

KI-gestützte MR-Klassifizierung von Zellen und zellulärer
Differenzierung

AI-assisted MR classification of cells and cellular
differentiation



Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der
Graduate School of Life Sciences, Julius-Maximilians-Universität Würzburg

Sektion Biomedizin

vorgelegt von

PHILIPP FEY

aus Lohr am Main

Würzburg 2023



eingereicht am:

Bürostempel

Mitglieder des Prüfungskomitees

Vorsitzender: Prof. Dr. Christian Janzen

1. Betreuer: Prof. Dr. Jan Hansmann

2. Betreuer: Prof. Dr. Peter Jakob

3. Betreuer: Dr. Karl-Heinz Hiller

Tag des Promotionskolloquiums:

Doktorurkunden ausgehändigt am:

*Ausdauer wird früher oder später belohnt –
meistens aber später.*

WILHELM BUSCH

Inhaltsverzeichnis

Abkürzungen	9
Glossar	11
Symbolverzeichnis	16
1 Zusammenfassung	17
1.1 Deutsche Zusammenfassung	17
1.2 English summary	19
2 Einleitung	20
2.1 Motivation	20
2.2 Stand der Technik	21
2.3 Zielsetzung	23
2.4 Methodische Grundlagen	24
2.4.1 Magnetresonanz (MR)	25
2.4.2 Aufbau und Funktion eines künstlichen neuronalen Netzwerks (KNN)	30
2.4.3 Zelluläre Entwicklung	33
3 Material und Methoden	38
3.1 Material	38
3.1.1 Chemikalien	38
3.1.2 Geräte	38
3.1.3 Medien, Puffer, Zusätze	39
3.1.4 Software	40
3.1.5 Verbrauchsmaterialien	40
3.1.6 Zellen	41
3.2 Methoden	42
3.2.1 Zellkultur	42

3.2.2	Validierung der MR-Parameter und Methoden	47
3.2.3	Aufbau und Funktion der Robotikplattform	49
3.2.4	Vorbereitung der Zellproben und Viabilitätsnachweis	54
3.2.5	Postprozessierung der Messdaten	56
3.2.6	Aufbau und Training der Support Vector Machine (SVM) . . .	58
3.2.7	Aufbau und Training des künstlichen neuronalen Netzwerks (KNN)	59
4	Ergebnisse	66
4.1	Validierung der MR-Parameter und -Methodik	66
4.1.1	Pulswinkelbestimmung	66
4.1.2	Phantommessungen	66
4.1.3	Erste Zell- und Medienmessungen	69
4.1.4	Zusammenfassung	73
4.2	Etablierung einer Robotikplattform	74
4.2.1	Aufbau der Robotikplattform	74
4.2.2	Validierung der Proben temperierung	76
4.2.3	Zusammenfassung	77
4.3	MR-Messung von Zelllinien	78
4.3.1	Messungen von drei Zelllinien	78
4.3.2	Erweiterung der Messungen auf zehn Zelllinien	79
4.3.3	Zellalterung und Viabilität	82
4.3.4	Zusammenfassung	88
4.4	MR-Messung von mesenchymalen Stromazellen	89
4.5	KI-gestützte Auswertung der Daten	90
4.5.1	Nutzung einer Support Vector Machine für die Klassifizierung	91
4.5.2	Etablieren einer Augmentierungsroutine	97
4.5.3	Etablieren und Testen eines künstlichen neuronalen Netzwerks zur Klassifikation von MSCs	98
4.5.4	Etablieren und Testen eines künstlichen neuronalen Netzwerks zur Klassifikation von Zelllinien	101
4.5.5	Zusammenfassung	113
4.6	Leistung der SVM im Vergleich zum KNN	113
5	Diskussion	115
5.1	Zuverlässigkeit der Messdaten	115
5.2	Die Sensitivität des Messsystems	116

5.3	Automatisierung eines Benchtop-Prozesses	117
5.4	Temperierung und Viabilität der Zellkultur	118
5.5	Ursprung des Zellsignals und seiner Unterschiede	119
5.6	KI für die Dateninterpretation	122
5.7	Benutzerdefinierte Datenaugmentierung	125
6	Ausblick	126
7	Danksagung	128
8	Publikationsliste	CXXX
9	Lebenslauf	CXXXI
9.1	Schulbildung	CXXXI
9.2	Universität	CXXXI
9.3	Beruflicher Werdegang	CXXXI
10	Appendix	CXXXII
10.1	Affidavit	CXXXII
10.2	Eidesstattliche Erklärung	CXXXII
	Abbildungsverzeichnis	CXXXIII
	Tabellenverzeichnis	CXXXVI
	Literaturverzeichnis	CXXXVII

Abkürzungen

2D	Zweidimensional
ADZP	Agarose-Dotagraf [®] -Zellphantome
°C	Grad Celsius
CI	Konfidenzintervall (engl. <i>confidence intervall</i>)
CPMG	Carr-Purcell-Meiboom-Gill-Sequenz
GUI	Grafische Benutzeroberfläche (engl. <i>graphical user interface</i>)
h	Stunde(n)
Hz	Hertz
K	Kelvin
KNN	künstliches neuronales Netzwerk
LNA	rauscharmer Signalverstärker (engl. <i>low noise amplifier</i>)
MaxTE	gibt maximalen Wert für TE an (in Sekunden)
MaxTI	gibt maximalen Wert für TI an (in Sekunden)
min	Minute(n)
mind.	mindestens
MinTI	minimaler Wert für TI (in Sekunden)
ml	Milliliter
ML	maschinelles Lernen (engl. <i>machine learning</i>)
MR	Magnetresonanz
MRT	Magnetresonanztomographie
NA	Anzahl an Mittlungen (engl. <i>number of averages</i>)
ND	Anzahl an Dummy-Messungen (engl. <i>number of dummy measurements</i>)
NR	Anzahl an Wiederholungen (engl. <i>number of repetitions</i>)
PBS	phosphatgepufferte Saline (engl. <i>phosphate buffered saline</i>)
P/S	Penicillin-Streptomycin
PCA	<i>principal component analysis</i> (dt. Hauptkomponentenanalyse)
PntsTI	Anzahl der Messpunkte in TI (engl. <i>points TI</i>)

RT	Raumtemperatur i. d. R. 21–22 °C
SLA	Stereolithographie
SNR	Signal-zu-Rausch-Verhältnis (engl. <i>signal-to-noise ratio</i>)
SVM	<i>Support Vector Machine</i>
T	Tesla
TE	Echozeit (in Sekunden)
TI	Inversionszeit (in Sekunden)
TR	Repetitionszeit (in Sekunden)
W	Watt
xg	relative Zentrifugalkraft in g (= Erdbeschleunigung)
μl	Mikroliter

Glossar

- Adam** Der Adam-Optimierungsalgorithmus ist ein adaptiver Gradientenabstiegsalgorithmus für neuronale Netzwerke, der sowohl Momentum- als auch Adaptionsterme verwendet, um die Lernrate anzupassen und das Training effizienter zu gestalten.
- AdamW** AdamW ist ein Optimierungsalgorithmus für neuronale Netzwerke, der auf dem Adam-Algorithmus basiert und zusätzlich eine Gewichtsverfallkomponente einführt, um das Auftreten von Overfitting zu verringern und die Stabilität des Trainings zu verbessern.
- Arduino** Eine quellenoffene Software und Hardwareplattform. Die Hardware besteht aus einem Mikrocontroller auf einer einfachen Platine mit Ein- und Ausgängen. Die Programmierung erfolgt in einer Sprache, welche C- bzw. C++ ähnelt.
- Amplifizierung** Amplifizierung im Kontext von Datenaugmentierung für KNN bezieht sich auf die Erhöhung der vorhandenen Datenmenge durch Anwendung verschiedener Transformationen und Modifikationen, um die Variation und Vielfalt der Trainingsdaten zu erhöhen und die Robustheit und Generalisierungsfähigkeit des neuronalen Netzwerks zu verbessern.
- ANOVA** ANOVA (engl. *Analysis of Variance*) ist eine statistische Methode zur Untersuchung von Unterschieden zwischen Gruppen und zur Bestimmung, ob diese Unterschiede auf zufällige Variation oder signifikante Effekte zurückzuführen sind.

Augmentierung	Datenaugmentierung für KNN bezieht sich auf den Prozess der künstlichen Erweiterung des Trainingsdatensatzes durch Anwendung verschiedener Transformationen wie Rotation, Skalierung und Spiegelung, um die Leistung, Robustheit und Generalisierungsfähigkeit des neuronalen Netzwerks zu verbessern.
Centroid	Der mathematisch berechnete Schwerpunkt eines Datenpeaks. Die entsprechende MATLAB [®] -Funktion ist in der Lage, diesen gewichtet d. h. nicht nur nach seiner Ausdehnung, sondern auch nach der Verteilung der Datenwerte zu berechnen.
Convolutional-Layer	Ein Convolutional-Layer in einem KNN ist eine Schicht, die Faltungsberechnungen auf den Eingabemerkmalsskizzen durchführt, um lokale Merkmale zu extrahieren und die räumliche Struktur der Daten beizubehalten.
Dense-Layer	Ein dense-Layer in einem KNN ist eine Schicht, in der jeder Eingabeneuron mit jedem Ausgabeneuron verbunden ist, um eine vollständig verbundene Struktur zu schaffen und komplexe Beziehungen zwischen den Merkmalen zu erlernen.
Drop-Out-Layer	Ein drop-out-Layer in einem KNN ist eine Schicht, die zufällig ausgewählte Neuronen während des Trainings deaktiviert, um Overfitting zu reduzieren und die Robustheit des Netzwerks zu verbessern.
Flatten-Schicht	Die Flatten-Schicht in einem KNN hat den Zweck, die räumliche Struktur der Daten aufzuheben und die Eingabe in einen eindimensionalen Vektor umzuwandeln, um sie mit den nachfolgenden vollständig verbundenen Schichten zu verbinden.
Fused Deposition Modeling (FDM)	FDM 3D-Druck (Fused Deposition Modeling) ist ein Additivfertigerungsverfahren, bei dem ein thermoplastisches Filament schichtweise aufgetragen und geschmolzen wird, um ein dreidimensionales Objekt zu erstellen.

Hyperparameter	Hyperparameter für KNN sind einstellbare Parameter, die vor dem Training festgelegt werden, um das Verhalten und die Leistung des neuronalen Netzwerks zu steuern, wie beispielsweise die Lernrate, die Anzahl der Schichten oder die Anzahl der Neuronen pro Schicht.
Instancenormalisierung	Ein KNN nutzt die Instanzennormalisierung, um die Eingabedaten in jedem Trainingsschritt zu normalisieren und so die Konvergenzgeschwindigkeit und Stabilität des Netzwerks zu verbessern.
Konfusionsmatrix	Die Konfusionsmatrix (engl. <i>confusion matrix</i>) ist eine Darstellung der Leistung eines Klassifikationsmodells, die die Anzahl der richtig und falsch zugeordneten Instanzen für jede Klasse in einer tabellarischen Form zeigt.
LNA	LNA (engl. <i>Low Noise Amplifier</i>) ist eine elektronische Verstärkerschaltung, die schwache Signale mit geringem Rauschen verstärkt, um die Empfindlichkeit und Leistung von Kommunikations- und Sensorgeräten zu verbessern.
MATLAB®	MATLAB® ist ein in den Natur- und Ingenieurwissenschaften weit verbreitetes, kommerzielles Softwarepaket des Unternehmens MathWorks. Es dient der mathematischen Problemlösung und der grafischen Darstellung von Ergebnissen.
machine learning	Beim maschinellen Lernen (dt. <i>Maschinelles-Lernen</i>) nutzen entsprechende Algorithmen statistische Methoden, um Gesetzmäßigkeiten und Muster in Daten zu erkennen. Durch diese statistischen Modelle ist es anschließend den Algorithmen möglich, Vorhersagen - basierend auf den erlernten Mustern - zu neuen, bisher ungesehenen Daten zu treffen.
Maxpooling-Layer	Ein Maxpooling-Layer in einem KNN ist eine Schicht, die die räumliche Dimension der Eingabemerkmale reduziert, indem der maximale Wert in jedem Bereich extrahiert wird, um die Dimensionalität zu verringern und die wichtigen Merkmale beizubehalten.

Normalisierungsschicht	Die Normalisierungsschicht für KNNs ist eine Schicht, die dazu dient, die Eingabedaten zu normalisieren und die Skalierung und Verteilung der Daten zu regulieren, um das Training zu beschleunigen und die Konvergenz des Netzwerks zu verbessern.
openMatlab	MATLAB [®] Framework, welches von der Pure Devices GmbH entwickelt wurde, um deren Geräte zu programmieren und anzusteuern.
PD	Pure Devices GmbH Hersteller des verwendeten Tabletop-MRTs inkl. des verwendeten Matlab [®] Frameworks.
Primärzellen	Zellen, welche direkt aus einem Gewebe i. d. R. Patientengewebe, z. B. durch eine Biopsie isoliert wurden.
shield	Unter einem <i>shield</i> versteht man eine Erweiterungsplatine z.B. für einen Arduino, um diesen mit zusätzlichen Funktionalitäten zu erweitern.
SNR	SNR (engl. <i>Signal-to-Noise Ratio</i>) ist ein Maß für das Verhältnis zwischen der Stärke eines Signals und dem Hintergrundrauschen und wird verwendet, um die Signalqualität zu bewerten.
Stereolithographie (SLA)	Bei der Stereolithographie handelt es sich um eine spezielle Methode des 3D-Druckens. Dabei wird das Werkstück Schicht für Schicht aus einem Becken mit Photopolymer generiert. Ein Laser belichtet und härtet damit jede Schicht an den entsprechenden Stellen. Die Basisplatte des Druckers entfernt sich mit jeder Schicht immer weiter vom Polymerbad, wodurch das Bauteil senkrecht aus dem Bad gezogen werden kann.
Stock	Die allgemeine Bezeichnung für einen vorbereiteten Vorrat einer fertigen Chemikalie, Medium oder Lösung. Dieser wird immer dann hergestellt, wenn die fertige Chemikalie/Lösung nicht bei jeder Verwendung neu zusammen gemischt werden soll.

Support Vector Machine (SVM)	Eine Methode des überwachten maschinellen Lernens. Hierbei werden zwei Datengruppen mittels eines Vektors voneinander getrennt. Dieser Vektor kann entweder linearer oder nicht-linearer Natur sein. Eine Vorhersage wird dann aufgrund des trainierten Vektors erzeugt und unge-sehene Daten einer der beiden Gruppen zugeordnet.
Tensorboard	Tensorboard ist ein Visualisierungstool von TensorFlow, das es ermöglicht, den Trainingsfortschritt, die Modellarchitektur und andere Metriken eines neuronalen Netzwerks interaktiv und anschaulich darzustellen.
TensorFlow	TensorFlow ist ein Open-Source-Framework für maschinelles Lernen und neuronale Netzwerke, das eine effiziente Entwicklung, Training und Bereitstellung von Modellen ermöglicht.
User-bias	Der User-Bias in einem KNN bezieht sich auf die individuelle Präferenz oder Neigung eines Benutzers, die in das Modell einfließt und sich auf die Vorhersagen oder Empfehlungen des Modells auswirken kann.
waschen	In der Zellkultur versteht man unter „waschen“ oder „spülen“ die Zugabe eines Waschpuffers, wie z. B. PBS oder Zellkulturmedium. Dies wird i. d. R. nach einer kurzen Einwirkzeit wieder abgesaugt.
Zelllinie	Immortale Zellen, welche entweder spontan mutiert oder genetisch modifiziert wurden. Zelllinien zeigen keine Alterserscheinungen und können folglich über viele Jahre hinweg kultiviert werden.
Zellkulturbedingungen	Wenn nicht anderweitig angegeben, entsprechen die Zellkulturbedingungen stets 37 °C bei 5% CO ₂ und >95 % relativer Luftfeuchte.

Symbolverzeichnis

Symbol	Beschreibung	Wert/[Einheit]
B_0	Stärke des externen Magnetfeldes	[T]
e	Eulersche Zahl	≈ 2.718
γ	gyromagnetisches Verhältnis	$42.576e6 \cdot \pi [\frac{MHz}{T}]$
ω_0	Präzessionsfrequenz	[Hz]
π	Kreiszahl	≈ 3.1415
T_1	longitudinale Relaxationszeit	[s]
T_2	transversale Relaxationszeit	[s]

1 Zusammenfassung

1.1 Deutsche Zusammenfassung

Für die Verwendung von zellbasierten Therapeutika ist vor allem die korrekt Identifikation sowohl vom Ausgangsmaterial wie auch dem produzierten Material von zentraler Wichtigkeit. In dieser Arbeit wurde eine Methodik entwickelt, welche eine nicht-invasive Klassifizierung von Zellen und zellulärer Entwicklung aufgrund ihrer zweidimensionalen Magnetresonanz-Korrelationsspektren ermöglichte.

Hierzu wurde ein mobiler MR-Scanner mit einer Feldstärke von 0.5 T und einem Isozentrum von 1 cm³ verwendet. Aufgrund der kompakten und leichten Bauweise war es möglich, das System in normalen Zellkulturelaboren zu verwenden. Von den Proben wurde ein zweidimensionales T₁/T₂ -Korrelationsspektrum aufgenommen, anhand dessen die Zellen klassifiziert werden sollten. Mithilfe von Agarose-Dotagraf[®] -Zell-Phantomen konnte die Stabilität und Reproduzierbarkeit des Messsystems und der verwendeten Sequenz validiert werden.

Aufgrund der unter Umständen recht langen Messzeiten der MR-Technologie war auch die Handhabung und Kultur der Zellproben während des Messprozesses von großer Bedeutung. Um hierfür den Durchsatz an Proben zu erhöhen, wurde eine kostengünstige und ebenfalls mobile Robotikanlage entwickelt. Diese basierte auf dem kommerziell erhältlichen Roboterarm *Braccio*, welcher durch einen Arduino Mega Mikrocontroller gesteuert wurde. Mit bis zu 24 Proben pro Tag konnte durch die Automatisierung der Durchsatz an Proben um den Faktor 3–4 gesteigert werden. Durch den entwickelten Prozess war es möglich, eine umfangreiche Datenbank bestehend aus 362 unabhängigen Messungen (biologische Replikate) aufzubauen. Die Datenbank enthielt Messungen von zehn unterschiedlichen Zelllinien. Zusätzlich wurden T₁/T₂ -Korrelationsspektren von mesenchymalen Stromazellen (MSCs) vor und nach deren Differenzierung zu Adipocyten aufgenommen, um ihre zelluläre Entwicklung nicht-invasiv charakterisieren zu können.

Die aufgenommenen Daten wurden mithilfe einer geeigneten *Support Vector Machine* wie auch angepassten künstlichen neuronalen Netzwerken klassifiziert. Mithilfe dieser Methoden konnten die Zelllinien und MSCs anhand ihrer aufgenommenen Korrelationsspektren mit einer Genauigkeit von bis zu 98% klassifiziert werden.

Diese hohe Treffsicherheit legte den Schluss nahe, dass die Kombination aus nicht-invasiver, zweidimensionaler T_1/T_2 -MR-Relaxometrie und der Verwendung von geeigneten Methoden des *machine learning* und der künstlichen Intelligenz eine effiziente Methodik für die nicht-invasive Klassifizierung von Zellen sowie zellulärer Entwicklung darstellt.

1.2 English summary

For the use of cell-based therapeutics, the correct identification of both the starting material and the produced material is of central importance. In this work, a methodology was developed that allowed non-invasive classification of cells and cellular development based on their two-dimensional magnetic resonance correlation spectra.

For this purpose, a mobile MR scanner with a field strength of 0.5 T and an isocenter of 1 cm³ was used. Due to its compact and lightweight design, it was possible to use the system in normal cell culture laboratories. A two-dimensional T₁/T₂ -correlation spectrum was recorded from the samples, which was used to classify the cells. Agarose-Dotagraf[®] cell phantoms were used to validate the stability and reproducibility of the measurement system and the sequence used.

Due to the possibly quite long measurement times of the MR technology, the handling and culture of the cell samples during the measurement process was also of great importance. To increase the throughput of samples for this purpose, a low-cost and also mobile robotic system was developed. This was based on the commercially available *Braccio* robotic arm, which was controlled by an Arduino Mega microcontroller. With up to 24 samples per day, the automation increased the throughput of samples by a factor of 3 - 4. Through the developed process it was possible to build an extensive database consisting of 362 independent measurements (biological replicates). The database contained measurements from ten different cell lines. In addition, T₁/T₂ -correlation spectra of mesenchymal stromal cells (MSCs) before and after their differentiation into adipocytes were recorded to characterize their cellular development non-invasively. The recorded data were classified using an appropriate *Support Vector Machine* as well as adapted artificial neural networks. Using these methods, the cell lines and MSCs could be classified based on their recorded correlation spectra with an accuracy of up to 98%. This high accuracy suggested that the combination of non-invasive, two-dimensional T₁/T₂ -MR relaxometry and the use of appropriate machine learning and artificial intelligence methods is an efficient methodology for the non-invasive classification of cells as well as cellular development.

2 Einleitung

2.1 Motivation

Die korrekte Identifizierung und Charakterisierung von Zellen stellen für die klinische und präklinische Forschung ein zentrale Fragestellungen dar. Häufig wird thematisiert, wie Zellen auf eine entsprechende Behandlung reagieren oder welche Medikation zu einer Reduktion der Tumorzellen führt. Im Bereich der regenerativen Medizin ist v. a. eine qualitative Aussage über den Entwicklungsstand von zellbasierten Transplantaten oder Arzneimitteln für neuartige Therapien (engl. *advanced therapy medicinal products* – *ATMPs*) von großem Interesse [1]. Für manche dieser Therapeutika muss sichergestellt werden, dass die Entwicklung der Zellen korrekt erfolgt ist, bevor sie für die Anwendung am Patienten freigegeben werden können [2]. Da die Wirkung der biologischen Therapeutika auf der Funktionsfähigkeit der applizierten Zellen basiert [3], ist die korrekte Identifizierung der gewünschten Zellen und ggf. deren Entwicklungs- bzw. Differenzierungszustandes von großer Bedeutung [4]. Aus der korrekten Zuordnung von zellulären Entwicklungen ergibt sich auch die Bedeutung der korrekten Identifizierung von unterschiedlichen Zell- und Gewebetypen. Beispielsweise in der klinischen Analyse von entnommenen Gewebebiopsien stellt eine korrekte Erkennung der Biopsie eine zentrale Rolle dar [5, 6].

Damit diese Anforderungen erfüllt werden können, werden in der vorgelegten Arbeit mehrere Ansprüche an eine entsprechende Methode gestellt: Sie sollte in der Lage sein, die gesamte Probe analysieren zu können, was häufig als *whole-mount* bezeichnet wird. Der Messprozess sollte unter sterilen Zellkulturbedingungen ablaufen, sodass die Viabilität der Zellen erhalten bleibt und keine Kontaminationen in die Kultur eingebracht werden. Des Weiteren sollte die Methode in der Lage sein, mehrere Proben hintereinander in relativ kurzen Abständen vermessen zu können, um einen hohen Probendurchsatz zu gewährleisten. Sie sollte mobil sein, damit sie in einem Routinelabor durchführbar ist, und sie sollte keine Hilfsstoffe – auch als *labels* bezeichnet – wie Farbstoffe oder Kontrastmittel benötigen, damit das Therapeutikum nicht modifiziert wird.

Die in dieser Arbeit vorgestellte Methodik ist keineswegs die erste oder einzige Möglichkeit, mit welcher Zellen zugeordnet werden können. Andere Möglichkeiten weisen

deutliche Vor- bzw. Nachteile im Vergleich zur vorgestellten Methodik auf. Nur unter der Berücksichtigung dieser wichtigen Unterschiede lässt sich fundiert beurteilen, welche Methodik für die erläuterte Fragestellung die beste Lösung liefert.

Um die Belastung für den oder die Patient*in so gering wie möglich zu halten, sollte die geforderte Methodik kein entnommenes Material verschwenden. Nur so kann gewährleistet werden, dass auch wirklich 100 % des generierten Transplantates oder Therapeutikums dem oder der Patient*in zu Gute kommen. Daraus schließt sich auch die Anforderung, dass die Untersuchung unter sterilen Bedingungen stattfinden muss. Nur so kann das untersuchte Therapeutikum anschließend auch für die Anwendung im Patienten bzw. der Patientin freigegeben werden. Für diese Freigabe ist es auch wichtig, dass die Untersuchung nicht auf der Zuführung von Fremdmaterial (z. B. Farbstoffen oder anderen Markern) basiert. Nur wenn das untersuchte Gewebe unverändert bleibt, kann davon ausgegangen werden, dass es später im Patienten oder der Patientin zu keinen unerwünschten Nebeneffekten kommt. Wenn mit lebenden Zellen gearbeitet wird, ist es natürlich auch von großer Bedeutung, dass diese durch den Messprozess nicht irreversibel beschädigt werden d. h. , dass keine DNA-schädigende Strahlung oder Chemikalien Verwendung finden. Andernfalls bestünde die Möglichkeit, dass die transplantierten Zellen anschließend entarten und tumoröses Gewebe entsteht. Die letzte Anforderung an die Methodik ist die Reproduzierbarkeit und der Durchsatz. Nur wenn die Methodik objektive und zuverlässig Ergebnisse erzielt und in einer angemessenen Zeit durchgeführt werden kann, eignet sie sich auch für die Anwendung am Patienten oder der Patientin.

2.2 Stand der Technik

Im klinischen Kontext sind v. a. histologische Färbungen die anerkannten Referenzmethoden [7, 8]. Hierfür wird eine Biopsie aus dem entsprechenden Gewebe entnommen und histologisch aufbereitet. Dieser Prozess beinhaltet das chemische Fixieren der Probe, um den enzymatischen Verdau des Gewebes zu verhindern, gefolgt von einer Einbettung in Harz oder Wachs [9]. Diese Methoden folgen i. d. R. standardisierten Arbeitsanweisungen (engl. *standardized operating procedures - SOP*). So wird gewährleistet, dass sie valide, reproduzierbare und vergleichbare Ergebnisse liefern. Jedoch obliegt die Interpretation der Ergebnisse dem Beobachter. Folglich ist das Ergebnis nicht objektiv und eine mögliche Diagnose kann variieren, je nachdem wer das

Ergebnis evaluiert [10, 11]. Außerdem basiert eine derartige Diagnose nur auf einem kleinen Ausschnitt der Gesamtheit und der Betrachter muss sich demnach darauf verlassen, dass die vorliegende Probe das gesamte Gewebe repräsentiert. Auch wird das Gewebe oder die Biopsie bei der Färbung irreversibel zerstört und steht somit einem weiteren Verwendungszweck nicht mehr zur Verfügung.

Als mehr oder weniger nicht-invasive Alternative hat sich im Bereich der präklinischen Zellkultur in den letzten Jahren u. a. die Raman-Spektroskopie als eine mögliche Methodik für diese Problemstellung herausgestellt. So konnte beispielsweise bereits 2008 in einer Studie gezeigt werden, dass die Raman-Spektroskopie kombiniert mit einem fluoreszenten Farbstoff genutzt werden kann, um unterschiedliche Bakterienspezies zu unterscheiden [12]. Durch die gute Auflösung der Raman-Spektroskopie kann diese als ‚label freie‘-Bildgebungsmethode verwendet werden, da es möglich ist, intrazelluläre Strukturen spektroskopisch nachzuweisen und sichtbar zu machen [13, 14]. Neuronale Netzwerke für die Klassifizierung oder Bilderkennung halfen dabei, die Raman-Aufnahmen zu beschleunigen und zu verbessern [15–18]. Es wurde gezeigt, dass die Raman-Spektroskopie in der Lage ist einzelne Zellen in einer Population korrekt zu identifizieren [14, 18–22]. All diese Aspekte machen Raman zu einem mächtigen Werkzeug bei der Identifizierung von Zellen.

Jedoch ist die Raman-Spektroskopie gerade aufgrund der beschriebenen Eigenschaften in ihrer Anwendung limitiert. Da es sich um eine mikroskopische Methode handelt, wird sie hauptsächlich für die Betrachtung und Analyse kleinerer Ausschnitte aus einer Probe bzw. einzelner Zellen verwendet. Auch ist Raman – wie die meisten mikroskopischen Methoden – in seiner Tiefenauflösung auf wenige μm beschränkt [23].

Neben chemisch-morphologischen Analysen von Zellen und Geweben liegt besonders die genetische Analyse der zu untersuchenden Zellpopulation nahe. Hierfür hat v. a. die Methode des single cell RNA sequencing (scRNA-seq) an Popularität gewonnen [24]. Mithilfe dieser Methode ist es möglich, genaue Informationen über die genetische Zusammensetzung einer Zellpopulation zu erhalten. Hierfür werden die exprimierten Gene bzw. deren RNA auf Einzelzellebene betrachtet und daraus entsprechende Karten erstellt [25, 26]. Diese können dann unter Zuhilfenahme von Algorithmen des maschinellen Lernens gruppiert und somit ähnliche Gruppen oder sogar spezielle Zelltypen identifiziert werden [27–29]. Dementsprechend erfüllt die scRNA-seq-Methode bereits die Anforderung, größere Populationen mit einer hohen Genauigkeit untersuchen zu können. Neue Studien konnte mRNA sogar in den Sekretionen von Hautzellen nachweisen, was evtl. die Erkennung und Diagnose von

Hautkrankheiten beschleunigen könnte [30]. Jedoch hat scRNA-seq im Großteil der Anwendungen den Nachteil, dass die Zellen bzw. das Gewebe durch die Analyse zerstört bzw. unbrauchbar werden [31].

Besser hierfür geeignet ist die Magnetresonanz (MR). Ein Vorteil der MR gegenüber dem zuvor erwähnten Raman ist v. a. die Möglichkeit, größere Proben von bis zu 70 cm Durchmesser (z. B. menschlicher Körper) vermessen zu können [32, 33]. Gegenüber scRNA-seq bietet sie den Vorteil, die Proben zerstörungsfrei und lebendig vermessen zu können. Mithilfe von quantifizierbaren MR-Methoden wie der MR-Relaxometrie [34–36] können genaue Daten für unterschiedliche Gewebetypen erhoben werden, um diese zu charakterisieren [37, 38]. Die Anwendung der MR-Relaxometrie reicht von der Untersuchung von Lebensmitteln wie Obst, Fleisch und Milchprodukten [39–44] über die Charakterisierung von Biopolymeren wie Proteinen in Wasser [45] bis hin zur zerstörungsfreien Untersuchung von Materialien [46]. Des Weiteren zeigten Studien die erfolgreiche Anwendung der Relaxometrie bei der Untersuchung von menschlichen Organen wie Leber, Niere oder Gelenkknorpel [47], Skelettmuskulatur [48] sowie der Differenzierung von Muskelvorläuferzellen [49].

Trotz der zahlreichen durchgeführten Studien zur Identifizierung von Geweben mittels MR-Methoden wird bei genauerer Betrachtung ein deutlicher Nachteil sichtbar: Im Vergleich zu anderen Methoden bietet die MR eine vergleichsweise schlechte Ortsauflösung. Während die Auflösung für klinische Zwecke i. d. R. ausreichend ist, stellt dies für die Laborumgebung - bei welcher häufig wenig biologisches Material und somit nur kleine Proben zur Verfügung stehen - eine große Herausforderung dar. Ansonsten erfüllt die MR alle der oben genannten Kriterien für eine passende Methodik: Sie bietet beispielsweise die Möglichkeit der whole-mount-Messung in einem sterilen Umfeld. Mit entsprechenden Messsequenzen ist es nicht notwendig, Kontrastmittel oder andere Stoffe für die Markierung der Zellen zu nutzen (‘label-freie’ Messung). Je nach Sequenz dauert eine Messung nur einige Minuten und es ist möglich, zahlreiche Proben hintereinander zu analysieren.

2.3 Zielsetzung

Ziel der vorgelegten Arbeit sollte es sein zu zeigen, dass die MR-Relaxometrie genutzt werden kann, um in einer sterilen Umgebung Zellen in einer nicht-invasiven Inprozesskontrolle zu charakterisieren. Die Arbeit lässt sich in drei Teile gliedern:

erstens die Etablierung und Validierung der Messmethode, zweitens die Etablierung einer Datenbank von Messungen basierend auf Zelllinien und differenzierenden Stammzellen und drittens die Etablierung einer KI-Methode zur Klassifizierung der aufgenommenen Daten.

Die Arbeit sollte in der Umgebung eines normalen Routinelabors erfolgen. Somit war ein großer klinischer MR-Scanner mit Helium-Kühlung für die Anwendung nicht praktikabel, sondern es sollte ein mobiler MR-Scanner genutzt werden. In Zusammenarbeit mit dem Hersteller PureDevices GmbH sollte ein entsprechender MR-Scanner für die Anwendung in der Zellkultur optimiert werden, welcher mit einem Fassungsvermögen von 1 cm^3 groß genug war, um entsprechende Zellkulturproben zu vermessen. Auch wenn der Scanner die Möglichkeit geboten hätte, die Probe orts aufgelöst zu vermessen, wurde entschieden, die Probe in ihrer Gesamtheit zu analysieren, um anschließend eine Aussage über die gesamte Zellpopulation zuzulassen. Um die lange Messdauer zu kompensieren, sollte ein automatisierter Probenwechsler etabliert werden, welcher den Durchsatz an messbaren Proben pro Tag erhöhte. Die bereits etablierte Messsequenz sollte auf den konkreten Anwendungsfall hin optimiert und mithilfe von Phantommessungen validiert werden.

Abschließend war zu untersuchen, unter welchen Bedingungen und in welchem Umfang die erhobenen Daten die notwendigen Informationen lieferten, um diverse Zellen und zelluläre Entwicklungsstadien unterscheiden zu können. Auf Basis der Messdaten sollte eine Datenbank aufgebaut werden, welche der Zuordnung der Proben dienete. Neben Messungen mit Zelllinien, wurde auch die Durchführung von Messungen mit sich differenzierenden mesenchymalen Stammzellen angestrebt. Für die Analyse sollte untersucht werden, welche Methoden des maschinellen Lernens für die Anwendung geeignet wären. Diese sollten anschließend genutzt werden, um geeignete Proben ihrem Zelltyp bzw. ihrem Entwicklungsstadium zuzuordnen zu können.

2.4 Methodische Grundlagen

Um die genauen Zusammenhänge der drei verwendeten Methoden besser verstehen zu können, werden in den nachfolgenden drei Sektionen theoretische Grundlagen zum Thema der Magnetresonanz, der künstlichen neuronalen Netzwerke und der zellulären Entwicklung erläutert.

2.4.1 Magnetresonanz (MR)

Die Magnetresonanztomographie oder auch Kernspintomographie stellt zusammen mit der Computer Tomographie (CT) zentrale Werkzeuge in der klinischen Diagnostik dar.

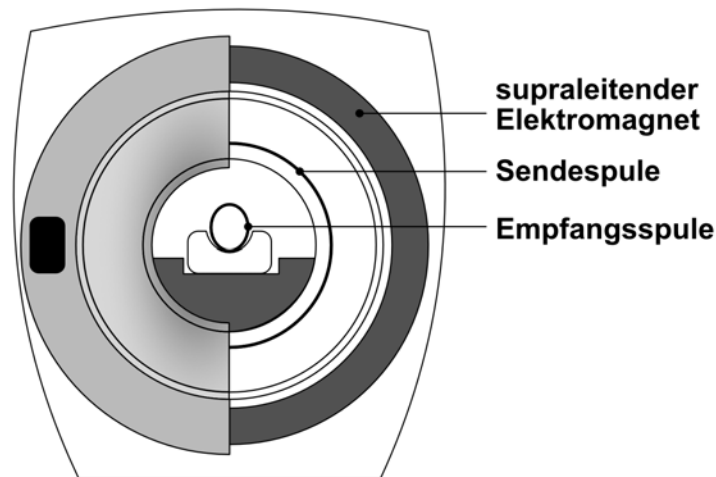


Abbildung 1 – Schematische Darstellung eines MRT: Ein konventioneller, klinischer MRT-Scanner besteht aus einem supraleitenden Elektromagneten, welcher i. d. R. Feldstärken von 1.5 bzw. 3 T aufweist. Der Patient wird mittig in den Scanner gefahren und für den Anwendungsfall optimierte Empfangsspulen an den entsprechenden Körperstellen angebracht. Die Sendespule des Gerätes ist typischerweise innerhalb des Scanners verbaut und von außen nicht sichtbar. Abbildung adaptiert von [50].

Im Gegensatz zur CT verwendet die MRT jedoch keine Röntgen- oder andere ionisierende Strahlung, sondern basiert auf Magnetfeldern. Herzstück eines MRT-Scanners stellt ein Elektromagnet dar. Dieser ist i. d. R. aus einem supraleitenden Material gebaut, welches durch flüssiges Helium auf unter 4,2 K (-269 °C) gekühlt wird. Durch die supraleitenden Eigenschaften des Materials weist es bei kalten Temperaturen keinen elektrischen Widerstand auf, sodass ein konstantes elektromagnetisches Feld mit hoher Feldstärke entsteht, sobald Strom durch das Material fließt. MRT-Scanner in klinischer Anwendung haben klassischerweise eine Feldstärke von 1.5 T bzw. 3 T. Zusätzlich zum Magneten sind noch Sende- Empfangs- und Gradientenspulen im MRT-Scanner verbaut. Die Sendespulen dienen dazu, das zu untersuchende Gewebe durch einen hochfrequenten Radiowellenpuls anzuregen. Das entstandene Signal wird wiederum durch die Empfangsspule aufgenommen und über mehrere Zwischenschritte in ein Bild umgewandelt. Die Gradientenspulen dienen der Ortskodierung. Anhand der aufgeprägten Magnetfeldgradienten ist es später möglich, Informationen über den Ursprung des Signals ermitteln zu können. Aus diesen Angaben kann dann am Computer ein digitales Bild der untersuchten Probe bzw. des Patienten erzeugt werden. So genannte Ganzkörperscanner tragen ihren Namen deshalb, weil der

Magnet eine Bohrung mit einem Durchmesser von etwa 60–70 cm hat, sodass sie groß genug sind, um mit ihnen einen erwachsenen Menschen komplett untersuchen zu können. Die entsprechenden Sende- und Empfangsspulen sind direkt im Scanner verbaut. Zusätzlich können weitere Spulen angeschlossen werden, da die Qualität der Aufnahme besser wird – d. h. eine höhere räumliche Auflösung und geringeres SNR hat –, wenn die Spule so an den zu untersuchenden Körperbereich angepasst ist, dass sie möglichst vollständig vom Untersuchungsobjekt ausgefüllt wird [51]. Beispielsweise wird bei Kopfuntersuchungen eine, an die Kopfform angepasste Spule verwendet, welche ein besseres Signal für diesen Anwendungsfall liefern.

Die MR-Technologie wird nicht nur in den klassischen klinischen MRT-Systemen zum Einsatz gebracht, sondern es gibt auch zahlreiche Entwicklungsprojekte zu mobilen MR-Systemen. Hierbei ist die Definition von ‚mobil‘ je nach Anwendungsfall jedoch unterschiedlich. Manche der Systeme sind konventionelle Ganzkörper-Systeme, welche in einem LKW-Container verbaut sind, sodass ihr Standort leicht verändert werden kann [52]. Andere Systeme sind deutlich kompakter und nutzen Elektromagneten, um die Mobilität zu erhöhen. Das System von Hyperfine wurde beispielsweise durch die US-Gesundheitsbehörde FDA für die Anwendung am Menschen zugelassen und ermöglicht es, mobile MR-Aufnahmen des Kopfes zu tätigen [53]. Die meisten anderen mobilen Systeme sind jedoch nicht für die Anwendung am Menschen zugelassen, sondern ausschließlich für akademische Zwecke gedacht. Sie bieten oft nur ein geringes Messvolumen, ermöglichen jedoch durch die Verwendung von OpenSource-Technologien eine gute Adaption an die gegebene Fragestellung [54–56].

Das Funktionsprinzip des MRT basiert auf dem (kern)physikalischen Effekt der Kernspinresonanz (engl. *nuclear magnetic resonance* - NMR). Aufgrund des Eigendrehimpulses (engl. *spin*) der Protonen und ihrer positiven Ladung erzeugen sie lokale Magnetfelder. Außerdem rotieren die Protonen zusätzlich auf einer gekippten Achse in einer Art ‚Kegelfigur‘. Diese Bewegung wird auch als Präzession bezeichnet (Abbildung 2 a). Sobald eine Probe (z. B. ein Patient) in ein MRT eingefahren wird, richten sich die Protonen aufgrund ihrer magnetischen Eigenschaften parallel oder antiparallel zum starken Magnetfeld des MRTs aus (Abbildung 2 b). Auch wenn sich nicht alle Protonen entlang des externen Magnetfeldes (B_0) ausrichten, sondern teilweise eine antiparallele Orientierung aufgrund der thermischen Energie einnehmen, entsteht ein magnetischer Vektor in der Probe parallel (longitudinal) zum externen Magnetfeld (Abbildung 2c-1 roter Pfeil). Hierbei korreliert die Präzessionsfrequenz der einzelnen Protonen (ω_0) direkt mit der Stärke des externen Magnetfeldes (B_0).

Außerdem ist sie direkt von einer für jede Teilchenart charakteristischen Konstanten, dem sog. gyromagnetischen Verhältnis (γ), abhängig. Diese Beziehungen werden in der Larmor-Gleichung (Gleichung (1)) ausgedrückt, einer der zentralen Gleichungen der MR-Physik. Daher wird die Präzessionsfrequenz auch als Larmorfrequenz bezeichnet:

$$\omega_0 = \gamma \cdot B_0 \quad (1)$$

Neben dem externen Magnetfeld erzeugt auch die chemische Umgebung der Protonen ein Magnetfeld und beeinflusst somit deren Frequenz.

Sobald Energie in Form eines hochfrequenten Radiowellenpulses (HF-Puls) mit der entsprechenden Larmorfrequenz in die Probe eingestrahlt wird, kommt es dazu, dass manche der Protonen in die energetisch höhere (antiparallele) Orientierung wech-

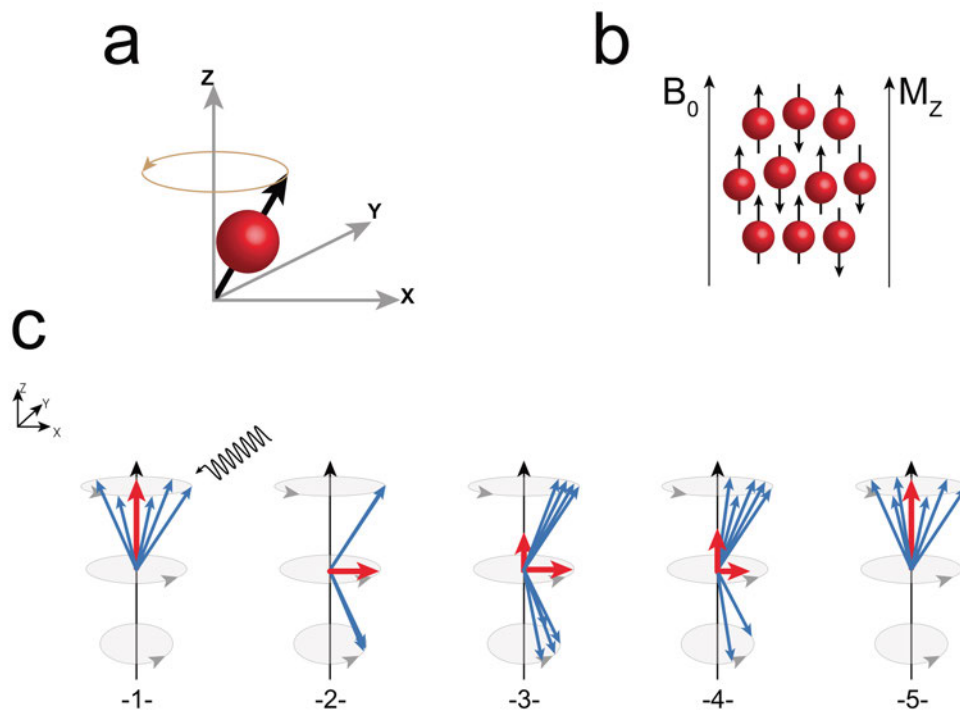


Abbildung 2 – Schematische Darstellung der physikalischen Grundlagen der MRT: Da jedes Proton eine positive Ladung trägt und um die eigene Achse rotiert, erzeugt es ein lokales Magnetfeld und kann als Stabmagnet verstanden werden. Außerdem ,taumelt‘ das Proton in einem bestimmten Winkel um eine Achse, was als Präzession bezeichnet wird (a). Die Frequenz der Präzession (ω_0) ist direkt von der Stärke des umgebenden Magnetfeldes abhängig. Diese Abhängigkeit ist in einer der zentralen Gleichungen der MR-Physik – als Larmor-Gleichung – zusammengefasst, weshalb die Präzessionsfrequenz auch als Larmorfrequenz bezeichnet wird. Wird ein Patient in ein MRT eingebracht, richten sich Protonen entlang des starken externen Magnetfeldes (B_0) aus (b). Durch diese Ausrichtung entsteht ein magnetischer Vektor, parallel (longitudinal) zum magnetischen Feld (M_Z). Wird nun Energie in Form einer gepulsten hochfrequenten Radiowelle (HF-Puls) eingestrahlt (c-1), führt dies einerseits dazu, dass manche der Protonen in die antiparallele Orientierung wechseln, andererseits synchronisiert sich die Präzessionsfrequenz der Protonen. Hierdurch ,kippt‘ der longitudinale Vektor z. B. um 90° in die transversale Ebene (c-2). Wird nun der HF-Puls abgeschaltet, kehren die Protonen langsam in den ursprünglichen, parallelen Zustand zurück (T_1) und dephasieren (T_2) (c-3/4). Hierdurch baut sich der transversale Vektor langsam ab (c-4), das System ,relaxiert‘. Daher werden T_1 und T_2 auch als Relaxationszeiten bezeichnet.

seln. Des Weiteren synchronisiert sich die Präzession der Protonen, wodurch der magnetische Vektor um z. B. 90° ‚kippt‘ (Abbildung 2 c-2). Aus diesem Grund wird ein entsprechender HF-Puls auch als 90° -Puls bezeichnet. Wenn ein Puls die doppelte Menge an Energie liefert, wird er als 180° -Puls bezeichnet, da er dazu führt, dass der longitudinale Ausgangsvektor um 180° gekippt wird. Der neu erzeugte, um 90° gekippte Vektor wird auch als transversaler Vektor bezeichnet. Diese Transversalmagnetisierung induziert eine Spannung in der im MR-Scanner verbauten Empfangsspule, welche als elektrisches Signal detektiert wird. Nach Abschalten des Pulses kehren die Protonen nach und nach wieder in den Ausgangszustand zurück, d. h. sie ‚relaxieren‘. Dies bedeutet einerseits, dass sie wieder ihre parallele Ausrichtung im Magnetfeld annehmen und außerdem abhängig von ihrer chemischen Umgebung mit unterschiedlichen Frequenzen rotieren, d. h. dephasieren. Hierdurch reduziert sich über eine bestimmte Zeit die transversale Magnetisierung und die longitudinale Magnetisierung baut sich graduell wieder auf (Abbildung 2 c-3/4). Schlussendlich ist das System vollkommen relaxiert und die transversale Magnetisierung ist vollständig dephasiert (Abbildung 2 c-5).

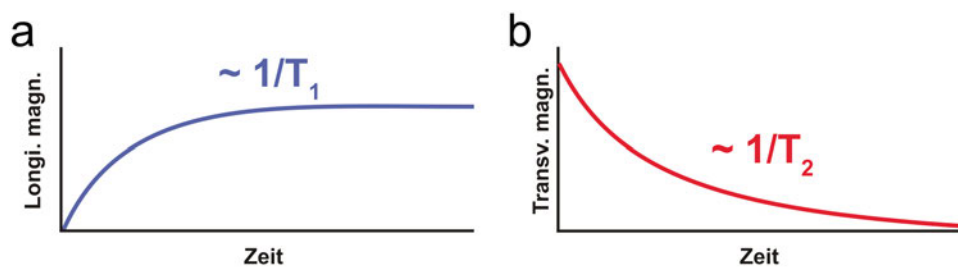


Abbildung 3 – Schematische Darstellung der T_1 - und T_2 -Relaxation: Sobald der HF-Puls abgeschaltet wurde, baut sich der erzeugte transversale Vektor langsam ab (b) und die longitudinale Magnetisierung baut sich wieder auf (a). Diese Verläufe werden durch die Relaxationszeitkonstanten T_1 und T_2 beschrieben. Die Werte für T_1 und T_2 stellen zwei zentrale Charakteristika in der MR-Physik dar.

Diese Zeitverläufe (Abbildung 3) werden durch die Relaxationszeitkonstanten T_1 und T_2 beschrieben. T_1 gibt dabei an, wie lange es dauert, bis die longitudinale Magnetisierung wieder auf $\frac{1}{e}$ des Ausgangswertes angestiegen ist, während T_2 beschreibt, wie lange es braucht, bis die transversale Magnetisierung auf $\frac{1}{e}$ abgefallen ist. Diese beiden Relaxationskonstanten T_1 und T_2 stellen zwei zentrale Merkmale des MR-Signals einer Probe dar. Wenn mittels entsprechender Gradientensysteme eine Ortskodierung des Signals und somit Bilder erzeugt werden, können anhand der T_1 - bzw. T_2 -Eigenschaften unterschiedliche Gewebe voneinander unterschieden werden. Beispielsweise hat Wasser im Gegensatz zu Fett einen längeren T_1 -Wert.

Wie bereits zuvor beschrieben, stellen HF-Pulse das zentrale Element eines

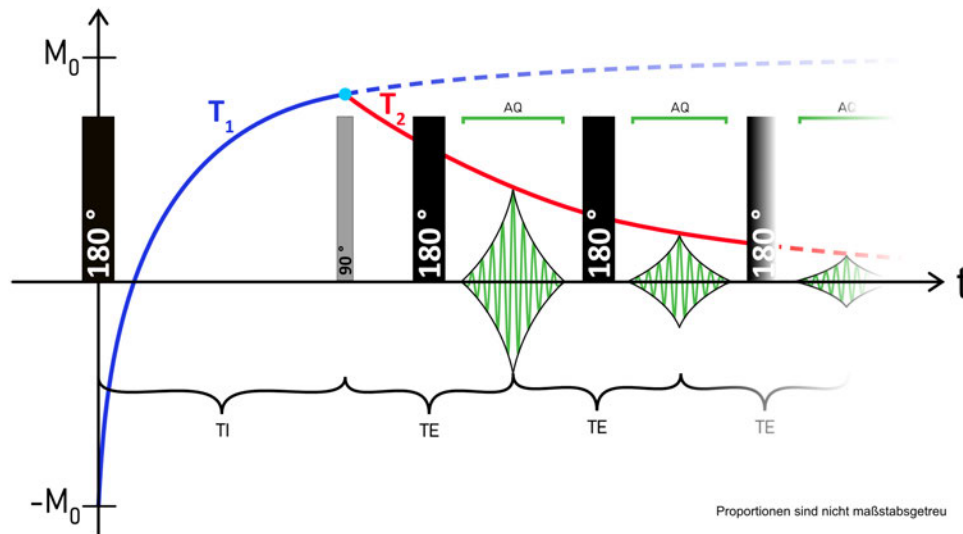


Abbildung 4 – Schematische Darstellung der verwendeten T_1/T_2 -Sequenz: Die verwendete Sequenz besteht zu Beginn aus einer ‚Inversion Recovery‘ (IR)-Sequenz (blauer T_1 -Verlauf), gefolgt von einem CPMG-Echozug zur Codierung des T_2 -Signales (roter T_2 -Verlauf). Die IR-Sequenz besteht aus einem 180° -Puls, gefolgt von einem 90° -Puls mit einem Abstand von T_I . Die Sequenz iteriert 32-mal über logarithmisch ansteigende T_I -Zeiten. Der CPMG-Echozug besteht aus 5000 180° -Pulsen, welche je einen Abstand von $T_E = 3\text{ ms}$ haben. **AQ:** Aquisitionsfenster; **T_I :** Inversionszeit; **T_E :** Echozeit.

MR-Experiments dar. Wenn mehrere Pulse hintereinander verwendet werden, spricht man von einer MR-Pulssequenz oder kurz MR-Sequenz. Abhängig davon, worauf bei der Messung der Fokus gelegt werden soll z. B. auf die Detektion einer Läsion im Gehirn oder die Untersuchung der Bandscheiben oder Gelenke wird eine andere, genau für diesen Zweck optimierte MR-Sequenz verwendet. In der vorliegenden Arbeit wurde eine Messsequenz verwendet, welche es erlaubt, simultan die T_1 - und T_2 -Verteilung für die untersuchte Probe zu ermitteln (Abbildung 4). Die in dieser Arbeit verwendete Art der Messung wird auch als 2D-Relaxometrie bezeichnet [57–59], da die verwendete MR-Sequenz und die Postprozessierungsschritte es ermöglichen, in einem Experiment gleichzeitig Informationen zur T_1 - als auch zur T_2 -Verteilung zu erheben. Die verwendete T_1/T_2 -Sequenz wurde erstmalig von Venkataraman et al. sowie Song et al. 2002 beschrieben [60, 61]. Die notwendigen Schritte für die Verarbeitung der Daten (Postprozessierung) wurden in den darauffolgenden Jahren u. a. von Marigheto et al. und Mitchell et al. veröffentlicht [62, 63]. Sofern nicht anders gekennzeichnet, wurden alle Informationen in diesem Abschnitt aus dem Buch ‚MRI made easy‘ von Prof. Dr. med. Hans H. Schild entnommen [64].

2.4.2 Aufbau und Funktion eines künstlichen neuronalen Netzwerks (KNN)

Wie bereits einleitend beschrieben, sollen künstliche neuronale Netzwerke (KNN) genutzt werden, um die aufgenommenen MR-Daten zu klassifizieren. Ähnlich wie beim menschlichen Nervensystem besteht ein KNN aus zahlreichen einzelnen Elementen, welche ebenfalls als Neurone bezeichnet werden. Diese sind typischerweise in Schichten (engl. *layer*) angeordnet (Abbildung 5).

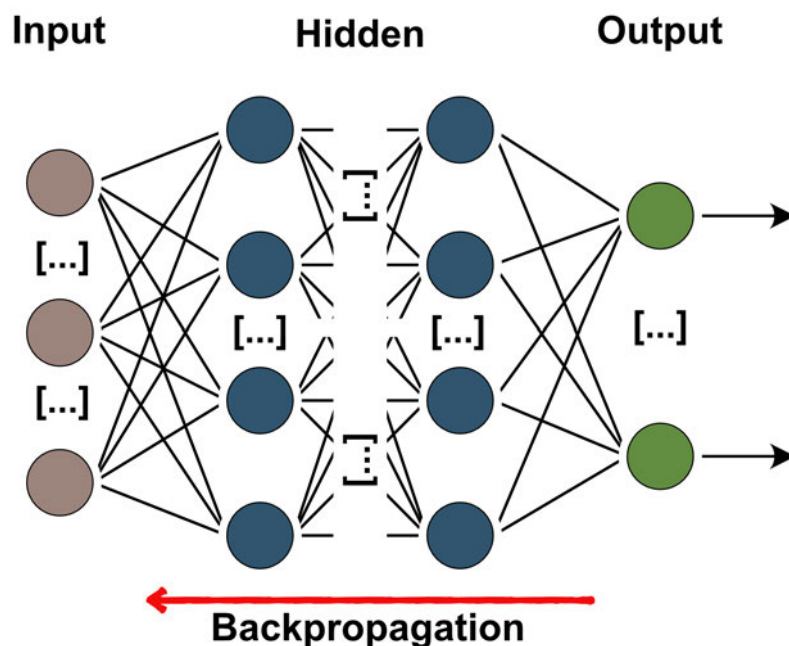


Abbildung 5 – Schematischer Aufbau eines KNN: In der hier dargestellten Version eines KNN ist jeder Knotenpunkt (=Neuron) mit jedem Neuron des nachfolgenden Layers (= Ebene / Schicht) verknüpft. Ein solcher Layer wird meist auch als ‚dense‘ oder ‚fully connected‘ layer bezeichnet. Die exemplarische Architektur besteht aus einem Eingabe (engl. Input)-Layer und zwei sog. versteckten (engl. hidden)-Layern. Die Ausgabe erfolgt dann über einen Ausgabe (engl. Output) -Layer. Als Backpropagation wird der Prozess bezeichnet, bei welchem die Neuronen nach dem einmaligen Durchlaufen der Layer die gelernte Information wieder zurück an die vorangegangenen Layer geben, um im nächsten Trainingsdurchlauf eine feinere Justage der Gewichte vornehmen zu können. Die abgebildete Architektur ist ein sehr vereinfachtes und generisches Beispiel. Die Struktur, d. h. die Anzahl der hidden-Layer und auch die Anzahl der verwendeten Neuronen in den Input-, Hidden- und Output-Schichten und auch die Art der Verknüpfung zwischen den einzelnen Neuronen wird an die jeweilige Fragestellung individuell angepasst. Dies wird durch die [...] in der Abbildung angedeutet.

Die Neuronen einer Schicht stehen in Verbindungen mit den Neuronen der nachfolgenden Schicht. Jedes Neuron führt eine entsprechende Berechnung durch. Übersteigt das Resultat einen definierten Schwellenwert, wird das Neuron aktiviert und gibt diese Information als Zahlenwert an die nachfolgenden Neuronen weiter. Dieser Prozess wird bei KNNs durch die Aktivierungsfunktion modelliert [65–67]. In biologischen Systemen basiert die Ausgabe bzw. die Weiterleitung eines Neurons auf dem ‚Alles-oder-Nichts-Prinzip‘. Metaphorisch wird die Rolle der Aktivierungsfunk-

tion in der Biologie vom sog. Axonhügel übernommen. Überschreitet das von den Dendriten eingehende Signal hier einen gewissen Schwellwert (= Schwellenpotential), ‚feuert‘ das Neuron, d. h. es leitet ein elektrisches Signal entlang seines Axons zu den Synapsen weiter. Liegt der eingehende Reiz am Axonhügel unterhalb des Schwellenpotentials, wird kein Reiz entlang des Neurons weitergeleitet [68]. In der Informatik war die Ausgabe bei älteren KNN-Modellen ebenfalls auf das ‚Alles-oder-Nichts-Prinzip‘ beschränkt, d. h. der Ausgabewert betrug entweder 1 oder 0. Neuere Aktivierungsfunktionen wie Sigmoid, tanh oder ReLU erlauben jedoch kontinuierliche Werte zwischen 1 und 0 [69].

