

Aus der Klinik für Kardiologie, Pneumologie und Angiologie
der Heinrich-Heine-Universität Düsseldorf

Direktor: Univ.-Prof. Dr. med. Malte Kelm

**Effizientes Screening auf
Aortenklappenstenosen mittels verständlicher
Künstlicher Intelligenz**

Dissertation

zur Erlangung des Grades eines Doktors der Medizin

der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Benita Isabel Kohlmann

2026

Als Inauguraldissertation gedruckt mit Genehmigung der
Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:

Dekan: Prof. Dr. med. Nikolaj Klöcker

Erstgutachterin: Prof. Dr. med. Hisaki Makimoto

Zweitgutachter: PD Dr. rer. nat. Sebastian Temme

Für meine Familie

Teile dieser Arbeit wurden veröffentlicht:

Makimoto H, Shiraga T, Kohlmann B, Magnisali CE, Gerguri S, Motoyama N, Clasen L, Bejinariu A, Klein K, Makimoto A, Jung C, Westenfeld R, Zeus T, Kelm M. Efficient screening for severe aortic valve stenosis using understandable artificial intelligence: a prospective diagnostic accuracy study. Eur Heart J Digit Health. 2022 May 16;3(2):141-152. doi: 10.1093/ehjdh/ztac029. PMID: 36713014; PMCID: PMC9707975.

Zusammenfassung

Die Bevölkerung wird durch den demographischen Wandel zunehmend älter, wodurch die Prävalenz von Herzklappenvitien ansteigt. Vor allem die Aortenklappenstenose als häufigstes behandlungsbedürftiges Vitium steht dabei im Fokus. Zeitgleich wird momentan kaum ein anderes Thema so kontrovers diskutiert wie die Chancen und Risiken der Künstlichen Intelligenz (KI). Dies betrifft insbesondere den medizinischen Bereich.

Ziel dieser Arbeit war es daher, eine transparente und nachvollziehbare Künstliche Intelligenz zu entwickeln, die anhand von digital aufgenommenen Auskultationsdaten den Schweregrad einer Aortenklappenstenose valide erkennen kann. Die Ergebnisse einer ergänzenden transthorakalen Echokardiographie dienten dabei als Basis für das Lernen sowie als Erfolgskontrolle der KI. Es wurden verschiedene *convolutional neural networks* (CNN) entwickelt, welche mit den Daten trainierten. Die drei Lokalisationen für die Auskultation waren dabei der zweite Interkostalraum rechts parasternal (2RSB), der dritte Interkostalraum links parasternal (Erbscher Punkt) sowie der fünfte Interkostalraum links medioclavicular (Apex).

Initial wurden 556 Datensätze in die Studie eingeschlossen. Für jeden der Auskultationsorte wurde ein separates Modell erstellt (L1-L3). Hier zeigte das Modell L1 mit den Auskultationsdaten vom zweiten Interkostalraum rechts (2RSB) die signifikant besten Ergebnisse mit einer Treffsicherheit von 93,4% sowie einem F1 Wert von 0,92 (p-Wert <0,001). Zur weiteren Leistungssteigerung wurden die Einzelmodelle in unterschiedlichen Gewichtungen kombiniert, sodass die Modelle E1-E3 entstanden. Hier konnte wiederum eine deutliche Überlegenheit des E1 Modells gezeigt werden. Zudem kann das Modell zuverlässig die verschiedenen Schweregrade der Aortenklappenstenose unterscheiden. Mit zusätzlichen Daten für eine klinische Validierung (n=132) konnte in Hinblick auf die schwere Aortenklappenstenose eine Sensitivität von 97,6% (41/42), eine Spezifität von 94,4% (85/90) sowie insgesamt eine Treffsicherheit von 95,7% (126/132) erreicht werden.

Abschließend wurde durch die Visualisierung der Entscheidungsfindung der Künstlichen Intelligenz diese nachvollziehbar und transparent gemacht. Dies geschah durch *Heatmaps*, welche mit Mel-Spektrogrammen verknüpft wurden. Dadurch konnte nachgewiesen werden, dass sich die KI bei der korrekten Beurteilung insbesondere der schweren Aortenklappenstenose auf die systolischen Phase der Herzaktion fokussiert, wie es bei der Auskultation der Aortenklappenstenose zu erwarten ist.

Zusammenfassend lässt sich festhalten, dass diese Künstliche Intelligenz in der Lage ist, die Schweregrade einer Aortenklappenstenose differenziert zu erkennen und dabei durch die Fokussierungsanalyse nachvollziehbare Ergebnisse liefert. Dies kann in Zukunft beispielsweise als Screeningtool verwendet werden.

Summary

The population is becoming increasingly older due to demographic change, which is increasing the prevalence of valvular heart disease. The focus is particularly on aortic valve stenosis, which is the most common type of valve requiring treatment. At the same time, hardly any other topic is currently being discussed as controversially as the opportunities and risks of artificial intelligence (AI). This applies in particular to the medical field.

The aim of this work was therefore to develop a transparent and comprehensible artificial intelligence that can validly recognize the severity of aortic valve stenosis based on digitally recorded auscultation data. The results of a complementary transthoracic echocardiography served as the basis for learning and as a success control for the AI. Various convolutional neural networks (CNN) were developed and trained with the data. The three localizations for auscultation were the second intercostal space on the right parasternal (2RSB), the third intercostal space on the left parasternal (Erb's point) and the fifth intercostal space on the left medioclavicular (apex).

Initially, 556 data sets were included in the study. A separate model was created for each of the auscultation locations (L1-L3). Here, model L1 with the auscultation data from the second right intercostal space (2RSB) showed the significantly best results with an accuracy of 93.4% and an F1 value of 0.92 (p-value <0.001). To further increase performance, the individual models were combined in different weightings to create the E1-E3 models. Here again, a clear superiority of the E1 model was demonstrated. In addition, the model can reliably differentiate between the various degrees of severity of aortic valve stenosis. With additional data for clinical validation (n=132), a sensitivity of 97.6% (41/42), a specificity of 94.4% (85/90) and an overall accuracy of 95.7% (126/132) were achieved regarding severe aortic valve stenosis.

Finally, the decision-making process of the artificial intelligence was visualized to make it comprehensible and transparent. This was done using heat maps, which were linked to Mel spectrograms. This made it possible to demonstrate that the AI focuses on the systolic phase of the cardiac action when correctly assessing severe aortic valve stenosis, as is to be expected when auscultating aortic valve stenosis.

In summary, it can be said that this artificial intelligence is able to differentiate between the degrees of severity of aortic valve stenosis and provide comprehensible results through focusing analysis. This can be used as a screening tool in the future, for example.

Abkürzungsverzeichnis

Abb	Abbildung
ACE	<i>Angiotensin-converting</i> Enzym
aHT	arterielle Hypertonie
ARB	Angiotensin II Rezeptorblocker
ARNI	Angiotensin-Rezeptor-Neprilysin-Inhibitor
AS	Aortenklappenstenose
AT1	Angiotensin 1
AUC	<i>area under the curve</i>
AV	<i>aortic valve</i>
BN	<i>Batch</i> -Normalisierung
BSA	<i>body surface area</i>
CABG	<i>coronary artery bypass graft</i>
CNN	<i>convolutional neural network</i>
Conv	<i>convolutional</i>
CRT	<i>cardiac resynchronization therapy</i>
CW-Doppler	<i>Continuous-Wave-Doppler</i>
EKG	Elektrokardiogramm
Grad-CAM++	<i>Gradient-Weighted Class Activation Mapping++</i>
Hz	Hertz
ICD	Implantierbarer Cardioverter-Defibrillator
IVSd	interventrikuläre Septumdicke (diastolisch)
KHK	Koronare Herzerkrankung
KI/AI	Künstliche Intelligenz / <i>artificial intelligence</i>
KÖF	Klappenöffnungsfläche
LA	linkes Atrium

LAD	linksatrialer Durchmesser
LAHB	linksanteriorer Hemiblock
LBBB	<i>left bundle branch block</i>
LPHB	linksposteriorer Hemiblock
LVDs	linksventrikulärer systolischer Durchmesser
LVEF	linksventrikuläre Ejektionsfraktion
LVOT	<i>left ventricular outflow tract</i>
LVPWd	linksventrikuläre posteriore Wanddicke (diastolisch)
MI	Mitralklappeninsuffizienz
NPV	negativ prädiktiver Wert
NYHA	<i>New York Heart Association</i>
PCI	perkutane Koronarintervention
PCSK9	Proteinkonvertase Subtilisin Kexin Typ 9
PPV	positiv prädiktiver Wert
P non sAS	Wahrscheinlichkeit einer nicht-schweren AS
P sAS	Wahrscheinlichkeit einer schweren AS
RBBB	<i>right bundle branch block</i>
ReLu	<i>Rectified Linear Unit</i>
RV	rechter Ventrikel
sec	Sekunden
TAVI	<i>transcatheter aortic valve implantation</i>
TIA	transitorische ischämische Attacke
TTE	transthorakale Echokardiographie
VHF	Vorhofflimmern
Vmax	maximale transvalvuläre Flussgeschwindigkeit
2RSB	zweiter Interkostalraum parasternal rechts

Inhaltsverzeichnis

1 Einleitung.....	1
1.1 Aortenklappenstenose.....	1
1.1.1 Übersicht und Epidemiologie der Aortenklappenstenose	1
1.1.2 Ätiologie	2
1.1.3 Pathophysiologie.....	3
1.1.4 Symptomatik	3
1.1.5 Diagnostik	4
1.1.6 Therapie	6
1.1.6.1 Medikamentöse Therapie.....	6
1.1.6.2 Interventionelle versus operative Therapie	7
1.1.7 Prognose.....	8
1.2 Künstliche Intelligenz.....	9
1.2.1 Einleitung und Funktionsweise.....	9
1.2.2 <i>Deep Learning</i> und <i>convolutional neural network</i>	10
1.2.3 Beispiele von Anwendungen in der Medizin	11
1.2.4 Bedeutung in der Kardiologie	12
1.3 Ziele der Arbeit/Fragestellung	13
2 Material und Methoden.....	15
2.1 Aufnahme von Patienten in die Datenbank	15
2.1.1 Einschlusskriterien	15
2.1.2 Ausschlusskriterien	15
2.1.3 Ethikvotum	16
2.2 Datenerhebung.....	16
2.2.1 Auskultation und EKG.....	16
2.2.2 Echokardiographie.....	17
2.2.3 Anamnese und Erhebung klinischer Daten.....	17
2.3 Patientencharakteristika.....	17
2.3.1 Trainings-/Validierungs-/Testgruppe	17
2.3.2 Klinische Validierung	21
2.4 Datenvorbereitung und Vorverarbeitung	24
2.4.1 Auskultationsdaten.....	24
2.4.1.1 Kurzzeit-Fourier-Transformation.....	24
2.4.1.2 Erstellung der Mel-Spektrogramme	25
2.4.2 Aufbereitung der Echokardiographiedaten mittels Kreuzvalidierung.....	26
2.5 Architektur und Struktur der KI.....	29
2.5.1 Einfaches <i>convolutional neural network</i>	29
2.5.2 Kombiniertes <i>convolutional neural network</i>	30
2.6 Training, Validierung und Testung der KI	31
2.6.1 Training der KI an den einzelnen Auskultationsorten.....	31
2.6.2 Zusammengesetzte Modelle	31
2.6.3 Parameter für die Überprüfung der Leistung der Modelle	32

2.6.4 Einstellung der Hyperparameter während des Trainings	32
2.6.5 Validierung im klinischen Kontext	33
2.7 Export auf eine Smartphone App	34
2.8 Visualisierung des Entscheidungsprozesses mittels Aktivierungskarten	34
2.9 Statistische Analysen	35
3 Ergebnisse.....	36
3.1 Leistung der Künstlichen Intelligenz im Test	36
3.1.1 2RSB	36
3.1.2 Erb-Punkt	38
3.1.3 Apex	40
3.1.4 Ergebnisse der zusammengesetzten Modelle	46
3.2 Klinische Validierung	49
3.2.1 Vergleich der Modelle	49
3.2.2 Ergebnisse des E1 Modells.....	50
3.3 Export auf eine Smartphone App	51
3.4 Visualisierung des Beurteilungsprozesses der KI.....	52
3.4.1 2RSB	53
3.4.2 Erb-Punkt	58
3.4.3 Apex	62
4 Diskussion	66
4.1 Ergebnisse dieser Studie.....	66
4.2 Kritische Evaluation der Methoden	69
4.3 Kritische Evaluation der Ergebnisse.....	72
4.4 Klinische Bedeutung dieser Studie und zukünftige Entwicklungen	74
4.5 Limitationen	76
5 Zusammenfassung und Schlussfolgerung.....	78
6 Literatur- und Quellenverzeichnis.....	80
7 Anhang.....	90
8 Tabellen- und Abbildungsverzeichnis	94

1 Einleitung

1.1 Aortenklappenstenose

1.1.1 Übersicht und Epidemiologie der Aortenklappenstenose

Die Aortenklappe ist eine der vier Herzklappen, sie ist im linken Herzen zwischen dem Ventrikel und der Aorta lokalisiert. Während der systolischen Herzaktion wird das Blut durch die Kontraktion des linken Ventrikels über die geöffnete Klappe in die Aorta und damit in den großen Körperkreislauf zur Versorgung der Peripherie mit Sauerstoff gepumpt. Durch den Klappenmechanismus verhindert sie mit ihrem Schluss während der Diastole den retrograden Fluss von Blut aus der Aorta ascendens zurück in den linken Ventrikel.

Mit zunehmendem Lebensalter steigt das Risiko für die Entwicklung einer Herzklappenerkrankung jedoch signifikant an (Bardon und Garelnabi, 2020). Epidemiologische Daten legen nahe, dass etwa 10-15% aller Personen über 75 Jahren an einem relevanten Herzklappenvitium leiden (Nkomo et al., 2006). Hiervon ist insbesondere die Aortenklappe betroffen. Die Aortenklappenstenose (AS) zählt zu den am häufigsten auftretenden Klappenfehlern weltweit (Kanwar et al., 2018; Flachskampf und Daniel, 2004). In den europäischen Ländern macht davon die Aortenklappenstenose mehr als 40% aus (Iung et al., 2007) und ist zudem das häufigste behandlungsbedürftige Vitium (Grimard et al., 2016). Neben der arteriellen Hypertonie (aHT) und der koronaren Herzerkrankung (KHK) gehört die Aortenklappenstenose sogar zu den drei häufigsten kardiovaskulären Erkrankungen (Lindman et al., 2016). Frauen und Männer sind in etwa gleich häufig betroffen (Billig et al., 2024).

Die Aortenklappenstenose ist dabei definiert als eine verringerte Klappenöffnungsfläche mit resultierend erschwerter Passage des Blutstroms aufgrund verschiedener nachfolgend dargestellter Pathomechanismen.

Durch den demographischen Wandel und die stetig älter werdende Bevölkerung wird ein Anstieg der Prävalenz der Aortenklappenstenose in den kommenden Jahren erwartet (Kanwar et al., 2018).

Gerade unter den älteren Patienten über 65 Jahre lassen sich zudem viele Betroffene mit noch unentdeckten Herzklappenvitien finden (d'Arcy et al., 2016). Da die Prognose einer höhergradigen Stenose mit einer deutlich reduzierten Lebenserwartung einher geht, sind eine frühzeitige Diagnose und damit ein rechtzeitiger Therapiebeginn entscheidend für das Überleben der Betroffenen (Henning, 2021; Vahanian et al., 2021).

1.1.2 Ätiologie

Die physiologische Öffnungsfläche der Aortenklappe beträgt bei Erwachsenen circa 2,6-4 Quadratzentimeter, diese nimmt jedoch bei Stenosierung der Aortenklappe deutlich ab. Parallel dazu nehmen die Strömungsgeschwindigkeiten des Blutes bei abnehmender Öffnungsfläche meist zu (Kanwar et al., 2018).

In den industrialisierten Ländern sind hauptsächlich zwei Pathomechanismen an der Genese der Aortenklappenstenose beteiligt. Die häufigste Form ist die degenerativ-arteriosklerotische Stenose, von welcher vorwiegend ältere Menschen betroffen sind. Dabei steht primär die progrediente Verkalkung der Klappe im Vordergrund, was die Beweglichkeit der Klappe herabsetzt (Thaden et al., 2014). Es wird dabei ein Zusammenspiel von Lipidablagerungen, einer chronischen Inflammation sowie einer Kalzifikation der Herzklappe ähnlich zu den Pathomechanismen bei der Arteriosklerose angenommen (Olszowska, 2011). Hier sind neben dem Alter insbesondere die Risikofaktoren des metabolischen Syndroms sowie der Nikotinkonsum als Ursachen zu nennen.

Neben den erworbenen Ursachen kann auch ein angeborener anatomischer Herzfehler, zum Beispiel eine bikuspid angelegte Aortenklappe, durch die vermehrte mechanische Belastung zu einem erhöhten Risiko für eine Stenosierung beitragen (Pedriali et al., 2020). Bis zu 2% der Bevölkerung sind von einer bikuspiden Aortenklappe betroffen (Chambers, 2005). Hier treten bereits in einem jüngeren Lebensalter häufig relevante Stenosen auf.

1.1.3 Pathophysiologie

Oftmals kann das kardiovaskuläre System die verminderte Öffnung der Aortenklappe noch über einen längeren Zeitraum gut kompensieren, sodass die Betroffenen wenig bis keine Beschwerden haben (Vahanian et al., 2012). Doch sobald progrediente Symptome auftreten, ist der Verlauf der Erkrankung meist rasch fortschreitend (Joseph et al., 2017).

Es kommt durch die Verengung der Klappe zu einem reduzierten Auswurf von Blut aus dem linken Ventrikel in die Aorta, daher muss der Ventrikel kompensatorisch einen höheren Druck generieren, damit eine adäquate Blutzirkulation weiterhin gewährleistet bleibt. Aus diesem Mechanismus resultiert konsekutiv eine konzentrische Linksherzhypertrophie durch Nachlasterhöhung, die sowohl mit einer systolischen als auch einer diastolischen Dysfunktion einher gehen kann (Grimard et al., 2016). Daher besteht bei fortgeschrittenen Befunden ebenso das Risiko der Entwicklung einer Herzinsuffizienz (Flachskampf und Daniel, 2004). Es kommt zudem zu einer eingeschränkten myokardialen Compliance der Ventrikel und zu einer reduzierten Perfusionsreserve bei Belastung, sodass ebenfalls ein Ischämierisiko besteht (Billig et al., 2024).

1.1.4 Symptomatik

Die klinischen Symptome einer Aortenklappenstenose können sehr heterogen sein. Viele Patienten mit progredienten Beschwerden sehen aufgrund dieser von einer körperlichen Betätigung ab und bemerken somit den Fortschritt ihrer Erkrankung nicht unmittelbar.

Patienten mit leicht- und mittelgradiger Aortenklappenstenose können die reduzierte Auswurfleistung durch die verengte Klappenöffnungsfläche oft, wie bereits erwähnt, durch einen kompensatorisch erhöhten Druck im linken Ventrikel noch ausgleichen (Vahanian et al., 2012). Es entsteht eine Latenzzeit, bevor die Symptomatik einsetzt. Dies geschieht häufig erst, wenn die Klappenöffnungsfläche schon weniger als ein Quadratzentimeter beträgt und somit als hochgradig einzustufen ist (Chambers, 2005). Hieraus resultiert die schlechte

Überlebensprognose mit einer Einjahressterblichkeit von ca. 30% (Strange et al., 2019).

Typisch für die fortgeschrittene Stenose ist eine Trias aus drei Kardinalsymptomen. Fast 80% der Patienten entwickeln als erstes Symptom eine progrediente Dyspnoe, insbesondere bei Belastung (Ristić-Andelkov et al., 2003). Bis zu 60% der Patienten geben im Verlauf pectanginöse Beschwerden an (Christen et al., 2006). Zudem treten kardiale Synkopen durch eine unzureichende zerebrale Blutversorgung auf (Flachskampf und Daniel, 2004). Sie lassen sich bei mehr als einem Drittel der Patienten finden (Ristić-Andelkov et al., 2003). All diese Symptome verstärken sich typischerweise bei Belastung. Weiterhin kann es durch die vermehrte Herzarbeit zu einer Linksherzhypertrophie mit nachfolgender Herzinsuffizienz kommen (Aronow, 2007). Dies betrifft vor allem die diastolische Dysfunktion in erheblichem Maße. Ein besonders gefürchtetes Risiko stellt der plötzliche Herztod dar, der vor allem bei symptomatischen Patienten gehäuft auftritt (Vahanian et al., 2012).

Doch selbst bei schon weit fortgeschrittenen Aortenklappenstenosen finden sich bis zu 20% asymptomatische Patienten (Ristić-Andelkov et al., 2003). Insgesamt lässt sich somit festhalten, dass die Klinik, die die Patienten bieten, individuell sehr divers ist. Dies verdeutlicht insbesondere den Stellenwert von regelmäßigen kardiologischen Verlaufskontrollen.

1.1.5 Diagnostik

Zu Beginn jeder Diagnostik steht neben der sorgfältigen Anamnese die ausführliche körperliche Untersuchung. Hierzu gehört die Auskultation des Herzens. Hier lässt sich bei vielen Betroffenen typischerweise ein charakteristisches spindelförmiges systolisches Herzgeräusch mit Punctum maximum im zweiten Intercostalraum rechts parasternal auskultieren, welches teils in die Karotiden fortgeleitet wird (Flachskampf und Daniel, 2004). Weiterhin sollte im Rahmen der Basisdiagnostik ein Elektrokardiogramm (EKG) geschrieben werden, hier lassen sich vielfach unspezifische Zeichen einer linksventrikulären Hypertrophie wie ein überdrehter Linkslagetyp oder ein positiver Sokolow-Lyon-Index sowie T-Negativierungen in den linkspräkordialen Ableitungen nachweisen.

Den Goldstandard zur Diagnostik stellt die Echokardiographie dar (Baumgartner et al., 2017). Hierbei wird beispielsweise die Öffnungsfläche ausgemessen oder der Druckgradient über den Herzklappen ermittelt. Dies kann sowohl in der transthorakalen als auch in der transösophagealen Ultraschalluntersuchung geschehen (Beneduce et al., 2020). Neben der Klappenöffnungsfläche sind jedoch noch weitere Faktoren wichtig für die Quantifizierung der Schwere des Vitiums. Modifiziert nach Hagendorff et al. (2020) lässt sich der Schweregrad der Aortenklappenstenose mittels Echokardiographie anhand verschiedener Charakteristika wie in Tabelle 1 dargestellt einteilen.

Tabelle 1: Einteilung der Schweregrade der Aortenklappenstenose anhand echokardiographisch erhobener Parameter

Schweregrad der Aortenklappenstenose	KÖF in Quadrat-zentimeter	KÖF/BSA in Quadrat-zentimeter pro Quadrat-meter	Mittlerer Gradient im CW-Doppler in Millimeter Quecksilbersäule	Vmax in Metern pro Sekunde	Vmax LVOT/ Vmax AV
Leichte AS (Grad I)	>1,5	> 0,85	< 20	2,6-2,9	> 0,5
Moderate AS (Grad II)	1,0-1,5	0,6-0,85	20-40	3,0-4,0	0,25-0,5
Schwere AS (Grad III)	< 1,0	< 0,6	> 40	> 4	< 0,25

KÖF: Klappenöffnungsfläche; KÖF/BSA: Klappenöffnungsfläche/body surface area; CW-Doppler: Continuous-Wave-Doppler; Vmax: maximale transvalvuläre Flussgeschwindigkeit; LVOT: left ventricular outflow tract; AV: aortic valve; AS: Aortenklappenstenose

Tabelle modifiziert nach Hagendorff et al. (2020)

Bei einer *low-flow-low-gradient* Aortenklappenstenose ist das Herz durch eine begleitende Herzinsuffizienz nicht in der Lage ein adäquates Schlagvolumen zu generieren und es bestehen geringere Druckgradienten (Clavel et al., 2016). Dies kann die Diagnosestellung erschweren.

Ergänzend werden im Rahmen der Diagnostik fakultativ weitere Untersuchungen durchgeführt. Dazu gehören neben einer Blutuntersuchung beispielsweise auf Troponin oder das *B-type natriuretic peptide* als Prognosemarker die kardiale Bildgebung mittels Magnetresonanz- oder Computertomograph, die Aufnahme eines Röntgenbildes des Thorax zur Abschätzung der Herzgröße sowie einer pulmonalvenösen Stauung oder eine Herzkatheteruntersuchung. Diese kann wichtige Hinweise für die Einteilung des Schweregrades der Stenose liefern, da in der transthorakalen Echokardiographie oft die gemessenen Werte vor allem bei reduzierter Pumpfunktion unterschätzt werden, während in der Herzkatheteruntersuchung die Druckgradienten sehr valide direkt im linken Ventrikel als auch in der Aorta gemessen werden können (Otto et al., 2021). Zusätzlich lässt sich so begleitend eine relevante KHK bestätigen oder ausschließen, was im Hinblick auf die Therapieoptionen entscheidend sein kann.

1.1.6 Therapie

1.1.6.1 Medikamentöse Therapie

Aktuell steht keine medikamentöse Therapie zur Verfügung, welche den Progress der Stenosierung verlangsamen oder aufhalten könnte (Goody et al., 2020). Daher beschränken sich die Therapieoptionen bisher nur auf mögliche Begleiterkrankungen, aber nicht die Stenosierung selbst. Hier ist beispielsweise die Einstellung eines arteriellen Hypertonus oder die Sekundär- beziehungsweise Tertiärprävention bei Hyperlipidämie durch die Gabe eines Statins zu nennen (Otto et al., 2021; Marquis-Gravel et al., 2016). Wenn eine Herzinsuffizienz vorliegt, dann sollte diese ebenfalls medikamentös bestmöglich eingestellt werden (Otto et al., 2021). Studien wie von Leon et al. (2010) zeigen jedoch, dass die alleinige medikamentöse Therapie den progredienten Verlauf der Erkrankung nicht verhindern kann. Daher sollte bei der Diagnosestellung der hochgradigen Aortenklappenstenose immer eine Indikation zur chirurgischen oder interventionellen Therapie geprüft werden. Bei den anderen Schweregrade sind regelmäßige echokardiographische Verlaufskontrollen indiziert.

