Qualitätskriterien zur Beurteilung von Sanden, mit Neusand erfasst, um den Regenerationsfortschritt anschließend anhand einer durchgeführten Versuchsreihe zu analysieren. Dabei lag der Fokus auf der Siebanalyse sowie Festigkeitsuntersuchungen von Biegeriegeln. Eine wichtige Erkenntnis dieser Untersuchungen lag in dem Grenzwert einer 5-minütigen Regenerationszeit, der nicht ausreichend war, um dem gebundenen Sand durch den Regenerationsprozess wieder eine brauchbare Qualität zu verleihen. Mit dieser Information konnte der Zusammenhang zu den akustischen Untersuchungen gezogen werden, um das Labeling der Audiodaten zu realisieren. Des Weiteren wurde deutlich, dass mit den granulometrischen Kennwerten, den Festigkeitswerten sowie mit weiterführenden Methoden wie den Möglichkeiten von REM-Analysen keine Inline-Prozessüberwachung der Sandqualität während der Sandregenerierung möglich war, weshalb sich die Arbeit auf die Überwachung der Prozessakustik konzentrierte.

Für die Analyse akustischer Daten wurden Convolutional Neural Networks verwendet, die sich bereits im Bereich der Audioklassifikation etabliert haben (vgl. Kapitel 3.2). Mit den CNNs sollte im Bezug auf die Problemstellung untersucht werden, inwieweit sich die Reproduktion des Qualitätskriteriums „Eigenschaften von Neusand“ akustisch erfassen und beschreiben lässt. Das detaillierte Vorgehen sowie die verwendeten Datenmengen wurde in Kapitel 5 vorgestellt. Die finalen Analyseergebnisse legten dar, dass es einem DL-Modell möglich ist, zwischen den Prozessgeräuschen von in der Anlage befindlichen Regeneraten und denen von Neusand zu unterscheiden. Zur Verifizierung der Modelle wurden dafür explizit Prozessaufnahmen mit Neusand sowie mit Regeneraten von dem Modelltraining sowie der -validierung ausgeschlossen. Darüber hinaus konnte auf Basis der verfügbaren Audiodaten belegt werden, dass das trainierte CNN eine Langzeitaufnahme von Regenerat zunächst als *Nicht-Neusand* klassifiziert und nach einer Übergangsphase letztlich als *Neusand* erkennt. Mit diesen Ergebnissen kann die oben genannte Hypothese bestätigt werden.

Mit weiterführenden Analysen wurde zusätzlich gezeigt, dass eine Unterscheidung zwischen einem störfreien Prozess und auftretenden Störgeräuschen mit Deep Learning-Modellen möglich ist. Der Einsatz eines solchen Modells im laufenden Betrieb bietet die Möglichkeit, verunreinigte Aufnahmen direkt zu identifizieren und diese nicht für eine Statusprognose zuzulassen. Abschließend wurde untersucht, ob CNNs auch in der Lage sind, anhand entsprechender Akustikdaten verschiedene Sandsorten zu differenzieren. Mit einer Prognosegenauigkeit von etwa 70% weisen die Ergebnisse trotz geringer Größe der Datenmenge auf ein vielversprechendes Trendverhalten hin.

## 6.2 Ausblick

Die im Rahmen der vorliegenden Dissertation erläuterte Machbarkeitsstudie beruht auf Prozessaufnahmen, die mit gebundenem Sand, jedoch nicht mit thermisch belastetem Sand durchgeführt wurden. Die Ergebnisse sind daher nur für Gegebenheiten nicht thermisch belasteter Sandsysteme interpretierbar. Zu diesen Systemen zählen beispielsweise die Wiederverwendung gebrochener Kerne, nicht verwendeter Modellkästen oder Kokillen. Ein weiterführender Schritt dieser Forschung ist demnach der Einsatz der entwickelten Methodik im betrieblichen Umfeld, um das Verfahren in der Praxis zu validieren. Zusätzlich ist zu empfehlen, die Forschung auf Formstoffsysteme, die durch Gießprozesse mit den hohen Temperaturen flüssiger Metalle in Kontakt kommen, auszuweiten. Dabei kann das gesamte Vorgehen identisch übernommen werden, sodass die Regenerationsprozesse lediglich mit thermisch belasteten Systemen aufzunehmen sind, um entsprechende DL-Modelle mit diesen Audiodaten neu zu trainieren und damit zu kalibrieren. Weitere Verbesserungen der Methodik können durch die Modelloptimierung erreicht werden. Bezogen auf die Deep-Learning-Modelle geht es dabei um die Hyperparameteroptimierung, um ggf. bessere Prognosen zu erzielen. Eine andere Möglichkeit liegt im Vergleich der Ergebnisse bei Verwendung von RNNs anstelle von CNNs, die in Kapitel 2.3.4 bereits kurz erläutert wurden.

Neben den klassisch hergestellten Sandformen bzw. -kernen der Gießereiindustrie setzt sich der 3D-Druck in der Branche immer weiter durch. Das patentierte *Binder Jetting* verklebt den Sand während des Druckvorgangs mit dem Binder, wodurch diese Technologie, bezogen auf die Teilegeometrie, ein hohes Maß an Freiheit bietet und sich ideal für die vielseitigen Gussstücke mit entsprechend vielfältigen Anforderungen der Gießereiindustrie eignet. Die Verwendung anorganischer Bindersysteme resultiert in sinkenden Emissionen und Kosten bei der Abfallentsorgung. [82] Die Firma ExOne<sup>36</sup> liefert mit ihren 3D-Druckern

---

<sup>36</sup>[www.exone.com](http://www.exone.com) (besucht am 06.06.2022)

für anorganische Formen und Kerne eine wirtschaftliche und ökologische Antwort auf aktuelle Nachhaltigkeitsanforderungen. Zusätzlich wird der nicht verdruckte, also der noch ungebundene überschüssige Sand wieder für folgende Druckprozesse eingesetzt [88]. Die Regenerierbarkeit wurde in den aufgeführten Quellen nicht thematisiert, sollte jedoch nicht vernachlässigt werden und könnte mit dem vorgestellten Verfahren dieser Dissertation ebenso zu den Zielen in Bezug auf die Nachhaltigkeit beitragen.