Jede Verbindung zwischen zwei Neuronen ist über ihre Gewichtung (engl. *weight*) charakterisiert (Abbildung 6). Dieses Gewicht beschreibt, wie wichtig die Verbindung für das finale Resultat des Netzwerks ist. Bei der Initialisierung des Netzwerks werden die Werte der Gewichte zufällig gewählt.

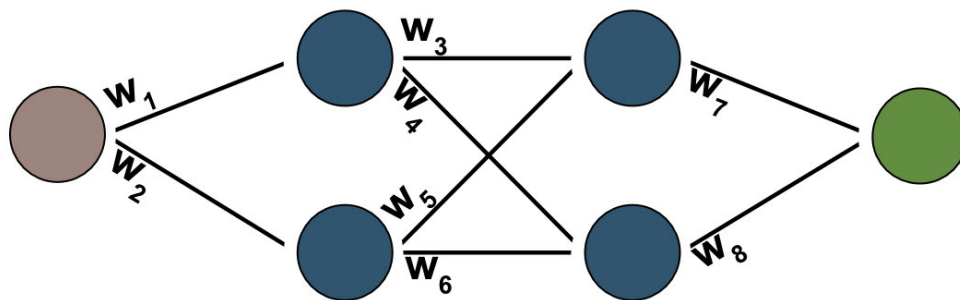


Abbildung 6 – Schematische Visualisierung der Gewichte eines KNNs: Die Verbindungen zwischen zwei Neuronen in einem KNN wird maßgeblich über das Gewicht (engl. *weight* - in der Abbildung mit ‚w‘ abgekürzt) charakterisiert. Die Werte dieser Gewichte werden anfänglich zufällig gewählt und dann mit jeder Trainingsiteration angepasst. Durch dieses Konzept ist das Netzwerk in der Lage zu ‚lernen‘, d. h. sich an die Fragestellung anzupassen.

Anschließend erfolgt eine mathematische Näherung an das vorgegebene Problem. Um den Erfolg eines Trainings quantifizieren zu können, wird eine Kostenfunktion verwendet, welche den Unterschied zwischen der erwarteten und der vorhergesagten Ausgabe des Netzwerks berechnet. Ziel des Trainings ist es, diese Kostenfunktion zu minimieren. Mit jeder neuen Trainingsiteration werden die Gewichte des Netzwerks so angepasst, dass die Kostenfunktion minimiert wird [70].

Ein essenzieller Erfolgsfaktor in der Geschichte der KNNs war die Implementierung der sog. Backpropagation im Jahr 1986 [71]. Diese führt dazu, dass basierend auf den Ergebnissen eines Durchlaufs die Gewichte für die nächste Trainingsiteration angepasst werden, d. h. das Netzwerk ‚lernt‘ von seinen Fehlern. Die Lernrate definiert hierbei, mit welchen Abständen bzw. Iterationen der Verlauf der Kosten-

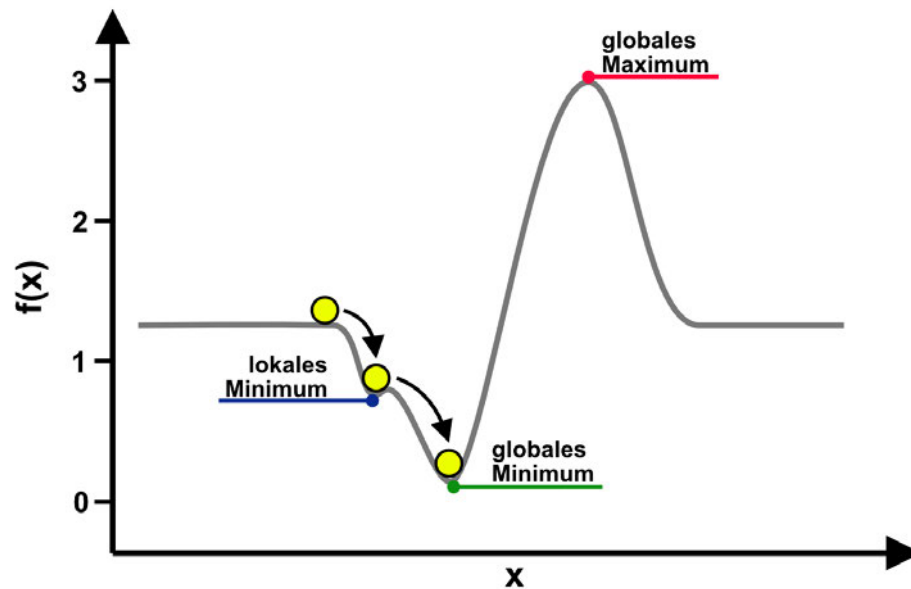


Abbildung 7 – Schematische Visualisierung der Kostenminimierung: Ziel des KNN Trainings ist es, die Kostenfunktion – auch Verlustfunktion (engl. loss function) genannt – zu minimieren. Dies wird von einer Optimierungsfunktion wie dem SGD (engl. stochastic gradient descend) oder adam (engl. adaptive moment estimation) durchgeführt. Häufig wird der Verlauf der Verlustfunktion mit einem rollenden Ball verglichen; daher wurde auch hier diese Analogie verwendet und die Verlustfunktion in der Abbildung durch eine gelbe Kugel angedeutet. Je nachdem, wie die Optimierungsfunktion arbeitet, kann es sein, dass sie nur ein lokales Minimum erreicht. Berücksichtigt sie jedoch ein gewisses ‚Trägheitsmoment‘, so kann sie das lokale Minimum überwinden und ihren Wert weiter reduzieren, bis sie das globale Minimum erreicht. Wenn das Trägheitsmoment jedoch zu klein bzw. groß ist, und das globale Minimum nicht so eindeutig eingrenzbar ist wie in dieser Abbildung, kann es dazu kommen, dass die Funktion das globale Minimum nie erreicht.

funktion abgetastet wird. Wird dieser Wert zu klein gewählt, braucht es eine sehr lange Zeit, bis das Minimum erreicht wird. Dahingegen kann eine zu große Lernrate dazu führen, dass das Minimum übersprungen wird. Um diesen Prozess zu optimieren, wird eine Optimierungsfunktion verwendet. Diese bestimmt unterschiedliche Faktoren wie die Lernrate und das Momentum, mit welchem abgetastet wird (Abbildung 7). Eines der ursprünglichsten Optimierungsverfahren ist das ‚stochastic gradient descend‘ Verfahren. Moderne Ansätze wie ‚adaptive moment estimation‘ oder kurz ‚adam‘ zeigen ebenfalls vielversprechende Ergebnisse [72].

Der Trainingsprozess wird typischerweise mehrere Male durchlaufen. Eine solche Iteration wird meist als Epoche bezeichnet. Zu Beginn des Trainings werden die Daten i. d. R. in Trainings- und Testdaten aufgeteilt. Dies bedeutet, dass ein Teil der Daten die Testdaten nicht für das eigentliche Training verwendet werden, sondern zur Evaluierung der Vorhersage des trainierten Modells. Innerhalb der Trainingsdaten erfolgt eine erneute Aufteilung in Trainings- und Validierungsdaten. Letztere erfüllen einen ähnlichen Zweck wie die Testdaten, nur dass sie den Prozess während des Trainings evaluieren und dazu dienen, die Gewichte möglichst effektiv anzupassen

[73, 74]. Um zu verhindern, dass das Netzwerk während des Trainings die gezeigten Daten schlichtweg ‚auswendig‘ lernt und somit nicht die tatsächlichen Muster und Zusammenhänge in den Daten erkennt, können so genannte ‚Drop-out‘ Schichten verwendet werden. Diese lassen in jeder Trainingsiteration einen Teil der Verbindungen zwischen den Neuronen zufällig weg. Dieser Anteil kann durch den Nutzer individuell vorgegeben werden. Durch das drop-out Verfahren ist das Netzwerk gezwungen, stärker zu generalisieren. Somit sinkt der ‚auswendig‘ gelernte Anteil im Training, wodurch die erzeugten Daten robuster d. h. zuverlässiger werden.

Die ‚traditionellen‘ KNNs sind v. a. in ihrer Performanz für die Verarbeitung von Bildern limitiert. Da jedoch die KI-basierte Verarbeitung von Bildern einen sehr interessanten Anwendungsfall für die KI-Forschung darstellt, wurden die sog. Convolutional Neural Networks (CNN) entwickelt. Die Grundlagen wurden in den 1980er und 1990er Jahren v. a. durch die Arbeit von Fukushima und LeCun gelegt [75–77]. Die CNNs dienen heute v. a. der Bild- und Audioerkennung und Verarbeitung [78]. Ihre besondere Eignung für die Bilderkennung rührt daher, dass sie in der Lage sind, mehrere Dimensionen an Daten aufzunehmen (Höhe, Breite und Tiefe). Des Weiteren erhöhen sie ihre Performanz dadurch, dass sie kleinere Regionen eines Bildes zusammenfassen und hierdurch die Komplexität der Daten reduzieren. Für Letzteres sind die sog. convolutional layer (dt. Faltungsschicht) verantwortlich. Diesen sind i. d. R. sog. pooling layer (dt. Bündelungsschichten) nachgeschaltet, welche die Daten der untersuchten Regionen zusammenfassen und somit deren Komplexität reduzieren. Diese Reduktion der Komplexität ermöglicht eine effektivere und beschleunigte Berechnung des gewünschten Ergebnisses. Abschließend verwendet auch das CNN voll-vernetzte Schichten (engl. *fully connected layer* oder *dense layer*) um eine Vorhersage für die erlernten Klassen zu tätigen [79] (Abbildung 8). Die CNN-Architektur hat sich in zahlreichen Tests und Challenges (z.B. MNIST oder ImageNet) als eine der effektivsten Methoden zur Bilderkennung herausgestellt [80]. Die CNNs dienen längst nicht mehr nur der Bilderkennung, sondern kommen auch in zahlreichen Aspekten der Bio-Wissenschaften zum Einsatz wie z. B. in der Radiologie, der Botanik, der Kardiologie oder der Psychologie [81–86].

2.4.3 Zelluläre Entwicklung

Neben der Aufnahmetechnik (Magnetresonanz) und der Analysemethode (künstliche neuronale Netzwerke) nehmen die verwendeten menschlichen und nicht-menschlichen

zellen eine zentrale Rolle in dieser Arbeit ein. Die Zelle ist die kleinste funktionelle Einheit aller lebenden Organismen. Manche Organismen wie Bakterien, manche Pilze oder Algen bestehen nur aus einer einzelnen Zelle. Die meisten komplexeren Lebewesen bestehen aus einer Vielzahl von Zellen und unterschiedlich spezialisierten Zelltypen. Laut einigen Schätzungen besteht der menschliche Körper aus etwa $3 \cdot 10^{13}$ Zellen, wobei die im Körper enthaltenen Bakterien nicht mitgezählt werden. Die Zahl an unterschiedlichen Zelltypen wird typischerweise mit etwa 200 angegeben [87, 88].

Der Körper besteht aus unterschiedlichen Geweben, die aus eben diesen unterschiedlichen Zelltypen aufgebaut sind. Im Zuge der menschlichen Embryonalentwicklung bilden sich die Gewebe aus den drei sog. Keimblättern: Entoderm, Mesoderm und Ectoderm. Das Entoderm bringt epitheliales Gewebe wie Verdauungstrakt, Leber, Schilddrüse und Atmungstrakt hervor. Das Mesoderm bildet u. a. die Knochen, Skelettmuskulatur, Bindegewebe, Herz, Blutgefäße, Nieren und Keimdrüsen. Das Ectoderm bringt die Epidermis, das Nervensystem, die Sinnesorgane und die Zähne hervor [89].

An dem Entstehungsprozess dieser Gewebe sind maßgeblich unterschiedliche Arten von Stammzellen beteiligt. Die menschliche Entwicklung beginnt mit der befruchteten Eizelle, welche sich immer wieder teilt, bis sich etwa am Tag sechs die sog. Blastocyste gebildet hat (Abbildung 9) [90]. Bis zu diesem Entwicklungszustand spricht man bei den Zellen von totipotenten Stammzellen. Der Begriff leitet sich in der klassischen Embryologie davon ab, dass die entsprechenden Zellen in der Lage sind bzw. das Potential haben, einen kompletten Organismus zu entwickeln [91]. Bei

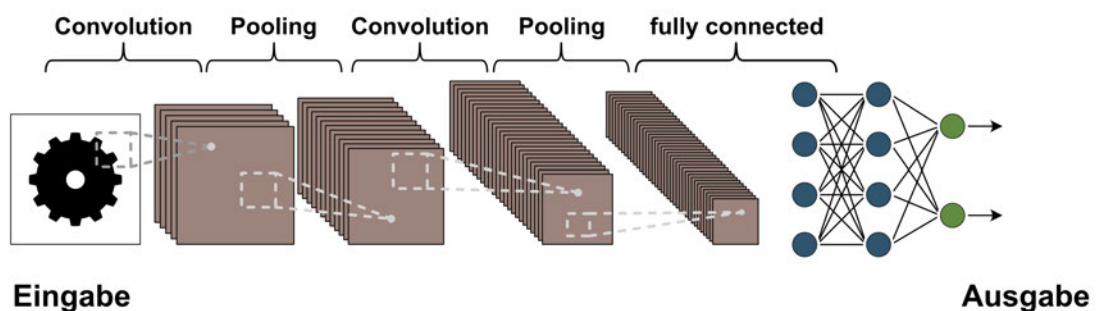


Abbildung 8 – Schematische Visualisierung des CNN Aufbaus: Ein convolutional neural network (CNN) ist vor allem für die Klassifikation von Bildern optimiert. Hierfür wird die Eingabe, sprich das zu untersuchende Bild, in seiner Komplexität reduziert, indem kleinere Bereiche zusammengefasst werden. Die Schichten (engl. layer), welche diese Berechnung durchführen, werden als Faltungsschichten (engl. convolutional layer) bezeichnet. I. d. R. folgen auf diese Schichten sog. pooling layer (dt. Bündelungsschichten) welche die berechneten Werte zu einer neuen, kleineren Schichten zusammenfassen. Dieser Prozess kann beliebig häufig wiederholt werden, wobei v. a. die Größe der zu kombinierenden Bereiche variiert werden kann. Abschließend nutzen die CNNs auch voll-vernetzte Schichten (engl. fully-connected / dense), um die so erzeugten Daten entsprechend der gewünschten Klassen vorherzusagen.

der Gewinnung von embryonalen Stammzellen werden die Zellen aus der inneren Zellmasse (Embryoblast) der Blastocyste entnommen [92]. Nachdem sich aus den isolierten Zellen kein vollständiger Organismus mehr bilden kann, bezeichnet man diese Zellen nicht mehr als toti- sondern als pluripotente Stammzellen. Aus pluripotenten Stammzellen lassen sich sämtliche Gewebe des menschlichen Körpers aus allen drei Keimblättern bilden. Im Laufe der weiteren Embryonalentwicklung verlieren die Zellen ihre Pluripotenz und werden als multipotente Stammzellen teilweise auch als Gewebestammzellen bezeichnet. Diese dienen i. d. R. dazu, abgestorbene Zellen des entsprechenden Gewebes wieder neu zu bilden, sind aber auf diese wenigen Zelltypen beschränkt [93].

Lange Zeit war man der Annahme, dass pluripotente Stammzellen nicht künstlich hergestellt werden können. Daher nutzte man hierfür meist überschüssige Embryonen aus künstlichen Befruchtungen. Dieses Verfahren ist in Deutschland jedoch durch § 2 des Embryonenschutzgesetz (ESchG) verboten [94]. Im Jahr 2012 erhielten John Gurdon und Shinya Yamanaka den Nobelpreis in Physiologie oder Medizin für die von ihnen beschriebenen Methoden, die es ermöglichen, bereits vollständig differenzierte Zellen wieder in einen pluripotenten d. h. Embryo-ähnlichen Zustand zurückzusetzen. Die so entstandenen Zellen wurden fortan als induzierte pluripotente Stammzellen (iPSCs) bezeichnet [95–97]. Die Etablierung der iPSCs stellte eine wissenschaftliche Revolution dar. Ab diesem Zeitpunkt war es möglich, auch

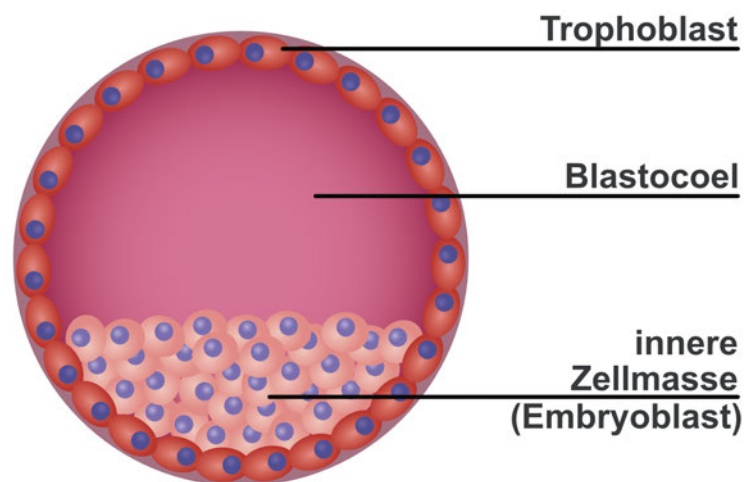


Abbildung 9 – Schematische Darstellung der Blastocyste: Im Gegensatz zu den vorherigen Entwicklungsstadien enthält die Blastocyste einen mit Flüssigkeit gefüllten Hohlraum, das Blastocoel. Die äußere Zellschicht der Blastocyste wird als Trophoblast bezeichnet. Der Trophoblast verwächst mit dem Endometrium der Mutter und bildet mit diesem die Placenta. Die im Inneren der Blastocyste befindliche Zellmasse (Embryoblast) enthält totipotente Zellen. Das bedeutet, dass aus diesen Zellen ein kompletter Organismus gebildet werden kann. Im Laufe der Embryonalentwicklung verlieren diese Zellen ihre Totipotenz und bilden erst pluripotente und anschließend multipotente Stammzellen. Hinweis: Die Größenverhältnisse in der Abbildung sind nicht maßstabsgetreu abgebildet.

komplexeres Gewebe wie Neuronen oder Zellen der Bauchspeicheldrüse aus einer einfachen Hautbiopsie herzustellen. Jedoch ist dieser Prozess sehr zeitintensiv und benötigt sehr teure Zellkulturmedien und Zusätze. Je nach Fragestellung wird aus diesem Grund auf eine andere Quelle an Stammzellen zurückgegriffen.

Laut Zahlen der OECD aus dem Jahr 2019 kam es in Deutschland zu 309 Hüfttransplantationen je 100.000 Einwohner [98]. Bei diesen Operationen gibt es für den entnommenen Hüftkopf i. d. R. keine weitere Verwendung. Daher dient das enthaltene Knochenmark häufig als Quelle für humane mesenchymale Stammzellen (hMSCs) für die Forschung. Teilweise werden diese Zellen auch als mesenchymale Stromazellen bezeichnet. Häufig werden diese Begrifflichkeiten austauschbar verwendet [99, 100]. Dabei bezieht sich ‚mesenchymal‘ auf den embryonalen Ursprung der Zellen, das Mesenchym. Dieses ist Teil des embryonalen Bindegewebes und entstammt größtenteils dem Mesoderm, jedoch auch teilweise dem Ekto- und Entoderm. Hierdurch erklärt sich auch das Differenzierungspotential der MSCs.

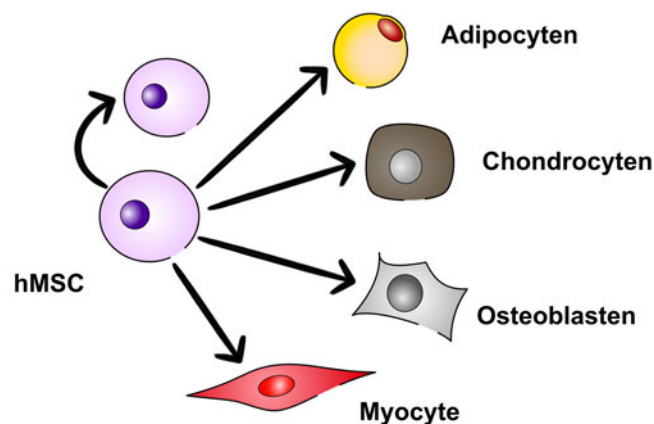


Abbildung 10 – Schematische Darstellung der MSC-Differenzierung: Kennzeichnend für eine Stammzelle ist ihr Potential, neue Stammzellen zu bilden (angedeutet durch gekrümmten Pfeil). Humane mesenchymale Stamm- oder Stromazellen können Zellen des Mesoderms bilden. Dazu gehören Zellen des Fettgewebes (Adipocyten), Knorpelzellen (Chondrocyten) und Knochenzellen (Osteoblasten). Teilweise geht man davon aus, dass MSCs auch unter bestimmten Umständen Zellen aus anderen Keimblättern wie Muskelzellen (Myocyten – Entoderm) oder teilweise sogar Nervenzellen (Neurone – Ectoderm) bilden können. Jedoch ist v. a. letzteres nicht eindeutig geklärt. Hinweis: Die Größenverhältnisse in der Abbildung sind nicht maßstabsgetreu abgebildet.

MSCs sind v. a. in der Lage, Zellen des Mesoderms zu bilden. Dies beinhaltet Zellen des Fettgewebes (Adipocyten), Knorpel- (Chondrocyten) und Knochengewebes (Osteoblasten). Mittlerweile geht man davon aus, dass MSCs auch Muskelzellen (Myocyten) bilden können [101], welche Teil des Entoderms sind. Im Bereich des Tissue Engineerings werden MSCs häufig dafür genutzt, entsprechende Gewebe als Testmodelle aufzubauen [102–106]. Neben ihrem Nutzen für die präklinische For-

schung wird ihre Wirkung im therapeutischen Kontext von unterschiedliche Erkrankungen wie Osteoporose [107], Arthrose [108], Verletzungen der Wirbelsäule [109], kardiovaskuläre Erkrankungen [110] sowie durch Diabetes verursachte, mikrovaskuläre Erkrankungen [111] seit Jahren diskutiert und erforscht. Auch wird ihr Potential für die Unterstützung von Krebstherapien untersucht [112]. Neben dem Knochenmark können MSCs auch aus der Nabelschnur und dem Fettgewebe sowie teilweise auch aus dem Lungengewebe und Blut isoliert werden [113–117].

Für die erwähnten therapeutischen Methoden ist die korrekte Identifizierung von Stammzellen bzw. von Zelltypen im Allgemeinen von zentraler Bedeutung. Hierfür stehen unterschiedliche Methoden zur Verfügung. Eine der am weitest verbreiteten Methoden zur zellulären Identifizierung sind histologische Färbungen. Diese können entweder unspezifisch Zellen färben, um die generelle Struktur des Gewebes sichtbar zu machen (z. B. Hämatoxylin-Eosin-Färbung) oder man verwendet Antikörper, um selektiv Zellen identifizieren zu können. V. a. die Antikörperfärbung findet häufige Anwendung bei der Identifizierung von Zellen, da sie abhängig von der Qualität der Antikörper Zellen mit hoher Genauigkeit markieren kann. Dazu werden Antikörper hergestellt, welche hoch spezifisch auf ein bestimmtes Antigen reagieren und sich bei Kontakt an dieses binden. Diese Antikörper sind zusätzlich mit einem Farbstoff gekoppelt, welcher Licht in einem definiertem Spektrum emittiert, wenn er mit einem Laser mit der passenden Wellenlänge angestrahlt wird. Hierdurch können Zellen hoch spezifisch angefärbt und lokalisiert werden. Teilweise werden auch mehrere Antikörper nacheinander gefärbt, um Zellen zu markieren, welche mehrere Marker auf ihrer Oberfläche besitzen. Zusätzlich werden normalerweise noch die Zellkerne angefärbt, um die einzelnen Zellkörper voneinander zu unterscheiden.

3 Material und Methoden

3.1 Material

3.1.1 Chemikalien

Tabelle 1: Übersicht über verwendete Chemikalien

Bezeichnung	Hersteller	Produktnummer
DescoseptAF	Dr. Schumacher	00-311-100
Ethanol (vergällt)	Carl Roth GmbH	4928.4
Incidin™ Plus	Ecolab	3011510

3.1.2 Geräte

Tabelle 2: Übersicht über verwendete Laborgeräte

Bezeichnung	Hersteller	Produktnummer
AccuBlock™ Mini-Compact	Labnet International	D0100
Arduino Mega 2560	Arduino	A000067
Autoklav	Systemec	VX-150
BBD 6220 Inkubator	Thermo Fisher Scientific Inc.	10423582
BioChem-VacuuCenter BVC Pumpe	Vacuubrand	20727201
Braccio Rescue kit	Arduino Tinkerkit	T0500001
Braccio Roboter	Arduino Tinkerkit	T050000
CoolCell® LX	Corning®	432001
DS18B20+ Linear IC - Temperatursensor	Maxim integrated™	DS18B20+
Einhängethermostat ET Basic	IKA-Werke	3163000
Eppendorf Research® plus 10 µl	Eppendorf AG	3124000032
Eppendorf Research® plus 100 µl	Eppendorf AG	3124000075
Eppendorf Research® plus 1000 µl	Eppendorf AG	3124000121
Fluoroptic® Thermometer	Luxtron corporation	m3300
Forma ULT Freezer	Thermo Fisher Scientific	TS586E 510L
Gefrierschrank	Liebherr	GNP 3376-20
Kleinschüttler KM CO2	Edmund Bühler GmbH + Co.KG	6136000
Mikroschalter MBB1 250 V/AC 5A	Hartmann Codier GmbH	62-263100008
Mikrowelle	Severin Elektrogeräte GmbH	MW 7890
Mikrozentrifuge 5417R	Eppendorf AG	FA45-24-11
MRT Konsole	Pure Devices GmbH	MMRT-L
MRT Magnet	Pure Devices GmbH	MAGSPEC
MRT Signalverstärker	Pure Devices GmbH	LNA23-24MHz
Multifuge X1R	Thermo Fisher Scientific	50146562
Multipette® M4	Eppendorf AG	4982000012
Pipettierhelfer accu-jet® pro	Brand	26304
Safe2020 Sterilbank	Thermo Fisher Scientific	51026640

Die Tabelle wird auf nächster Seite fortgesetzt.

Bezeichnung	Hersteller	Fortsetzung von Tabelle 2
		Produktnummer
Stickstofftank	Chart Industries	MVE TEC 3000
Wasserbad 006T	Lauda Dr.R.Wobser GmbH	2470-145-01
Zählkammer, Neubauer improved	Brand+ Co.KG	717805

3.1.3 Medien, Puffer, Zusätze

Tabelle 3: Übersicht über verwendete Zellkulturmedien

Bezeichnung	Hersteller	Produktnummer
DMEM 4,5g/l Glukose + GlutaMAX [®]	Thermo Fisher Scientific	61965-026
DMEM/F-12 (1:1) (1x) + GlutaMAX [®]	Thermo Fisher Scientific	31331-028
RPMI 1640 + GlutaMAX [®]	Thermo Fisher Scientific	61870-010

Tabelle 4: Übersicht über verwendete Medienzusätze

Bezeichnung	Hersteller	Produktnummer
D-(+)-Glucose -Lösung	Merck	G8769-100ML
Dexamethason	Merck	D4902-100MG
FCS	PanBiotech	P30-3306
hbFGF-2	R& D Systems	233-FB-025
IBMX	AppliChem	A0695.0001
Indomethacin	Merck	I8280-5G
Insulin	Merck	I9278-5ML
Lipid Mix 1	Merck	L0288-100ML
Penicillin-Stroptomycin	Merck	P4333+100ML

Tabelle 5: Übersicht über weitere Chemikalien und Puffer für die Zellkultur

Bezeichnung	Hersteller	Produktnummer
Dimethylsulfoxid (DMSO)	Merck	D8418-50ML
PBS ⁻	Merck	D8537-500ML
Trypanblau (0,4%)	Thermo Fisher Scientific	15250-061
0,5 Trypsin/EDTA (10X)	Thermo Fisher Scientific	15400-054

Tabelle 6: Übersicht über verwendete Puffer und Lösungen für Histologie

Bezeichnung	Hersteller	Produktnummer
2-Propanol	Carl Roth GmbH + Co.KG	6752.2
Agarose LE	Genaoxx bioscience GmbH	M3044.0500
Ölrot O	Merck	O0625-25G
ROTI [®] Histofix (4%)	Carl Roth GmbH	P087.3

Tabelle 7: Übersicht über verwendete Kontrastmittel

Bezeichnung	Hersteller	Produktnummer
Dotagraf 0,5 mmol/ ml	Jenapharm GmbH & Co. KG	203793/01

3.1.4 Software

Tabelle 8: Übersicht über verwendete Software

Bezeichnung	Hersteller	Version
Affinity Designer	Serif (Europe) Ltd	2
Affinity Publisher	Serif (Europe) Ltd	2
Arduino IDE	Arduino.cc	1.8.9
Citavi	Swiss Academic Software GmbH	6
Prism	GraphPad Software	9
ImageJ	NIH	1.52
Inventor	Autodesk	2019
JupyterNotebook	Project Jupyter	6.0.1
LaTeX	LaTeX Project Team	3
MATLAB®	The MathWorks, Inc.	R2019b
MindManager	Corel Corporation	18.0.284
MS Office	Microsoft	2016
Netron	Lutz Roeder	6.1.7
Overleaf	Overleaf	v2
Python	Python Software Foundation	3.8
Sourcetree	Atlassian, Sydney	3.1.2
TexMaker	Pascal Brachet	5.9.6

3.1.5 Verbrauchsmaterialien

Tabelle 9: Übersicht über verwendete Verbrauchsmaterialien

Bezeichnung	Ausführung	Hersteller	Produktnummer
Combitips®	2,5 ml	Eppendorf AG	0030089650
	5,0 ml	Eppendorf AG	0030089669
Einweg-Kulturröhrchen	10x75 mm	Duran	23 172 03 065
Green-TEC Pro Filament (2,85 mm)	Natur	Filamentworld	GTP300X800NAT
Kryoröhrchen	2,0 ml	Greiner bio-one	126280
Microplatten	96-well	Greiner bio-one	655073
Nunc EasyFlask	T-75	Thermo Fisher Scientific	156800
Pipettenspitzen	0,1 - 20 µl	Nerbe Plus GmbH	07-362-2005
	0 - 200 µl	Nerbe Plus GmbH	07-366-2005
	100 - 1250 µl	Nerbe Plus GmbH	07-369-2005
Reaktionsgefäße	0,2 ml	BioZym Biotech	711088
	1,5 ml	Sarstedt	72.690.001
	2,0 ml	Sarstedt	72.695.500
Serologische Pipetten	5 ml	Greiner bio-one	606180
	10 ml	Greiner bio-one	607180

Die Tabelle wird auf nächster Seite fortgesetzt.

Bezeichnung	Ausführung	Hersteller	Fortsetzung von Tabelle 9
			Produktnummer
	25 ml	Greiner bio-one	760180
	50 ml	Greiner bio-one	768180
Silikongummistopfen	20 mm	VWR	217-9503
Zellkulturflaschen	T-75	TPP	90026
	T-150	TPP	90151
Zellkulturplatten	6-well	TPP	92006
	12-well	TPP	92012
	24-well	TPP	92012
	96-well	TPP	92096
Zentrifugenröhrchen	15 ml	Greiner bio-one	188271
	50 ml	Greiner bio-one	227261

3.1.6 Zellen

3.1.6.1 Primärzellen

Alle Arbeiten entsprachen den Bestimmungen der Ethikkommission der Julius-Maximilians-Universität Würzburg (Abstimmung 280/18 und 182/10).

Tabelle 10: Übersicht über verwendete Primärzellen

Bezeichnung	Beschreibung	Herkunft
humane mesenchymale Stromazellen (hMSC)	Stromazellen isoliert aus Knochenmark	<i>homo sapiens</i>

3.1.6.2 Zelllinien

Alle verwendeten Zelllinien entstammten der Biobank des Lehrstuhls für Tissue Engineering und Regenerativer Medizin der Uniklinik Würzburg.

Tabelle 11: Übersicht über verwendete Zelllinien

Zelllinie	Spezies	Gewebetyp	Morphologie	Herkunft
A549	<i>Homo sapiens</i>	Lunge	Epithel	ATCC® CCL-185™
C2C12	<i>Mus musculus</i>	Muskel	Myoblast	ATCC® CRL-172™
CHO-k1	<i>Cricetulus griseus</i>	Ovar	Epithel	ATCC® CCL-61™
HEK293T	<i>Homo sapiens</i>	Embryonale Niere	Epithel	ATCC® CRL-3216™
HeLa	<i>Homo sapiens</i>	Cervixkarzinom	Epithel	ATCC® CCL-2™
K562	<i>Homo sapiens</i>	Knochenmark	Lymphoblast	ATCC® CCL-243™
L929	<i>Mus musculus</i>	Unterhautfettgewebe	Fibroblast	ATCC® CCL-1™
MDA-MB-231	<i>Homo sapiens</i>	Brust/Milchdrüse (Metastase)	Epithel	ATCC® HTB-26™
THP-1	<i>Homo sapiens</i>	Blut (Akute monozytische Leukämie)	Monozyten	DSMZ ACC 16
Vero	<i>Cercopithecus aethiops</i>	Niere	Epithel	ATCC® CCL-81™

3.2 Methoden

3.2.1 Zellkultur

Um die Viabilität von Zellen unter Laborbedingungen zu gewährleisten, ist es von zentraler Bedeutung geeignete Methoden zu verwenden. Kritische Punkte in der Zellkultur sind das Auftauen, das Passagieren und das erneute Einfrieren der Zellen. Sollen die kultivierten Zellen in einen anderen Zelltyp differenziert werden, sind hierfür entsprechende Protokolle notwendig, welche i. d. R. in einem Nachweis zur erfolgreichen Differenzierung enden.

3.2.1.1 Auftauen von Zellen

Eine Methode, um Zellen über einen sehr langen Zeitraum in ihrer Entwicklung und Alterung zu stoppen, ist, sie einzufrieren. Hierzu wurden alle verwendeten Zellen in dafür vorgesehenen Boxen in flüssigem Stickstoff gelagert. Vor dem Auftauen wurden 5 ml des entsprechenden Zellkulturmediums in 15 ml-Zentrifugenröhrchen vorgelegt und im Wasserbad auf 37 °C angewärmt. Die verwendeten C2C12-, L929-, HEK293T- und HeLa-Zellen wuchsen in DMEM-Medium mit GlutaMAX und 4,5 g/l Glukose supplementiert mit 10 % FCS und 1 % Penicillin-Streptomycin. Vero-Zellen wuchsen im selben Medium, jedoch war hier zusätzlich 1 % Natriumpyruvat zugesetzt. Die MDA231-, K562- und A-549-Zellen wuchsen in RPMI 1640-Medium mit GlutaMAX, supplementiert mit 1 % Penicillin-Streptomycin und 10 % FCS. Die verwendeten CHO-Zellen wuchsen in DMEM/F-12-Medium mit GlutaMAX supplementiert mit 1 % Penicillin-Streptomycin und 10 % FCS. THP-1-Zellen wurden in RPMI-1640 Medium kultiviert, supplementiert mit 1 % Penicillin-Streptomycin und 10 % FCS.

Die Zellkulturflaschen wurden ebenfalls mit der entsprechenden Menge Medium befüllt (T-150 mit 20 ml; T-75 mit 10 ml) und im Inkubator vorgewärmt. Die Zellen wurden aus dem Stickstofftank entnommen und mittels Trockeneis gekühlt. Unter sterilen Bedingungen wurde das Kryoröhrchen geöffnet und 1 ml des vorgewärmten Zellkulturmediums aus dem Zentrifugenröhrchen mittels 1000 µl Pipette hinzugefügt. Durch vorsichtiges Auf- und Abpipettieren wurde ein Teil der gefrorenen Zellsuspension aufgetaut. Diese wurde in das vorbereitete 15 ml-Zentrifugenröhrchen, welches bereits Medium enthielt, überführt. Diese Schritte wurden so oft wiederholt, bis der gesamte Inhalt des Kryoröhrchens aufgetaut und überführt war. Anschließend wurden die Zellen bei 300 xg für 3 min zentrifugiert. Der Überstand wurde vor-

sichtig mit einer sterilen Pasteurpipette abgesaugt und das Pellet in 1 ml frischem Medium gründlich resuspendiert. Die Suspension wurde anschließend in die vorbereitete(n) Zellkulturflasche(n) überführt. Es wurden $1 \cdot 10^6$ Zellen pro 150 cm^2 Wachstumsfläche ausgebracht bzw. $5 \cdot 10^5$ pro 75 cm^2 . Die Zellen wurden unter Standard-Zellkulturbedingungen kultiviert, bis sie eine Konfluenz von ca. 80 % erreichten. In diesem Zeitraum wurde das Medium alle 3–4 Tage gewechselt. Für Primärzellen fand dies i. d. R. dreimal pro Woche (montags, mittwochs und freitags) statt, während für Zelllinien ein Medienwechsel zweimal pro Woche (montags und freitags) ausreichte.

3.2.1.2 Passagieren von Zellen und Zellzahlbestimmung

Um die Zellen in einer kontinuierlichen Wachstumsphase zu halten, ist es von zentraler Bedeutung, die Kultur, sobald sie die zur Verfügung stehende Wachstumsfläche bis zu einem definierten Prozentanteil (Konfluenz) bewachsen haben, in ihrer Zellzahl zu reduzieren. Dieser Prozess wird als Passagieren bezeichnet. Dabei unterscheidet sich das Passagieren von Zellen, welche auf dem Boden der Kulturflasche anwachsen (adhärente Zellen) vom Passagieren von Zellen welche nicht anwachsen (Suspensionszellen) deutlich. Ein wichtiges Werkzeug für das Passagieren von Zellen ist eine korrekte Bestimmung der vorhandenen Zellzahl. Diese wird i. d. R. mit einer Neubauerzählkammer bestimmt.

Adhärenz wachsende Zellen: Sobald die Zellen eine Konfluenz von ca. 80 % erreicht hatten, wurde die T-150-Flasche mit 10 ml (5 ml bei T-75) PBS⁻ gewaschen. Anschließend wurden $20 \mu\text{l}/\text{cm}^2$ auf 37°C vorgewärmtes Trypsin/EDTA zugegeben. Die Zellen wurden für 3 min unter Zellkulturbedingungen inkubiert. Anschließend wurde das Ablösen der Zellen unter dem Mikroskop kontrolliert und die Zellen ggf. durch leichtes Klopfen an den Rand der Flasche vollständig abgelöst. Mit der Zugabe von 10 ml serumhaltigen Mediums wurde die Trypsin-Reaktion abgestoppt und die Zellen in ein frisches, beschriftetes 15 ml-Zentrifugenröhrchen überführt. Anschließend wurde dieses bei $300 \times g$ für 3 min zentrifugiert. Danach wurde der Überstand steril abgesaugt und das Zellpellet in 5 ml frischem Zellkulturmedium aufgenommen und gründlich resuspendiert. In einer 96-well-Platte wurden $40 \mu\text{l}$ Trypanblau vorgelegt und $10 \mu\text{l}$ der Zellsuspension zugegeben. Dieses Verhältnis wurde entsprechend der Größe des Zellpellets angepasst. $10 \mu\text{l}$ dieser Mischung wurden auf eine vorbereitete Neubauerzählkammer aufgebracht und die vitale Zellzahl unter dem Mikroskop

bestimmt. Abschließend wurde der anfängliche Verdünnungsfaktor mit eingerechnet und so die Zellzahl pro ml Medium bestimmt. Zur weiteren Kultivierung der Zellen wurde eine definierte Anzahl an Zellen aus der Suspension entnommen und zusammen mit frischem Medium in eine entsprechend vorbereitete Zellkulturflasche überführt.

Da es sich bei den verwendeten Zelllinien um Zellen handelte, welche sich bereits seit mehreren Jahrzehnten in Kultur befinden, wurde davon abgesehen, eine absolute Passagenzahl anzugeben. Daher wurde am Anfang der Arbeit ein größerer Vorrat an Zellen eingefroren und dieser als Passage 0 (P0) definiert. Ab dem Auftauen wurden die Passagen in P+1 usw. gezählt.

Suspensionszellen: Da Suspensionszellen nicht anwachsen, mussten sie auch nicht mittels Trypsin abgelöst werden. Als Kulturflasche wurde eine spezielle T-75-Flasche verwendet, welche für die Kultur von Suspensionszellen geeignet war. Die Zellzahl wurde ebenfalls mittels Trypanblau und einer Neubauerzählkammer bestimmt (siehe oben). Die Zellen wurden weiter in derselben Flasche kultiviert. Beim Passagieren der Zellen wurden die Zellen in einem Zentrifugenröhrchen pelletiert (300 xg, 3 min) und die Flasche mit PBS⁻ gespült. Anschließend wurde frisches Medium in die Flasche gegeben und die entsprechende Zellzahl zugeführt. Die Suspensionszellen sollten eine Zelldichte von $1 \cdot 10^6$ Zellen/ml nicht übersteigen und wurden normalerweise mit $1 \cdot 10^5$ Zellen/ml ausgesät.

Lebensdauer von Zelllinien: Auch wenn i. d. R. davon ausgegangen wird, dass es sich bei Zelllinien um immortale Zellen handelt, wurden die verwendeten Zelllinien nicht über die Passage P+15 hinaus verwendet. Hierbei handelt es sich um eine gängige Praxis in der Zellkultur, um zu verhindern, dass mögliche Kultureffekte das Verhalten und Wachstum der Zellen beeinflussen. Durch diese Maßnahme kann eine gleichbleibende Qualität der gewonnenen Zellen sichergestellt werden.

3.2.1.3 Kryokonservierung von Zellen

Die Zellen lagerten für die langfristige Aufbewahrung im gefrorenen Zustand. Zum Einfrieren der Zellen wurde das CoolCell[®] der Firma Corning verwendet. Der Metallring der CoolCell[®] wurde im Inkubator auf 37 °C vorgewärmt. Die Zellen wurden entsprechend des obigen Protokolls kultiviert, von der Zellkulturplatte abgelöst

und die vitale Zellzahl bestimmt. Pro Kryoröhrchen wurden i. d. R. $1 \cdot 10^6$ Zellen eingefroren. Es wurden so viele Zellen entnommen, wie Kryoröhrchen benötigt wurden. Diese Zellen wurden für 3 min bei 300 xg pelletiert und in so viel Medium aufgenommen, dass final eine doppelt so hohe Zellkonzentration wie die gewünschte entstand. Parallel dazu wurde eine Vorratslösung aus 40 Volumen-% (Vol.-%) FCS, 40 Vol.-% des entsprechenden Zellkulturmediums und 20 Vol.-% DMSO vorbereitet. Diese Mischung wurde fortan als Kryomedium bezeichnet. Pro Kryovial wurden 500 µl des Kryomediums kalkuliert. Die Kryoröhrchen wurden entsprechend beschriftet und steril unter der Sicherheitswerkbank geöffnet. Mittels der passenden Eppendorf Combitip[®] wurde die Zellsuspension aufgenommen und 500 µl pro Kryoröhrchen verteilt. Danach wurden hierzu 500 µl des Kryomediums gegeben, sodass sich final eine Zellzahl von $1 \cdot 10^6$ Zellen/ml mit 20 Vol.-% FCS und 10 Vol.-% DMSO ergab. Der Metallring der CoolCell[®] wurde wieder eingesetzt und die verschlossenen Kryoröhrchen in die entsprechenden Löcher gegeben. Die CoolCell[®] wurde hiernach verschlossen und für 24 h bei -80°C eingefroren. Bedingt durch die Bauweise der CoolCell[®], nahm die Temperatur dabei mit $1^\circ\text{C}/\text{min}$ ab. Anschließend wurden die Zellen in den Kryotank überführt, wo sie bis zur Verwendung in der Gasphase über flüssigem Stickstoff bei -196°C lagerten.

3.2.1.4 Kultur und Differenzierung von humanen mesenchymalen Stromazellen (hMSCs)

Kultur: Die verwendeten MSCs wurden aus dem Knochenmark menschlicher Hüftköpfe (*caput femoralis*) entnommen. Da in dieser Arbeit keine neuen Stammzellen isoliert wurden, wird daher auch nicht auf das entsprechende Protokoll diesbezüglich eingegangen. Alle Arbeiten entsprachen den Bestimmungen der Ethikkommission der Julius-Maximilians-Universität Würzburg (Abstimmung 280/18 und 182/10). Die Zellen wurden entsprechend des oben beschriebenen Protokolls aufgetaut (3.2.1.1) und für die Wachstumsphase in DMEM-F12 (1:1)-Medium supplementiert mit 10 Vol.-% FCS und 1 Vol.-% P/S kultiviert. Für eine schnellere Expansion wurden 10 ng/ml FGF-2 zugegeben. Die Zellen wurden für die Expansion bis zu einer Konfluenz von 60–70 % kultiviert und dann entsprechend dem oben beschriebenen Protokoll (3.2.1.2) passagiert.

Differenzierung: Für die Differenzierung wurden die Zellen bis zu einer Konfluenz von 100 % kultiviert, anschließend ebenfalls abgelöst und die Zellzahl bestimmt. Danach wurde die entsprechende Zellzahl in das dafür vorgesehene Kulturgefäß gegeben und bis zu einer Konfluenz von 100 % kultiviert. An Tag 0 der Differenzierung wurde das Expansionsmedium abgenommen, die Zellen mit PBS⁻ gewaschen und nachfolgend mit frisch zubereitetem Expansionsmedium überschichtet. Für die adipogene Differenzierung bestand das Differenzierungsmedium aus DMEM-Medium mit 4,5 g/l Glukose und GlutaMax[®], supplementiert mit 10 Vol.-% FCS, 1 µM Dexamethason, 1 µg/ml Insulin, 100 µM Indomethacin, 500 µM IBMX, 1 Vol.-% D-Glukose und 0,1 Vol.-% Lipidmix [105]. Das Differenzierungsmedium wurde i. d. R. montags, mittwochs und freitags gewechselt. Für die adipogene Differenzierung wurden die Zellen für 21 Tage mit dem Differenzierungsmedium kultiviert. Die Differenzierung wurde unter dem Mikroskop kontrolliert und manifestierte sich durch die Bildung von Zellen mit einer deutlichen Ansammlung von intrazellulären Fetttröpfchen.

3.2.1.5 Histologische Färbung der differenzierten hMSCs

Für den Nachweis der erfolgreichen Differenzierung der MSCs wurden diese selektiv histologisch angefärbt. Hierfür wurden undifferenzierte hMSCs in eine 6-well-Platte ausgesät und solange im Expansionsmedium kultiviert, bis sie eine Konfluenz von 100 % erreicht hatten. Danach wurden sie wie schon im Abschnitt 3.2.1.4 beschrieben differenziert. Nach 21 Tagen wurde das Kulturmedium abgenommen und die Zellen zweimal mit PBS⁻ gewaschen. Anschließend wurden sie mit 2 ml ROTI[®] Histofix 4 % bei RT für 10 min fixiert. Die Zellen wurden erneut gewaschen, diesmal jedoch mit vollentsalztem Wasser (VE H₂O). In der Zwischenzeit wurde eine Stammlösung für die Färbung angesetzt. Diese bestand aus 0,5 g Ölrot O und 100 ml 2-Propanol. Für die Färbelösung wurde die Stammlösung 6:4 mit VE H₂O verdünnt. Die Zellen wurden für fünf Minuten mit 2 ml 60 % 2-Propanol bei RT inkubiert und mit 2 ml der vorbereiteten Ölrot O Färbelösung für 10 min bei RT gefärbt. Abschließend wurden die gefärbten Zellen erst mit 2 ml 60 % 2-Propanol und dann mit 2 ml VE H₂O gewaschen. Die Zellen wurden am Mikroskop bei entsprechender Vergrößerung zur Dokumentation fotografiert.

3.2.2 Validierung der MR-Parameter und Methoden

Aufgrund der vielen variablen Parameter, welche Einfluss auf das Ergebnis der MR-Messung nehmen können, ist es wichtig, deren Werte eindeutig zu bestimmen und anschließend die korrekte Funktionsweise des Experiments zu validieren. Grundlage jedes MR-Experimentes ist die verwendete Messsequenz. Diese definiert die Pulse und deren Abstände zueinander. Die Anregung der Probe mit den geeigneten HF-Pulsen stellt das zentrale Werkzeug dar. Hierfür ist die Bestimmung der korrekten Pulswinkel wichtig. Ebenfalls ist es wichtig, die Shim-Werte des Systems zu bestimmen, um möglich Inhomogenitäten auszugleichen. Für die Validierung der Messung bieten sich Phantome an, welche eine standardisierte, reproduzierbare und stabile Probe darstellen. Zur Evaluation der Signalqualität stellt das Signal-zu-Rausch-Verhältnis eine gute Kenngröße dar.

3.2.2.1 Aufbau der MR-Sequenz

Die zugrundeliegende Architektur der Messsequenz wurde von Philipp Mörchel (Fraunhofer IIS) in die vom Hersteller des MR-Scanners bereitgestellte MATLAB-Umgebung openMatlab eingepflegt – sie wird im Folgenden als T_1/T_2 -Sequenz bezeichnet. Dabei handelt es sich um eine T_1 -gewichtete T_2 -Messung, bei welcher in n -Wiederholungen (hier 32 PntsTI) die T_1 -Wichtung durch Veränderung der Inversionszeit (TI) erhöht wird (Unterabschnitt 2.4.1). Diese Inversionszeit wurde mit entsprechend vielen Punkten logarithmisch zwischen 3 ms und 15 s (MaxTE) skaliert. In jedem Iterationsschritt erfolgt eine T_2 -Abtastung mittels einer CPMG-Sequenz, welche aus 5000 rechteckigen 180° -Refokussierungspulsen besteht. Die verwendete Interechozeit (TE_i) lag bei 3 ms. Jedes Echo wurde mit einer Bandbreite von 200 kHz und 40 Messpunkten aufgenommen. Aus den angegebenen Parametern resultiert eine finale Datenstruktur von $32 \times 40 \times 5000$ Punkten (32 Messpunkte in T_1 , 40 Messpunkte pro Echo, 5000 Messpunkte in T_2).

Zusammengefasst ergaben sich als Standardmessparameter folgende Werte:

TR: 40 s; **TE_i:** 3 ms; **MaxTE:** 15 s; **NR:** 1; **Repetitions:** 1; **NA:** 1; **ND:** 0; **MinTI:** 5 ms; **MaxTI:** 15 s; **PntsTI:** 32; **DeadTime:** 5 μ s.

3.2.2.2 Pulswinkelbestimmung

Die Pulswinkel wurden automatisch durch die vom Hersteller implementierte Routine vor jedem Experiment neu bestimmt. Für die Festlegung der Dauer der 90° - und

180°-Pulse wurde ein entsprechender Bereich inkrementell abgetastet. Zuerst wurde der Wert für den 180°-Puls als doppelter Wert des 90°-Pulses definiert. Dann wurde dieser so lange in einem Spin-Echo-Experiment durchlaufen, bis eine max. Echoamplitude erreicht war. Anschließend wurde dieser Wert für den 90°-Puls gesetzt und der Wert für den 180°-Puls auf dieselbe Art nachjustiert. Die Amplitude der Pulse blieb dabei stets konstant.

3.2.2.3 Bestimmung der Shim-Werte

Die Shim-Werte wurden an einem vom Hersteller mitgeliefertem Ölphantom einmalig für das System mit der entsprechenden Herstellerfunktion bestimmt und anschließend als feste Werte für die Messungen übernommen. Die Shim-Werte wurden mit dem Ölphantom erhoben, da die kleinen 50 µl-Proben, welche für die durchgeführten Zellmessungen verwendet wurden, keine zuverlässigen Shim ermöglichten.

3.2.2.4 Präparation der Agarose-Dotagraf®-Zellphantome (ADCP)

Zur Überprüfung der Reproduzierbarkeit und Stabilität der MR-Messung wurden Phantommessungen durchgeführt. Für die Phantome wurden 0,6 Gew.-% Agarose in RPMI-1640-Medium gelöst. Zusätzlich wurden 0,25 µmol/ml Dotagraf® (= 1/2000 aus 0,5 mmol/ml Stammlösung) zugegeben. Die Agarose diente dazu, die T₁-Zeit einzustellen, während Dotagraf® für das Justieren der T₂-Zeit verantwortlich war. Da Dotagraf® auf dem Gadoliniumchelat Gadotersäure basiert, hat es eine starke Wechselwirkung mit den Protonen in seiner Umgebung. Dies führt zu einer schnelleren Relaxation des MR-Signals. Durch diesen Effekt konnte das Kontrastmittel verwendet werden, um die T₂-Zeit auf den gewünschten Wert einzustellen. Das Gemisch wurde für wenige Sekunden in der Mikrowelle bei 700 W aufgeköcht, bis sich die komplette Agarose gelöst hatte. Da sich das so hergestellte ADCP über einige Monate hielt, wurde eine Stammlösung von 20 ml angesetzt und diese bei 4 °C gelagert. Die Stammlösung wurde bei Bedarf erneut in der Mikrowelle so lange bei 700 W erhitzt, bis sich das ADCP wieder vollständig verflüssigt hatte. Die heiße Lösung wurde kurz in einem Eisbad abgekühlt, bis sie etwa handwarm war. Aus der abgekühlten Lösung wurde steril das Volumen der Proben zum Messen entnommen und in ebenfalls sterile 0,2 ml-Zentrifugenröhrchen überführt. Diese kühlten bei Raumtemperatur (RT) mindestens 10 min ab, bevor sie verwendet wurden. Die Stammlösung wurde anschließend wieder bei 4 °C gelagert.

3.2.2.5 Bestimmung des Signal-zu-Rauschen Verhältnisses (SNR)

Um das SNR zu bestimmen, wurde eine ADCP-Probe hergestellt und im MR-Scanner mit den Standardparameterwerten vermessen. Für die Ermittlung des Signalanteils wurde der Mittelwert des Betrages der jeweils ersten Messpunkte der letzten sechs Inversionen gebildet, da hier der Signalanteil am höchsten ist. Für den Rauschanteil wurde die Standardabweichung der letzten $2 \cdot 10^5$ Messpunkte verwendet, aufgrund der Annahme, dass das Signal komplett dephasiert war und daher nur noch Rauschen aufgenommen wurde. Für die Berechnung des SNR wurde der Quotient aus dem berechneten Signalwert und der berechneten Rausch-Standardabweichung gebildet.

3.2.3 Aufbau und Funktion der Robotikplattform

Aufgrund der langen Messdauer der Proben war die Etablierung eines Probenwechslers von großer Wichtigkeit. So sollte die Zahl an messbaren Proben pro Arbeitstag signifikant erhöht werden. Aufgrund der vielfältigen Einsatzmöglichkeiten wurde eine Robotikplattform entwickelt, welche aus 3D-gedruckten Komponenten und einem Roboterarm bestand. Neben dem erhöhten Durchsatz war auch die Temperierung der Proben eine zentrale Aufgabe der Plattform.

3.2.3.1 Aufbau der Probenbox

Um das Einstellen der Proben an den Computer weiterleiten zu können, wurde eine Probenbox entwickelt. Diese beinhaltete 12 Positionen für die zu vermessenden Proben. Die Box war am PC entwickelt und mittels FDM 3D-Druck gedruckt worden. An jeder Position der späteren Proben war ein D2F-FL Mikroschalter verbaut, welcher beim Einstellen der Probe gedrückt wurde (Abbildung 11). Jeder Schalter war mit einem $4,7 \text{ k}\Omega$ Vorwiderstand mit einem digitalen Pin eines Arduino Megas verbunden. Um den Aufbau der Robotikplattform modular zu halten, waren die Kabel der Schalter nicht direkt mit dem Arduino verbunden, sondern durch einen DB-25 Stecker getrennt. So konnte die Box bei Bedarf abgesteckt werden bzw. zukünftig mit einer weiteren Probenbox ergänzt werden, wodurch die Kapazität der Anlage von 12 parallelen Proben deutlich erweitert werden konnte.

Neben den Mikroschaltern war auch ein DS18B20 Tempepraturesensor in der Probenbox verbaut, welcher ebenfalls mit dem Arduino verbunden war und so die Temperatur innerhalb der Probenbox an den Computer weitergeben konnte. Neben dem

Temperatursensor in der Probenbox, waren noch zwei weitere Sensoren angeschlossen. Einer wurde in den Heizblock gesteckt und einer nahm die Umgebungstemperatur auf. So konnten alle wichtigen Temperaturen kontinuierlich überwacht werden. Jeder Temperatursensor war ebenfalls mit einem $4,7\text{k}\Omega$ Vorwiderstand mit einem der digitalen Pins der Arduino Megs verbunden.