1.1.6.2 Interventionelle versus operative Therapie

Früher galt der biologische oder mechanische Aortenklappenersatz mittels Sternotomie und extrakorporaler Zirkulation als Verfahren der Wahl. Komorbiditäten wie eine stark verkalkte Aorta, eine eingeschränkte linksventrikuläre Ejektionsfraktion oder eine chronische Niereninsuffizienz steigern das perioperative Risiko deutlich. Daher ist der konventionelle Klappenersatz für bis zu ein Drittel der Patienten nicht indiziert. Es wird zudem von einer perioperativen Sterblichkeit von bis zu 3% berichtet (Rajput und Zeltser, 2023). Aufgrund dessen wurde diese Operation gerade bei den älteren und multimorbiden Patienten zunehmend durch den kathetergestützten perkutanen Aortenklappenersatz (*transcatheter aortic valve implantation*, TAVI) abgelöst (Widder und Bauersachs, 2014). Mittlerweile wird bei mittelgradigem und sogar niedrigem präoperativem Risiko der Patienten ein gleichwertiges Outcome der TAVI zum kardiochirurgischen Aortenklappenersatz erreicht (Van Mieghem et al., 2020; Siontis et al., 2019). Wie die Metaanalyse von Siontis et al. (2019) zudem zeigte, lag die Zweijahres-Gesamt mortalität der Patienten mit TAVI 17% unter der des konventionellen Aortenklappenersatzes. Wenn jedoch ein längerer Follow-up-Zeitraum erfasst wird, dann nivellieren sich die Mortalitätsraten aufgrund des meist älteren Patientenkollektivs bei der TAVI-Methode (Siontis et al., 2019).

Doch auch die TAVI birgt Risiken. Eine schwerwiegende Komplikation hierbei ist die Störung der Reizleitung der elektrischen Impulse insbesondere am Atrioventrikular-Knoten, welche bei höhergradigen Blockierungen eine dringliche Schrittmacherindikation darstellt (Mirolo et al., 2018). Bis zu ein Drittel der Patienten entwickelt nach dem Eingriff zusätzlich einen Linksschenkelblock (Mangieri et al., 2018). Ferner ist ein trans- oder paravalvuläres Leck eine typische Nebenwirkung, es kommt in bis zu 40% der Fälle zu einem unzulänglichen Schluss der Aortenklappe mit dem Risiko der Ausbildung einer mindestens leichtgradigen Aortenklappeninsuffizienz (Reardon et al., 2017, Kodali et al., 2012).

Ein weiteres Risiko nach einer TAVI-Implantation besteht im Auftreten eines Schlaganfalls. Metaanalysen wie die von Eggebrecht et al. (2012) legen nahe, dass circa 3,3% der Patienten in den ersten 30 Tagen nach dem Eingriff einen Schlaganfall erleiden.

1.1.7 Prognose

Insgesamt lässt sich festhalten, dass die hochgradige Aortenklappenstenose unbehandelt eine palliative Diagnose darstellt, da insbesondere bei den symptomatischen Patienten die Letalität sehr hoch ist. Hier versterben circa 50% der Betroffenen ohne adäquate Therapie innerhalb der ersten zwei Jahre nach Symptombeginn (Berglund et al., 2018; Leon et al., 2010). Oft ist diese Phase zudem von kardialen Dekompensationen, Dyspnoe und pectanginösen Beschwerden begleitet. Die Letalität liegt im Durchschnitt höher als bei vielen malignen Tumorerkrankungen. Nicht zu vernachlässigen sind dabei die zusätzlich auftretenden Komorbiditäten, welche Einfluss auf die Prognose nehmen. Hier ist als Beispiel die begleitende Herzinsuffizienz zu nennen, die bei bis zu einem Viertel der Patienten mit schwerer Aortenklappenstenose zeitgleich auftritt und so das Krankheitsbild oft verkompliziert (Kamperidis et al., 2016). Die fortgeschrittene Herzinsuffizienz hat ebenfalls eine sehr schlechte Überlebensrate. In Metaanalysen lag sie sowohl bei erhaltener als auch bei reduzierter linksventrikulärer Pumpfunktion im Mittel bei 2,1 Jahren nach Diagnosestellung (Shah et al., 2017). Auch andere Erkrankungen wie ein Schlaganfall oder eine KHK sind mit einer Aortenklappenstenose vergesellschaftet (Di Minno et al., 2018). Zudem zeigen Studien auch für die asymptomatischen Patienten mit hochgradiger Aortenklappenstenose einen Benefit bei einem frühzeitigen Therapiestart (Campo et al., 2019).

Außerdem lässt sich festhalten, dass bereits die moderate Aortenklappenstenose mit einem bis zu 2,6-fach erhöhten Sterberisiko gegenüber der Normalbevölkerung einher geht (Stassen et al., 2023).

Vor diesem Hintergrund ist eine zeitnahe Diagnose der Aortenklappenstenose, möglichst noch im Anfangsstadium, wichtig. So kann man rechtzeitig über eine Therapie entscheiden, um Komorbiditäten zu verhindern, die Lebensqualität zu erhalten und der palliativen Diagnose eine kurative Therapieoption wie beispielsweise die TAVI entgegenzustellen.

1.2 Künstliche Intelligenz

1.2.1 Einleitung und Funktionsweise

Die Künstliche Intelligenz (KI), im englischen Sprachraum als *artificial intelligence* (AI) bezeichnet, ist ein Teilbereich der Informatik. Mit Hilfe der Künstlichen Intelligenz sollen Computerprogramme „lernen zu denken“ und sich anhand neuer Informationen weiterentwickeln, in Analogie zum Lernen von Menschen (Mintz und Brodie, 2019). Bis heute ist eine genaue Definition des Begriffes in Fachkreisen umstritten.

Die Erstbeschreibung des Begriffes „Künstliche Intelligenz“ von John McCarthy schon im Jahr 1956 zeigt die lange Historie der Vorstellung von Menschen, dass sie Programme entwickeln können, die am Ende ihre eigene Intelligenz übertreffen (Amisha et al., 2019). Die Vorstellung der Visionäre war, dass die Computerprogramme differenziert Fragestellungen lösen und diese auf zukünftige Problemkomplexe anwenden (Larentzakis und Lygeros, 2021).

Schon 1950 gab es durch Alan Turing erste Gedankenexperimente zur Denkfähigkeit von Maschinen (Scherk et al., 2017). Jedoch verging seither viel Zeit, bis sich das Konzept durch stetigen technischen Fortschritt nach und nach durchsetzen konnte. Primär fehlte es an den grundsätzlichen Voraussetzungen in der Technik, sodass das Projekt nicht den durchschlagenden Erfolg hatte.

Fortschritte konnten erstmals Ende der 1990er Jahre verbucht werden, als es einer KI gelang den amtierenden Schachweltmeister im direkten Duell zu besiegen (Scherk et al., 2017). Weiterhin stellten jedoch die enormen Rechenleistungen der Computer ein großes Problem dar. Der Aufschwung ging dann Anfang des 21. Jahrhunderts weiter, als das *Deep Learning* erfunden wurde (Kaul et al., 2020).

Eine sehr beeindruckende Erfolgswelle erlebt die KI momentan. Fast täglich erscheinen neue Artikel zu Funktionen und Chancen, aber auch Risiken der Künstlichen Intelligenz. Hierfür sind neben der steigenden Rechenleistung der Computer die hohe Menge an Daten und die Verbesserung der Algorithmen verantwortlich (Scherk et al., 2017). Erst durch diesen Fortschritt gelingt aktuell wohl die endgültige Etablierung der KI.

Viele Bereiche des menschlichen Lebens sind mittlerweile schon durch die Entwicklung der Künstlichen Intelligenz revolutioniert worden. Beispielsweise funktionieren Spracherkennungen auf Smartphones durch die Benutzung, da die KI stetig neue Datensätze erhält und mit Hilfe dieser lernen kann (Mintz und Brodie, 2019). Auch die Entwicklung von selbstfahrenden Autos oder die Gesichtserkennung wäre ohne eine KI nicht möglich (Krittanawong et al., 2017; Chan et al., 2020).

Die Künstliche Intelligenz ist dabei nicht von sich aus intelligent, sondern lernt durch neue Datensätze. Dadurch ist sie einerseits limitiert durch den von Menschen programmierten Algorithmus sowie andererseits durch die Daten, die der KI zur Verfügung gestellt werden (Schütze und Schlieter, 2019). Daher kann die KI nur so gut werden, wie die Daten sind, mit denen sie lernt. Üblicherweise wird dabei der größere Teil der Datensätze für das Training und die Validierung verwendet, während ein kleinerer Teil für den anschließenden Testdurchlauf genutzt wird (Zaharchuk et al., 2018).

1.2.2 Deep Learning und convolutional neural network

Deep Learning bezeichnet eine Lernstrategie der Künstlichen Intelligenz und beruht dabei auf der Verwendung von neuronalen Netzwerken (*convolutional neural networks, CNN*). Es ist dabei als Teilbereich des maschinellen Lernens zu verstehen (Zaharchuk et al., 2018). Maschinelles Lernen hat die Besonderheit, dass das Programm selbstständig lernt und sich weiterentwickelt, ohne dass jeder Zwischenschritt anhand von Algorithmen vorgegeben werden muss.

Dabei funktioniert das neuronale Netzwerk in Analogie zum Gehirn eines Menschen, indem Verbindungen zwischen Neuronen dynamisch auf- und wieder abgebaut werden. Demnach gibt es je nach Anforderungen an das Netzwerk eine unterschiedlich große Anzahl an Neuronen. Das Netzwerk erlernt dann aus Datenmengen eigenständig Algorithmen zu erkennen, welche auf neue Daten angewendet werden können (LeCun et al., 2015).

Es gibt dabei eine Eingabe- und eine Ausgabeschicht, zwischen welche die Neuronen geschaltet und untereinander verknüpft werden (Zaharchuk et al., 2018).

Hierbei unterscheidet man unter anderem die *convolutional layers*, die *pooling layers* und die *fully-connected layers*, welche hintereinandergeschaltet werden und die Informationen auf verschiedene Art und Weisen weiterverarbeiten (Yasaka et al., 2018).

Die letzte Schicht wird als Ausgabeschicht bezeichnet und gibt eine Wahrscheinlichkeit für eine vorher definierte Fragestellung an. Dabei werden während des Trainingsprozesses die Gewichtungen zwischen den Neuronen so modifiziert, dass korrekte Verbindungen gestärkt und fehlerhafte verändert werden. Wichtig ist zudem, dass diese Verbindungen nicht durch Programmierungen verändert werden, sondern dass das Modell eigenständig lernt, Muster zu erkennen und seine Ergebnisse zu verbessern (Zaharchuk et al., 2018).

Dabei werden verschiedene Trainingsformen verwendet. Unterschieden wird vor allem zwischen dem überwachten und dem nicht-überwachten Lernen. Dies charakterisiert sich dadurch, dass die Ergebnisse beim überwachten Lernen bekannt sind und man sie mit den Trainingsresultaten der KI vergleicht, während die KI beim nicht-überwachten Lernen selbstständig Zusammenhänge und Lösungsstrategien erkennen muss (Choy et al., 2018).

1.2.3 Beispiele von Anwendungen in der Medizin

Die Künstliche Intelligenz ist seit einiger Zeit vermehrt in der Medizin präsent und wird in vielen Fachbereichen intensiv erforscht. Besonders in der Diagnostik ist es als Screeningtool anerkannt (Larentzakis und Lygeros, 2021).

Besonderen Stellenwert hat die KI bereits in der Radiologie. Hier findet bei der Auswertung von verschiedenen bildgebenden Verfahren mittlerweile eine Vielzahl von KI-gestützten Programmen Anwendung, die den Radiologen bei der Bewertung helfen (Hosny et al., 2018, Pianykh et al., 2020). Aufgrund der immer wiederkehrenden Muster, die sich je nach vorliegender Erkrankung bildmorphologisch charakteristisch präsentieren, ist dieser Teilbereich der Medizin für die Anwendung der Künstlichen Intelligenz geradezu prädestiniert.

Als andere Anwendungsgebiete in der Medizin sind exemplarisch die Anästhesie und die Gynäkologie genannt. In der Anästhesie ist es durch den Einsatz einer KI

gelingen, die Narkosetiefe sinnvoll zu steuern, was vorteilhaft sowohl für die Patienten als auch für die Ärzte ist (Hashimoto et al., 2020).

Im Bereich der Gynäkologie wird beispielsweise eine KI im Rahmen der Reproduktionsmedizin eingesetzt, um bei der *in-vitro* Fertilisation die Keimzellen herauszufiltern, welche ideal zusammenpassen (Wang et al., 2019).

Bisher wird die KI in vielen Gebieten noch in Studien untersucht, es ist aber absehbar, dass sich diese Technologie in der Medizin durch die weitreichenden Erfolge immer mehr durchsetzen und daher eine größere Bedeutung erlangen wird.

1.2.4 Bedeutung in der Kardiologie

In der Kardiologie wird die Verwendung von Künstlicher Intelligenz ebenfalls intensiv analysiert und kommt dabei zu guten Ergebnissen (Zippel-Schultz et al., 2021). Es wurde beispielsweise eine KI trainiert, die gelernt hat, wie sich ein akuter Myokardinfarkt im EKG charakterisiert. In anschließenden Test-EKGs konnte diese KI dann Myokardinfarkte zuverlässiger erkennen als erfahrene Kardiologen (Makimoto et al., 2020). Es gibt zudem bereits erste Untersuchungen im Bereich der Erkennung von Herzklappenvitien durch aufgezeichnete Herztöne mittels KI (Chorba et al., 2021; Maknickas und Maknickas, 2017; Huai et al., 2021).

1.3 Ziel der Arbeit/Fragestellung

Da die Menschen durchschnittlich immer älter werden, wird die Prävalenz der Herzklappenerkrankungen und damit auch die der Aortenklappenstenose steigen (Yi et al., 2021). Wie bereits beschrieben ist die schnelle und zuverlässige Detektion insbesondere von höhergradigen Stenosen von Bedeutung, da diese eine schlechte Prognose aufweisen und viele Komplikationen sowie einen Verlust von Lebensqualität verursachen können.

Aufgrund des epidemiologischen Wandels werden zukünftig Ärzte aller Fachrichtungen häufiger Patienten mit Herzklappenvitien und der Aortenklappenstenose im Speziellen betreuen. Dabei ist die Auskultation von Patienten eine der traditionellsten Diagnosemethoden der Medizin und Grundlage der Diagnostik der Klappenvitien.

Da die Digitalisierung auch vor der Medizin nicht Halt macht, ist der grundlegende Gedanke hinter der hier vorgestellten Untersuchung die Erstellung einer Künstlichen Intelligenz zur vereinfachten und zeitnahen Diagnostik von Aortenklappenstenosen mittels digital aufgezeichneter Auskultationsdaten. Dabei soll es zu einer Verknüpfung von altbekannten, gut erprobten Methoden wie der Herzauskultation mit neuen, stetig verbesserten Technologien wie der Künstlichen Intelligenz kommen.

Dies kann insbesondere dabei helfen, dass Ärzte anderer Fachabteilungen, mit weniger Erfahrung oder nicht-akademisches Personal im Rettungsdienst oder in Ambulanzen die Herzklappenvitien leichter und zuverlässiger detektieren können. Angesichts immer knapper werdender menschlicher Ressourcen und zunehmendem Zeitdruck kann die Entwicklung einer KI als effizientes Screeningtool wertvolle Zeit sparen. Durch die möglichst einfache Anwendbarkeit beispielsweise durch eine App-basierte Benutzung ist ebenfalls eine Verwendung im Rahmen einer digitalisierten Telemedizin anzustreben.

Das Ziel der Studie war die Entwicklung einer Künstlichen Intelligenz, die anhand von Auskultationsaufnahmen lernt, verschiedene Schweregrade der Aortenklappenstenose korrekt zuzuordnen. Dies sollte zudem unter der Koexistenz von anderen schwerwiegenden Herzklappenvitien wie der Mitralklappeninsuffizienz

möglich sein. Die Erfolgskontrolle geschah dabei durch eine ergänzende Echokardiographie aller Patienten, um den realen Schweregrad der Klappenvitien zu erfassen.

Zudem war ein weiteres Ziel, dass die KI in ihrer Entscheidungsfindung nachvollziehbar wird, man folglich anhand ihrer Fokussierung visualisieren kann, wieso die KI welche Entscheidung getroffen hat. Letztlich soll diese KI dann als App auf einem Smartphone funktionieren, sodass die Untersuchung von Patienten im klinischen Alltag vereinfacht wird.

2 Material und Methoden

2.1 Aufnahme von Patienten in die Datenbank

Im Rahmen dieser Untersuchung konnten grundsätzlich alle Patienten eingeschlossen werden, die eine vollständige transthorakale Echokardiographie (TTE) im Universitätsklinikum Düsseldorf erhalten haben und nicht unter die im Folgenden erläuterten Ausschlusskriterien fallen. Die Echokardiographie wurde dabei nach den aktuell gültigen Standards der Deutschen Gesellschaft für Kardiologie durch fachkundige Ärzte durchgeführt (Hagendorff et al., 2020).

2.1.1 Einschlusskriterien

Alle Patienten, die nicht eines der nachfolgenden Ausschlusskriterien aufwiesen und sich nach einer ausführlichen Aufklärung und Information über die Methoden und Ziele der Studie schriftlich dazu bereit erklärten, an der Untersuchung teilzunehmen, konnten eingeschlossen werden. Die Einwilligung wurde per Unterschrift bestätigt. Hierbei war ein Rückzug der Einwilligung jederzeit möglich. Es wurde dabei ein Mindestalter von 18 Jahren festgelegt, es gab keine Altershöchstgrenze.

2.1.2 Ausschlusskriterien

Aufgrund der möglichen Verfälschungen des aufgenommenen EKGs durch eine potenzielle Schrittmacherstimulierung und der Echokardiographie durch einliegende Schrittmachersonden wurden jegliche Patienten mit Herzschrittmachern oder implantierbaren Cardioverter-Defibrillatoren (ICD) nicht berücksichtigt. Einzig Eventrecorder waren erlaubt.

Zudem wurden keine Patienten nach einer Herzklappenoperation oder -intervention (zum Beispiel TAVI) berücksichtigt, da dies im Besonderen durch die veränderte Hämodynamik die Auskultationsergebnisse beeinflusst, weswegen es zu fehlerhaften Ergebnissen kommen konnte.

2.1.3 Ethikvotum

Die Studie wurde durch die medizinische Ethikkommission des Universitätsklinikums Düsseldorf genehmigt (Aktenzeichen 2019-763).

2.2 Datenerhebung

In der hier dargestellten Untersuchung wurden Daten von Patienten in den Jahren 2020 und 2021 erhoben, die aufgrund verschiedener Indikationen im stationären oder ambulanten Setting in der Klinik für Kardiologie, Pneumologie und Angiologie des Universitätsklinikums Düsseldorf eine transthorakale Echokardiographie durch geschultes ärztliches Personal erhalten haben. Dabei wurden die Resultate dieser Herzultraschalluntersuchungen als Grundlage und Ergebniskontrolle für das maschinelle Lernen der Künstlichen Intelligenz genutzt.

2.2.1 Auskultation und EKG

Nach ausführlicher Aufklärung und nachfolgender schriftlicher Einwilligung wurde bei allen eingeschlossenen Patienten ein 12-Kanal-EKG geschrieben. Zudem wurden die Herztöne durch ein digitales Stethoskop an jedem Auskultationsort 15 Sekunden lang abgehört, aufgezeichnet und anschließend gespeichert. Dafür wurde das EKO Duo System verwendet (Eko Devices, Oakland, Kalifornien).

Die Auskultationsorte waren dabei der zweite Intercostalraum rechts parasternal (2RSB), der dritte Intercostalraum links parasternal (Erb-Punkt) sowie der fünfte Intercostalraum links in der Medioclavicularlinie (Apex). Die Herztöne wurden im 4000-Hertz Waveformat aufgenommen. Dies ist ein gängiges Speicherformat für digitale Rohaudiodateien.

Die Aufzeichnung des EKGs und der Herztöne mit dem digitalen Stethoskop fand zeitnah vor oder nach der Echokardiographie (+/- ein Tag) statt.

2.2.2 Echokardiographie

Im Rahmen der vollständigen Echokardiographie wurden die wichtigsten Parameter mit in die Untersuchung aufgenommen. Hierunter fielen beispielsweise die Bestimmung der linksventrikulären Ejektionsfraktion (LVEF in %), die Messung des linksventrikulären systolischen Durchmessers (LVDs in Millimeter) oder die Beurteilung der Schweregrade der Herzklappenvitien vor allem der Mitral-, Trikuspidal- und Aortenklappe.

2.2.3 Anamnese und Erhebung klinischer Daten

Abschließend wurde eine ausführliche Anamnese erhoben, bei der beispielsweise die regelmäßige Medikamenteneinnahme oder die Vorerkrankungen abgefragt wurden (siehe Fragebogen im Anhang).

Zudem wurde die *New York Heart Association* (NYHA) Klassifikation abgefragt, welche unter anderem Rückschlüsse über die Dyspnoe der Patienten unter körperlicher Betätigung oder in Ruhe zulässt. Dadurch ergaben sich Hinweise auf die Belastbarkeit oder Einschränkungen der Betroffenen.

Im Rahmen der gesamten Untersuchungen wurden alle Patientendaten stets anonym und vertraulich behandelt und es war jederzeit möglich, die Einwilligung in die Studie zurückzuziehen.

2.3 Patientencharakteristika

2.3.1 Trainings-/Validierungs-/Testgruppe

Insgesamt wurden im Rahmen der Trainings-, Validierungs- und Testphase der Künstlichen Intelligenz 556 Datensätze verwendet. Hierbei lag das Durchschnittsalter der Patienten bei 68,9 Jahre (+/-14,8 Jahre). Es nahmen 251 Frauen (45,1%) und 305 Männer (54,9%) teil. Beeinträchtigt von ihrer Erkrankung waren zusammengekommen 326 der Betroffenen, was anhand der NYHA Klassifikation Grad II-IV deutlich wird (58,6%). Häufige Komorbiditäten waren ein arterieller Hypertonus (69,4%), ein Diabetes mellitus (23,7%) sowie eine Dyslipidämie (35,4%). Auch vorherige Erkrankungen wurden abgefragt, wie

beispielsweise der stattgehabte Schlaganfall beziehungsweise TIA (transitorische ischämische Attacke, 13%), der vorangegangene Myokardinfarkt (12,7%) oder eine kardiale Stentimplantation (31,8%) und Bypass-Operation (*coronary artery bypass graft*, CABG, 6,7%).

Zum Zeitpunkt der Aufnahme des EKGs und der Herztöne fand sich bei 101 Patienten ein Vorhofflimmern (18,2%). Die genauen Patientencharakteristika finden sich in Tabelle 2.

Tabelle 2: Charakteristika der Patienten der Trainings-, Validierungs- und Testgruppe

	Training-/Validierung-/Testdatensätze (n=556)
Alter (Jahre)	68,9 (+/-14,8)
Geschlecht	Frauen: 251 (45,1%), Männer: 305 (54,9%)
NYHA-Klassifikation	Keine: 112; I: 118; II: 154; III: 168; IV: 4
Arterielle Hypertonie	386 (69,4%)
Diabetes mellitus	132 (23,7%)
Dyslipidämie	197 (35,4%)
Vorheriger Schlaganfall/TIA	70 (13%)
Stattgehabter Myokardinfarkt	71 (12,7%)
Kardiale Stentimplantation	177 (31,8%)
CABG	37 (6,7%)
LVDs in Millimeter	37,4 (+/-10,1)
LVEF in %	57,9 (+/-10,7)
IVSd in Millimeter	11,1 (+/-2,9)
LVPWd in Millimeter	10,2 (+/-2,5)
LA Fläche in Quadratzentimeter	19,7 (+/-7,3)
Mitralklappeninsuffizienz	Keine: 211 (38%); I: 195 (35,1%); II: 96 (17,3%); III: 54 (9,6%)
Mitralklappenstenose	Keine: 508 (91,4%); I: 39 (7%); II: 9 (1,6%); III: 0 (0%)
Aortenklappeninsuffizienz	Keine: 361 (64,9%); I: 157 (28,2%); II: 36 (6,5%); III: 2 (0,4%)
Aortenklappenstenose	Keine: 340 (61,1%); I: 51 (9,2%); II: 51 (9,2%); III: 114 (20,5%)
Trikuspidalklappeninsuffizienz	Keine: 228 (41%); I: 243 (43,7%); II: 65 (11,7%); III: 20 (3,6%)
Trikuspidalklappenstenose	Keine: 556 (100%); I: 0 (0%); II: 0 (0%); III: 0 (0%)
VHF zum Zeitpunkt der Aufnahme	101 (18,2%)

Herzfrequenz (Schläge pro Minute)	72,7 (+/-13,9)
P-Wellen Dauer in Millisekunden	109,8 (+/-18,7)
PQ-Intervall in Millisekunden	171,3 (+/-37,9)
QRS-Dauer in Millisekunden	100,3 (+/-22,6)
QT-Intervall in Millisekunden	405 (+/- 37)
QRS-Achse (Grad)	32 (+/-52)
Einnahme von Antiarrhythmika	18 (3,2%)
Einnahme von Calcium-Antagonisten	142 (25,5%)
Einnahme von ACE-Hemmern	168 (30,2%)
Einnahme von AT1-Blockern	156 (28,1%)
Einnahme von Beta-Blockern	316 (56,8%)
Einnahme von Aldosteronantagonisten	67 (12,1%)
Einnahme von ARNI	16 (2,9%)
Einnahme von Statinen	296 (53,2%)
Einnahme von Alpha-Blockern	28 (5%)

NYHA: *New York Heart Association*; TIA: transitorische ischämische Attacke; CABG: *coronary artery bypass graft*; LVDs: linksventrikulärer systolischer Durchmesser; LVEF: linksventrikuläre Ejektionsfraktion; IVSd: interventrikuläre Septumdicke (diastolisch); LVPWd: linksventrikuläre posteriore Wanddicke (diastolisch) ; LA: linkes Atrium; VHF: Vorhofflimmern; ACE: *Angiotensin-Converting-Enzym*; AT1: Angiotensin 1; ARNI: Angiotensin-Rezeptor-Neprilysin-Inhibitoren
Tabelle modifiziert nach Makimoto et al. (2022)

Von den ursprünglich 886 eingeschlossenen Datensätzen wurden vor der Kreuzvalidierung wie nachfolgend genauer beschrieben insgesamt 330 ausgeschlossen, die keine Aortenklappenstenose enthielten. Bei diesen Datensätzen finden sich annähernd gleich viele Männer und Frauen. Die betroffenen Patienten waren im Durchschnitt etwas jünger (60,5 Jahre +/-16 Jahre) und weniger vorerkrankt als in der Trainings-, Validierungs- und Testgruppe. Zudem ist die NYHA-Klassifikation geringer. Nur knapp 45% der Patienten geben eine Beeinträchtigung der Leistungsfähigkeit an (NYHA Score Grad II-IV). Auch die linksventrikuläre Pumpfunktion ist geringfügig besser als bei den oben gezeigten Datensätzen (61,5% versus 57,9%). Zudem nahmen die Patienten durchschnittlich weniger Medikamente ein. Die genauen Ergebnisse lassen sich Tabelle 3 entnehmen.