Die Mehrwerte für die Gießereiindustrie im Kontext der Ressourceneffizienz sollten gezielt und kontextspezifisch herausgearbeitet werden, denn gerade im Rahmen der ganzheitlichen Prozessanalyse können diese von großer Bedeutung sein. In der ganzheitlich betrachteten Formstoffsteuerung werden die einzelnen Stationen häufig noch isoliert verwaltet; eine intelligente Vernetzung der Teilprozesse fehlt, wodurch die Anforderungen an eine effiziente und beherrschbare Produktion nicht erfüllt werden können. Um auch langfristig Nachhaltigkeit zu erzielen, sind nicht nur die Potentiale der Einbindung von KI-Methoden zu untersuchen, sondern auch die Entwicklung notwendiger Soft- und Hardwarelösungen in Bezug auf sämtliche Prozessphasen zu unterstützen. So könnten wechselseitige Abhängigkeiten von Teilprozessen und mögliche Ursachen für kritische Situationen erkannt und dem entsprechenden Teilprozess zugeordnet werden. Auch die in dieser Arbeit entwickelte Methodik zur akustischen Bestimmung eines Prozessfortschrittes könnte Anwendung in solch einem Gesamtsystem finden. Dennoch gibt nach Aussagen einer Studie des VDI im Anwendungsbereich KI-gestützter Modelle, besonders in KMU und demnach auch in der Gießereiindustrie, noch erhebliche Schwachstellen und Entwicklungspotentiale: Die aufgeführte Erhebung macht deutlich, dass 60% der befragten Unternehmen auf externe Unterstützung bei der Implementierung von KI-Anwendungen zurückgreifen, weil fehlendes Fachwissen (54,2%) sowie Mangel an Personal (54,2%) und Zeit (41,7%) vorherrscht, die als Hauptgründe gegen eine Eigenentwicklung angeführt werden. [27, 78]

Im Zusammenhang mit der vierten industriellen Revolution ist *Smart Factory* ein wichtiges Stichwort, gerade im Bereich von Fertigungsprozessen. Eine intelligente Fabrik kann als Fertigungssystem mit Produktionseinheiten definiert werden, die an ein cloudbasiertes, zentralisiertes System angeschlossen sind und interaktive Informationsaustauschfunktionen über das Internet und die Cloud bieten. Im Bereich audioüberwachter Systeme bedeutet das, dass zur Erkennung des Betriebsstatus einer jeden Produktionseinheit des gesamten Fertigungssystems ein zugehöriges CNN entwickelt wird, welches nur die Geräusche der eigenen Einheit klassifiziert und alle anderen Prozessgeräusche als Rauschen erkennt. Der Vorteil eines solchen Überwachungssystems ist, dass es selbst bei gleichzeitigem Betrieb mehrerer Einheiten möglich ist, eine Einheit und ihren Status zu identifizieren, Betriebsgeräusche zu erkennen und eine unbegrenzte Anzahl von Einheiten gleichzeitig zu überwachen. Kim, Lee et al. (2021) entwickelten so ein schallbasiertes System zur Überwachung des Betriebszustands mehrerer Einheiten innerhalb eines gesamten Fertigungsprozesses und testeten es in einer Werkstatt und einer Fabrik. Die Weiterentwicklung zu einem System, welches wiederum den Fortschritt des gesamten Fertigungsprozesses analysiert, stellen sie in Aussicht. [66] In Bezug auf die vorliegenden Ergebnisse könnten neben einem Regenerationsprozess auch andere Anlagen eines Gießereibetriebes akustisch überwacht und in ein zusammenhängendes System überführt werden. Zu beachten ist, dass eine der größten Schwachstelle solcher Systeme laut Henze, Gorishti et al. (2019) der Mangel an Trenddaten ist. Es ist entscheidend, das System in einer produktiven Umgebung und nicht nur in einer Demo-Umgebung zu testen bzw. zu kalibrieren, um nicht nur künstlich erzeugte Fehlerzustände zu berücksichtigen, sondern auch reale [54].



# Literatur

- [1] *Abreißen, loslassen*. Diogenes Kalender. Donnerstag, 3. Februar. 2022.
- [2] E. Alpaydin. *Maschinelles Lernen*. 2. Auflage. De Gruyter Studium. Berlin und Boston: De Gruyter Oldenbourg, 2019. ISBN: 9783110617887. URL: [http://www.degruyter.com/search?f\\_0=isbnissn&q\\_0=9783110617887&searchTitles=true](http://www.degruyter.com/search?f_0=isbnissn&q_0=9783110617887&searchTitles=true).
- [3] *Bedienungsanleitung: Schwingungstransmitter*. URL: <https://autosen.com/de/Motion-Control/Schwingungssensoren/Schwingungstransmitter-ISO10816-AV001> (besucht am 26.04.2021).
- [4] O. Beucher. *Signale und Systeme: Theorie, Simulation, Anwendung*. 3. Aufl. 2019. Berlin, Heidelberg: Springer, 2019. ISBN: 978-3-662-58043-1. URL: <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1620088>.
- [5] W. Bian, J. Wang, B. Zhuang, J. Yang, S. Wang und J. Xiao. „Audio-Based Music Classification with DenseNet and Data Augmentation“. *PRICAI 2019: Trends in Artificial Intelligence*. Hrsg. von A. C. Nayak und A. Sharma. Bd. 11672. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, S. 56–65. ISBN: 978-3-030-29893-7. DOI: 10.1007/978-3-030-29894-4\_5.
- [6] L. Bifano, A. Fischerauer und G. Fischerauer. „Investigation of complex permittivity spectra of foundry sands“. *tm - Technisches Messen* 87.5 (2020). S. 372–380. ISSN: 0171-8096. DOI: 10.1515/teme-2019-0121.
- [7] L. Bifano, A. Fischerauer, G. Fischerauer und A. Liedtke. „Ansätze für eine Prozessüberwachung bei der Sandregeneration“. *Giesserei* 107 5/6 (2020). S. 23–28.
- [8] I. Bindernagel, Hrsg. *Formstoffe und Formverfahren in der Gießereitechnik*. Bd. 12. VDG-Taschenbuch. Düsseldorf: Giesserei-Verlag GmbH, 1983. ISBN: 3-87260-057-5.
- [9] J. R. Brown. *Foseco Ferrous Foundryman's Handbook*. Oxford: Butterworth-Heinemann, 2000. ISBN: 0 7506 4284 X.
- [10] E. Brunhuber, Hrsg. *Giesserei-Lexikon*. 15. Aufl. Berlin: Schiele & Schön, 1991. ISBN: 3794905164.
- [11] S. L. Brunton und J. N. Kutz. *Data-Driven Science and Engineering*. Cambridge University Press, 2019. ISBN: 9781108380690. DOI: 10.1017/9781108380690.
- [12] A. Bührig-Polaczek, W. Michaeli und G. Spur. *Handbuch Urformen*. 1. Aufl. Handbuch der Fertigungstechnik. München: Hanser, 2013. ISBN: 978-3-446-42035-9. DOI: 10.3139/9783446434066.
- [13] A. Bührig-Polaczek u. a. „Digitalisierung in der Gießereitechnik, Teil 1: Herausforderungen und Chancen“. *Giesserei* 108 6 (2021). S. 37–43.