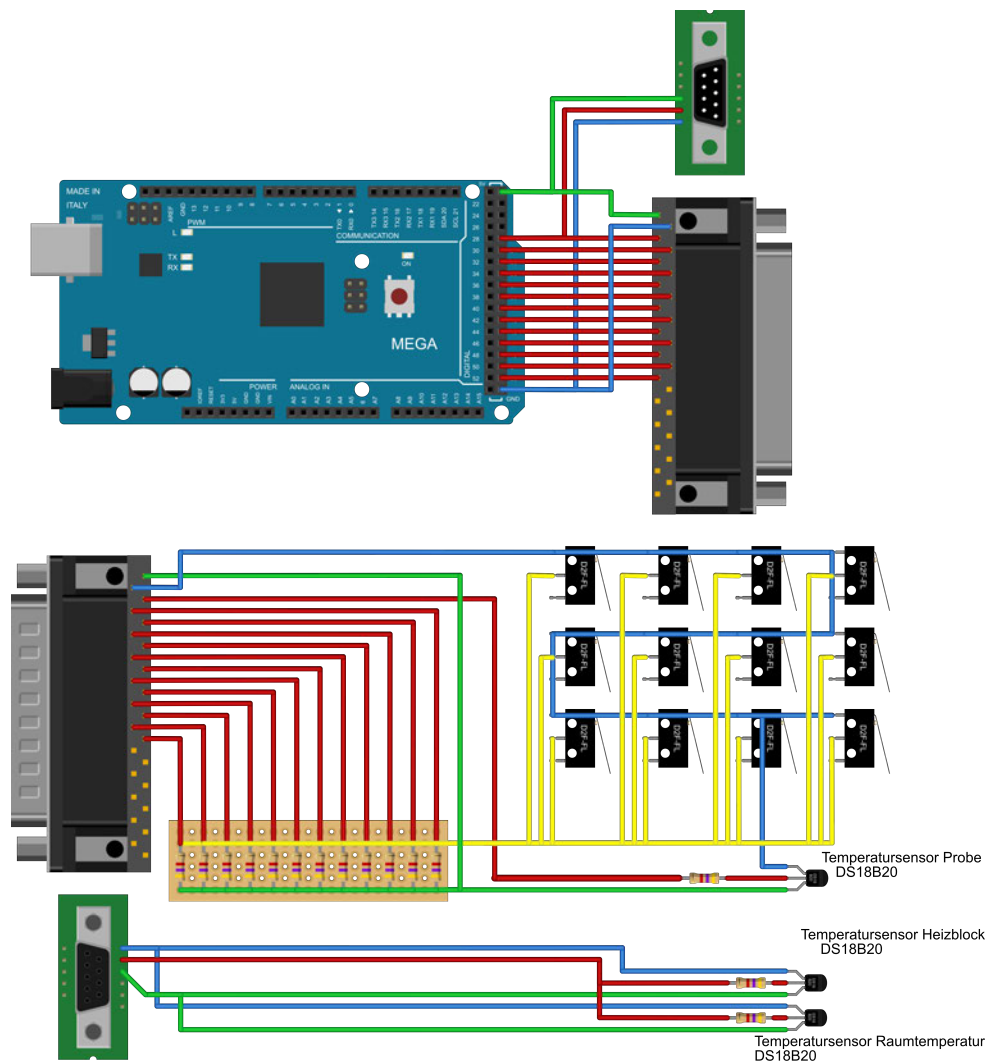


Abbildung 11 – Aufbau und Verkabelung der Probenbox: Um das Einstellen von bis zu 12 Proben in die Robotikplattform zu monitoren, waren in der Probenbox 12 D2F-FL Mikroschalter verbaut und über einen $4,7\text{k}\Omega$ Vorwiderstand mit den digitalen Pins des Arduino megas verbunden. Sobald einer der Schalter gedrückt wurde, gab der Arduino das entsprechende Signal an den Computer weiter. Neben den Mikroswitchern waren auch DS18B20 Temperatursensoren in der Probenbox und dessen Umgebung verbaut. Einer der Sensoren zeichnete die Temperatur innerhalb der Probenbox auf, wohingegen ein weiterer die Raumtemperatur aufzeichnete und ein dritter die Temperatur des Heizblockes. Die Temperatursensoren waren ebenfalls über einen $4,7\text{k}\Omega$ Vorwiderstand mit der Arduino verbunden.

3.2.3.2 Bedienung der Plattform

Die zentrale Schnittstelle für den MR-Scanner und die Roboterplattform wurde mittels MATLAB® erstellt. Um dies so nutzerfreundlich wie möglich zu gestalten, wurde eine graphische Oberfläche (engl. *Graphical User Interface* GUI) entwickelt, in welcher der Nutzer alle relevanten Informationen ablesen und eingeben konnte.

Die Kommunikation mit dem Arduino lief über eine serielle Schnittstelle via USB,

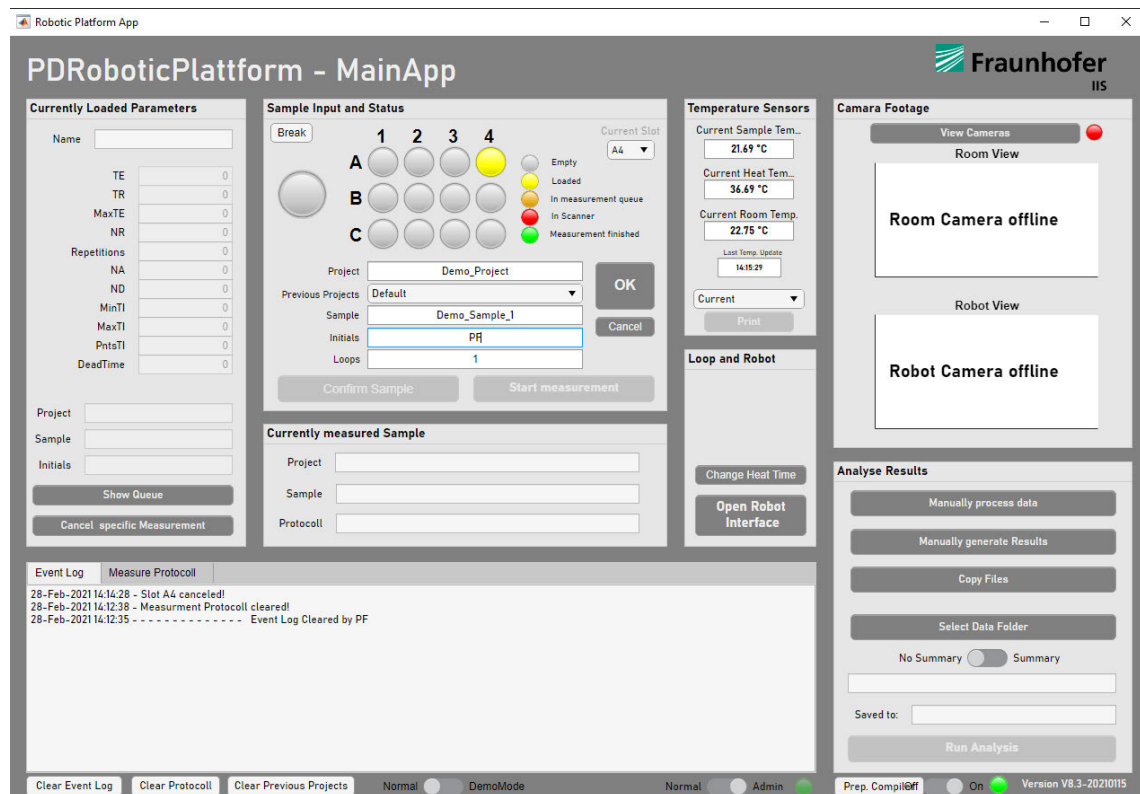


Abbildung 12 – Exemplarische Darstellung der verwendeten MATLAB® GUI. In der erstellten GUI wurde dem Nutzer über ein 3x4 Raster angezeigt, welche Probe gerade bearbeitet wird. Ein entsprechendes Ampelsystem zeigte dabei den Status der Probe an. Außerdem konnte der Nutzer unterschiedliche Projekt- und Proben-spezifische Daten eingeben. Es war dem Nutzer möglich die aktuellen Temperaturen der unterschiedlichen Messfühler aus zu lesen und ggf. deren Verlauf graphisch auftragen zu lassen. Über das 'Event Log' konnte der Nutzer jederzeit die zu letzt ausgegebenen Statusmeldungen einsehen. Sofern Justagen an den Einstellungen des Roboters notwendig waren, konnte der Nutzer über 'Open Robot Interface' direkt auf die gespeicherten Daten zugreifen und diese ggf. anpassen. Es war dem Nutzer auch möglich die eingestellten MR-Parameter ein zu sehen, um sicher zu stellen, dass die richtige Probe auch mit den richtigen MR-Parametern vermessen wurde.

ebenso wie die Kommunikation mit dem MR-Scanner. Der Nutzer platzierte die entsprechende Probe in der Probenbox und gab die relevanten Probeninformationen in der GUI ein. Dazu zählten eine Probenbezeichnung, eine Projektbezeichnung, Informationen zu den gemessenen Zellen (Zelltyp, Passage, Zellzahl, verwendetes Medium und Zusätze) und die relevanten Parameter für die MR-Messung (TE, TR, MaxTE, NR, Repetitions, NA, ND, MinTI, MaxTI, PntsTI und DeadTime). Die Parameter der MR-Messung konnten als Voreinstellungen gespeichert werden. Alle weiteren

MR-Parameter waren durch das System vorgegeben und konnten durch den Nutzer nicht editiert werden. Ein genauerer Ablauf der vorzunehmenden Schritte kann dem Prozessdiagramm in Abbildung 13 entnommen werden.

Der Roboter füllte den Heizblock mit max. vier Proben auf. Wenn mehr als vier

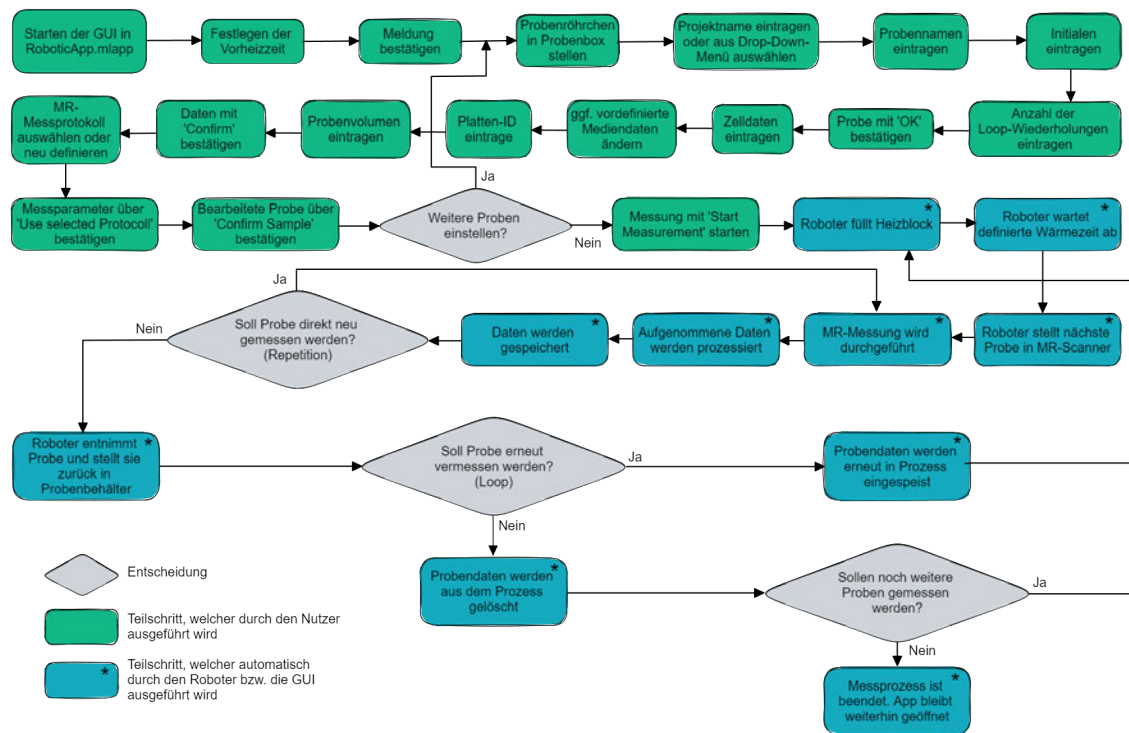


Abbildung 13 – Prozessdiagramm für den Programmablauf der verwendeten MATLAB® GUI: Die GUI führte den Nutzer durch die Eingabe der relevanten Probenparameter, wie z. B. Probenname, Projektbezeichnung so wie die Eingabe der relevanten Zelldaten, wie z. B. Konfluenz, Bezeichnung, Zellzahl oder Passage. Die gewünschten Sequenzparameter konnten als Vorlage gespeichert und bei Bedarf abgerufen werden. Ein wichtiger Aspekt war die Unterscheidung zwischen ‚Repetition‘ und ‚Loop‘. Als ‚Repetition‘ wurde eine Wiederholung bezeichnet, bei welcher die Probe nach der Messung nicht aus dem Scanner genommen wurde und direkt n-mal neu vermessen wurde. Als ‚Loop‘ wurde die Wiederholung bezeichnet, bei welcher die Probe nach der Messung entnommen wurde und an das Ende der Messreihenfolge wieder angehängt wurde. Die Anzahl der Repetitions bzw. Loops konnte individuell angepasst werden und war beliebig kombinierbar.

Proben in der Probenbox vorbereitet waren, blieben diese vorerst dort stehen. Im Anschluss an die Wärmezeit im Heizblock positionierte der Roboter die Probe im MR-Scanner, wo sie erneut für i. d. R. 10 min temperiert wurden. Anschließend wurde die eigentliche MR-Messung gestartet. Wenn diese abgeschlossen war, entnahm der Roboter die Probe aus dem MR-Scanner und platzierte sie auf ihrer ursprünglichen Position in der Probenbox. Abschließend füllte der Roboter den Heizblock mit weiteren Proben auf, die zur Messung vorgesehen waren, jedoch zuvor im Heizblock noch keinen Platz gefunden hatten.

Wenn vom Nutzer gewünscht, bestand die Möglichkeit, eine Probe über mehrere Messzyklen zu messen, d. h. die Probe wurde vorgewärmt und anschließend ver-

messen und dann wieder in die Probenbox zurückgestellt. Die Probeninformationen wurden automatisch übernommen und die Probe erneut zum Vermessen vorgemerkt. Diese Form der Wiederholung wurde als Loop bezeichnet. Außerdem war es möglich, dass die Probe mehrmals direkt hintereinander im MR-Scanner vermessen wurde, ohne dass sie zwischen den Messungen herausgenommen wurde. Normalerweise war hierfür eine gewisse Wartezeit zwischen den Messungen vorgesehen, um eine vollständige Relaxation des MR-Signals zu gewährleisten. Diese Wiederholungen wurden als Repetitions bezeichnet. In regelmäßigen Abständen wurden die Positionen des Roboters überprüft, da sich diese durch Abnutzungs- und Ermüdungserscheinungen am Kunststoff des Roboters im Laufe der Zeit veränderten.

3.2.3.3 Validierung der Temperierungsroutine

Um sicherzustellen, dass die Proben zuverlässig auf 37 °C temperiert waren, musste der notwendige Temperierungsprozess zuerst validiert werden. Für die Messungen wurde ein optisches Thermometer der Firma Luxtron (m3300 Fluoroptic[®]) mit den entsprechenden Glasfasersensoren verwendet. Das Thermometer wurde auf 17 °C kalibriert. Dazu wurde eine entsprechende Kalibrierung in entsalztem Wasser durchgeführt, welches über Nacht in einem Heizblock auf 17 °C temperiert wurde. Für die Messung wurde in den Deckel eines 0,2 ml-Zentrifugenröhrchens ein feines Loch gebohrt, genauso wie in den Silikonstopfen, welcher die verwendeten Glasröhrchen verschloss. Durch diese Löcher wurde die entsprechende Glasfaser-sonde geführt. Das Reaktionsgefäß wurde mit 50 µl RPMI-1640-Medium befüllt und die Messsonde in die Mitte des Mediums platziert, ohne dass sie dabei den Rand berührte. Anschließend wurde das so präparierte Messsystem mittels Eis auf nahezu 0 °C gekühlt. Für die Durchführung der Messung wurde ein entsprechendes MATLAB[®]-Skript verwendet.

Die Probe wurde aus dem Eis genommen und kurz abgetrocknet. Danach wurde sie direkt in den Heizblock mit 37 °C überführt. Nachdem die Probe dort die Maximaltemperatur angenommen hatte, wurde sie in den MR-Scanner überführt und dort ebenfalls für einige Minuten equilibrieren lassen. Final wurde die Probe wieder auf Raumtemperatur (RT) herunter gekühlt. Dieser Prozess wurde mehrere Male unabhängig voneinander wiederholt. Alternativ dazu wurde die Probe auch direkt aus dem Eis in den Scanner überführt und dort direkt equilibriert.

3.2.4 Vorbereitung der Zellproben und Viabilitätsnachweis

Aufgrund des geringen Volumens der Zellsuspension ist es von großer Wichtigkeit, bei der Präparation der Proben akkurat zu arbeiten. Der hierfür entwickelte Präparationsprozess soll sicherstellen, dass jede Probe aus der entsprechenden Anzahl an Zellen besteht. Um sicherzustellen, dass weder der Präparationsprozess noch die Prozessierung durch den Roboter bzw. die lange Messdauer dazu führen, dass ein Großteil der Zellen sterben, ist die Viabilität der Zellen zu überprüfen.

3.2.4.1 Herstellung der Zellproben

Die Zellen wurden entsprechend 3.2.1.2 und 3.2.1.2 abgelöst und die Zellzahl bestimmt. Die entsprechende Zellzahl i. d. R. $5 \cdot 10^6$ Zellen wurde entnommen und entsprechend dem Volumen entweder in ein steriles 1,5 ml - oder ein 2,0 ml -Reaktionsgefäß überführt. Falls das Volumen mehr als 2 ml entsprach, wurden die Zellen erst in ein 15 ml-Zentrifugenröhrchen überführt, für 3 min bei 300 xg pelletiert und dann mit 1,5 ml frischem Medium in ein entsprechendes Reaktionsgefäß gegeben. Anschließend wurden alle Proben in einer Tischzentrifuge bei 300 xg für 3 min pelletiert und das Medium vorsichtig mit einer 100 µl Pipettenspitze bis auf ein Minimum abgenommen. Anschließend wurden die Zellen in 50 µl frischem Medium (passend für jede Zellart, s. Unterabschnitt 3.2.1) resuspendiert. Beim Aufnehmen der Zellen wurden nicht nur die zugegebenen 50 µl aufgenommen, sondern das komplette Volumen. Die Differenz entstand durch das Restvolumen des Zellpellets. Dazu wurde die Pipettenspitze einer 100 µl-Pipette in das Reaktionsgefäß getaucht und mittels des Volumeneinstellrings der Pipette das komplette Volumen aufgenommen. Anhand der Volumenanzeige der Pipette konnte das exakte Volumen abgelesen werden. Die aufgenommene Zellsuspension wurde in ein steriles 0,2 ml-Zentrifugenröhrchen überführt und das ursprüngliche 1,5 ml/2,0 ml -Reaktionsgefäß mit weiteren 50 µl frischem Medium gespült, welche anschließend ebenfalls in das 0,2 ml -Gefäß gegeben wurden. Das 0,2 ml -Röhrchen wurde verschlossen und die Deckelanhänge mittels einer kleinen Schere abgeschnitten, damit es in das Glasröhrchen passte, welches für die Messung verwendet wurde. Das Glasröhrchen wurde mit einem Silikongummistopfen verschlossen und bei 300 xg für 2 min zentrifugiert. Anschließend wurde das kleine Kunststoffröhrchen wieder aus dem Glasröhrchen entnommen und die zuvor notierte Menge Volumen mit einer 100 µl-Pipette vorsichtig abgenommen, sodass ein Restvolumen von exakt 50 µl im Reaktionsgefäß verblieb, zusammen mit der gewünschten Anzahl an Zellen. Das Röhrchen wurde verschlossen, in das Glasröhrchen gegeben

und erneut mit dem Silikonstopfen fest verschlossen.

3.2.4.2 Viabilitätsbestimmung der Zellen nach der MR-Messung

Durchführung der Viabilitätsbestimmung: Für die Bestimmung der Viabilität wurde der NucleoCounter NC-200 zusammen mit den dafür vorgesehenen Via-1-Kassetten verwendet. Der NucleoCounter nutzte das System der Mikrofluidik, um ein minimales Volumen an Zelle zu vermessen. In den Via-1-Kassetten wurden die Zellen mit den Farbstoffen Acridinorange (AO) und 4,6-Diamidino-2-Phenylindol (DAPI) gemischt. Von den angefärbten Zellen wurden dann fluoreszenzmikroskopische Aufnahmen gemacht, welche automatisiert mittels der entsprechenden Software ausgewertet wurden. Da AO sowohl lebende als auch tote Zellen und DAPI ausschließlich tote Zellen anfärbte, konnte über das entsprechende Verhältnis eine Aussage über die Viabilität der gemessenen Zellen getroffen werden.

Die entsprechende Probe befand sich i. d. R. zuvor mit einem Volumen von 50 μl in einem 200 μl -Zentrifugenröhrchen. Mit der entsprechenden Pipette wurden 50 μl frisches Zellkulturmedium zugegeben und darin das Zellpellet vorsichtig resuspendiert; anschließend wurden die gesamten 100 μl in ein frisches 1,5 ml-Zentrifugenröhrchen überführt. Das 200 μl -Röhrchen wurde zweimal mit je 100 μl frischem Medium gespült, welches jeweils in das 1,5 ml-Röhrchen hinzugefügt wurden. Das Volumen im 1,5 ml-Röhrchen wurde abschließend mit 700 μl frischen Mediums auf 1 ml aufgefüllt. Aus dieser Zellsuspension, i. d. R. $5 \cdot 10^6$ Zellen pro ml, wurden 100 μl entnommen und in einem frischen 1,5 ml-Zentrifugenröhrchen mit 900 μl frischem Medium verdünnt. Die so verdünnte Zellsuspension wurde anschließend mit dem voreingestellten Standardprotokoll am NucleoCounter vermessen. Die Messdaten wurden als Tabelle für die Datenanalyse exportiert, zusammen mit einer graphischen Darstellung der eigentlichen Messwerte als Scatterplot.

Negativkontrolle: Um die Messmethode zu validieren, war die Verwendung einer geeigneten Negativkontrolle wichtig. Als Negativkontrolle wurden die Zellen wie oben beschrieben aufbereitet. Ein Heizblock wurde auf 90 °C vorgeheizt. Anschließend wurde das Reaktionsgefäß mit den Zellen für 30 min in den Heizblock transferriert. Dieser schüttelte die Probe bei 400 RPM. Für die Messung wurden die Zellen erneut, wie bereits oben beschrieben, mit entsprechendem Medium verdünnt und am Nucleocounter NC-200 ihre Viabilität ermittelt.

Positivkontrolle: Als Positivkontrolle wurde die Viabilität der präparierten Proben direkt nach der Aufarbeitung vermessen. Der Ablauf folgte dem im vorangegangenen Paragraphen beschriebenen Protokoll.

Unterschiedliche Kulturbedingungen: Um unterschiedliche Kulturbedingungen zu untersuchen, wurden die vorbereiteten Proben für 5 Stunden bei unterschiedlichen Bedingungen kultiviert. Pro Kulturbedingung wurden drei unabhängig präparierte Proben verwendet und anschließend deren Viabilität bestimmt. Drei Proben wurden bei RT auf der Werkbank kultiviert. Drei Proben wurden im Heizblock bei 37 °C kultiviert. Drei Proben wurden im Inkubator unter Standardbedingungen kultiviert. Für diese Proben wurde das Plastikkulturgefäß geöffnet, um einen Gasaustausch zu ermöglichen. Drei weitere Proben wurden ebenfalls im Inkubator kultiviert, jedoch in einem Gesamtvolumen von 1 ml Kultumedium auf die verwendeten $5 \cdot 10^6$ Zellen. Um den Einfluss von CO_2 zu untersuchen, wurden drei weitere Proben bei RT in einen luftdichten Plastikcontainer gestellt, die Deckel der Kulturgefäße geöffnet und für 5 min kontinuierlich mit einer Gasmischung aus 5 % CO_2 , 21 % Sauerstoff und 74 % Stickstoff begast.

3.2.5 Postprozessierung der Messdaten

Nachdem die Zellen im MR-Scanner vermessen wurden, mussten die aufgenommenen Daten prozessiert werden, um die entsprechenden Signalverläufe zu berechnen. Die so entstandenen Korrelationsspektren wurden auf die Form und Position der Signalcluster / Signalpeak hin untersucht. Die Zahl der so erzeugten Daten wurde dann für eine bessere Weiterverarbeitung amplifiziert. Dieser Amplifizierungsprozess wird als Augmentierung bezeichnet. Die durchgeführten Schritte werden nachfolgend beschrieben.

3.2.5.1 Postprozessierung der aufgenommenen Rohdaten

Wie bereits oben beschrieben, lagen die resultierten Daten als komplexer Datensatz in der Form von 32x40x5000 Datenpunkten vor. Zuerst wurden die gemessenen 40 Datenpunkte pro Echo gemittelt, um auf ein 32x5000 Format zu kommen. Aufgrund der Messesequenz waren alle Daten logarithmisch skaliert. Zur Reduzierung

der Komplexität wurde die Anzahl der Punkte in T_2 von 5000 auf 8 mal PntsTI reduziert. Da in dieser Studie mit 32 PntsTI gearbeitet wurde, entsprach dies 256 Datenpunkten. Anschließend wurde die Amplitude der Daten auf den Wert eins normalisiert. Eine ‚Prinicipal Component Analysis‘ (PCA) wurde mittels einer Singulärwertzerlegung (SVD) durchgeführt (Regularisierungsparameter $\lambda = 1,5$). Abschließend wurde eine inverse 2D Laplace-Transformation durchgeführt, bei welcher die Daten auf ein logarithmisches 300x300-Raster im Bereich von $1 \cdot 10^{-3}$ s bis 10 s für T_1 und T_2 transformiert wurde.

3.2.5.2 Berechnung des gewichteten Schwerpunktes eines Signalpeaks

Um die Position eines Signalpeaks besser beurteilen zu können, wurde dessen gewichteter Schwerpunkt berechnet. Die Signalcluster jeder Messung wurden mit der MATLAB®-Funktion *regionprops()* auf ihre geometrische Eigenschaften hin untersucht. Die *regionprops()*-Funktion erkennt und analysiert zusammenhängende Bildbereiche. Für jeden detektierten Bereich berechnete sie unterschiedliche Eigenschaften wie z. B. die Anzahl der Pixel im entsprechenden Bereich, die Zirkularität, die räumliche Orientierung oder den geometrischen Schwerpunkt (engl. *Centroid*). Mit der Eigenschaft ‚WeightedCentroid‘, welche sowohl eine binäre (schwarz-weiß) Matrix für die Form der Peaks als auch eine gewichtete Matrix mit den reellen Werten erwartete, wurde der gewichtete Schwerpunkt (engl. *Centroid*) für jede detektierte Region bestimmt. Er wurde als ‚gewichtet‘ (engl. *weighted*) bezeichnet, da er sowohl über die Geometrie als auch über die Verteilung der Datenpunkte berechnet wurde. Die berechneten Schwerpunkte konnten in einem Scatter-Plot gegeneinander aufgetragen werden.

3.2.5.3 Augmentierung der prozessierten Daten

Der Augmentierungsalgorithmus wurde von Daniel Weber (Fraunhofer IIS) entwickelt und durch mich an die konkrete Anwendung angepasst. Die Augmentierung erfolgte auf dem logarithmischen 300x300 Raster. Zuerst wurden die Werte auf ein lineares Raster interpoliert, welches aus zehnmal so vielen Stützpunkten wie das logarithmische Raster bestand. Anschließend konnte jeder beliebige Signalpeak, basierend auf seinem gewichteten Schwerpunkt (s. Unterunterabschnitt 3.2.5.2), selektiert werden. I. d. R. wurde der größte und zweitgrößte Peak selektiert, was den Medien- bzw. Zellpeaks entsprach. Diese konnten dann über zwei Operatoren modifi-

ziert werden: Ein Operator bewegte den Peak zufällig in T_1 bzw. T_2 -Richtung. Der andere Operator stauchte bzw. streckte den selektierten Peak zufällig. Der Nutzer konnte den Wert der Augmentierungsparameter individuell für T_1 und T_2 definieren. Dabei entsprachen die definierten Werte der gewünschten Standardabweichung der Verteilung. Der angegebene Wert der Standardabweichung wurde anschließend mit einem zufälligen Wert (basierend auf der Normalverteilung `randn()`-Funktion von Matlab[®]), multipliziert und der selektierte Peak um diesen erzeugten Wert in T_1 - bzw. T_2 -Richtung gestaucht/gestreckt bzw. bewegt. Abschließend wurden die so modifizierten Daten wieder auf das entsprechende logarithmische Raster angepasst und als *.png Bilddateien exportiert.

Um die Menge der Daten um den Faktor x zu erhöhen, wurde der Augmentierungsprozess x -mal durchlaufen. Die Anzahl an durchlaufenen Zyklen wurde als Augmentierungsfaktor bezeichnet.

3.2.6 Aufbau und Training der Support Vector Machine (SVM)

Um die erzeugten Daten korrekt klassifizieren zu können, wurde u. a. eine SVM auf den Daten trainiert und deren Ergebnisse evaluiert.

3.2.6.1 Aufbereitung der Trainingsdaten

Da die SVM darauf ausgelegt ist, einzelne Datenpunkt durch Vektoren zu trennen, wurden die erzeugten 2D-Spektren auf ihre gewichteten Schwerpunkte (s. Unterunterabschnitt 3.2.5.2) reduziert. Die Koordinaten des Schwerpunktes wurden als T_1 - und T_2 -Werte zusammen mit dem Namen der entsprechenden Zelllinien gespeichert. Außerdem wurde nur der Schwerpunkt des Zellpeaks berücksichtigt, da davon ausgegangen wurde, dass dieser die gesuchte Information trägt.

3.2.6.2 Training der SVM

Die vorbereiteten Daten wurden als CSV-Datei in Python eingelesen. Für das Training der SVM wurde das scikitlearn-Paket von Python verwendet [118, 119]. Die eingelesenen Daten wurden zufällig durch die entsprechende scikitlearn Funktion in einem Verhältnis von 50 % in Trainings- und Testdaten aufgeteilt. Diese Aufteilung wurde für jeden neuen Trainingsdurchlauf erneut durchgeführt. Die SVM wurden

anschließend auf den generierten Trainingsdaten mit einem polynomialen Kernel, einem ‚degree‘ von 2 und einem ‚coefficient‘ von 1 durchgeführt. Die Qualität der trainierten SVM wurde anschließend an den zuvor generierten Testdaten evaluiert und in Genauigkeit (%) bemessen. Diese gab an, wieviel Prozent der zu evaluierenden Datenpunkte korrekt ihrem entsprechenden label, sprich dem gesuchten Zelltyp, zugeordnet wurden. Als Visualisierung hiervon konnten die Entscheidungsgrenzen, welche die SVM verwendet hatte, graphisch ausgegeben werden.

3.2.7 Aufbau und Training des künstlichen neuronalen Netzwerks (KNN)

Die verwendete KNN-Architektur basiert auf der Architektur des VGG-Netzwerks [120]. Das KNN wurde für die Klassifizierung der Zelllinien und der Klassifizierung der MSC-Differenzierung individuell optimiert.

3.2.7.1 Aufbau

Das verwendete KNN wurde in der weit verbreiteten Python-Programmiersprache entwickelt und trainiert. Hierfür wurde v. a. das Pytorch-Paket verwendet [121]. Wie bereits erwähnt, wurde das Netzwerk für den jeweiligen Verwendungszweck (Zelllinien (Abbildung 14 und Abbildung 15) bzw. MSC-Entwicklung (Abbildung 16 und Abbildung 17)) entsprechend angepasst.

Architektur für die Klassifizierung von Zelllinien: Das KNN wurde durch eine Faltungsschicht mit 16 Eingabeneuronen initiiert. Dieser folgte eine maxpooling-Schicht, gefolgt von einer weiteren Faltungsschicht mit 64 Eingabeneuronen. Hier-nach folgte erneut eine maxpooling-Schicht. Nach einer weiteren Faltungsschicht mit 128 Eingabeneuronen und einer weiteren maxpooling-Schicht folgten abschließend eine Faltungsschicht mit 256 Eingabeneuronen. Auch wenn hier nicht jedes Mal erwähnt, folgte jeder Faltungsschicht eine Normalisierungsschicht, welche mittels Instancenormalization ausgeführt wurde. Alle Faltungsschichten nutzten eine 3x3-Kernel und ReLU als Aktivierungsfunktion. Abschließend wurde eine *Flatten*-Schicht eingebaut, bevor das Ergebnis klassifiziert wurde. Für die Klassifizierung nutze das Netz drei fully connected / dense layer mit 1024, 1024 und 512 Eingabeneuronen. Auch hier wurde ReLu als Aktivierungsfunktion verwendet. Für die

Klassifikation wurde eine drop-out-Rate von 0,25 verwendet, um die Überangepasstheit (engl. *Overfitting*) soweit wie möglich zu reduzieren. Für das gesamte Training wurde AdamW mit einer Lernrate von $1 \cdot 10^{-3}$ und einem weight decay von 0,15 als Optimierungsfunktion verwendet. Abschließend nutzte das Netzwerk die Softmax-Aktivierungsfunktion, um die Ausgabewerte in eine Prozentangabe umzuwandeln.

Architektur für die Klassifizierung von MSCs: Das KNN wurde durch eine Faltungsschicht mit 10 Eingabeneuronen initiiert. Dieser folgte eine maxpooling-Schicht, gefolgt von zwei weiteren Faltungsschichten, mit je 32 Eingabeneuronen. Hiernach folgte erneut eine maxpooling-Schicht. Nach zwei weiteren Faltungsschichten mit je 64 Eingabeneuronen, und einer weiteren maxpooling-Schicht folgten abschließend eine Faltungsschicht mit 128 Eingabeneuronen und eine mit 256 Eingabeneuronen. Auch wenn hier nicht jedes Mal erwähnt, folgte jeder Faltungsschicht eine Normalisierungsschicht, welche mittels Instancenormalization ausgeführt wurde. Alle Faltungsschichten nutzten eine 3x3-Kernel und ReLU als Aktivierungsfunktion. Abschließend wurde eine Flatten-Schicht eingebaut, bevor das Ergebnis klassifiziert wurde. Für die Klassifizierung nutze das Netz drei fully connected / dense layer mit 512, 512 und 256 Eingabeneuronen. Auch hier wurde ReLU als Aktivierungsfunktion verwendet. Für das gesamte Training wurde AdamW mit einer Lernrate von $1 \cdot 10^{-3}$ und einem weight decay von 0,15 als Optimierungsfunktion verwendet. Abschließend nutze das Netzwerk die Softmax-Aktivierungsfunktion, um die Ausgabewerte in eine Prozentangabe umzuwandeln.

3.2.7.2 Training des KNN

Python lud die Daten als *.png Dateien in die erzeugte Umgebung. Dort konnten die Daten nach Zelltypen gefiltert werden. Anschließend teilte das System die Daten in Trainings (75 %) und Testdaten (25 %) auf. Die Trainingsdaten wurden erneut in Trainings (80 %) und Validierungsdaten (20 %) aufgeteilt. Durch eine eigens hierfür generierte Funktion konnte sichergestellt werden, dass jeder enthaltene Zelltyp in allen Datensätzen vorkam. Wenn notwendig, war es auch möglich, die Datenmatrize von 300x300-Datenpunkten zu beschneiden, um z. B. den Medienpeak zu entfernen. Außerdem war es möglich, die Anzahl der verwendeten Augmentierungen einzustellen. Wenn folglich Daten geladen wurden, welche 40 Mal augmentiert wurden, konnte jeder Augmentierungsfaktor zwischen 0 und 40 gewählt werden. Es konnte definiert werden, wie häufig die Genauigkeit des Trainings erhoben werden sollte, ob dies nach

jeder Trainingsepoche erfolgen sollte oder nach einer definierten Zahl an Iterationen.

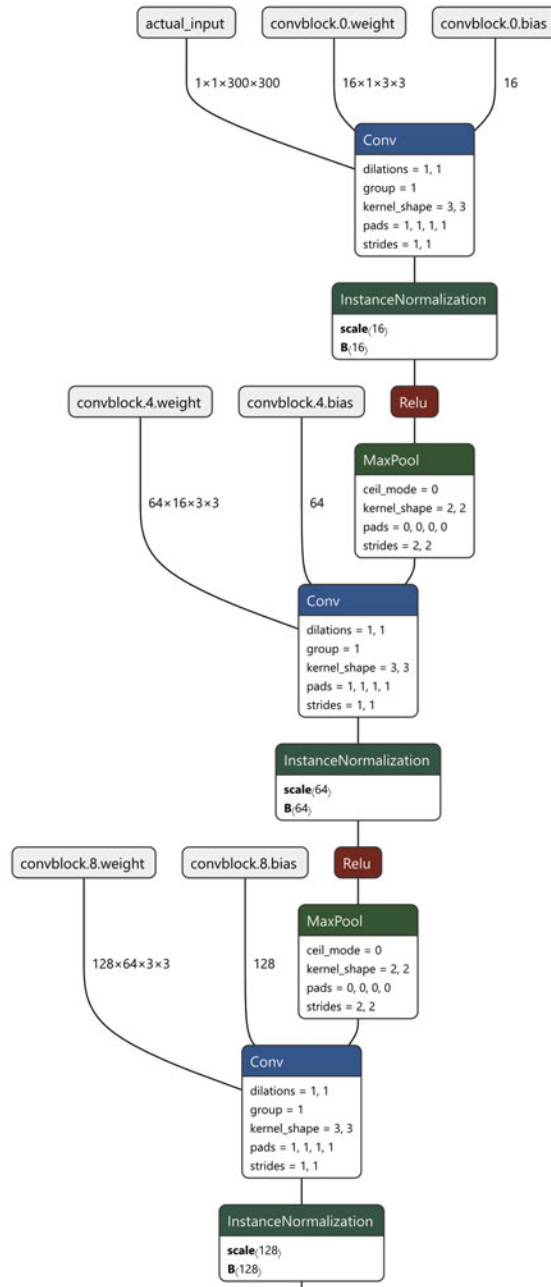


Abbildung 14 – Etablierte KNN-Architektur zur Klassifikation von Zelllinien (Teil 1): Für die Klassifikation der Zelllinien wurde ein modifiziertes VGG-Netzwerk verwendet, welches eine Kombination aus vier Faltungsschichten (engl. convolutional layer) verwendete. Diese waren durch Normalisierungsschichten voneinander getrennt. Anschließend wurden die analysierten Daten durch vier voll-verbundene (engl. fully-connected / dense) Schichten mit einer drop-out-Rate von 0,75 entsprechend der gewünschten Klassen klassifiziert. Für alle Schichten wurde ReLU als Aktivierungsfunktion verwendet und Adamax mit einer Lernrate von $1 \cdot 10^{-3}$ als Optimierungsfunktion.

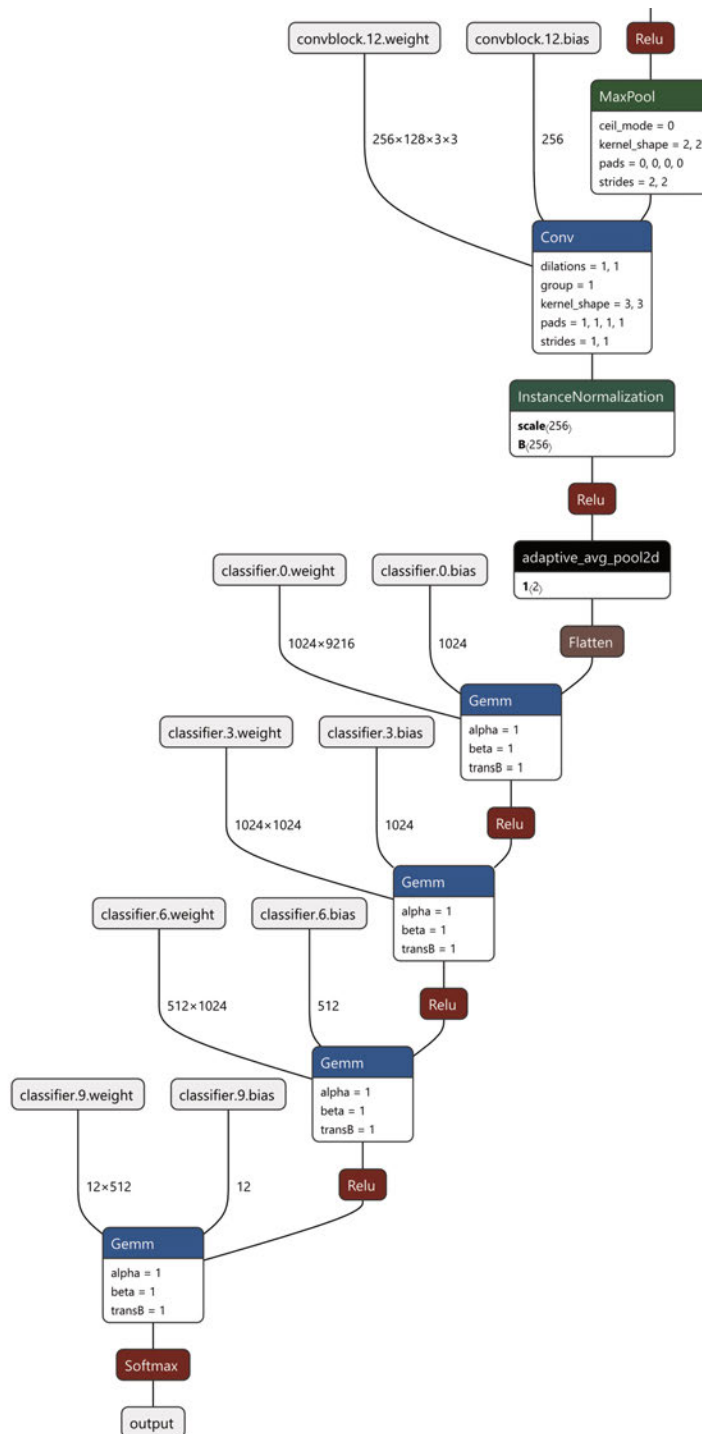


Abbildung 15 – Etablierte KNN-Architektur zur Klassifikation von Zelllinien (Teil 2): Für die Klassifikation der Zelllinien wurde ein modifiziertes VGG-Netzwerk verwendet, welches eine Kombination aus vier Faltungsschichten (engl. convolutional layer) verwendete. Diese waren durch Normalisierungsschichten voneinander getrennt. Anschließend wurden die analysierten Daten durch vier voll-verbundene (engl. fully-connected / dense) Schichten, mit einer drop-out-Rate von 0,75 entsprechend der gewünschten Klassen klassifiziert. Für alle Schichten wurde ReLu als Aktivierungsfunktion verwendet und Adamax mit einer Lernrate von $1 \cdot 10^{-3}$ als Optimierungsfunktion.

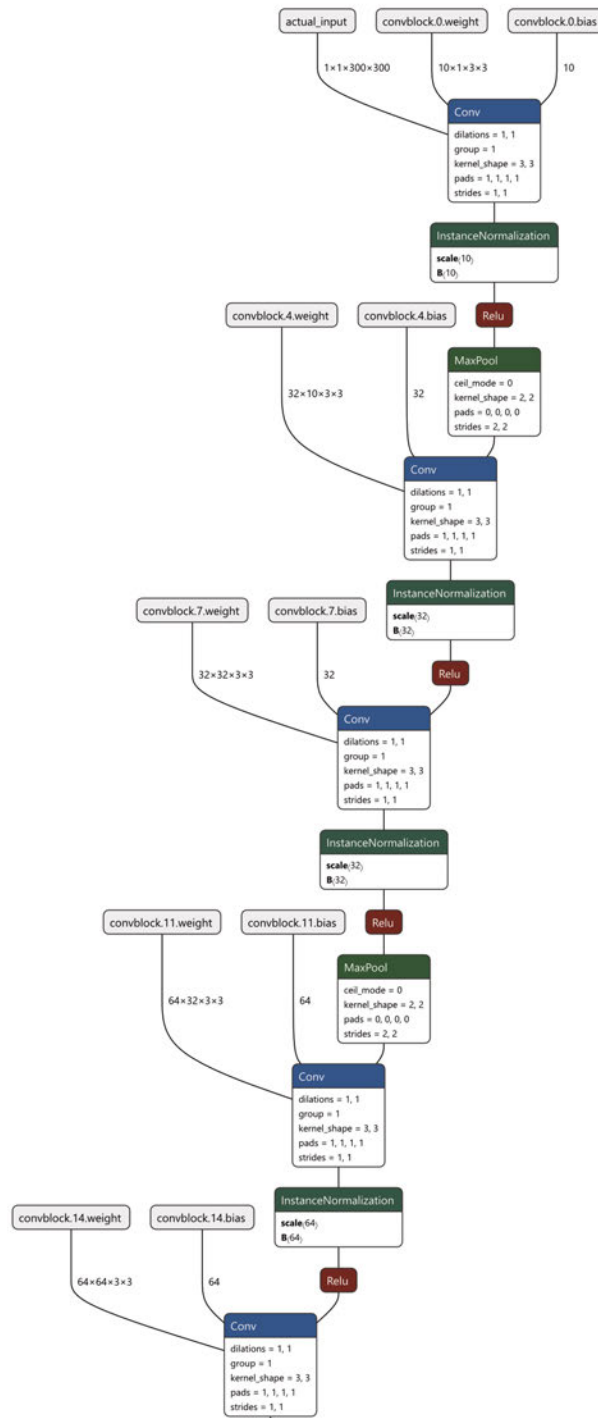


Abbildung 16 – Etablierte KNN-Architektur zur Klassifikation von MSCs (Teil 1): Für die Klassifikation der MSCs wurde ein modifiziertes VGG-Netzwerk verwendet, welches eine Kombination aus sieben Faltungsschichten (engl. convolutional layer) verwendete. Diese waren durch Normalisierungsschichten voneinander getrennt. Anschließend wurden die analysierten Daten durch vier voll-verbundene (engl. fully-connected / dense) Schichten entsprechend der gewünschten Klassen klassifiziert. Für alle Schichten wurde ReLu als Aktivierungsfunktion verwendet und Adamax mit einer Lernrate von $1 \cdot 10^{-3}$ als Optimierungsfunktion.

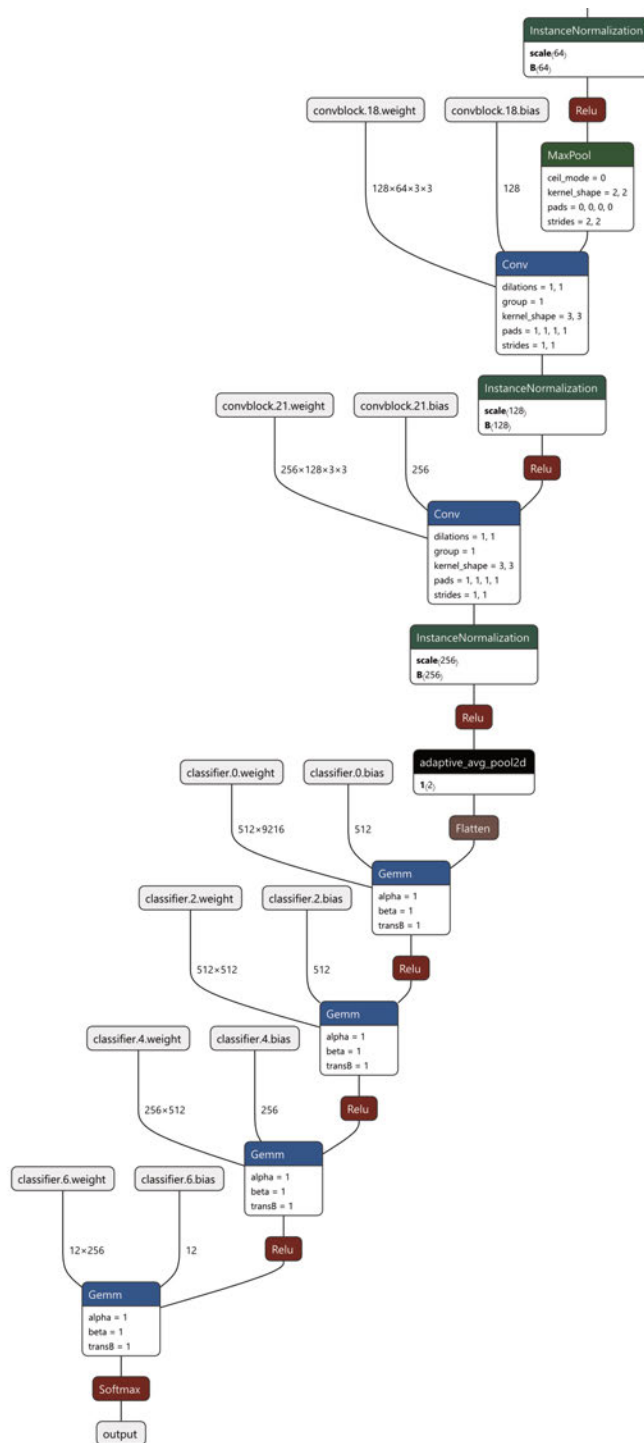


Abbildung 17 – Etablierte KNN-Architektur zur Klassifikation von MSCs (Teil 2): Für die Klassifikation der MSCs wurde ein modifiziertes VGG-Netzwerk verwendet, welches eine Kombination aus sieben Faltungsschichten (engl. convolutional layer) verwendete. Diese waren durch Normalisierungsschichten voneinander getrennt. Anschließend wurden die analysierten Daten durch vier voll-verbundene (engl. fully-connected / dense) Schichten entsprechend der gewünschten Klassen klassifiziert. Für alle Schichten wurde ReLu als Aktivierungsfunktion verwendet und Adamax mit einer Lernrate von $1 \cdot 10^{-3}$ als Optimierungsfunktion.

Das Training für die Zelllinien wurde mit einer Batchgröße von 75 und für die MSCs mit 5 durchgeführt. Es wurden Trainings sowohl mit 12 als auch mit 25 Epochen durchgeführt. Alle Durchläufe wurden als Tensorboard-Dateien exportiert, was es ermöglichte, den Trainingsverlauf auch im Nachhinein noch nachvollziehen zu können.

Die Performanz des trainierten Modells wurde durch den Verlauf von dessen Verlustfunktion ebenso wie dessen Genauigkeit erhoben. Nach dem Training wurde das Netz auf den zuvor extrahierten Trainingsdaten evaluiert und die Genauigkeit als finales Bewertungskriterium herangezogen.

4 Ergebnisse

4.1 Validierung der MR-Parameter und -Methodik

4.1.1 Pulswinkelbestimmung

Wie bereits in der Einleitung erwähnt, wurde eine kombinierte T_1/T_2 -Sequenz verwendet, um zweidimensionale Korrelationsspektren der zu untersuchenden Probe aufzunehmen. Neben den speziellen zeitlichen Parametern der Sequenz, wie TE oder TI, waren v. a. die Dauer der 90°- und 180°-Pulse (Tabelle 12) von charakteristischer Bedeutung für die Messung. Die Werte für die Pulswinkel wurden anhand eines vom Hersteller des MR-Scanners bereitgestellten Öl-Phantoms erhoben und anschließend für sämtliche weiteren Messungen verwendet. Basierend auf diesen Werten wurde vor jeder Messung automatisch eine Pulswinkelbestimmung für die jeweilige Probe durchgeführt.

Tabelle 12: Erhobene Werte für die Pulsdauer und Amplitude der MR-Messung.

Bezeichnung	Wert
Pulsamplitude	120 μ T
Dauer 90°-Puls	37,27 μ s
Dauer 180°-Puls	74,63 μ s

Die *Shim*-Werte wurden ebenfalls mit dem Ölphantom erhoben und anschließend für sämtliche weitere Messungen global verwendet.

4.1.2 Phantommessungen

Um die Stabilität des Messsystems sowie dessen Reproduzierbarkeit zu evaluieren, wurden Messungen mit standardisierten Agarose-Kontrastmittel-Phantomen durchgeführt (siehe 3.2.2.4). Da diese auf die zu erwartenden Messwerte der verwendeten Proben abgestimmt sein mussten, wurden mehrere Versuchsreihen angestellt, bis die passende Kombination aus Agarose und dem Kontrastmittel Dotagraf[®] bestimmt

wurde (Abbildung 18 b). Als Vergleich dienten Messungen mit CHO (*chinese hamster ovary*)-Zellen (Abbildung 18-a). Die etablierten Phantome wurden fortan als Agarose-Dotagraf-Zellphantome (ADZP) bezeichnet.

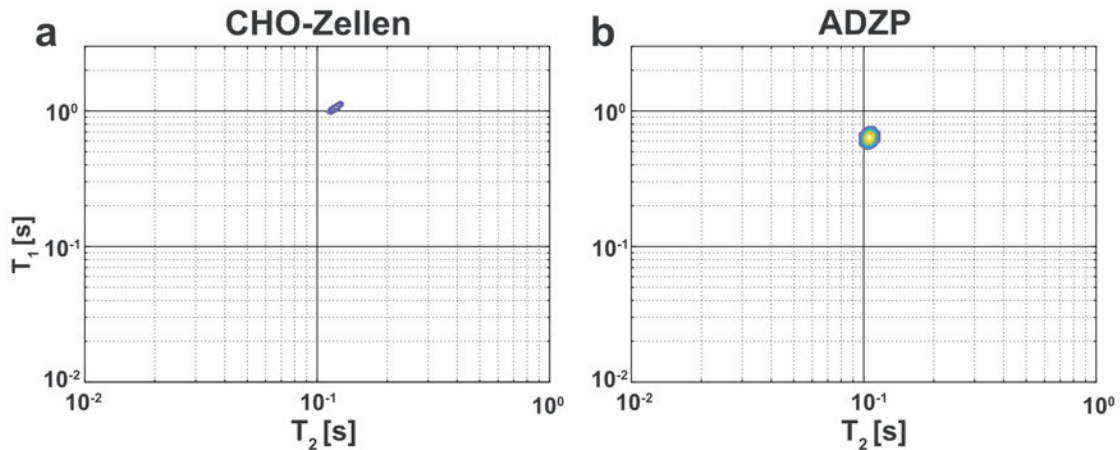


Abbildung 18 – Etablierung von Agarose-Dotagraf[®]-Zellphantomen (ADZP): Als Richtwert dienten Messungen mit CHO-Zellen (a). Basierend darauf wurden mehrere Versuchsreihen durchgeführt, bis die passende Konzentration von Agarose und Dotagraf gefunden war. Das erzeugte Signal des ADZPs (b) lag bei etwas kürzeren T_1 - bzw. T_2 -Zeiten als die Vergleichswerte der CHO-Messung. Die verwendeten Konzentrationen lagen bei 0,6 Gew.-% Agarose und 0,25 μM Dotagraf[®], welche in RPMI-1640 Medium gelöst wurden.

Damit sich das Signal-zu-Rausch-Verhältnis (engl. *signal to noise ratio* SNR) verbesserte, wurde ein rauscharmer Signalverstärker (engl. *low noise amplifier* LNA) verwendet. Um dessen Effekt zu quantifizieren, wurden ADZPs mit und ohne LNA vermessen und jeweils deren SNR bestimmt (siehe 3.2.2.5). Beim Vergleich vom SNR mit LNA zum SNR ohne LNA ergab sich ein Verhältnis von 1,46. Hierdurch ergab sich eine Zeitersparnis um den Faktor 2,13. Aufgrund dieser deutlichen Einsparung wurde soweit nicht anders gekennzeichnet oder erwähnt bei jeder Messung der LNA verwendet.

Mithilfe der etablierten ADZPs wurde anschließend das Messsystem genauer charakterisiert. Dazu wurden zwölf separate ADZPs mit je 50 μl hergestellt und nacheinander vermessen, um die Stabilität des Messsignals zu evaluieren. Für die bessere Vergleichbarkeit wurde das 2D-Signal in der T_1 - bzw. T_2 -Dimension separat aufgetragen und die entsprechenden Replikate als Zeilen in einem sog. mesh plot dargestellt (Abbildung 19 a). Diese Art der Darstellung erleichterte es, Unterschiede zwischen den aufgetragenen Proben eindeutiger und schneller zu identifizieren. Hierdurch wurde ersichtlich, dass die ADZPs in T_1 und T_2 jeweils einen deutlichen Signalpeak zeigten und bei kürzeren Relaxationszeiten nur wenige Signalpeaks aufwiesen.