Tabelle 3: Charakteristika der exkludierten Patienten

	Exkludierte Datensätze (n=330)
Alter (Jahre)	60,5 (+/-16)
Geschlecht	Frauen: 153 (46,4%), Männer: 177 (53,6%)
NYHA-Klassifikation	Keine: 132; I: 90; II: 74; III: 34; IV: 0
Arterielle Hypertonie	174 (52,7%)
Diabetes mellitus	47 (14,2%)
Dyslipidämie	90 (27,3%)
Vorheriger Schlaganfall/TIA	45 (13,6%)
Statthabter Myokardinfarkt	29 (8,8%)
Kardiale Stentimplantation	55 (16,6%)
CABG	5 (1,5%)
LVDs in Millimeter	34,3 (+/-9)
LVEF in %	61,5 (+/-7,8)
IVSd in Millimeter	10,3 (+/-2,6)
LVPWd in Millimeter	9,5 (+/-2,3)
LA Fläche in Quadratzentimeter	16,4 (+/-5,2)
Mitralklappeninsuffizienz	Keine: 210 (63,6%); I: 116 (35,2%); II: 4 (1,2%); III: 0 (0%)
Mitralklappenstenose	Keine: 323 (97,9%); I: 6 (1,8%); II: 0 (0%); III: 1 (0,3%)
Aortenklappeninsuffizienz	Keine: 270 (81,8%); I: 46 (13,9%); II: 14 (4,2%); III: 0 (0%)
Aortenklappenstenose	Keine: 330 (100%); I: 0 (0%); II: 0 (0%); III: 0 (0%)
Trikuspidalklappeninsuffizienz	Keine: 190 (57,6%); I: 124 (37,6%); II: 12 (3,6%); III: 4 (1,2%)
Trikuspidalklappenstenose	Keine: 329 (99,7%); I: 0 (0%); II: 0 (0%); III: 1 (0,3%)
VHF zum Zeitpunkt der Aufnahme	29 (8,8%)
Herzfrequenz (Schläge pro Minute)	72 (+/-13,7)
P-Wellen Dauer (in Millisekunden)	108,3 (+/-14,7)
PQ-Intervall in Millisekunden	162 (+/- 31,8)
QRS-Dauer in Millisekunden	94,8 (+/-19)
QT-Intervall in Millisekunden	398,2 (+/-39,8)
QRS-Achse (Grad)	40,2 (+/-44,6)
Einnahme von Antiarrhythmika	4 (1,2%)
Einnahme von Calcium-Antagonisten	75 (22,7%)

Einnahme von ACE Hemmern	65 (19,7%)
Einnahme von AT1 Blockern	72 (21,8%)
Einnahme von Beta-Blockern	123 (37,3%)
Einnahme von Aldosteronantagonisten	20 (6,1%)
Einnahme von ARNI	5 (1,5%)
Einnahme von Statinen	116 (35,1%)
Einnahme von Alpha-Blockern	13 (3,9%)

NYHA: *New York Heart Association*; TIA: transitorische ischämische Attacke; CABG: *coronary artery bypass graft*; LVDs: linksventrikulärer systolischer Durchmesser; LVEF: linksventrikuläre Ejektionsfraktion; IVSd: interventrikuläre Septumdicke (diastolisch); LVPWd: linksventrikuläre posteriore Wanddicke (diastolisch) ; LA: linkes Atrium; VHF: Vorhofflimmern; ACE: *Angiotensin-Converting-Enzym*; AT1: Angiotensin 1; ARNI: Angiotensin-Rezeptor-Neprilysin-Inhibitoren
Tabelle modifiziert nach Makimoto et al. (2022)

2.3.2 Klinische Validierung

Die Kohorte für die klinische Validierung bestand aus 132 Patienten, wovon 60 Frauen (45,5%) und 72 Männer (54,5%) waren. Das Durchschnittsalter lag bei 75 Jahren (+/-12,5 Jahre). Als hauptsächliche Komorbiditäten wurden bei diesen Patienten insbesondere die arterielle Hypertonie (81,1%), eine Dyslipidämie (46,2%) sowie ein Vorhofflimmern zum Zeitpunkt der Aufnahme in die Studie (31,1%) detektiert.

Die in der Echokardiographie erhobenen Daten zeigten eine durchschnittliche linksventrikuläre Pumpfunktion von 54,5% (+/- 13,1%). Die weiteren Messwerte und die Bestimmung der Schweregrade der Herzklappenvitien finden sich in Tabelle 4.

Tabelle 4: Charakteristika der Patienten der klinischen Validierung

	Klinische Validierung (n=132)
Alter (Jahre)	75,0 (+/-12,5)
Geschlecht	Frauen: 60 (45,5%), Männer: 72 (54,5%)
NYHA-Klassifikation	Keine: 10; I: 21; II: 36; III: 58; IV: 7
Arterielle Hypertonie	107 (81,1%)
Diabetes mellitus	33 (25%)
Dyslipidämie	61 (46,2%)
Vorheriger Schlaganfall/TIA	16 (12,1%)
Statthabter Myokardinfarkt	19 (16,5%)
Kardiale Stentimplantation	56 (42,4%)
CABG	11 (8,3%)
LVDs in Millimeter	40,2 (+/-10,5)
LVEF in %	54,5 (+/-13,1)
IVSd in Millimeter	11,6 (+/-3,4)
LVPWd in Millimeter	10,1 (+/-2,2)
LA Fläche in Quadratzentimeter	22 (+/-7,3)
Mitralklappeninsuffizienz	Keine: 31 (32,5%); I: 56 (42,4%); II: 27 (20,5%); III: 18 (13,6%)
Mitralklappenstenose	Keine: 117 (86,7%); I: 13 (9,8%); II: 2 (1,5%); III: 0 (0%)
Aortenklappeninsuffizienz	Keine: 63 (47,7%); I: 46 (34,9%); II: 23 (17,4%); III: 0 (0%)
Aortenklappenstenose	Keine: 34 (25,8%); I: 28 (21,2%); II: 28 (21,2%); III: 42 (31,8%)
Trikuspidalklappeninsuffizienz	Keine: 39 (29,5%); I: 77 (58,4%); II: 14 (10,6%); III: 2 (1,5%)
Trikuspidalklappenstenose	Keine: 132 (100%); I: 0 (0%); II: 0 (0%); III: 0 (0%)
VHF zum Zeitpunkt der Aufnahme	41 (31,1%)
Herzfrequenz (Schläge pro Minute)	74,4 (+/-17,3)
P-Wellen Dauer in Millisekunden	109,5 (+/-19,7)
PQ-Intervall in Millisekunden	181,2 (+/-39,3)
QRS-Dauer in Millisekunden	103 (+/-26)
QT-Intervall in Millisekunden	417 (+/-46)
QRS-Achse (Grad)	16 (+/-49)
Einnahme von Antiarrhythmika	7 (5,3%)

Einnahme von Calcium-Antagonisten	40 (30,3%)
Einnahme von ACE Hemmern	53 (40,2%)
Einnahme von AT1 Blockern	43 (32,6%)
Einnahme von Beta-Blockern	85 (64,4%)
Einnahme von Aldosteronantagonisten	30 (22,7%)
Einnahme von ARNI	4 (3%)
Einnahme von Statinen	82 (62,1%)
Einnahme von Alpha-Blockern	4 (3%)

NYHA: *New York Heart Association*; TIA: transitorische ischämische Attacke; CABG: *coronary artery bypass graft*; LVDs: linksventrikulärer systolischer Durchmesser; LVEF: linksventrikuläre Ejektionsfraktion; IVSd: interventrikuläre Septumdicke (diastolisch); LVPWd: linksventrikuläre posteriore Wanddicke (diastolisch) ; LA: linkes Atrium; VHF: Vorhofflimmern; ACE: *Angiotensin-Converting-Enzym*; AT1: Angiotensin 1; ARNI: Angiotensin-Rezeptor-Neprilysin-Inhibitoren
Tabelle modifiziert nach Makimoto et al. (2022)

2.4 Datenvorbereitung und Vorverarbeitung

2.4.1 Auskultationsdaten

Zur Vorverarbeitung der Auskultationsdaten wurden die aufgenommen Herztöne in vier Sekunden lange Abschnitte geteilt. Frühere Untersuchungen haben gezeigt, dass es bei zu langen Dateien vermehrt zu störenden Hintergrundgeräuschen kommen kann und andererseits zu kurze Dateien nicht genügend Erkennungspotential bieten (Zhou et al., 2021).

Zudem wurde ein *Oversampling* durchgeführt. Dies wird immer dann notwendig, wenn die Datensätze unausgewogen sind und somit bestimmte Charakteristika häufig auftreten und andere nur selten vorkommen. Die Künstliche Intelligenz kann daher Schwierigkeiten haben, diese seltenen Parameter korrekt zu detektieren (Gnip et al., 2021). Dies lässt sich auf diese Studie übertragen, da hier die Fälle ohne Klappenvitium häufiger vertreten sind. Darum wurde ein *Oversampling* durchgeführt, wobei die vorhandenen Datensätze erweitert wurden und insgesamt balancierter sind. Somit werden die Zuverlässigkeit und die Erkennungsfähigkeit der KI durch diese Methode besser.

2.4.1.1 Kurzzeit-Fourier-Transformation

Bei der Kurzzeit-Fourier-Transformation handelt es sich um eine mathematische Methode mit Hilfe derer das Frequenzspektrum eines akustischen Signals, in diesem Fall die Frequenzänderung der Herztöne, in Abhängigkeit von der Zeit dargestellt werden kann. Hierzu wurden die im Wave-Format gespeicherten Töne dementsprechend so aufbereitet, dass ihre Frequenzvariabilität sichtbar wurde. Dies ist notwendig, da die KI auf diese Weise für den Menschen nachvollziehbare Ergebnisse liefert, indem eine Analyse der Frequenzspektren durchgeführt werden kann.

2.4.1.2 Erstellung der Mel-Spektrogramme

Die Mel-Skala beschreibt den Zusammenhang zwischen der Frequenz eines Tones und der wahrgenommenen Tonhöhe. Der Mensch kann Unterschiede in der Tonhöhe nicht bei jeder Frequenz gleich gut wahrnehmen. Niedrigere Frequenzen werden deutlich besser differenziert als höhere.

Auf Grundlage dessen lässt sich anhand der komplexen Audiodateien ein Spektrogramm ableiten, welches auf der Mel-Skala beruht. Ein Spektrogramm verbildlicht dabei die Frequenzen des akustischen Signals in Abhängigkeit des zeitlichen Verlaufs.

Dafür wurden die vier Sekunden langen Datensätze mit der Kurzzeit-Fourier-Transformation als graphische Aufarbeitung der aufgenommenen Herztöne gegen die Zeit aufgetragen. Anschließend wurde eine Mel-skalierte Filterbank mit 128 Filtern angewendet. Basierend auf der Datenerweiterung mittels *Oversampling* wurden danach insgesamt 25344 Log-Mel-Spektrogramme aus den Aufzeichnungen erstellt, mit denen die KI trainiert werden konnte.

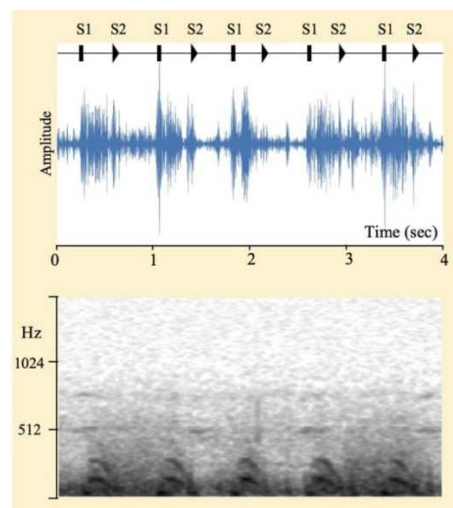


Abb. 1: Visualisierung der Tonspuren

Hz: Hertz; sec: Sekunden

Aufzeichnung der Tonspuren in der Waveform (oben) als visuelle Darstellung der Töne mittels Ausschlägen in der Amplitude sowie als Mel-Spektrogramm (unten) zur Demonstration der Frequenzintensität der aufgezeichneten Töne in Hertz aufgetragen jeweils über die Zeit in Sekunden; S1 und S2 jeweils als erster und zweiter Herzton

Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Abbildung (Abb.) 1 ist ein Beispiel für einen aufgenommen Datensatz (4 Sekunden lang), welcher einmal im oberen Teil der Abbildung in der Waveform und im unteren Teil umgewandelt im Mel-Spektrogramm visualisiert wurde.

2.4.2 Aufbereitung der Echokardiographiedaten mittels Kreuzvalidierung

Um die bestmöglichen Ergebnisse mit Hilfe der Künstlichen Intelligenz zu erreichen, gibt es verschiedene Methoden. In diesem Fall wurde ein stratifiziertes fünffaches Kreuzvalidierungsverfahren eingesetzt. Hintergrund dieser Methode ist es zu verhindern, dass sich die KI besonders auf ein bestimmtes Merkmal konzentriert, welches aber lediglich einen Ausreißer darstellt. Die Kreuzvalidierung schließt daher immer einen Teil der Datenmenge aus dem Trainingsprozess aus, um diesen für die Testung zu verwenden. Hierfür werden die Daten in Trainings-, Entwicklungs- und Testdatensätze aufgeteilt. Dieser Prozess wird mehrfach wiederholt, sodass am Ende ein Mittelwert der Ergebnisse gebildet werden kann.

Bei der klinischen Datenerhebung wurde anhand von Vorbefunden schon eine Vorauswahl an Patienten getroffen. Dies war notwendig, da es eine höhere Prävalenz von Menschen ohne Herzklappenerkrankungen gibt und so bevorzugt diejenigen eingeschlossen werden konnten, bei denen höhergradige Vitien vermutet wurden. Die Auskultatoren waren jedoch verblindet in Hinblick auf den Schweregrad.

Wie in Abbildung 2 gezeigt, wurden im ersten Schritt prospektiv 886 Patienten in die Studie eingeschlossen, von denen 114 eine schwere und 772 keine schwere Aortenklappenstenose aufwiesen. Darunter fielen sowohl Patienten ohne AS (n=670), mit einer leichten (n=51) oder einer moderaten (n=51) Aortenklappenstenose. Diese Befunde wurden auf Grundlage der Echokardiographie-Daten erhoben.

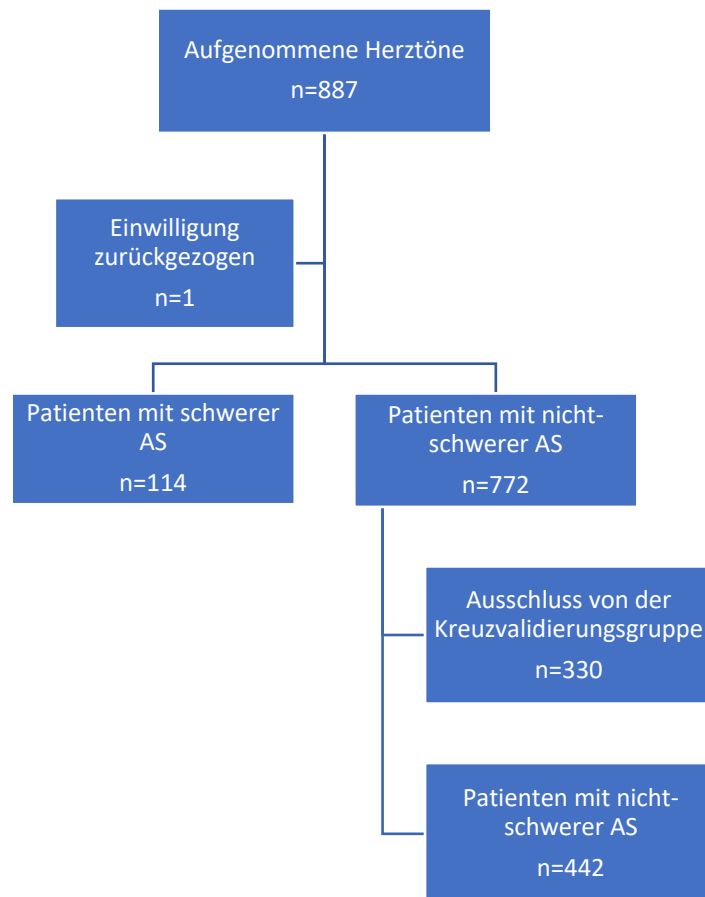


Abb. 2: Flowchart der rekrutierten Patienten

Von den 772 ursprünglich eingeschlossenen Patienten mit nicht-schwerer AS wurden insgesamt 330 Datensätze ohne nachweisbare Aortenklappenstenose von der Kreuzvalidierungskohorte ausgeschlossen.

Da simultan auftretende Herzklappenfehler wie vor allem die ebenfalls in der systolischen Phase auftretende Mitralklappeninsuffizienz die Erkennungsfähigkeit der Künstlichen Intelligenz in Bezug auf die Aortenklappenstenose beeinträchtigen können, sind zusätzlich zu den Daten der Aortenklappenstenose ebenso der Schweregrad der Mitralklappeninsuffizienz mit in das Training eingeflossen (siehe dazu Tabelle 11 im Anhang).

Das Ziel war es daher mit Hilfe der Stratifizierung die Zahl der Mitral- und Aortenklappenvitien in den verschiedenen Datensätzen auszugleichen, um eine Verzerrung zu vermeiden. Auf dieser Grundlage wurden dann Trainingsgruppen erstellt, ohne dass es in diesen Gruppen zu Überlappungen kam, während die Testfälle in allen Gruppen gleichblieben.

Bei den Datensätzen mit koexistenter hochgradiger Mitralklappeninsuffizienz wurden beispielsweise von 54 Datensätzen 16 als Testdatensätze verwendet. Begleitend kam dabei sowohl keine als auch eine leichte, moderate und schwere Aortenklappenstenose vor. Von den verbliebenen 38 Datensätzen wurden wiederum jeweils 16 Fälle für den Trainings- und 16 für den Entwicklungsdatsatz ausgewählt. Somit konnten Überschneidung der Fälle innerhalb einer Gruppe vermieden werden.

Bei den Auskultationsdaten mit koexistenter moderater, leichter oder nicht vorhandener Mitralklappeninsuffizienz war die Auswahl der Datensätze größer. Auch hier wurde wieder eine Kohorte für die Testung festgelegt. Aus den verbliebenen Fällen konnten ebenfalls Gruppen für die fünffache stratifizierte Kreuzvalidierung erstellt werden, ohne dass es zu einer Überschneidung der Fälle kam. Daher konnte das Training mithilfe der stratifizierten fünffachen Kreuzvalidierung dazu beitragen, dass die Resultate der KI am Ende verbessert werden konnten, indem eine Verzerrung vermieden wurde.

Das Training der KI erfolgte mit 442 Datensätzen mit nicht-schwerer und 114 mit schwerer Aortenklappenstenose. Diese wurden zu Trainings-, Validierungs- und Testzwecken verwendet. Insgesamt wurden dabei 100 Datensätze als Testdatensätze verwendet, jedes Trainingsset enthielt 356 Trainingsfälle sowie 100 Entwicklungsdatsätze. Hierbei gab es keine Überschneidungen zwischen den Gruppen. Die Testdaten waren in allen fünf Durchgängen gleich.

2.5 Architektur und Struktur der KI

Basierend auf Vorerfahrungen bei der Erstellung einer Künstlichen Intelligenz wurden zwei verschiedene *convolutional neural networks* (CNNs), im deutschen Sprachgebrauch als neuronale Faltungsnetzwerke bezeichnet, erstellt. Der Vorteil eines neuronalen Netzwerkes ist es, dass dieses Netzwerk während des Trainings eigenständig herausfiltert, welche Bereiche für die Lösung der Aufgabe wichtig sind, sich auf diese fokussiert und korrekte Verknüpfungen verstärkt.

Da der gesamte Vorgang der Entwicklung der Künstlichen Intelligenz sehr komplex ist und hauptsächlich in den informatischen Fachbereichen abgehandelt wird, soll dieses im Rahmen einer medizinischen Doktorarbeit nur grundsätzlich erklärt werden.

2.5.1 Einfaches *convolutional neural network*

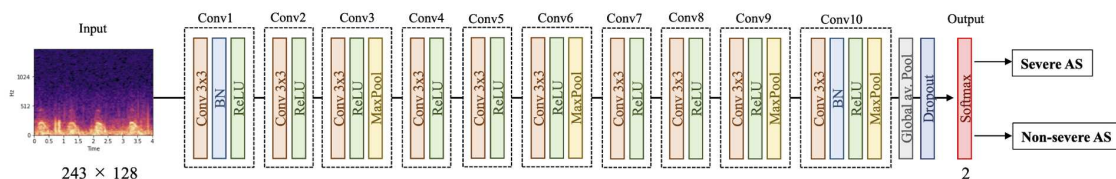


Abb. 3: Beispiel eines einfachen *convolutional neural network*

ReLu: *Rectified Linear Unit*, Conv: *Convolutional*, BN: *Batch-Normalisierung*, Hz: Hertz

Beispielhafter Aufbau eines einfachen *convolutional neural networks* vom *Input* der visualisierten Herzöne als Mel-Spektrogramm mit 128 Filtern über die zehn hintereinander gereihten Faltungsschichten bis hin zum *Output* und der Entscheidungsfindung, ob eine schwere oder nicht-schwere AS vorliegt

Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Wie Abbildung 3 zeigt, besteht dieses Modell aus zehn Faltungsschichten (Conv, *Convolutional*), wobei jeder Schicht eine Aktivierungsfunktion folgt (ReLU, *Rectified Linear Unit*). Zudem findet man in der ersten und der letzten Schicht eine *Batch-Normalisierung* (BN), im dritten, sechsten, neunten und zehnten Faltungsblock außerdem eine *Poolingschicht* (*MaxPool*). Am Ende des Modells wurde eine globale *Poolingschicht* zwischengeschaltet, ehe dann die letzte Schicht nach dem *Dropout* mit einer *Softmax-Funktion* aktiviert wurde. Hier entsteht die Entscheidung, ob eine schwere Aortenklappenstenose vorliegt oder nicht.

2.5.2 Kombiniertes *convolutional neural network*

Abbildung 4 zeigt ein Modell bestehend aus der Kombination von mehreren Kernlängen und -breiten. Es umfasst parallele Faltungsblöcke mit alternierenden Kernen, die zusammengefasst werden, bevor sich zwei weitere Faltungsschichten anschließen. Auch hier wird nach einer globalen *Poolingschicht* und einem *Dropout* durch die *Softmax-Funktion* die letzte Schicht aktiviert.

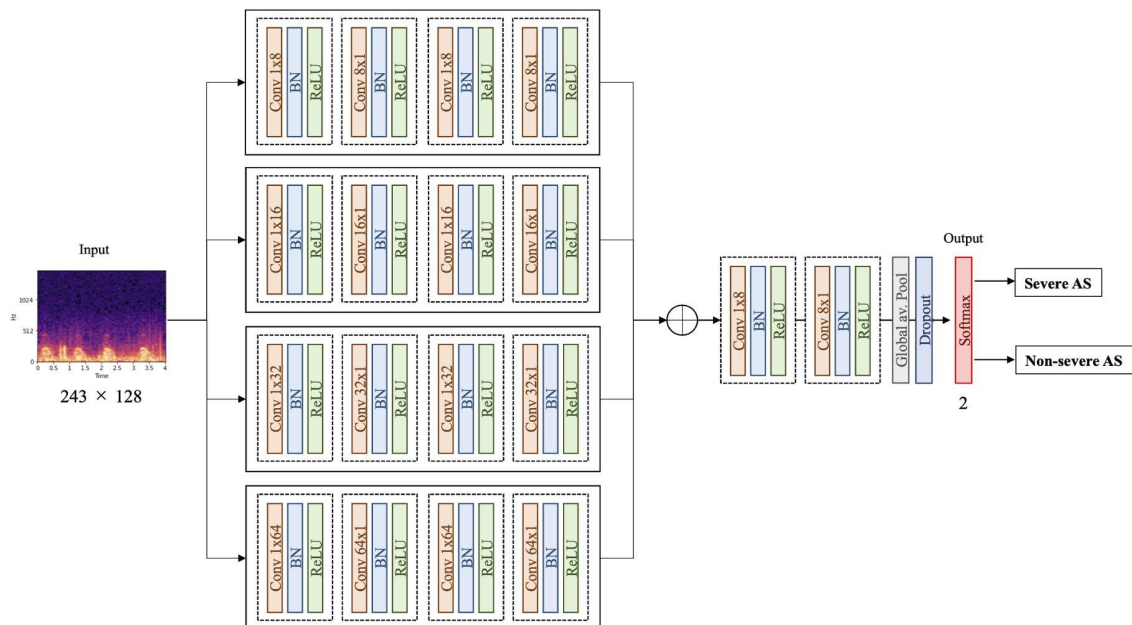


Abb. 4: Beispiel eines kombinierten *convolutional neural network*

ReLU: *Rectified Linear Unit*, Conv: *Convolutional*, BN: *Batch-Normalisierung*, Hz: Hertz

Beispielhafter Aufbau eines kombinierten *convolutional neural networks* vom Input der visualisierten Herztöne als Mel-Spektrogramm mit 128 Filtern über die parallelen Faltungsschichten bis hin zum Output und der Entscheidungsfindung über eine schwere oder nicht-schwere AS
Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Mit beiden Netzwerken lässt sich die Grundlage für das Künstliche Lernen erstellen, welche für die Detektion der schweren Aortenklappenstenose trainiert werden kann. Beide CNN besitzen am Ende eine Ausgabeschicht der Neuronen, die sogenannten *fully connected layer*, und verwenden eine sigmoidale Aktivierungsfunktion, welche Wahrscheinlichkeiten für die korrekte Detektion der Aortenklappenstenose berechnet. Hierfür sind die Programme Python 3.7 und TensorFlow 2.3 (Google, Mountain View, Kalifornien) verwendet worden.

2.6 Training, Validierung und Testung der KI

Nachdem die Künstliche Intelligenz mit dem neuronalen Netzwerk designt worden ist, muss sie anschließend mit Daten trainiert werden, damit sie sich in ihrer Funktion verbessert. Als Ergebniskontrolle werden die Daten aus der Echokardiographie verwendet. Durch die Rückmeldung, ob sie mit ihrer Einschätzung korrekt oder falsch lag, lernt die KI, welche Verbindungen zielführend sind und somit beibehalten und verstärkt werden müssen und welche verändert werden sollten. Dies ändert etwas an der Gewichtung der Verbindungen innerhalb des neuronalen Netzwerkes, gut vergleichbar mit dem dynamischen Auf- und Abbau von Neuronen im menschlichen Gehirn. Die Modelle wurden am Ende jedes Durchlaufs auf der Grundlage der Verlustfunktion und der Treffsicherheit validiert. Die Modelle mit dem niedrigsten Wert der Verlustfunktion wurden als die trainierten CNN-Modelle in jedem Kreuzvalidierungsdurchgang definiert.