- [14] H. Chaves. *Vorlesung Messtechnik: Kreuzkorrelation*. Hrsg. von T. Freiberg. URL: <https://tu-freiberg.de/sites/default/files/media/institut-fuer-mechanik-und-fluiddynamik-15832/Lehre/lehrveranstaltungen/fluid/MT/kreuzkorrelation.pdf> (besucht am 30.05.2022).
- [15] S. Chen und D. Prof. Hartmann. „KI-Anwenderumgebung für die Analyse komplexer Fertigungsprozesse“. *Giesserei Special 1* (2021). S. 58–59.
- [16] F. Chollet. *Deep Learning mit Python und Keras: Das Praxis-Handbuch; vom Entwickler der Keras-Bibliothek*. 1. Auflage. Frechen: mitp, 2018. ISBN: 9783958458383. URL: <http://www.mitp.de/IT-WEB/Programmierung/Deep-Learning-mit-Python-und-Keras.html?listtype=search&searchparam=python>.
- [17] S.F. Crone. *Neuronale Netze zur Prognose und Disposition im Handel: Zugl.: Hamburg, Univ., Fakultät für Wirtschafts- und Sozialwissenschaften, Diss., 2008*. 1. Aufl. Gabler Research Betriebswirtschaftliche Forschung zur Unternehmensführung. Wiesbaden: Gabler, 2010. ISBN: 978-3-8349-1174-2.
- [18] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan und Q. Le V. *AutoAugment: Learning Augmentation Policies from Data*. URL: <http://arxiv.org/pdf/1805.09501v3>.
- [19] M. Curilem, J.P. Canario, L. Franco und R. A. Rios. „Using CNN To Classify Spectrograms of Seismic Events From Llaima Volcano (Chile)“. *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2018, S. 1–8. ISBN: 978-1-5090-6014-6. DOI: 10.1109/IJCNN.2018.8489285.
- [20] P. Czapla und R. Danko. *The State of Art of the Mechanical Reclamation of Used Foundry Sands*. 2013. URL: [http://www.lodlewnictwo.agh.edu.pl/inne/konferencje/konferencja\\_2013/data/sesja\\_studencka/Czapla.pdf](http://www.lodlewnictwo.agh.edu.pl/inne/konferencje/konferencja_2013/data/sesja_studencka/Czapla.pdf) (besucht am 31.05.2022).
- [21] J. S. Dańko und R. Dańko. „Mechanical Reclamation Treatment of Used Silica Foundry Sand at Extra-low and Ambient Temperatures“. *Material Science* 4 (2005). S. 344–347.
- [22] J. S. Dańko, M. Holtzer und R. Dańko. „Factors influencing selection of effective reclamation techniques and assessment methods of the reclaimed material quality“. *Archives of Foundry Engineering* 7 (2007). S. 29–32. ISSN: 1897-3310.
- [23] J. S. Dańko, M. Holtzer und R. Dańko. „Problems of scientific and development research concerning the reclamation of used foundry sands“. *Archives of Foundry Engineering* 10 (2010). S. 29–34. ISSN: 1897-3310.
- [24] N. Davis und K. Suresh. „Environmental Sound Classification Using Deep Convolutional Neural Networks and Data Augmentation“. *2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*. IEEE, 2018, S. 41–45. ISBN: 978-1-5386-7336-2. DOI: 10.1109/RAICS.2018.8635051.
- [25] J. Dettelbacher und W. Schlüter. „Transparenter Betrieb durch digitales Produktionsabbild“. *Giesserei* 107 11 (2020). S. 34–37.
- [26] E. Dietrich und S. Conrad. *Statistische Verfahren zur Maschinen- und Prozessqualifikation*. 8., aktualisierte Auflage. Hanser eLibrary. München: Carl Hanser Verlag, 2022. ISBN: 978-3-446-46447-6. DOI: 10.3139/9783446465046. URL: <https://www.hanser-elibrary.com/doi/book/10.3139/9783446465046>.