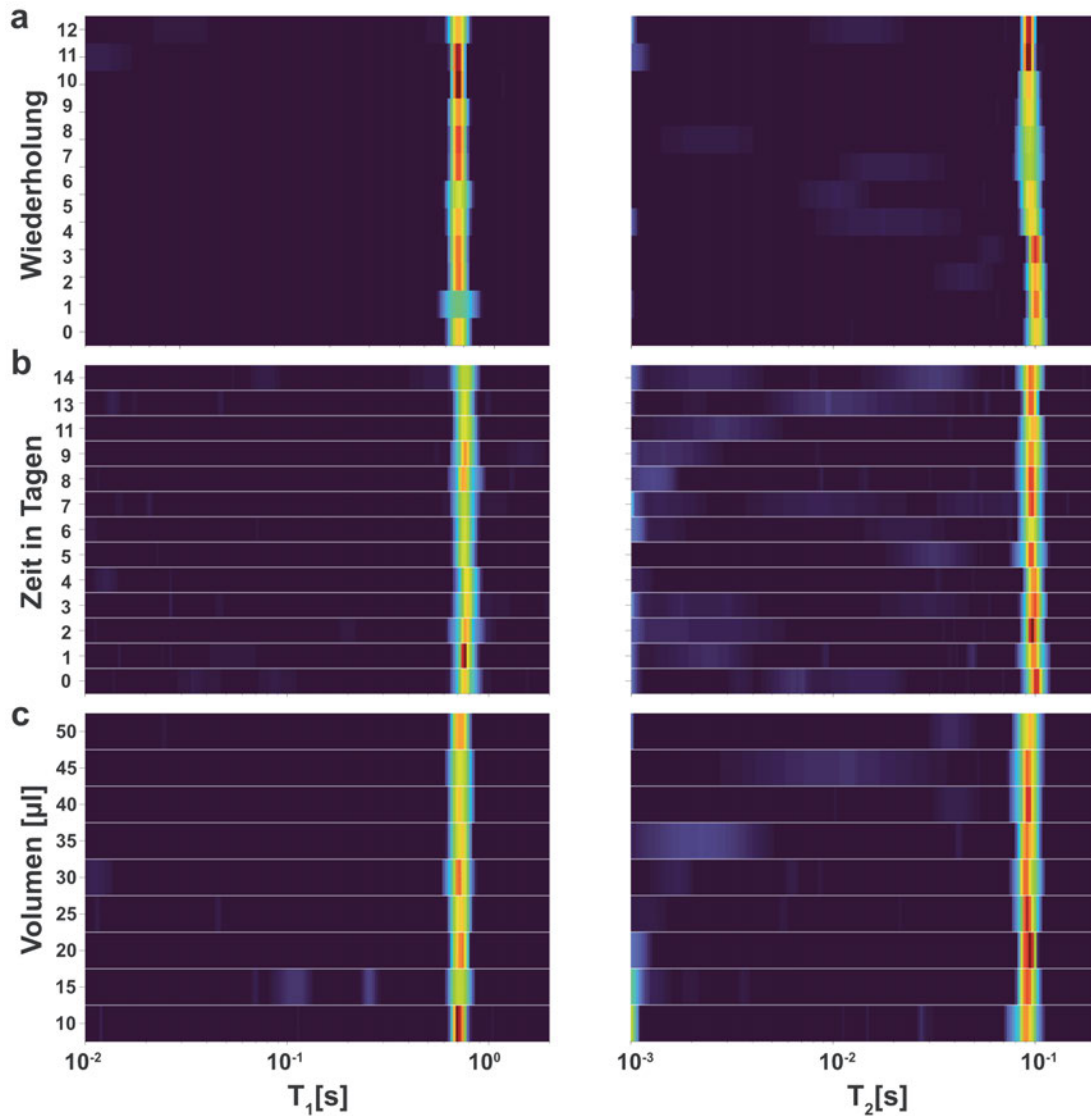


Abbildung 19 – Stabilitätsmessungen am verwendeten MR-Messsystem anhand der entwickelten ADZPs: Für die Evaluation der Reproduzierbarkeit der Messungen wurden zwölf ADZPs mit je 50 µl hergestellt und nacheinander vermessen. Die Darstellung der Ergebnisse in der T_1 - bzw. T_2 -Dimension zeigte eine leichte Fluktuation des Signals um einen stabilen Mittelwert (a). Für die Untersuchung der Langzeitreproduzierbarkeit wurden mehrere ADZPs hergestellt und über einen Zeitraum von zwei Wochen täglich vermessen (b). In der Darstellung wurden die Replikate pro Messtag kombiniert, sodass die resultierende Darstellung die Gesamtheit der Messungen widerspiegelt. Wie schon bei a, zeigte sich hier ein stabiler Mittelwert für beide Messdimensionen, um welchen die Breite des Signalpeaks leicht fluktuierte. Basierend auf diesen Ergebnissen wurde eine gute Reproduzierbarkeit der Daten über einen Zeitraum von mindestens zwei Wochen prognostiziert. Um die Unabhängigkeit des Messsignals vom verwendeten Probenvolumen zu belegen, wurde eine Reihe an ADZPs hergestellt, deren Volumen in 5 µl-Schritten zwischen 10 µl und 50 µl lagen (c). Pro Volumen wurden drei unabhängige Replikate vermessen und deren Signale ebenfalls kombiniert. Das Signal zeigte erneut einen positionsstabilen Mittelwert, welcher in der Breite leicht fluktuierte. Daraus bestätigte sich die ursprüngliche Annahme, dass das Signal unabhängig vom Probenvolumen war. Ebenso bestätigte diese die Verwendung der definierten Pulse und Shim-Werte als vordefinierte Parameter. Da die Messung von Proben mit weniger als 10 µl keine verwertbaren Ergebnisse erzeugte, lag das minimal detektierbare Volumen bei < 10 µl. Abbildung adaptiert von [122].

Um auch die Langzeitstabilität des Messsystems zu evaluieren, wurden frische ADZPs hergestellt und über einen Zeitraum von 14 Tagen täglich neu vermessen

(Abbildung 19 b). Für die Darstellung des Signals wurden die Replikate pro Tag kombiniert, sodass das resultierende Signal die Gesamtheit aller Proben widerspiegelt. Es zeigte sich, dass sich der große Signalpeak in seiner Position nur schwach veränderte, auch wenn seine Ausdehnung in T_1 bzw. T_2 leicht fluktuierte. Ebenfalls fluktuierte das Signal im kürzeren Messbereich. Die Stabilität des Systems begründet sich in den geringen Schwankungen und verdeutlicht die Reproduzierbarkeit über einen Zeitraum von 14 Tagen.

Neben der Reproduzierbarkeit wurde auch die Abhängigkeit des Signals zum Probenvolumen untersucht. Dazu wurden ADZPs mit unterschiedlichen Volumina hergestellt. Pro Volumen wurden je drei unabhängige Replikate vermessen und das Signal ebenfalls kombiniert. Auch hier zeigte sich ein gleichbleibendes Signal über alle erhobenen Messpunkte hinweg (Abbildung 19 c). Das Signal fluktuierte ebenfalls leicht um einen stabilen Mittelpunkt. Basierend auf den Beobachtungen wurde keine Abhängigkeit des Signals vom Probenvolumen festgestellt. Unterhalb von 10 μl konnten keine verwertbaren Daten aufgenommen werden, wodurch die Volumensensitivität des Systems bei $< 10 \mu\text{l}$ begründet wird.

Damit die qualitativen Aussagen mit quantitativen belegt werden konnten, wurden die gewichteten Schwerpunkte für die aufgenommenen Spektren berechnet. Die hieraus resultierten Werte für T_1 und T_2 wurden dann mittels Varianzanalyse (engl. *analysis of variance* - ANOVA) auf ihre Signifikanz hin untersucht (Alpha-Niveau 5%). Diese Analyse ergab, dass die Spektren, welche über einen Zeitraum von 14 Tagen aufgenommen wurden, keine signifikanten Unterschiede aufwiesen, weder für T_1 (Abbildung 20 a1) noch für T_2 (Abbildung 20 a2). Auch die Analyse der Volumenstabilität ergab weder für T_1 (Abbildung 20 b1) noch für T_2 (Abbildung 20 b2) signifikante Unterschiede. Diese Analyse belegte, dass die zuvor getroffenen qualitativen Aussagen valide waren. Demnach lieferte die untersuchte Methode belegbar stabile Messwerte.

4.1.3 Erste Zell- und Medienmessungen

Nachdem die ADZP Messungen die Stabilität und Reproduzierbarkeit des Messsystems erfolgreich bestätigt hatten, wurden die ersten Zell- und Medientestungen durchgeführt. Bedingt durch die Sequenzparameter waren die erzeugten Daten auf einem Gitter zwischen 1 ms und 15 s in beiden Dimensionen T_1 und T_2 skaliert. Die Messungen zeigten, dass die relevanten Daten jedoch im Bereich zwischen 1 und 5 Sekunden für T_1 und 0,1 bis 4 Sekunden für T_2 lagen. Daher wurden alle nachfolgen-

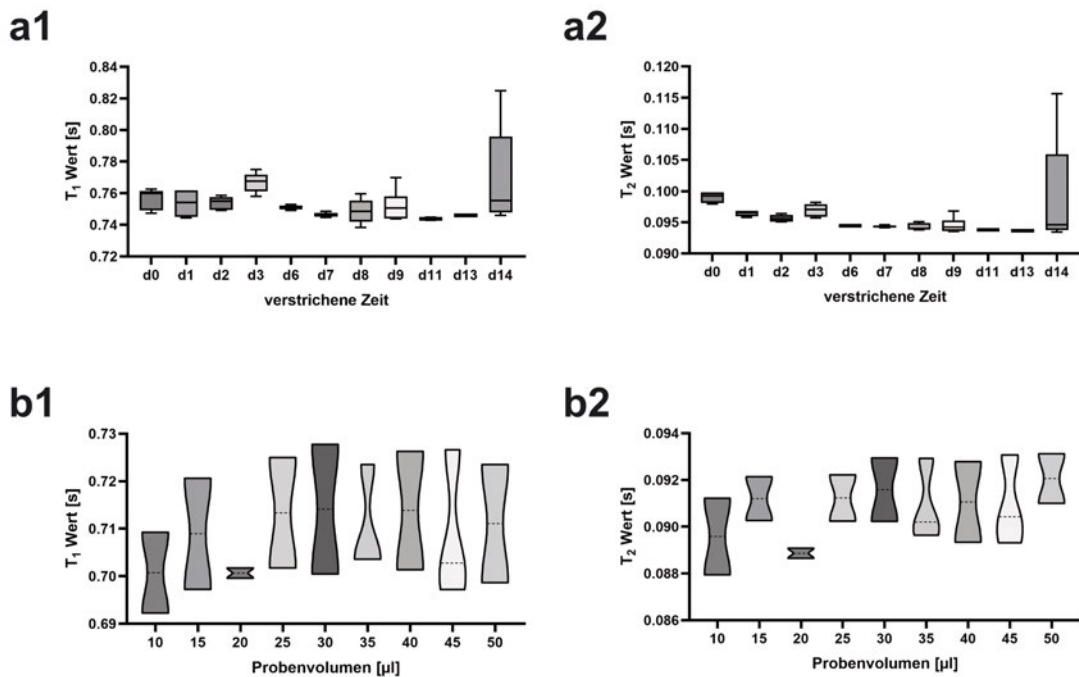


Abbildung 20 – Statistische Analyse zu den durchgeführten ADZP Stabilitätsmessungen: Für die in Abbildung 19 aufgeführten Stabilitätsmessungen wurden die gewichteten Schwerpunkte berechnet, um die qualitativen Aussagen zu quantifizieren. Für die Messung der Langzeitstabilität (a) ergaben sich weder für die T₁-(a-1) noch für die T₂-Werte(a-2) statistisch signifikante Unterschiede. Auch für die Analyse des Probenvolumen (b) ergaben sich weder für die T₁-(b-1) noch für die T₂-Werte(b-2) signifikante Unterschiede. Dies bestätigte die zuvor aufgestellte Hypothese, dass die untersuchte Messmethode zeitlich stabile Ergebnisse lieferte. Auch bestätigte sich hierdurch, dass das Probenvolumen keinen Einfluss auf die erhobenen Messwerte nahm. Für die statistische Analyse wurde eine Varianzanalyse (engl. analysis of variance - ANOVA) gepaart mit einem Tukey's test für den direkten Vergleich der zu untersuchenden Gruppen durchgeführt. Es wurde ein Alpha-Niveau von 5% angenommen. Abbildung adaptiert von [122].

den Darstellungen zum Zweck der Übersichtlichkeit der 2D-Korrelationspektren auf diesen Bereich zugeschnitten. Beim Vergleich zwischen Proben, welche nur das Zellkulturmedium enthielten (Abbildung 21 a), und Proben, welche sowohl Medium als auch Zellen enthielten (Abbildung 21 b), wurde ersichtlich, dass die Medienproben nur einen einzelnen großen Signalpeak erzeugten, welcher nachfolgend als Medienpeak bezeichnet wurde. Proben, welche neben dem Kulturmedium auch Zellen enthielten, wiesen neben dem Medienpeak auch weitere Peaks bei kürzeren T₁- bzw. T₂-Zeiten auf. Diese wurden fortan als Zellpeak(s) bezeichnet. Die Medienpeaks lagen typischerweise im Bereich zwischen 3 und 4 Sekunden für T₁ und 2 und 3 Sekunden für T₂. Die Medienpeaks lagen i. d. R. im Bereich von 0,8 und 2 Sekunden für T₁ und 0,1 bis 0,4 Sekunden für T₂.

Anhand von mehreren Messungen mit CHO-Zellen wurde anschließend die minimal mögliche Zellzahl ermittelt, bis zu welcher das Messsignal stabil blieb. Dazu wurden mehrere Proben mit Zellzahlen zwischen $1 \cdot 10^6$ und $1 \cdot 10^7$ vermessen und deren

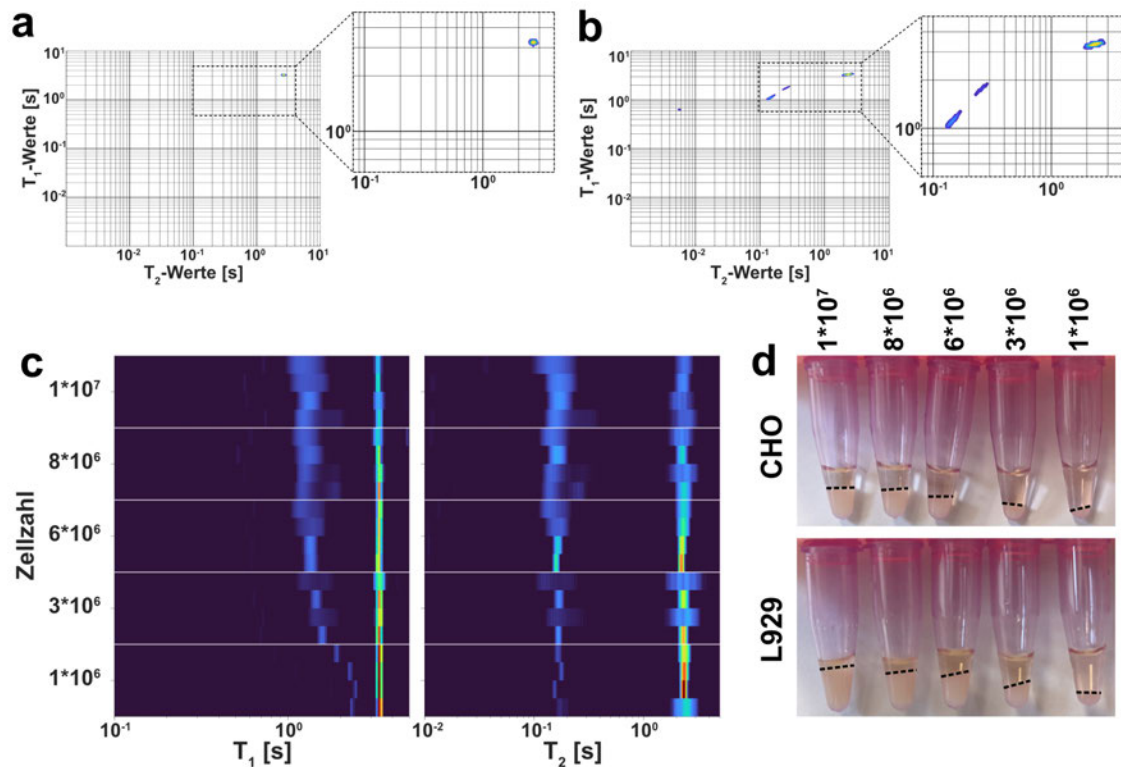


Abbildung 21 – Erste Zellmessungen zeigten zwei Gruppen an Signalpeaks und ermöglichten die Bestimmung der minimal messbaren Zellzahl: Erste Messungen mit C2C12-Zellen zeigten, dass sich das detektierbare Signal in zwei deutlich abgegrenzten Bereichen gliedert. Ein Bereich lag bei 3 - 4 Sekunden in T_1 und 2 - 3 Sekunden in T_2 . Das Signal trat bei jeder Messung auf, auch wenn keine Zellen enthalten waren (a). Aus diesem Grund repräsentiert es die Mediumskomponente der Probe und wurde dementsprechend als Mediumspeak bezeichnet. Sobald Zellen in der Probe enthalten waren (b), traten zusätzlich zum Mediumspeak weitere Signalpeaks auf, welche jedoch bei kürzeren T_1 - (1 - 3 Sekunden) und T_2 - ($1 \cdot 10^{-1}$ - $4 \cdot 10^{-1}$ Sekunden) Zeiten lagen. Da diese Signalanteile ausschließlich dann auftraten, wenn Zellen in der Probe enthalten waren, wurden sie als ‚Zellpeak(s)‘ bezeichnet. Anhand einer Messreihe mit CHO-Zellen wurde die minimal detektierbare Zellzahl bestimmt. Dazu wurden Proben mit $1 \cdot 10^6$, $3 \cdot 10^6$, $6 \cdot 10^6$, $8 \cdot 10^6$ und $1 \cdot 10^7$ Zellen je Probe vermessen und deren Signal in T_1 und T_2 verglichen. Der Signalverlauf zeigte (c), dass das Signal bei $> 3 \cdot 10^6$ in seiner Position sehr stabil blieb, auch wenn die Streuung in der Breite zunahm. Bei Proben mit $\geq 3 \cdot 10^6$ Zellen streuten die Zellpeaks weniger in T_1 - und T_2 -Richtung, bekamen in T_1 jedoch deutlich längere Relaxationszeiten. Außerdem nahm die Amplitude des Mediumspeak deutlich zu. Beim Vergleich unterschiedlicher Zelllinien z. B. CHO und L929 (d), wurde deutlich, dass die unterschiedlichen Zellzahlen bei gleichem Probenvolumen unterschiedlich große Zellpellets und dementsprechend unterschiedlich großem Medienüberstand erzeugten (gestrichelte Linien in d markieren Zellen-Medien-Grenze). Basierend auf diesen Beobachtungen wurde die minimal detektierbare Zellzahl auf $\geq 3 \cdot 10^6$ festgesetzt. Abbildung adaptiert von [122].

Signale in T_1 und T_2 aufgetragen (Abbildung 21 c). Dabei zeigte sich, dass die beobachteten Zell- und Medienpeaks ab einer Zellzahl von $3 \cdot 10^6$ augenscheinlich stabil blieben (Abbildung 21 c). Unterhalb dieser kritischen Zellzahl driftete die Zellpeaks deutlich in Richtung des Medienpeaks, dessen Intensität zudem sichtbar zunahm. Daher wurde eine Zellzahl von $\geq 3 \cdot 10^6$ Zellen pro Probe festgelegt. Wenn nicht anderweitig definiert, wurden dementsprechend $5 \cdot 10^6$ Zellen pro Messung verwendet, da dies dem Kriterium von $\geq 3 \cdot 10^6$ Zellen entsprach und außerdem für alle üblichen Zelltypen eine realistische Zellzahl darstellt, welche in der Zellkultur erreicht werden kann. Abhängig vom verwendeten Zelltyp und dessen Größe, korreliert diese

minimale Zellzahl mit dem zuvor beschriebenen minimalen Detektionsvolumen (siehe 4.1.2).

Um die qualitativen Aussagen quantitativ zu belegen, wurden auch hier die gewichteten Schwerpunkte der betrachteten Spektren berechnet und die entsprechenden T_1 - bzw. T_2 -Zeiten statistisch miteinander verglichen. Auch hier wurde eine ANOVA (Alpha-Niveau 5 %) gepaart mit einem Tukey's Test für diese Analyse verwendet. Die Analyse zeigte, dass die T_1 -Werte für Proben, welche $1 \cdot 10^6$ Zellen enthielten, signifikant länger waren, als für Proben, welche mehr als $1 \cdot 10^6$ Zellen enthielten (Abbildung 22a). Zwischen den anderen untersuchten Zellzahlen konnte kein statistisch signifikanter Unterschied detektiert werden. Bei der Untersuchung der T_2 -Werte, konnten im Gegensatz zu T_1 keine signifikanten Unterschiede detektiert werden (Abbildung 22b). Diese Analysen bestätigten die zuvor getroffenen, qualitativen Aussagen und legten den Schluss nahe, dass eine Zellzahl von $\geq 3 \cdot 10^6$ stabile, reproduzierbare Messungen zuließen.

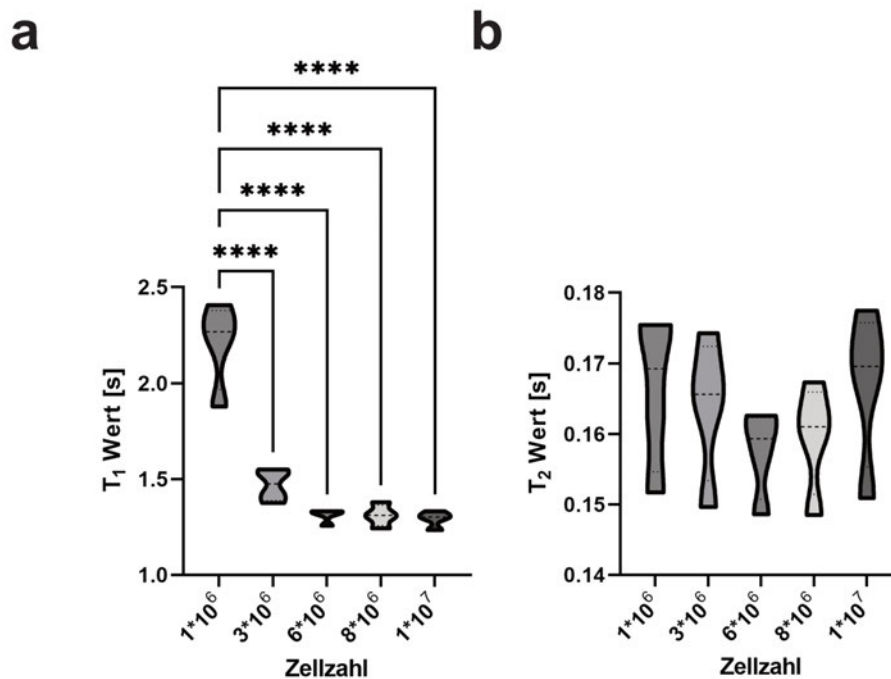


Abbildung 22 – Statistische Analyse der Signalstabilität relativ zur gemessenen Zellzahl: Für die Analyse wurden die gewichteten Schwerpunkte der untersuchten Spektren berechnet und mittels ANOVA (Alpha-Niveau 5 %) gepaart mit einem Tukey's Test statistisch verglichen. Diese Analysen zeigten, dass Proben, welche $1 \cdot 10^6$ Zellen enthielten, signifikant längere T_1 -Zeiten aufwiesen im Vergleich zu Proben mit $\geq 3 \cdot 10^6$ Zellen (a). Proben mit $\geq 3 \cdot 10^6$ Zellen, zeigten in der statistischen Analyse keine signifikanten Unterschiede (b). Beim Vergleich der T_2 -Zeiten konnten keine statistisch signifikanten Unterschiede detektiert werden. Diese Ergebnisse legten den Schluss nahe, dass die zuvor getroffenen, qualitativen Aussagen zutreffend waren und eine Zellzahl von $\geq 3 \cdot 10^6$ reproduzierbare Ergebnisse lieferte. $n_b = 4$; ns: nicht signifikant; *: p-Wert $\leq 0,05$; **: p-Wert $\leq 0,01$; ***: p-Wert $\leq 0,001$; ****: p-Wert $\leq 0,0001$

4.1.4 Zusammenfassung

Die gewählten Sequenz- und Messparameter erzeugten ein Signal, welches validiert durch Phantommessungen (siehe 4.1.2) sowohl reproduzierbar als auch stabil war. Erste Zellmessungen zeigten einen dem Medienanteil zuordenbaren Bereich, welcher als Medienpeak bezeichnet wurde. Der Signalbereich des Zellanteils wurde als Zellpeak(s) bezeichnet. Durch vergleichende Messungen wurde die Stabilität des Signals auf Zellzahlen von $\geq 3 \cdot 10^6$ Zellen pro Probe festgesetzt (siehe 4.1.3). Daraus resultierte eine Zellzahl von $5 \cdot 10^6$ Zellen, welche soweit nicht anderweitig gekennzeichnet für jede durchgeführte Zellmessung galt. Um die Darstellung der Daten übersichtlicher zu gestalten, wurde fortan das Korrelationsspektrum in den Abbildungen auf den relevanten Signalbereich von $T_1 = 1 - 5$ Sekunden und $T_2 = 0,1 - 4$ Sekunden zugeschnitten.

4.2 Etablierung einer Robotikplattform

Bedingt durch die Messsequenz dauerte die Messung einer Probe ca. 30 min. Unter optimalen Bedingungen limitierte dies die maximal mögliche Probenzahl pro 8 h Arbeitstag auf 16. Unter realistischen Bedingungen reduzierte sich diese Zahl auf etwa fünf bis sechs Proben pro Tag. Für eine zuverlässige Auswertung der erhobenen Daten mittels maschinellem Lernen war es jedoch wichtig, einen möglichst großen Datensatz an zellbasierten MR-Messungen zu erheben. Es gibt unterschiedliche theoretische Annahmen wie groß ein Datensatz sein sollte, um die enthaltenen Klassen mittels maschinellem Lernen gut vorherzusagen zu können. Manche Autoren gehen dabei von etwa dem Faktor 50 - 1000 der vorherzusagenden Klassen aus, was im Fall dieser Arbeit (10 vorherzusagende Zelltypen) mindestens 500 Datenpunkte bedeuten würde [123, 124]. Andere Quellen gehen wiederum vom 10- [125 127] oder sogar 50- [128] Fachen der im neuronalen Netzwerk enthaltenen Gewichte aus. Das heißt, hier würde die Größe des Datensatzes direkt mit dem der Architektur des verwendeten KNN korrelieren.

Um diesen Größenordnungen an Daten gerecht werden zu können, wurde entschieden, einen autonom operierenden Probenwechsler zu entwickeln, welcher die Proben eigenständig prozessiert und somit den möglichen Durchsatz von Proben pro Arbeitstag signifikant steigert.

4.2.1 Aufbau der Robotikplattform

Zentrale Komponente der Robotikplattform war der Roboterarm Braccio der Firma Arduino. Die Steuerung erfolgte ebenso wie die des MR-Scanners über die wissenschaftliche Programmiersprache MATLAB[®] (siehe 3.2.3). Die Probenröhrchen wurden in die Probenbox gestellt, in der an jeder Position ein Mikroschalter verbaut war (Abbildung 23 c). Dieser gab das Signal an MATLAB[®] weiter. Durch konische Aufsätze, wie z. B. beim Heizblock (Abbildung 23 d), wurden die Ungenauigkeiten des Roboterarmes ausgeglichen.

Durch das einfache Design der Plattform war es möglich, jeden auftretenden Fehler schnell und effizient zu beheben. Hierdurch erreichte der Roboterarm eine Zuverlässigkeit von 100 %, d. h. er verlor zu keinem Zeitpunkt eine Probe bzw. prozessierte alle eingegebenen Proben erfolgreich. Dies ermöglichte es, den Roboter ohne Aufsicht laufen zu lassen. Die Probenbox bot Platz für bis zu 12 Proben, deren vollständige Prozessierung ca. 6 h dauerte. Wenn die Aufbereitung der kultivierten Zellen entspre-

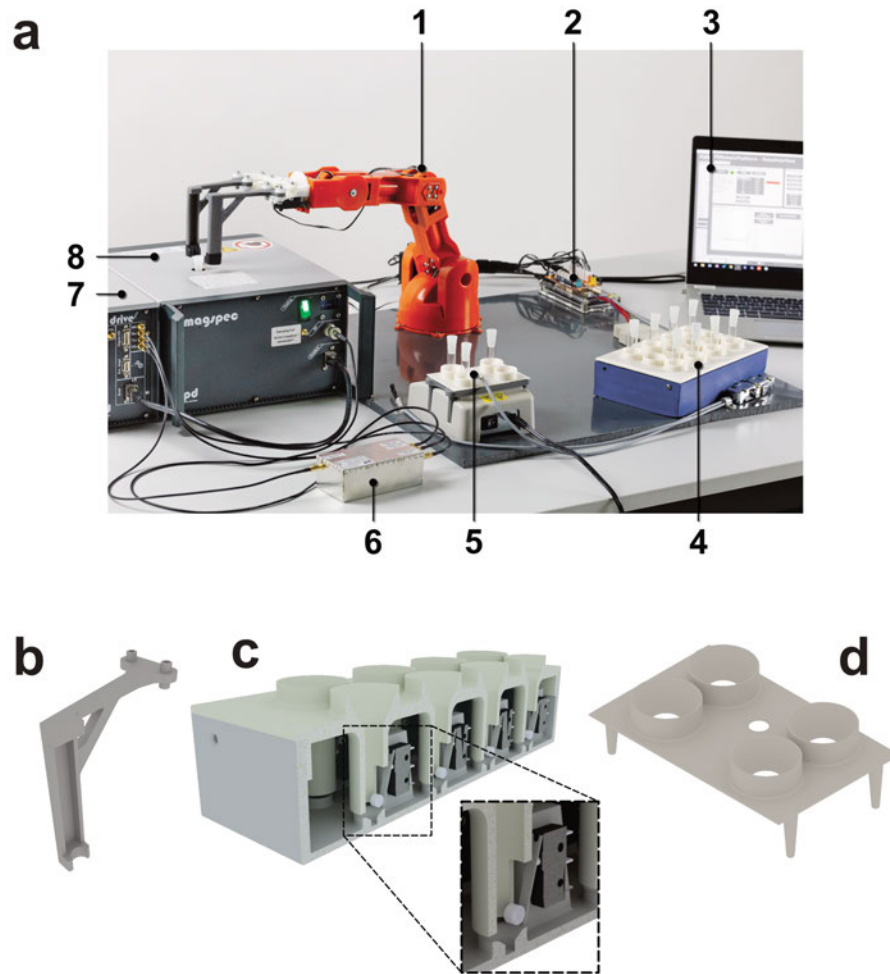


Abbildung 23 – Aufbau der verwendeten Robotikplattform: Als zentraler Roboterarm wurde der Braccio-Roboter der Firma Arduino verwendet (a-1). Dieser war – wie alle anderen Bauteile – auf einer Plastikplatte verschraubt. Der Greifer des Roboterarmes wurde neu entwickelt und mittels SLA-3D-Druck hergestellt (b). Gesteuert wurde der Roboterarm über einen Arduino Mega Mikrokontroller (a-2), welcher per USB mit einem Laptop verbunden war (a-3). Ebenso wie die Konsole des MR-Scanners (a-7), welche wiederum den Magneten des Scanners (a-8) steuerte, wurde auch der Mikrokontroller über MATLAB® angesteuert. Die Proben wurden in die Probenbox (a-4) eingegeben, in welcher Mikroschalter (c) verbaut waren, die das Signal an den Arduino Mega weiterleiteten. Um die Proben bereits auf 37 °C vorzutemperieren, wurden sie nach der Probenbox vom Roboterarm automatisch in einen Heizblock (a-5) gestellt. Auf diesem waren Führungskoni (d) angebracht, welche die Ungenauigkeiten des Roboterarmes kompensierten. Als einzige Komponente war der Signalverstärker (LNA) nicht auf der Plastikplatte verankert (a-6) sondern nur über Kabel mit der Konsole bzw. dem Magneten des MR-Scanners verbunden. Abbildung adaptiert von [122].

chend geplant wurde, war es so möglich, an einem Arbeitstag anstelle der vorherigen fünf bis sechs bis zu 24 Proben zu vermessen. Dies entsprach – verglichen mit einer manuellen Prozessierung – einer Steigerung um den Faktor 4 - 5.

4.2.2 Validierung der Probentemperatur

Da die Relaxationszeiten in der MR sehr temperaturabhängig sind, war es von essenzieller Bedeutung, dass die Proben temperatur der des Magneten entsprach, um mögliche Störeffekte zu minimieren. Zudem war es wichtig, dass die Probe entsprechend vortemperiert wurde, damit ein mögliches Equilibrieren der Probe nicht während der Messung stattfand und somit das Signal veränderte.

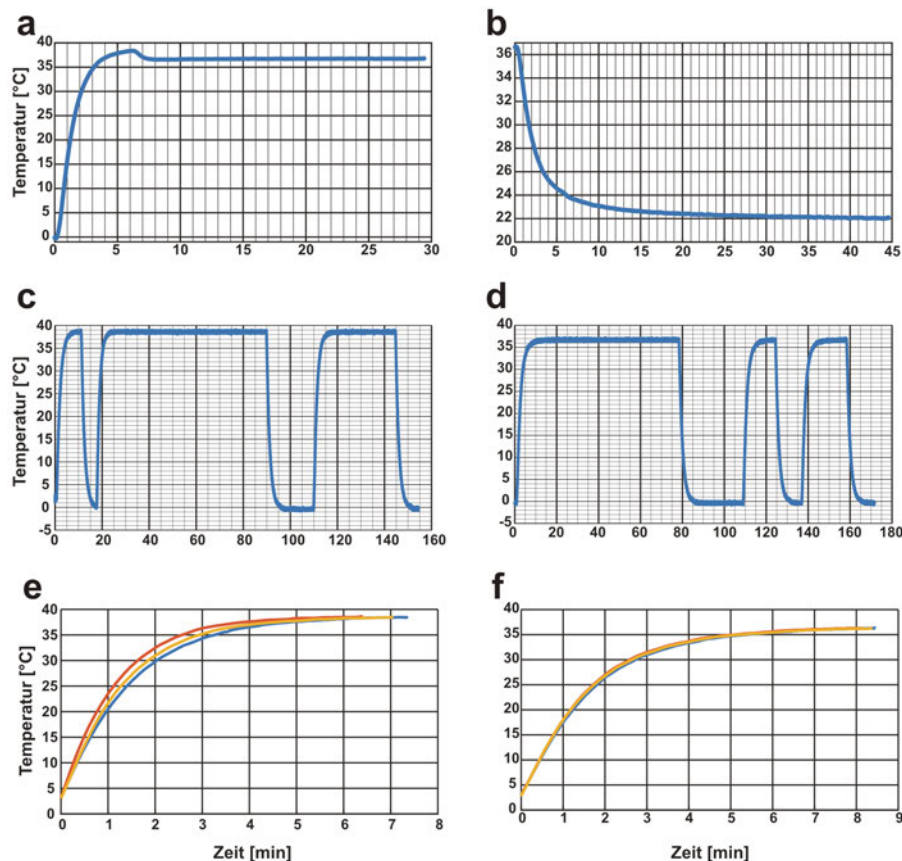


Abbildung 24 – Validierung der Probentemperatur im Heizblock und dem Magneten des MR-Scanners: Um mögliche Temperatureffekte während der Messung ausschließen zu können, war eine vollständige Temperierung der Probe auf 37 °C wichtig. Dazu wurde ein optisches Laserthermometer mit einer Glasfasersonde verwendet. Die entsprechend präparierte Probe wurde von Eiswasser (= 0 °C) auf 37 °C im Heizblock vorgewärmt und, sobald die Zieltemperatur erreicht war, in den MR-Scanner transferiert (a). Dabei wurde festgestellt, dass es eine Temperaturdifferenz von 1 °C zwischen MR-Magneten und Heizblock gab, woraufhin der Heizblock entsprechend angepasst wurde. Wenn die Probe aus dem Magneten genommen wurde, fiel die Temperatur innerhalb von 15–25 min auf RT (ca. 22 °C) ab (b). Um genauere Einblicke in den Prozess zu erhalten, wurde dieselbe Probe mehrere Male nacheinander vom Eiswasser aufgewärmt und anschließend wieder abgekühlt. Dies wurde sowohl für den Heizblock (c) als auch den Magneten des MR-Scanners durchgeführt (d). Aus diesen Daten wurde anschließend ausgewertet, wie lange es dauerte, bis die Probe die Zieltemperatur erreichte. Sowohl für den Heizblock (e) als auch für den Magneten (f) blieb diese Zeit bei unter 10 min, auch wenn der Heizblock mit ca. 7 min die Probe schneller vorheizte als der Magnet (f) mit etwa 8,5 min. Basierend auf diesen Beobachtungen wurde die Zeit zum Vortemperieren auf mind. 10 min festgelegt.

Da die optimale Kulturtemperatur von Zellen bei 37 °C liegt, waren auch die Magneten des MR-Scanners für die Arbeit mit Zellen auf diese Temperatur eingestellt.

Um mögliche Temperatureffekte zu vermeiden, wurden auch die Proben vor den Messungen auf 37 °C geheizt. Wenn die Probe zuerst im Heizblock vorgewärmt wurde, zeigte sich, dass die Probe im Heizblock bereits nach 6–7 Minuten auf die Zieltemperatur angewärmt war. Wenn sie die gewünschten 37 °C erreicht hatte und in den MR-Magneten transferiert wurde, sank die Temperatur auf 36 °C innerhalb von weniger als einer Minute ab und blieb anschließend stabil bei dieser Temperatur (Abbildung 24 a). Wenn die Probe aus dem Scanner genommen und bei RT gelagert wurde, sank die Temperatur innerhalb von 15–25 min auf die entsprechenden 22 °C ab (Abbildung 24 b). Um genauere Einblicke in den Temperierungsvorgang zu erhalten, wurde dieselbe Probe mehrere Male nacheinander vom Eiswasser im Heizblock (Abbildung 24 c) bzw. den MR-Magneten (Abbildung 24 d) aufgeheizt. Die entstandenen Daten wurden daraufhin untersucht, wie lange es dauerte, bis die Probe von 0 °C auf das angestrebte Ziel von 37 °C gekommen war. Dabei wurde deutlich, dass beide Methoden – das Vorheizen im Heizblock (Abbildung 24 e) oder das Vorheizen im MR-Magneten (Abbildung 24 f) – in weniger als 10 min abgeschlossen waren, auch wenn die Temperierung im Heizblock mit ca. 7 min etwas schneller war als im Magneten mit 8,5 min. Aufgrund dieser Erkenntnis wurde im Automatisierungsprozess festgelegt, dass die Probe zuerst mind. 10 min im Heizblock vorgeheizt werden sollte und anschließend erneut 10 min im MR-Magneten. Dadurch wurde sichergestellt, dass es zu keinen Temperatureffekten während der Messung kommen würde.

4.2.3 Zusammenfassung

Um den Messdurchsatz von fünf bis sechs Proben pro Tag auf ein vier- bis fünffaches davon zu steigern, wurde eine kostengünstige Automatisierungslösung entwickelt und validiert (siehe 3.2.3.2). Diese basierte auf dem kommerziell verfügbaren Roboterarm Braccio. Der Greifer des Roboters wurde für den konkreten Anwendungsfall neu entwickelt und mittels SLA-3D-Druck hergestellt (siehe 4.2.1). Die Temperierung der Proben wurde validiert und alles in einem einheitlichen Prozessablauf festgehalten (siehe 4.2.2). Durch diesen war es der Roboterplattform möglich, bis zu 24 Proben an einem Arbeitstag autonom zu vermessen.

4.3 MR-Messung von Zelllinien

Im Vergleich zu primären Zellen werden Zelllinien als potenziell unsterblich angesehen. Das bedeutet, sie zeigen weniger Alterungseffekte als primäre Zellen und können folglich länger und reproduzierbarer kultiviert werden. Daher wurde entschieden, die ersten Messungen mit Zelllinien und nicht mit primären Zellen durchzuführen.

4.3.1 Messungen von drei Zelllinien

Nachdem bereits erste, vorläufige Messungen mit Zelllinien durchgeführt worden waren (siehe 4.1.2 und 4.1.3), wurde entschieden, mit drei Zelllinien – namentlich den CHO-, MDA231- und K562- Zellen – einen größeren Datensatz aufzubauen, um daran eine Klassifizierung testen zu können. Wie bereits erwähnt, wurden pro Probe $5 \cdot 10^6$ Zellen in einem Volumen von 50 μl vermessen (siehe 3.2.4.1 und 4.1.3).

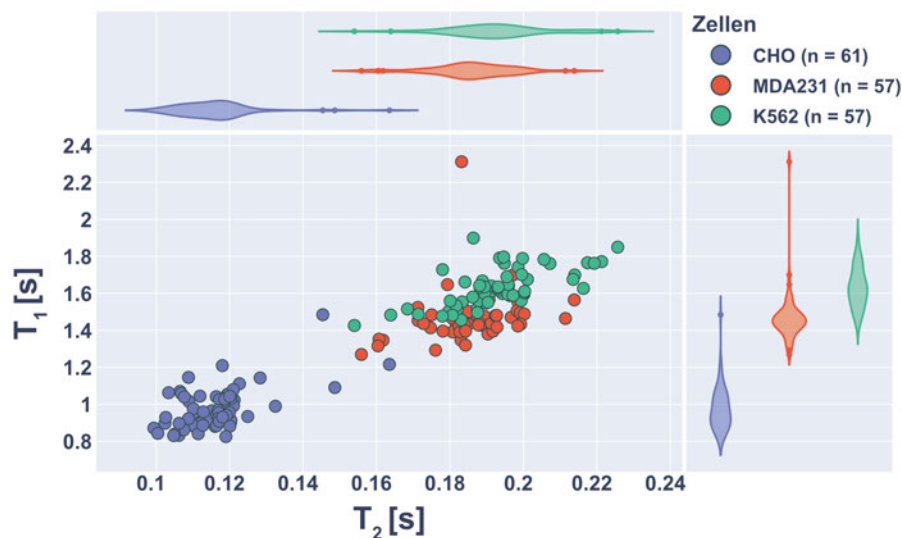


Abbildung 25 – Vergleich der Zellpeak-Schwerpunkte von den drei Zelllinien CHO, MDA231 und K562: Die Zelllinien wurden mehrfach gemessen (CHO $n_b = 61$; MDA231 $n_b = 57$; K562 $n_b = 57$) und die Schwerpunkte von den entsprechenden Zellpeaks gegeneinander aufgetragen. Dabei zeigte sich, dass sich v. a. die CHO-Zellen durch ihre kürzeren T_1 - und T_2 -Zeiten deutlich gegen die anderen beiden abgrenzten. Die MDA231 und K562 waren nicht so deutlich abgrenzbar. Sie ergaben zwar gut erkennbare Cluster, diese verschwammen jedoch leicht ineinander, was eine eindeutige Abgrenzung erschwerte. Wenn die Cluster in T_1 (violin plot rechts) bzw. T_2 (violin plot oben) separat geplottet wurden, zeigte sich, dass sich die MDA231 und K562 in ihrer Verteilung v. a. in T_2 ähnelten, während die MDA231 in T_1 deutlich enger gruppiert lagen und nur wenige Punkte mit längeren T_1 -Zeiten aufwiesen.

Für den Vergleich der gemessenen Daten der CHO-, MDA231- und K562-Zellen wurden die Schwerpunkte der Zellpeaks bestimmt und deren Positionen in einem Scatterplot miteinander verglichen. Die Medienpeaks waren nicht mitberücksichtigt, da

sie ausschließlich die Medienkomponente der Probe repräsentierten und folglich nur wenig Veränderungen zwischen den Zelllinien zeigen würden.

Der Vergleich der dargestellten Schwerpunkte zeigte erste Unterschiede zwischen den untersuchten Zellen. Diese Unterschiede waren v. a. bei den CHO-Zellen am offensichtlichsten. Dort clusterten die Medienpeaks deutlich abgegrenzt von den MDA231- und K562-Zellen. Einige wenige Messungen streuten in längere T_1 - und T_2 -Zeiten und damit näher an die übrigen Zellen heran. Dieser Unterschied war für die MDA231 und K562 nicht so deutlich erkennbar. Für diese beiden Zelllinien waren zwar ebenfalls zwei distinkte Cluster erkennbar, jedoch lagen diese sehr nah aneinander und die Grenzlinie wurde unscharf (Abbildung 25-Mitte). Wenn der 2D-Plot auf die einzelnen Dimensionen projiziert wurde und als violin-plot dargestellt wurde, waren die Ähnlichkeiten bzw. Unterschiede zwischen den Zellen in Bezug auf die Position der Zellpeak-Schwerpunkte ebenfalls deutlich sichtbar. Hier konnte man erkennen, dass sich die Zelllinien MDA231 und K562 in T_2 (Abbildung 25-Oben) stärker ähnelten als in T_1 (Abbildung 25-Rechts).

4.3.2 Erweiterung der Messungen auf zehn Zelllinien

Die beschriebenen Beobachtungen würden eine erste Klassifizierung basierend auf den erhobenen MR-Korrelationsspektren bereits zulassen. Daher wurde beschlossen, die Messungen auf bis zu zehn Zelllinien zu erweitern. Dabei war die Auswahl der Zelllinien hauptsächlich dadurch begründet, welche in der Zellbank des Lehrstuhls für Tissue Engineering vorhanden waren. Bei der Auswahl wurde auch darauf geachtet, adhärent wachsende und nicht-adhärent wachsende Zelllinien (sog. Suspensionszelllinien) sowie Zelllinien aus unterschiedlichen Spezies und mit unterschiedlicher Morphologie zu verwenden (Abbildung 26).

Aufgrund von experimentellen Bedingungen war der abschließende Datensatz an erhobenen Messdaten nicht ausgeglichen. Das bedeutet, dass nicht jede Zelllinie gleich häufig enthalten war (Abbildung 27). Es wurde jedoch sichergestellt, dass von jeder Zelllinie mindestens 16 unabhängige, biologische Replikate gemessen worden waren. Die abschließende Datenbank bestand aus 339 unabhängigen Messungen, welche über die zehn verwendeten Zelllinien verteilt erhoben wurden.

Um die erhobenen Daten besser vergleichen zu können, wurde für jede Zelllinie ein separater contour plot angefertigt, welcher alle erhobenen 2D-Korrelationsspektren zugeschnitten auf den relevanten Bereich darstellte (Abbildung 28). Die Replikate

wurden dabei übereinander projiziert. Dabei war auffällig, dass bei allen gemessenen Zelllinien der Mediumspeak ähnlich ausgeprägt war und außerdem in T_2 stärker streute als in T_1 -Richtung. Alle Zellen zeigten einen deutlich abgegrenzten Bereich bei kürzeren T_1 - und T_2 -Zeiten, in welchem die Zellpeaks lokalisiert waren. Hier zeigten sich deutliche Unterschiede zwischen den unterschiedlichen Zelllinien. Die CHO- und C2C12-Zellen zeigten einen eng gruppierten Bereich für Zellpeaks mit wenigen bis keinen Ausreißern. Alle weiteren Zellen wiesen eine deutlich größere Streuung der Messdaten auf. Die L929-, THP1-, HeLa-, HEK293T- und Vero-Zellen wiesen häufig mehr als einen diskreten Zellpeak auf, was sich dadurch kennzeichnete, dass deutliche Ausreißer bei längeren T_1 - und T_2 -Zeiten lagen. Einige Zelllinien, darunter u. a. K562-, MDA231-, L929- und HEK293T-Zellen, wiesen sehr breite Zellpeaks auf, welche v. a. in T_2 -Richtung eine Vielzahl an Messwerten enthielten.

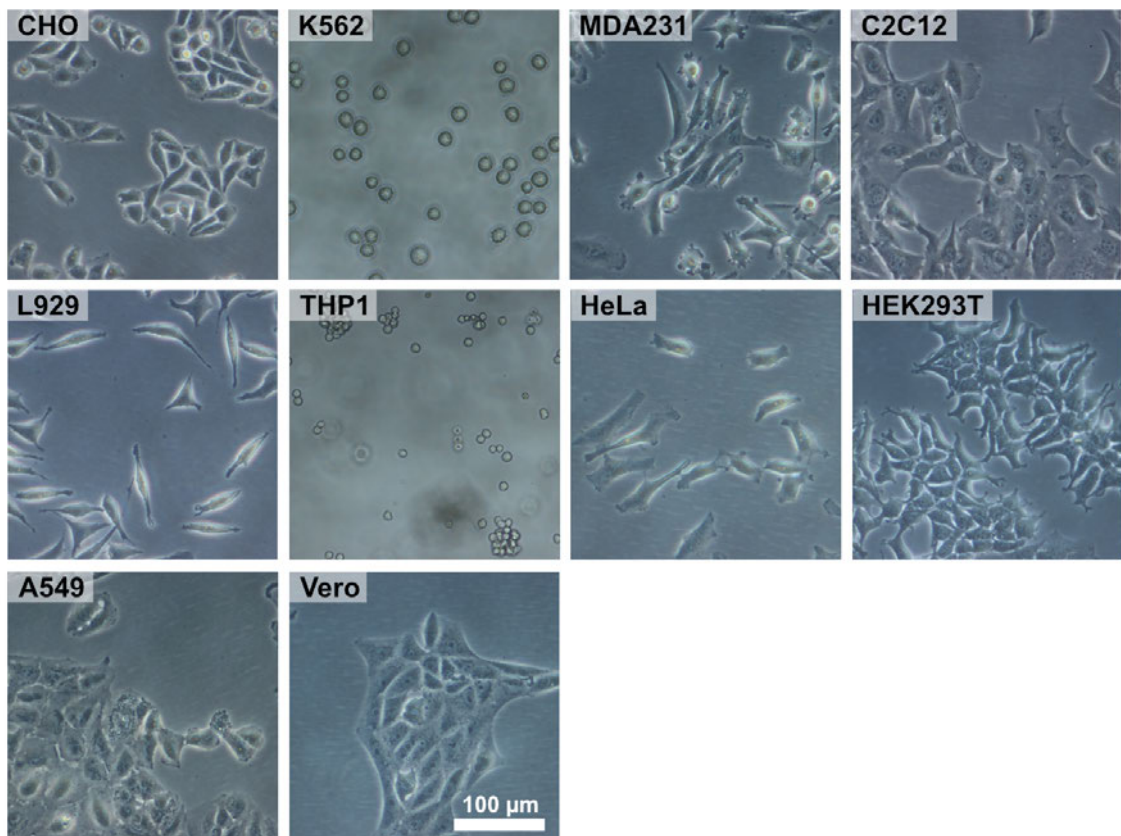


Abbildung 26 – Mikroskopische Aufnahmen der verwendeten Zelllinien: Die Zellen unterschieden sich in der ursprünglichen Spezies, aus welcher sie isoliert wurden, sowie aus dem entsprechenden Gewebetyp. Daher zeigten sie unter dem Mikroskop auch deutlich unterschiedliche Morphologien. Während die meisten Zellen adhärent wuchsen, waren die THP1 und die K562 sog. Suspensionzellen, d. h. , sie wuchsen nicht auf dem Boden der Kulturschale an. Manche der adhärent wachsenden Zellen tendierten zur Koloniebildung, wie z. B. die Vero-, HEK293T-, CHO- oder C2C12-Zellen, während andere, wie z. B. die HeLa- oder L929-Zellen, je nach Konfluenz der Kultur weniger zur Koloniebildung tendierten. Unter dem Mikroskop konnte verifiziert werden, dass die Kultur nicht kontaminiert oder anderweitig verunreinigt war. Abbildung adaptiert von [122].

Da diese Darstellung zwar einen detaillierten Einblick in die Zusammensetzung der Korrelationsspektren für jeden Zelltyp ermöglichte, es dem Beobachter jedoch erschwerte, die unterschiedlichen Zellen miteinander zu vergleichen, wurde hier genau wie bei den ersten drei Messungen in Abschnitt 4.3.1 der Schwerpunkt für die Zellpeaks bestimmt und gemeinsam in einem scatter plot dargestellt (Abbildung 29). So war es möglich, einen besseren Überblick über die Unterschiede zwischen den verschiedenen Zelllinien zu erhalten. Während bei der Darstellung mit nur drei Zelllinien (Abbildung 25) die Unterschiede noch recht deutlich erkennbar waren, war dies bei dem vollen Umfang der zehn Zelllinien nicht mehr möglich. Die meisten Zelllinien wiesen zwar deutliche Cluster auf, jedoch waren diese eng gruppiert, wie bei den CHO-Zellen, oder über einen größeren T_2 -Bereich verteilt, wie bei den L929- oder K562-Zellen. Auch bei der separaten Betrachtung der T_1 - und T_2 -Dimension wurden lokale Cluster in T_1 sichtbar, welche deutlich abgegrenzt zu anderen lagen wie bei CHO oder MDA231, oder aber sehr ähnlich wie Vero, L929, A549 und C2C12. Durch diesen Vergleich wurde offensichtlich, dass eine rein optische Klassifizierung der Zellen durch den Beobachter unmöglich ist.



Abbildung 27 – Anzahl der im finalen Datensatz enthaltenen, unabhängigen Messungen, aufgelistet für jede Zelllinie: Da vorerst die Messungen mit nur drei Zelllinien getestet worden waren und erst anschließend die weiteren Zelllinien gekommen waren, führte das dazu, dass der Datensatz nicht ausgeglichen war. Das bedeutet, nicht von jeder Zelllinie waren gleichviele Messungen enthalten. Mit mind. 16 unabhängigen, biologischen Replikaten waren jedoch für jede Zelllinie ausreichende Datenpunkte vorhanden, um statistisch signifikante Aussagen treffen zu können. Die drei ursprünglich gemessenen Zelllinien CHO, MDA231 und K562 waren am häufigsten vertreten, gefolgt von den Zelllinien L929, C2C12 und THP1 mit mehr als 25 biologischen Replikaten pro Zelllinie. Mit über 20 biologischen Replikaten waren die Zelllinien HeLa und HEK293T ebenfalls ausreichend häufig vertreten. Abschließend waren die A549- und die Vero-Zellen mit 17 bzw. 16 biologischen Replikaten am wenigsten häufig vertreten.

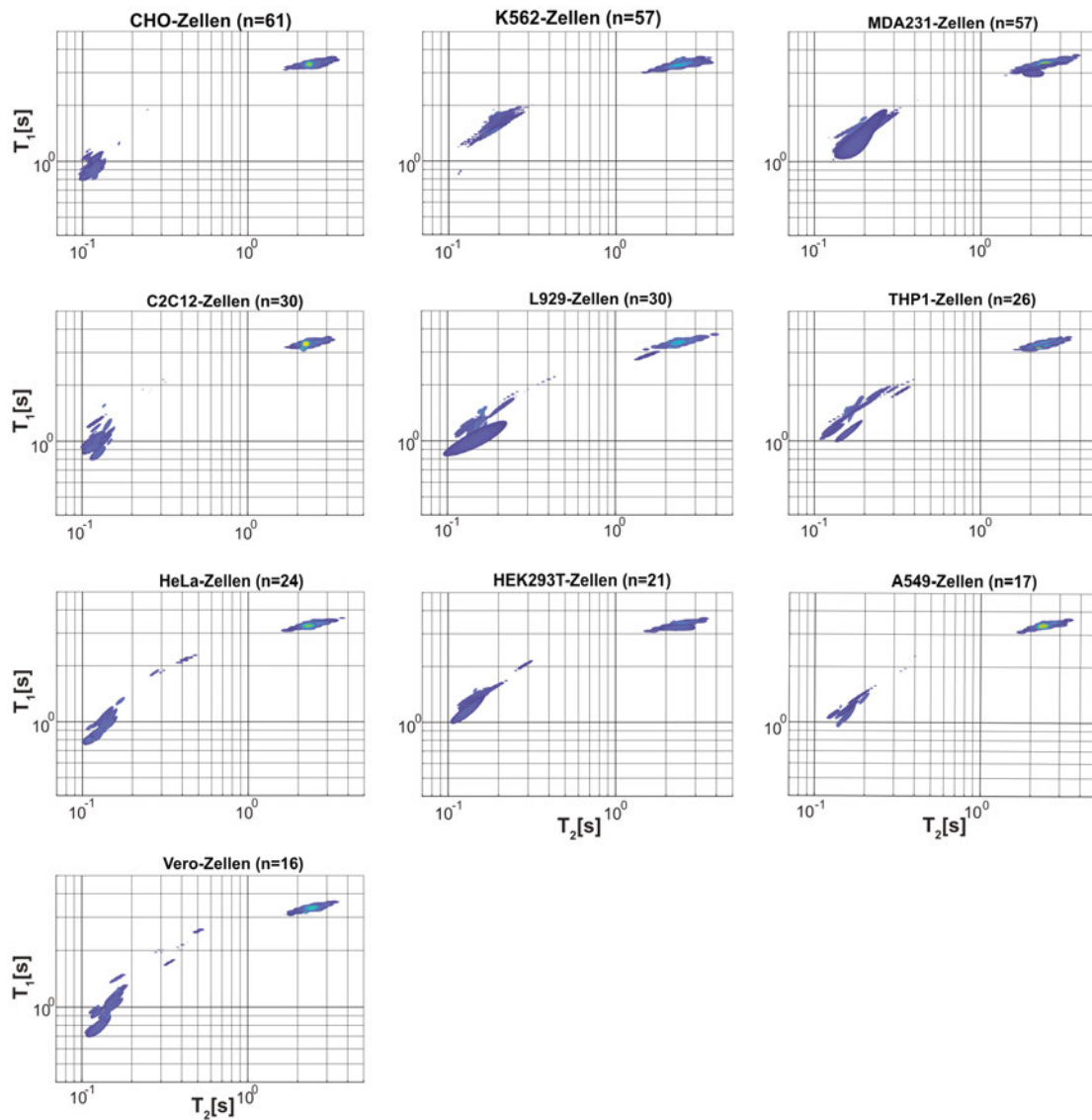


Abbildung 28 – Separate 2D-Korrelationsspektren für jeden gemessenen Zelltyp, kombiniert für alle erhobenen Replikate: Bei der Betrachtung der Korrelationsspektren wurde ersichtlich, dass die Medienpeaks aller erhobenen Messungen sehr stabil in ihrer Position waren. Sie streuten gleichmäßig über alle Zelllinien hinweg, v. a. in der T_2 -Dimension. Die Zellpeaks der CHO- und C2C12-Zellen lagen – im Vergleich zu allen anderen Zellen – deutlich enger gruppiert, d. h. sie wiesen eine geringere Streuung in T_1 und T_2 auf. Andere Zelllinien streuten deutlich stärker in ihren T_2 -Werten, wie z. B. die K562-, MDA231-, L929- oder THP1-Zellen. Manche der Korrelationsspektren wiesen auch zusätzliche Cluster auf, welche dem Zellsignal zugeordnet wurden, wie z. B. bei den HeLa-, HEK293T- oder Vero-Zellen. Abbildung adaptiert von [122].