2.6.1 Training der KI an den einzelnen Auskultationsorten

Die Modelle wurden sowohl mit Daten von allen drei Auskultationsorten als auch mit einem separaten Training nur mit Auskultationsergebnissen von einer Lokalisation trainiert. So wurde das beste Modell gefunden. Diese Modelle wurden als L1 (zweiter Intercostalraum parasternal rechts, 2RSB), L2 (dritter Intercostalraum parasternal links, Erb-Punkt) und L3 (fünfter Intercostalraum medioclavicular links, Apex) bezeichnet.

2.6.2 Zusammengesetzte Modelle

Aus den Modellen L1-3 wurden dann durch verschiedene Kombinationen zusammengesetzte Modelle entwickelt. Hierdurch entstand das Modell E1 in der gleichwertigen Gewichtung der Modelle L1-3 (im Verhältnis 1:1:1). Die Modelle E2 und E3 gewichteten vor allem L1 stärker (E2 6:2:2 und E3 8:1:1).

Bei den kombinierten Modellen wurde die Treffsicherheit mit den kompletten aufgenommenen Herztönen (15 Sekunden) bestimmt. Für die Bestimmung der Treffsicherheit wurde die Zahl der korrekt vorhergesagten Schweregrade der AS durch die Anzahl aller analysierten Patienten geteilt.

2.6.3 Parameter für die Überprüfung der Leistung der Modelle

Bei der Überprüfung der KI wurde als Leistungsparameter vor allem die Treffsicherheit beziehungsweise Genauigkeit der Vorhersage (*Accuracy*) betrachtet. Daneben waren die Sensitivität und die Spezifität von Bedeutung. Darauf aufbauend wurden ebenfalls der positiv und negativ prädiktive Wert (PPV; NPV) bestimmt. Zusätzlich erfolgte die Bestimmung der Fläche unter der Kurve (*area under the curve, AUC*). Zudem wurde der F1-Wert berechnet, das harmonische Mittel aus der Sensitivität und dem positiv prädiktiven Wert.

2.6.4 Einstellungen der Hyperparameter während des Trainings

Im Rahmen der Abstimmung der neuronalen Netzwerke während des Trainings mussten bestimmte Voreinstellungen gefunden werden, um die bestmöglichen Ergebnisse zu erreichen. Diese sogenannten Hyperparameter wurden vor dem Training festgesetzt und blieben während des gesamten Trainingsprozesses gleich. Auf Grundlage von vorherigen Untersuchungen wurden diese Hyperparameter angepasst.

Anders verhielt es sich bei den Gewichtungen der neuronalen Netzwerke, da sich diese durch immer mehr Trainingsdatensätze verändern und vor allem verbessern. Die Gewichtung der Netzwerke wurde an Vorarbeiten von He et al. (2015) orientiert. Da es in dieser Studie mehr Datensätze ohne schwere AS gibt, wurden die Verbindungen für die schwere AS im Verhältnis zu den Datensätzen ohne schwere AS 1,2:1 gewichtet.

2.6.5 Validierung im klinischen Kontext

Um das Modell mit weiteren Auskultationsdaten zu testen, wurden für die Validierung im klinischen Kontext anschließend erneut Patientendaten prospektiv in die Studie eingeschlossen. Damit es zu aussagekräftigen Ergebnissen kam, musste eine ausreichend große Gruppe rekrutiert werden. Hierbei wurde wie bei den vorherigen Datensätzen nach keiner, leichter, moderater und schwerer Aortenklappenstenose unterschieden. Das Ziel dieser Phase war somit die Weiterentwicklung der KI mittels einer von den initialen Trainings-, Validierungs- und Testfällen unabhängigen Kohorte. Zudem sollten vor allem die Resultate bei der Detektion der schweren Aortenklappenstenose getestet werden, da diese klinisch die größte Relevanz hat.

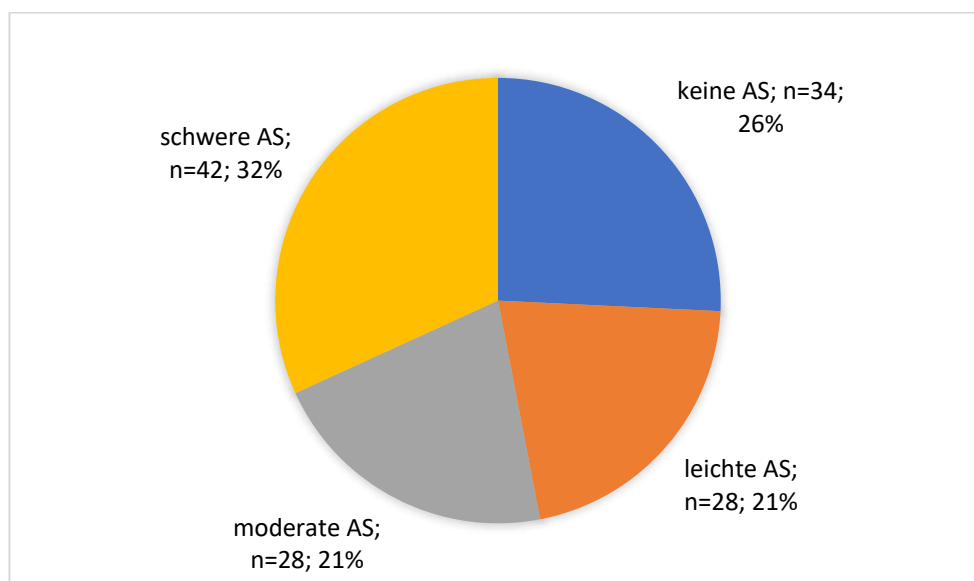


Abb. 5: Rekrutierung neuer Datensätze im Rahmen der klinischen Validierung

Anteil der Schweregrade der Aortenklappenstenosen an denen für die klinische Validierung rekrutierten Datensätzen

Wie Abbildung 5 zeigt, wurden insgesamt erneut 132 Datensätze rekrutiert und im Rahmen der klinischen Validierung verwendet. Diese wurden an den drei schon erwähnten Auskultationsorten aufgenommen. Im zweiten Schritt wurden noch die 330 oben genannten Patienten ohne Aortenklappenstenose hinzugefügt, die vor der Kreuzvalidierung ausgeschlossen worden waren, um die Prävalenz der Schweregrade der Aortenklappenstenose an die Bevölkerung anzugleichen.

2.7 Export auf eine Smartphone App

Um die Effizienz der KI im Alltag noch weiter zu verbessern, ist es notwendig, dass sie möglichst leicht zugänglich ist für die Anwender. Dies gelingt über eine Smartphone App, mit Hilfe derer evaluiert wird, ob eine schwere Aortenklappenstenose vorliegt. Hierfür wurde die Anwendung „*AudioClassification*“ verwendet. Die getesteten Modelle wurden mittels *TensorFlowLite* in diese Anwendung übertragen. Es wurde ein iPhone 11 Pro mit iOS14.4 (Apple, Cupertino, Kalifornien) als Testobjekt verwendet.

Dafür wurden die Herztöne von allen drei Auskultationsorten in die Anwendung übertragen. Im Klassifikationsmodus konnten dann verschiedene Parameter wie die Gewichtung oder die Fourier-Transformation für die Daten individuell festgelegt werden. Anschließend wurde die Analyse in der App durchgeführt. Am Ende des Prozesses stand die Entscheidung, ob eine schwere AS vorliegt oder nicht.

2.8 Visualisierung des Entscheidungsprozesses mittels Aktivierungskarten

Um die entwickelte KI für den Menschen transparent und nachvollziehbar zu machen, ist es wichtig zu verstehen, auf Grundlage welcher Parameter die Künstliche Intelligenz Entscheidungen trifft. Dies ist möglich durch das sogenannte *Gradient-Weighted Class Activation Mapping++* (Grad-CAM++). Hierbei wird die Aktivierung der letzten Faltungsschicht des neuronalen Netzwerkes dargestellt (Chattopadhyay et al., 2018). Es entstehen Aktivierungskarten (*Heatmaps*), wann welche Neurone des Netzwerkes aktiv sind. Diese Visualisierung wurde dem Mel-Spektrogramm überlagert. Somit wurde anhand der unterschiedlichen Farbskalierungen sichtbar, welche Phase der Herzaktion von den Modellen bei der Entscheidungsfindung fokussiert wird. Auf dieser Grundlage entscheidet die KI für den Menschen transparent und nachvollziehbar, welcher Schweregrad der AS vorliegt. Darüber hinaus konnte eine Fokusscoreanalyse durchgeführt werden.

2.9 Statistische Analysen

Die Analysen wurden unter Verwendung der JMP-Software Version 14 (SAS Institut) und mit Python-Skripten durchgeführt. Für die statistischen Analysen der Leistungen der erstellten Modelle mit den Testdaten und während der klinischen Validierung wurde die *Bootstrapping* Technik verwendet.

Alle kontinuierlich erhobenen Daten wurden darauf geprüft, ob sie normalverteilt oder nicht-normalverteilt sind. Normalverteilte Daten werden als Mittelwert mit Standardabweichung angegeben, nicht normal-verteilte Daten werden als Medianwerte mit den entsprechenden Quartilen angegeben. Kategoriale Daten werden als Zahlen und Prozente dargestellt. Der Chi-Quadrat-Test, der Kruskal-Wallis-Test, der Student's t-Test oder der Fisher's Exact Test wurden durchgeführt, wenn dies angemessen war. Für die globalen Teststatistiken wurde ein Signifikanzniveau von 5 % verwendet.

3 Ergebnisse

3.1 Leistung der Künstlichen Intelligenz im Test

Die neuronalen Netzwerke wurden mit den Datensätzen erst an den drei Auskultationsorten (2RSB, Erb, Apex) separat erprobt. Hierbei wurden die Aufzeichnungen in vier Sekunden lange Abschnitte geteilt, um das Training zu validieren. Anschließend wurden die Modelle auch mit Auskultationsdaten von mehreren Orten trainiert. Es wurden jeweils mehrere verschiedene CNNs verglichen. Als Vergleichskriterien zwischen den Modellen wurden die Treffsicherheit (*Accuracy*), der F1 Wert und die *area under the curve* (AUC) herangezogen. Zudem wurden die Sensitivität, Spezifität, der positive und negative prädiktive Wert sowie die Verlustfunktion betrachtet. Die Treffsicherheit kann dabei vom Begriff her mit der Richtigkeit oder Genauigkeit gleichgesetzt werden.

3.1.1 2RSB

Tabelle 5: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten von 2RSB

	Alle Modelle	CNN10L-T	cCNN-T	CNN10L-L1	cCNN-L1	p-Wert
F1 Wert	0,892 (+/- 0,025)	0,869 (+/- 0,015)	0,881 (+/- 0,024)	0,92 (+/- 0,016)	0,9 (+/- 0,012)	0,0014
Treffsicherheit	0,909 (+/- 0,023)	0,888 (+/- 0,011)	0,896 (+/- 0,02)	0,934 (+/- 0,014)	0,917 (+/- 0,01)	0,0005
AUC	0,964 (+/- 0,015)	0,949 (+/- 0,014)	0,96 (+/- 0,011)	0,975 (+/- 0,01)	0,972 (+/- 0,0006)	0,0035
Sensitivität	0,89 (+/- 0,028)	0,876 (+/- 0,032)	0,906 (+/- 0,033)	0,902 (+/- 0,015)	0,876 (+/- 0,019)	0,16
Spezifität	0,923 (+/- 0,034)	0,897 (+/- 0,005)	0,888 (+/- 0,018)	0,957 (+/- 0,017)	0,948 (+/- 0,019)	<0,0001
PPV	0,896 (+/- 0,042)	0,863 (+/- 0,002)	0,857 (+/- 0,022)	0,94 (+/- 0,023)	0,926 (+/- 0,024)	<0,0001

NPV	0,92 (+/- 0,019)	0,908 (+/- 0,021)	0,928 (+/- 0,024)	0,93 (+/- 0,01)	0,912 (+/- 0,012)	0,18
Verlustfunktion	0,262 (+/- 0,045)	0,308 (+/- 0,021)	0,285 (+/- 0,042)	0,225 (+/- 0,027)	0,229 (+/- 0,011)	0,0003

AUC = *area under the curve*; PPV = positiv prädiktiver Wert, NPV = negativ prädiktiver Wert
Vergleich der erreichten Leistungen in Bezug auf den F1 Wert, die Treffsicherheit, die AUC, die Sensitivität und Spezifität, den PPV sowie NPV sowie der Verlustfunktion sowohl zusammengefasst von allen Modellen als auch der einzelnen Modelle, der jeweils beste erreichte Wert pro Zeile ist fett markiert

Tabelle modifiziert nach Makimoto et al. (2022)

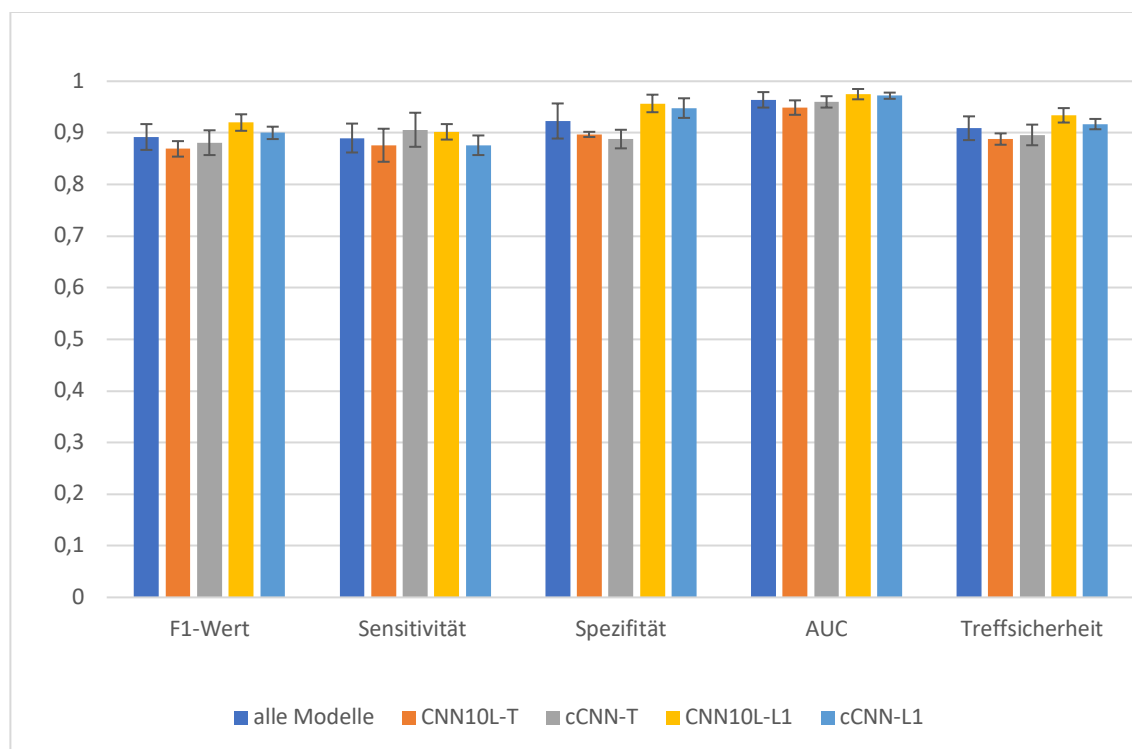


Abb. 6: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten von 2RSB

Grafische Darstellung der in Tabelle 5 dargestellten Leistungsdaten der verschiedenen CNNs in Hinblick auf den F1-Wert, die Sensitivität, Spezifität sowie AUC und Treffsicherheit mit Überlegenheit des CNN10L-L1 Modells

Abbildung modifiziert nach Makimoto et al. (2022)

Wie in Tabelle 5 und Abbildung 6 dargestellt, erzielte das Modell CNN10L-L1 die besten Testergebnisse. Hierbei handelt es sich um ein 10-schichtiges neuronales Netzwerk, welches spezifisch nur mit den Auskultationsergebnissen von dieser Lokalisation trainiert wurde. Dieses CNN zeigt signifikant bessere Ergebnisse als die anderen Modelle in fast allen Bereichen außer der Sensitivität und dem negativ

prädiktiven Wert. Vor allem die Spezifität mit 95,7% (p-Wert <0,0001) sowie der positiv prädiktive Wert mit 93% (p-Wert <0,0001) sind signifikant besser als in den anderen Vergleichsmodellen (Makimoto et al., 2022).

3.1 Erb-Punkt

Tabelle 6: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten des Erb-Punktes

	Alle Modelle	CNN10L-T	cCNN-T	CNN10L-L2	cCNN-L2	p-Wert
F1-Wert	0,803 (+/- 0,029)	0,817 (+/- 0,033)	0,804 (+/- 0,037)	0,804 (+/- 0,027)	0,786 (+/- 0,017)	0,46
Treffsicherheit	0,838 (+/- 0,021)	0,851 (+/- 0,023)	0,840 (+/- 0,025)	0,839 (+/- 0,019)	0,824 (+/- 0,011)	0,26
AUC	0,897 (+/-0,01)	0,9 (+/- 0,01)	0,9 (+/- 0,013)	0,893 (+/- 0,009)	0,874 (+/- 0,008)	0,53
Sensitivität	0,778 (+/- 0,0047)	0,787 (+/- 0,055)	0,778 (+/- 0,059)	0,784 (+/- 0,047)	0,764 (+/- 0,037)	0,89
Spezifität	0,883 (+/- 0,016)	0,898 (+/- 0,005)	0,886 (+/- 0,015)	0,879 (+/- 0,006)	0,868 (+/- 0,019)	0,014
PPV	0,83 (+/- 0,02)	0,85 (+/- 0,01)	0,833 (+/- 0,02)	0,827 (+/- 0,009)	0,811 (+/- 0,019)	0,0081
NPV	0,845 (+/- 0,028)	0,852 (+/- 0,033)	0,845 (+/- 0,035)	0,848 (+/- 0,029)	0,834 (+/- 0,02)	0,79
Verlustfunktion	0,442 (+/- 0,037)	0,434 (+/- 0,042)	0,421 (+/- 0,031)	0,463 (+/- 0,039)	0,45 (+/- 0,031)	0,3

AUC = *area under the curve*; PPV = positiv prädiktiver Wert, NPV = negativ prädiktiver Wert
 Vergleich der erreichten Leistungen in Bezug auf den F1 Wert, die Treffsicherheit, die AUC, die Sensitivität und Spezifität, den PPV sowie NPV sowie der Verlustfunktion sowohl zusammengefasst von allen Modellen als auch der einzelnen Modelle, der jeweils beste erreichte Wert pro Zeile ist fett markiert

Tabelle modifiziert nach Makimoto et al. (2022)

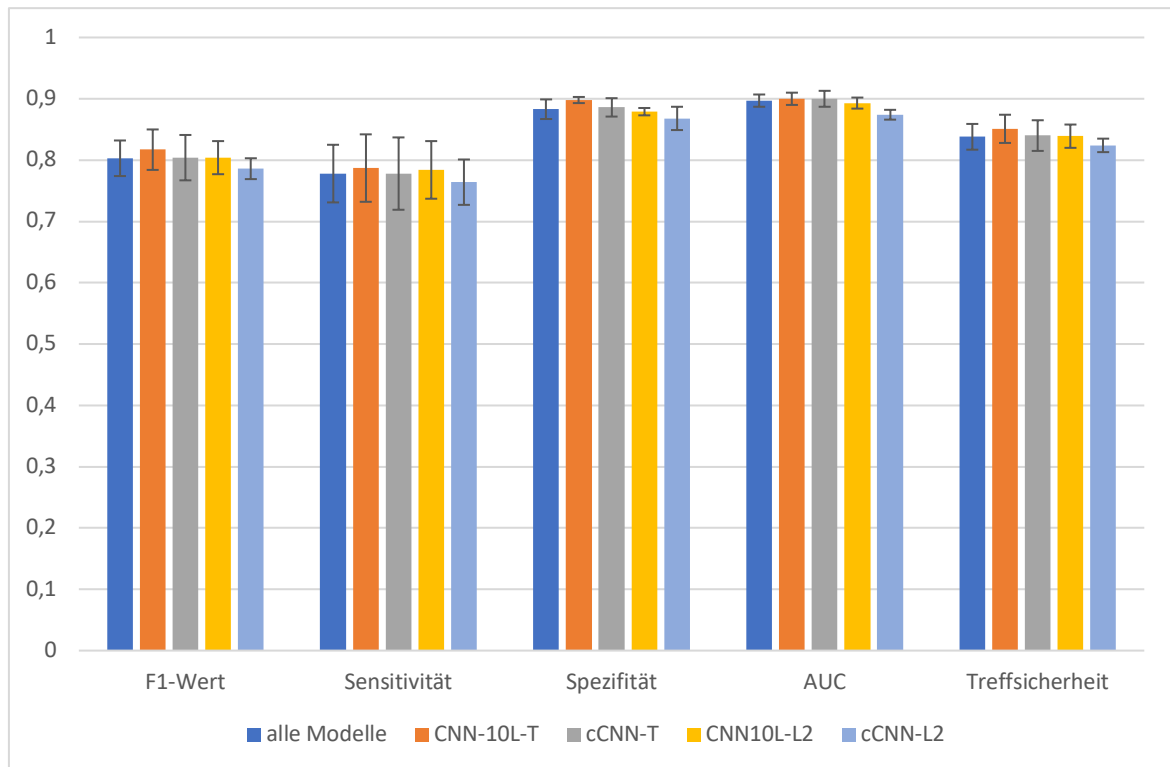


Abb. 7: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten des Erb-Punktes

Grafische Darstellung der in Tabelle 6 dargestellten Leistungsdaten der verschiedenen CNNs in Hinblick auf den F1-Wert, die Sensitivität, Spezifität sowie AUC und Treffsicherheit mit Überlegenheit des CNN10L-T Modells

Abbildung modifiziert nach Makimoto et al. (2022)

Wenn man sich in Tabelle 6 und Abbildung 7 die Ergebnisse für die Auskultationsdaten mit Fokus auf den Erb-Punkt (L2) anschaut, dann fällt auf, dass hier das CNN10L-T Modell am besten abschneidet. Dieses neuronale Netzwerk hat dabei mit Daten von allen drei Auskultationsorten trainiert. Insbesondere in Hinblick auf die Spezifität (89,8%, p-Wert 0,014) und den positiv prädiktiven Wert wurden hier die besten Ergebnisse erreicht (85%, p-Wert 0,0081; Makimoto et al., 2022).

3.1.3 Apex

Tabelle 7: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten des Apex

	Alle Modelle	CNN-10L-T	cCNN-T	CNN10-L3	cCNN-L3	p-Wert
F1-Wert	0,8 (+/- 0,013)	0,798 (+/- 0,01)	0,809 (+/- 0,014)	0,795 (+/- 0,009)	0,796 (+/- 0,017)	0,33
Treffsicherheit	0,833 (+/-0,01)	0,835 (+/- 0,006)	0,838 (+/- 0,012)	0,831 (+/- 0,006)	0,828 (+/- 0,015)	0,5
AUC	0,91 (+/- 0,014)	0,896 (+/- 0,01)	0,923 (+/- 0,008)	0,902 (+/- 0,011)	0,917 (+/- 0,009)	0,001
Sensitivität	0,785 (+/- 0,029)	0,769 (+/- 0,026)	0,808 (+/- 0,022)	0,773 (+/- 0,036)	0,791 (+/-0,02)	0,12
Spezifität	0,868 (+/- 0,023)	0,883 (+/- 0,017)	0,860 (+/- 0,019)	0,874 (+/- 0,029)	0,856 (+/- 0,009)	0,22
PPV	0,816 (+/- 0,023)	0,829 (+/- 0,017)	0,810 (+/- 0,019)	0,821 (+/- 0,028)	0,802 (+/- 0,022)	0,27
NPV	0,846 (+/- 0,016)	0,839 (+/- 0,014)	0,859 (+/- 0,013)	0,84 (+/- 0,018)	0,848 (+/- 0,013)	0,14
Verlustfunktion	0,385 (+/- 0,033)	0,414 (+/- 0,026)	0,358 (+/- 0,019)	0,405 (+/- 0,029)	0,363 (+/- 0,017)	0,0023

AUC = *area under the curve*; PPV = positiv prädiktiver Wert, NPV = negativ prädiktiver Wert
 Vergleich der erreichten Leistungen in Bezug auf den F1 Wert, die Treffsicherheit, die AUC, die Sensitivität und Spezifität, den PPV sowie NPV sowie der Verlustfunktion sowohl zusammengefasst von allen Modellen als auch der einzelnen Modelle, der jeweils beste erreichte Wert pro Zeile ist fett markiert

Tabelle modifiziert nach Makimoto et al. (2022)

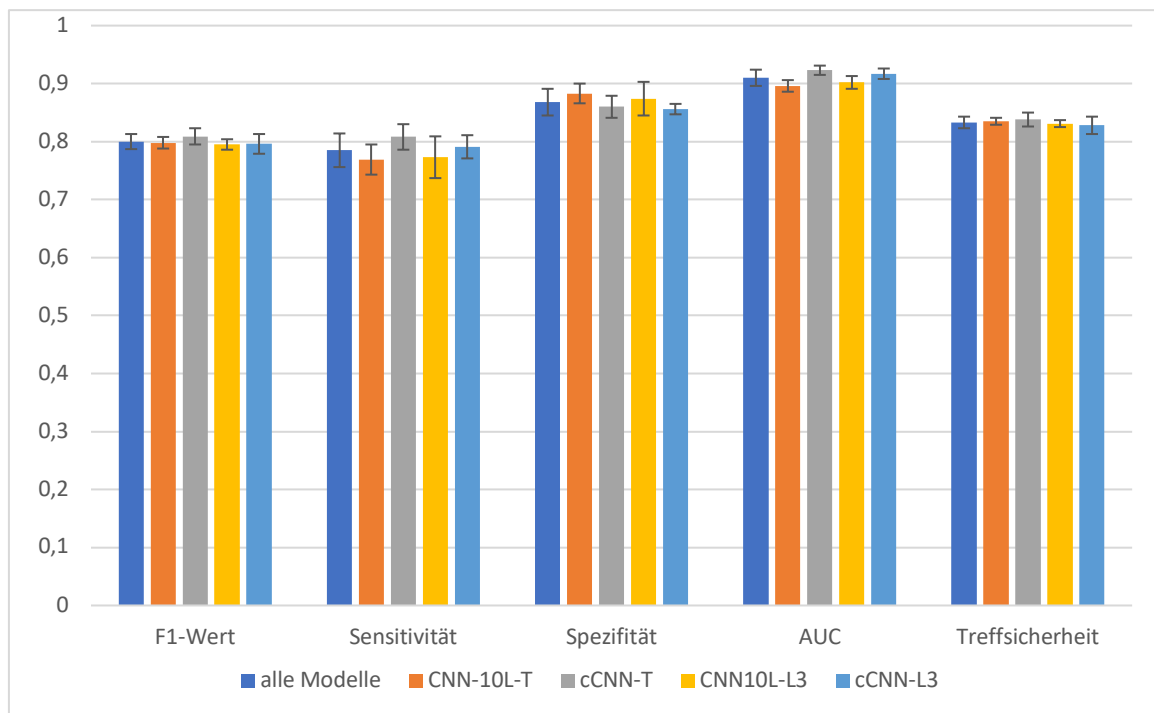


Abb. 8: Vergleich der Leistungsdaten der verschiedenen *convolutional neural networks* anhand der Auskultationsdaten des Apex

Grafische Darstellung der in Tabelle 7 dargestellten Leistungsdaten der verschiedenen CNNs in Hinblick auf den F1-Wert, die Sensitivität, Spezifität sowie AUC und Treffsicherheit mit Überlegenheit des cCNN-T Modells

Abbildung modifiziert nach Makimoto et al. (2022)

Tabelle 7 sowie Abbildung 8 präsentieren die Ergebnisse der Auskultationsdaten mit Schwerpunkt auf die Apex-Lokalisation. Hier schneidet das cCNN-T Modell am besten ab. Dabei wurde das neuronale Netzwerk mit Kombinationskernen mit dem gesamten Datensatz von allen drei Auskultationsorten trainiert. Insbesondere die Ergebnisse bei der Fläche unter der Kurve und der Verlustfunktion sind den anderen Modellen signifikant überlegen.