- [27] T. Disselhoff u. a. *Potenziale der schwachen künstlichen Intelligenz für die betriebliche Ressourceneffizienz*. Hrsg. von VDI Zentrum Ressourceneffizienz GmbH. 2021. URL: [https://www.ressource-deutschland.de/fileadmin/user\\_upload/downloads/studien/VDI-ZRE\\_Studie\\_KI-betriebliche-Ressourceneffizienz\\_Web\\_bf.pdf](https://www.ressource-deutschland.de/fileadmin/user_upload/downloads/studien/VDI-ZRE_Studie_KI-betriebliche-Ressourceneffizienz_Web_bf.pdf) (besucht am 31.05.2022).
- [28] Dr. rer. nat. Hans-Joachim Blankenburg, Prof. Dr. rer. nat. habil. Hans Jürgen Rösler. *Quarzrohstoffe*. 1. Aufl. Leipzig: VEB Deutscher Verlag für Grundstoffindustrie, 1978.
- [29] H. Elsner. *Quarzrohstoffe in Deutschland*. Stand: Januar 2016. Hannover: Bundesanstalt für Geowissenschaften und Rohstoffe, Januar 2016. ISBN: 9783943566314.
- [30] W. Ernst. „Schalt doch mal ab...“. *Giesserei* 107 10 (2020). S. 50–56.
- [31] F. Ettetmeyer u. a. „Entkernverhalten anorganisch gebundener Sandkerne: Charakterisierung, Klassifizierung, Modellierung“. *Giesserei Special* 1 (2021). S. 18–23.
- [32] L. Fahrmeir, C. Heumann, R. Künstler, I. Pigeot und G. Tutz. *Statistik: Der Weg zur Datenanalyse*. 8., überarbeitete und ergänzte Auflage. Springer-Lehrbuch. Berlin und Heidelberg: Springer Spektrum, 2016. ISBN: 978-3-662-50371-3. DOI: 10.1007/978-3-662-50372-0.
- [33] U. Fayyad, G. Piatetsky-Shapiro und P. Smyth. „From Data Mining to Knowledge Discovery in Database“. *AI Magazine* 17 3 (1996). S. 37–54. URL: <https://doi.org/10.1609/aimag.v17i3.1230> (besucht am 28.12.2021).
- [34] F.B. Fitch. „Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics, vol. 5 (1943), pp. 115–133“. *Journal of Symbolic Logic* 9.2 (1944). S. 49–50. ISSN: 0022-4812. DOI: 10.2307/2268029.
- [35] E. Flemming und W. Tilch. *Formstoffe und Formverfahren: Mit 118 Tabellen*. 1. Aufl. Leipzig: Dt. Verl. für Grundstoffindustrie, 1993. ISBN: 3342005319.
- [36] S. Franke, Hrsg. *Giesserei Lexikon*. Ausgabe 2019, 20. Auflage. Berlin: Schiele & Schön, 2019. ISBN: 9783794909162.
- [37] S. Franke, Hrsg. *Taschenbuch der Gießerei-Praxis 2020*. 1. Auflage. Berlin: Schiele & Schön, 2019. ISBN: 9783794909544.
- [38] T. Frey und M. Bossert. *Signal- und Systemtheorie: Mit 117 Abbildungen, 26 Tabellen, 64 Aufgaben mit Lösungen und 84 Beispielen*. 2., korrigierte Auflage. Aus dem Programm Nachrichtentechnik. Wiesbaden: Vieweg + Teubner, 2008. ISBN: 978-3-8351-0249-1.
- [39] A. H. Fritz, Hrsg. *Fertigungstechnik*. 12., neu bearbeitete und ergänzte Auflage. Springer-Lehrbuch. Berlin und Heidelberg: Springer Vieweg, 2018. ISBN: 978-3-662-56534-6. DOI: 10.1007/978-3-662-46535-3. URL: <http://www.springer.com/>.
- [40] J. Frochte. *Maschinelles Lernen: Grundlagen und Algorithmen in Python*. 3., überarbeitete und erweiterte Auflage. Plus.Hanser-Fachbuch. München: Carl Hanser Verlag, 2021. ISBN: 978-3-446-46144-4.
- [41] A. P. Fros und M. Fehlbier. „Anwendungsbeispiele zur Prozessoptimierung: Fachgebiet Gießereitechnik, Uni Kassel - KI driven Automated Optical Inspection, Use Case Ultradünnwand-Druckgießen“. *Giesserei* 108 7 (2021). S. 34–42.

- [42] H. Gemming. *Prozesssicherheit in der Formstoffaufbereitung mit Hilfe der Fuzzy-Logik*. 2003. URL: [http://webdoc.sub.gwdg.de/ebook/serien/aa/Freiberger\\_Diss\\_Online/185.pdf](http://webdoc.sub.gwdg.de/ebook/serien/aa/Freiberger_Diss_Online/185.pdf) (besucht am 27.06.2022).
- [43] T. Giannakopoulos und A. Pikrakis. *Introduction to audio analysis: A Matlab approach*. 1. ed. MATLAB examples. Amsterdam: Academic Press an imprint of Elsevier, 2014. ISBN: 978-0-08-099388-1. URL: <http://www.loc.gov/catdir/enhancements/fy1606/2015473171-d.html>.
- [44] L. Göllmann u. a. *Mathematik für Ingenieure: Verstehen - Rechnen - Anwenden: Band 1: Vorkurs, Analysis in einer Variablen, Lineare Algebra, Statistik*. Berlin: Springer Vieweg, 2017. ISBN: 978-3-662-53866-1. DOI: 10.1007/978-3-662-53867-8.
- [45] I. Goodfellow. *Deep learning: Das umfassende Handbuch: Grundlagen, aktuelle Verfahren und Algorithmen, neue Forschungsansätze*. 1. Auflage. Frechen: mitp Verlags GmbH & Co, 2018. ISBN: 978-3-95845-702-7.
- [46] G. Görz, U. Schmid und T. Braun, Hrsg. *Handbuch der Künstlichen Intelligenz*. 6. Auflage. Berlin und Boston: De Gruyter Oldenbourg, 2021. ISBN: 9783110659849. DOI: 10.1515/9783110659948.
- [47] J. Götze und M. Göbbels. *Einführung in die Angewandte Mineralogie*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. ISBN: 978-3-662-50264-8. DOI: 10.1007/978-3-662-50265-5.
- [48] P. Grunert. *Machine Learning und Neuronale Netze: Der verständliche Einstieg mit Python*. Landslut: BMU Verlag, 2021. ISBN: 978-3-96645-073-7.
- [49] D. Guicking. *Schwingungen*. Wiesbaden: Springer Fachmedien, 2016. ISBN: 978-3-658-14135-6. DOI: 10.1007/978-3-658-14136-3.
- [50] A. Guzhov, F. Raue, J. Hees und A. Dengel. *AudioCLIP: Extending CLIP to Image, Text and Audio*. URL: <http://arxiv.org/pdf/2106.13043v1>.
- [51] V. Haanappel, T. Linke, E. Schulte und M. Jendrock. „Regenerierung anorganisch gebundener Sandsysteme“. *Giesserei 109 2* (2022). S. 44–49.
- [52] T. Hastie, R. Tibshirani und J. H. Friedman. *The elements of statistical learning: Data mining, inference, and prediction*. Second edition, corrected at 12th printing 2017. Springer series in statistics. New York, NY: Springer, 2017. ISBN: 9780387848570. DOI: 10.1007/b94608.
- [53] K. He, X. Zhang, S. Ren und J. Sun. *Deep Residual Learning for Image Recognition*. URL: <http://arxiv.org/pdf/1512.03385v1>.
- [54] D. Henze, K. Gorishti, B. Bruegge und J.-P. Simen. „AudioForesight: A Process Model for Audio Predictive Maintenance in Industrial Environments“. *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. IEEE, 2019, S. 352–357. ISBN: 978-1-7281-4550-1. DOI: 10.1109/ICMLA.2019.00066.
- [55] S. Hershey u. a. *CNN Architectures for Large-Scale Audio Classification*. URL: <http://arxiv.org/pdf/1609.09430v2>.
- [56] K. Hildebrand, M. Gebauer und M. Mielke, Hrsg. *Daten- und Informationsqualität: Die Grundlage der Digitalisierung*. 5., erweiterte und aktualisierte Auflage. Wiesbaden und Heidelberg: Springer Vieweg, 2021. ISBN: 9783658309909.