4.3.3 Zellalterung und Viabilität

Um sicher zu stellen, dass die beobachteten Effekte nicht in der Viabilität der Zellen begründet lagen, wurde eine Versuchsreihe zur Bestimmung von eben jenem Wert durchgeführt. Hierfür wurden die Zellen in der Robotikplattform prozessiert und ihre Viabilität nach 5 h mittels NucleoCounter gemessen. Dabei zeigte sich,

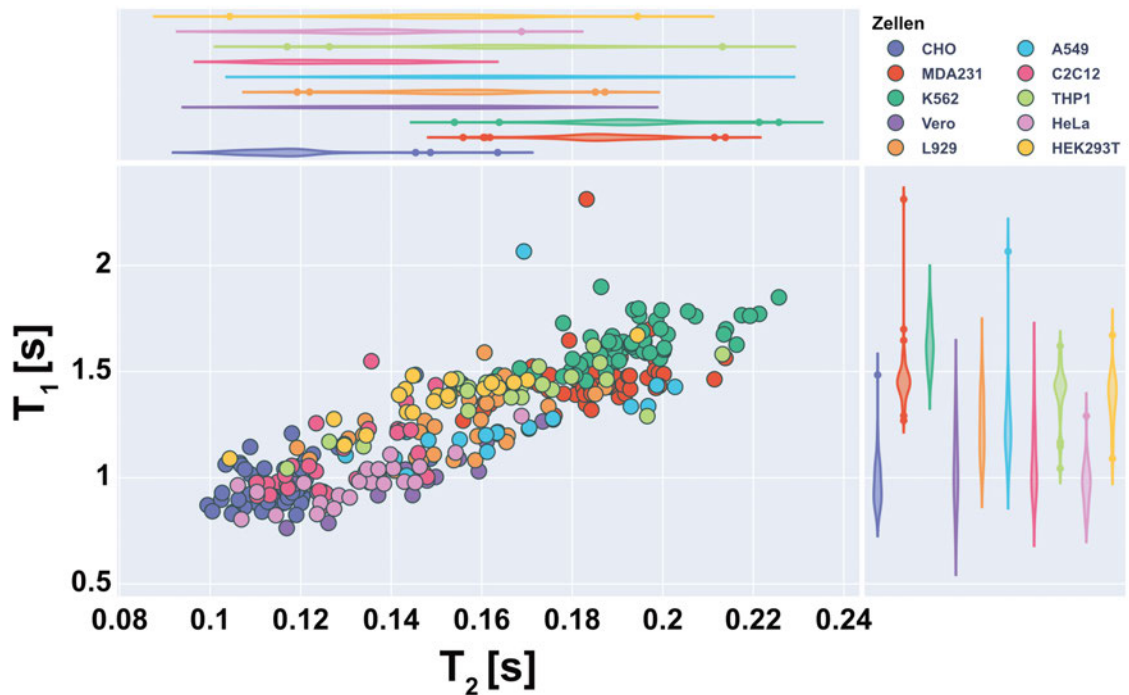


Abbildung 29 – Scatter-plot Darstellung der Zellpeaks für den vollen Messumfang von zehn unterschiedlichen Zelllinien: Für den besseren Vergleich wurden die Schwerpunkte der Zellpeaks gegeneinander aufgetragen. Außerdem wurde die Projektion in T_1 - bzw. T_2 -Dimension seitlich als violin plot dargestellt. Bei den Projektionen wurde deutlich, dass die Zellen in T_2 -Richtung nur wenige, mit dem Auge kaum differenzierbare Unterschiede aufwiesen. In der T_1 -Dimension waren manche Unterschiede deutlicher sichtbar. In der 2D-Darstellung wiesen alle Zellen zusammenhängende Cluster auf, welche jedoch – abhängig von der Zelllinie – ein geringere Streuung, wie z. B. die CHO- oder K562-Zellen, oder eine deutlich erhöhte Streuung zeigten, wie z. B. die L929- oder HeLa-Zellen. Aufgrund der unklaren Grenzen zwischen den Clustern war eine visuelle Klassifizierung, basierend auf den dargestellten Daten, nicht möglich.

dass mit Ausnahme der HEK293T- und Vero-Zellen alle Zelllinien eine Viabilität von $>90\%$ aufwiesen (Abbildung 30a). Die statistische Analyse der Daten mittels ANOVA (Alpha-Niveau = 5%) gepaart mit Tukey's Test zeigte, dass keine der untersuchten Zellen, außer HEK293T und Vero, signifikante Unterschiede zur Positivkontrolle aufwiesen. Diese Beobachtungen legen den Schluss nahe, dass die verwendete Methodik keinen Einfluss auf einen Großteil der Zellen hatte.

Um weiter zu untersuchen, welche der möglichen Kulturfaktoren (Temperatur, CO_2 , Mediumversorgung) für die schwankende Viabilität verantwortlich sein könnten, wurden diese individuell getestet (Abbildung 30b). Dabei zeigte sich, dass die Kultur der Zellen bei RT signifikante Unterschiede zu den anderen Kulturbedingungen aufwies. Sobald die Zellen bei $37^\circ C$ kultiviert wurden, was auch bei der Kultur im MR-Scanner der Fall war, zeigten sich keine signifikanten Unterschiede mehr. Dies galt auch für den Vergleich zur Positivkontrolle, bei welcher die Zellen direkt nach dem Ablösen aus der Kulturflasche auf ihre Viabilität hin untersucht worden waren.

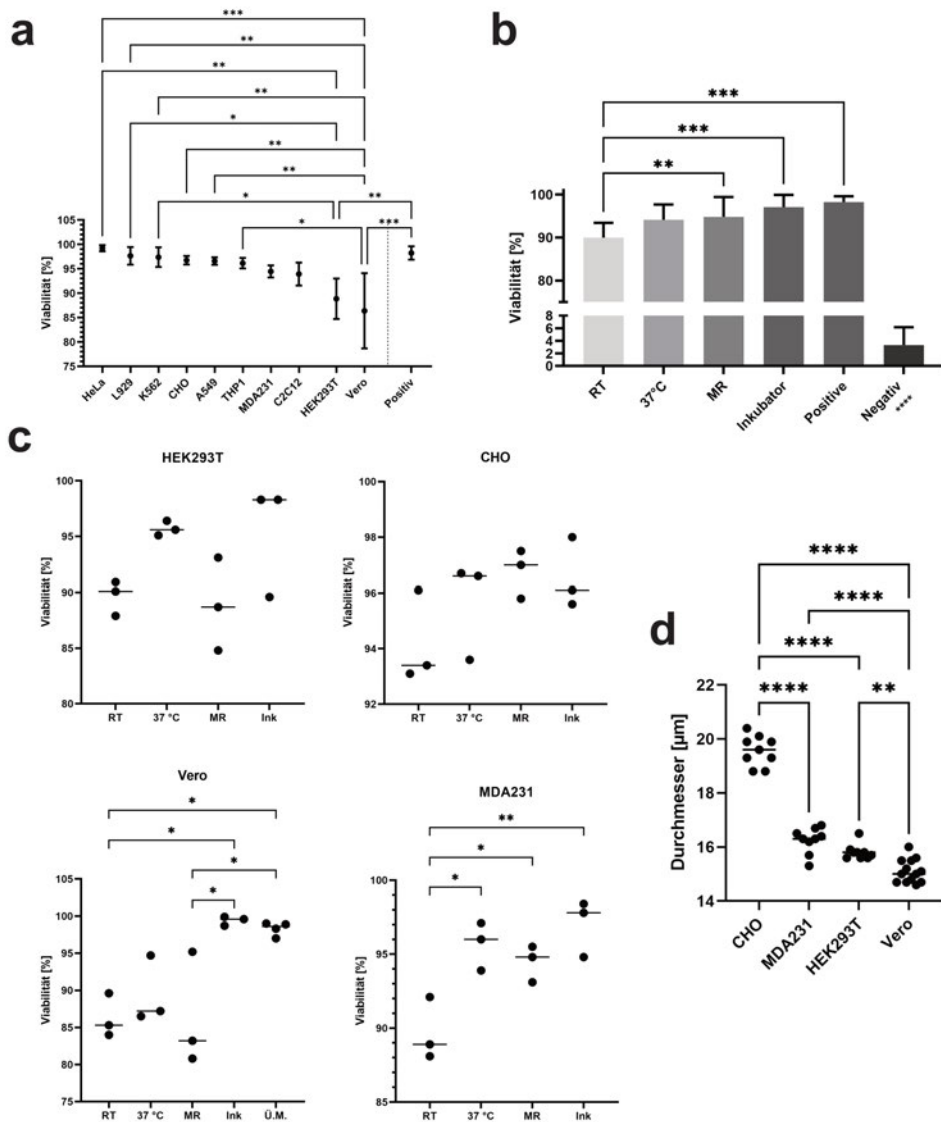


Abbildung 30 – Untersuchung der Viabilität der verwendeten Zelllinien: Die Zellen wurden in der Roboterplattform prozessiert und anschließend ihre Viabilität mithilfe des NucleoCounters gemessen ($n_b = 3$) (3.2.4.2). Die statistische Analyse der Ergebnisse (ANOVA gepaart mit Tukey's Test; Alpha-Niveau 5%) zeigte, dass die gemessenen Werte für Vero- und HEK293T-Zellen signifikant unter dem Vergleich zu den anderen dokumentierten Werten lagen. Keine der anderen untersuchten Zellen wiesen hingegen signifikante Unterschiede zur verwendeten Positivkontrolle auf (a). Eine genauere Untersuchung der möglichen Ursachen für den Abfall der Viabilität zeigte, dass nur die RT einen signifikanten Unterschied im Vergleich zu den anderen untersuchten Kulturbedingungen aufwies (b). Alle untersuchten Proben zeigten einen hoch signifikanten Unterschied zur untersuchten Negativkontrolle. Bei der genaueren Betrachtung der gemessenen Einzelwerte aus b für die unterschiedlichen Zelllinien (c) zeigten sich für HEK293T und CHO keine signifikanten Unterschiede zwischen den untersuchten Kulturbedingungen. Bei den Vero-Zellen zeigten sich signifikante Unterschiede zwischen den Proben, welche im Inkubator (Ink) kultiviert wurden und den Proben bei RT, ebenso wie zu den Proben, welche mit überschüssigem Medium (Ü.M.) kultiviert wurden. Bei den MDA231-Zellen zeigten sich signifikante Unterschiede zwischen RT Proben und allen anderen untersuchten Kulturbedingungen. Die Analyse der Zelldurchmesser, welche vom Nucleocounter neben der Viabilität ebenfalls gemessen wurde, zeigte signifikante Unterschiede zwischen unterschiedlichen Zellen (d). Diese korrelierten jedoch nicht mit den generellen Viabilitäten aus a, da große Zelldurchmesser, wie bei CHO-Zellen, nicht zu kleineren Viabilitäten führten und umgekehrt, kleine Zelldurchmesser wie bei Vero, durchaus kleinere Viabilitäten hervorriefen. Abbildung adaptiert von [122]. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$

Da die Ergebnisse auf der Kombination aus vier unterschiedlichen Zelllinien beruhen (HEK293T, CHO, Vero und MDA231), wurde eine statistische Analyse auch auf Ebene der einzelnen Zelllinien durchgeführt (Abbildung 30c). Hier zeigte sich, dass es keine statistisch signifikanten Unterschiede zwischen den unterschiedlichen Kulturbedingungen für HEK293T und CHO gab. Bei den Vero-Zellen gab es v. a. im Vergleich zu den Zellen, welche im Inkubator kultiviert wurden ebenso wie bei den Zellen, welche mit überschüssigem Medium (Ü.M.) kultiviert wurden, signifikante Unterschiede im Vergleich zu den restlichen Kulturbedingungen. Bei den MDA231-Zellen waren diese Unterschiede nur im Vergleich zwischen der Kultur bei RT und allen anderen Bedingungen detektierbar. Diese Beobachtungen bestätigen, dass es zwischen der Kultur im MR-Scanner und der optimalen Kultur im Inkubator keine statistisch detektierbaren Unterschiede für keine der Zellen gab. Das heißt der Unterschied in den Viabilitätswerten war andersweitig begründet. Eine weitere Möglichkeit war die Unterversorgung mit Zellkulturmedium.

Da die Zellen nach ihrer Zellzahl abgemessen wurden, jedoch unterschiedliche Größen, d. h. Durchmesser, besaßen, wurde auch diese Option betrachtet. Die Analyse des NucleoCounters gab u. a. auch einen Wert für den durchschnittlichen Durchmesser der untersuchten Probe zurück (Abbildung 30d). Die statistische Analyse zeigte, dass die untersuchten Durchmesser signifikante Unterschiede aufwiesen. Durch die gleichbleibende Zellzahl in den untersuchten Proben korrelierten diese Durchmesser folglich auch mit der Größe der erzeugten Zellpellets und demnach auch mit der Menge an überschüssigem Medium in der Probe.

Auch wenn die beobachteten Effekte nicht der Viabilität der Zellen geschuldet waren, wurde doch untersucht, wie sich eine Zellalterung auf das Signal auswirkte. Dabei ist zu beachten, dass hier der Begriff der Zellalterung flexibler verwendet wird, als er typischerweise gebraucht wird. Normalerweise bezieht sich der Begriff der Zellalterung in der Zellkultur auf die fortschreitende Passage der Zellen, was z. B. bei primären Zellen Einfluss auf die Teilungsraten nehmen kann. In dieser Arbeit wurde der Begriff der Zellalterung so definiert, dass eine Probe hergestellt und ohne Medienwechsel bei RT gelagert wurde. Ohne Medienwechsel ist anzunehmen, dass die Zellen über die Zeit hinweg sterben und die Viabilität der Probe deutlich abnimmt. Hierfür wurden mehrere Replikate von vorbereiteten Zellproben über einen Zeitraum von bis zu 168 h oder 192 h zu mehreren Zeitpunkten im MR-Scanner untersucht. Vermessen wurden die C2C12- (Abbildung 31 a), die K562- (Abbildung 31 b), die HeLa- (Abbildung 31 c) und die L929- (Abbildung 31 d) Zellen. Die Ergebnisse der Replikate wurden, wie schon bei Abbildung 19, kombiniert, damit das Resultat über-

sichtlicher war und alle gemessenen Proben repräsentierte. Dabei zeigte sich bei allen untersuchten Proben, dass die Medienpeaks über die Zeit hinweg zu kürzeren T_1 - und T_2 -Zeiten drifteten und – mit Ausnahme der K562-Zellen – an Intensität verloren und gleichzeitig in der Breite zunahmen. Die Zellpeaks tendierten zu längeren T_1 - und T_2 -Zeiten, wohingegen dieser Trend bei den Medienpeaks nicht zu beobachten war. Im Zuge der Alterung war keine Tendenz zu erkennen, dass systematisch neue Peaks entstanden. Lediglich die schon existierenden Peaks änderten ihre Form und Position. Aus den dokumentierten Daten lies sich der Schluss ziehen, dass die Alterung der Probe einen sichtbaren Einfluss auf die Zusammensetzung des Signals hatte und daher Messdaten nur unter der Berücksichtigung des Probenalters evaluiert werden sollten.

Als Negativkontrolle für die Viabilität der Zellen wurden Proben bei 95 °C für 30 min und 300 RPM gekocht (siehe 3.2.4.2). Bei mikroskopischer Betrachtung der gekochten Proben – fortan als Negativkontrolle bezeichnet – waren kaum noch intakte Zellkörper sichtbar (Abbildung 32 a), was vergleichbar mit einem Endzustand nach langer Kulturdauer ist. Bei entsprechender Vergrößerung waren deutlich die Überreste zerstörter Zellen zu erkennen (Abbildung 32 a: rote Pfeile).

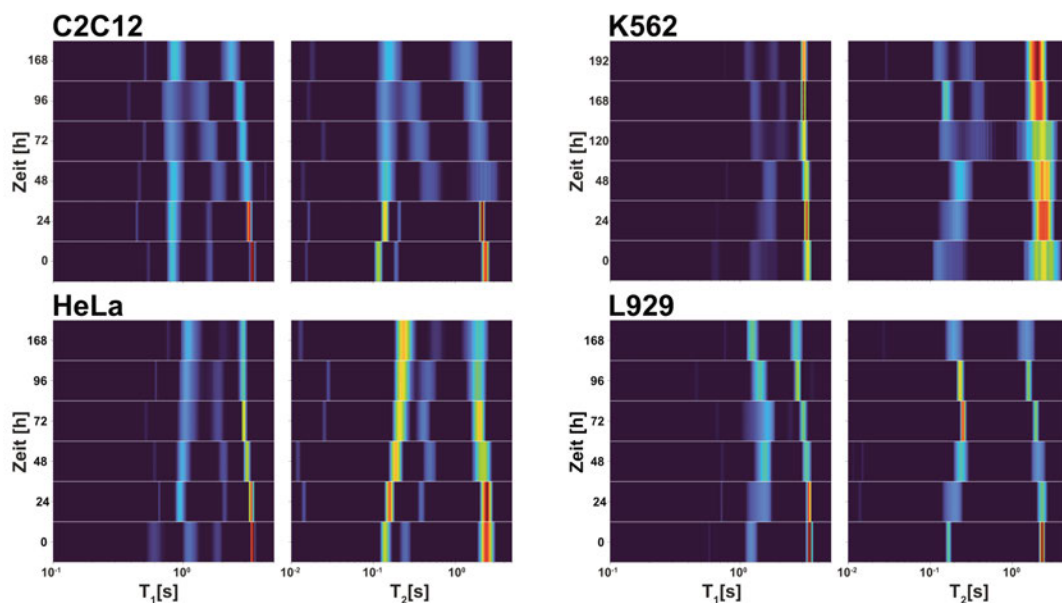


Abbildung 31 – Visualisierung der Zellalterung anhand einer Messreihe über bis zu 192 h: Um auftretende Alterungseffekte zu untersuchen, wurden bei vier Zelllinien Proben über einen Zeitraum von bis zu 192 h regelmäßig vermessen. Dabei zeigte sich, dass die Medienpeaks über den Zeitverlauf hin zu kürzeren T_1 - und T_2 -Zeiten drifteten und dabei i. d. R. auch an Intensität verloren. Gleichzeitig stieg die Streuung innerhalb des Medienpeaks, wodurch sie an Breite in beiden Dimensionen zunahm. Die Zellpeaks änderten ebenfalls ihre Position, tendierten jedoch zu längeren T_1 - und T_2 -Zeiten. Hier war der Drift v. a. unter Berücksichtigung der logarithmischen Skalierung der Achsen jedoch nicht so deutlich wie die des Medienpeaks.

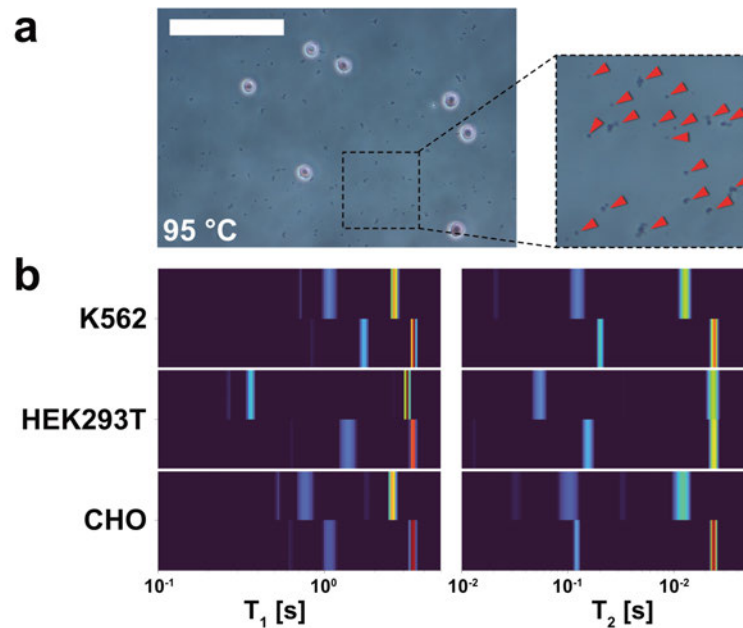


Abbildung 32 – Etablierung und Messung von Negativkontrollen: Hierfür wurden Zellen bei 95 °C und 300 RPM für 30 min gekocht. Die Methodik wurde anhand von NucleoCounter-Messung validiert. Die gekochten Proben – fortan als Negativkontrollen bezeichnet – zeigten nur noch wenige intakte Zellkörper (a). Bei entsprechender mikroskopischer Vergrößerung waren deutliche Fragmente von geplatzten Zellen zu erkennen (a – **Ausschnitt: rote Pfeile**). Bei der Vermessung der so hergestellten Negativkontrollen, zeigte sich, dass die Medienpeaks zwar ebenfalls zu kürzeren Relaxationszeiten tendierten, dabei ihre Breite jedoch nicht veränderten (b). Die Zellpeaks der drei untersuchten Zelllinien K562, HEK293T und CHO änderten ebenfalls ihre Position. Jedoch verlängerten sich nicht ihre T_1 - bzw. T_2 -Zeiten, wie in Abbildung 31 beschrieben, sondern verkürzten diese. Damit konnte zwar ein Effekt bestätigt werden, jedoch entsprach dieser nicht den zuvor beschriebenen Beobachtungen. Der Skalierungsbalken in a entspricht 100 μm . In b entspricht die jeweils untere Linie der Lebendkontrolle und die jeweils obere Linie der Negativkontrolle für jede gemessene Zelllinie.

Eine Überprüfung mithilfe des Nucleocounters ergab eine durchschnittliche Viabilität von 3,3 %. Dies bestätigte, dass die gekochte Kultur als valide Negativkontrolle verwendet werden konnte.

Bei der MR-Messung von entsprechend vorbereiteten Negativkontrollen (Abbildung 32 b) bestätigte sich die bereits beschriebene Tendenz der Medienpeaks, ihre Position in Richtung kürzerer Relaxationszeiten zu verschieben. Jedoch war dies, anders als zuvor beschrieben, nicht bei allen Proben von einer Verbreiterung der Peaks begleitet. Auch war im Gegensatz zu Abbildung 31 die Tendenz der Zellpeaks nicht in Richtung von längeren, sondern von kürzeren Relaxationszeiten gerichtet.

Nachdem die Negativkontrollen im Gegensatz zu lang kultivierten Proben (Abbildung 31) größtenteils nur noch Zellfragmente enthielten, wurde angenommen, dass das resultierende Signal zwar durch die Viabilität der Zellen beeinflusst wird, jedoch auch maßgeblich von der strukturellen Integrität der Zellen abhängt. Durch diese Hypothese kann erklärt werden, warum sowohl Abbildung 32 als auch

Abbildung 31 aus inaktiven bzw. toten Zellen bestehen, jedoch unterschiedliche Alterungserscheinungen aufweisen.

4.3.4 Zusammenfassung

Anhand erster Zellmessungen konnte gezeigt werden, dass die Schwerpunkte der Zellpeaks möglicherweise für die Klassifizierung der Zelllinien verwendet werden können (siehe 4.3.1). Je mehr Zelltypen in die Analyse miteinbezogen wurden, desto schwieriger gestaltete sich die Klassifizierung (siehe 4.3.2). Hier war es visuell nicht mehr möglich, die Zellen eindeutig voneinander zu unterscheiden. Bei der Untersuchung der Viabilität konnte gezeigt werden, dass alle Zellen nach einer durchschnittlichen Prozesszeit von 5 h mit einer Viabilität von $> 80\%$ und in den meisten Fällen sogar $> 90\%$ als lebendig einzustufen waren (siehe 4.3.3). Die Schwankungen in den Viabilitätswerten konnte v. a. über die Unterversorgung mit Medium erklärt werden (Unterabschnitt 4.3.3). Bei der Vermessung von gealterten Proben, welche entweder inaktiv oder tot waren, konnte jedoch gezeigt werden, dass längere Zeiträume zu deutlichen Veränderungen in den Korrelationsspektren führten. Durch die Vermessung von etablierten Negativkontrollen konnte zwar ein Effekt bestätigt werden, jedoch entsprach dieser nicht den dokumentierten Effekten der Zeitreihenmessung. Daraus wurde geschlossen, dass einerseits die Viabilität bei stark gealterten Proben Einfluss auf das resultierende MR-Signal hatte, dass dieses jedoch auch von der zellulären Integrität der untersuchten Proben abhing.

4.4 MR-Messung von mesenchymalen Stromazellen

Da die verwendeten Zelllinien zwar ein gut reproduzierbares Messsystem darstellten, jedoch für die Demonstration der Methodik für die klinische Anwendung nur eingeschränkt geeignet waren, wurden für diese Demonstration mesenchymale Stromazellen (MSCs) verwendet. Die MSCs werden, wie eingangs beschrieben, in zahlreichen Studien und Therapieansätzen als Behandlung von unterschiedlichen Krankheiten verwendet. Die verwendeten Zelllinien werden i. d. R. jedoch nicht für therapeutische Zwecke verwendet. Hierdurch stellen die verwendeten MSCs ein geeigneteres Beispiel für die klinische Anwendung dar.

Durch ihren Stammzellcharakter konnten die MSCs in unterschiedliche Zelltypen differenziert werden. Nachdem wie bereits in der Einleitung beschrieben (Unterabschnitt 2.4.1) v. a. Fettgewebe physikalisch gesehen deutliche Unterschiede im MR-Signal zeigen sollte, wurden die MSCs in Fettzellen (Adipocyten) differenziert.

Die Zellen wurden wie in 3.2.1.4 beschrieben kultiviert. Dann wurden sie aus der Kulturflasche abgelöst, als MR-Probe aufbereitet und durch die Robotikplattform automatisch vermessen. Anschließend wurden die vermessenen Zellen wieder in Kultur genommen und entsprechend des beschriebenen Protokolls in Adipocyten differenziert. Im Anschluss wurden die Zellen erneut im MR-Scanner vermessen.

Bei der Betrachtung unter dem Mikroskop zeigte sich, dass die differenzierten Zellen im Vergleich zu den undifferenzierten Zellen (Abbildung 33-b Undiff.) intrazelluläre Ansammlungen von Tropfen aufwiesen (Abbildung 33-b Diff.). Diese konnten mittels der Ölrot O Färbung als Lipidtropfen und die Zellen somit als Adipocyten identifiziert werden (Abbildung 33-b Ölrot O). Beim Vergleich der aufgenommenen MR-Spektren fiel auf, dass die undifferenzierten Zellen einen Medienpeak (Abbildung 33-a₁ grüne Pfeilspitze) und einen Zellpeak (Abbildung 33-a₁ gelbe Pfeilspitze) aufwiesen. Die differenzierten Zellen wiesen hingegen einen Medienpeak (Abbildung 33-a₂ grüne Pfeilspitze) und zwei Zellpeaks (Abbildung 33-a₁ grüne und rote Pfeilspitze) auf.

Zusammen mit den mikroskopisch / histologischen Beobachtungen legte dies den Schluss nahe, dass die gebildeten Lipidanteile der Zellen zu dieser Veränderung im MR-Signal beigetragen hatten.

4.5 KI-gestützte Auswertung der Daten

Die Auswertung und Analyse von Daten mithilfe von Methoden der künstlichen Intelligenz (KI) bzw. des Machine Learnings (ML), gewinnt immer mehr an Bedeutung. Da KI sehr gut darin ist Regeln und Strukturen in Daten erkennen zu können, und diese somit effektiv und zuverlässig zu prozessieren und klassifizieren, wurde auch in dieser Arbeit auf entsprechende Methoden zurückgegriffen. Vor allem die Objektivität der Beurteilung der Daten macht die KI zu einem wichtigen Werkzeug um die gemessenen Daten möglichst ohne eine Voreingenommenheit des Nutzers (engl. *bias*) zu klassifizieren.

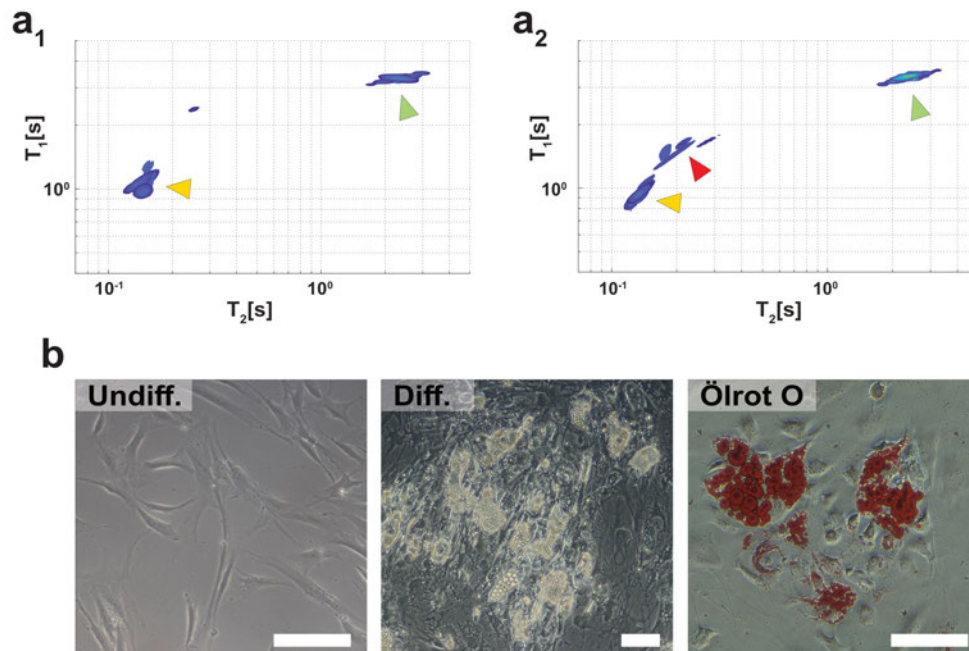


Abbildung 33 – Adipogene Differenzierung und MR-Messung von mesenchymalen Stromazellen: Wie in Unterunterabschnitt 3.2.1.4 beschrieben, wurden die mesenchymalen Zellen kultiviert und anschließend vermessen. Zum Vergleich wurden die Zellen vor der Differenzierung (b - 'Undiff.')

und nach der Differenzierung (b - 'Diff.')

präpariert und vermessen. Mittels der Ölröt O Färbung konnten die gebildeten Lipidtröpfchen in den MSCs angefärbt und dadurch identifiziert werden. Auch wenn für die Differenzierung ein adaptiertes Differenzierungsprotokoll verwendet wurde (Unterunterabschnitt 3.2.1.4), war die Differenzierung nicht zu 100 % effektiv. Beim Vermessen der Proben wurde deutlich, dass die undifferenzierten Zellen (a₁) einen Medienpeak (grüner Pfeilkopf) und einen Zellpeak (gelber Pfeilkopf) aufwiesen. Die differenzierten Zellen hingegen wiesen reproduzierbar zwei Zellpeaks auf (gelber und roter Pfeilkopf). Es wurden n_b = 8 undifferenzierte und n_t = 7 differenzierte Zellproben vermessen. Die Skalierungsbalken in b betragen 100 μm. Abbildung adaptiert von [122].

4.5.1 Nutzung einer Support Vector Machine für die Klassifizierung

Wie bereits in 4.3.2 beschrieben, war es möglich, die erhobenen Korrelationsspektren anhand der Position ihrer Zellpeaks einzuordnen. Um den Nutzer-Bias zu eliminie-

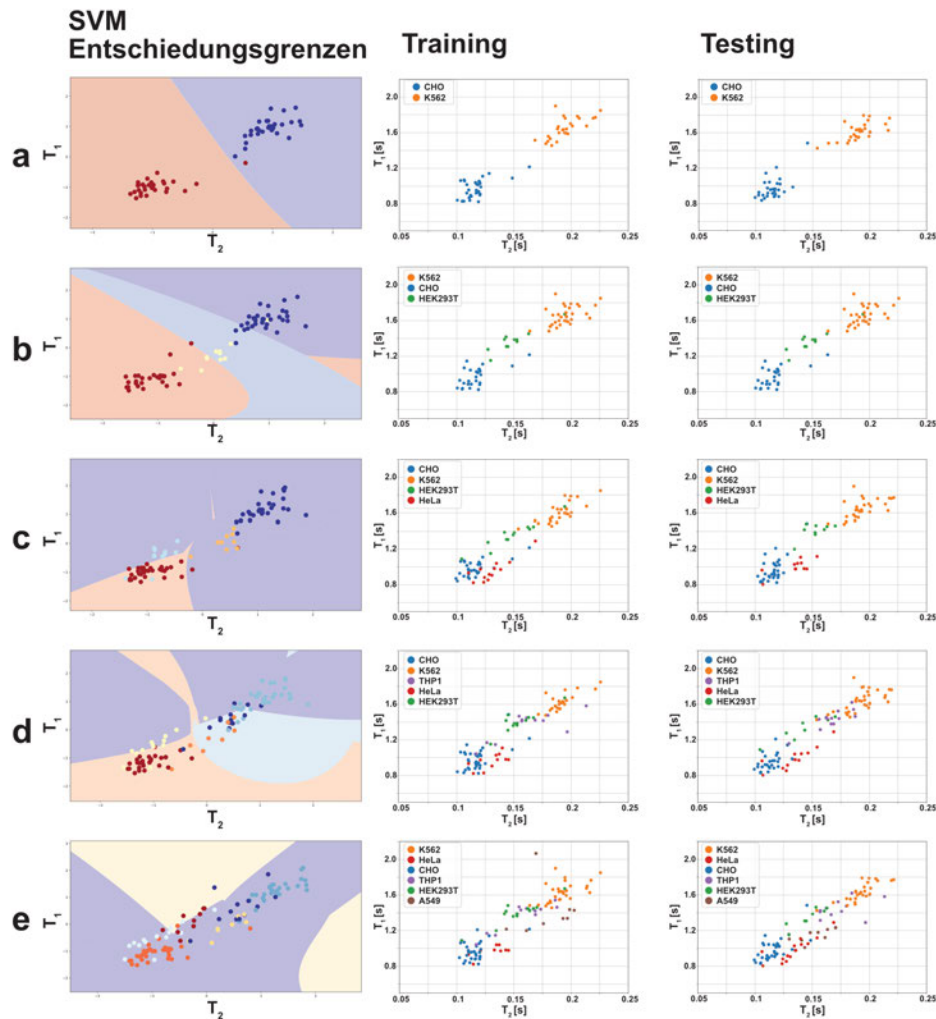


Abbildung 34 – Klassifizierung von Zellpeaks mithilfe einer SVM - Teil 1: Für die Klassifizierung mittels SVM wurde der Schwerpunkt der Zellpeaks berechnet und – zusammen mit der Bezeichnung der Zelllinie – als Label in den SVM-Algorithmus eingespeist. Die abgebildeten Entscheidungsgrenzen (farbige Bereiche) visualisieren die vom SVM angelegten Grenzen, welche für die Separation der Punktwolken automatisch berechnet wurden. Es war zu erkennen, dass mit steigender Zahl an berücksichtigten Zellen die Komplexität der Entscheidungsgrenzen stieg. Für jeden Durchlauf wurden 50 % der Daten als Trainings- und 50 % als Testdaten verwendet, damit sichergestellt wurde, dass für die Evaluation des trainierten Modells ausschließlich Datenpunkte verwendet wurden, die nicht für das Training verwendet worden waren. Die Visualisierung der Entscheidungsgrenzen basierte auf der Auswertung der Testdaten. Diesem stand die Visualisierung der Trainings- und Testdaten gegenüber, um einen direkt Vergleich zu erhalten. Aufgrund der schnellen Berechnung des SVM-Modells konnten 300 unabhängige Replikate berechnet werden ($n_t = 300$). Die Abbildung zeigt für jede Kombination an Zellen (a - e) einen der 300 durchgeführten Durchläufe. Die Zusammensetzung der Trainings- und Testdaten und demnach auch die Form der Entscheidungsgrenzen variierte zwischen jedem Durchlauf. Die Daten wurden vor der Analyse normalisiert, wodurch die Achsen der Entscheidungsgrenzen einheitslose Werte angaben. Für die Visualisierung der Trainings- bzw. Testdaten wurden die nicht-normalisierten Werte mit den entsprechenden T_1 und T_2 verwendet. Abbildung adaptiert von [122].

ren, wurde hierfür auf eine ML-Methode zurückgegriffen, welche als Support Vector Machine (SVM) bezeichnet wird. Der SVM-Algorithmus versucht, die Daten mit-

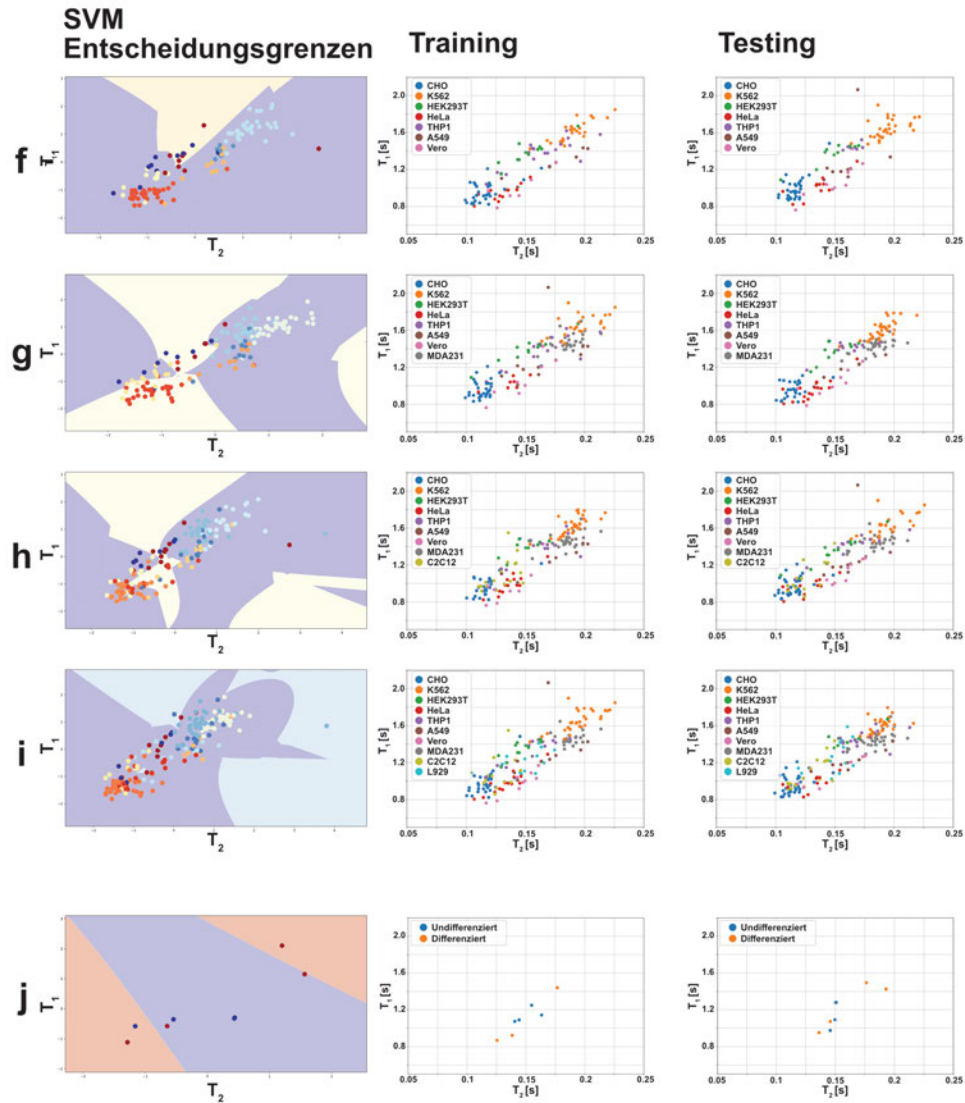


Abbildung 35 – Klassifizierung von Zellpeaks mithilfe einer SVM - Teil 2: Für die Klassifizierung mittels SVM wurde der Schwerpunkt der Zellpeaks berechnet und zusammen mit der Bezeichnung der Zelllinie als Label in den SVM-Algorithmus eingespeist. Die abgebildeten Entscheidungsgrenzen (farbige Bereiche) visualisieren die vom SVM angelegten Grenzen, welche für die Separation der Punktwolken automatisch berechnet wurden. Es war zu erkennen, dass mit steigender Zahl an berücksichtigten Zellen die Komplexität der Entscheidungsgrenzen stieg. Für jeden Durchlauf wurden 50 % der Daten als Trainings- und 50 % als Testdaten verwendet, damit sichergestellt wurde, dass für die Evaluation des trainierten Modells ausschließlich Datenpunkte verwendet wurden, die nicht für das Training verwendet worden waren. Die Visualisierung der Entscheidungsgrenzen basierte auf der Auswertung der Testdaten. Diesem stand die Visualisierung der Trainings- und Testdaten gegenüber, um einen direkt Vergleich zu erhalten. Aufgrund der schnellen Berechnung des SVM-Modells konnten 300 unabhängige Replikate berechnet werden ($n_t = 300$). Die Abbildung zeigt für jede Kombination an Zellen (a - i) einen der 300 durchgeführten Durchläufe. Da für die MSCs zwei Zellpeaks detektierbar waren, wurden nur die höchsten / prominentesten Signalpeaks für die SVM Auswertung verwendet (j). Die Zusammensetzung der Trainings- und Testdaten und demnach auch die Form der Entscheidungsgrenzen variierte zwischen jedem Durchlauf. Die Daten wurden vor der Analyse normalisiert, wodurch die Achsen der Entscheidungsgrenzen einheitslose Werte angeben. Für die Visualisierung der Trainings- bzw. Testdaten wurden die nicht-normalisierten Werte mit den entsprechenden T_1 und T_2 verwendet. Abbildung adaptiert von [122].

hilfe einer Linie bzw. eines Vektors entsprechend der vorgegebenen Kennung (engl. *Label*) zu separieren.

Die SVM wurde mittels der Programmiersprache *Python* durchgeführt wobei die Schwerpunkte der Zellpeaks als Datenpunkte zum tragen kamen (siehe 3.2.5.2). Dabei wurde für jede Probe jedoch nur ein einzelner Peak betrachtet. Falls eine Probe mehr als einen Zellpeak aufwies, fand dieser in der Analyse keine Berücksichtigung. Der SVM-Algorithmus erwartete eine Normalisierung der Daten, wodurch die Achsen der Plots nicht wie bisher in Sekunden angegeben wurden, sondern in normalisierten und damit einheitslosen Werten (Abbildung 34, Abbildung 35).

Um die Vielseitigkeit der SVM untersuchen- und demonstrieren zu können, wurde der Algorithmus zuerst an den bereits etablierten Zelllinien trainiert bevor er an den MSCs getestet werden konnte. Ein möglicher Algorithmus zur Klassifizierung der Zellen anhand ihrer 2D-T₁/T₂ -Zeiten muss in der Lage sein, so viele Zelltypen wie möglich zu berücksichtigen. Daher wurde bei der Etablierung der SVM die Zahl der inkludierten Zellen inkrementell erhöht (13). Zuerst wurden nur zwei Zelllinien verwendet: MDA231 und CHO. Diese waren mittels der SVM mit einer Genauigkeit von 97,98 % sehr gut voneinander zu trennen (Abbildung 34-a, Abbildung 36). Jedoch führt bereits eine einzige zusätzliche Zelllinie (K562) dazu, dass die Genauigkeit auf 89,3 % abfiel (Abbildung 36, Abbildung 34-b). Wenn die Zahl der inkludierten Zellen graduierlich erhöht wurde, sank die Genauigkeit der SVM auf bis zu 50,82 % ab (10 Zelllinien) (Abbildung 36, Abbildung 34-c bis i).

Tabelle 13: Verwendete Zelllinien: Für das Training der KNNs wurden unterschiedliche Kombinationen an Zelllinien verwendet.

Anzahl an Zelllinien	Beinhaltete Zelllinien
2	CHO, K562
3	CHO, K562, HEK293T
4	CHO, K562, HEK293T, HeLa
5	CHO, K562, HEK293T, HeLa, THP1
6	CHO, K562, HEK293T, HeLa, THP1, A549
7	CHO, K562, HEK293T, HeLa, THP1, A549, Vero
8	CHO, K562, HEK293T, HeLa, THP1, A549, Vero, MDA231
9	CHO, K562, HEK293T, HeLa, THP1, A549, Vero, MDA231, C2C12
10	CHO, K562, HEK293T, HeLa, THP1, A549, Vero, MDA231, C2C12, L929

Nachdem die SVM konzeptionell so aufgebaut ist, dass pro Probe nur ein Datenpunkt berücksichtigt werden kann, lieferte die Klassifizierung der MSC-Differenzierung, bei der wie schon in Abschnitt Abschnitt 4.4 beschrieben zwei Medi-

enpeaks auftraten, sehr variable Ergebnisse (Abbildung 36, Abbildung 34-j). Zwar lag die Genauigkeit bei durchschnittlich 62,67%, jedoch streuten die erzielten Genauigkeiten von 100% bis hin zu 0%. Neben dem Aspekt, dass die SVM für eine Fragestellung, welche auf zwei Datenpunkte zurückgeführt werden kann, ungeeignet ist, spielte auch die geringe Datenmenge für die MSCs eine entscheidende Rolle. Demnach stellte die SVM zwar eine durchaus geeignete Klassifizierungsmethode für die untersuchten Zelllinien, jedoch nicht für die Klassifizierung von Entwicklungsstadien der MSCs dar.

Statistische Auswertungen mittels ANOVA und Tukey's multiple comparison test zeigten, dass nur der Vergleich von 5 vs. 6 Zelllinien, 7 Zelllinien vs. MSCs und 8 Zelllinien vs. MSCs keine signifikanten Unterschiede zeigte. Alle anderen Kombinationen wiesen statistisch signifikante Unterschiede auf. Da dies in der Abbildung 36 zu unübersichtlich wäre, wurden die formellen Ergebnisse der Analyse nicht visualisiert, sondern nur in Textform in der Tabelle 14 angegeben.

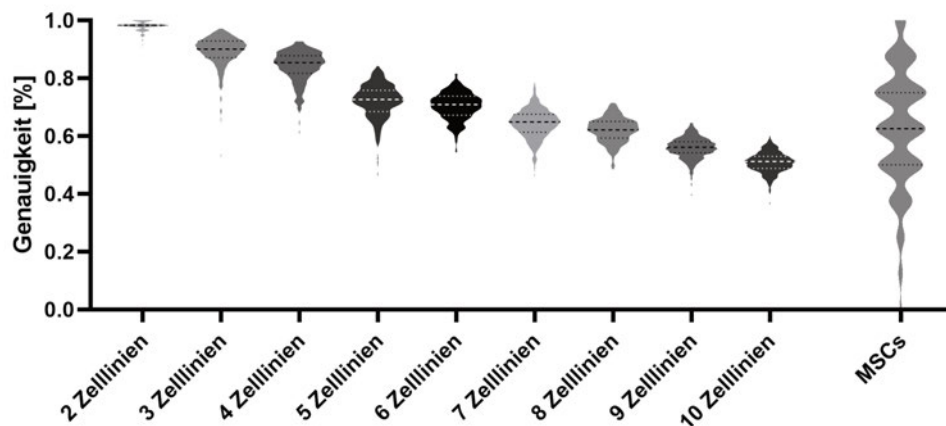


Abbildung 36 – Performanz der trainierten SVM: Um die Leistung der verwendeten SVM bei unterschiedlichen Zellzahlen zu testen, wurden diese inkrementell erhöht. Für jede getestete Zellzahl wurden $n_t = 300$ unabhängige Wiederholungen berechnet. Während die Genauigkeit bei zwei Zelllinien (CHO, K562) im Mittel noch bei 97,98% lag, reduzierte sie sich bei drei Zelllinien (+ HEK293T) auf 89,3%, bei vier Zelllinien (+ HeLa) auf 84,04%, bei fünf Zelllinien (+ THP1) auf 72,16%, bei sechs Zelllinien (+ A549) auf 70,34%, bei sieben Zelllinien (+ Vero) auf 64,25%, bei acht Zelllinien (+ MDA231) auf 62,03%, bei neun Zelllinien (+ C2C12) auf 62,03% und bei 10 Zelllinien (+ L929) auf 50,82%. Für die untersuchten MSC-Proben ergab sich im Mittel eine Genauigkeit von 62,67%, jedoch mit einer deutlich größeren Streuung im Vergleich zu allen untersuchten Zelllinien. Die MSCs streuten von 100% bis 0%. Der statistische Vergleich zwischen den dargestellten Proben ergab in fast allen Fällen einen signifikanten Unterschied (Tabelle 14). Abbildung adaptiert von [122].

Tabelle 14: Werte der statistischen Analyse der SVM-Performanz. Die Werte basieren auf einem Tukey's multiple comparison test mit einem Alpha Niveau von 5%. CI: Konfidenzintervall

Tukey's multiple comparisons Test	Mittlerer Unterschied	95,00 % CI des Unterschieds	Unter Schwellwert?	Angepasster p-Wert
2 Zelllinien vs. 3 Zelllinien	0,08673	0,06647 bis 0,1070	Ja	<0,0001
2 Zelllinien vs. 4 Zelllinien	0,1394	0,1192 bis 0,1597	Ja	<0,0001
2 Zelllinien vs. 5 Zelllinien	0,2582	0,2379 bis 0,2785	Ja	<0,0001
2 Zelllinien vs. 6 Zelllinien	0,2764	0,2561 bis 0,2966	Ja	<0,0001
2 Zelllinien vs. 7 Zelllinien	0,3373	0,3170 bis 0,3575	Ja	<0,0001
2 Zelllinien vs. 8 Zelllinien	0,3594	0,3392 bis 0,3797	Ja	<0,0001
2 Zelllinien vs. 9 Zelllinien	0,4214	0,4011 bis 0,4416	Ja	<0,0001
2 Zelllinien vs. 10 Zelllinien	0,4716	0,4513 bis 0,4918	Ja	<0,0001
2 Zelllinien vs. MSCs	0,3531	0,3329 bis 0,3734	Ja	<0,0001
3 Zelllinien vs. 4 Zelllinien	0,05268	0,03243 bis 0,07294	Ja	<0,0001
3 Zelllinien vs. 5 Zelllinien	0,1715	0,1512 bis 0,1917	Ja	<0,0001
3 Zelllinien vs. 6 Zelllinien	0,1896	0,1694 bis 0,2099	Ja	<0,0001
3 Zelllinien vs. 7 Zelllinien	0,2506	0,2303 bis 0,2708	Ja	<0,0001
3 Zelllinien vs. 8 Zelllinien	0,2727	0,2525 bis 0,2930	Ja	<0,0001
3 Zelllinien vs. 9 Zelllinien	0,3346	0,3144 bis 0,3549	Ja	<0,0001
3 Zelllinien vs. 10 Zelllinien	0,3849	0,3646 bis 0,4051	Ja	<0,0001
3 Zelllinien vs. MSCs	0,2664	0,2461 bis 0,2866	Ja	<0,0001
4 Zelllinien vs. 5 Zelllinien	0,1188	0,09853 bis 0,1390	Ja	<0,0001
4 Zelllinien vs. 6 Zelllinien	0,1370	0,1167 bis 0,1572	Ja	<0,0001
4 Zelllinien vs. 7 Zelllinien	0,1979	0,1776 bis 0,2181	Ja	<0,0001
4 Zelllinien vs. 8 Zelllinien	0,2200	0,1998 bis 0,2403	Ja	<0,0001
4 Zelllinien vs. 9 Zelllinien	0,2820	0,2617 bis 0,3022	Ja	<0,0001
4 Zelllinien vs. 10 Zelllinien	0,3322	0,3119 bis 0,3524	Ja	<0,0001
4 Zelllinien vs. MSCs	0,2137	0,1934 bis 0,2340	Ja	<0,0001
5 Zelllinien vs. 6 Zelllinien	0,01818	-0,002075 bis 0,03844	Nein	0,1231
5 Zelllinien vs. 7 Zelllinien	0,07909	0,05883 bis 0,09934	Ja	<0,0001
5 Zelllinien vs. 8 Zelllinien	0,1012	0,08099 bis 0,1215	Ja	<0,0001
5 Zelllinien vs. 9 Zelllinien	0,1632	0,1429 bis 0,1834	Ja	<0,0001
5 Zelllinien vs. 10 Zelllinien	0,2134	0,1931 bis 0,2336	Ja	<0,0001
5 Zelllinien vs. MSCs	0,09491	0,07466 bis 0,1152	Ja	<0,0001
6 Zelllinien vs. 7 Zelllinien	0,06091	0,04065 bis 0,08116	Ja	<0,0001
6 Zelllinien vs. 8 Zelllinien	0,08306	0,06281 bis 0,1033	Ja	<0,0001
6 Zelllinien vs. 9 Zelllinien	0,1450	0,1247 bis 0,1652	Ja	<0,0001
6 Zelllinien vs. 10 Zelllinien	0,1952	0,1749 bis 0,2155	Ja	<0,0001
6 Zelllinien vs. MSCs	0,07673	0,05648 bis 0,09699	Ja	<0,0001
7 Zelllinien vs. 8 Zelllinien	0,02216	0,001904 bis 0,04241	Ja	0,0193
7 Zelllinien vs. 9 Zelllinien	0,08408	0,06383 bis 0,1043	Ja	<0,0001
7 Zelllinien vs. 10 Zelllinien	0,1343	0,1140 bis 0,1546	Ja	<0,0001
7 Zelllinien vs. MSCs	0,01583	-0,004430 bis 0,03608	Nein	0,2831
8 Zelllinien vs. 9 Zelllinien	0,06192	0,04167 bis 0,08218	Ja	<0,0001
8 Zelllinien vs. 10 Zelllinien	0,1121	0,09188 bis 0,1324	Ja	<0,0001
8 Zelllinien vs. MSCs	-0,006333	-0,02659 bis 0,01392	Nein	0,9929
9 Zelllinien vs. 10 Zelllinien	0,05021	0,02996 bis 0,07047	Ja	<0,0001
9 Zelllinien vs. MSCs	-0,06826	-0,08851 bis -0,04800	Ja	<0,0001
10 Zelllinien vs. MSCs	-0,1185	-0,1387 bis -0,09822	Ja	<0,0001

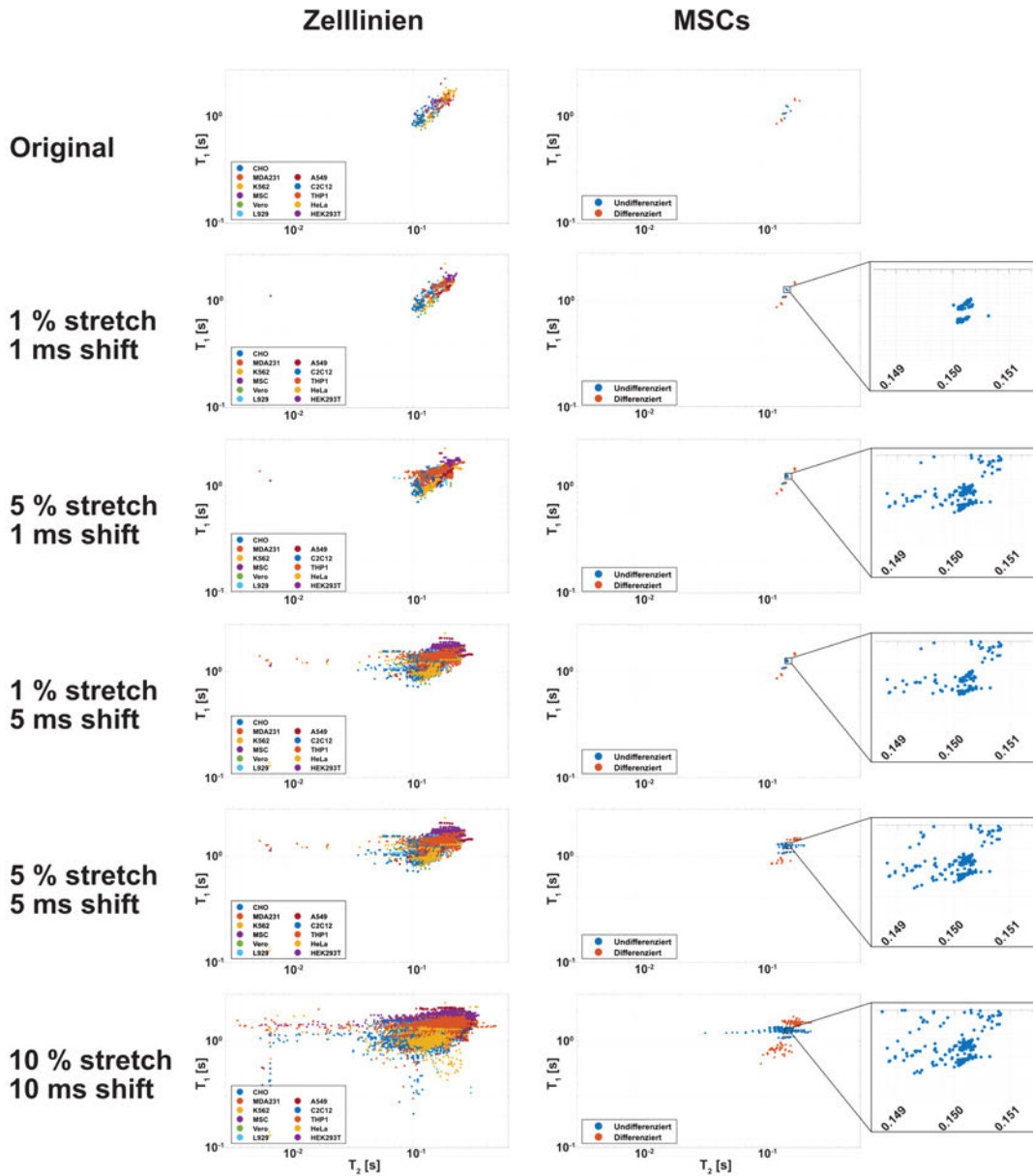


Abbildung 37 – Vergleich unterschiedlicher Augmentierungsparameter: Um den Einfluss der unterschiedlichen Augmentierungsparameter zu untersuchen, wurden die Daten mit einem Augmentierungsfaktor von 40 und unterschiedlichen Augmentierungsparametern vervielfältigt. Anschließend wurden aus den generierten *.png-Dateien erneut der gewichtete Schwerpunkt berechnet. Als Referenz wurden die nicht-augmentierten Daten verwendet. Bei einem Strecken / Stauchen von 1 % und einem Verschieben (engl. shift) von 1 ms resultierte dies in lokal stark gruppierten Datenwolken. Wurde der Faktor zum Strecken / Stauchen der Daten auf 5 % erhöht, während der shift weiterhin bei 1 ms lag, führte dies bereits zu einem leichten Versatz der einzelnen Datenpunkten. Dieser Versatz wurde deutlich erhöht, wenn das Strecken / Stauchen bei 1 % belassen wurde, aber der shift auf 5 ms erhöht wurde. Wurden die Datenpeaks zusätzlich noch mit 5 % gestreckt / gestaucht, führte dies zu keiner erkennbaren Veränderung der Datenpunkte im Vergleich zu 1 % Strecken / Stauchen. Wenn beide Parameter auf 10 erhöht wurden, sorgt dies dafür, dass die – ursprünglich lokal gruppierten – Datenpunkte sehr stark abgelenkt wurden. Da bei den MSCs nur zwei Gruppen betrachtet wurden und diese auch deutlich weniger ursprüngliche Daten beinhaltet hatten, waren die beschriebenen Effekte zwar genauso zu beobachten, jedoch auf die gesamten Proben bezogen, nicht so ausgeprägt wie bei den Zelllinien. Abbildung adaptiert von [122].