Ein Vergleich der Modelle (Abbildungen 9 bis 11) zeigt, dass in sämtlichen untersuchten Parametern die Modelle, die mit den Auskultationsdaten vom zweiten Interkostalraum rechts parasternal (2RSB) trainiert wurden, signifikant bessere Ergebnisse erzielten als die Modelle L2 (Erb) und L3 (Apex). Der p-Wert liegt in allen Fällen unter 0,001. Daher lässt sich der Schluss ziehen, dass mit den Daten von diesem Auskultationsort die signifikant besten Ergebnisse erzielt werden (Makimoto et al., 2022).

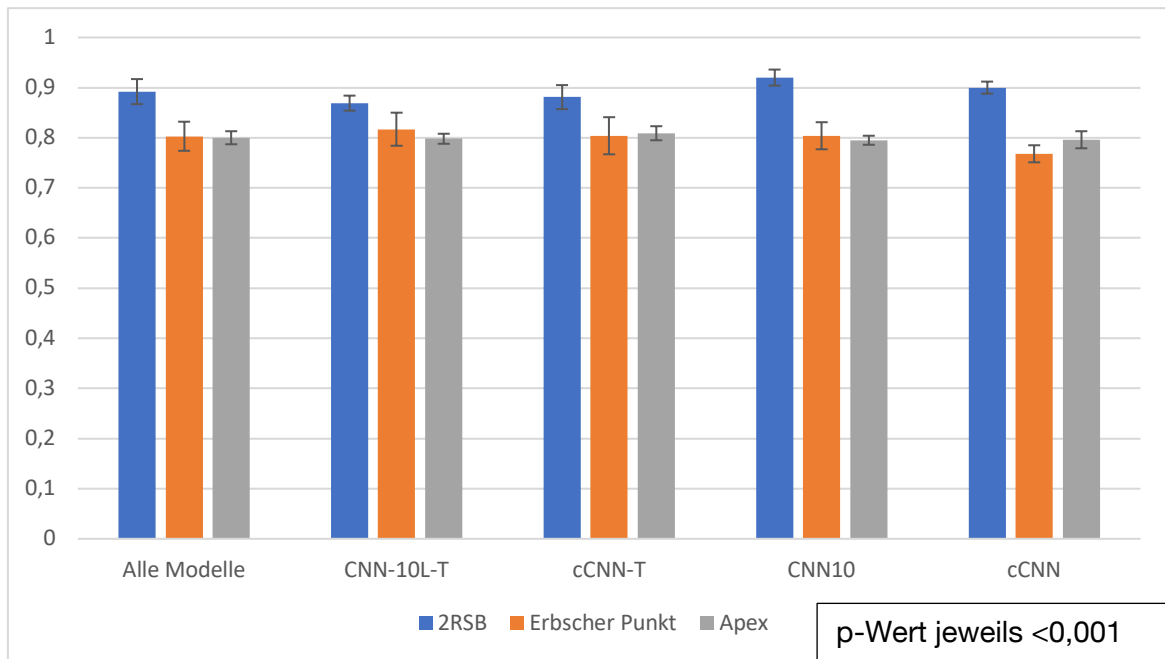


Abb. 9: Vergleich der Ergebnisse der *convolutional neural networks* anhand des F1-Wertes
 Die signifikant besten Ergebnisse werden mit den Auskultationsdaten vom zweiten Interkostalraum rechts parasternal (2RSB) erreicht (p-Wert <0,001)
 Abbildung modifiziert nach Makimoto et al. (2022)

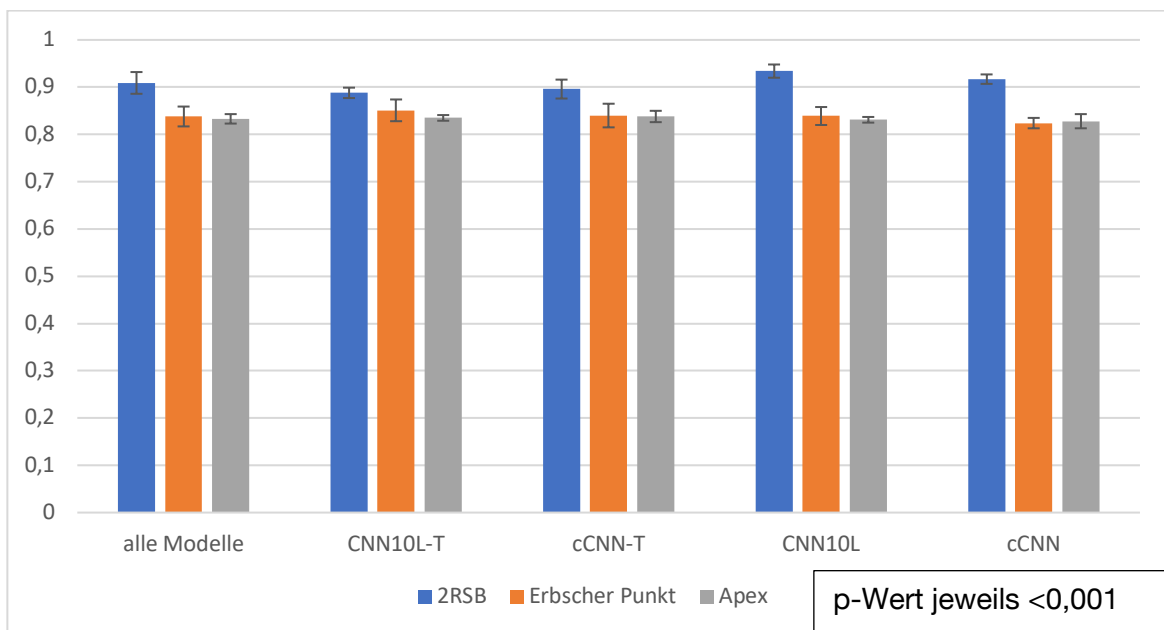


Abb. 10: Vergleich der Ergebnisse der *convolutional neural networks* anhand der Treffsicherheit
 Die signifikant besten Ergebnisse werden mit den Auskultationsdaten vom zweiten Interkostalraum rechts parasternal (2RSB) erreicht (p-Wert <0,001)
 Abbildung modifiziert nach Makimoto et al. (2022)

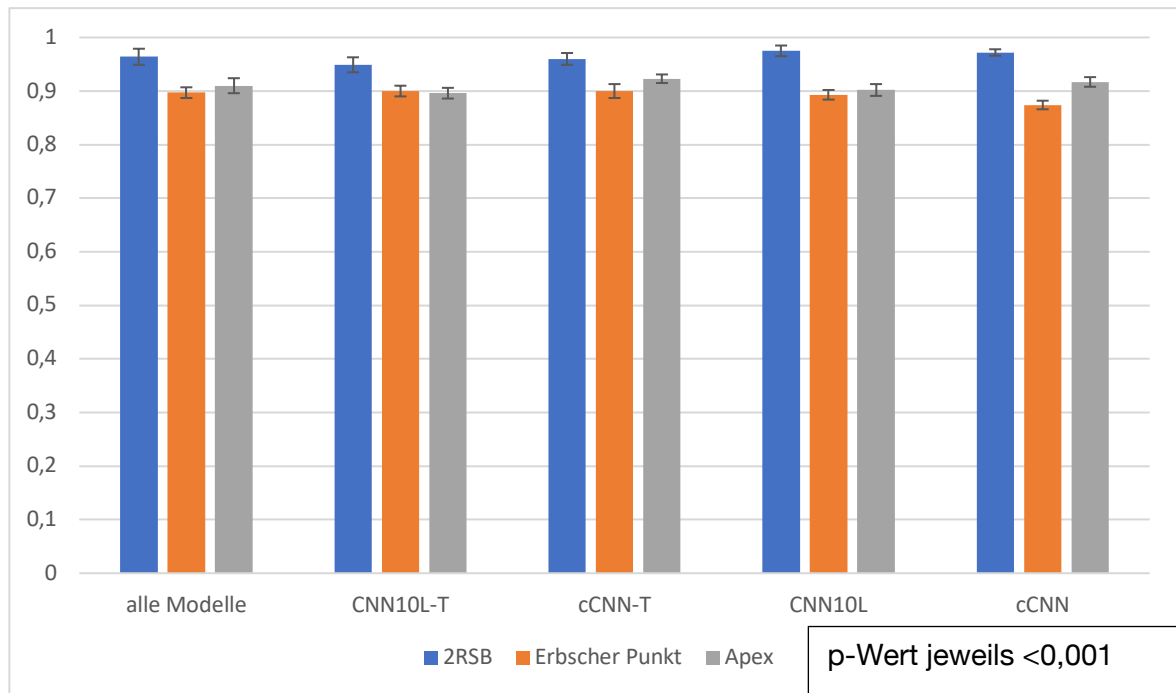


Abb. 11: Vergleich der Ergebnisse der *convolutional neural networks* anhand der *area under the curve*

Die signifikant besten Ergebnisse werden mit den Auskultationsdaten vom zweiten Interkostalraum rechts parasternal (2RSB) erreicht (p-Wert <0,001)

Abbildung modifiziert nach Makimoto et al. (2022)

Bei den unterschiedlichen neuronalen Netzwerken wurde schlussendlich immer jenes für die weitere Verwendung vorgesehen, welches den besten F1 Wert (siehe Abbildung 9) aufweisen konnte. Dieses war in allen Fällen auch das Modell, welches in den meisten untersuchten Leistungsparametern die signifikant besten Ergebnisse zeigen konnte (Makimoto et al., 2022).

Werden nun im nächsten Schritt die gesamten aufgenommenen Herztöne (15 Sekunden) zu Grunde gelegt, dann kann man bei den verwendeten Modellen L1-3 die Treffsicherheit (Genauigkeit) in Hinblick auf die Detektion des Schweregrades der Aortenklappenstenose auswerten. Dies zeigt in nachfolgenden Abbildung 12.

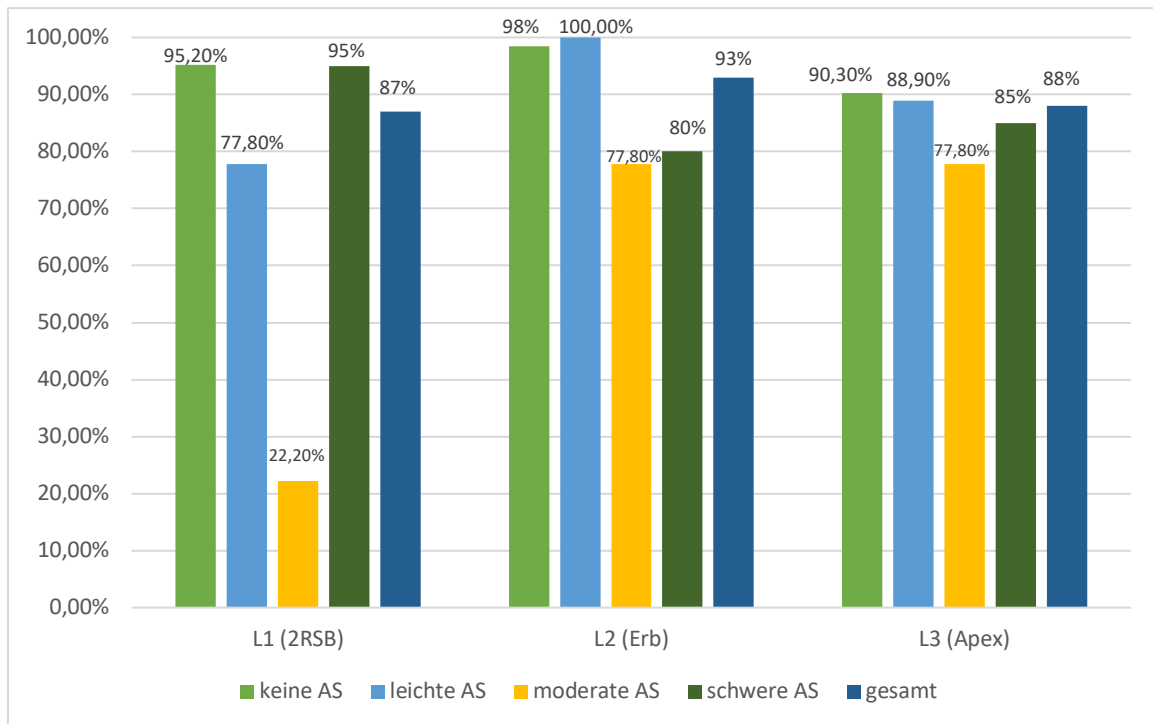


Abb. 12: Vergleich der Treffsicherheit der Modelle L1-L3 in Bezug auf den Schweregrad der Aortenklappenstenose

Unterscheidung der Treffsicherheit bei keiner, einer leichten, einer moderaten sowie einer schweren AS durch die Modelle L1, L2 und L3

Abbildung modifiziert nach Makimoto et al. (2022)

Hierbei lässt sich feststellen, dass die Treffsicherheit im Gesamten gemittelt über alle Schweregrade der AS bei den Modellen L1-3 bei über 85% lag. Vor allem die schweren Aortenklappenstenosen sowie die Fälle ohne AS wurden zuverlässig in allen Modellen mit hoher Wahrscheinlichkeit korrekt zugeordnet. Hier schneidet erneut das Modell L1 besonders gut ab, welches 95% der schweren Aortenklappenstenosen korrekt erkennt. Beim moderaten Schweregrad gibt es teils Unsicherheiten in der Erkennung (Makimoto et al., 2022).

Neben der Treffsicherheit können auch die Sensitivität und die Spezifität als Parameter für die Überprüfung der erstellten Modelle herangezogen werden, wie man in Abbildung 13 erkennen kann.

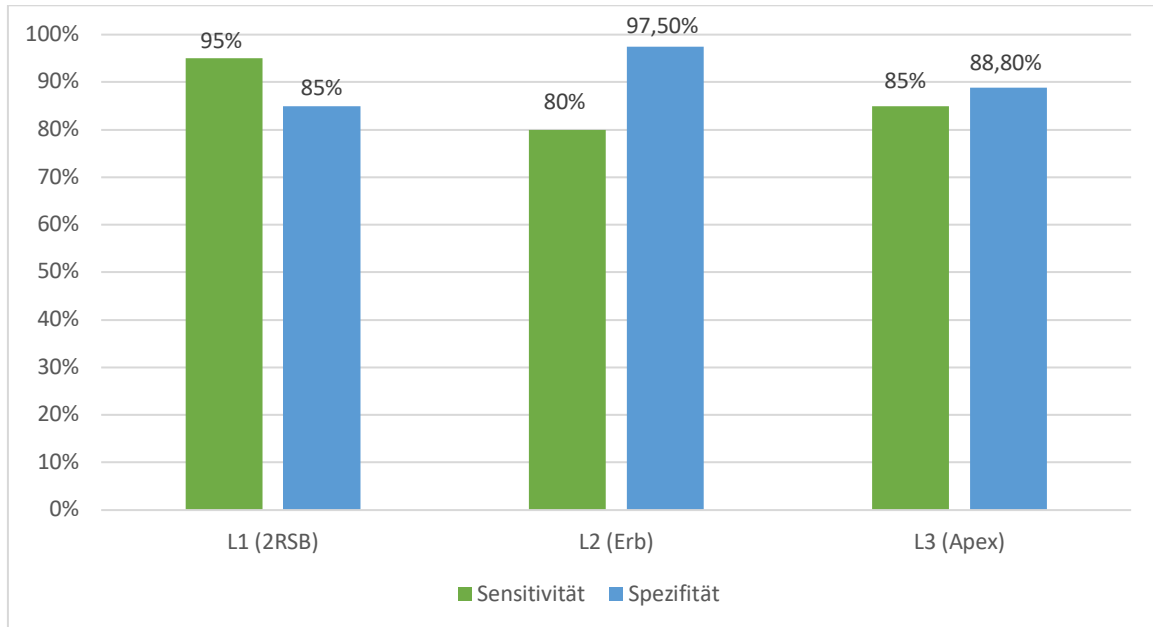


Abb. 13: Vergleich der Sensitivität und der Spezifität der Modelle L1-3

Abbildung modifiziert nach Makimoto et al. (2022)

Das L1 Modell erzielte dabei mit 95% die höchste Sensitivität für die Erkennung der schweren AS, während das L2 Modell mit 97,5% die beste Spezifität aller Datensätze aufwies (Makimoto et al., 2022). Insgesamt lagen alle drei Modelle kontinuierlich in beiden Kategorien über 80%.

3.1.4 Ergebnisse der zusammengesetzten Modelle

Um die Leistung der Modelle L1-3 noch weiter zu verbessern, wurden im nächsten Schritt die Modelle miteinander kombiniert. Es entstanden dadurch die Modelle E1-3 (*Ensemble*). Diese basieren auf L1-3 in jeweils unterschiedlich gewichteten Anteilen. Das E1 Modell wurde in gleichen Anteilen erstellt (1:1:1). Die anderen beiden Modelle gewichten jeweils L1 stärker als L2 und L3, da sich wie bereits erwähnt zeigte, dass die Auskultationsergebnisse von 2RSB signifikant bessere Ergebnisse lieferten. Es findet sich bei E2 die Gewichtung 6:2:2 und bei E3 8:1:1 (L1:L2:L3; Makimoto et al., 2022).

Mit diesen neu zusammengesetzten Modellen wurde erneut das Training durchgeführt. Die Ergebnisse finden sich in den folgenden Abbildungen 14 und 15.

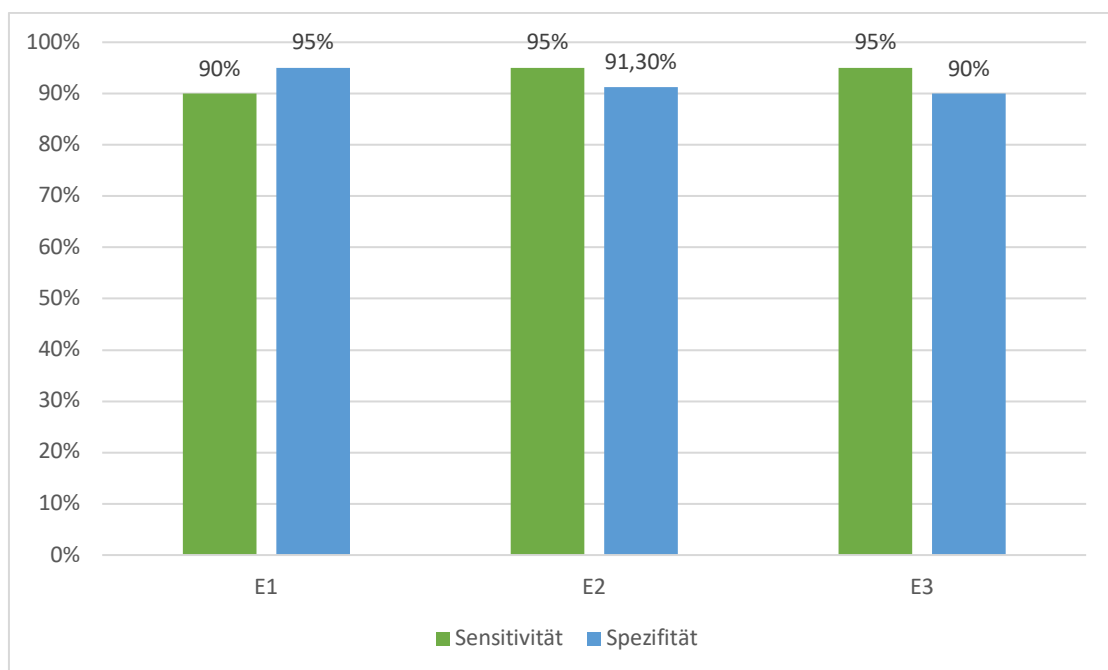


Abb. 14: Vergleich der Sensitivität und Spezifität der zusammengesetzten Modelle E1-E3
Abbildung modifiziert nach Makimoto et al. (2022)

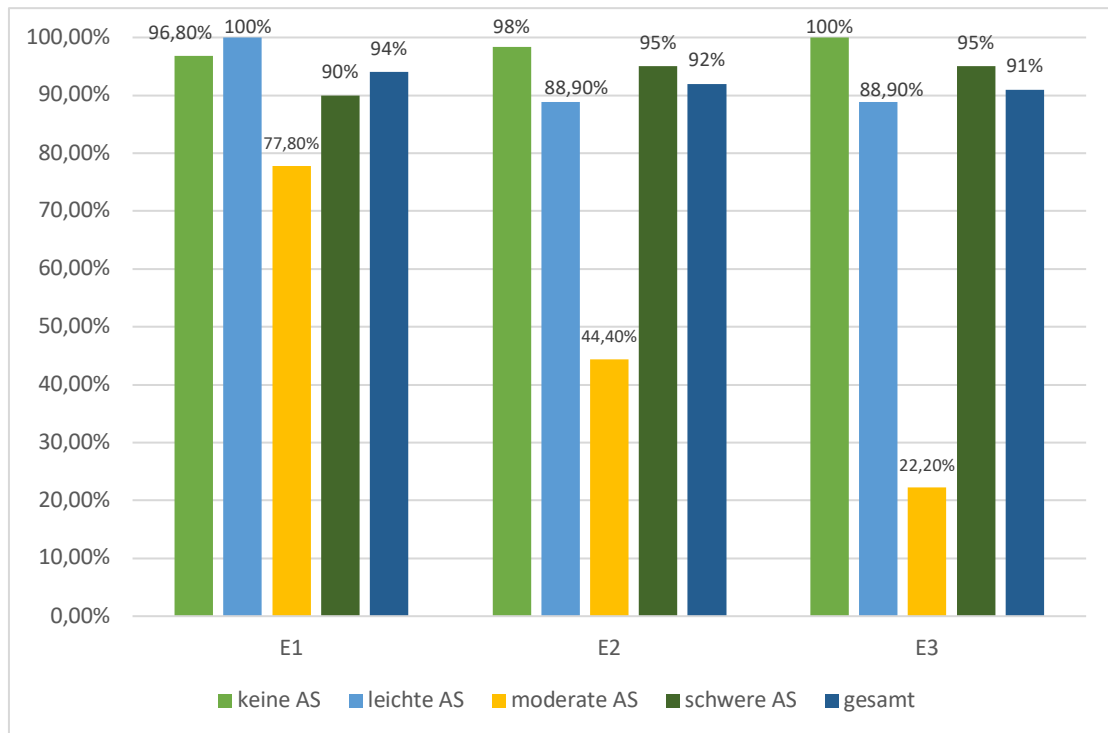


Abb. 15: Vergleich der Treffsicherheit der Modelle E1-E3 in Bezug auf den Schweregrad der Aortenklappenstenose

Unterscheidung der Treffsicherheit bei keiner, einer leichten, einer moderaten sowie einer schweren AS durch die Modelle E1, E2 und E3

Abbildung modifiziert nach Makimoto et al. (2022)

Die Ergebnisse zeigen nun eine weitere Verbesserung der Sensitivitäts- und Spezifitätswerte. Sämtliche erhobene Kennzahlen liegen nun über 90%. Die höchste Sensitivität erreichten die Modelle E2 und E3 mit jeweils 95%, die höchste Spezifität wies das E1 Modell ebenfalls mit 95% auf. Darüber hinaus ist der F1 Wert sowie die Treffsicherheit mit 94% beim E1 Modell signifikant höher ($p < 0,001$; Makimoto et al., 2022).

Die Genauigkeit der Vorhersage variierte zudem je nach Schweregrad der Aortenklappenstenose. Insgesamt zeigte das E1 Modell die besten Ergebnisse mit einer korrekten Vorhersage von 94%, die Detektion der schweren AS gelang dabei in 90% der Fälle. Unsicherheiten bei der Zuordnung verblieben bei der moderaten Aortenklappenstenose. Schließlich wurde Modell E1 für die weitere klinische Validierung ausgewählt.

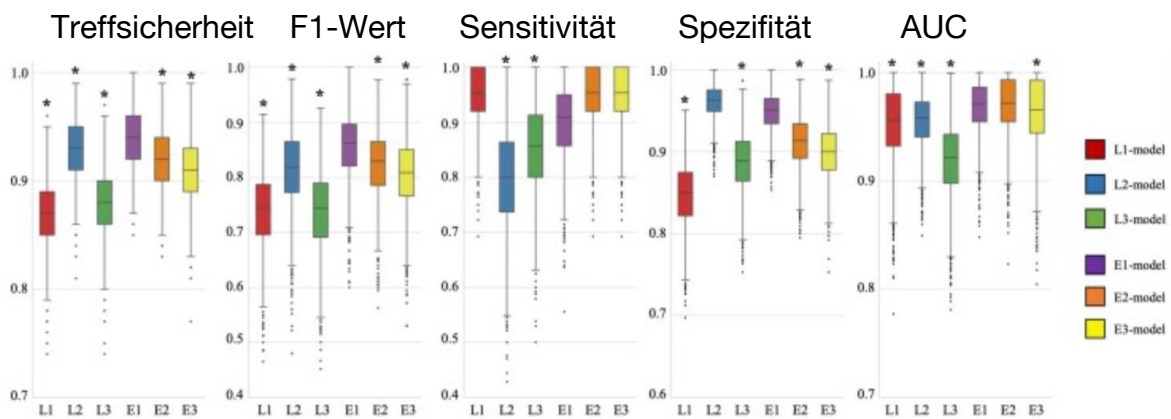


Abb. 16: Vergleich zwischen den Modellen L1-L3 und E1-E3 in Hinblick auf die Treffsicherheit, F1 Wert, Sensitivität, Spezifität und area under the curve (AUC)

Mit Sternchen markiert sind die Modelle, die signifikant schlechtere Ergebnisse erzielen als das E1 Modell (p-Wert <0,001). Insgesamt erzielt das E1 Modell die signifikant besten Ergebnisse.