- [57] T. Hope, Y. S. Resheff und I. Lieder. *Einführung in TensorFlow: Deep-Learning-Systeme programmieren, trainieren, skalieren und deployen*. 1. Auflage. Heidelberg: O'Reilly, 2018. ISBN: 978-3-96009-074-8.
- [58] A. G. Howard u. a. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. URL: <http://arxiv.org/pdf/1704.04861v1>.
- [59] T. Inoue u. a. „Shuffling and Mixing Data Augmentation for Environmental Sound Classification“. *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*. New York University, 25.10.2019 - 26.10.2019, S. 109–113. ISBN: 978-0-578-59596-2. DOI: 10.33682/wgyb-bt40.
- [60] S. Ioffe und C. Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. URL: <http://arxiv.org/pdf/1502.03167v3>.
- [61] I. Izdebska-Szanda, A. Baliński und M. Angrecki. „Evaluation of Reclamability of Molding Sands with New Inorganic Binders“. *Archives of Foundry Engineering* 12.2 (2012). ISSN: 1897-3310. DOI: 10.2478/v10266-012-0032-y.
- [62] D. S. Johnson und S. Grollmisch. „Techniques Improving the Robustness of Deep Learning Models for Industrial Sound Analysis“. *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, S. 81–85. ISBN: 978-9-0827-9705-3. DOI: 10.23919/Eusipco47968.2020.9287327.
- [63] B. Y. Kang, K. H. Kim, D. W. Park und M. S. Lee. „Effect of Regenerated-Foundry Sand on the Mechanical Properties of Core“. *Materials Science Forum* 922 (2018). S. 149–154. DOI: 10.4028/www.scientific.net/MSF.922.149.
- [64] S. Kayal und B. K. Chakrabarti. „Reclamation and Utilisation of Foundry Waste Sand“. *High Temperature Materials and Processes* 27.1 (2008). ISSN: 0334-6455. DOI: 10.1515/HTMP.2008.27.1.51.
- [65] U. Kiencke, M. Schwarz und T. Weickert. *Signalverarbeitung: Zeit-Frequenz-Analyse und Schätzverfahren*. München: Oldenbourg Verlag, 2008. ISBN: 978-3-486-58668-8.
- [66] J. Kim, H. Lee, S. Jeong und S.-H. Ahn. „Sound-based remote real-time multi-device operational monitoring system using a Convolutional Neural Network (CNN)“. *Journal of Manufacturing Systems* 58 (2021). S. 431–441. ISSN: 02786125. DOI: 10.1016/j.jmsy.2020.12.020.
- [67] K. H. Kim, M. A. Bae, M. S. Lee, H.-S. Park und J. H. Baek. „Regeneration of used sand with sodium silicate binder by wet method and their core manufacturing“. *Journal of Material Cycles and Waste Management* 23.1 (2021). S. 121–129. ISSN: 1438-4957. DOI: 10.1007/s10163-020-01103-5.
- [68] E. Y. Ko, K. H. Kim, J. H. Baek, I. Hwang und M. S. Lee. „Wet regeneration of waste artificial sand used in sand casting using chemical solutions“. *Environmental Engineering Research* 26.6 (2021). ISSN: 1226-1025. DOI: 10.4491/eer.2020.421.
- [69] W. König. „Lernende Kernschieß-Maschine und intelligente Schmelzöfen“. *Giesserei* 107 11 (2020). S. 60–64.
- [70] D. Kreß und B. Kaufhold. *Signale und Systeme verstehen und vertiefen: Denken und Arbeiten im Zeit- und Frequenzbereich; mit 34 Übungsaufgaben*. 1. Aufl. Studium Informations- und Kommunikationstechnik. Wiesbaden: Vieweg + Teubner, 2010. ISBN: 978-3-8348-1019-9.