4.5.2 Etablieren einer Augmentierungsroutine

Der nächste Schritt in der Klassifizierung der Zelllinien war die Verwendung von künstlichen neuronalen Netzwerken (KNNs). Ein häufig diskutierter Aspekt in der Performanz der KNNs ist deren Bedarf an ausreichend Trainingsdaten. Um diesem Bedarf nachkommen zu können, war es notwendig, die Menge der verwendbaren Daten zu maximieren. Da die Erzeugung von biologischen Replikaten zeit- und materialaufwendig war und somit trotz der Etablierung der Robotikplattform nicht die notwendige Anzahl von mehreren hunderten oder tausenden Messungen getätigt werden konnten, wurde hierfür eine Methode des ML verwendet, welche häufig im Zusammenhang mit KNNs Anwendung findet, nämlich die Datenaugmentierung.

Um den Augmentierungsprozess zu visualisieren, wurden von den augmentierten Daten erneut die gewichteten Schwerpunkte berechnet und in einem Scatterplot (separiert nach Zelllinien und MSCs) dargestellt. Es wurden unterschiedliche Werte für die Augmentierungsparameter gewählt. Für jeden Durchlauf wurden die Daten mit einem Augmentierungsfaktor von 40 vervielfacht. Die Visualisierung (Abbildung 37) zeigte, wie die Augmentierung mit 1 % Stauchen / Strecken (engl. *stretch*) und 1 ms verschieben (engl. *shift*) sehr lokal gruppierte Signalcluster bildete. Die gegebenen Werte wurden hierbei gleichmäßig für T_1 und T_2 angewendet. Wenn das Strecken / Stauchen auf 5 % erhöht wurde und das Verschieben gleich blieb, führte dies dazu, dass die Datenpunkte anfangen stärker auseinander zu driften. Dies wurde durch das Erhöhen des *shift* (Verschieben) Parameters auf 5 ms noch deutlich verstärkt. Hier war kaum eine Änderung zu erkennen, wenn zusätzlich auch das Strecken / Stauchen auf 5 % angehoben wurde. Bei 10 % Strecken/Stauchen und 10 ms *shift* traten zahlreiche Datenpunkt deutlich außerhalb des ursprünglichen Signalbereiches auf, wodurch die Signalcluster deutlich verschwammen.

Dieselben Effekte konnten auch für die MSCs beobachtet werden, auch wenn hier die Signalcluster weniger fusionierten, da die Zahl der Signalepunkte kleiner war.

Durch diese Beobachtungen lies sich bereits eine erste Vermutung für die Performanz der KNN basierend auf diesen unterschiedlichen Augmentierungsparametern aufstellen. Da v. a. der *shift*-Parameter Einfluss auf die visuelle Trennbarkeit der Datenpunkte hatte, wurde angenommen, dass eine Augmentierung der Daten mit mehr als 5 ms *shift* vermutlich einen signifikanten Einfluss auf die Trennbarkeit der Proben haben könnte.

Aufgrund der unterschiedlichen Einflüsse wurden nachfolgend die Augmentierungs-

parameter und -faktoren für den konkreten Anwendungsfall optimiert.

4.5.3 Etablieren und Testen eines künstlichen neuronalen Netzwerks zur Klassifikation von MSCs

Für die zuvor verwendete SVM musste die Komplexität der Daten stark reduziert werden. Die Datenmatrix von 300x300 Datenpunkten wurde durch die Schwerpunktberechnung auf einen einzelnen Wert herunter gerechnet. Dies ermöglichte auf der einen Seite die effiziente Kalkulation und Klassifizierung der Daten, allerdings vernachlässigte diese Form der Klassifizierung auch die Anzahl, Form und Ausrichtung der individuellen Signalpeaks.

Daher war der logische nächste Schritt die komplette Datenmatrix mit 300x300 Werten zu klassifizieren. Hierfür wurden künstliche neuronale Netzwerke (KNN) verwendet.

Die Architektur des verwendeten VGG-Netzwerks nutzte zu Beginn mehrere Faltungsschichten (engl. *convolutional layer*), um die Daten zu interpretieren. Anschließend verwendet es klassische voll-vernetzte (engl. *full-connected / dense*) Schichten, um die analysierten Daten zu klassifizieren.

Basierend auf dieser Theorie wurde in Python durch Jannik Stebani (Fraunhofer IIS) eine entsprechende Umgebung erstellt, in welcher basierend auf der Prinzip des VGG-Netzwerks sämtliche Parameter, wie z. B. die Anzahl der convolutional layer, deren Ausprägung, die verwendeten Aktivierungsfunktionen, die Art der Normalisierung, die Anzahl und Größe der Klassifikationsschichten, die drop-out-Rate oder die Position und Anzahl an Pooling Schichten. Außerdem ermöglichte es, die erstellte Umgebung, die Datenmatrix, auf einen beliebigen Abschnitt zu zuschneiden und lieferte eine Möglichkeit, die eingelesenen Daten auf die gewünschte Anzahl an Augmentierungen zu reduzieren, ebenso wie nur definierte Zellen in den Trainingsprozess zu inkludieren. Sie ermöglichte auch die Aufteilung der Daten in Trainings- und Validierungsdaten, wobei sie berücksichtigte, dass jeder erstellte Datensatz auch die gewünschten Zellen enthielt und nicht durch Zufall ein Zelltyp im Training nicht berücksichtigt wurde.

Um die gegebene Fragestellung optimal bearbeiten zu können, war es notwendig, für die untersuchten Zelllinien und die MSCs separate Modelle für die Klassifizierung zu trainieren.

Aufgrund der Komplexität der verwendeten neuronalen Netzwerke und der damit verbundenen Vielzahl an (Hyper-) Parametern wurde von einer individuellen Eva-

luation der einzelnen Parameteroptimierungen in dieser Arbeit abgesehen, da dies keinen inhaltlichen Mehrwert für die hier gestellte Fragestellung geliefert hätte. Auch ist diese Anpassung niemals zu 100 % abgeschlossen, sondern könnte immer weiter verfeinert werden. Da die hier verwendeten Modelle gute bis sehr gute Ergebnisse lieferten, wurde deren Architektur zu einem bestimmten Zeitpunkt eingefroren und weiterhin so verwendet (Unterunterabschnitt 3.2.7.1).

4.5.3.1 Einfluss des Medienpeaks auf die Trainingsgenauigkeit des KNN für die Klassifizierung von MSCs

Um den Einfluss des Medienpeaks auf das Signal zu untersuchen, wurde die 300x300 Datenmatrix auf den Bereich des Zellpeaks zugeschnitten. Hierzu wurde der Medienpeak weggeschnitten (engl. *crop*), so dass die Daten nicht mehr aus dem vollen 300x300 Raster bestanden ($T_1 = 0,30079$ s; $T_2 = 0,04062$ s). Anschließend wurde ein KNN mit der etablierten Architektur auf diesen Daten trainiert und das Ergebnis mit dem der nicht zugeschnittenen Daten verglichen.

Dies wurde exemplarisch für 1 % Strecken / Stauchen und 5 ms Versatz getestet. Hierfür ergab sich im Vergleich ein signifikanter Unterschied zwischen den Daten (Tabelle 15). Demnach beeinflusste der Medienpeak der untersuchten MSC-Proben die Leistung des KNN deutlich. Folglich wurde für die weitere Analyse der MSC-Daten der Medienpeak, sofern nicht anderweitig erwähnt, immer entfernt und das Training nur auf den Signalbereich des / der Zellpeaks aufgebaut.

Tabelle 15: Vergleich der KNN-Leistung bei zugeschnittenen vs. nicht zugeschnittenen Daten von MSCs. Der statistische Vergleich mittels t-Test zeigte, dass die das Wegschneiden des Medienpeaks eine signifikante Verbesserung der KNN Genauigkeit zur Folge hatte.

	1 % Strecken / Stauchen 1 ms Versatz Augmentierungsfaktor = 5 $n_t = 10$
Mittelwert (n = 10) (Kein crop / crop)	38,89 % / 85 %
Standardabweichung	15,04 % / 22,5 %
Ungepaarter t-Test	
P Wert	< 0,0001
P Wert Zusammenfassung	****
Unterschiede zwischen Mittelwerten	- 0,4611 ± 0,08559
95 % Konfidenzintervall	- 0,6409 bis - 0,2813

4.5.3.2 Testen der KNN-Performanz für die Klassifikation von MSCs

Für die Optimierung des KNN-Trainings wurden für jede getestete Kondition zehn unabhängige Trainingsdurchläufe ($n_t = 10$) trainiert. Für die Analyse wurden die resultierenden Genauigkeiten der Trainingsdurchläufe mittels ANOVA und Tukey's multiple comparison Test statistisch auf Unterschiede hin untersucht.

Bei der Analyse der Augmentierungsfaktoren (Abbildung 38-a) zeigte sich, dass ausschließlich der Vergleich von Daten, welche nicht augmentiert wurden, signifikant niedrigere Genauigkeiten lieferte als Modelle, welche auf einer Augmentierung mit einem Augmentierungsfaktor von > 0 basierten. Dies bekräftigte die Notwendigkeit für eine Augmentierung der gemessenen Daten. Da es statistisch keinen Unterschied machte, wie häufig die Daten augmentiert wurden (Abbildung 38), wurde ein Augmentierungsfaktor von fünf als Standardwert für die Augmentierung der MSC-Daten für die weiteren Analysen angenommen.

Beim Vergleich der unterschiedlichen Augmentierungsparameter konnten keine signifikanten Unterschiede gemessen werden (Abbildung 38-b). Demnach war es für die Genauigkeit des trainierten Modells unerheblich, wie die Werte für die Augmentierungsparameter gewählt wurden. Dies passte zur Beobachtung aus 4.5.2, in der beschrieben wurde, dass aufgrund der geringen Probenzahl die Gruppen der gewichteten Schwerpunkte durch die Augmentierung nicht so stark ineinander verschwommen. Aufgrund der durchschnittlichen Genauigkeit der Parameterkombinationen (Tabelle 16) wurde 5 % Strecken / Stauchen und 1 ms Versatz als Standardparameter für die Augmentierung der MSC-Daten gewählt.

Für die Evaluation des KNN-Trainings mit den zuvor definierten Augmentierungswerten (Augmentierungsfaktor = 5; Strecken / Stauchen = 5 %; T_1 / T_2 -Versatz = 1 ms), wurden die Verlustfunktion der Trainingsdurchläufe (Abbildung 38-c) und die Genauigkeit der Testdurchläufe (Abbildung 38-d) verwendet. Die Verlustfunktion war bereits nach drei bis fünf Epochen auf ein Minimum abgefallen, ebenso wie die Genauigkeit der Testdurchläufe nach der selben Zeit auf ein Maximum angestiegen war. Während die Testgenauigkeit im Mittel ($n_t = 10$) bei etwa 85 % lag, war deren Standardabweichung über den kompletten Trainingsdurchlauf hinweg relativ hoch (Abbildung 38-d - leicht transparenter Bereich). Durch die geringe Ausgangszahl an gemessenen Proben lag die Vermutung nahe, dass die wechselhafte Leistung auf eine Varianz in der Zuweisung der Trainings- und Testdaten zurückzuführen war. Je nachdem wie diese durch den Algorithmus zusammengestellt worden war, führte dies zu einem vorteilhaften oder nicht vorteilhaften Testergebnis. Dies wurde auch in der Konfusionsmatrix (Abbildung 38-e) deutlich. Zwar war hier zu erkennen, dass ein

Großteil der untersuchten Proben korrekt zugeordnet worden waren, jedoch gab es auch einige undifferenzierte Proben, welche fälschlicherweise als differenzierte Proben vorhergesagt worden waren. Auch diese Werte basierten auf einer Mittelung aus $n_t = 10$.

Die beschriebenen Metriken und Beobachtungen legen den Schluss nahe, dass es möglich war, differenzierte von undifferenzierten MSCs mit einer Genauigkeit von durchschnittlich 85 % nicht-invasiv voneinander zu trennen.

Tabelle 16: Durchschnittliche Genauigkeit für den Vergleich von unterschiedlichen Augmentierungsparameter Kombinationen. Auch wenn keine statistisch signifikanten Unterschiede detektierbar waren, lag die Kombination aus 1 % Strecken / Stauchen und 5 ms Versatz mit 85 % Genauigkeit am höchsten. STD = Standardabweichung

Parameterkombinationen	Durchschnittliche Genauigkeit ($n_t = 10$)
1 % / 1 ms	83,33 % (STD = 17,57 %)
1 % / 5 ms	85 % (STD = 22,50 %)
5 % / 1 ms	73,89 % (STD = 26,20 %)
5 % / 5 ms	76,67 % (STD = 20,92 %)
5 % / 10 ms	74,44 % (STD = 18,92 %)
10 % / 5 ms	81,11 % (STD = 13,15 %)
10 % / 10 ms	81,11 % (STD = 20,89 %)

4.5.4 Etablieren und Testen eines künstlichen neuronalen Netzwerks zur Klassifikation von Zelllinien

Wie bereits zuvor für die Klassifizierung der MSC-Entwicklung wurde für die Klassifizierung der Zelllinien ein modifiziertes-VGG Netzwerk verwendet. Bereits in der Analyse mittels SVM war deutlich geworden, dass schon ein Zellpeak ausreichte, um die aufgenommenen Daten effizient klassifizieren zu können, je nach Kombination und Anzahl der zu klassifizierenden Zelllinien. Mit diesem Vorwissen war es möglich, die Anzahl der Faltungsschichten im KNN auf vier anstelle von sieben zu reduzieren und so den Trainingsprozess zu beschleunigen.

4.5.4.1 Testen der KNN-Performanz für die Klassifikation von Zelllinien

Um wie bereits zuvor bei den MSCs die Leistung des trainierten KNN-Modells quantifizieren zu können, wurde die finale Genauigkeit basierend auf den Testdaten berechnet. Diese gab an, wie wahrscheinlich es war, dass das trainierte Modell den korrekten Zelltyp vorhersagte. Hierfür wurden für jede getestete Kondition zehn un-

abhängige, technische Replikate ($n_t = 10$) berechnet und die gemessene Genauigkeit mittels ANOVA und Tukey's multiple comparison Test auf signifikante Unterschiede hin untersucht.

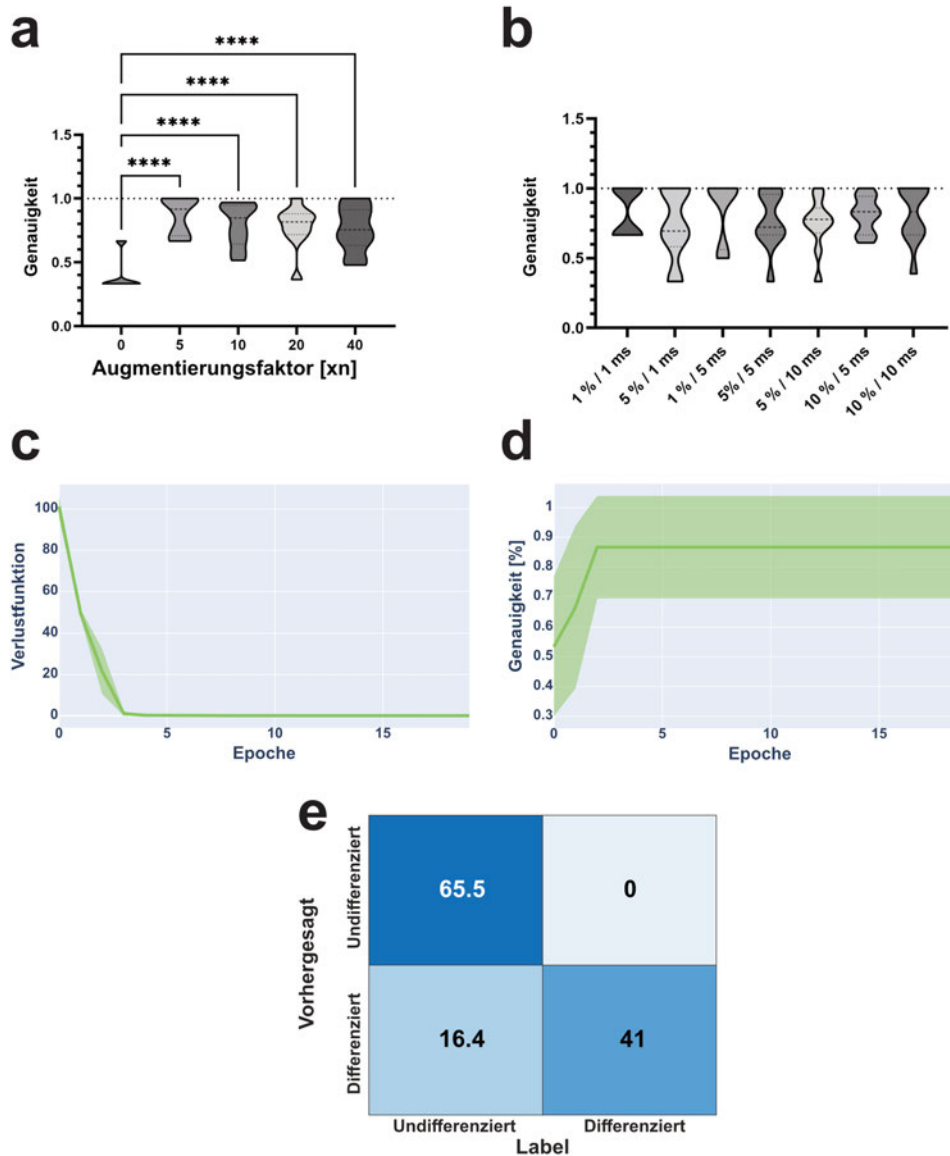


Abbildung 38 – Optimierung der Augmentierungsroutine und Evaluation der Leistung des KNN zur Klassifizierung von MSCs: Die Optimierung der Augmentierungsfaktoren (a) zeigte, dass die Menge der Augmentierung unerheblich war. Statistisch signifikante Unterschiede waren nur für den Vergleich zwischen nicht-augmentierten und augmentierten Daten messbar. Bei der Evaluation der Augmentierungsparameter (b) waren keine signifikanten Unterschiede detektierbar. Demnach war es für die Leistung des KNN unerheblich, wie stark die Datenpeaks im angegebenen Bereich gestaucht / gestreckt bzw. verschoben wurden. Die durchschnittliche Genauigkeit bei 5% Strecken / Stauchen und 1 ms T_1 - bzw. T_2 -Versatz lag im Mittel mit 85% am höchsten. Als Metrik für die Analyse des KNN-Trainingserfolges dienten v. a. die Trainings-Verlustfunktion (c) wie auch die Testgenauigkeit (d). Beide Verläufe erreichten ihren finalen Wert nach etwa drei bis fünf Durchläufen (Epochen). Im Vergleich zur Verlustfunktion wies die Genauigkeit jedoch eine relativ hohe Standardabweichung auf, was ggf. durch die Zuordnung der Trainings- und Testdaten begründet werden kann. Diese Abweichung war auch in der Konfusionsmatrix (e) erkennbar. Zwar war der Großteil der Proben korrekt zugeordnet, jedoch war auch ein Teil der undifferenzierten Zellen fälschlicherweise als differenzierte vorhergesagt. $n_t = 10$ für alle Teilabbildungen. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$; *Abbildungadaptiert von [122].*

Bei der Analyse der unterschiedlichen Augmentierungsfaktoren bei 10 Zelllinien (Abbildung 39-a) zeigte der Vergleich, dass die unaugmentierten Werte im Vergleich zu den augmentierten Werten signifikant schlechtere Ergebnisse lieferten. Dabei war es unerheblich, wie häufig die Daten augmentiert wurden. Keine der untersuchten, augmentierten Proben lieferte signifikante Unterschiede im Vergleich zu anderen, augmentierten Datensätzen. V. a. der Unterschied zu den nicht augmentierten Daten begründete die Bedeutung der Augmentierung für die Leistung des trainierten KNN zur Klassifizierung der MSCs. Diese Beobachtung konnte auch für eine geringere Zahl an untersuchten Zelllinien bestätigt werden (Abbildung 39-c). Jedes Mal wies nur der Vergleich zwischen den nicht augmentierten Daten und allen weiteren Proben signifikante Unterschiede auf.

Für die Analyse der Augmentierungsparameter wurden ebenfalls je Kondition zehn unabhängige Netzwerke trainiert ($n_t = 10$) und deren Genauigkeiten miteinander verglichen. Es zeigten sich statistisch signifikante Unterschiede beim Vergleich zwischen den Proben, welche mit 1 ms Versatz augmentiert wurden (sowohl für 1 % als auch 5 % strechen / stauchen) im Vergleich zu allen anderen Konditionen. Die anderen Konditionen (5 ms, 10 ms, 5 %, 10 %) zeigten untereinander keine signifikanten Unterschiede. Dies legte den Schluss nahe, dass v. a. der Versatz in T_1 und T_2 ausschlaggebend für die Genauigkeit des trainierten Netzwerks war. Dies unterstützte auch die Hypothese, welche in 4.5.2 aufgestellt wurde, nämlich dass v. a. der Versatz zu einer gestiegenen Unschärfe in den Daten führte. Da die Genauigkeit bei mehr als 1 ms Versatz signifikant sank, wurden die optimalen Augmentierungsparameter für die Spektren der Zelllinien auf 5 % Strecken / Stauchen und 1 ms Versatz festgelegt.

Wie schon bei der Klassifizierung der Zelllinien mittels SVM wurde auch hier für die unterschiedliche Anzahl an zu klassifizierenden Zellen unabhängige Modelle berechnet ($n_t = 10$). Die unterschiedlichen Durchläufe zeigten die sinkende Leistung des Netzwerks mit zunehmender Zahl an zu klassifizierenden Zelllinien (erkennbar an der steigenden Zahl der Missklassifizierungen bzw. der sinkenden Zahl der korrekten Klassifizierungen) (Abbildung 40 / Abbildung 41 / Abbildung 42). Auch war dies an der Anzahl an Epochen abzulesen, welche es brauchte, um den finalen Wert der verwendeten Metriken (Verlustfunktion und Genauigkeit) zu erreichen. Für die Zusammensetzung der Zellproben wurden für nahezu jede Zahl an zu klassifizierenden Zelllinien alle möglichen Kombinationen gemessen. I. d. R. zeigte der statistische Vergleich zwischen den Ergebnissen jedoch keine detektierbaren Unterschiede. In diesen

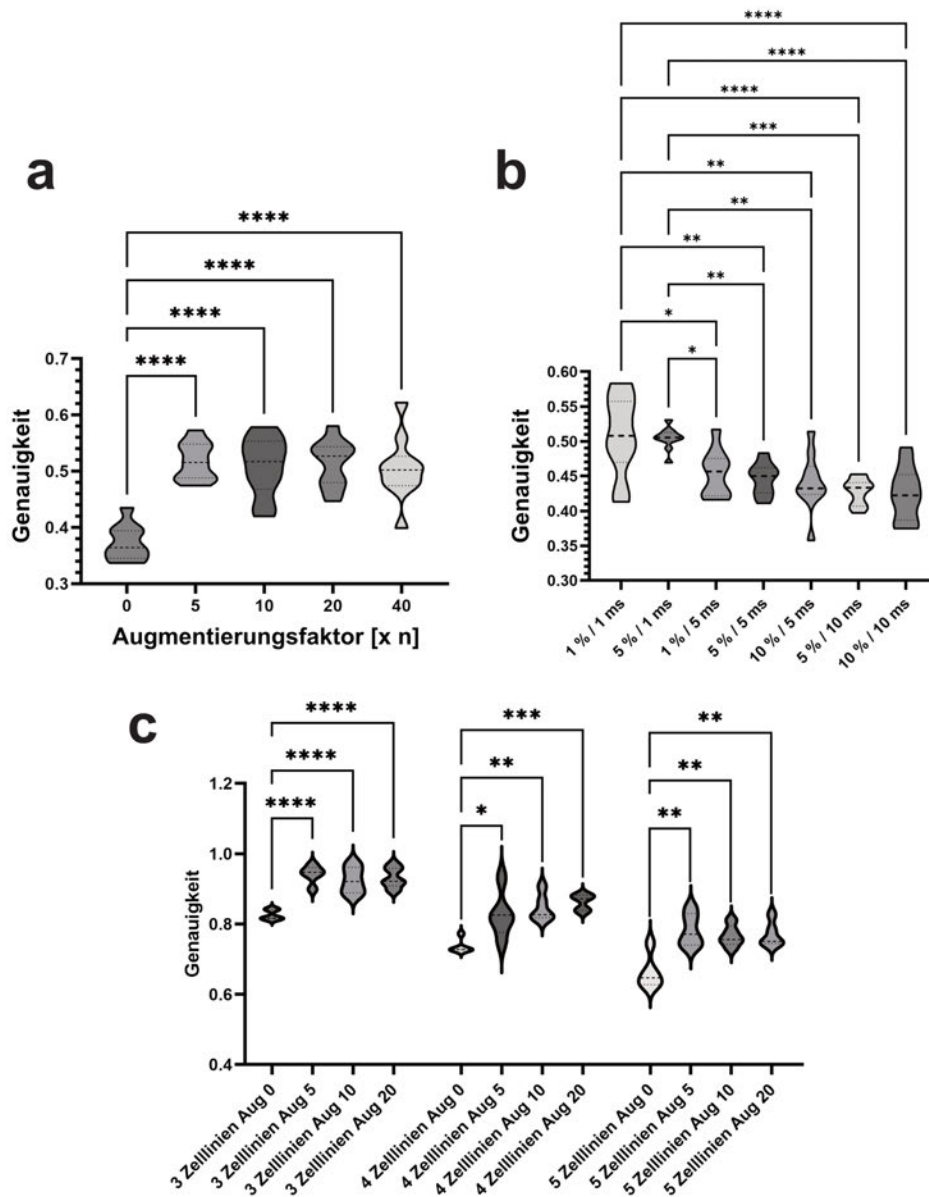


Abbildung 39 – Testen von unterschiedlichen Augmentierungsfaktoren und -parametern auf die Performanz der etablierten KNN-Architektur zur Klassifizierung von Zelllinien: Für die Evaluation von unterschiedlichen Augmentierungsfaktoren wurden je Augmentierungsfaktor zehn unabhängige Netzwerke trainiert (10 Zelllinien; $n_t = 10$) und deren resultierende Genauigkeit statistisch mit einander verglichen (ANOVA und Tukey's Test mit Alpha-Niveau von 5%) (a). Dabei zeigte sich, dass es keine signifikanten Unterschiede zwischen allen verwendeten Augmentierungsfaktoren gab. Jedoch gab es einen signifikanten Unterschied im Vergleich zu den nicht augmentierten Daten (Augmentierungsfaktor = 0). Auch beim Vergleich von anderen verwendeten Zelllinien blieb diese Beobachtung bestehen (c). Dasselbe Vorgehen wurde für die Evaluation der Augmentierungsparameter verwendet (b). Für jeden Parameter wurden ebenfalls zehn unabhängige Trainingdurchläufe durchgeführt ($n_t = 10$). Die statistische Analyse zeigte, dass es v. a. zwischen den Proben mit 1 ms Versatz und allen anderen zu statistisch relevanten Unterschieden kam. Das Strecken und Stauchen hatte keinen signifikanten Einfluss auf die Genauigkeit des trainierten Modells. Basierend auf diesen Ergebnissen wurden die Augmentierungsparameter für die Spektren der Zelllinien auf 5% Strecken / Stauchen und 1 ms Versatz festgelegt. Abbildung adaptiert von [122]. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$

Fällen wurde dann die Kombination an Zellen verwendet, welche im Mittel ($n_t = 10$) den höchsten Durchschnittswert aufwies.

Aus den verwendeten Metriken (Testgenauigkeit und Trainingsverlustfunktion) war ersichtlich, dass mit steigender Zahl an zu klassifizierenden Zellen zwar der Zeitpunkt des Erreichens des Genauigkeitsmaximums mehr oder weniger gleich blieb, dieses im Wert jedoch deutlich abnahm. Sobald mehr als drei Zelllinien klassifiziert werden sollten, erreichte die Trainingsverlustfunktion nicht mehr vor dem Ende des Trainingsprozesses ihr Minimum. Außerdem reduzierte sich die Steigung der Verlustfunktion mit zunehmender Zahl an Zelllinien, sodass auch der finale Wert der Verlustfunktion stetig größer wurde. In weiterführenden Untersuchungen (4.5.4.3) wurde jedoch deutlich, dass die Verdopplung der Trainingszeit von 12 auf 25 Epochen, tendenziell bessere Ergebnisse lieferte, diese Unterschiede jedoch statistisch nicht signifikant belegbar waren. Demnach führte zwar die gestiegene Zahl an zu klassifizierenden Zelllinien zu einem vermeintlichem Bedarf an längerem Training, jedoch hätte dieses auch keine signifikant höheren Genauigkeiten erzielt.

4.5.4.2 Einfluss der drop-out-Rate auf die Leistung des KNN

Die drop-out-Rate konnte im verwendeten Framework für die Klassifikation individuell angepasst werden. Die verwendeten Werte spiegelten relative Werte in Prozent wider. Um die optimalen Werte zu definieren, wurde das Training für die Klassifizierung der MSCs und Zelllinien mit insgesamt vier unterschiedlichen drop-out-Werten durchgeführt (0 % = Kein drop-out; 25 %; 50 %; 75 %). Für jeden Wert wurden zehn unabhängige KNNs trainiert ($n_t = 10$) und deren resultierende Genauigkeit gemittelt. Diese Mittelwerte wurden anschließend mittels ANOVA (Alpha-Niveau = 5 %) und Tukey's Multiple Comparison Test verglichen.

Die Auswertung zeigte für die getesteten MSCs, dass ab einer drop-Out Rate von 75 % die Genauigkeit des trainierten KNNs signifikant absank. Auch wenn die restlichen drop-out-Raten keine signifikanten Unterschiede zeigten, lag der Mittelwert bei 25 % drop-out (59 %; $n_t = 10$) jedoch deutlich unter dem von ohne (w/o) drop-out (81,55 %; $n_t = 14$) und 50 % drop-out (78,33 %; $n_t = 10$) (Abbildung 44-a).

Bei der Analyse der Daten der Zelllinien zeigten sich ebenfalls nur im Vergleich zu 75 % drop-out signifikante Unterschiede. Im Gegensatz zu den MSCs, waren diese Unterschiede jedoch im Vergleich zu allen anderen drop-out-Raten (w/o: 91,62 %; 25 %: 92,81 %; 50 %: 91,71 %) mit einem p -Wert von $< 0,001$ messbar (Abbildung 44-b).

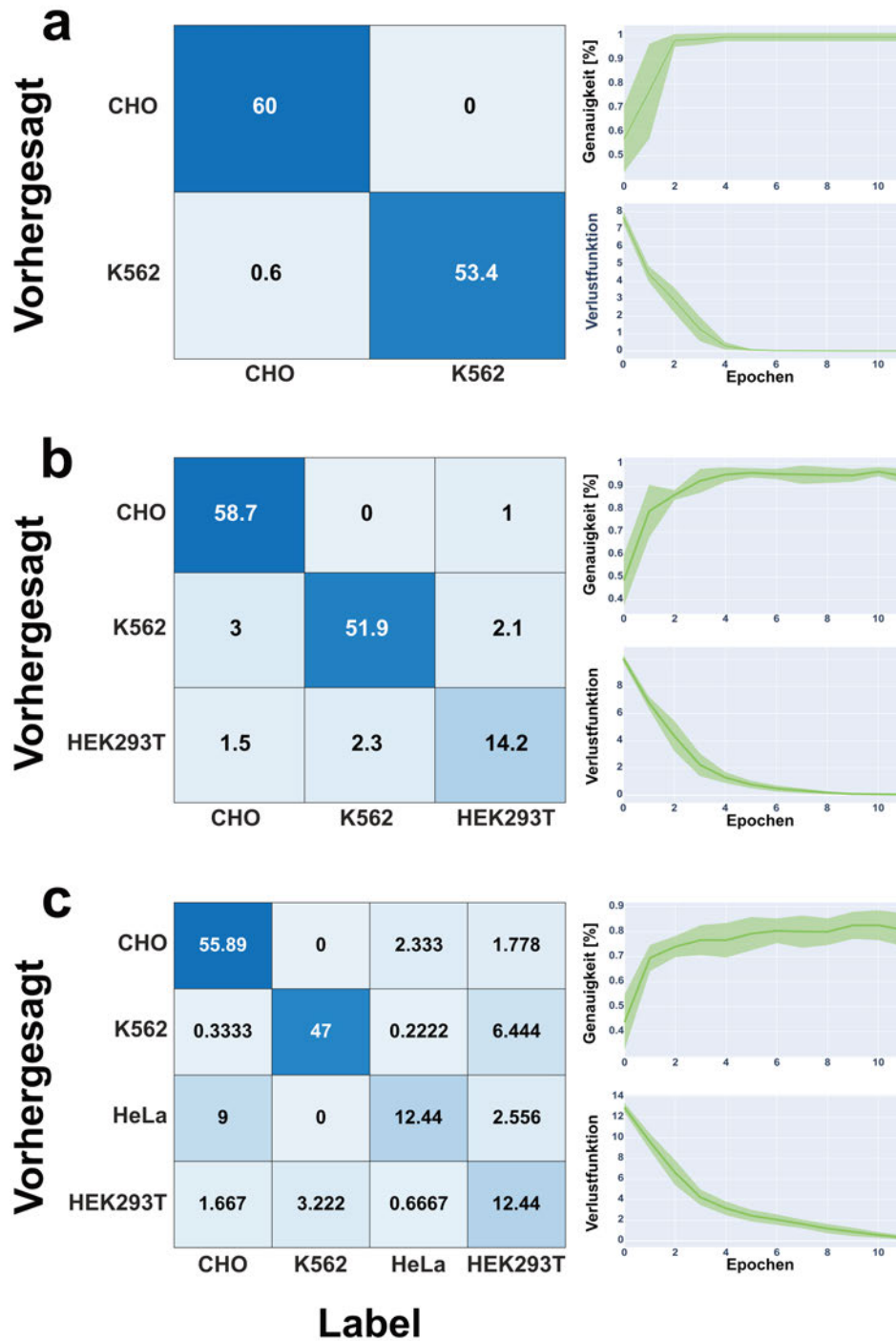


Abbildung 40 – Performanz des KNN zur Klassifikation von Zelllinien (Teil 1): Zur Evaluation der Leistung des KNNs wurden die Konfusionsmatrix, die Testgenauigkeit und Trainingsverlustfunktion verwendet. Bei der Klassifizierung von nur zwei Zelllinien (a) wurden nahezu in allen der durchgeführten Testläufe ($n_t = 10$) alle Zellen korrekt zugeordnet. Die gemessene Genauigkeit erreichte bereits nach zwei Trainingsepochen ihr Maximum und die Verlustfunktion war nach vier bis sechs Epochen auf ein Minimum abgefallen. Bei der Klassifikation von drei Zelllinien (b) wurden auch nahezu alle der Zellen in den durchgeführten zehn Wiederholungen korrekt klassifiziert. Im Gegensatz zur Klassifizierung von zwei Zelllinien erreichte die Trainingsgenauigkeit erst nach vier Trainingsepochen ihr Maximum und auch die Verlustfunktion war etwa nach acht Trainingsepochen auf ihr Minimum abgefallen. Dieser Trend führte sich auch bei der Klassifizierung von vier Zelllinien (c) weiter. Hier stieg die Genauigkeit nach etwa sechs bis acht Epochen auf einen relativ konstanten Wert an. Auch wenn die Verlustfunktion nach zehn Epochen augenscheinlich noch nicht auf ihren finalen Wert abgefallen war, wurde zum Zwecke der Vergleichbarkeit das Training nach 12 Epochen abgebrochen. Die Menge an Missklassifizierungen lag, basierend auf der Konfusionsmatrix, in etwa auf demselben Niveau wie bei der Klassifizierung von vier Zelllinien. Die Werte der Konfusionsmatrix repräsentieren den Mittelwert nach $n_t = 10$. Abbildung adaptiert von [122].

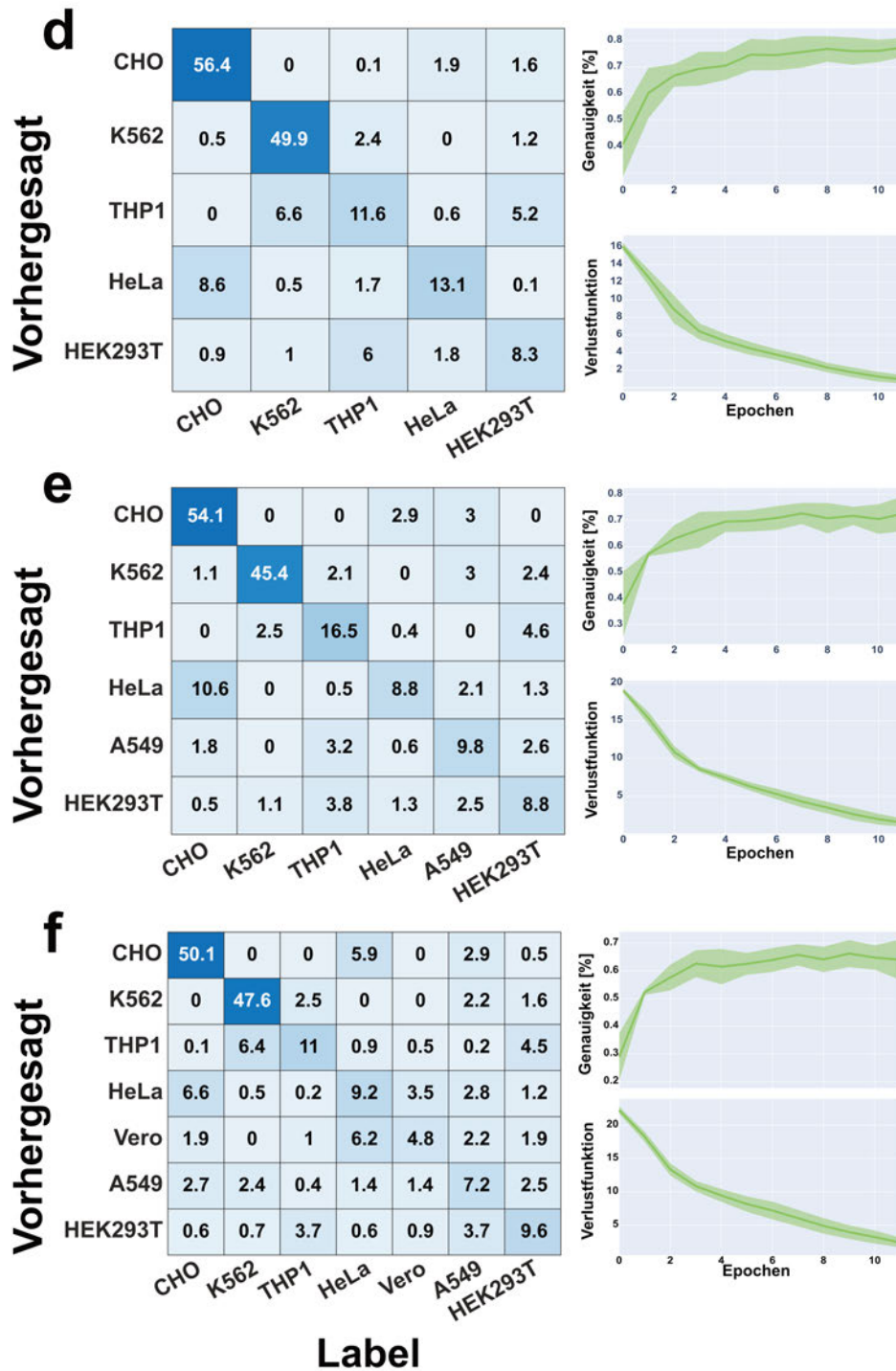


Abbildung 41 – Performanz des KNN zur Klassifikation von Zelllinien (Teil 2): Bei der Klassifizierung von fünf Zelllinien stieg die Anzahl an missklassifizierten Zellen in der Konfusionsmatrix etwas an (d). Dies wurde auch durch eine kleinere finale Trainingsgenauigkeit deutlich. Im Vergleich zur Klassifizierung von vier Zelllinien, blieb der Zeitpunkt in etwa gleich, bei welchem der finale Wert der Genauigkeit erreicht wurde. Die Verlustfunktion viel hingegen etwas weniger steil ab und war, wie schon bei c, nach 12 Epochen nicht auf ihren minimalen Wert abgefallen. Der beschriebene Trend setzte sich auch für die Klassifizierung von sechs (e) und sieben (f) Zelllinien fort. Mit steigender Anzahl zu klassifizierenden Zellen stieg auch die Zahl an missklassifizierten Zellen in der Konfusionsmatrix. Die Testgenauigkeit erreichte ihr Plateau immer um denselben Zeitpunkt (ca. sechs bis acht Epochen), jedoch mit immer weiter sinkendem Maximum und die Steigung der Trainingsverlustfunktion flachte immer weiter ab. Die Werte der Konfusionsmatrix repräsentieren den Mittelwert nach $n_t = 10$. Abbildung adaptiert von [122].

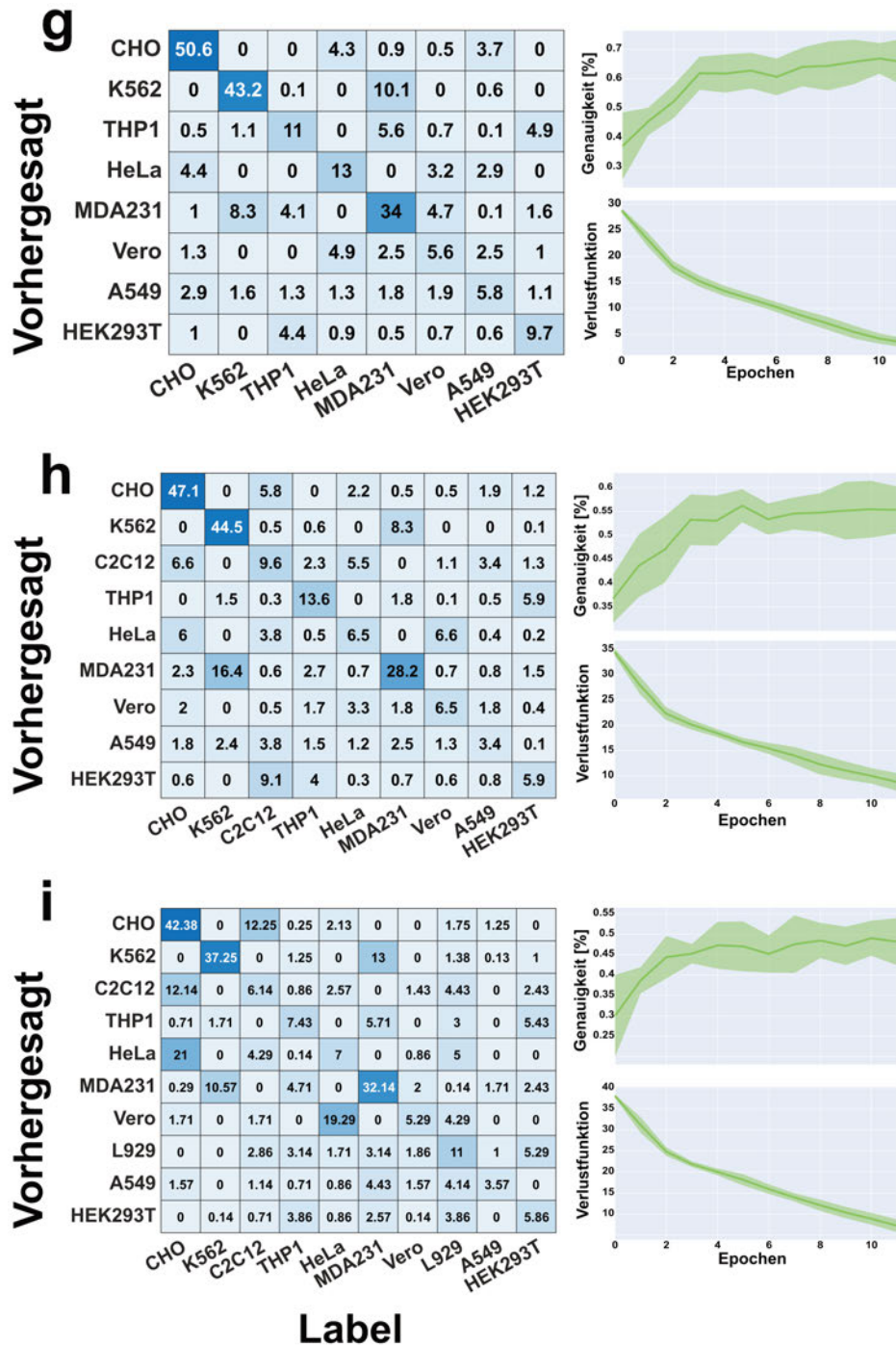


Abbildung 42 – Performanz des KNN zur Klassifikation von Zelllinien (Teil 3): Wie bereits in Abbildung 41 beschrieben, setzte sich auch für die Klassifizierung von acht (g), neun (h) und zehn (i) Zelllinien der Trend fort, dass mit steigender Zahl an zu klassifizierenden Zelllinien zwar die Testgenauigkeit ihr Maximum nach etwa sechs bis acht Epochen erreicht hatte, dieses jedoch immer geringer wurde. Die sinkenden Genauigkeit spiegelte sich auch in der Zahl an missklassifizierten Zellen in der Konfusionsmatrix wider. Außerdem flachte die Steigung der Trainingsverlustfunktion immer weiter ab, sodass deren Wert bei 12 Epochen mit zunehmender Anzahl an zu klassifizierenden Zelllinien immer größer wurde. Die Werte der Konfusionsmatrix repräsentieren den Mittelwert nach $n_t = 10$. Abbildung adaptiert von [122].

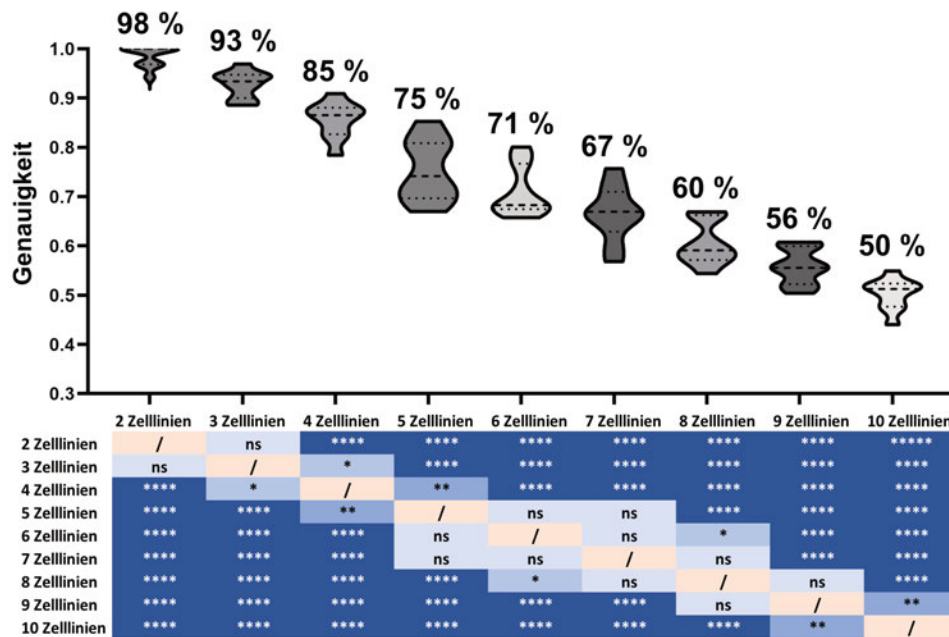


Abbildung 43 – Vergleich der resultierenden Genauigkeiten der trainierten KNNs bei unterschiedlicher Anzahl von Zelllinien: Die Evaluation der finalen Testgenauigkeit der trainierten KNNs zeigte, dass diese mit der Anzahl an zu klassifizierenden Zelllinien, wie bereits in Abbildung 40, Abbildung 41, Abbildung 42 gezeigt, sank. Anfänglich lag die Testgenauigkeit bei 98 % (2 Zelllinien) und sank dann bis auf 50 % (10 Zelllinien) ab. Der statistische Vergleich von zehn unabhängigen Trainingsdurchläufen ($n_t = 10$) zeigte, dass die meisten Kombinationen an Zelllinien signifikante Unterschiede zueinander aufwiesen. Abbildung adaptiert von [122]. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$

Die beschriebenen Beobachtungen legten den Schluss nahe, dass in der vorliegenden Fragestellung, eine zu hohe drop-out-Rate die Leistung des KNN signifikant beeinflussen konnte. Da für die Klassifizierung der MSCs die gewählten drop-out-Raten offenbar keinen Einfluss bzw. ggf. sogar einen negativen Einfluss auf die finale Leistung des KNNs haben konnten und die graphische Auswertung der Trainingsläufe zeigte, dass es zu keinem erkennbaren Overfitting kam, wurde für deren Klassifizierung entschieden, keine drop-out-Layer zu verwenden. Für die Klassifikation der Zelllinien zeigten sich keine signifikanten Unterschiede zwischen den drop-out-Raten $< 75\%$. Da hier jedoch eine drop-out-Rate mit 25% im Mittel ($n_t = 10$) die höchste Genauigkeit lieferte, wurde entschieden für die Klassifikation der Zelllinien fortan drop-out-Schichten mit einer drop-out-Rate von 25% zu verwenden.

4.5.4.3 Einfluss der Anzahl an Trainingsepochen auf die Leistung des KNN

Die Leistung eines KNN hängt maßgeblich auch von der Anzahl an Trainingsepochen ab, welche verwendet wurde, um das Modell zu trainieren. Ab einem gewissen

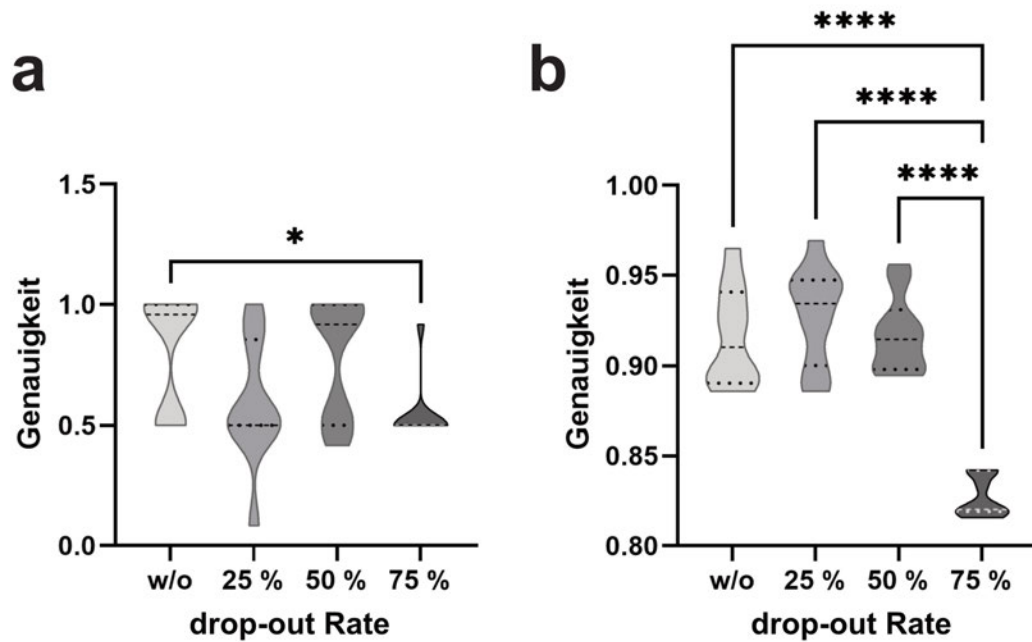


Abbildung 44 – Einfluss der drop-out-Rate auf die Performanz des KNN: Für die Evaluation des Einflusses der drop-out-Rate wurden, je getestetem Wert, zehn unabhängige Trainingsdurchläufe für die etablierten KNN-Architekturen durchgeführt. Für die MSCs (a) zeigte der statistische Vergleich der resultierenden Genauigkeiten, dass eine drop-out-Rate von 75 % zu signifikant schlechteren Ergebnissen führte – im Vergleich zu einem Training ohne drop-out-Layer. Für die Klassifizierung von drei Zelllinien (b) zeigten sich erneut für eine drop-out-Rate von 75 % signifikant schlechtere Ergebnisse der KNN-Trainings. Hier war jedoch der Vergleich zu jeder getesteten drop-out-Rate signifikant unterschiedlich. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$

Zeitpunkt, der je nach Fragestellung, Trainingsdaten und verwendeten Architektur, sehr variabel ist, verändern sich die verwendeten Metriken nicht mehr stark. Ab diesem Zeitpunkt kann dann das Training gestoppt werden. Dies impliziert auch, dass ein längeres Training nicht immer ein zuverlässigeres Netzwerk erzeugt. Daher sollte für ein KNN-Training auch immer die Anzahl der Trainingsepochen validiert werden.

Hierfür wurde ein Netzwerk mit der etablierten Architektur auf den Daten trainiert. Dabei wurden die Ergebnisse für die augmentierten Daten (10 Zelllinien, 5 % Strecken / Stauchen und 1 ms Versatz) gemittelt ($n_t = 10$) und über unterschiedliche Augmentierungsfaktoren hinweg für 12 und 25 Trainingsepochen aufgetragen und statistisch verglichen (Abbildung 45-a).

Es zeigte sich dabei, dass zwar für jeden getesteten Augmentierungsfaktor das Training mit 25 Epochen tendenziell bessere Ergebnisse (d. h. höhere Genauigkeiten) lieferte, jedoch zeigte die statistische Analyse keine signifikanten Unterschiede (Abbildung 45-a).

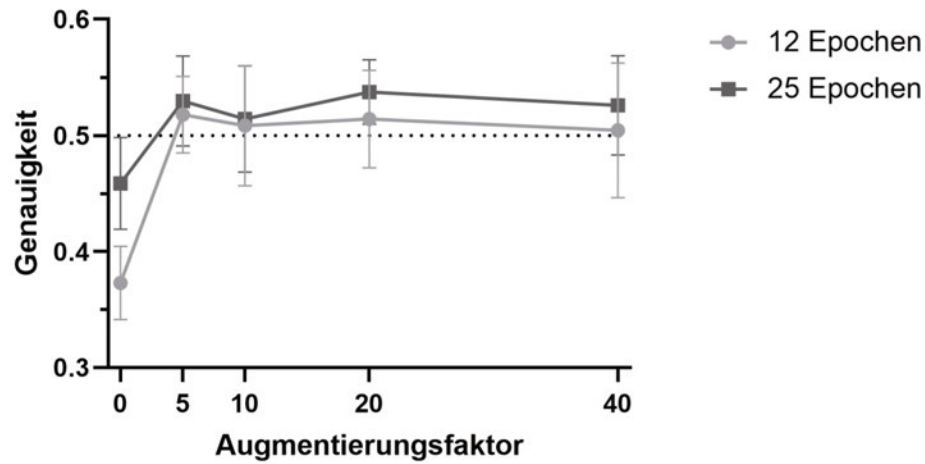


Abbildung 45 – Einfluss der Anzahl an Trainingsepochen auf die Performanz des KNN: Der Datensatz aus 10 unterschiedlichen Zelllinien wurde mit 5 % Strecken / Stauchen und 1 ms Versatz unterschiedlich häufig augmentiert und die resultierende Genauigkeit bei einer Trainingsdauer von 12 und 25 Epochen verglichen. Außer bei einem Augmentierungsfaktor von 0 zeigten sich keine statistisch signifikanten Unterschiede zwischen den unterschiedlichen Trainingslängen (Tabelle 17). Abbildung adaptiert von [122].

Nur der Vergleich der nicht augmentierten Daten zeigte, dass hier ein längeres Training mit 25 Epochen signifikant bessere Ergebnisse lieferte (Tabelle 17). Dies belegt erneut die Beobachtungen, welche bereits in 4.5.4.1 beschrieben wurden, nämlich dass jede Form der Augmentierung zu signifikant verbesserten Ergebnissen führte.

Tabelle 17: Durchschnittliche Werte und Ergebnisse des t-tests je Augmentierungsfaktor für den Vergleich von 12 und 25 Trainingsepochen. Beim Vergleich der gemessenen Genauigkeiten für 10 Zelllinien und unterschiedlichen Augmentierungsfaktoren bei einer Trainingsdauer von 12 bzw. 25 Epochen zeigte sich, dass nur die nicht augmentierten Daten signifikante Unterschiede aufwiesen. Bei allen weiteren getesteten Augmentierungsfaktoren waren keine signifikanten Unterschiede zwischen den unterschiedlichen Epochenzahlen festzustellen.

	Aug. Faktor = 0	Aug. Faktor = 5	Aug. Faktor = 10	Aug. Faktor = 20	Aug. Faktor = 40
Mittelwert (n = 10) (12 / 25 Epochen)	37,28 % / 45,87 %	51,78 % / 52,95 %	50,85 % / 51,42 %	51,41 % / 53,73 %	50,44 % / 52,58 %
Standardabweichung (12 / 25 Epochen)	3,161 % / 3,962 %	3,324 % / 3,876 %	5,177 % / 4,597 %	4,197 % / 2,784 %	5,805 % / 4,283 %
Ungepaarter t-Test					
p-Wert	<0,0001	0,4752	0,7965	0,1635	0,3619
p-Wert Zusammenfassung	****	ns	ns	ns	ns
Unterschiede zwischen Mittelwerten	0,08587 ±0,01603	0,01178 ±0,01615	0,005731 ±0,02189	0,02314 ±0,01593	0,02134 ±0,02281
95 % Konfidenzintervall	0,05220 bis 0,1195	-0,02215 bis 0,04570	-0,04027 bis 0,05173	-0,01032 bis 0,05660	-0,02659 bis 0,06927

4.5.4.4 Einfluss von Medienpeaks auf die Leistung des KNN

Neben der Architektur des Netzwerks spielte auch die Aufbereitung der Daten eine wichtige Rolle in der Genauigkeit der Klassifizierung. Hiermit ist nicht die Aufbereitung der Messdaten durch das Messsystem gemeint, sondern welche Anpassungen an den Messdaten vorgenommen werden können, nachdem sie final durch das Messsystem ausgegeben wurden. Hierzu zählte v. a. das Zuschneiden der Daten auf definierte Bereiche. Mit dieser Methode sollte untersucht werden, inwiefern die einzelnen Signalpeaks die Leistung des KNNs positiv oder negativ beeinflussen.