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

In der Abbildung 16 zeigt jedes der mit einem Sternchen markierten Graphen die signifikante Unterlegenheit der Modelle gegen das Modell E1 an. Wenn man die Werte von E1 mit denen von L1-L3 sowie E2 und E3 vergleicht, erreicht das Modell E1 beispielsweise bei der Treffsicherheit gegenüber allen anderen Modellen ein signifikant besseres Ergebnis (p-Wert <0,001). Dies ist auch beim F1 Wert der Fall. Es zeigt sich dementsprechend die deutliche Überlegenheit des Modells E1 (Makimoto et al., 2022).

3.2 Klinische Validierung

3.2.1 Vergleich der Modelle

Verglichen wurden im Rahmen der klinischen Validierung mit den additiv rekrutierten Datensätzen (n=132) die Modelle L1-L3 von den oben genannten Auskultationsorten mit dem E1 Modell, da sich dieses bei den Testdatensätzen als das zusammengesetzte Modell mit der besten Leistung durchgesetzt hatte. Dies lässt sich Abbildung 17 entnehmen. Das E1 Modell ist dabei in der gleichen Gewichtung der Modelle L1-3 zusammengesetzt (1:1:1).

Die Modelle L1-3 lagen bei der Treffsicherheit insgesamt allesamt knapp unter 90%, E1 jedoch deutlich darüber (95,7%). Auch bei der Sensitivität (E1: 97,6%) sowie der Spezifität (E1: 94,4%) zeigt das E1 Modell jeweils die besten Ergebnisse. Gleiches gilt für den F1 Wert. Die Ergebnisse bestätigen somit auch in der klinischen Validierungsgruppe eine Überlegenheit von E1.

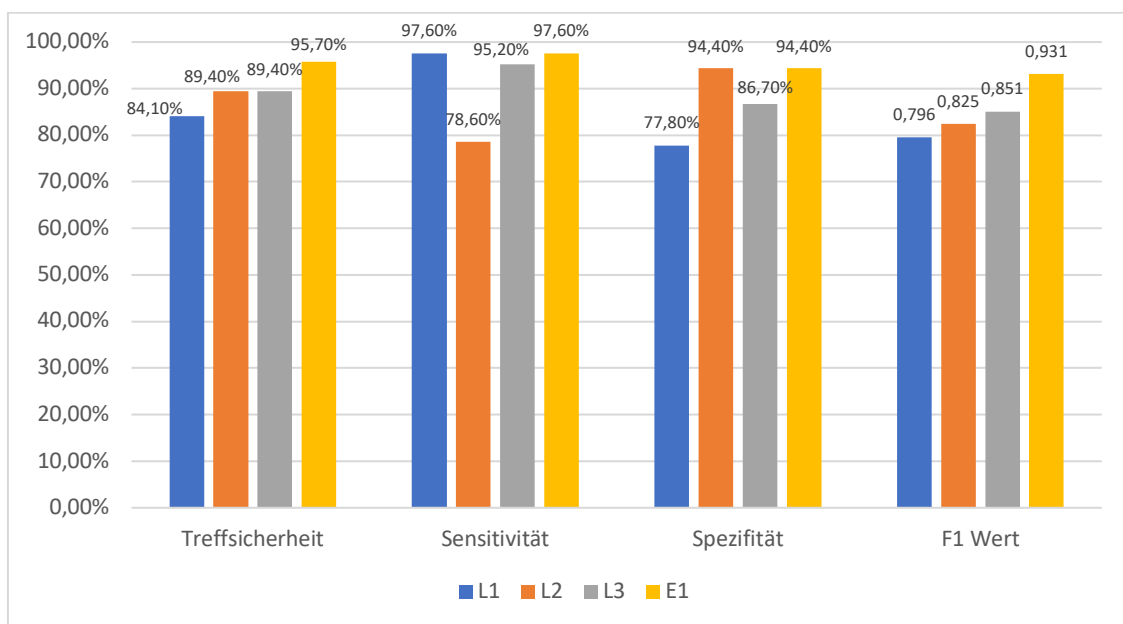


Abb. 17: Vergleich der Modelle L1, L2, L3 und E1 in Hinblick auf die Treffsicherheit, die Sensitivität, die Spezifität und den F1 Wert bei der klinischen Validierung
Abbildung modifiziert nach Makimoto et al. (2022)

3.2.2 Ergebnisse des E1 Modells

Das E1 Modell zeigte bei der klinischen Validierung sehr zuverlässige Ergebnisse. Wie in Abbildung 18 ersichtlich, wurden insgesamt 126 der 132 Testdaten korrekt zugeordnet (95,7%). Wenn man dies aufgeschlüsselt für die einzelnen Schweregrade der AS betrachtet, dann lässt sich feststellen, dass das E1 Modell in allen Kategorien deutlich über 90% Treffsicherheit liegt. Die einzige Ausnahme bildet hierbei erneut der moderate Schweregrad. Hier wurden 82,1% der Datensätze korrekt zugeordnet. Dafür sind die Zuordnungen bei der schweren (97,6%), leichten (100%) und nicht vorhandener AS (100%) sehr zuverlässig gelungen (Makimoto et al., 2022).

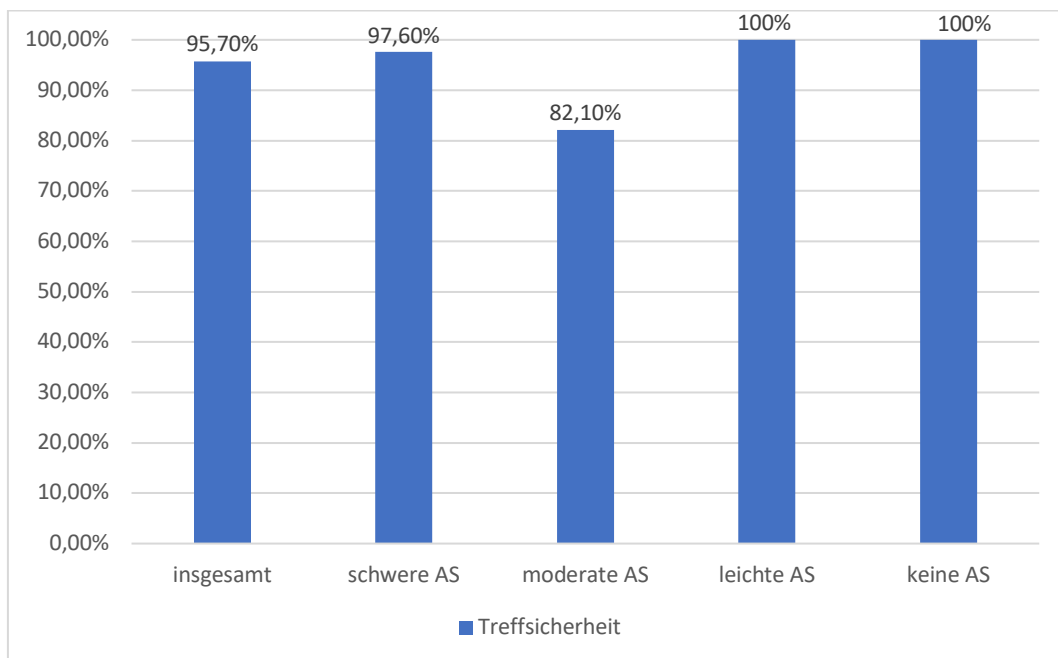


Abb. 18: Vergleich der erreichten Treffsicherheit des E1 Modells bei den verschiedenen Schweregraden der Aortenklappenstenose

Abbildung modifiziert nach Makimoto et al. (2022)

Somit lässt sich festhalten, dass das E1 Modell die signifikant besten Ergebnisse auch bei der klinischen Validierung erzielt und als das Modell für die weitere Verwendung bestimmt wurde.

3.3 Export auf eine Smartphone App

Durch die Implementierung der trainierten Modelle in der App „*AudioClassification*“ wurde eine Anwendung auf dem Smartphone realisiert. Hierbei wurde das Modell E1 verwendet, da es die besten Ergebnisse aufwies. Die App gab die Ergebnisse als binäre Klassifikation zwischen nicht-schwerer Aortenklappenstenose als „0“ und schwerer AS als „1“ an. Zudem wurde für jeden der Auskultationsorte separat eine Vorhersage getroffen.

Die im Rahmen der klinischen Validierung zusätzlich aufgenommenen Datensätze wurden verwendet. Von den 42 Fällen mit schwerer Aortenklappenstenose wurden von der Smartphone Anwendung insgesamt 41 korrekt erkannt, was einer Sensitivität von 97,6% entspricht. Da 85 von 90 Datensätzen ohne schwere Aortenklappenstenose korrekt zugeordnet worden sind, entspricht dies einer Spezifität von 94,4%. Hier zeigte sich zudem eine besonders zuverlässige Erkennung von Datensätzen ohne jegliche Aortenklappenstenose (34 von 34 Fälle, 100%) oder mit nur einer leichtgradigen AS (ebenfalls 100% richtig erkannt). Die moderate AS konnte in 23 von 28 Fällen korrekt zugeordnet werden. Somit ergibt sich insgesamt eine Treffsicherheit von 95,7% sowie ein F1 Wert von 0,932 (Makimoto et al., 2022).

In einem zusätzlichen Schritt wurden die Datensätze der 330 initial exkludierten Patienten ohne Aortenklappenstenose ebenfalls mit in die Beurteilung eingeschlossen. Dadurch wurden die Verhältnisse zwischen Datensätzen mit und ohne AS realitätsnäher. Hier konnten insgesamt 414 von 415 nicht-schweren AS korrekt zuordnet werden (99,8%). Lediglich eine nicht-schwere AS wurde als hochgradig eingestuft.

3.4 Visualisierung des Beurteilungsprozesses der KI

Die Anwendung von KI-Modellen in der Medizin setzt ein grundsätzliches Verständnis dieser voraus. Hierbei steht besonders die Entscheidungsfindung der KI im Mittelpunkt. Damit der Beurteilungsprozess nachvollziehbar wird, wurde die letzte Faltungsschicht der neuronalen Netzwerke, in denen die Entscheidung stattfindet, mittels Grad-CAM++ visualisiert. Hieraus entstand eine *Heatmap*, welche mit den Mel-Spektrogrammen verknüpft wurde. Somit wurden die Bereiche farblich hervorgehoben, auf welchen der Fokus der KI lag. Hierbei stellte sich für die KI vordergründig die Frage, ob es sich um eine schwere AS handelt oder nicht. Es wurden die einzelnen Modelle der Auskultationsorte verwendet (L1-3). Die finale Entscheidung geschah aufgrund der gemittelten Vorhersagen an den drei Auskultationsorten. Dabei wurden bei den Analysen 10 Vorhersagen berechnet und daraus der Mittelwert gebildet. Betrachtet wurden sowohl Beispiele, in denen die Zuordnung der KI korrekt gelang als auch diejenigen, bei denen es zu einem fehlerhaften Ergebnis kam, um daraus Rückschlüsse zur zukünftigen Fehlervermeidung und Optimierung zu erhalten. In den folgenden Abbildungen 19 bis 30 ist jeweils im oberen Abschnitt die Waveform der Daten dargestellt, also eine grafische Aufzeichnung der Herztöne mit Amplitudenausschlag und der Einzeichnung der Herztöne S1 und S2. Darunter befindet sich das daraus erstellte Mel-Spektrogramm (jeweils in Hertz), welchem in der untersten Grafik die *Heatmap* überlagert wurde, anhand dessen die Fokussierung der KI visualisiert wird. Die besonders beachteten Bereiche sind dabei auf der Farbskala in rötlichen Farben gehalten. Zudem sind im oberen Abschnitt die Wahrscheinlichkeiten angegeben, mit denen die KI den vorliegenden Schweregrad der AS bestimmt hat (siehe Abbildung 19 bis 30 P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P nonsAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose). Hier werden Beispiele aus der klinischen Validierungsgruppe gezeigt. Dabei wurde zwischen dem ersten Herzton (S1), dem zweiten Herzton (S2) sowie der systolischen (S1-S2) sowie der diastolischen Phase (S2-S1) unterschieden. Es wurden jeweils vier Sekunden lange Abschnitte betrachtet.

3.4.1 2RSB

Wenn das L1 Modell mit den Auskultationsdaten vom zweiten Intercostalraum parasternal rechts eine schwere Aortenklappenstenose korrekt klassifiziert hatte, dann wird mittels der in Abbildung 19 gezeigten *Heatmap* sichtbar, dass sich die KI vor allem auf das systolische Herzgeräusch sowie den Bereich zwischen dem ersten und zweiten Herzton fokussierte. Die erhöhte neuronale Aktivität wird dabei durch die rötlichen Farbmarkierungen visualisiert (Makimoto et al., 2022).

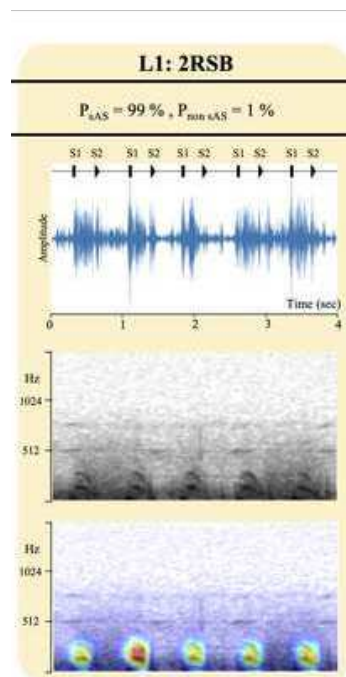


Abb. 19: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* bei korrekt erkannter schwerer AS beim L1 Modell

Die KI gibt als Wahrscheinlichkeit für das Vorhandensein einer schweren Aortenklappenstenose (P sAS) 99% an, hier liegt auch eine schwere AS vor. Anhand der Fokussierungsanalyse mittels des Mel-Spektrogramms lässt sich die Entscheidungsfindung der KI nachvollziehbar machen, hier im Beispiel als Fokussierung vor allem auf die systolische Herzphase

P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P non sAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; Hz: Hertz, sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Tabelle 8: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L1 Modell

	Nicht-schwere Aortenklappenstenose	Schwere Aortenklappenstenose	p-Wert
Fokussierung auf	Fokusscore/Herzschlag	Fokusscore/Herzschlag	
- S1	0,279	0,534	<0,0001
- S1-S2	0,39	0,989	<0,001
- S2	1,011	0,148	<0,001
- S2-S1	0,071	0	0,0021

Tabelle modifiziert nach Makimoto et al. (2022)

Der Fokusscore wurde anhand eines Punktesystems ermittelt. Hier wurden für eine moderate Fokussierung ein Punkt und für eine starke Fokussierung zwei Punkte vergeben. Diese Punkte wurden anschließend durch die Anzahl der Herzschläge pro Aufzeichnung geteilt, wodurch der Fokusscore errechnet wurde. Wie in Tabelle 8 dargestellt, zeigt sich bei der Analyse des Fokusscores pro Herzschlag eine Signifikanz für die Entscheidungsfindung der KI. Wie bereits erwähnt, findet man bei der schweren Aortenklappenstenose eine signifikant stärkere Fokussierung auf die systolische Phase (S1 und S1-2) als bei der nicht-schweren Aortenklappenstenose.

Abbildung 20 visualisiert wiederum exemplarisch die durch die KI korrekt identifizierte leichtgradige Aortenklappenstenose links sowie den ebenfalls korrekt zugeordneten Fall ohne Stenosierung. Bei den als nicht-schwere AS erkannten Auskultationsdaten konzentrierte sich die KI auf die späte systolische Phase und insbesondere den zweiten Herzton (S2, p-Wert <0,001; Makimoto et al., 2022).

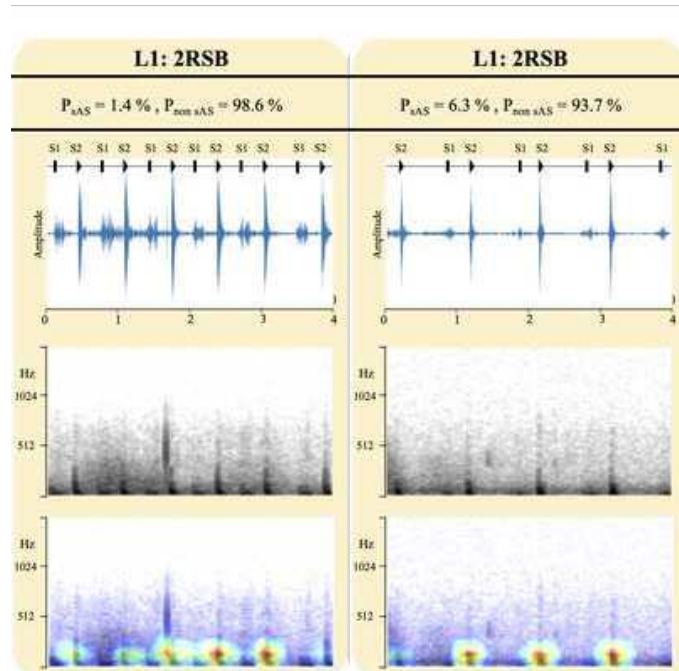


Abb. 20: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* bei korrekt erkannter leichter (links) sowie keiner (rechts) AS beim L1 Modell

In der Abbildung zeigt die KI jeweils nur sehr geringe Wahrscheinlichkeiten für das Vorliegen einer schweren AS an, damit liegt sie richtig, da hier eine leichte beziehungsweise keine AS vorliegt.

Zudem lässt sich die Fokussierung vor allem auf den zweiten Herzton (S2) nachvollziehen.

P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P non sAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; Hz:

Hertz, sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Eine Analyse der vom Modell fehlerhaft klassifizierten Fälle zeigt, dass der Fokus auf denselben Abschnitten lag wie bei den korrekt klassifizierten Fällen. Dementsprechend war dies bei der fälschlicherweise als schwere AS erkannten Aufzeichnung die systolische Phase und bei der inkorrekt als nicht-schwere AS eingestuften Auskultation das zweite Herzgeräusch (Makimoto et al., 2022).

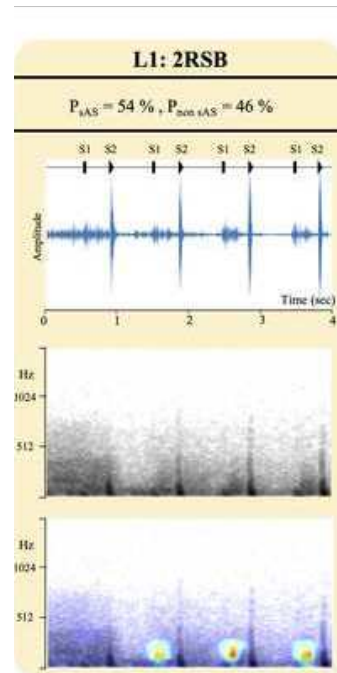


Abb. 21: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit Heatmap als Beispiel bei Schwierigkeiten bei der Bestimmung der schweren Aortenklappenstenose

Es wird eine schwere AS nur mit 54% Wahrscheinlichkeit auch als diese erkannt, es findet vor allem eine Fokussierung auf die spätsystolische Phase und den zweiten Herzton statt
 P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie $P_{non sAS}$ als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; Hz: Hertz, sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Abbildung 21 illustriert ein Beispiel für die Schwierigkeit bei der Bestimmung, da es sich hier um eine schwere AS handelte, diese jedoch nur mit einer Wahrscheinlichkeit von 54% als solche erkannt wurde. Somit lag hier eine Unsicherheit des Modells vor. Die KI fokussierte sich dabei insbesondere auf die späte systolische Phase bis zum zweiten Herzton. Obwohl die schwere Aortenklappenstenose im L1 Modell als solche erkannt wurde, wurde am Ende aufgrund der Mittelung der Daten mit den Modellen L2 und L3 eine nicht-schwere AS diagnostiziert (Makimoto et al., 2022).

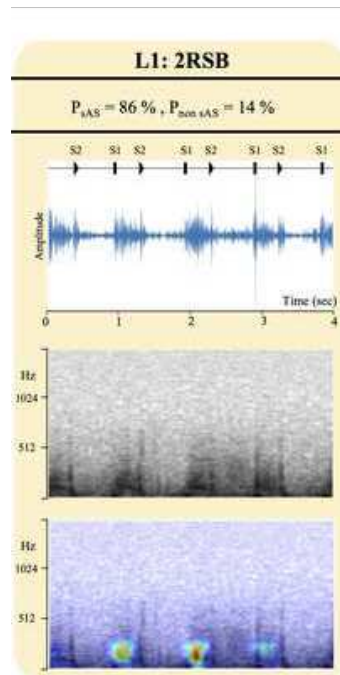


Abb. 22: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* bei einer fälschlicherweise als schwer klassifizierten moderaten Aortenklappenstenose

Schwierigkeiten bei der Bestimmung der moderaten AS, da diese aufgrund der Fokussierung auf die systolische Herzaktion der schweren AS zugeordnet worden ist

P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P non sAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; Hz: Hertz, sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Der Datensatz in Abbildung 22 stammt von einer moderaten Aortenklappenstenose. Von der KI wurde jedoch detektiert, dass es sich mit hoher Wahrscheinlichkeit um eine schwere AS handelt, da der Fokus erneut auf der systolischen Phase der Herzaktion lag. Man kann dabei festhalten, dass es eine Unsicherheit des Modells L1 bei der Unterscheidung zwischen der moderaten und schweren AS gab, da teils moderate Aortenklappenstenosen aufgrund der Fokussierung auf die systolische Phase der Herzaktion dem höchsten Schweregrad zugeordnet wurden (Makimoto et al., 2022).

3.4.2 Erb-Punkt

Das Modell L2 für den Erb-Punkt (dritter Intercostalraum links) ähnelte in der Fokussierung dem Modell L1. Auch hier wurde, wie in Abbildung 23 dargestellt, für die schwere AS der Fokus auf die systolische Phase der Herzaktion gelegt, für die nicht-schwere AS lag er auf dem zweiten Herzton sowie zwischen dem zweiten und ersten Herzton. Wie anhand von Tabelle 9 gezeigt wird, ist die Fokusanalyse am Erb-Punkt jedoch weniger signifikant. Ebenso wie beim L1 Modell ist die Fokussierung auch bei der fehlerhaften Erkennung der Herztöne ähnlich. Der Unterschied zu L1 besteht jedoch darin, dass hier zuverlässiger die moderate von der schweren AS abgegrenzt werden konnte und es diesbezüglich zu weniger Fehleinschätzungen kam. Hierbei fokussierte sich das L2 Modell im Gegensatz zum L1 Modell auf die späte systolische Phase bis zum zweiten Herzton (S2). Die gleiche Fokussierung galt ebenfalls für die leichte und keine AS, wie in Abbildung 24 veranschaulicht wird (Makimoto et al., 2022).

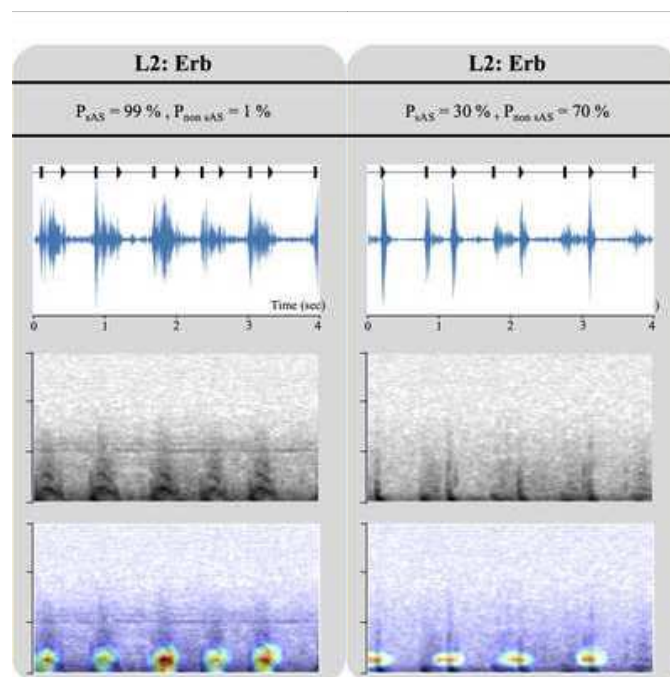


Abb. 23: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit Heatmap bei schwerer (links) und moderater (rechts) AS beim L2 Modell

In beiden Fällen kann die KI die Schweregrade der Aortenklappenstenose durch die korrekte Fokussierung nachvollziehbar zuordnen.

P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P_{non sAS} als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose, sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

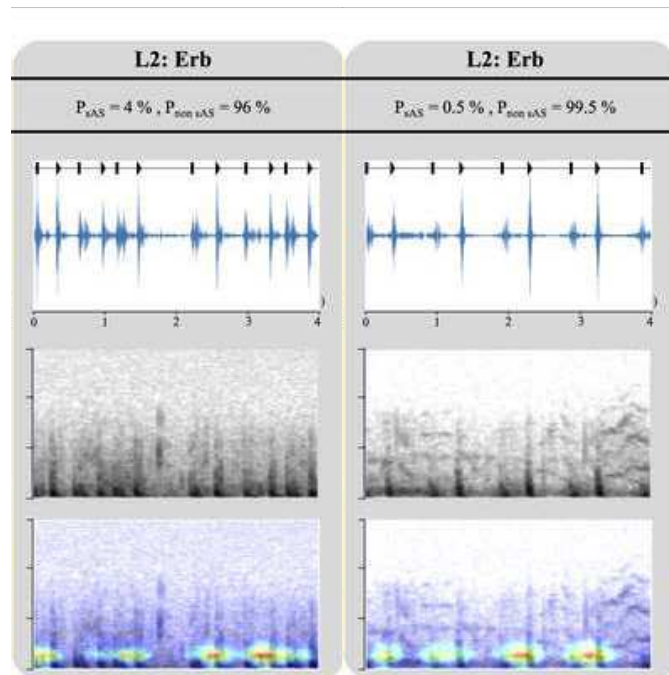


Abb. 24: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* bei leichter (links) sowie keiner (rechts) AS beim L2 Modell

In beiden Fällen kann die KI die Schweregrade der Aortenklappenstenose durch die korrekte Fokussierung nachvollziehbar zuordnen

P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P_{non sAS} als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Tabelle 9: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L2 Modell

	Nicht-schwere Aortenklappenstenose	Schwere Aortenklappenstenose	P- Wert
Fokussierung auf	Fokusscore/Herzschlag	Fokusscore/Herzschlag	
- S1	0,375	0,215	0,014
- S1-S2	1,322	1,283	0,55
- S2	0,832	0,187	<0,001
- S2-S1	0,228	0,011	<0,001

Tabelle modifiziert nach Makimoto et al. (2022)

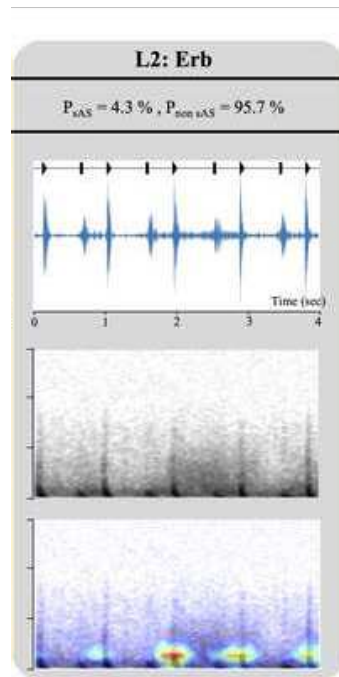


Abb. 25: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* bei fälschlicher Einordnung einer schweren als nicht-schwere Aortenklappenstenose beim L2 Modell

Hier wird mit großer Wahrscheinlichkeit das Vorliegen einer nicht-schweren Aortenklappenstenose präferiert, hier liegt jedoch eine schwere AS vor.