- [71] A. Krizhevsky, I. Sutskever und G. E. Hinton. „ImageNet classification with deep convolutional neural networks“. *Communications of the ACM* 60.6 (2017). S. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386.
- [72] T. Kuttner und A. Rohnen. *Praxis der Schwingungsmessung: Messtechnik und Schwingungsanalyse mit MATLAB®*. 2., überarbeitete und erweiterte Auflage. Wiesbaden: Springer Vieweg, 2019. ISBN: 978-3-658-25047-8.
- [73] U. Lämmel und J. Cleve. *Künstliche Intelligenz: Wissensverarbeitung - neuronale Netze*. 5., überarbeitete Auflage. München: Hanser, 2020. ISBN: 978-3-446-45914-4.
- [74] J. Lange und T. Lange. *Fourier-Transformation zur Signal- und Systembeschreibung: Kompakt, visuell, intuitiv verständlich*. essentials. Wiesbaden: Springer Fachmedien Wiesbaden, 2019. ISBN: 978-3-658-24849-9. DOI: 10.1007/978-3-658-24850-5.
- [75] Y. le Cun. *Generalization and Network Design Strategies: Technical Report CRG-TR-89-4*. 1989. URL: <http://yann.lecun.com/exdb/publis/pdf/lecun-89.pdf> (besucht am 07.02.2022).
- [76] J. Li, X. Niu, L. Yang, G. Lu und H. Qiu. „Optimierung der Herstellungsausrüstung für Sandkerne“. *Giesserei* 109 3 (2022). S. 26–30.
- [77] T. Li, M. Ogihara und G. Tzanetakis. *Music data mining*. Chapman & Hall / CRC data mining and knowledge discovery series. Boca Raton: CRC Press, 2012. ISBN: 978-1-4398-3555-5.
- [78] A. Liedtke. „Formsand 4.0: Formsandaufbereitung im Zeitalter von Industrie 4.0 oder ”Big-Data“ für den Anwendbar nutzbar machen“. *Giesserei* 104 7 (2017). S. 70–75. URL: [https://www.mic henfelder.com/wp-content/uploads/2017/09/100\\_Formsand\\_4\\_0\\_dt.pdf](https://www.mic henfelder.com/wp-content/uploads/2017/09/100_Formsand_4_0_dt.pdf).
- [79] J. Lunze. *Künstliche Intelligenz für Ingenieure: Methoden zur Lösung ingenieurtechnischer Probleme mit Hilfe von Regeln, logischen Formeln und Bayesnetzen*. 3., überarbeitete Auflage. De Gruyter Studium. Berlin und Boston: De Gruyter Oldenbourg, 2016. ISBN: 9783110448962.
- [80] G. Maguolo, M. Paci, L. Nanni und L. Bonan. „AudioGmenter: a MATLAB toolbox for audio data augmentation“. *Applied Computing and Informatics* (2021). ISSN: 2210-8327. DOI: 10.1108/ACI-03-2021-0064.
- [81] J. Mariscal-Harana, V. Alarcón, F. González, J. J. Calvente, F. J. Pérez-Grau, A. Viguria und A. Ollero. „Audio-Based Aircraft Detection System for Safe RPAS BVLOS Operations“. *Electronics* 9.12 (2020). S. 2076. DOI: 10.3390/electronics9122076.
- [82] A. M. Martin Kaiser. „Neuste Entwicklungen bei Sandformen und -kernen“. *Giesserei* 108 11 (2021). S. 70–73.
- [83] M. Mauch und S. Ewert. „The Audio Degradation Toolbox And Its Application To Robustness Evaluation“. *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*. 2013, S. 83–88. ISBN: ISBN 978-0-615-90065-0.
- [84] B. Meffert und O. Hochmuth. *Werkzeuge der Signalverarbeitung*. Humboldt-Universität zu Berlin, 2018. DOI: 10.18452/19595.
- [85] M. Meyer, Hrsg. *Signalverarbeitung: Analoge und digitale Signale, Systeme und Filter*. 8. Aufl. Wiesbaden: Springer Vieweg, 2017. ISBN: 978-3-658-18320-2.
- [86] M. Möser. *Technische Akustik*. 6., erw. und aktualisierte Aufl. VDI-/Buch. Berlin und Heidelberg: Springer, 2005. ISBN: 3-540-22510-2.

- [87] M. Möser, Hrsg. *Digitale Signalverarbeitung in der Messtechnik*. Fachwissen Technische Akustik. Berlin, Germany: Springer Vieweg, 2018. ISBN: 978-3-662-56612-1.
- [88] A. Müller. „Schrittweise in die grüne Gießerei“. *Giesserei* 109 4 (2022). S. 60–61.
- [89] A. C. Müller und S. Guido. *Einführung in Machine Learning mit Python: Praxiswissen Data Science*. 1. Auflage. Heidelberg: O'Reilly, 2017. ISBN: 978-3-96010-111-6.
- [90] M. Müller. *Fundamentals of Music Processing: Using Python and Jupyter Notebooks*. 2nd ed. 2021. Springer eBook Collection. Cham: Springer International Publishing und Imprint Springer, 2021. ISBN: 9783030698072. DOI: 10.1007/978-3-030-69808-9.
- [91] J. Niemine und R. P. Jung. „Innovatives Service-Konzept bringt Kosten- und Prozessoptimierung“. *Giesserei* 109 2 (2022). S. 32–39.
- [92] W. Nitsche und A. Brunn. *Strömungsmesstechnik*. 2., aktualisierte und bearb. Aufl. VDI-Buch. Berlin und Heidelberg: Springer, 2006. ISBN: 3-540-20990-5.
- [93] S. Ntalampiras, D. Kosmin und J. Sanchez. „Acoustic classification of individual cat vocalizations in evolving environments“. *2021 44th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2021, S. 254–258. ISBN: 978-1-6654-2933-7. DOI: 10.1109/TSP52935.2021.9522660.
- [94] Olympus. „Deep Learning optimiert die industrielle Bildanalyse“. *Giesserei* 108 11 (2021). S. 96.
- [95] K. Palanisamy, D. Singhanian und A. Yao. *Rethinking CNN Models for Audio Classification*. URL: <http://arxiv.org/pdf/2007.11154v2>.
- [96] D. S. Park, W. Chan, Y. Zhang, C.-C. Chiu, B. Zoph, E. D. Cubuk und Q. Le V. „SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition“ (2019). S. 2613–2617. DOI: 10.21437/Interspeech.2019-2680. URL: <http://arxiv.org/pdf/1904.08779v3>.
- [97] B. Pfister und T. Kaufmann. *Sprachverarbeitung*. Berlin, Heidelberg: Springer, 2017. ISBN: 978-3-662-52837-2. DOI: 10.1007/978-3-662-52838-9.
- [98] M. Plaue. *Data Science: Grundlagen, Statistik und maschinelles Lernen*. Lehrbuch. Berlin: Springer Spektrum, 2021. ISBN: 9783662634882.
- [99] R. O. Pohl, Hrsg. *Pohls Einführung in die Physik: Mechanik, Akustik und Wärmelehre: Mit 498 Abbildungen, 84 Aufgaben und 77 Videofilmen*. 21. Aufl. Berlin Heidelberg: Springer Spektrum, 2017. ISBN: 978-3-662-48662-7.
- [100] M. Polke. „Erforderliche Informationsstrukturen für die Qualitätssicherung“. *Chemie Ingenieur Technik* 65.7 (1993). S. 791–796. ISSN: 0009286X. DOI: 10.1002/cite.330650703.
- [101] H. Polzin. *Anorganische Binder zur Form- und Kernherstellung in der Gießerei: Teilw. zugl.: Freiberg, TU Bergakad., Habil.-Schr., 2012*. 1. Aufl. Berlin: Schiele & Schön, 2012. ISBN: 9783794908240.
- [102] F. Puente León und S. Bauer. *Praxis der Digitalen Signalverarbeitung*. KIT Scientific Publishing, 2017. DOI: 10.5445/KSP/1000067012.
- [103] F. Puente León und H. Jäkel. *Signale und Systeme*. 7., überarbeitete Auflage. De Gruyter Studium. Berlin und Boston: De Gruyter Oldenbourg, 2019. ISBN: 978-3-11-062631-5.
- [104] E. Radtke und M. Vogt. „Ein einfacher Fußabdruck und seine komplizierten Implikationen“. *Giesserei* 108 1 (2021). S. 17–18.