Hierzu wurde der Medienpeak weggeschnitten (engl. *crop*), sodass die Daten nicht mehr aus dem vollen 300x300 Raster bestanden ($T_1 = 0 - 3,0079\text{ s}$; $T_2 = 0 - 0,4062\text{ s}$). Anschließend wurden erneut zehn unabhängige Modelle je Zellkombination trainiert und deren Genauigkeiten statistisch mit denen der Modelle verglichen, welche auf nicht zugeschnittenen Daten basierten (Abbildung 46).

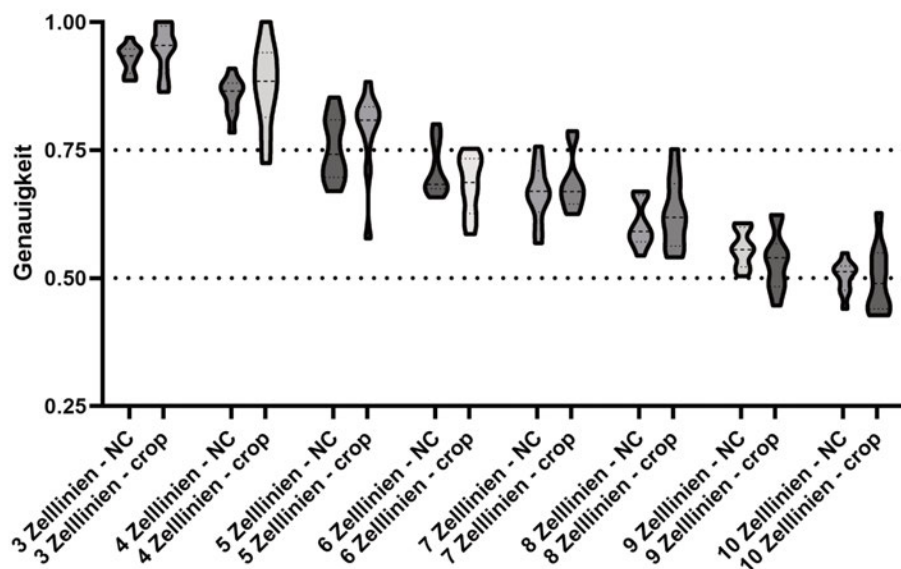


Abbildung 46 – Einfluss des Medienpeaks auf die Leistung des etablierten KNN: Um den Einfluss des Medienpeaks auf die Genauigkeit des trainierten Modells zu testen, wurden die Trainings- und Testdaten auf den Bereich zwischen $T_1 = 0 - 3,0079\text{ s}$ und $T_2 = 0 - 0,4062\text{ s}$ zugeschnitten (engl. *crop*). Dies wurde für jede Kombination an Zellen einzeln getestet und das Ergebnis aus zehn unabhängigen Trainingsdurchläufen ($n_t = 10$) statistisch mittels ANOVA (Alpha Niveau = 5 %) und Tukey’s multiple comparison Test verglichen. Dabei zeigten sich keine statistisch signifikanten Unterschiede zwischen den zugeschnittenen und nicht zugeschnittenen Trainings- und Evaluationsdurchläufen. Der Vergleich der jeweiligen Mittelwerte zeigte zwar, dass i. d. R. die Modelle, welche auf den zugeschnittenen Daten beruhten, höhere Genauigkeiten erzielt hatten, jedoch waren diese Unterschiede statistisch nicht messbar. Abbildung adaptiert von [122].

Die Analyse (ANOVA mit Alpha-Niveau von 5 % kombiniert mit Tukey’s multiple comparison Test) zeigte keine signifikanten Unterschiede. Zwar lagen die durchschnittlichen Genauigkeiten der Modelle, welche auf den zugeschnittenen Daten ba-

sierten, höher im Vergleich zu den nicht zugeschnittenen Modellen, jedoch ohne detektierbare Signifikanz.

Da das Zuschneiden der Daten auf den Medienpeak das Risiko mit sich brachte, in zukünftigen Experimenten ggf. mögliche Einflüsse von Signalen außerhalb des aktuell definierten Bereiches zu vernachlässigen, wurde entschieden, die Daten vorerst nicht weiter zu zuschneiden.

4.5.5 Zusammenfassung

Anhand von Daten, welche von weit verbreiteten Zelllinien wie auch mesenchymalen Stromazellen erhoben wurden, wurde untersucht, ob diese mittels Methoden des Machine Learnings (ML) klassifiziert werden können. Hierfür wurden unterschiedliche Hyperparameter, wie z. B. die drop-out-Rate oder die Anzahl der verwendeten Trainingsepochen optimiert und statistisch miteinander verglichen. Außerdem wurde der Einfluss der erwähnten Zell- und Medienpeaks beschrieben. Die finale Architektur des KNN und der SVM ermöglichte eine Klassifizierung der Zelllinien bzw. der MSCs mit einer Genauigkeit von bis zu 98 %, je nach Zusammensetzung der Daten.

4.6 Leistung der SVM im Vergleich zum KNN

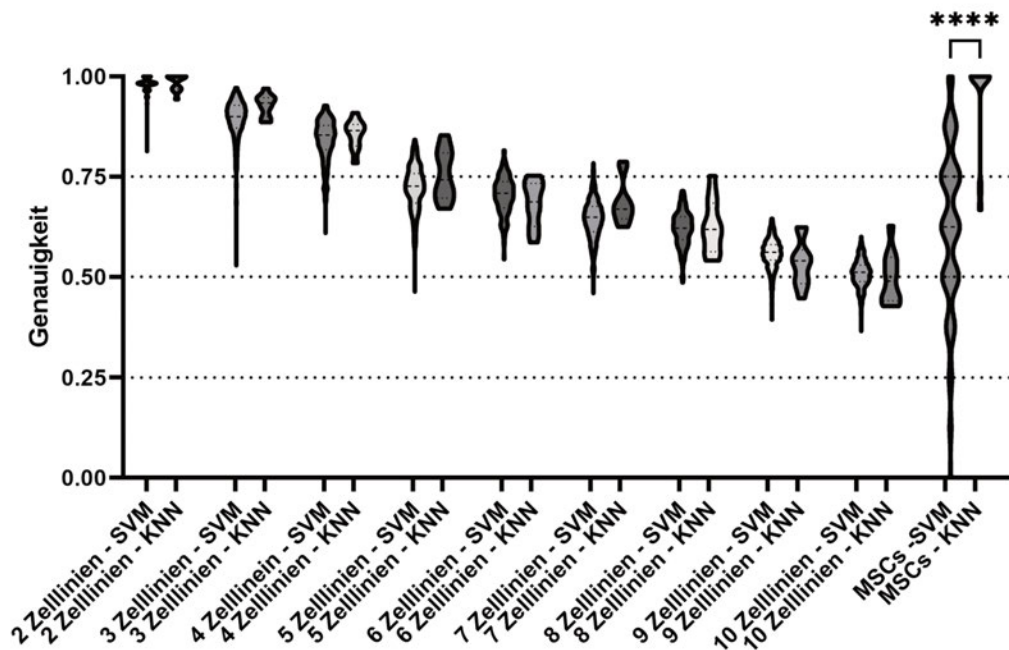
Nachdem im vorangegangenen Abschnitt (4.5) die Architektur und das Training eines KNN bzw. SVM zur Klassifizierung von Zelllinien und MSCs angepasst wurde, sollte abschließend untersucht werden, inwiefern sich die aufgezeigten Klassifizierungslösungen statistisch voneinander unterscheiden.

Hierfür wurden die erzielten Genauigkeiten der verwendeten Kompositionen von Zelllinien und MSCs statistisch miteinander verglichen (Abbildung 47). Außer beim Vergleich der MSCs kam es zu keinen statistisch signifikanten Unterschieden. Sowohl bei der SVM wie auch beim KNN sank die Genauigkeit mit zunehmender Zahl an berücksichtigten Zelllinien von anfänglich fast 100 % auf ca. 50 % ab. Auch wenn die angewandten statistischen Tests (ANOVA mit 5 % Signifikanzniveau gekoppelt mit Tukey's multiple comparison Test) für die Zelllinien keine Signifikanzen anzeigten, war v. a. für eine geringe Anzahl an Zelllinien die Streuung der Daten für die SVM erkennbar größer als für die entsprechenden KNN-Trainingsdurchläufe ($n_t = 10$).

Nur beim Vergleich der MSC-Proben zeigten die statistischen Tests signifikante Un-

terschiede mit $p < 0,0001$ an. Demnach war hier die Genauigkeit des KNN gegenüber dem der SVM signifikant höher.

Diese Beobachtungen legten den Schluss nahe, dass es für die Klassifikation der Zelllinien keinen statistischen Unterschied machte, welcher ML-Algorithmus verwendet wurde. Bei der Klassifikation der MSCs war das KNN demnach der SVM deutlich überlegen.



*Abbildung 47 – Abschließender Vergleich der Genauigkeit der Klassifikation mittels SVM und KNN: Der statistische Vergleich der Leistung von KNN und SVM zeigte, dass es für die Klassifizierung von Zelllinien keine signifikanten Unterschiede gab, d. h. dass in beiden Fällen im Mittel die Leistung der Klassifizierung sehr ähnlich war. Für die Klassifizierung der MSC-Differenzierungen, wies das verwendete KNN im Vergleich zur verwendeten SVM signifikant höhere Genauigkeiten auf. Auch war die Streuung der gemessenen Genauigkeiten beim KNN im Vergleich zur verwendeten SVM deutlich geringer. Demnach stellte das KNN für diese Fragestellung die geeignetere Herangehensweise dar. Alle Daten basieren auf $n_t = 10$; Abbildung adaptiert von [122]. ns: nicht signifikant; *: p -Wert $\leq 0,05$; **: p -Wert $\leq 0,01$; ***: p -Wert $\leq 0,001$; ****: p -Wert $\leq 0,0001$*

5 Diskussion

5.1 Zuverlässigkeit der Messdaten

Im Jahr 2018 gab es weltweit etwa 36.000 klinische MR-Scanner, zu denen jährlich etwa 2.500 hinzukommen [129]. Aufgrund dieser Vielzahl an Messungen sind die Vergleichbarkeit und Reproduzierbarkeit von Messungen eines der höchsten Ziele in der MR. Weltweit werden passende Phantome verwendet, um die korrekte Funktionsweise der MR-Scanner zu verifizieren [130]. Diese Phantome können entweder der Verifizierung von Strukturen dienen [131] oder bei quantitativer MR der Verifizierung der erhobenen Messwerte [132, 133]. Ähnlich wie in Abschnitt 4.1, bestehen die Phantome meist aus einem Geliermittel, versetzt mit chemischen Substanzen, um die T_1 - und T_2 -Werte entsprechend anzupassen [134]. Wie auch in der vorliegenden Arbeit werden im klinischen Kontext i. d. R. Routinemessungen mit entsprechenden Phantomen durchgeführt, um sicher zu stellen, dass die Hard- und Software nach wie vor zuverlässig arbeiten. Daher ist es wichtig, dass das Phantom seine chemischen und biophysikalischen Eigenschaften nicht verändert [135–137]. Nur unter Berücksichtigung dieser Phantommessungen kann ein möglicher Effekt in den Messungen richtig eingeordnet werden.

Für eine Beurteilung der Daten wurden auch in dieser Studie Phantommessungen durchgeführt (siehe 4.1.2). Die Phantommessungen dienten hier nicht der Vergleichbarkeit unterschiedlicher Systeme – wie dies in anderen Studien häufig der Fall ist [138–140] –, sondern der Evaluation der Stabilität der Messungen. Anhand der Phantome war auch das Rauschverhalten des Systems gut zu erkennen. Wie in Abbildung 19 dargestellt, fluktuierten die Daten sichtbar um einen konstanten Mittelwert. Dieser Fluktuation könnten unterschiedliche Ursachen zugrunde liegen. Eine mögliche Ursache wäre ein Fehler in der Vorbereitung der Proben. Durch den Hersteller wird die Abweichung der verwendeten Pipetten mit 2% angegeben [141]. Die verwendeten Pipetten wurden stets durch eine qualifizierte Firma justiert, sodass eine Abweichung von mehr als 2% auf Fehler durch den Präparator begründet wären. Da die Messungen mit kurzen Abständen (d. h. wenigen Minuten) (Abbildung 19-a) und längeren Abständen (Abbildung 19-b) – d. h. 24 h zwischen jeder Messung – vergleichbar waren, kann auch von einer Stabilität des verwendeten Messphan-

toms ausgegangen werden. Unter Berücksichtigung des Präparationsfehlers bleibt dann nur eine inhärente Instabilität des Messsystems als mögliche Erklärung für die leichten Schwankungen. Nachdem die Schwankungen jedoch im Vergleich zu den Signalwerten der durchgeführten Messungen zu vernachlässigen sind, lässt sich die Stabilität des Systems über diese Phantommessungen bestätigen. Demnach wurde das Ziel, ein reproduzierbares und zuverlässiges Messsystem zu etablieren und qualifizieren, erfüllt.

5.2 Die Sensitivität des Messsystems

Wie bereits zuvor in Abbildung 19-c und Abbildung 21-c dargestellt, lag die Sensitivitätsgrenze des Systems bei etwa $10 \mu\text{l}$ bzw. $< 3 \cdot 10^6$ Zellen. Die Sensitivität eines MR-Scanners hängt indirekt von seiner Feldstärke ab. Die Feldstärke des Systems korreliert direkt mit dem Signal-zu-Rausch-Verhalten (SNR) [142, 143] und mit einem höheren SNR kann bei gleicher Messzeit sowohl eine höhere räumliche oder spektroskopische Auflösung als auch eine stabilere Rekonstruktion der Daten erreicht werden [144]. In der hier präsentierten Arbeit wurde mit einem 0.5 T Magneten gearbeitet. Viele weitere Studien, v. a. im Bereich der hochauflösenden Spektroskopie und Bildgebung, arbeiten im klinischen Bereich mit Feldstärken zwischen 1.5 und 3 T, moderne Geräte teilweise auch mit 7 T [145] und manche Laborsysteme erreichen Feldstärken von mehr als 20 T. Mit diesen Systemen ist es unter gewissen Umständen und unter Zuhilfenahme von Kontrastmitteln möglich, eine Auflösung in der Bildgebung bis auf eine Zellpopulation von 100 Zellen zu erreichen [146].

Nachdem die Fragestellung dieser Arbeit beinhaltet hat, die Messungen an einem mobilen System durchzuführen (siehe 2.1), limitiert dies die Arbeiten auf geringe Feldstärken. Damit geht eine vergleichsweise niedrige Auflösung einher, was die Anwendung auf Populationen von $\geq 3 \cdot 10^6$ Zellen limitiert. Für eine Verbesserung der Sensitivität bei gleichbleibender Feldstärke wäre es notwendig, den Abstand der Empfangsspule zur Probe zu reduzieren. Dies ist mit sog. Microcoils möglich. Diese Art von Empfangsspule kann sowohl in der Bildgebung als auch in der Spektroskopie verwendet werden, um die Ergebnisse zu verbessern [147, 148]. Außerdem wurde in einer Studie von 2021 berichtet, dass es möglich ist, Microcoils zu transplantieren [149] und im Gewebe Volumina von bis zu 450 nl zu vermessen [150]. Unter Zuhilfenahme solcher Microcoils wäre es ggf. möglich, die Sensitivität signifikant zu steigern. Durch die verbesserte Datengrundlage wäre es damit auch möglich mit-

tels Methoden aus dem Feld der KI – die untersuchten Zell- oder ggf. zukünftig auch Gewebetypen akkurater voneinander zu trennen und genauere Vorhersagen treffen zu können. In der präsentierten Arbeit wurden keine Microcoils entwickelt oder verwendet, da dies den Rahmen einer Arbeit aus der Sektion Biomedizin gesprengt hätte.

Auch wenn die erläuterten Möglichkeiten die Sensitivität des Systems steigern könnten, bestätigt die hohe Genauigkeit der trainierten KNNs die Validität der Messmethode. Demnach war die Sensitivität des Messsystems ausreichend, um die untersuchten Zellen mit hoher Genauigkeit voneinander zu trennen.

5.3 Automatisierung eines Benchtop-Prozesses

Um den Probendurchsatz pro Tag zu erhöhen, wurde ein Probenwechsler entwickelt (siehe 4.2). Aufgrund des flexiblen Roboterarms – welcher das Herzstück der Anlage bildet – war es möglich, diverse 3D-gedruckte Komponenten zu integrieren. Da der Roboterarm nur aus sechs Modellbau-Servomotoren besteht, waren Reparaturen am System ohne größeren Aufwand und tiefe Sachkenntnisse möglich. Dies ermöglichte, die Anlage auch ohne Vorwissen im Bereich der Robotik etablieren zu können. Nach entsprechenden Optimierungsschritten arbeitete die Roboteranlage sehr zuverlässig und es war möglich, bis zu 24 Proben an einem Arbeitstag zu vermessen während per Hand realistisch Weise maximal ca. 5 Proben möglich gewesen wären. Durch die Integration eines Lagersystems war es möglich, dass das System mit bis zu 12 Proben gleichzeitig bestückt wurde und anschließend diese autonom durch das System prozessiert wurden, was essentiell zur Beschleunigung des Messprozesses beitrug.

Die Verwendung von Servomotoren und die einfache Bauweise limitierte die Anlage jedoch in ihrer Präzision. Um diese Ungenauigkeiten zu kompensieren, wurden auf alle Öffnungen konische Aufsätze angebracht (s. Abbildung 23), um die Probenröhrchen beim Platzieren zu führen und zu zentrieren. Die Optimierung der Bewegungsabläufe und die Definition der korrekten Positionen gestaltete sich als sehr aufwendig, da die verwendeten Servomotoren kein Feedback über ihre aktuelle Position an den Arduino Mega gaben. Durch die Verwendung von komplexeren Motoren oder Winkelsensoren könnte eine Möglichkeit implementiert werden, bei der der Nutzer dem Roboterarm durch geführte Bewegungen direkt die gewünschten Positionen und Bewegungsabläufe einlernt. Dies würde nicht nur die Nutzerfreundlichkeit, sondern auch die Vielseitigkeit des Systems deutlich erhöhen. Da die Bewegungsabläufe

jedoch vergleichsweise simpel sind, würde dasselbe Ergebnis vermutlich mit einem Linearroboter schneller erzielt werden. Diese werden mit großem Erfolg in der Fertigung von z. B. Computer-Leiterplatten oder der Bestückung von Maschinen eingesetzt. Das Grundprinzip ist ähnlich wie bei einem FDM 3D-Drucker und ließe sich gut auf den konkreten Anwendungsfall anpassen. Im Vergleich zu anderen Automatisierungslösungen stellte die hier präsentierte Anlage jedoch eine sehr kostengünstige Lösung dar. Die gesamten Materialkosten beliefen sich auf etwa 350–400 €. Andere Roboterarme wie z. B. der Niryo Ned2 sind technisch ausgereifter, stellen jedoch mit einem Preis von ca. 3500 € [151] auch einen deutlich höheren Kostenfaktor dar. Auch wenn andere Systeme mögliche Vorteile versprechen, lieferte die entwickelte Automatisierungslösung zuverlässige Ergebnisse. Alle Proben wurden erfolgreich prozessiert und das Ziel, den Durchsatz an Proben pro Tag zu erhöhen, wurde erfolgreich umgesetzt.

5.4 Temperierung und Viabilität der Zellkultur

Da das Messsetup als Inprozesskontrolle gedacht ist, war es wichtig die Viabilität der Zellen so hoch wie möglich zu halten. Als zentraler Kulturparameter diente dabei die Temperatur. Daher wurde die Betriebstemperatur des MR-Scanners auf 37 °C angepasst, um der Kulturtemperatur der Zellen zu entsprechen. Außerdem war ein zusätzlicher Heizblock im Aufbau integriert, um bis zu vier Proben gleichzeitig vorzuheizen. Wie in 4.2.2 beschrieben, wurde die korrekte Temperierung durch mehrfache Kontrollmessungen verifiziert. Dabei wurde ersichtlich, dass die Zieltemperatur bereits nach ca. 5–8 min erreicht war. Um jedoch trotzdem eine entsprechende Temperatur garantieren zu können, wurde im Prozessablauf vorgesehen, dass die Zellen mind. 10 min im Heizblock und anschließend erneut 10 min im MR-Scanner geheizt wurden. So konnte zweifelsfrei eine korrekte Temperierung sichergestellt werden.

Die Probenbox, in welcher die Proben in den Messprozess eingegeben wurden, war nicht beheizt. Daher wurden die Zellen für die Evaluation der Viabilität zeitweise geheizt und anschließend bei RT gelagert, um reelle Arbeitsbedingungen abzubilden. Die anschließenden Viabilitätsmessungen zeigten (Abbildung 30), dass mit Ausnahme der HEK293T- und Vero-Zellen alle eine Viabilität von > 90 % aufwiesen. Die statistische Analyse zeigte lediglich bei den HEK293T- und Vero-Zellen eine signifikante Abweichung. Demnach war das Ziel, die Zellen lebendig zu vermessen, erfüllt. Die Abweichungen der HEK293T- und Vero-Zellen lassen sich evtl. mit Mess-

oder Präparationsfehlern erklären. Da sie mit $> 80\%$ immer noch im unbedenklichen Bereich lagen [152], widersprechen diese Ergebnisse in keiner Weise der Annahme, dass die gewählten Kulturbedingungen ausreichend für den gewählten Messzeitraum waren.

Um eine langfristige Kultur gewährleisten zu können, wäre es notwendig mehrere Faktoren der Kultur konstant zu halten. Es wäre wichtig, die Kultur unter geeigneten atmosphärischen Bedingungen zu kultivieren. Dies entspräche $\geq 95\%$ relativer Luftfeuchtigkeit, um ein Verdunsten des Kulturmediums zu verhindern. Nachdem der aktuelle Aufbau der Probenröhrchen einen entsprechenden Gasaustausch nicht zulässt, wäre hierfür eine Anpassung der Probenröhrchen notwendig. Diese Anpassung müsste dafür sorgen, dass das Kulturmedium im Röhrchen in direktem Austausch mit der ihm umgebenen Atmosphäre stünde. Des Weiteren ist es wichtig, den pH-Wert der Kultur bei etwa 7.4 zu stabilisieren [153]. Da die meisten Kulturmedien mit Hydrogencarbonat gepuffert sind [154], benötigt es eine Konzentration von 5% CO_2 , um den pH-Wert über die Kulturdauer hinweg konstant zu halten; Außerdem wäre, wie bereits in dieser Arbeit beschrieben, ein Medienaustausch innerhalb des Messröhrchens notwendig. Dies könnte mit PDMS-basierten Mikrofluidiksystemen erreicht werden [155–158]. Dieses System müsste jedoch so konstruiert sein, dass es das Wegspülen der Zellen verhindert. Dies wäre entweder durch eine physikalische Barriere wie einem Filter möglich oder über einen sehr geringen Medienfluss in der Probe, so dass die Kräfte, welche auf die Zellprobe wirken, die physikalische Integrität der Probe nicht nachhaltig stören würden. Auch wenn PDMS im MR-Spektrum nachgewiesen werden kann [159, 160], liegen die Relaxationszeiten mit wenigen Millisekunden deutlich außerhalb der Messwerte, welche von Medium und Zellen zu erwarten sind und würde folglich die Zellkulturbedingungen nicht beeinflussen.

5.5 Ursprung des Zellsignals und seiner Unterschiede

Eine immer wiederkehrende Fragestellung in diesem Projekt war diejenige nach der Herkunft des MR-Signals. Diese Frage konnte auch nach dieser Arbeit nicht eindeutig geklärt werden. Jedoch zeigten die Daten, dass mehrere Faktoren Einfluss auf das Signal nahmen. Es konnte gezeigt werden, dass die Zellzahl und damit zusammenhängend die Pelletgröße das Ergebnis nicht nur in seiner Intensität sondern, wenn es unter eine kritische Zellzahl fiel, auch in seiner Position beeinflusste. Nachdem

darauf geachtet wurde, immer dasselbe Probenvolumen unabhängig der Zellzahl zu verwenden, führte eine wechselnde Zellzahl oder auch eine andere Zellgröße bei gleicher Zellzahl zu unterschiedlich großen Pellets (Abbildung 21). Bei großen Zellen z. B. CHO, war folglich ein relativ geringer Anteil an Medium in der Probe enthalten, während kleinere Zellen, wie z. B. Vero-Zellen, einen größeren Anteil an Medium und einen kleineres Zellpellet aufwiesen. Jedoch korrelierten diese Beobachtungen nicht mit den zuvor dokumentierten Werten für die Viabilitäten. Vero-Zellen wiesen eine geringere Viabilität auf, obwohl der relative Anteil des Zellkulturmediums höher war als bei CHO-Zellen, welche jedoch eine höhere Zellzahl aufwiesen. Wenn jedoch die Menge an überschüssigen Medium bei den Vero-Zellen auf 1 ml erhöht wurde (Abbildung 30), wiesen auch diese eine Viabilität von nahezu 100 % auf. Daraus wurde geschlossen, dass die geringen Unterschiede an Mediumsmenge in den Zellproben keinen Einfluss auf die Viabilität hatten. Eine signifikante Steigerung der Viabilität würde nur durch eine deutlich höhere Versorgung mit Zellkulturmedium gewährleistet werden. Folglich war die Unterversorgung mit Medium ausschlaggebend für die Veränderung der Viabilität bei den Zellen.

Der relative Anteil des Mediums und infolgedessen überlagerte ab einem Schwellenwert das Mediensignal das der Zellen, wodurch die dokumentierte Veränderung des Zellpeaks bei $< 3 \cdot 10^6$ Zellen begründet werden kann. Aufgrund dieser Abhängigkeiten wurden die Messungen immer mit derselben Zellzahl durchgeführt, um diesen Faktor in der Signalanalyse konstant halten zu können.

Neben der Pelletgröße beeinflusst das Volumen der Zellen und deren Aufbau auch die Diffusion innerhalb der Probe. Die Zellen stellen Diffusionsbarrieren für freie Moleküle wie Wasser dar. Mit geeigneten Sequenzen kann der Diffusionskoeffizient der Probe bestimmt werden [161]. Die Messung basiert auf der Anregung der Probe, einer Rephasierung der Spins und einer danach geschalteten Warte- bzw. Diffusionszeit. Anschließend wird die Probe rephasiert und das Signal verglichen. Der Signalunterschied kommt aufgrund der Diffusion der ursprünglich angeregten Protonen zustande. Eben dieser Effekt könnte auch die hier präsentierten Ergebnisse beeinflusst haben. Eben dieser Effekt wird genutzt, um die zelluläre Struktur von Geweben nicht-invasiv zu untersuchen [162–164]. Auch wenn diskutiert wird, ob unterschiedliche Zelltypen anhand ihres Diffusionskoeffizienten unterschieden werden können oder nicht, konnte doch gezeigt werden, dass es einen direkten Zusammenhang zwischen Zellzahl und Diffusionssignal gab [165]. Aufgrund der langen Messzeit, der Temperatur von 37 °C und der damit einhergehenden höheren Molekularbewegung und Diffusion könnten unterschiedlich große Zellen zu unterschiedlich stark

ausgeprägten Diffusionseffekten geführt haben. Unterschiedliche Zellgrößen und -zusammensetzungen könnten auch zu unterschiedlichen intra- und extrazellulären Diffusionsprozessen geführt haben. Mögliche Diffusionseffekte können sich wiederum auf das gemessene 2D-Korrelationsspektrum der Zellprobe ausgewirkt haben. Nachdem die Diffusion auch stark von der Temperatur abhängt [166–169], bekräftigt die vorangegangene Betrachtung der Diffusion auch die Notwendigkeit der Temperierung der zu untersuchenden Probe (Unterabschnitt 4.2.2). Sollte es während der Messung zu Temperaturschwankungen kommen, könnten diese zu Veränderungen in der Diffusion innerhalb der Probe führen und folglich auch zu einem veränderten MR-Signal.

Ein weiterer Faktor, welcher das Signal beeinflussen könnte, ist die chemische Zusammensetzung der Zellen. Ein wesentlicher Teil der Zelle besteht aus Lipiden. Beispielsweise besteht die Zellmembran, welche den Zellkörper umschließt, sowie die Plasmamembranen, welche sämtliche Zellorganellen abgrenzen, aus einer Phospholipiddoppelschicht [170]. Nachdem Lipide deutlich kürzere Relaxationszeiten im Vergleich zu Wasser besitzen, ist davon auszugehen, dass diese einen deutlichen Anteil am zellulären MR-Signal ausmachen. Nachdem die Zusammensetzung der Membranen auf den jeweiligen Zelltyp und dessen Funktion angepasst ist [171–173], besteht die Möglichkeit, dass auch diese Unterschiede das Signal beeinflussen.

Abschließend lässt sich nicht exakt bestimmen, durch welche Strukturen das Zellsignal erzeugt wird. Vermutlich ist es eine Kombination aus den chemischen und physikalischen Eigenschaften des Zellpellets und dessen Wechselwirkung mit dem Zellkulturmedium.

Die Zielsetzung und der zeitliche Rahmen dieser Arbeit waren auf die Verwendung der T_1/T_2 -Relaxometrie beschränkt. Jedoch bieten auch andere MR-Methoden die Möglichkeit, Proben nicht-invasiv untersuchen zu können. Beispielsweise kann die Diffusion einer Probe untersucht werden. Durch diese Untersuchung lassen sich Rückschlüsse auf die zelluläre Mikroumgebung ziehen [174]. Anwendung findet dies u. a. in der Onkologie, bei der Untersuchung von Brustkrebs oder in der Neurologie bei der Untersuchung von neuronalem Gewebe [175]. Da sich bei der Tumorbildung auch die Struktur des umliegenden Gewebes ändert und somit dessen Diffusion, kann die Messung von eben dieser dazu dienen, benigne von malignen Tumoren zu unterscheiden [176].

Neben der Diffusion können auch kleinste chemische Strukturen mittels der *chemical exchange saturation transfer* (CEST) Methode nachgewiesen werden. Je nach Fragestellung ermöglicht es das entsprechende MR-Experiment beispielsweise biolo-

gische Metabolite wie Kreatin oder Laktat nicht-invasiv in der untersuchten Probe nachzuweisen und u. U. sogar zu quantifizieren [177–179]. Dies dient in der Onkologie der nicht-invasiven Identifizierung von tumorösem Gewebe [180, 181]. Mittels dieser Methode ist es auch möglich, den pH-Wert einer Probe zu ermitteln, was ebenfalls wieder Rückschlüsse auf die biologische Zusammensetzung und Funktion des Gewebes zulässt [182–186].

Während die meisten MR-Experimente auf der Anregung von Wasserstoffatomen beruhen, besteht auch die Möglichkeit mittels entsprechender Experimente den Natrium- und Kalium-Gehalt einer Probe zu messen. Beide Salze sind wichtige Bestandteile von diversen Geweben wie z. B. Muskeln und neuronalem Gewebe. Somit stellen beide wichtige Bioparameter für die klinische Diagnostik dar [187, 188].

Die erläuterten Parameter wie Diffusion, CEST, Natrium und Kalium-Gehalt der Probe würden zweifelsfrei wertvolle Zusatzinformationen über die untersuchten Zellen liefern. Mit den Werten könnten v. a. Differenzierungsprozess genauer abgebildet werden, da diese i. d. R. von morphologischen Veränderungen der Zellen begleitet werden. Diese morphologischen Veränderungen haben mit hoher Wahrscheinlichkeit Einfluss auf die Diffusion der Zellmembran. Auch die veränderte chemische Zusammensetzung bzw. eine gesteigerte oder reduzierte Konzentration entsprechender Metabolite können mit der geeigneten Methode nachgewiesen werden. Eine methodische Berücksichtigung dieser Methoden in der vorgelegten Arbeit hätte jedoch den zeitlichen Rahmen überschritten. Daher war von Anfang an für diese Arbeit die T_1/T_2 -Relaxometrie als aussagekräftiger Read-Out Parameter vorgesehen. Auch war das Ziel dieser Arbeit eine Methode zu etablieren, welche flexibel und mobil einsetzbar ist. Die oben beschriebenen Methodiken werden i. d. R. bei sehr hohen Feldstärken durchgeführt, welche bisher nicht mit mobilen MR-Scannern erreichbar sind.

5.6 KI für die Dateninterpretation

Seit der Etablierung von benutzerfreundlichen Softwareumgebungen für den Umgang mit Methoden der künstlichen Intelligenz, wie linearer Regression, SVM und KNN, stieg auch deren Bedeutung in der wissenschaftlichen Datenanalyse. Hierfür hat sich v. a. die *Python*-Programmiersprache durchgesetzt, da sie einfachen Zugang zu Softwarepaketen wie *scikitlearn* [119], *Keras* [189], *TensorFlow* [190] und *PyTorch* [121] ermöglicht. Mithilfe dieser Pakete und der hohen Rechenleistung von modernen

Grafikkarten ist es heute so einfach wie nie zuvor, künstliche Intelligenz zu entwickeln und einzusetzen. KNNs werden mittlerweile in vielen klinischen Bereichen wie der radiologischen Diagnostik [191, 192] oder auch in der Entdeckung und Optimierung von neuen Medikamenten [193, 194] eingesetzt. Ebenso wurde auch in der vorliegenden Studie ein KNN als zentrale Methode für die Datenklassifizierung verwendet. In allen Fällen basiert das ‚Wissen‘ eines KNN ausschließlich auf den gelernten Daten. Gute Netzwerke zeichnen sich dadurch aus, dass sie in der Lage sind, das Problem zu generalisieren. Das bedeutet, dass ein Klassifizierungsalgorithmus auch dann Objekte korrekt in die gelernten Klassen einordnet, wenn er den konkreten Gegenstand zuvor noch nie gesehen hat. Daher ist die Qualität des Netzwerks sehr stark von den Trainings- und Testdaten abhängig. Große Sorgfalt sollte darauf verwandt werden, hochqualitative Daten für das Training zu erzeugen. Im konkreten Fall dieser Arbeit bedeutet dies, die Qualität der MR-Messungen so hoch wie möglich zu halten z. B. durch lange Messzeiten oder den Einsatz von optimierter Hardware. Auch spielt natürlich die Zusammensetzung des verwendeten Datensatzes eine wichtige Rolle. Nur wenn die einzelnen Datenpunkte auch korrekt betitelt sind, kann das Netzwerk die korrekten Zusammenhänge erlernen. Dies ist v. a. bei klinischen Daten wichtig. Sollten z. B. Bilder von histologischen Färbungen klassifiziert werden, unterliegen die zuordneten Labels den subjektiven Einschätzungen des Betrachters. Nur wenn diese Labels in allen Fällen richtig zugeordnet wurden, ist es für das KNN auch möglich korrekte Vorhersagen zu treffen. Ebenso ist es wichtig, dass der Datensatz so ausgeglichen wie möglich ist. Das bedeutet, dass jede Klasse in etwa gleich häufig im Datensatz vertreten ist. Sollte dies nicht der Fall sein, könnte es alleine durch die Überrepräsentation einzelner Klassen dazu kommen, dass die Genauigkeit des gesamten Netzwerks niedrig ist [195–197].

Mithilfe der zuvor beschriebenen Automatisierungslösung war es in der vorliegenden Arbeit möglich, reproduzierbare Ergebnisse in großem Umfang zu erzeugen. Dies ermöglichte es, Zelllinien bzw. differenzierte MSCs mit einer Genauigkeit von bis zu 98 % voneinander zu trennen. Die Leistung des KNN bzw. der SVM hing dabei jedoch stark von der Zusammensetzung der Daten ab. Sobald die Anzahl an zu klassifizierenden Zellen erhöht wurde, sank die erreichbare Genauigkeit. Trotz aller vorgenommenen Optimierungen war es im Rahmen dieser Arbeit nicht möglich, die Genauigkeit bei 10 Zelllinien signifikant auf $> 50\%$ zu steigern. Da bei der Klassifikation der MSCs lediglich die Unterscheidung zwischen differenzierten und nicht-differenzierten Zellen getroffen wurde, war hier eine deutlich höhere Genauigkeit möglich (93.89%). Sollten zukünftige Studien eine detaillierte Unterscheidung

von mehreren Differenzierungszuständen anstreben, wäre es auch hier denkbar, dass die Genauigkeit der verwendeten Algorithmen signifikant sinken könnte.

In dieser Arbeit war es nicht möglich, die Genauigkeit des KNN bei der Klassifizierung von 10 Zelllinien auf $> 50\%$ zu steigern. Da die verwendeten Algorithmen eher einem Proof-of-Principle dienen sollten, ist es nicht auszuschließen, dass zukünftige Arbeiten zu diesem Thema auch die Genauigkeit noch steigern. Hier könnte u. a. auch die zuvor beschriebene Sensitivität des Messsystems ausschlaggebend dafür sein, bessere Daten zu erzeugen, welche dann wiederum durch ein KNN effizienter klassifiziert werden können.

Auch wäre es denkbar die Methode des *Transfer Learnings* zu nutzen, um bereits vortrainierte KNNs für die Klassifizierung zu nutzen. Transfer Learning findet in vielen Bereichen Anwendung [198–201]. Hierbei wird ein bereits trainiertes Netzwerk d. h. inkl. aller optimierten Gewichte nur geringfügig adaptiert, um darauf neue, bisher unbekannte Daten zu klassifizieren. Dieses Konzept wird i. d. R. bei der Bilderkennung verwendet, um aufwändige, teilweise tagelange Trainingsdurchläufe auf sehr großen Datensätzen zu vermeiden. Da jedoch bereits die Klassifikation von MSCs im Vergleich zu Zelllinien eine Anpassung der Netzwerkstruktur voraussetze, wäre auch für die entsprechende Adaption eines vortrainierten Netzwerks ein nicht zu vernachlässigender Arbeitsaufwand nötig. Weil im Voraus nicht abzusehen ist, ob das Training von Erfolg gekrönt sein wird, wurde in dieser Studie darauf verzichtet. Stattdessen wurde ein simples Netzwerk verwendet, um daran zu demonstrieren, dass eine nicht-invasive Klassifizierung von Zellen anhand ihrer 2D MR-Spektren generell möglich ist.

Nachdem gezeigt werden konnte, dass eine SVM vergleichbare Genauigkeiten lieferte, wurde daraus geschlossen, dass v. a. die Position der Zell- und Medienpeaks von zentraler Bedeutung ist. Dabei scheint die Orientierung und Form des Signals eine eher untergeordnete Wichtigkeit zu spielen. Auch wenn die SVM zwar erfolgreich die Zelllinien klassifizierte, stieß sie bei den MSCs an ihre Grenzen. Da die SVMs durch ihre Architektur auf einen Datenpunkt limitiert sind, fiel es den KNNs leichter, Daten zu klassifizieren, deren Identität auf mehr als nur einem einzelnen Signalpunkt zurückzuführen ist. Dies belegt die Wichtigkeit, mehrere Algorithmen für die zukünftige Klassifizierung von Zellmaterial in Betracht zu ziehen.

5.7 Benutzerdefinierte Datenaugmentierung

Eine Methode, um die Qualität eines KNN zu steigern, ist die Augmentierung. Häufig liegt eine schwache KNN-Leistung an einem zu kleinen oder zu spezialisierten Trainingsdatensatz [202–204]. Bei der Objekterkennung in Bildern werden die Trainingsdaten künstlich vergrößert, indem ein Bild u. a. gedreht, vergrößert, gestaucht oder mit Rauschen überlagert wird. Hierdurch wird ein ‚virtueller‘ Datensatz an zusätzlichen Bildern generiert. Anhand des größeren Datensatzes lernt das neuronale Netzwerk besser zu generalisieren, ist robuster gegenüber Unregelmäßigkeiten in den Daten und braucht ggf. weniger Trainingsepochen, um gute Genauigkeiten zu erreichen [202]. Im weit verbreiteten Keras Paket für Python [189] sind solche Augmentierungsmöglichkeiten bereits integriert. Die dort verwendeten Methoden sind jedoch sehr stark auf die Bilderkennung optimiert. Aufgrund dieser Spezialisierung werden in vielen weiteren Teilbereichen angepasste Augmentierungsalgorithmen entwickelt und verwendet [205–208]. Auch in der vorliegenden Arbeit konnten die Keras-Funktionen nicht auf die verwendeten Daten angewandt werden. Es erfolgte zwar eine Klassifizierung von Matrices, welche als Bilder betrachtet werden können, jedoch ist die Position der Signale von zentraler Bedeutung. Wenn ein entsprechender Algorithmus das Bild drehen oder dehnen würde, würde dies dazu führen, dass Daten entstünden, welche auf physikalischem Weg nicht reproduzierbar wären.

Wie in 4.5.3.2 und 4.5.4.1 dargestellt, reichte die Menge an unaugmentierten Daten nicht aus, um ein zuverlässiges Netzwerk zu trainieren. Aus diesem Grund wurde ein benutzerdefinierter Augmentierungsalgorithmus entwickelt und dessen Parameter im Zuge dieser Arbeit für den konkreten Anwendungsfall optimiert. Die Optimierung basierte hierbei auf empirisch erhobenen Daten. Der Erfolg des KNN-Trainings bestätigte die Effektivität der Augmentierung. Es konnte gezeigt werden, dass die Augmentierung signifikant zum erfolgreichen KNN-Training beitrug (z. B. Steigerung von 37 % auf 51 % bei den getesteten Zelllinien und 40 % auf 86 % bei den getesteten MSCs). Dabei spielte es oft keine große Rolle, wie häufig die Daten augmentiert worden waren. Bereits ein Augmentierungsfaktor von 5 war i. d. R. ausreichend, um ein gutes Ergebnis zu erzielen. Bei der Ausprägung der Augmentierung war zu erkennen, dass ein zu ausgeprägtes Verändern der Daten eher zu einer Verschlechterung der Ergebnisse führte.

Trotzdem lässt sich abschließend die Notwendigkeit und Bedeutung der Augmentierung für das erfolgreiche trainieren der KNNs bekräftigen. Ohne diese Schritte wären die unterschiedlichen Zelltypen nicht mit der entsprechenden Genauigkeit voneinander zu trennen gewesen.

6 Ausblick

In dieser Arbeit konnte gezeigt werden, dass mit der 2D-MR-Relaxometrie, kombiniert mit einem künstlichen neuronalen Netzwerk (KNN), unterschiedliche Zelllinien mit einer Genauigkeit von bis zu 98 % voneinander getrennt werden können. Außerdem konnte gezeigt werden, dass die beschriebene Methodik differenzierte von undifferenzierten MSCs ebenfalls mit einer hohen Genauigkeit (93.89 %) unterscheiden kann.

Um die Anwendung weiter voranzutreiben, ist es zwingend notwendig, eine noch umfangreichere Datenbank aufzubauen. Mithilfe einer solchen umfangreichen Datenbank wäre es möglich, ein KNN zu trainieren, welches deutlich mehr Zelltypen und z.B. auch primäre Zellen mit einer hohen Genauigkeit voneinander unterscheiden könnte. V. a. Letzteres wäre ein wichtiger Aspekt für die Anwendung im präklinischen und klinischen Kontext.

Nachdem in dieser Arbeit die Differenzierung der MSCs nur binär als differenziert bzw. undifferenziert klassifiziert wurde, sollten zukünftige Arbeiten Zeitreihen beinhalten. Basierend auf diesen Messungen könnte dann ein entsprechendes KNN dafür genutzt werden, anhand der MR-Korrelationsspektren den Differenzierungsstand der Zellen zu bestimmen. Auch sollten weitere zelluläre Entwicklungen bzw. Reaktionen vermessen werden, wie z.B. die *chondrogene* und *osteogene* Differenzierung, welche ebenfalls anhand der MR-Relaxometrie-Daten klassifiziert werden könnten.

Die hier verwendete Sequenz wurde entwickelt, um eine korrelierte Quantifizierung der Relaxationszeiten zu ermöglichen. Dafür war es notwendig, am Ende jeder TI-Iteration eine entsprechende Pause zu machen, um sicherzustellen, dass das angeregte Signal komplett relaxiert war. Um die Messzeit zu verkürzen, sollten zukünftige Arbeiten in Betracht ziehen, diese Pause auf ein Minimum zu reduzieren. Hierdurch entsteht der Nachteil, dass die Ergebnisse nicht mehr quantifizierbar sind, jedoch besteht die Möglichkeit, dass die entstehenden Daten trotzdem mithilfe eines KNN klassifiziert werden können. Durch die so entstehende kürzere Messzeit ergäbe sich die Möglichkeit, deutlich schnellere zelluläre Reaktionen und Veränderungen untersuchen zu können. Dies würde das Anwendungsfeld der Methodik zweifelsfrei erweitern.

In der hier vorgelegten Arbeit betrug der maximale Zeitraum, in welchem die Zellen

im Messsystem kultiviert wurden, 5 bis 7 Stunden. Dies limitierte die Anwendung für manche Fragestellungen (Kapitel 5). Daher sollte in zukünftigen Arbeiten ein Bioreaktor-Setup entwickelt werden, welches eine Kultur von Zellproben im Messsystem für mehrere Stunden oder ggf. sogar Tage ermöglicht. Dieser Bioreaktor sollte sowohl eine Temperierung der Proben, als auch eine konstante Umgebungsatmosphäre für das Kulturmedium aus 5 % CO₂ und ≥ 95 % relativer Luftfeuchtigkeit gewährleisten können. Idealerweise sollte er in eine Automatisierung integrierbar sein, sodass eine Langzeitmessung von unterschiedlichen Proben erreicht werden kann. Mit solch einem Setup könnten die bereits erwähnten Zeitreihen direkt im Messsystem kultiviert und vermessen werden, ohne dass die Proben zwischenzeitlich zurück in die Zellkultur überführt werden müssten.

Eine weitere Erleichterung der Messung wäre der Methodentransfer auf ein Messsystem, welches keinen Volumenscanner, sondern einen Oberflächenscanner verwendet. Mithilfe dessen wäre es möglich, die Zellen direkt in der ursprünglichen Zellkulturflasche zu vermessen. Damit könnte man Differenzierungen und zelluläre Entwicklungsprozesse direkt in der Flasche überwachen. Ein solches System könnte in eine größere Automatisierungsanlage eingebunden und ggf. als zusätzlicher Readout-Parameter für die Beurteilung der Zellkultur herangezogen werden. Wie in der Diskussion bereits beschrieben, würde auch die Erhebung von zusätzlichen MR-Parametern wie die der Diffusion, des pH-Wertes, biologischen Metaboliten oder dem Natrium bzw. Kalium-Gehalt die Genauigkeit der Vorhersage erhöhen.

7 Danksagung

Hier beginnt der sehr persönliche Teil dieser Arbeit. Jeder, der mich kennt, weiß, wie wichtig mir Freundschaften und kollegiales Miteinander sind. Darum möchte ich mir die Zeit nehmen und mich gebührend bei allen Menschen bedanken, die im weitesten Sinne Teil dieser Arbeit waren.

Ich möchte mich am Anfang vor allem bei **Daniel Weber** bedanken. Ohne dich wäre nichts von all dem hier möglich gewesen. Du hattest immer ein offenes Ohr für Fragen und hast mich so sehr wie kaum eine andere Person in meinem wissenschaftlichem Leben inspiriert. Mein Dank an **Daniel Haddad** steht diesem um nichts nach. Du hast mich ins Projekt geholt. Ich hätte mir keine bessere Person als Vorgesetzten und „Paten“ wünschen können. Danke für die zahlreichen motivierenden und inspirierenden Gespräche. Ohne euch beide, Daniel und Daniel, wäre ich vermutlich an dieser Herausforderung gescheitert.

Mein Dank gilt selbstverständlich auch den Betreuern meiner Doktorarbeit. Allen voran **Jan Hansmann**. Danke, dass du mich 2018 für dieses Projekt begeistert hast. Du hast immer an mich und das Projekt geglaubt und als 1. Gutachter der Doktorarbeit essentiell zum Erfolg beigetragen. Großen Dank an **Prof. Peter Jakob**. Durch Ihre kritische Betrachtung der Ergebnisse und das ständige Hinterfragen der Methoden haben Sie immer wieder die richtigen Impulse zum richtigen Zeitpunkt gegeben, die dazu geführt haben, dass das Projekt erfolgreich voranging. Vielen Dank auch an **Karl-Heinz Hiller**. Danke, dass du mich 2018 so herzlich bei Fraunhofer willkommen geheißen und meine berufliche und persönliche Entwicklung immer gefördert hast. Ohne euch drei, Jan, Prof. Jakob und Karl-Heinz, hätte ich mich sicherlich in den Irrungen und Wirrungen und den unendlichen Möglichkeiten einer Doktorarbeit verloren. Danke für eure Führung und Inspiration.

Ich hatte das große Glück in drei Teams arbeiten zu dürfen in denen ich mich sehr wohl gefühlt habe. Daher geht mein Dank an das Team des **Fraunhofer EZRT** sowie des **Lehrstuhls für Experimentalphysik V**. Ihr habt mich in der Welt der MR-Physik willkommen geheißen und unterstützt. Ich möchte ich mich vor allem bei **Heidrun Pfeuffer, Sabine Voll, Martin Blaimer, Markus Hildenbrand, Jannik Stebani, Philipp Mörchel, Felix Breuer, Florian Fidler und Michael Eder** bedanken. Vielen Dank für eine unvergessliche und sehr lehrreiche Zeit. Ohne den **Lehrstuhl für Tissue Engineering und Regenerative Medizin** wäre ich heute nicht die Person, die ich bin. In all den Jahren meiner Master- und Doktorarbeit fand ich hier nicht nur Kollegen, sondern gute Freunde. Auch wenn es viele sind, möchte ich mich namentlich bei ihnen bedanken. Allen voran gilt mein Dank meinen Bürokollegen und Freunden: **Marius, Hanna, Sebi und Christina**. Ohne euch hätte ich das nicht geschafft. Außerdem geht mein Dank an **Özlem, Anna-Kristina, Alevtina, Sabine, Heide, Heidi, Philipp, Tobi, Verena, Florian, Maxi, Chris, Christian, Christoph, Lisa, Marc, Thomas, Lukas, Christa, Joachim, Maria, Marco, Olli, Antje, Daniela, Tom, Consti, Matthias, Sanjana, Ji, Tamás, Sofia, Gudrun, Franzi, Katharina, Anna, Sarah und Sabrina**. Ohne euch wäre meine Zeit am TERM nicht dieselbe gewesen.

Durch den Stress einer Doktorarbeit vergisst man gerne auch mal, dass noch andere Dinge in der Welt gibt. Man verliert sich und überarbeitet sich. Darum ist es so wichtig Freunde und Familie zu haben, die einen wieder auf den Boden der Tatsachen zurückholen. Danke **Kadda und Hobbit**. Ohne euch wäre die Zeit in Würzburg sehr langweilig gewesen. Danke für die vielen schönen Stunden und inspirierenden Diskussion beim Abendessen.

Den größten Anteil an dem Erfolg dieser Arbeit hast du, **Caro**. Ohne den Ausgleich und deine Liebe, weiß ich nicht, ob ich durchgehalten hätte.

Zu guter Letzt möchte ich mich bei meiner **Familie** bedanken: dafür, dass ihr mich immer unterstützt habt, mich immer gefördert habt, meine Liebe zur Wissenschaft und Forschung immer gesehen und mich stets motiviert habt, meinen Weg zu gehen.

Und Danke an **alle, die ich hier nicht genannt habe**, die mich aber trotzdem in den letzten **11 Jahren** an der Uni Würzburg begleitet und diese Zeit so unvergesslich gemacht haben. Danke euch allen.

8 Publikationsliste

P. Fey, D.L. Weber, J. Stebani, P. Mörchel, P. Jakob, J. Hansmann, K.-H. Hiller, D. Haddad. „Non-destructive classification of unlabeled cells: Combining an automated benchtop magnetic resonance scanner and artificial intelligence“- *PLoS computational biology*, 2023

F. Karl, M. Wußmann, L. Kreß, T. Malzacher, **P. Fey**, F. Groeber-Becker, N Üçeyler. „Patient-derived in vitro skin models for investigation of small fiber pathology“- *Annals of Clinical and Translational Neurology*, 2019

10 Appendix

10.1 Affidavit

I hereby confirm that my thesis entiteled 'AI-assisted MR classification of cells and cellular differentiation' is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, 15.06.2023

Place, Date

Signature

10.2 Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „KI-gestützte MR-Klassifizierung von Zellen und zellulärer Differenzierung“ eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 15.06.2023

Ort, Datum

Unterschrift

Abbildungsverzeichnis

Abbildung 1	Schematische Darstellung eines MRT.	25
Abbildung 2	Schematische Darstellung der physikalischen Grundlagen der MRT.	27
Abbildung 3	Schematische Darstellung der T_1 - und T_2 -Relaxation.	28
Abbildung 4	Schematische Darstellung der verwendeten T_1/T_2 - Sequenz.	29
Abbildung 5	Schematischer Aufbau eines KNN.	30
Abbildung 6	Schematische Visualisierung der Gewichte eines KNNs.	31
Abbildung 7	Schematische Visualisierung der Kostenminimierung.	32
Abbildung 8	Schematische Visualisierung des CNN Aufbaus.	34
Abbildung 9	Schematische Darstellung der Blastocyste.	35
Abbildung 10	Schematische Darstellung der MSC-Differenzierung.	36
Abbildung 11	Aufbau und Verkabelung der Probenbox.	50
Abbildung 12	Exemplarische Darstellung der verwendeten MATLAB® GUI.	51
Abbildung 13	Prozessdiagramm für den Programmablauf der verwendeten MATLAB® GUI.	52
Abbildung 14	Etablierte KNN-Architektur zur Klassifikation von Zelllinien (Teil 1).	61
Abbildung 15	Etablierte KNN-Architektur zur Klassifikation von Zelllinien (Teil 2).	62
Abbildung 16	Etablierte KNN-Architektur zur Klassifikation von MSCs (Teil 1).	63
Abbildung 17	Etablierte KNN-Architektur zur Klassifikation von MSCs (Teil 2).	64
Abbildung 18	Etablierung von Agarose-Dotagraf®-Zellphantomen.	67
Abbildung 19	Stabilitätsmessungen am verwendeten MR-Messsystem anhand der entwickelten ADZPs.	68

Abbildung 20	Statistische Analyse zu den durchgeführten ADZP Stabilitätsmessungen	70
Abbildung 21	Erste Zellmessungen zeigten zwei Gruppen an Signalpeaks und ermöglichten die Bestimmung der minimal messbaren Zellzahl.	71
Abbildung 22	Statistische Analyse der Signalstabilität relativ zur gemessenen Zellzahl.	72
Abbildung 23	Aufbau der verwendeten Robotikplattform.	75
Abbildung 24	Validierung der Proben temperierung im Heizblock und dem Magneten des MR-Scanners.	76
Abbildung 25	Vergleich der Zellpeak-Schwerpunkte der drei Zelllinien CHO, MDA231 und K562.	78
Abbildung 26	Mikroskopische Aufnahmen der verwendeten Zelllinien.	80
Abbildung 27	Anzahl der im finalen Datensatz enthaltenen, unabhängigen Messungen; aufgelistet für jede Zelllinie. . .	81
Abbildung 28	Separate 2D-Korrelationsspektren für jeden gemessenen Zelltyp, kombiniert für alle erhobenen Replikate. . .	82
Abbildung 29	Scatterplot der Zellpeaks für den vollen Messumfang von zehn unterschiedlichen Zelllinien.	83
Abbildung 30	Untersuchung der Viabilität der verwendeten Zelllinien.	84
Abbildung 31	Visualisierung der Zellalterung anhand einer Messreihe über bis zu 192 h.	86
Abbildung 32	Etablierung und Messung von Negativkontrollen. . .	87
Abbildung 33	Adipogene Differenzierung und MR-Messung von mesenchymalen Stromazellen.	90
Abbildung 34	Klassifizierung von Zellpeaks mithilfe einer SVM. - Teil 1	91
Abbildung 35	Klassifizierung von Zellpeaks mithilfe einer SVM. - Teil 2	92
Abbildung 36	Performanz der trainierten SVM.	94
Abbildung 37	Vergleich unterschiedlicher Augmentierungsparameter.	96

Abbildung 38	Optimierung der Augmentierungsroutine und Evaluation der Leistung des KNN zur Klassifizierung von MSCs.	102
Abbildung 39	Testen von unterschiedlichen Augmentierungsfaktoren und -parametern auf die Performanz der etablierten KNN-Architektur zur Klassifizierung von Zelllinien.	104
Abbildung 40	Performanz des KNN zur Klassifikation von Zelllinien (Teil 1).	106
Abbildung 41	Performanz des KNN zur Klassifikation von Zelllinien (Teil 2).	107
Abbildung 42	Performanz des KNN zur Klassifikation von Zelllinien (Teil 3).	108
Abbildung 43	Vergleich der resultierenden Genauigkeiten der trainierten KNNs bei unterschiedlicher Anzahl von Zelllinien.	109
Abbildung 44	Einfluss der drop-out-Rate auf die Performanz des KNN.	110
Abbildung 45	Einfluss der Anzahl an Trainingsepochen auf die Performanz des KNN.	111
Abbildung 46	Einfluss des Medienpeaks auf die Leistung des etablierten KNN.	112
Abbildung 47	Abschließender Vergleich der Genauigkeit der Klassifikation mittels SVM und KNN.	114

Tabellenverzeichnis

Tabelle 1	Übersicht über verwendete Chemikalien	38
Tabelle 2	Übersicht über verwendete Laborgeräte	38
Tabelle 3	Übersicht über verwendete Zellkulturmedien	39
Tabelle 4	Übersicht über verwendete Medienzusätze	39
Tabelle 5	Übersicht über weitere Chemikalien und Puffer für die Zellkultur	39
Tabelle 6	Übersicht über verwendete Puffer und Lösungen für Histologie	39
Tabelle 7	Übersicht über verwendete Kontrastmittel	40
Tabelle 8	Übersicht über verwendete Software	40
Tabelle 9	Übersicht über verwendete Verbrauchsmaterialien	40
Tabelle 10	Übersicht über verwendete Primärzellen	41
Tabelle 11	Übersicht über verwendete Zelllinien	41
Tabelle 12	Erhobene Werte für die Pulsdauer und Amplitude der MR-Messung.	66
Tabelle 13	Verwendete Zelllinien.	93
Tabelle 14	Werte der statistischen Analyse der SVM-Performanz.	95
Tabelle 15	Vergleich der KNN-Leistung bei zugeschnittenen vs. nicht zugeschnittenen Daten von MSCs.	99
Tabelle 16	Durchschnittliche Genauigkeit für den Vergleich von unterschiedlichen Augmentierungsparameter Kombinationen.	101
Tabelle 17	Durchschnittliche Werte und Ergebnisse des t-Tests je Augmentierungsfaktor für den Vergleich von 12 und 25 Trainingsepochen.	111

Literaturverzeichnis

- [1] Paul-Ehrlich-Institut. *ATMP*. 2022. URL: <https://www.pei.de/DE/arzneimittel/atmp/atmp-node.html>.
- [2] Marina Radrizzani, Sabrina Soncin, Viviana Lo Cicero et al. „Quality Control Assays for Clinical-Grade Human Mesenchymal Stromal Cells: Methods for ATMP Release“. In: *Methods in molecular biology (Clifton, N.J.)* 1416 (2016), Seiten 313–337. DOI: 10.1007/978-1-4939-3584-0{\textunderscore}19.
- [3] Daniela Celeste Profico, Maurizio Gelati, Daniela Ferrari et al. „Human Neural Stem Cell-Based Drug Product: Clinical and Nonclinical Characterization“. In: *International journal of molecular sciences* 23.21 (2022). DOI: 10.3390/ijms232113425.
- [4] Patrick Bittorf, Thorsten Bergmann, Simone Merlin et al. „Regulatory-Compliant Validation of a Highly Sensitive qPCR for Biodistribution Assessment of Hemophilia A Patient Cells“. In: *Molecular therapy. Methods & clinical development* 18 (2020), Seiten 176–188. ISSN: 2329-0501. DOI: 10.1016/j.omtm.2020.05.029.
- [5] Veronica Ortega, Christina Mendiola und Gopalrao V. N. Velagaleti. „Identification of Chromothripsis in Biopsy Using SNP-Based Microarray“. In: *Methods in molecular biology (Clifton, N.J.)* 1769 (2018), Seiten 85–117. DOI: 10.1007/978-1-4939-7780-2{\textunderscore}7.
- [6] Jacob D. Grodsky, Randall D. Craver und Isa F. Ashoor. „Early identification of transplant glomerulopathy in pediatric kidney transplant biopsies: A single-center experience with electron microscopy analysis“. In: *Pediatric transplantation* 23.5 (2019), e13459. DOI: 10.1111/ptr.13459.
- [7] Hani A. Alturkistani, Faris M. Tashkandi und Zuhair M. Mohammedsaleh. „Histological Stains: A Literature Review and Case Study“. In: *Global journal of health science* 8.3 (2015), Seiten 72–79. ISSN: 1916-9736. DOI: 10.5539/gjhs.v8n3p72.
- [8] James C. E. Underwood. „More than meets the eye: the changing face of histopathology“. In: *Histopathology* 70.1 (2017), Seiten 4–9. DOI: 10.1111/his.13047.
- [9] Arslaan Javaeed, Shanza Qamar, Sundus Ali et al. „Histological Stains in the Past, Present, and Future“. In: *Cureus* 13.10 (2021), e18486. ISSN: 2168-8184. DOI: 10.7759/cureus.18486.
- [10] Martin J. van den Bent. „Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician’s perspective“. In: *Acta neuropathologica* 120.3 (2010), Seiten 297–304. DOI: 10.1007/s00401-010-0725-7.
- [11] Mieke R. van Bockstal, Martine Berlière, Francois P. Duhoux et al. „Interobserver Variability in Ductal Carcinoma In Situ of the Breast“. In: *American journal of clinical pathology* 154.5 (2020), Seiten 596–609. DOI: 10.1093/ajcp/aqaa077.