P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P_{non sAS} als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden

Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Mit hoher Wahrscheinlichkeit von 95,7% wurde in Abbildung 25 das Vorliegen einer nicht-schweren AS angegeben. Dieser Datensatz stammte jedoch von einer schweren Aortenklappenstenose. Die Fokussierung lag dabei fälschlicherweise vor allem auf dem zweiten Herzton (Makimoto et al., 2022).

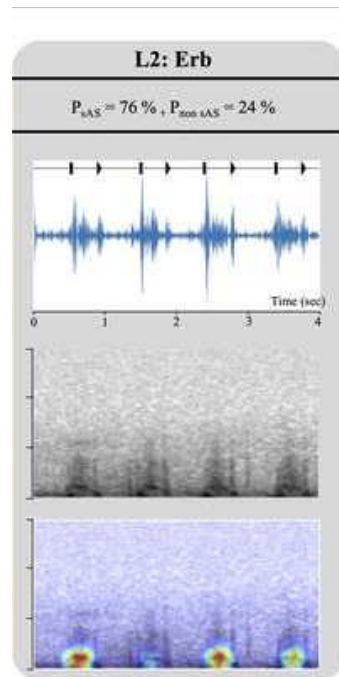


Abb. 26: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* mit fälschlicher Einordnung einer moderaten als schwere Aortenklappenstenose beim L2 Modell

Hier wird mit hoher Wahrscheinlichkeit das Vorliegen einer schweren Aortenklappenstenose präferiert, hier liegt jedoch ein moderater Schweregrad vor

P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P non sAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden

Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Der Fall in Abbildung 26 stellt sich umgekehrt dar, hierbei handelte es sich um eine moderate AS, die KI gab jedoch eine höhere Wahrscheinlichkeit für das Vorliegen einer schweren AS an. Dabei kam es zu einer inkorrekten Fokussierung auf die systolische Phase und damit dem Schluss der KI, dass es sich hier um eine schwere AS handelt (Makimoto et al., 2022).

3.4.3 Apex

Das L3 Modell mit den Auskultationsdaten von der Herzspitze (fünfter Intercostalraum links medioclavicular) zeigte nicht den einen spezifischen Fokus für den Entscheidungsprozess wie er bei den anderen beiden Modellen beobachtet werden konnte. Dies zeigen die Ergebnisse in Tabelle 10.

Tabelle 10: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L3 Modell

	Nicht-schwere Aortenklappenstenose	Schwere Aortenklappenstenose	P-Wert
Fokussierung auf	Fokusscore/Herzschlag	Fokusscore/Herzschlag	
- S1	0,966	0,979	0,82
- S1-S2	1,205	0,551	<0,001
- S2	0,984	0,581	<0,001
- S2-S1	1,185	1,324	0,0054

Tabelle modifiziert nach Makimoto et al. (2022)

Am deutlichsten ist der Fokus bei der korrekten Klassifikation der schweren AS erkennbar, hier ließ sich in den *Heatmaps* teils eine Konzentrierung auf die enddiastolische Phase vor dem ersten Herzton finden. Bei der nicht-schweren AS wurde eine große Bandbreite an Foki festgestellt, die zudem zwischen den verschiedenen Auskultationsergebnissen eine Variabilität aufwies. Trotz dieser variablen und nicht klar nachvollziehbaren Fokussierung konnte L3 ebenfalls sehr zuverlässig zwischen den Schweregraden der AS differenzieren, wie die Abbildungen 27 und 28 zeigen (Makimoto et al., 2022).

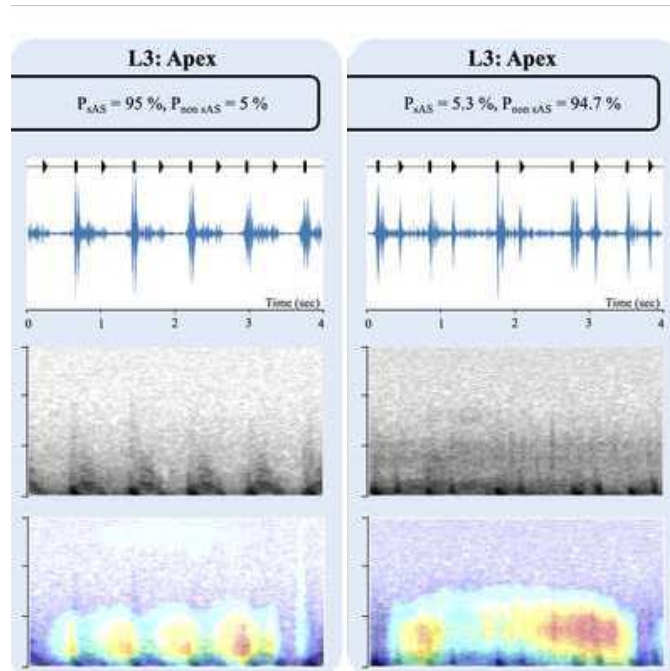


Abb. 27: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit Heatmap mit korrekter Erkennung der schweren (links) und moderaten AS (rechts) beim L3 Modell

In beiden Fällen wird der Schweregrad korrekt zugeordnet, bei der schweren AS lässt sich eine Fokussierung auf die späte diastolische Herzaktion vor dem ersten Herzton erkennen, bei der moderaten AS ist eine Fokussierung nicht eindeutig auszumachen. P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie $P_{non sAS}$ als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden
Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

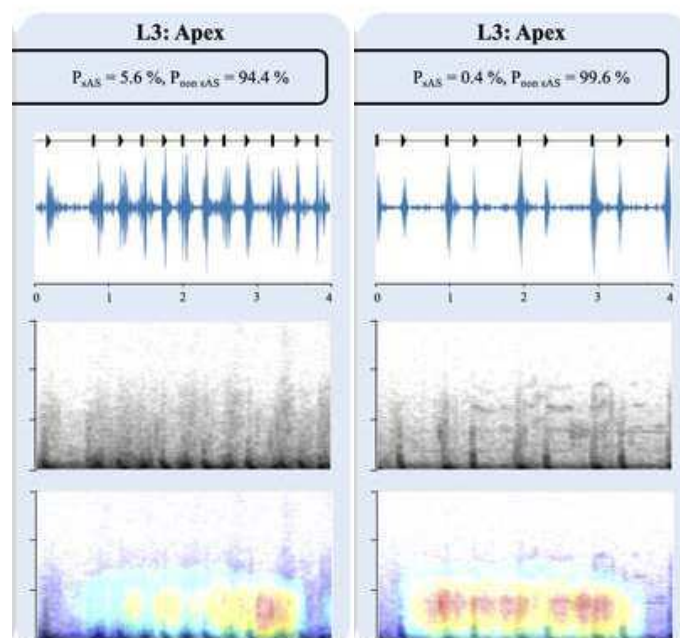


Abb. 28: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit Heatmap mit korrekter Erkennung der leichten (links) und keiner AS (rechts) beim L3 Modell

In beiden Fällen wird der Schweregrad korrekt zugeordnet. P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie $P_{non sAS}$ als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden
Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

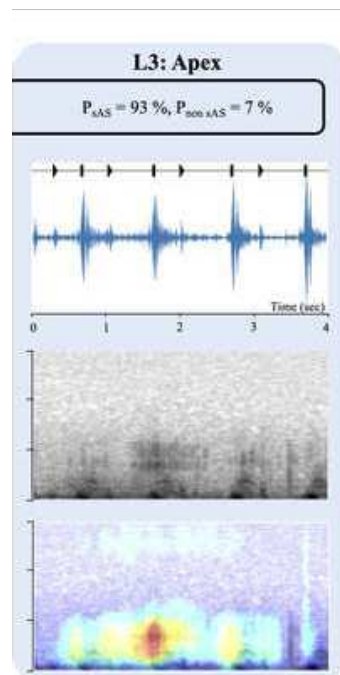


Abb. 29: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* mit fälschlicher Erkennung einer moderaten als schwere AS beim L3 Modell

Hier wird durch die KI mit großer Wahrscheinlichkeit das Vorliegen einer schweren AS angegeben, jedoch handelt es sich um eine moderate AS.

P_{sAS} in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie $P_{non\ sAS}$ als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden

Abbildung publiziert im European Heart Journal - Digital Health (Makimoto et al., 2022)

Die in Abbildung 29 dargestellte fälschliche Erkennung der moderaten als schwere AS lässt eher wenig Rückschlüsse zu. Hier kam es am ehesten zu einer Konzentration auf die enddiastolischen Bereiche vor dem ersten Herzton bis zur systolischen Herzaktion. Jedoch ist die genaue Fokussierung in der *Heatmap* schwer nachzuvollziehen (Makimoto et al., 2022).

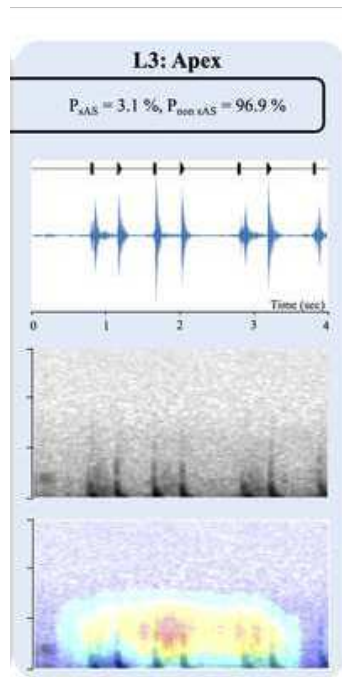


Abb. 30: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit *Heatmap* mit fälschlicher Erkennung einer schweren AS als nicht-schwere AS beim L3 Modell

Hier wird durch die KI mit großer Wahrscheinlichkeit das Vorliegen einer nicht-schweren AS angegeben, jedoch handelt es sich um eine schwere AS. P sAS in % als Vorhersage der Wahrscheinlichkeit einer schweren Aortenklappenstenose sowie P non sAS als Vorhersage der Wahrscheinlichkeit einer nicht-schweren Aortenklappenstenose; sec: Sekunden
Abbildung publiziert im European Heart Journal- Digital Health (Makimoto et al., 2022)

Bei der in Abbildung 30 als nicht-schwer klassifizierten Aortenklappenstenose liegt ursprünglich eine schwere Form vor. Hier lässt sich keine genaue Analyse der Fokussierung durchführen, da diese sehr breit gefasst war und somit die Transparenz einschränkte (Makimoto et al., 2022).

Zusammenfassend lässt sich sagen, dass die Visualisierung der Entscheidungen besonders bei den Modellen L1 und L2 gut nachvollziehbar ist, während das Modell L3 eine breitere Fokussierung aufweist. Doch trotzdem kann auch dieses Modell sehr zuverlässig korrekte Entscheidungen treffen. Auch anhand der falsch zugeordneten Fälle lassen sich Rückschlüsse ziehen.

4 Diskussion

4.1 Ergebnisse dieser Studie

Das Ziel dieser Arbeit war es, eine Künstliche Intelligenz zu entwickeln, die anhand von aufgezeichneten Herztönen realer Patienten die Erkennung und Zuordnung verschiedener Schweregrade der Aortenklappenstenose erlernen kann. Hierfür wurden die Herztöne von Patienten mit und ohne Aortenklappenstenose digital aufgenommen, während ergänzend eine Echokardiographie zur Evaluation des Schweregrades des Herzklappenvitiums durchgeführt wurde. Auf Basis dieser Daten konnte die Künstliche Intelligenz die akustischen Charakteristika der verschiedenen Schweregrade der Aortenklappenstenose erlernen. Dabei wurden zur Auskultation verschiedene Orte verwendet und diese primär einzeln, im Verlauf jedoch auch kombiniert getestet und trainiert, um anschließend die bestmögliche Leistung zu erzielen.

Zusammenfassend zeigte sich hierbei, dass der Künstlichen Intelligenz die zuverlässige Detektion von Aortenklappenstenosen gelingt. Dabei korrelierte die Treffsicherheit in Hinblick auf den Schweregrad der AS insbesondere damit, an welcher Lokalisation die Herztöne auskultiert und aufgezeichnet wurden. Insgesamt wurden im Rahmen dieser Untersuchung drei Auskultationsstellen verwendet. Dies waren der zweite Intercostalraum rechts parasternal (2RSB), der dritte Intercostalraum links parasternal (Erb-Punkt) sowie der fünfte Intercostalraum links medioclavicular (Apex). Für jeden der Auskultationsorte wurden mehrere neuronale Netzwerke getestet und die Ergebnisse in Hinblick auf die Sensitivität, Spezifität, Treffsicherheit, *area under the curve*, den F1 Wert, den positiv und negativ prädiktiven Wert sowie die Verlustfunktion verglichen.

Von den drei Lokalisationen erzielte das Modell L1, welches mit den Aufzeichnungen des ersten Auskultationspunktes (2RSB) trainiert wurde, die signifikant besten Ergebnisse (p-Wert <0,001 beim Vergleich der F1 Werte, der Treffsicherheit und der *area under the curve*). Hier ließ sich zudem eine Gesamtreffsicherheit von über 90% erreichen, insbesondere zeigte sich das

CNN10L-L1 Modell im Bereich der Spezifität mit 95,7% sowie dem positiv prädiktiven Wert mit 93% (jeweils p-Wert $<0,0001$) deutlich überlegen.

Doch auch die anderen Modelle erreichten überzeugende Ergebnisse. Das CNN10L-T Modell imponierte im Besonderen bei den Auskultationsdaten des Erb-Punktes, im Bereich des Apex zeigte das cCNN-T die signifikant besten Ergebnisse.

Bei der differenzierten Analyse der Treffsicherheit der Modelle L1-L3 fiel auf, dass insgesamt ein Ergebnis von 87% beim L1 Modell (2RSB), 93% beim Modell L2 (Erb) und 88% beim Modell L3 (Apex) erreicht wurde. Dies bedeutet, dass der genaue Schweregrad der Aortenklappenstenose korrekt zugeordnet werden konnte. Hierbei wurde weiterhin deutlich, dass die Entscheidungsfindung der Künstlichen Intelligenz insbesondere dann gelingt, wenn keine beziehungsweise eine leichte oder eine schwergradige Aortenklappenstenose vorliegt, während bei der moderaten AS etwas häufiger Unsicherheiten auftraten.

Um die Erkennungsfähigkeit im nächsten Schritt zu verbessern, wurden die Modelle von den einzelnen Auskultationsorten in unterschiedlichen Gewichtungen kombiniert. Hieraus entstanden die zusammengesetzten Ensemble-Modelle E1-3. So konnte die Treffsicherheit der Künstlichen Intelligenz bei der Erkennung der schweren Aortenklappenstenose beim E1 Modell auf über 96% gesteigert werden, auch insgesamt über alle Schweregrade gemittelt lag die Treffsicherheit mit 94% beim E1 Modell signifikant am höchsten (p-Wert $<0,001$). So wurde die Leistung im Rahmen einer klinischen Validierung mit 132 zusätzlich rekrutierten, prospektiv aufgenommenen Datensätzen erneut getestet. Hierbei gelang ebenfalls eine äußerst zuverlässige Detektion der verschiedenen Schweregrade der Aortenklappenstenose mit Hilfe der neuronalen Netzwerke. Beispielsweise konnte mit dem E1 Modell eine Gesamttreffsicherheit von über 95% erreicht werden, mit Ausnahme des moderaten Schweregrades lagen alle Werte bei über 90%. Daher wurde dieses Modell als Grundlage für die Entwicklung einer Smartphone App verwendet, um die Ergebnisse weiter zu evaluieren.

Zudem konnte verdeutlicht werden, dass die Entscheidungen der Künstlichen Intelligenz für den Menschen mithilfe von Anwendungen wie Grad-CAM++ nachvollziehbar werden, wenn sie in *Heatmaps* und Mel-Spektrogrammen für das menschliche Auge sichtbar und somit transparent dargestellt wurden. Dadurch konnte individuell festgelegt werden, auf welche spezifischen Bereiche der Herztöne sich die Künstliche Intelligenz fokussiert und damit deren Beurteilungskriterien bestimmen. Beispielsweise imponierte bei den Daten des zweiten Intercostalraums rechts, dass sich das neuronale Netzwerk bei schweren Aortenklappenstenosen vor allem auf systolische Herzgeräusche fokussierte, während es bei der leichten oder keinen Stenosierung eine Konzentration auf die späte systolische Phase und insbesondere auf den zweiten Herzton gab. Hier konnten signifikante Ergebnisse bei der Analyse der Fokussierungen erreicht werden. Auch durch die Interpretation von fehlerhaften Entscheidungen der KI wurden Rückschlüsse gezogen.

Zusammenfassend lässt sich festhalten, dass es mit Hilfe der hier trainierten neuronalen Netzwerke möglich ist, ein effizientes Screening auf Aortenklappenstenosen anhand einer verständlichen Künstlichen Intelligenz durchzuführen.

Weiterhin wird deutlich, dass die Inklusion aller Auskultationsorte eine entscheidende Bedeutung hat, da in einigen Fällen der Schweregrad in einem Modell falsch zugeordnet worden ist, dies durch die Mittelung der Ergebnisse von allen Auskultationsorten ausgeglichen werden konnte.

4.2 Kritische Evaluation der Methoden

Der Hintergrund dieser Studie war es mithilfe von aufgezeichneten Auskultationsdaten eine KI zu trainieren, die in der Lage ist, den Schweregrad einer Aortenklappenstenose zuverlässig zu detektieren. Dies geschah mit Hilfe von additiv durchgeführten transthorakalen Echokardiographien, die als Trainings- und Leistungskontrolle dienten, da sie für die Schweregradbestimmung der AS den Goldstandard darstellen.

Die Auskultation zur Diagnostik von Herzklappenerkrankungen ist schon lange bekannt und wird viel verwendet (Montinari und Minelli, 2019). Positiv hervorzuheben sind hierbei vor allem die schnelle Durchführbarkeit und die geringe Belastung der Patienten durch die strahlen- und schmerzfreie Untersuchung. Zudem können erfahrene Ärzte, die schon viele Auskultationen durchgeführt haben, hierbei sehr valide Ergebnisse erzielen (McGee, 2010). Nachteile sind die große Untersucherabhängigkeit sowie die notwendige Erfahrung in Bezug auf die Interpretation der Ergebnisse. In Untersuchungen wie von Gardezi et al. (2018) wurde gezeigt, dass die Auskultation eher eine geringe Treffsicherheit bei der Detektion von Herzklappenvitien aufweist, wenn sie als alleiniges Tool verwendet wird. Teilweise wird nur eine Sensitivität von 70% selbst durch erfahrene Anwender erreicht, wenn sie durch keine weitere Diagnostik ergänzt wird (Roldan et al., 1996).

Doch die in dieser Untersuchung trainierte Künstliche Intelligenz hat durch die Auskultationsdaten gelernt und dabei dennoch sehr gute Ergebnisse erzielt. Besonders hervorheben lässt sich dabei die zuverlässige Detektion der schweren sowie nicht vorhandenen und leichten Aortenklappenstenose. Daraus kann geschlossen werden, dass die Unterscheidungen für das menschliche Gehör sehr viel schwieriger sind als für eine Künstliche Intelligenz. Diese kann sich genauer auf die wesentlichen Abschnitte fokussieren und dieses immer wieder rekapitulieren, was sich anhand der Visualisierungen mittels Grad-CAM++ nachvollziehen lässt.

Bezüglich des Trainings und der Leistungskontrolle mit den erhobenen Daten der Echokardiographie ergaben sich mitunter Schwierigkeiten. Die Resultate aus dieser Ultraschalluntersuchung dienten dabei als Grundlage für den Lernprozess der KI.

Gerade bei adipösen Patienten ist die Schallqualität mittels Echokardiographie teils eingeschränkt, sodass die Bestimmung der Klappenvitien nicht immer ganz trivial erscheint, weil die Klappen erschwert darstellbar sind. Zudem befinden sich zuweilen Patienten zwischen zwei Schweregraden der Aortenklappenstenose, bei denen echokardiographisch die Unterscheidung schwerfällt und teilweise nicht eindeutig möglich ist.

Ferner gibt es zusätzlich andere Erkrankungen, die für Artefakte bei den Herzgeräuschen sorgen, die denen der schweren Aortenklappenstenose ähneln. Dazu zählt beispielsweise die Obstruktion des linksventrikulären Ausflusstraktes. Diese wurden hier jedoch nicht gesondert erfasst. Dies könnte daher als weiterführendes Forschungsthema in nachfolgenden Studien analysiert werden.

Dem gegenüber stehen die wesentlichen Vorteile dieser Studie. Einerseits wurden die Daten real am Menschen im klinischen Alltag aufgenommen, wodurch die Ergebnisse zukünftig auf die Allgemeinbevölkerung im Rahmen eines Screenings übertragbar sind. Die KI hat somit direkt gelernt mit Störgeräuschen wie Stimmen im Hintergrund oder der Atmung umzugehen, da diese bei einer späteren Verwendung in der Praxis ebenfalls auftreten werden. Es wurde insgesamt kein Datensatz aufgrund der Tonqualität aussortiert. Gerade störende Atemgeräusche sind dabei kaum zu vermeiden, da vor allem die älteren und schwerer betroffenen Patienten selten in der Lage sind, für die gesamte Aufzeichnungsdauer der Herztöne von 15 Sekunden die Luft anzuhalten. Wie bereits diskutiert sind jedoch gerade die Patienten, die unter einer hochgradigen Aortenklappenstenose leiden, oft der NYHA Klasse III oder IV zuzuordnen.

In ähnlichen Studien wie der von Chorba et al. (2021) wurden durch die große zur Verfügung stehende Datenmenge alle uneindeutigen oder zu sehr durch Störgeräusche verunreinigten Aufzeichnungen aussortiert. Somit konnten mit der dort verwendeten KI sehr gute Ergebnisse bezüglich der korrekten Zuordnung der Herzgeräusche erzielt werden. Wenn dann jedoch die qualitativ schlechteren Aufzeichnungen mit einbezogen wurden, wurde die Leistung jener KI deutlich unzuverlässiger (Chorba et al., 2021). Somit lassen sich mit reinen Tönen für das

Training vorübergehend die besseren Ergebnisse erzielen, dies halten jedoch schwer der Überprüfung unter realen Bedingungen stand, wie es im Rahmen dieser Studie der Fall war.

Zum anderen wurde die hier vorgestellte KI mit allen Schweregraden der Aortenklappenstenose trainiert, sie lernte ebenfalls die leichte oder moderate Form von der schweren AS abzugrenzen. Diese Unterscheidung bringt zusätzlich einen großen Vorteil, da es in der klinischen Praxis einen bedeutenden Unterschied in Hinblick auf Therapie und Lebenserwartung macht, welchen Schweregrad der Stenosierung die Aortenklappe aufweist (Fortuni und Delgado, 2020). Erneut kann an dieser Stelle der Vergleich zu der Studie von Chorba et al. (2021) gezogen werden, in welcher es insbesondere um die Unterscheidung zwischen relevanter (in dem Fall wurde dazu die moderate oder schwere AS gezählt) und keiner AS ging. Diese Studie ist somit als Weiterentwicklung schon publizierter Untersuchungen wie der von Chorba et al. (2021) zu verstehen, indem die KI die Schweregrade der Aortenklappenstenose genauer differenzieren kann. Zudem konnten selbst alle *low-flow low-gradient* Stenosen in der klinischen Validierungsgruppe korrekt zugeordnet werden.

Weiterhin sind in dieser Untersuchung Patienten mit zusätzlichen Klappenvitien, allen voran der Mitral- und Trikuspidalklappeninsuffizienz miteingeschlossen worden. Diese Erkrankungen können ebenfalls systolische Herzgeräusche hervorrufen (Salazar et al., 2012). Durch die Koexistenz anderer Vitien wird die Beurteilung der Aortenklappenstenose durch mögliche akustische Verfälschungen schwieriger. Aber auch dies ist sehr realitätsnah, da viele Patienten neben einer Aortenklappenstenose noch andere Herzklappenerkrankungen aufweisen und somit die Unterscheidung der verschiedenen Herzgeräusche eine essenzielle Voraussetzung für den späteren alltäglichen klinischen Einsatz als Screeningtool ist.

4.3 Kritische Evaluation der Ergebnisse

Insgesamt haben die Ergebnisse gezeigt, dass es möglich ist, eine Künstliche Intelligenz für das Screening von Aortenklappenstenosen zu entwickeln, welche die Diagnostik im klinischen Alltag erleichtern kann.

Anhand der Datenanalyse wurde deutlich, dass die Auskultation an verschiedenen Lokalisationen sinnvoll ist. In dieser Studie wurden bei den einzelnen Modellen L1-3 die besten Ergebnisse bei den Datensätzen des zweiten Intercostalraums rechts parasternal (2RSB) erzielt. Dies ist insofern nachvollziehbar, da hier ebenso bei der routinemäßigen Auskultation der Aortenklappe in der Praxis das Hauptaugenmerk liegt. Doch durch die zusätzliche Integration der anderen Auskultationsorten in das Modell wurde die Treffsicherheit wie auch die Sensitivität und Spezifität weiter verbessert, was sich in den Resultaten der zusammengesetzten Modelle E1-3 erkennen lässt. Dies ist mit der klinischen Praxis zu vergleichen, da hier ebenfalls die anderen Auskultationsorte in die Untersuchung mit einbezogen werden.

Diese Kombination ist demnach entscheidend für die Weiterentwicklung und das Verständnis der KI als hilfreiches Screeninginstrument. Dafür ist eine möglichst hohe Sensitivität wichtig, damit viele Erkrankte durch die Künstliche Intelligenz erkannt werden. Diese können im nächsten Schritt dann weiterer Diagnostik zugeführt werden. Hierfür ist es von Bedeutung, dass mit dem Screening so wenig betroffene Patienten wie möglich übersehen werden.

Neben den Leistungen, welche die KI bei der Zuordnung der Schweregrade der AS erreicht, war die Visualisierung der Entscheidungsfindung ein Ziel. Hier gelang es, die Fokussierung zu analysieren und somit den Beurteilungsprozess nachvollziehbar zu gestalten. Dies geschah mithilfe der Umwandlung der Herztöne in Mel-Spektrogramme, welchen *Heatmaps* überlagert wurden. Dadurch konnte der Fokus der KI bei der Beurteilung analysiert werden.