- [105] K. D. Rao und M. Swamy. *Digital Signal Processing: Theory and Practice*. Singapore: Springer, 2018. ISBN: 9789811080807. DOI: 10.1007/978-981-10-8081-4.
- [106] S. Raschka und V. Mirjalili. *Machine Learning mit Python und Keras, TensorFlow 2 und Scikit-Learn*. 3rd edition. Boston: mitp Verlag und Safari, 2021. ISBN: 9783747502150.
- [107] T. Rashid. *Neuronale Netze selbst programmieren: Ein verständlicher Einstieg mit Python*. 1. Auflage. Heidelberg: O'Reilly, 2017. ISBN: 978-3-96009-043-4.
- [108] M. Reichert und E. Schulte. „Neues Verfahren für anorganisch gebundene Gießereisande“. *Giesserei* 108 5 (2021). S. 57–58.
- [109] A. Risse. *Fertigungsverfahren der Mechatronik, Feinwerk- und Präzisionsgerätetechnik*. Wiesbaden: Vieweg+Teubner Verlag, 2012. ISBN: 978-3-8348-1519-4. DOI: 10.1007/978-3-8348-8312-4.
- [110] P. Ryan, S. Takafuji, C. Yang, N. Wilson und C. McBride. „Using Self-Supervised Learning of Bird-song for Downstream Industrial Audio Classification“. *ICML 2020 Workshop on Self-supervision in Audio and Speech*. 2020. URL: [https://openreview.net/forum?id=\\_P9LyJ5pMDb](https://openreview.net/forum?id=_P9LyJ5pMDb) (besucht am 05.07.2022).
- [111] J. Salamon und J.P. Bello. „Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification“. *IEEE Signal Processing Letters* 24.3 (2017). S. 279–283. ISSN: 1070-9908. DOI: 10.1109/LSP.2017.2657381.
- [112] J. Schlüter und T. Grill. *Exploring Data Augmentation For Improved Singing Voice Detection With Neural Networks*. 2015. DOI: 10.5281/zenodo.1417745.
- [113] H. Y. Shi, H. Liu und S. L. Guo. „Casting Recycling of Animal Protein-Based Binder Sand“. *Advanced Materials Research* 573-574 (2012). S. 1156–1160. DOI: 10.4028/www.scientific.net/AMR.573-574.1156.
- [114] K. Siebertz, D. van Bebber und T. Hochkirchen. *Statistische Versuchsplanung: Design of Experiments (DoE)*. 2. Auflage. VDI-Buch. Berlin und Heidelberg: Springer Vieweg, 2017. ISBN: 9783662557426. DOI: 10.1007/978-3-662-55743-3.
- [115] K. Simonyan und A. Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. URL: <http://arxiv.org/pdf/1409.1556v6>.
- [116] G. R. Sinambari und S. Sentpali. *Ingenieurakustik: Physikalische Grundlagen und Anwendungsbeispiele*. 5., völlig überarb. und erw. Aufl. Wiesbaden: Springer Vieweg, 2014. ISBN: 978-3-658-05071-9. DOI: 10.1007/978-3-658-05072-6.
- [117] E. Sprengel, M. Jaggi, Y. Kilcher und T. Hofmann. *Audio Based Bird Species Identification using Deep Learning Techniques*. 2016. URL: <http://ceur-ws.org/Vol-1609/16090547.pdf> (besucht am 29.06.2022).
- [118] M. Stachowicz und K. Granat. „Possibilities of Reclamation Microwave-Hardened Molding Sands with Water Glass“. *Archives of Metallurgy and Materials* 59.2 (2014). S. 757–760. DOI: 10.2478/amm-2014-0127.
- [119] M. Stachowicz und K. Granat. „Influence of wet activation of used inorganic binder on cyclically refreshed water glass moulding sands hardened by microwaves“. *China Foundry* 13.6 (2016). S. 427–432. ISSN: 1672-6421. DOI: 10.1007/s41230-016-5077-z.