-
- [12] Mario Krause, Petra Rösch, Benno Radt et al. „Localizing and identifying living bacteria in an abiotic environment by a combination of Raman and fluorescence microscopy“. In: *Analytical chemistry* 80.22 (2008), Seiten 8568–8575. DOI: 10.1021/ac8014559.
- [13] Katharina Klein, Alexander M. Gigler, Thomas Aschenbrenner et al. „Label-free live-cell imaging with confocal Raman microscopy“. In: *Biophysical journal* 102.2 (2012), Seiten 360–368. DOI: 10.1016/j.bpj.2011.12.027.
- [14] Eva Brauchle, Anne Knopf, Hannah Bauer et al. „Non-invasive Chamber-Specific Identification of Cardiomyocytes in Differentiating Pluripotent Stem Cells“. In: *Stem cell reports* 6.2 (2016), Seiten 188–199. DOI: 10.1016/j.stemcr.2015.12.007.
- [15] Hao He, Mengxi Xu, Cheng Zong et al. „Speeding Up the Line-Scan Raman Imaging of Living Cells by Deep Convolutional Neural Network“. In: *Analytical chemistry* 91.11 (2019), Seiten 7070–7077. DOI: 10.1021/acs.analchem.8b05962.
- [16] Charalambos Kallepitis, Mads S. Bergholt, Manuel M. Mazo et al. „Quantitative volumetric Raman imaging of three dimensional cell cultures“. In: *Nature communications* 8 (2017), Seite 14843. DOI: 10.1038/ncomms14843.
- [17] Murali K. Maruthamuthu, Amir Hossein Raffiee, Denilson Mendes de Oliveira et al. „Raman spectra-based deep learning: A tool to identify microbial contamination“. In: *Microbiology-Open* 9.11 (2020), e1122. DOI: 10.1002/mbo3.1122.
- [18] Sara Mattana, Maurizio Mattarelli, Lorena Urbanelli et al. „Non-contact mechanical and chemical analysis of single living cells by microspectroscopic techniques“. In: *Light, science & applications* 7 (2018), Seite 17139. DOI: 10.1038/lsa.2017.139.
- [19] Julia Marzi, Eva M. Brauchle, Katja Schenke-Layland et al. „Non-invasive functional molecular phenotyping of human smooth muscle cells utilized in cardiovascular tissue engineering“. In: *Acta biomaterialia* 89 (2019), Seiten 193–205. DOI: 10.1016/j.actbio.2019.03.026.
- [20] I. Notingher, S. Verrier, S. Haque et al. „Spectroscopic study of human lung epithelial cells (A549) in culture: living cells versus dead cells“. In: *Biopolymers* 72.4 (2003), Seiten 230–240. ISSN: 0006-3525. DOI: 10.1002/bip.10378.
- [21] Iwan W. Schie und Thomas Huser. „Methods and applications of Raman microspectroscopy to single-cell analysis“. In: *Applied spectroscopy* 67.8 (2013), Seiten 813–828. DOI: 10.1366/12-06971.
- [22] Rachael Smith, Karen L. Wright und Lorna Ashton. „Raman spectroscopy: an evolving technique for live cell studies“. In: *The Analyst* 141.12 (2016), Seiten 3590–3600. DOI: 10.1039/c6an00152a.
- [23] Neil J. Everall. „Confocal Raman microscopy: performance, pitfalls, and best practice“. In: *Applied spectroscopy* 63.9 (2009), 245A–262A. DOI: 10.1366/000370209789379196.
- [24] Eva Hedlund und Qiaolin Deng. „Single-cell RNA sequencing: Technical advancements and biological applications“. In: *Molecular aspects of medicine* 59 (2018), Seiten 36–46. DOI: 10.1016/j.mam.2017.07.003.

-
- [25] Byungjin Hwang, Ji Hyun Lee und Duhee Bang. „Single-cell RNA sequencing technologies and bioinformatics pipelines“. In: *Experimental & molecular medicine* 50.8 (2018), Seiten 1–14. DOI: 10.1038/s12276-018-0071-8.
- [26] Thale Kristin Olsen und Ninib Baryawno. „Introduction to Single-Cell RNA Sequencing“. In: *Current protocols in molecular biology* 122.1 (2018), e57. DOI: 10.1002/cpmb.57.
- [27] Jose Alquicira-Hernandez, Anuja Sathe, Hanlee P. Ji et al. „scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data“. In: *Genome biology* 20.1 (2019), Seite 264. DOI: 10.1186/s13059-019-1862-5.
- [28] Jian Hu, Xiangjie Li, Gang Hu et al. „Iterative transfer learning with neural network for clustering and cell type classification in single-cell RNA-seq analysis“. In: *Nature machine intelligence* 2.10 (2020), Seiten 607–618. ISSN: 2522-5839. DOI: 10.1038/s42256-020-00233-7.
- [29] Benoît Schmauch, Alberto Romagnoni, Elodie Pronier et al. „A deep learning model to predict RNA-Seq expression of tumours from whole slide images“. In: *Nature communications* 11.1 (2020), Seite 3877. DOI: 10.1038/s41467-020-17678-4.
- [30] Yusuke Inoue, Ana Nikolic, Dylan Farnsworth et al. „Extracellular signal-regulated kinase mediates chromatin rewiring and lineage transformation in lung cancer“. In: *eLife* 10 (2021). DOI: 10.7554/eLife.66524.
- [31] Álvaro Quintanal-Villalonga, Joseph M. Chan, Ignas Masilionis et al. „Protocol to dissociate, process, and analyze the human lung tissue using single-cell RNA-seq“. In: *STAR protocols* 3.4 (2022), Seite 101776. DOI: 10.1016/j.xpro.2022.101776.
- [32] Mario Morone, Maria Antonietta Bali, Nina Tunariu et al. „Whole-Body MRI: Current Applications in Oncology“. In: *AJR. American journal of roentgenology* 209.6 (2017), W336–W349. DOI: 10.2214/AJR.17.17984.
- [33] J. M. Winfield, M. D. Blackledge, N. Tunariu et al. „Whole-body MRI: a practical guide for imaging patients with malignant bone disease“. In: *Clinical radiology* 76.10 (2021), Seiten 715–727. DOI: 10.1016/j.crad.2021.04.001.
- [34] H. Peemoeller und M.M Pintar. „Two-dimensional time-evolution approach for resolving a composite free-induction decay“. In: *Journal of Magnetic Resonance (1969)* 41.2 (1980), Seiten 358–360. ISSN: 00222364. DOI: 10.1016/0022-2364(80)90085-2.
- [35] H. Peemoeller, R. K. Shenoy, M. M. Pintar et al. „Improved characterization of healthy and malignant tissue by NMR line-shape relaxation correlations“. In: *Biophysical journal* 38.3 (1982), Seiten 271–276. DOI: 10.1016/S0006-3495(82)84558-X.
- [36] H. Peemoeller, R.K Shenoy und M.M Pintar. „Two-dimensional nmr time evolution correlation spectroscopy in wet lysozyme“. In: *Journal of Magnetic Resonance (1969)* 45.2 (1981), Seiten 193–204. ISSN: 00222364. DOI: 10.1016/0022-2364(81)90116-5.
- [37] C. Badve, A. Yu, S. Dastmalchian et al. „MR Fingerprinting of Adult Brain Tumors: Initial Experience“. In: *AJNR. American journal of neuroradiology* 38.3 (2017), Seiten 492–499. DOI: 10.3174/ajnr.A5035.

-
- [38] Vera Catharina Keil, Stilyana Peteva Bakoeva, Alina Jurcoane et al. „A pilot study of magnetic resonance fingerprinting in Parkinson’s disease“. In: *NMR in biomedicine* 33.11 (2020), e4389. DOI: 10.1002/nbm.4389.
- [39] Celso F. Balthazar, Jonas T. Guimarães, Ramon S. Rocha et al. „Nuclear magnetic resonance as an analytical tool for monitoring the quality and authenticity of dairy foods“. In: *Trends in Food Science & Technology* 108 (2021), Seiten 84–91. ISSN: 09242244. DOI: 10.1016/j.tifs.2020.12.011.
- [40] Laurance D. Hall, Stephen D. Evans und Kevin P. Nott. „Measurement of textural changes of food by MRI relaxometry“. In: *Magnetic resonance imaging* 16.5-6 (1998), Seiten 485–492. DOI: 10.1016/S0730-725X(98)00116-7.
- [41] Martin D. Hürlimann, Lauren Burcaw und Yi-Qiao Song. „Quantitative characterization of food products by two-dimensional D-T2 and T1-T2 distribution functions in a static gradient“. In: *Journal of colloid and interface science* 297.1 (2006), Seiten 303–311. ISSN: 0021-9797. DOI: 10.1016/j.jcis.2005.10.047.
- [42] François Mariette. „NMR Relaxometry and Imaging of Dairy Products“. In: *Modern Magnetic Resonance*. Herausgegeben von Graham A. Webb. Cham: Springer International Publishing, 2017, Seiten 1–23. ISBN: 978-3-319-28275-6. DOI: 10.1007/978-3-319-28275-6{\textunderscore}38-1.
- [43] N. Marigheto, Luca Venturi und B. Hills. „Two-dimensional NMR relaxation studies of apple quality“. In: *Postharvest Biology and Technology* 48.3 (2008), Seiten 331–340. ISSN: 09255214. DOI: 10.1016/j.postharvbio.2007.11.002.
- [44] Sirvan Sultan Uguz, E. Burcin Ozvural, Maria Jardim Beira et al. „Use of NMR Relaxometry to identify frankfurters of different meat sources“. In: *Molecular Physics* 117.7-8 (2019), Seiten 1015–1019. ISSN: 0026-8976. DOI: 10.1080/00268976.2018.1542162.
- [45] L. Venturi, N. Woodward, D. Hibberd et al. „Multidimensional Cross-Correlation Relaxometry of Aqueous Protein Systems“. In: *Applied Magnetic Resonance* 33.3 (2008), Seiten 213–234. ISSN: 0937-9347. DOI: 10.1007/s00723-008-0066-z.
- [46] Daniel Haddad, Philipp Mörchel, Markus Hildenbrand et al. „Selected Magnetic Resonance applications for non-destructive material testing“. In: *tm - Technisches Messen* 88.1 (2021), Seiten 17–23. ISSN: 0171-8096. DOI: 10.1515/teme-2019-0149.
- [47] Joshua Warner, Simon Donell, Kevin Wright et al. „The characterisation of mammalian tissue with 2D relaxation methods“. In: *Magnetic resonance imaging* 28.7 (2010), Seiten 971–981. DOI: 10.1016/j.mri.2010.03.015.
- [48] G. Saab, R. T. Thompson, G. D. Marsh et al. „Two-dimensional time correlation relaxometry of skeletal muscle in vivo at 3 Tesla“. In: *Magnetic resonance in medicine* 46.6 (2001), Seiten 1093–1098. ISSN: 0740-3194. DOI: 10.1002/mrm.1304.
- [49] Natalie C. Chuck, Fahd Azzabi Zouraq, Markus Rottmar et al. „MR imaging relaxometry allows noninvasive characterization of in vivo differentiation of muscle precursor cells“. In: *Radiology* 274.3 (2015), Seiten 800–809. DOI: 10.1148/radiol.14140483.

-
- [50] Dominik Weishaupt, Victor D. Köchli und Borut Marincek. *Wie funktioniert MRI?* Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. ISBN: 978-3-642-41615-6. DOI: 10.1007/978-3-642-41616-3.
- [51] T. Feuillet, M-J Seurin, O. Leveneur et al. „Coil optimization for low-field MRI: a dedicated process for small animal preclinical studies“. In: *Laboratory animals* 49.2 (2015), Seiten 153–167. DOI: 10.1177/0023677214558103.
- [52] Siemens healthineers. *Mobile MRI Scanner eco: Eine komplette und flexible Lösung*. 2022. URL: <https://www.siemens-healthineers.com/de/refurbished-systems-medical-imaging-and-therapy/ecoline-refurbished-systems/magnetic-resonance-imaging-ecoline/mobile-mri-scanner-eco>.
- [53] Hyperfine. *Portable MR Imaging. Here. Now*. 2022. URL: <https://hyperfine.io/product/>.
- [54] Clarissa Zimmerman Cooley, Jason P. Stockmann, Thomas Witzel et al. „Design and implementation of a low-cost, tabletop MRI scanner for education and research prototyping“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 310 (2020), Seite 106625. ISSN: 1090-7807. DOI: 10.1016/j.jmr.2019.106625.
- [55] Pure Devices GmbH. *Benchtop MRI spectrometers: high performance - low-cost - portable*. 2022. URL: <https://www.pure-devices.com/>.
- [56] TNW Faculty. *Tabletop MRI*. 2021. URL: <https://www.utwente.nl/en/tnw/mdi/Research/Magnetic%20Imaging/tabletop/>.
- [57] Ruiliang Bai, Alexander Cloninger, Wojciech Czaja et al. „Efficient 2D MRI relaxometry using compressed sensing“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 255 (2015), Seiten 88–99. ISSN: 1090-7807. DOI: 10.1016/j.jmr.2015.04.002.
- [58] Dan Benjamini und Peter J. Basser. „Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 271 (2016), Seiten 40–45. ISSN: 1090-7807. DOI: 10.1016/j.jmr.2016.08.004.
- [59] Luca Venturi, Joshua Warner und Brian Hills. „Multisliced ultrafast 2D relaxometry“. In: *Magnetic resonance imaging* 28.7 (2010), Seiten 964–970. DOI: 10.1016/j.mri.2010.03.024.
- [60] L. Venkataramanan, Yi-Qiao Song und M. D. Hurlimann. „Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions“. In: *IEEE Transactions on Signal Processing* 50.5 (2002), Seiten 1017–1026. ISSN: 1053587X. DOI: 10.1109/78.995059.
- [61] Y-Q Song, L. Venkataramanan, M. D. Hurlimann et al. „T(1)–T(2) correlation spectra obtained using a fast two-dimensional Laplace inversion“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 154.2 (2002), Seiten 261–268. ISSN: 1090-7807. DOI: 10.1006/jmre.2001.2474.
- [62] N. Marigheto, L. Venturi, D. Hibberd et al. „Methods for peak assignment in low-resolution multidimensional NMR cross-correlation relaxometry“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 187.2 (2007), Seiten 327–342. ISSN: 1090-7807. DOI: 10.1016/j.jmr.2007.04.016.

-
- [63] J. Mitchell, T. C. Chandrasekera und L. F. Gladden. „Numerical estimation of relaxation and diffusion distributions in two dimensions“. In: *Progress in nuclear magnetic resonance spectroscopy* 62 (2012), Seiten 34–50. DOI: 10.1016/j.pnmrs.2011.07.002.
- [64] Hans H. Schild. *MRI made easy*. Schering-Diagnostika. Berlin: Schering, 1990. ISBN: 3-921817-41-2.
- [65] Bin Ding, Huimin Qian und Jun Zhou. „Activation functions and their characteristics in deep neural networks“. In: *Proceedings of the 30th Chinese Control and Decision Conference (2018 CCDC)*. Piscataway, NJ: IEEE, 2018, Seiten 1836–1841. ISBN: 978-1-5386-1244-6. DOI: 10.1109/CCDC.2018.8407425.
- [66] Jianli Feng und Shengnan Lu. „Performance Analysis of Various Activation Functions in Artificial Neural Networks“. In: *Journal of Physics: Conference Series* 1237.2 (2019), Seite 022030. ISSN: 1742-6588. DOI: 10.1088/1742-6596/1237/2/022030.
- [67] Yingying Wang, Yibin Li, Yong Song et al. „The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition“. In: *Applied Sciences* 10.5 (2020), Seite 1897. DOI: 10.3390/app10051897.
- [68] Claire Yu-Mei Huang und Matthew N. Rasband. „Axon initial segments: structure, function, and disease“. In: *Annals of the New York Academy of Sciences* 1420.1 (2018), Seiten 46–61. DOI: 10.1111/nyas.13718.
- [69] Siddharth Sharma, Simone Sharma und Anidhya Athaiya. „ACTIVATION FUNCTIONS IN NEURAL NETWORKS“. In: *International Journal of Engineering Applied Sciences and Technology* 04.12 (2020), Seiten 310–316. DOI: 10.33564/IJEAST.2020.v04i12.054.
- [70] Nikolaus Kriegeskorte und Tal Golan. „Neural network models and deep learning“. In: *Current biology : CB* 29.7 (2019), R231–R236. DOI: 10.1016/j.cub.2019.02.034.
- [71] David E. Rumelhart, Geoffrey E. Hinton und Ronald J. Williams. „Learning representations by back-propagating errors“. In: *Nature* 323.6088 (1986), Seiten 533–536. DOI: 10.1038/323533a0.
- [72] Diederik P. Kingma und Jimmy Ba. *Adam: A Method for Stochastic Optimization*. URL: <http://arxiv.org/pdf/1412.6980v9>.
- [73] Anita Rácz, Dávid Bajusz und Károly Héberger. „Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification“. In: *Molecules (Basel, Switzerland)* 26.4 (2021). DOI: 10.3390/molecules26041111.
- [74] Jimin Tan, Jianan Yang, Sai Wu et al. *A critical look at the current train/test split in machine learning*. URL: <https://arxiv.org/pdf/2106.04525>.
- [75] K. Fukushima. „Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position“. In: *Biological cybernetics* 36.4 (1980), Seiten 193–202. ISSN: 0340-1200. DOI: 10.1007/BF00344251.
- [76] Y. LeCun, B. Boser, J. S. Denker et al. „Backpropagation Applied to Handwritten Zip Code Recognition“. In: *Neural Computation* 1.4 (1989), Seiten 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541.

-
- [77] Y. LeCun, L. Bottou, Y. Bengio et al. „Gradient-based learning applied to document recognition“. In: *Proceedings of the IEEE* 86.11 (1998), Seiten 2278–2324. ISSN: 00189219. DOI: 10.1109/5.726791.
- [78] Saad Albawi, Tareq Abed Mohammed und Saad Al-Zawi. „Understanding of a convolutional neural network“. In: *Proceedings of 2017 International Conference on Engineering & Technology (ICET'2017)*. Piscataway, NJ: IEEE, 2017, Seiten 1–6. ISBN: 978-1-5386-1949-0. DOI: 10.1109/ICEngTechnol.2017.8308186.
- [79] Keiron O’Shea und Ryan Nash. *An Introduction to Convolutional Neural Networks*. URL: <https://arxiv.org/pdf/1511.08458>.
- [80] Olga Russakovsky, Jia Deng, Hao Su et al. „ImageNet Large Scale Visual Recognition Challenge“. In: *International Journal of Computer Vision* 115.3 (2015), Seiten 211–252. ISSN: 0920-5691. DOI: 10.1007/s11263-015-0816-y.
- [81] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do et al. „Convolutional neural networks: an overview and application in radiology“. In: *Insights into imaging* 9.4 (2018), Seiten 611–629. ISSN: 1869-4101. DOI: 10.1007/s13244-018-0639-9.
- [82] Zhi Li, Dengshi Zhou, Li Wan et al. „Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram“. In: *Journal of electrocardiology* 58 (2020), Seiten 105–112. DOI: 10.1016/j.jelectrocard.2019.11.046.
- [83] Teja Kattenborn, Jens Leitloff, Felix Schiefer et al. „Review on Convolutional Neural Networks (CNN) in vegetation remote sensing“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 173 (2021), Seiten 24–49. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2020.12.010.
- [84] Jinzhu Lu, Lijuan Tan und Huanyu Jiang. „Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification“. In: *Agriculture* 11.8 (2021), Seite 707. DOI: 10.3390/agriculture11080707.
- [85] Pedro R. A. S. Bassi und Romis Attux. „A deep convolutional neural network for COVID-19 detection using chest X-rays“. In: *Research on Biomedical Engineering* 38.1 (2022), Seiten 139–148. ISSN: 2446-4732. DOI: 10.1007/s42600-021-00132-9.
- [86] Xinru Kong, Yan Yao, Cuiying Wang et al. „Automatic Identification of Depression Using Facial Images with Deep Convolutional Neural Network“. In: *Medical science monitor : international medical journal of experimental and clinical research* 28 (2022), e936409. DOI: 10.12659/MSM.936409.
- [87] Eva Bianconi, Allison Piovesan, Federica Facchin et al. „An estimation of the number of cells in the human body“. In: *Annals of human biology* 40.6 (2013), Seiten 463–471. DOI: 10.3109/03014460.2013.807878.
- [88] Ron Sender, Shai Fuchs und Ron Milo. „Revised Estimates for the Number of Human and Bacteria Cells in the Body“. In: *PLoS biology* 14.8 (2016), e1002533. DOI: 10.1371/journal.pbio.1002533.

-
- [89] Neil A. Campbell und Jane B. Reece. *Biologie*. 8., akt. Aufl., [Nachdr.] Pearson Studium - Biologie. München: Pearson Studium, 2011. ISBN: 9783827372871. URL: <http://swbplus.bsz-bw.de/bsz333360648cov.htm>.
- [90] Scott F. Gilbert. *Developmental biology*. 8. ed., 2. print. Sunderland, Mass.: Sinauer, 2006. ISBN: 9780878932504.
- [91] Vimal K. Singh, Abhishek Saini, Manisha Kalsan et al. „Describing the Stem Cell Potency: The Various Methods of Functional Assessment and In silico Diagnostics“. In: *Frontiers in cell and developmental biology* 4 (2016), Seite 134. ISSN: 2296-634X. DOI: 10.3389/fcell.2016.00134.
- [92] J. A. Thomson, J. Itskovitz-Eldor, S. S. Shapiro et al. „Embryonic stem cell lines derived from human blastocysts“. In: *Science (New York, N.Y.)* 282.5391 (1998), Seiten 1145–1147. ISSN: 0036-8075. DOI: 10.1126/science.282.5391.1145.
- [93] Stammzellnetzwerk.NRW e.V. *Was für Arten von Stammzellen gibt es?* 2022. URL: <https://www.stammzellen.nrw.de/informieren/stammzellen-verstehen/was-fuer-arten-von-stammzellen-gibt-es>.
- [94] Bundestag. „Gesetz zum Schutz von Embryonen: Embryonenschutzgesetz (ESchG)“. In: (). URL: <https://www.gesetze-im-internet.de/eschg/BJNR027460990.html>.
- [95] J. B. Gurdon. „The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles“. In: *Development* 10.4 (1962), Seiten 622–640. ISSN: 0950-1991. DOI: 10.1242/dev.10.4.622.
- [96] Kazutoshi Takahashi und Shinya Yamanaka. „Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors“. In: *Cell* 126.4 (2006), Seiten 663–676. ISSN: 0092-8674. DOI: 10.1016/j.cell.2006.07.024.
- [97] Nobel Prize Outreach. *The Nobel Prize in Physiology or Medicine 2012: jointly to John B. Gurdon and Shinya Yamanaka for the discovery that mature cells can be reprogrammed to become pluripotent*. 8.10.2012. URL: <https://www.nobelprize.org/prizes/medicine/2012/press-release/>.
- [98] OECD. *Health at a Glance 2019*. OECD, 2019. ISBN: 9789264382084. DOI: 10.1787/4dd50c09-en.
- [99] James A. Ankrum, Joon Faii Ong und Jeffrey M. Karp. „Mesenchymal stem cells: immune evasive, not immune privileged“. In: *Nature biotechnology* 32.3 (2014), Seiten 252–260. DOI: 10.1038/nbt.2816.
- [100] Christian Horst Tonk, Markus Witzler, Margit Schulze et al. „Mesenchymal Stem Cells“. In: *Essential Current Concepts in Stem Cell Biology*. Herausgegeben von Beate Brand-Saberi. Learning Materials in Biosciences. Cham: Springer International Publishing, 2020, Seiten 21–39. ISBN: 978-3-030-33922-7. DOI: 10.1007/978-3-030-33923-4₂.

- [101] Shivangi Mishra, Jayesh Kumar Sevak, Anamica Das et al. „Publisher Correction: Umbilical cord tissue is a robust source for mesenchymal stem cells with enhanced myogenic differentiation potential compared to cord blood“. In: *Scientific reports* 11.1 (2021), Seite 13825. DOI: 10.1038/s41598-021-93047-5.
- [102] Michael J. Chen, Jonathan P. Whiteley, Colin P. Please et al. „Identifying chondrogenesis strategies for tissue engineering of articular cartilage“. In: *Journal of tissue engineering* 10 (2019), Seite 2041731419842431. ISSN: 2041-7314. DOI: 10.1177/2041731419842431.
- [103] Davide Confalonieri, Andrea Schwab, Heike Walles et al. „Advanced Therapy Medicinal Products: A Guide for Bone Marrow-derived MSC Application in Bone and Cartilage Tissue Engineering“. In: *Tissue engineering. Part B, Reviews* 24.2 (2018), Seiten 155–169. DOI: 10.1089/ten.TEB.2017.0305.
- [104] Franziska Ehlicke, Jonathan Berndt, Nina Marichikj et al. „Biomimetic in vitro test system for evaluation of dental implant materials“. In: *Dental materials : official publication of the Academy of Dental Materials* 36.8 (2020), Seiten 1059–1070. DOI: 10.1016/j.dental.2020.04.020.
- [105] Christoph Malkmus, Shabnam Jawahar, Nancy Tricoche et al. „Preliminary evaluations of 3-dimensional human skin models for their ability to facilitate in vitro the long-term development of the debilitating obligatory human parasite *Onchocerca volvulus*“. In: *PLoS neglected tropical diseases* 14.11 (2020), e0008503. DOI: 10.1371/journal.pntd.0008503.
- [106] Tobias Weigel, Christoph Malkmus, Verena Weigel et al. „Full synthetic 3D fibrous scaffolds for stromal tissues - replacement of animal-derived scaffold materials demonstrated by multilayered skin“. In: *Advanced materials (Deerfield Beach, Fla.)* (2021), e2106780. DOI: 10.1002/adma.202106780.
- [107] Leili Aghebati-Maleki, Sanam Dolati, Reza Zandi et al. „Prospect of mesenchymal stem cells in therapy of osteoporosis: A review“. In: *Journal of cellular physiology* 234.6 (2019), Seiten 8570–8578. DOI: 10.1002/jcp.27833.
- [108] Yves-Marie Pers, Lars Rackwitz, Rosanna Ferreira et al. „Adipose Mesenchymal Stromal Cell-Based Therapy for Severe Osteoarthritis of the Knee: A Phase I Dose-Escalation Trial“. In: *Stem cells translational medicine* 5.7 (2016), Seiten 847–856. ISSN: 2157-6564. DOI: 10.5966/sctm.2015-0245.
- [109] Fabio Cofano, Marina Boido, Matteo Monticelli et al. „Mesenchymal Stem Cells for Spinal Cord Injury: Current Options, Limitations, and Future of Cell Therapy“. In: *International journal of molecular sciences* 20.11 (2019). DOI: 10.3390/ijms20112698.
- [110] Yajun Guo, Yunsheng Yu, Shijun Hu et al. „The therapeutic potential of mesenchymal stem cells for cardiovascular diseases“. In: *Cell death & disease* 11.5 (2020), Seite 349. DOI: 10.1038/s41419-020-2542-9.
- [111] Yi Wang, Su-Kang Shan, Bei Guo et al. „The Multi-Therapeutic Role of MSCs in Diabetic Nephropathy“. In: *Frontiers in endocrinology* 12 (2021), Seite 671566. ISSN: 1664-2392. DOI: 10.3389/fendo.2021.671566.

-
- [112] Chenghai Li, Hua Zhao und Bin Wang. „Mesenchymal stem/stromal cells: Developmental origin, tumorigenesis and translational cancer therapeutics“. In: *Translational oncology* 14.1 (2021), Seite 100948. ISSN: 1936-5233. DOI: 10.1016/j.tranon.2020.100948.
- [113] Jin-Yang Chen, Xiao-Zhou Mou, Xiao-Chun Du et al. „Comparative analysis of biological characteristics of adult mesenchymal stem cells with different tissue origins“. In: *Asian Pacific journal of tropical medicine* 8.9 (2015), Seiten 739–746. DOI: 10.1016/j.apjtm.2015.07.022.
- [114] Marwa El-Sayed, Mohamed Ali El-Feky, Mostafa I. El-Amir et al. „Immunomodulatory effect of mesenchymal stem cells: Cell origin and cell quality variations“. In: *Molecular biology reports* 46.1 (2019), Seiten 1157–1165. DOI: 10.1007/s11033-018-04582-w.
- [115] Yen-Ting Ho, Takashi Shimbo, Edward Wijaya et al. „Chromatin accessibility identifies diversity in mesenchymal stem cells from different tissue origins“. In: *Scientific reports* 8.1 (2018), Seite 17765. DOI: 10.1038/s41598-018-36057-0.
- [116] Noushin Ouryazdanpanah, Shahriar Dabiri, Ali Derakhshani et al. „Peripheral Blood-Derived Mesenchymal Stem Cells: Growth Factor-Free Isolation, Molecular Characterization and Differentiation“. In: *Iranian journal of pathology* 13.4 (2018), Seiten 461–466. ISSN: 1735-5303.
- [117] Wenyan Zhou, Junxin Lin, Kun Zhao et al. „Single-Cell Profiles and Clinically Useful Properties of Human Mesenchymal Stem Cells of Adipose and Bone Marrow Origin“. In: *The American journal of sports medicine* 47.7 (2019), Seiten 1722–1733. DOI: 10.1177/0363546519848678.
- [118] *Scikit-learn: Machine learning in Python*. 2011. URL: <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://githubhelp.com>.
- [119] Lars Buitinck, Gilles Louppe, Mathieu Blondel et al. „API design for machine learning software: experiences from the scikit-learn project“. In: *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 2013, Seiten 108–122.
- [120] Karen Simonyan und Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. URL: <https://arxiv.org/pdf/1409.1556>.
- [121] Adam Paszke, Sam Gross, Francisco Massa et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. URL: <https://arxiv.org/pdf/1912.01703>.
- [122] Philipp Fey, Daniel Ludwig Weber, Jannik Stebani et al. „Non-destructive classification of unlabeled cells: Combining an automated benchtop magnetic resonance scanner and artificial intelligence“. In: *PLoS computational biology* 19.2 (2023), e1010842. DOI: 10.1371/journal.pcbi.1010842.
- [123] *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012. ISBN: 978-1-4673-1490-9.
- [124] Junghwan Cho, Kyewook Lee, Ellie Shin et al. *How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?* 2015. DOI: 10.48550/arXiv.1511.06348.

-
- [125] Eric B. Baum und David Haussler. „What Size Net Gives Valid Generalization?“ In: *Neural Computation* 1.1 (1989), Seiten 151–160. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.1.151.
- [126] Y. S. Abu-Mostafa. „Hints“. In: *Neural Computation* 7.4 (1995), Seiten 639–671. ISSN: 0899-7667. DOI: 10.1162/neco.1995.7.4.639.
- [127] Simon S. Haykin. *Neural networks and learning machines*. 3. ed. New York: Pearson, 2009. ISBN: 978-0-13-147139-9.
- [128] Ahmad Alwosheel, Sander van Cranenburgh und Caspar G. Chorus. „Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis“. In: *Journal of Choice Modelling* 28 (2018), Seiten 167–182. ISSN: 17555345. DOI: 10.1016/j.jocm.2018.07.002.
- [129] Godwin Inalegwu Ogbale, Adekunle Olakunle Adeyomoye, Augustina Badu-Peprah et al. „Survey of magnetic resonance imaging availability in West Africa“. In: *The Pan African medical journal* 30 (2018), Seite 240. DOI: 10.11604/pamj.2018.30.240.14000.
- [130] Kathryn E. Keenan, Maureen Ainslie, Alex J. Barker et al. „Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom“. In: *Magnetic resonance in medicine* 79.1 (2018), Seiten 48–61. ISSN: 0740-3194. DOI: 10.1002/mrm.26982.
- [131] Robba Rai, Yu Feng Wang, David Manton et al. „Development of multi-purpose 3D printed phantoms for MRI“. In: *Physics in medicine and biology* 64.7 (2019), Seite 075010. DOI: 10.1088/1361-6560/ab0b49.
- [132] K. A. Kraft, P. P. Fatouros, G. D. Clarke et al. „An MRI phantom material for quantitative relaxometry“. In: *Magnetic resonance in medicine* 5.6 (1987), Seiten 555–562. ISSN: 0740-3194. DOI: 10.1002/mrm.1910050606.
- [133] Hirokazu Kato, Masahiro Kuroda, Koichi Yoshimura et al. „Composition of MRI phantom equivalent to human tissues“. In: *Medical physics* 32.10 (2005), Seiten 3199–3208. ISSN: 0094-2405. DOI: 10.1118/1.2047807.
- [134] M. J. Firbank, R. M. Harrison, E. D. Williams et al. „Quality assurance for MRI: practical experience“. In: *The British journal of radiology* 73.868 (2000), Seiten 376–383. ISSN: 0007-1285. DOI: 10.1259/bjr.73.868.10844863.
- [135] Matthew F. Bieniosek, Brian J. Lee und Craig S. Levin. „Technical Note: Characterization of custom 3D printed multimodality imaging phantoms“. In: *Medical physics* 42.10 (2015), Seiten 5913–5918. ISSN: 0094-2405. DOI: 10.1118/1.4930803.
- [136] Els Fieremans und Hong-Hsi Lee. „Physical and numerical phantoms for the validation of brain microstructural MRI: A cookbook“. In: *NeuroImage* 182 (2018), Seiten 39–61. DOI: 10.1016/j.neuroimage.2018.06.046.
- [137] Kengo Hattori, Yusuke Ikemoto, Wataru Takao et al. „Development of MRI phantom equivalent to human tissues for 3.0-T MRI“. In: *Medical physics* 40.3 (2013), Seite 032303. ISSN: 0094-2405. DOI: 10.1118/1.4790023.

- [138] Sergio Curto, Bassim Aklan, Tim Mulder et al. „Quantitative, Multi-institutional Evaluation of MR Thermometry Accuracy for Deep-Pelvic MR-Hyperthermia Systems Operating in Multi-vendor MR-systems Using a New Anthropomorphic Phantom“. In: *Cancers* 11.11 (2019). ISSN: 2072-6694. DOI: 10.3390/cancers11111709.
- [139] Gergő Veres, János Kiss, Norman Félix Vas et al. „Phantom Study on the Robustness of MR Radiomics Features: Comparing the Applicability of 3D Printed and Biological Phantoms“. In: *Diagnostics (Basel, Switzerland)* 12.9 (2022). ISSN: 2075-4418. DOI: 10.3390/diagnostics12092196.
- [140] Roberta Matheoud, Ronald Boellaard, Lucy Pike et al. „EFOMP’s protocol quality controls in PET/CT and PET/MR“. In: *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB)* 105 (2023), Seite 102506. DOI: 10.1016/j.ejmp.2022.11.010.
- [141] Eppendorf AG. *Eppendorf Research plus: Werksjustierung*. 2013. URL: https://www.eppendorf.com/product-media/doc/de/327080/Eppendorf_Liquid-Handling_Adjustment_Research-plus_Factory-Adjustment_Eppendorf-Research-plus-Factory-Adjustment.pdf.
- [142] H. R. Hart, P. A. Bottomley, W. A. Edelstein et al. „Nuclear magnetic resonance imaging: contrast-to-noise ratio as a function of strength of magnetic field“. In: *AJR. American journal of roentgenology* 141.6 (1983), Seiten 1195–1201. DOI: 10.2214/ajr.141.6.1195.
- [143] A. J. Maubon, J. M. Ferru, V. Berger et al. „Effect of field strength on MR images: comparison of the same subject at 0.5, 1.0, and 1.5 T“. In: *Radiographics : a review publication of the Radiological Society of North America, Inc* 19.4 (1999), Seiten 1057–1067. ISSN: 0271-5333. DOI: 10.1148/radiographics.19.4.g99j1281057.
- [144] Satish Mulleti, Amrinder Singh, Varsha P. Brahmkhatri et al. „Super-Resolved Nuclear Magnetic Resonance Spectroscopy“. In: *Scientific reports* 7.1 (2017), Seite 9651. DOI: 10.1038/s41598-017-09884-w.
- [145] Oliver Kraff und Harald H. Quick. „7T: Physics, safety, and potential clinical applications“. In: *Journal of magnetic resonance imaging : JMRI* 46.6 (2017), Seiten 1573–1589. DOI: 10.1002/jmri.25723.
- [146] Albrecht Stroh, Cornelius Faber, Thomas Neuberger et al. „In vivo detection limits of magnetically labeled embryonic stem cells in the rat brain using high-field (17.6 T) magnetic resonance imaging“. In: *NeuroImage* 24.3 (2005), Seiten 635–645. DOI: 10.1016/j.neuroimage.2004.09.014.
- [147] A. G. Webb. „Radiofrequency microcoils for magnetic resonance imaging and spectroscopy“. In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 229 (2013), Seiten 55–66. ISSN: 1090-7807. DOI: 10.1016/j.jmr.2012.10.004.
- [148] Hans Weber, Nicoleta Baxan, Dominik Paul et al. „Microcoil-based MRI: feasibility study and cell culture applications using a conventional animal system“. In: *Magma (New York, N.Y.)* 24.3 (2011), Seiten 137–145. DOI: 10.1007/s10334-011-0244-0.

- [149] Magdalena Couty, Anne Rubin, Marion Woytasik et al. „Implantable wireless microcoils for 7Tesla Magnetic Resonance Imaging of the rat brain: Optimization of the PDMS packaging“. In: *2012 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*. IEEE, 2012, Seiten 37–40. ISBN: 978-1-4577-1136-7. DOI: 10.1109/BioWireless.2012.6172728.
- [150] Justine Deborne, Noël Pinaud und Yannick Crémillieux. „Implantable NMR Microcoils in Rats: A New Tool for Exploring Tumor Metabolism at Sub-Microliter Scale?“ In: *Metabolites* 11.3 (2021). ISSN: 2218-1989. DOI: 10.3390/metabo11030176.
- [151] Niryo. *Niryo Ned2*. 2022. URL: <https://niryo.com/product/ned-2-english/>.
- [152] OECD. *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*. OECD, 2021. ISBN: 9789264265295. DOI: 10.1787/9789264265295-en.
- [153] Shannon G. Klein, Alexandra Steckbauer, Samhan M. Alsolami et al. „Toward Best Practices for Controlling Mammalian Cell Culture Environments“. In: *Frontiers in cell and developmental biology* 10 (2022), Seite 788808. ISSN: 2296-634X. DOI: 10.3389/fcell.2022.788808.
- [154] MERCK. *Dulbecco’s Modified Eagle’s Medium (DME) Formulation*. 2022. URL: <https://www.sigmaaldrich.com/DE/de/technical-documents/technical-article/cell-culture-and-cell-culture-analysis/mammalian-cell-culture/dulbecco-modified-eagle-medium-formulation>.
- [155] Andre D. Castiaux, Dana M. Spence und R. Scott Martin. „Review of 3D Cell Culture with Analysis in Microfluidic Systems“. In: *Analytical methods : advancing methods and applications* 11.33 (2019), Seiten 4220–4232. ISSN: 1759-9660. DOI: 10.1039/C9AY01328H.
- [156] Philip Dettinger, Tino Frank, Martin Etzrodt et al. „Automated Microfluidic System for Dynamic Stimulation and Tracking of Single Cells“. In: *Analytical chemistry* 90.18 (2018), Seiten 10695–10700. DOI: 10.1021/acs.analchem.8b00312.
- [157] Melikhan Tanyeri und Savaş Tay. „Viable cell culture in PDMS-based microfluidic devices“. In: *Methods in cell biology* 148 (2018), Seiten 3–33. ISSN: 0091-679X. DOI: 10.1016/bs.mcb.2018.09.007.
- [158] Jihyoung Choi, Sanjana Mathew, Sabrina Oerter et al. „Online Measurement System for Dynamic Flow Bioreactors to Study Barrier Integrity of hiPSC-Based Blood-Brain Barrier In Vitro Models“. In: *Bioengineering (Basel, Switzerland)* 9.1 (2022). ISSN: 2306-5354. DOI: 10.3390/bioengineering9010039.
- [159] B. R. Cherry und T. M. Alam. „Relaxation nuclear magnetic resonance imaging (R-NMRI) of PDMS/PDPS siloxane copolymer desiccation“. In: *Polymer* 45.16 (2004), Seiten 5611–5618. ISSN: 00323861. DOI: 10.1016/j.polymer.2004.05.065.
- [160] R. Samuel, H. J. Sant, F. Jiao et al. „Microfluidic laminate-based phantom for diffusion tensor-magnetic resonance imaging (DT-MRI)“. In: *Journal of micromechanics and micro-engineering : structures, devices, and systems* 21.9 (2011), Seiten 950271–9502711. ISSN: 0960-1317. DOI: 10.1088/0960-1317/21/9/095027.

-
- [161] Carsten Funck, Frederik Bernd Laun und Andreas Wetscherek. „Characterization of the diffusion coefficient of blood“. In: *Magnetic resonance in medicine* 79.5 (2018), Seiten 2752–2758. ISSN: 0740-3194. DOI: 10.1002/mrm.26919.
- [162] Ganesh Adluru, Yaniv Gur, Jeffrey S. Anderson et al. „Assessment of white matter microstructure in stroke patients using NODDI“. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2014* (2014), Seiten 742–745. DOI: 10.1109/EMBC.2014.6943697.
- [163] Nian Wang, Jieying Zhang, Gary Cofer et al. „Neurite orientation dispersion and density imaging of mouse brain microstructure“. In: *Brain structure & function* 224.5 (2019), Seiten 1797–1813. ISSN: 1863-2653. DOI: 10.1007/s00429-019-01877-x.
- [164] Xiaoyu Jiang, Hua Li, Sean P. Devan et al. „MR cell size imaging with temporal diffusion spectroscopy“. In: *Magnetic resonance imaging* 77 (2021), Seiten 109–123. DOI: 10.1016/j.mri.2020.12.010.
- [165] Hamsanandini Radhakrishnan, Sepideh Kiani Shabestari, Mathew Blurton-Jones et al. „Using Advanced Diffusion-Weighted Imaging to Predict Cell Counts in Gray Matter: Potential and Pitfalls“. In: *Frontiers in neuroscience* 16 (2022), Seite 881713. ISSN: 1662-4548. DOI: 10.3389/fnins.2022.881713.
- [166] A. Einstein. „Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen“. In: *Annalen der Physik* 322.8 (1905), Seiten 549–560. ISSN: 00033804. DOI: 10.1002/andp.19053220806.
- [167] D. Le Bihan. „Molecular diffusion, tissue microdynamics and microstructure“. In: *NMR in biomedicine* 8.7-8 (1995), Seiten 375–386. DOI: 10.1002/nbm.1940080711.
- [168] J. MacFall, D. M. Prescott, E. Fullar et al. „Temperature dependence of canine brain tissue diffusion coefficient measured in vivo with magnetic resonance echo-planar imaging“. In: *International journal of hyperthermia : the official journal of European Society for Hyperthermic Oncology, North American Hyperthermia Group* 11.1 (1995), Seiten 73–86. ISSN: 0265-6736. DOI: 10.3109/02656739509004949.
- [169] V.Ya. Rudyak, S. N. Dubtsov und A. M. Baklanov. „Measurements of the temperature dependent diffusion coefficient of nanoparticles in the range of 295–600K at atmospheric pressure“. In: *Journal of Aerosol Science* 40.10 (2009), Seiten 833–843. ISSN: 00218502. DOI: 10.1016/j.jaerosci.2009.06.006.
- [170] Bruce Alberts, Alexander Johnson, Julian Lewis et al. *Molecular biology of the cell*. Sixth edition. New York, NY: Garland Science Taylor and Francis Group, 2015. ISBN: 9780815344643.
- [171] Alexander Dospil. *Biochemie des Menschen: Das Lehrbuch für das Medizinstudium*. 5., korrigierte Aufl. Stuttgart: Thieme, 2012. ISBN: 978-3-13-130885-6.
- [172] Garth L. Nicolson. „The Fluid-Mosaic Model of Membrane Structure: still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years“. In: *Biochimica et biophysica acta* 1838.6 (2014), Seiten 1451–1466. ISSN: 0006-3002. DOI: 10.1016/j.bbamem.2013.10.019.

-
- [173] Sjoerd van Deventer, Abbey B. Arp und Annemiek B. van Spriel. „Dynamic Plasma Membrane Organization: A Complex Symphony“. In: *Trends in cell biology* 31.2 (2021), Seiten 119–129. DOI: 10.1016/j.tcb.2020.11.004.
- [174] Nita Amornsiripanitch, Sebastian Bickelhaupt, Hee Jung Shin et al. „Diffusion-weighted MRI for Unenhanced Breast Cancer Screening“. In: *Radiology* 293.3 (2019), Seiten 504–520. DOI: 10.1148/radiol.2019182789.
- [175] Matthew Cieslak, Philip A. Cook, Xiaosong He et al. „QSIPrep: an integrative platform for preprocessing and reconstructing diffusion MRI data“. In: *Nature methods* 18.7 (2021), Seiten 775–778. DOI: 10.1038/s41592-021-01185-5.
- [176] Mami Iima, Maya Honda, Eric E. Sigmund et al. „Diffusion MRI of the breast: Current status and future directions“. In: *Journal of magnetic resonance imaging : JMRI* 52.1 (2020), Seiten 70–90. DOI: 10.1002/jmri.26908.
- [177] Phillip Zhe Sun. „Quasi-steady-state CEST (QUASS CEST) solution improves the accuracy of CEST quantification: QUASS CEST MRI-based omega plot analysis“. In: *Magnetic resonance in medicine* 86.2 (2021), Seiten 765–776. ISSN: 0740-3194. DOI: 10.1002/mrm.28744.
- [178] Joshua M. Goldenberg und Mark D. Pagel. „Assessments of tumor metabolism with CEST MRI“. In: *NMR in biomedicine* 32.10 (2019), e3943. DOI: 10.1002/nbm.3943.
- [179] B. Wu, G. Warnock, M. Zaiss et al. „An overview of CEST MRI for non-MR physicists“. In: *EJNMMI physics* 3.1 (2016), Seite 19. ISSN: 2197-7364. DOI: 10.1186/s40658-016-0155-2.
- [180] Michal Rivlin, Judith Horev, Ilan Tsarfaty et al. „Molecular imaging of tumors and metastases using chemical exchange saturation transfer (CEST) MRI“. In: *Scientific reports* 3 (2013), Seite 3045. DOI: 10.1038/srep03045.
- [181] Jianpan Huang, Zilin Chen, Se-Weon Park et al. „Molecular Imaging of Brain Tumors and Drug Delivery Using CEST MRI: Promises and Challenges“. In: *Pharmaceutics* 14.2 (2022). ISSN: 1999-4923. DOI: 10.3390/pharmaceutics14020451.
- [182] Vipul R. Sheth, Yuguo Li, Liu Qi Chen et al. „Measuring in vivo tumor pHe with CEST-FISP MRI“. In: *Magnetic resonance in medicine* 67.3 (2012), Seiten 760–768. ISSN: 0740-3194. DOI: 10.1002/mrm.23038.
- [183] D. Delli Castelli, Giuseppe Ferrauto, Juan Carlos Cutrin et al. „In vivo maps of extracellular pH in murine melanoma by CEST-MRI“. In: *Magnetic resonance in medicine* 71.1 (2014), Seiten 326–332. ISSN: 0740-3194. DOI: 10.1002/mrm.24664.
- [184] Xing Yang, Xiaolei Song, Sangeeta Ray Banerjee et al. „Developing imidazoles as CEST MRI pH sensors“. In: *Contrast media & molecular imaging* 11.4 (2016), Seiten 304–312. DOI: 10.1002/cmri.1693.
- [185] Peng Xiao, Jianpan Huang, Xiongqi Han et al. „Monitor Tumor pHe and Response Longitudinally during Treatment Using CEST MRI-Detectable Alginate Microbeads“. In: *ACS applied materials & interfaces* 14.49 (2022), Seiten 54401–54410. DOI: 10.1021/acsmi.2c10493.

-
- [186] Dario Livio Longo, Elisa Pirotta, Riccardo Gambino et al. „Tumor pH Imaging Using Chemical Exchange Saturation Transfer (CEST)-MRI“. In: *Methods in molecular biology (Clifton, N.J.)* 2614 (2023), Seiten 287–311. DOI: 10.1007/978-1-0716-2914-7{\textunderscore}18.
- [187] Guillaume Madelin, Jae-Seung Lee, Ravinder R. Regatte et al. „Sodium MRI: methods and applications“. In: *Progress in nuclear magnetic resonance spectroscopy* 79 (2014), Seiten 14–47. DOI: 10.1016/j.pnmrs.2014.02.001.
- [188] Olgica Zaric, Vladimir Juras, Pavol Szomolanyi et al. „Frontiers of Sodium MRI Revisited: From Cartilage to Brain Imaging“. In: *Journal of magnetic resonance imaging : JMRI* 54.1 (2021), Seiten 58–75. DOI: 10.1002/jmri.27326.
- [189] François Chollet. *Keras*. 2015. URL: <https://keras.io>.
- [190] Martín Abadi, Paul Barham, Jianmin Chen et al. *TensorFlow: A system for large-scale machine learning*. 2015. URL: <https://www.tensorflow.org>.
- [191] Shelly Soffer, Avi Ben-Cohen, Orit Shimon et al. „Convolutional Neural Networks for Radiologic Images: A Radiologist’s Guide“. In: *Radiology* 290.3 (2019), Seiten 590–606. DOI: 10.1148/radiol.2018180547.
- [192] Gökçen Eraslan, Žiga Avsec, Julien Gagneur et al. „Deep learning: new computational modelling techniques for genomics“. In: *Nature reviews. Genetics* 20.7 (2019), Seiten 389–403. DOI: 10.1038/s41576-019-0122-6.
- [193] Marek Wesolowski und Bogdan Suchacz. „Artificial neural networks: theoretical background and pharmaceutical applications: a review“. In: *Journal of AOAC International* 95.3 (2012), Seiten 652–668. ISSN: 1060-3271. DOI: 10.5740/jaoacint.SGE{\textunderscore}Wesolowski{\textunderscore}ANN.
- [194] Yinqiu Xu, Hequan Yao und Kejiang Lin. „An overview of neural networks for drug discovery and the inputs used“. In: *Expert opinion on drug discovery* 13.12 (2018), Seiten 1091–1102. DOI: 10.1080/17460441.2018.1547278.
- [195] Justin M. Johnson und Taghi M. Khoshgoftaar. „Survey on deep learning with class imbalance“. In: *Journal of big data* 6.1 (2019). ISSN: 2196-1115. DOI: 10.1186/s40537-019-0192-5.
- [196] Kushankur Ghosh, Colin Bellinger, Roberto Corizzo et al. „The class imbalance problem in deep learning“. In: *Machine Learning* (2022). ISSN: 0885-6125. DOI: 10.1007/s10994-022-06268-8.
- [197] Yee Liang Thian, Dian Wen Ng, James Thomas Patrick Decourcy Hallinan et al. „Effect of Training Data Volume on Performance of Convolutional Neural Network Pneumothorax Classifiers“. In: *Journal of digital imaging* 35.4 (2022), Seiten 881–892. DOI: 10.1007/s10278-022-00594-y.
- [198] Deevyankar Agarwal, Gonçalo Marques, Isabel de La Torre-Díez et al. „Transfer Learning for Alzheimer’s Disease through Neuroimaging Biomarkers: A Systematic Review“. In: *Sensors (Basel, Switzerland)* 21.21 (2021). DOI: 10.3390/s21217259.

-
- [199] Shaoya Guan, Tianmiao Wang, Kai Sun et al. „Transfer Learning for Nonrigid 2D/3D Cardiovascular Images Registration“. In: *IEEE journal of biomedical and health informatics* 25.9 (2021), Seiten 3300–3309. DOI: 10.1109/JBHI.2020.3045977.
- [200] Shuteng Niu, Meryl Liu, Yongxin Liu et al. „Distant Domain Transfer Learning for Medical Imaging“. In: *IEEE journal of biomedical and health informatics* 25.10 (2021), Seiten 3784–3793. DOI: 10.1109/JBHI.2021.3051470.
- [201] Yanshan Xiao, Fei Liang und Bo Liu. „A Transfer Learning-Based Multi-Instance Learning Method With Weak Labels“. In: *IEEE transactions on cybernetics* 52.1 (2022), Seiten 287–300. DOI: 10.1109/TCYB.2020.2973450.
- [202] Isidro Cortes-Ciriano und Andreas Bender. „Improved Chemical Structure-Activity Modeling Through Data Augmentation“. In: *Journal of chemical information and modeling* 55.12 (2015), Seiten 2682–2692. DOI: 10.1021/acs.jcim.5b00570.
- [203] Shih-Gu Huang, Moo K. Chung und Anqi Qiu. „Fast mesh data augmentation via Chebyshev polynomial of spectral filtering“. In: *Neural networks : the official journal of the International Neural Network Society* 143 (2021), Seiten 198–208. DOI: 10.1016/j.neunet.2021.05.025.
- [204] Adan Lin, Junhao Wu und Xuan Yang. „A data augmentation approach to train fully convolutional networks for left ventricle segmentation“. In: *Magnetic resonance imaging* 66 (2020), Seiten 152–164. DOI: 10.1016/j.mri.2019.08.004.
- [205] Yong Won Jin, Shuo Jia, Ahmed Bilal Ashraf et al. „Integrative Data Augmentation with U-Net Segmentation Masks Improves Detection of Lymph Node Metastases in Breast Cancer Patients“. In: *Cancers* 12.10 (2020). ISSN: 2072-6694. DOI: 10.3390/cancers12102934.
- [206] Yun Luo, Li-Zhen Zhu, Zi-Yu Wan et al. „Data augmentation for enhancing EEG-based emotion recognition with deep generative models“. In: *Journal of neural engineering* 17.5 (2020), Seite 056021. DOI: 10.1088/1741-2552/abb580.
- [207] Connor Shorten, Taghi M. Khoshgoftaar und Borko Furht. „Text Data Augmentation for Deep Learning“. In: *Journal of big data* 8.1 (2021), Seite 101. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00492-0.
- [208] Yulin Wang, Gao Huang, Shiji Song et al. „Regularizing Deep Networks with Semantic Data Augmentation“. In: *IEEE transactions on pattern analysis and machine intelligence* PP (2021). DOI: 10.1109/TPAMI.2021.3052951.