- [120] J. Stange-Elbe. *Computer und Musik: Grundlagen, Technologien und Produktionsumgebungen der digitalen Musik*. De Gruyter Studium. Berlin: De Gruyter, 2015. ISBN: 978-3-486-59031-9.
- [121] M. Stieß. *Mechanische Verfahrenstechnik - Partikeltechnologie 1*. Berlin, Heidelberg: Springer, 2009. ISBN: 978-3-540-32551-2. DOI: 10.1007/978/3-540-32552-9.
- [122] T. C. Stocker und I. Steinke. *Statistik: Grundlagen und Methodik*. 2., korrigierte Auflage. Berlin und Boston: De Gruyter Oldenbourg, 2022. ISBN: 978-3-11-074409-5. DOI: 10.1515/9783110744194-202.
- [123] Q. Z. Sun, J. Jing, P. Q. Zhang, Z. K. Zhao und J. Wang. „Principle and Applications of Grinding Wheel Regenerator“. *Applied Mechanics and Materials* 121-126 (2011). S. 301–305. DOI: 10.4028/www.scientific.net/AMM.121-126.301.
- [124] Y. Sun, T. M. Maeda, C. Solis-Lemus, D. Pimentel-Alarcon und Z. Burivalova. *Classification of animal sounds in a hyperdiverse rainforest using Convolutional Neural Networks*. URL: <http://arxiv.org/pdf/2111.14971v1>.
- [125] N. Takahashi, M. Gygli, B. Pfister und L. van Gool. *Deep Convolutional Neural Networks and Data Augmentation for Acoustic Event Detection*. URL: <http://arxiv.org/pdf/1604.07160v2>.
- [126] D. Tomkins und E. Nichols. *DCASE 2019 CHALLENGE TASK 5: CNN+VGGISH: Technical Report*. 2019. URL: [https://dcase.community/documents/challenge2019/technical\\_reports/DCASE2019\\_Tompkins\\_85.pdf](https://dcase.community/documents/challenge2019/technical_reports/DCASE2019_Tompkins_85.pdf) (besucht am 30.06.2022).
- [127] B. T. Tong. „Transfer learning and its applications in image and audio recognition“. Bachelorarbeit. Duisburg: Universität Duisburg-Essen, 2021.
- [128] E. Tsalera, A. Papadakis und M. Samarakou. „Comparison of Pre-Trained CNNs for Audio Classification Using Transfer Learning“. *Journal of Sensor and Actuator Networks* 10.4 (2021). S. 72. DOI: 10.3390/jsan10040072.
- [129] VDG-Merkblatt P 27. *Prüfung von Formgrundstoffen: Bestimmung der granulometrischen Kennwerte*. Hrsg. von VDG - Verein Deutscher Gießereifachleute. 1999.
- [130] VDG-Merkblatt P 72. *Bindemittelprüfung: Prüfung von kalthärtenden, kunstharzgebundenen feuchten Formstoffen mit Härterzusatz*. Hrsg. von VDG - Verein Deutscher Gießereifachleute. 1999.
- [131] VDI. „Intelligent Manufacturing: Beste digitale Lösung der Fertigungsindustrie ausgezeichnet“. *VDI-Z* 4 (2022). S. 56–57.
- [132] H. Wang, Y. Zou und W. Wang. *SpecAugment++: A Hidden Space Data Augmentation Method for Acoustic Scene Classification*. URL: <http://arxiv.org/pdf/2103.16858v3>.
- [133] R. Wartala. *Praxiseinstieg Deep Learning: Mit Python, Caffe, TensorFlow und Spark eigene Deep-Learning-Anwendungen erstellen*. 1. Auflage. Heidelberg: O'Reilly, 2018. ISBN: 978-3-96009-054-0.
- [134] S. Weidman. *Deep Learning - Grundlagen und Implementierung: Neuronale Netze mit Python und PyTorch programmieren*. 1. Auflage. Heidelberg: O'Reilly, 2020. ISBN: 9783960091363.
- [135] M. Werner. *Digitale Signalverarbeitung mit MATLAB: Grundkurs mit 16 ausführlichen Versuchen*. 5., durchges. und aktualisierte Aufl. Studium Nachrichtentechnik. Wiesbaden: Vieweg + Teubner, 2012. ISBN: 978-3-8348-1473-9.

- [136] M. Werner. *Digitale Signalverarbeitung mit MATLAB®: Grundkurs mit 16 ausführlichen Versuchen*. 6., vollständig überarbeitete Auflage. Lehrbuch. Wiesbaden und Heidelberg: Springer Vieweg, 2019. ISBN: 9783658186463.
- [137] E. Westhoff. *On-line-Analyse von Formsandsystemen: Zugl.: Aachen, Techn. Hochsch., Diss., 1998*. Als Ms. gedr. Bd. 2. Forschung, Entwicklung, Ergebnisse / Gießerei-Institut. Aachen: Shaker, 1998. ISBN: 3826535820.
- [138] J. Xie, X. Aubert, X. Long, J. von Dijk, B. Arsenali, P. Fonseca und O. Sebastiaan. „Audio-based snore detection using deep neural networks“. *Computer Methods and Programs in Biomedicine* 200 (2020). URL: <https://doi.org/10.1016/j.cmpb.2020.105917> (besucht am 13.07.2021).
- [139] M. Yang, C. Peng und Z. Li. „An Audio-based Intelligent Fault Classification System for Belt Conveyor Rollers“. *2021 40th Chinese Control Conference (CCC)*. IEEE, 2021, S. 4647–4652. ISBN: 978-9-8815-6380-4. DOI: 10.23919/CCC52363.2021.9550449.
- [140] A. Zach, B. Kleinert, U. Nitsch und G. Wolf. „Anorganische Kernbinderreste in bentonitischen Formstoffsystemen“. *Giesserei 107* 10 (2020). S. 34–40.
- [141] A. Zai und B. Brown. *Einstieg in Deep Reinforcement Learning: KI-Agenten mit Python und PyTorch programmieren*. München: Hanser, 2020. ISBN: 9783446459007.
- [142] B. Zhang, J. Leitner und S. Thronton. *Audio Recognition using Mel Spectrograms and Convolution Neural Networks*. 2019. URL: [http://noiselab.ucsd.edu/ECE228\\_2019/Reports/Report38.pdf](http://noiselab.ucsd.edu/ECE228_2019/Reports/Report38.pdf) (besucht am 29.06.2022).