Hierbei wurde sichtbar, dass sich die KI bei der Erkennung der schweren AS vor allem auf das systolische Herzgeräusch fokussierte. Wenn jedoch eine nicht-schwere AS erfasst wurde, dann lag der Fokus eher auf dem zweiten Herzton (S2).

Dies galt sowohl bei der Auskultation des zweiten Intercostalraums rechts parasternal (2RSB, L1) als auch beim Erb-Punkt (L2). Dabei ließ sich feststellen, dass die Fokussierung der KI vor allem durch die unterschiedlichen Architekturen der neuronalen Netzwerke erklärbar wird.

Wie schon lange bekannt ist, sind die besonderen Charakteristika der schweren Aortenklappenstenose vor allem das raue Systolikum mit Punctum maximum im zweiten Intercostalraum rechts parasternal, die Abschwächung des zweiten Herztons und die Fortleitung des Geräusches in die Carotiden (Billig et al., 2024). Die Fortleitung wurde in dieser Studie nicht mit einbezogen, aber die anderen beiden Charakteristika lassen sich hier genau so finden. Es ist davon auszugehen, dass es während des Trainings zu einem Lernprozess gekommen ist. Die KI hat erkannt, dass diese Bereiche bei der schweren AS besonders verändert sind im Vergleich zur Kontrollgruppe ohne AS und hat daraus eigenständig Rückschlüsse für die Verarbeitung der nächsten Datensätze gezogen. Genau das ist die Stärke der KI, da sie selbst erkennt, worauf sie sich fokussieren muss, ohne dass dies programmiert werden muss. Es ist davon auszugehen, dass die KI durch das Training mit den Auskultationsdatensätzen immer zuverlässiger geworden ist, wie es im Rahmen von Lernprozessen von Menschen erwartet wird.

Beim L3 Modell hingegen wurde ein neuronales Netzwerk verwendet, welches eine Kombination aus mehreren Kernen aufweist. Hierbei ließ sich feststellen, dass der Fokus nicht so lokalisiert ausgewiesen werden konnte wie bei den Modellen L1 und L2. Insgesamt bezog sich das Modell auf einen recht großen Bereich. Trotzdem konnte das Modell ebenfalls zuverlässig zwischen schwerer und nicht-schwerer Aortenklappenstenose unterscheiden. Hier ist in Zukunft zu evaluieren, welche Architektur der neuronalen Netzwerke für weitere Untersuchungen die beste ist.

Weiterführend stellt sich die Frage, wie die Detektion der moderaten Aortenklappenstenose in Zukunft verbessert werden kann, da diese durch die KI in Abgrenzung sowohl zur schweren als auch zur leichten oder keinen AS am schwierigsten zu detektieren war und die Treffsicherheit hier noch ausbaufähig ist.

4.4 Klinische Bedeutung dieser Studie und zukünftige Entwicklungen

Die Aortenklappenstenose wird aufgrund der zunehmenden Alterung der Gesellschaft eine immer höhere Prävalenz in der Gesellschaft aufweisen (Thaden et al., 2014). Dabei sind vor allem die frühe Erkennung und daraus resultierende Therapie entscheidend. In Zeiten immer knapper und wertvoller werdenden menschlichen Ressourcen sind dabei Hilfsmittel unabdingbar, um die Arbeitskräfte zu entlasten. Dabei kann und wird die Künstliche Intelligenz einen großen Stellenwert einnehmen.

Die hier vorgestellte KI hat das Potential Patienten herauszufiltern, die bisher nicht entdeckte Aortenklappenstenosen haben und diesen zu einem schnelleren Therapiebeginn verhelfen. Dies kann insbesondere für Betroffene ohne ausgeprägte Symptomatik hilfreich sein. Bei allen eingeschlossenen Studienteilnehmern wurde der NYHA-Score abgefragt. Mithilfe dieses Scores kann abgeschätzt werden, welche Patienten wie stark durch ihre kardiale Erkrankung im Alltag eingeschränkt sind (Caraballo et al., 2019). Wenn nun die Daten aus der klinischen Validierungsgruppe berücksichtigt werden, ergeben sich insgesamt 132 Datensätze. Davon waren 42 Datensätze von Betroffenen mit schwerer Aortenklappenstenose. Dabei hatten zusammengekommen 15 Patienten keine (NYHA I, n=4) oder nur leichte Symptome (NYHA II, n=11). Bei diesen Betroffenen besteht die Gefahr, dass sie aufgrund fehlender Krankheitszeichen zu wenig weiterführende Diagnostik erhalten. Bei all diesen 15 Patienten hat die KI jedoch zuverlässig die schwere Aortenklappenstenose erkannt. Somit könnte in Zukunft deutlich schneller eine Diagnose gestellt und damit auch eine Therapie eingeleitet werden, da die hochgradige Aortenklappenstenose eine schlechte Prognose aufweist.

Weiterhin ist die KI als Screeningtool ebenfalls im nicht-ärztlichen Bereich anwendbar, wie beispielsweise im Rettungsdienst. Hierbei kann die KI bei der primären Diagnostik hilfreich sein, schon bei Patienten zu Hause oder auf dem Weg

ins Krankenhaus, beispielsweise bei kardialer Dekompensation oder unklaren Synkopen.

Neben der Aortenklappenstenose ist die Mitralklappeninsuffizienz ebenfalls eine häufig auftretende Erkrankung, welche in fortgeschrittenen Stadien ebenfalls eine sehr schlechte Prognose mit sich bringt (Enriquez-Sarano et al., 2009). Daher sollte in Folgeuntersuchungen der Fokus noch mehr auf die Mitralklappe gelegt werden. Hier wurde schon in vorherigen Studien wie von Chorba et al. (2021) gezeigt, dass die Auskultation der Mitralklappe mithilfe eines digitalen Stethoskops und anschließende Analyse mit einer Künstlichen Intelligenz nicht so valide Ergebnisse hervorbringt wie die der Aortenklappe. Ein Ziel wäre es hier die Resultate weiter zu verbessern.

Ferner könnte untersucht werden, inwiefern die Kombination mit anderen Diagnoseverfahren wie beispielsweise dem EKG die Detektion von Herzklappenerkrankungen verbessern könnte. Im EKG lassen sich spezifische Parameter finden, die Hinweise auf eine Herzklappenerkrankung liefern können. Hierzu zählen beispielsweise ein Links- oder Rechtsschenkelblock sowie eine linksventrikuläre Hypertrophie (Kwon et al., 2020). Untersuchungen von Cohen-Shelly et al. (2021) oder Kwon et al. (2020) haben mittlerweile gezeigt, dass es möglich ist, anhand von aufgenommen EKGs Hinweise auf eine Aortenklappenstenose zu erhalten. Denkbar wäre, dass zukünftig das EKG und die aufgezeichneten Herztöne kombiniert werden, um die Leistungsfähigkeit zu verbessern.

Weiterhin sollten Überlegungen angestellt werden bezüglich der Architektur der CNNs. Hier sind in den Modellen L1-3 unterschiedliche Architekturen verwendet worden. Somit sollte weiter analysiert werden, welche Architektur die besten Ergebnisse erzielen kann.

Parallel stellt sich zudem die Frage nach den ethischen Dimensionen einer KI, besonders im Gesundheitssektor. Neben den bereits hervorgehobenen Vorteilen in Bezug auf die ubiquitäre Verfügbarkeit und einfache Anwendung ergeben sich daher ebenfalls kritische Aspekte. Es müssen geltende Regeln für den Umgang mit

der KI definiert werden. Problematisch wird häufig der Datenschutz angesehen (Radanliev et al., 2024). Ferner stellt sich die Frage nach der Haftbarkeit von Ärzten durch mögliche Fehlentscheidungen der KI (Kotter und Pinto dos Santos, 2024). Daher sind Grundbedingungen unbedingt einzuhalten, hierzu gehört neben einer transparenten Verständlichkeit vor allem die Wahrung der Autonomie der Patienten sowie ein Verantwortungsbewusstsein der mit der KI arbeitenden Personen (Küster und Schultz, 2023).

4.5 Limitationen

Insgesamt lässt sich nun festhalten, dass die Studie einige Limitationen in ihrer Aussagekraft aufweist. Einerseits ergibt sich das Problem der eher geringen Patientenzahlen. Die KI benötigt gerade für das Training viele Datensätze, mit denen die Leistung immer wieder verbessert werden kann. Da das Team für die Aufnahme der Patienten in die Studie aus geschulten Ärzten und Medizinstudierenden bestand, waren die Kapazitäten dieser begrenzt. Dies war jedoch wichtig, um den hohen Standard bei der Datenerfassung beizubehalten. Zudem waren gerade die Zahlen von Betroffenen mit Herzklappenvitien geringer als diejenigen ohne, sodass deutlich mehr Patienten ohne schwere Aortenklappenstenose in die Studie aufgenommen wurden.

Aufgrund dieser kleineren Anzahl an eingeschlossenen Patienten konnte die Leistungsberechnung nur hypothetisch geschehen. Wenn noch größere Datenmengen zur Verfügung stehen, dann würde sich die Leistung des Modells weiter verbessert.

Ferner lässt sich anmerken, dass insgesamt nur Aufzeichnungen von drei Auskultationsorte genutzt wurden. Wenn zusätzlich weitere Lokalisationen verwendet werden, könnten sich hierdurch ebenfalls die Ergebnisse verbessern.

Zudem gab es teils Datensätze, die echokardiographisch zwischen zwei Schweregraden der Aortenklappenstenose eingeordnet wurden, hier ergab sich das Problem der korrekten Zuordnung.

Weiterhin sollte nicht außer Acht gelassen werden, dass die KI nur die reinen Auskultationsdaten auswertet, aber keinerlei Aussagekraft über den klinischen Zustand der Patienten gibt. Die KI kann hier als Screeningtool dienen, aber die ärztliche Kompetenz in Hinblick auf das ganzheitliche Betrachten der Menschen nicht ersetzen. Zudem können mithilfe eines transthorakalen Ultraschalls beispielsweise zusätzlich die Pumpfunktion oder die Kontraktilität des Herzens bestimmt werden. Dies ist durch die Künstliche Intelligenz in Form eines Modells für eine Smartphone App nicht möglich. Daher ist es wie schon erwähnt vor allem als Methode für die ersten Schritte der Diagnostik und Unterstützung im klinischen Alltag beispielsweise im Rahmen einer digitalisierten Telemedizin zu verstehen. Die ärztliche Tätigkeit kann es in keinerlei Hinsicht ersetzen.

5 Zusammenfassung und Schlussfolgerung

Zusammenfassend lässt sich feststellen, dass die Technologie der Künstlichen Intelligenz im Bereich der Erkennung von Aortenklappenstenosen mithilfe der Auskultation sehr vielversprechend ist. Durch die einfache Anwendbarkeit der KI auf einem Smartphone kann es in den klinischen Alltag als Screeningtool integriert werden, insbesondere um die Effizienz zu erhöhen. So können gerade die Patienten, bei denen ein relevantes Herzklappenvitium vermutet wird, weiterer Diagnostik zugeführt werden. Weitere Vorteile der Künstlichen Intelligenz sind zudem die ständige Bereitschaft der Technologie, welche allzeit verfügbar ist. Somit kann sich die KI in Zukunft als ideale Ergänzung zur ärztlichen Tätigkeit im medizinischen Sektor durchsetzen.

Als weiteren wichtigen Anwendungsbereich in einer digitalen und vernetzten Welt kann man sich die Telemedizin vorstellen. In dieser Untersuchung wurde schon der Export des KI-Modells auf eine Smartphone App durchgeführt. Wenn dieses Modell für Patienten als Tool für zu Hause zur Verfügung gestellt wird, dann könnte dies in einer telemedizinischen Sprechstunde genutzt werden, in welcher sich Arzt und Patient nicht persönlich an einem Ort treffen, sondern digital kommunizieren.

In Zukunft wird die Weiterentwicklung der Künstlichen Intelligenz in der gesamten Medizin, aber auch speziell im Bereich der Detektion von Herzklappenvitien im Vordergrund stehen. Dafür könnte die KI mit anderen Klappenvitien trainiert werden und diese mit in die Untersuchungen einbeziehen. Eine andere Möglichkeit wäre die Verknüpfung mit EKGs oder klinischen Untersuchungsmethoden wie Laborparametern.

Die Künstliche Intelligenz hat das Potential, die Medizin, aber auch weite Teile unserer Welt nachhaltig zu verändern. Es gibt viele Studien und Erfolgsmeldungen zu Anwendungen der KI in den letzten Jahren. Viele Errungenschaften und Fortschritte wären ohne diese Technologie gar nicht möglich gewesen. Wenn man diese als Chance und Unterstützung begreift, dann bieten sich dadurch unzählige neue Möglichkeiten.

Gleichzeitig sollte dabei bedacht werden, dass es die Idee dieser Technologie schon seit vielen Jahren gibt, diese sich in der Medizin trotzdem bisher noch nicht vollständig durchgesetzt hat. Gerade im Bereich der Behandlung von Patienten sind sich Ärzte oft unsicher bezüglich der Sicherheit neuer Methoden und vertrauen auf altbekannte Verfahren. Doch auch hier lassen sich durch Studien wie dieser wichtige Fortschritte erzielen, da die Ergebnisse mithilfe der Analyse der Fokussierung der Künstlichen Intelligenz sichtbar und nachvollziehbar geworden sind.

Durch dieses Verständnis kann es gelingen, dass die Künstliche Intelligenz mehr Anerkennung erfährt, da die Funktionsweise transparent wird. Dies kann zu einem stärkeren Aufschwung der Technologie beitragen, weil es für eine Entlastung von Arbeitskräften in den Krankenhäusern und in der ganzen Medizin sorgen kann. Gerade aktuell erlebt die Künstliche Intelligenz einen so massiven Aufschwung wie noch nie zuvor. Man darf auf die weiteren Entwicklungen sehr gespannt sein, da das Potenzial gewaltig ist.

6 Literatur- und Quellenverzeichnis

1. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Family Med Prim Care*. 2019 Jul;8(7):2328-2331. doi: 10.4103/jfmprc.jfmprc_440_19. PMID: 31463251; PMCID: PMC6691444.
2. Aronow WS. Valvular aortic stenosis in the elderly. *Cardiol Rev*. 2007 Sep-Oct;15(5):217-25. doi: 10.1097/CRD.0b013e31805f6796. PMID: 17700380.
3. Bardon KM, Garelnabi M. The impact of altered mechanobiology on aortic valve pathophysiology. *Arch Biochem Biophys*. 2020 Sep 30;691:108463. doi: 10.1016/j.abb.2020.108463. Epub 2020 Jun 23. PMID: 32590066.
4. Baumgartner H Chair, Hung J Co-Chair, Bermejo J, Chambers JB, Edvardsen T, Goldstein S, Lancellotti P, LeFevre M, Miller F Jr, Otto CM. Recommendations on the echocardiographic assessment of aortic valve stenosis: a focused update from the European Association of Cardiovascular Imaging and the American Society of Echocardiography. *Eur Heart J Cardiovasc Imaging*. 2017 Mar 1;18(3):254-275. doi: 10.1093/ehjci/jew335. PMID: 28363204.
5. Beneduce A, Capogrosso C, Moroni F, Ancona F, Falasconi G, Pannone L, Stella S, Ingallina G, Melillo F, Ancona MB, Romano Rt V, Palmisano A, Latib A, Colombo A, Montorfano M, Esposito A, Agricola E. Aortic valve area calculation using 3D transesophageal echocardiography: Implications for aortic stenosis severity grading. *Echocardiography*. 2020 Dec;37(12):2071-2081. doi: 10.1111/echo.14883. Epub 2020 Oct 7. PMID: 33026122.
6. Berglund V, Mattsson G, Magnusson P. Aortastenos – en vanlig sjukdom som kräver individualiserad handläggning [Aortic stenosis is a common disease which requires individualized treatment]. *Lakartidningen*. 2018 Apr 23;115:E3DR. Swedish. PMID: 29688568.
7. Billig S, Hein M, Uhlig M, Schumacher D, Thudium M, Coburn M, Weisheit CK. Anästhesie bei Aortenklappenstenose : Anästhesiologisches Management von Patienten mit Aortenklappenstenose bei nichtkardiochirurgischen Eingriffen [Anesthesia for aortic valve stenosis : Anesthesiological management of patients with aortic valve stenosis during noncardiac surgery]. *Anaesthesiologie*. 2024 Mar;73(3):168-176. German. doi: 10.1007/s00101-024-01380-x. Epub 2024 Feb 9. PMID: 38334810; PMCID: PMC10920418.
8. Campo J, Tsoris A, Kruse J, Karim A, Andrei AC, Liu M, Bonow RO, McCarthy P, Malaisrie SC. Prognosis of Severe Asymptomatic Aortic Stenosis With and Without Surgery. *Ann Thorac Surg*. 2019 Jul;108(1):74-79. doi: 10.1016/j.athoracsur.2019.01.031. Epub 2019 Mar 21. PMID: 30905426.

9. Caraballo C, Desai NR, Mulder H, Alhanti B, Wilson FP, Fiuzat M, Felker GM, Piña IL, O'Connor CM, Lindenfeld J, Januzzi JL, Cohen LS, Ahmad T. Clinical Implications of the New York Heart Association Classification. *J Am Heart Assoc.* 2019 Dec 3;8(23):e014240. doi: 10.1161/JAHA.119.014240. Epub 2019 Nov 27. PMID: 31771438; PMCID: PMC6912957.
10. Chambers J. Aortic stenosis. *BMJ.* 2005 Apr 9;330(7495):801-2. doi: 10.1136/bmj.330.7495.801. PMID: 15817530; PMCID: PMC556059.
11. Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep Learning in Medical Image Analysis. *Adv Exp Med Biol.* 2020;1213:3-21. doi: 10.1007/978-3-030-33128-3_1. PMID: 32030660; PMCID: PMC7442218.
12. Chattopadhyay A, Sarkar A, Howlader P, Balasubramanian VN. GradCAM++: improved visual explanations for deep convolutional networks. arXiv:1710.11063. doi: 10.1109/WACV.2018.00097
13. Chorba JS, Shapiro AM, Le L, Maidens J, Prince J, Pham S, Kanzawa MM, Barbosa DN, Currie C, Brooks C, White BE, Huskin A, Paek J, Geocaris J, Elnathan D, Ronquillo R, Kim R, Alam ZH, Mahadevan VS, Fuller SG, Stalker GW, Bravo SA, Jean D, Lee JJ, Gjergjindrea M, Mihos CG, Forman ST, Venkatraman S, McCarthy PM, Thomas JD. Deep Learning Algorithm for Automated Cardiac Murmur Detection via a Digital Stethoscope Platform. *J Am Heart Assoc.* 2021 May 4;10(9):e019905. doi: 10.1161/JAHA.120.019905. Epub 2021 Apr 26. PMID: 33899504; PMCID: PMC8200722.
14. Christen, T., Lerch, R., & Trigo Trindade, P. (2006, June). Die Aortenklappenstenose beim Erwachsenen. Teil 1. Ätiologie, Pathophysiologie und Diagnose. In *Swiss Medical Forum* (Vol. 6, No. 26, pp. 626-632). EMH Media.
15. Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, Geis JR, Pandharipande PV, Brink JA, Dreyer KJ. Current Applications and Future Impact of Machine Learning in Radiology. *Radiology.* 2018 Aug;288(2):318-328. doi: 10.1148/radiol.2018171820. Epub 2018 Jun 26. PMID: 29944078; PMCID: PMC6542626.
16. Clavel MA, Magne J, Pibarot P. Low-gradient aortic stenosis. *Eur Heart J.* 2016 Sep 7;37(34):2645-57. doi: 10.1093/eurheartj/ehw096. Epub 2016 Mar 31. PMID: 27190103; PMCID: PMC5030681.
17. Cohen-Shelly M, Attia ZI, Friedman PA, Ito S, Essayagh BA, Ko WY, Murphree DH, Michelena HI, Enriquez-Sarano M, Carter RE, Johnson PW, Noseworthy PA, Lopez-Jimenez F, Oh JK. Electrocardiogram screening for aortic valve stenosis using artificial intelligence. *Eur Heart J.* 2021 Aug 7;42(30):2885-2896. doi: 10.1093/eurheartj/ehab153. PMID: 33748852.

18. d'Arcy JL, Coffey S, Loudon MA, Kennedy A, Pearson-Stuttard J, Birks J, Frangou E, Farmer AJ, Mant D, Wilson J, Myerson SG, Prendergast BD. Large-scale community echocardiographic screening reveals a major burden of undiagnosed valvular heart disease in older people: the OxVALVE Population Cohort Study. *Eur Heart J*. 2016 Dec 14;37(47):3515-3522. doi: 10.1093/eurheartj/ehw229. Epub 2016 Jun 26. PMID: 27354049; PMCID: PMC5216199.
19. Di Minno MND, Di Minno A, Ambrosino P, Songia P, Pepi M, Tremoli E, Poggio P. Cardiovascular morbidity and mortality in patients with aortic valve sclerosis: A systematic review and meta-analysis. *Int J Cardiol*. 2018 Jun 1;260:138-144. doi: 10.1016/j.ijcard.2018.01.054. PMID: 29622430.
20. Eggebrecht H, Schmermund A, Voigtländer T, Kahlert P, Erbel R, Mehta RH. Risk of stroke after transcatheter aortic valve implantation (TAVI): a meta-analysis of 10,037 published patients. *EuroIntervention*. 2012 May 15;8(1):129-38. doi: 10.4244/EIJV8I1A20. PMID: 22391581.
21. Enriquez-Sarano M, Akins CW, Vahanian A. Mitral regurgitation. *Lancet*. 2009 Apr 18;373(9672):1382-94. doi: 10.1016/S0140-6736(09)60692-9. Epub 2009 Apr 6. PMID: 19356795.
22. Flachskampf FA, Daniel WG. Aortenstenose [Aortic valve stenosis]. *Internist (Berl)*. 2004 Nov;45(11):1281-90; quiz 1291-2. German. doi: 10.1007/s00108-004-1285-5. PMID: 15459762.
23. Fortuni F, Delgado V. Assessment of aortic valve stenosis severity: multimodality imaging may be the key. *Eur Heart J Cardiovasc Imaging*. 2020 Oct 1;21(10):1103-1104. doi: 10.1093/ehjci/jeaa129. PMID: 32533186.
24. Gardezi SKM, Myerson SG, Chambers J, Coffey S, d'Arcy J, Hobbs FDR, Holt J, Kennedy A, Loudon M, Prendergast A, Prothero A, Wilson J, Prendergast BD. Cardiac auscultation poorly predicts the presence of valvular heart disease in asymptomatic primary care patients. *Heart*. 2018 Nov;104(22):1832-1835. doi: 10.1136/heartjnl-2018-313082. Epub 2018 May 24. PMID: 29794244.
25. Gnip P, Vokorokos L, Drotár P. Selective oversampling approach for strongly imbalanced data. *PeerJ Comput Sci*. 2021 Jun 18;7:e604. doi: 10.7717/peerj-cs.604. PMID: 34239981; PMCID: PMC8237317.
26. Goody PR, Hosen MR, Christmann D, Niepmann ST, Zietzer A, Adam M, Bönner F, Zimmer S, Nickenig G, Jansen F. Aortic Valve Stenosis: From Basic Mechanisms to Novel Therapeutic Targets. *Arterioscler Thromb Vasc Biol*. 2020 Apr;40(4):885-900. doi: 10.1161/ATVBAHA.119.313067. Epub 2020 Mar 12. PMID: 32160774.

27. Grimard BH, Safford RE, Burns EL. Aortic Stenosis: Diagnosis and Treatment. *Am Fam Physician*. 2016 Mar 1;93(5):371-8. PMID: 26926974.
28. Hagendorff, A, Fehske, W, Flachskampf, F, Helfen, A, Kreidel, F, Kruck, S, Knebel, F (2020). Manual zur Indikation und Durchführung der Echokardiographie–Update 2020 der Deutschen Gesellschaft für Kardiologie. *Die Kardiologie*, 14(5), 396-431.
29. Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations. *Anesthesiology*. 2020 Feb;132(2):379-394. doi: 10.1097/ALN.0000000000002960. PMID: 31939856; PMCID: PMC7643051.
30. He K, Zhang X, Ren S, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *IEEE International Conference on Computer Vision (ICCV) 2015*;1026–1034. doi: 10.1109/ICCV.2015.123
31. Henning RJ. The current diagnosis and treatment of patients with aortic valve stenosis. *Future Cardiol*. 2021 Sep;17(6):1143-1160. doi: 10.2217/fca-2020-0140. Epub 2021 Mar 17. PMID: 33728942.
32. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018 Aug;18(8):500-510. doi: 10.1038/s41568-018-0016-5. PMID: 29777175; PMCID: PMC6268174.
33. Huai X, Kitada S, Choi D, Siriaraya P, Kuwahara N, Ashihara T. Heart sound recognition technology based on convolutional neural network. *Inform Health Soc Care*. 2021 Sep 2;46(3):320-332. doi: 10.1080/17538157.2021.1893736. Epub 2021 Apr 4. PMID: 33818274.
34. lung B, Baron G, Tornos P, Gohlke-Bärwolf C, Butchart EG, Vahanian A. Valvular heart disease in the community: a European experience. *Curr Probl Cardiol*. 2007 Nov;32(11):609-61. doi: 10.1016/j.cpcardiol.2007.07.002. PMID: 17976510.
35. Joseph J, Naqvi SY, Giri J, Goldberg S. Aortic Stenosis: Pathophysiology, Diagnosis, and Therapy. *Am J Med*. 2017 Mar;130(3):253-263. doi: 10.1016/j.amjmed.2016.10.005. Epub 2016 Nov 1. PMID: 27810479.
36. Kamperidis V, Delgado V, van Mieghem NM, Kappetein AP, Leon MB, Bax JJ. Diagnosis and management of aortic valve stenosis in patients with heart failure. *Eur J Heart Fail*. 2016 May;18(5):469-81. doi: 10.1002/ejhf.466. Epub 2016 Jan 4. PMID: 26725980.
37. Kanwar A, Thaden JJ, Nkomo VT. Management of Patients With Aortic Valve Stenosis. *Mayo Clin Proc*. 2018 Apr;93(4):488-508. doi: 10.1016/j.mayocp.2018.01.020. PMID: 29622096.

38. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc.* 2020 Oct;92(4):807-812. doi: 10.1016/j.gie.2020.06.040. Epub 2020 Jun 18. PMID: 32565184.
39. Kodali SK, Williams MR, Smith CR, Svensson LG, Webb JG, Makkar RR, Fontana GP, Dewey TM, Thourani VH, Pichard AD, Fischbein M, Szeto WY, Lim S, Greason KL, Teirstein PS, Malaisrie SC, Douglas PS, Hahn RT, Whisenant B, Zajarias A, Wang D, Akin JJ, Anderson WN, Leon MB; PARTNER Trial Investigators. Two-year outcomes after transcatheter or surgical aortic-valve replacement. *N Engl J Med.* 2012 May 3;366(18):1686-95. doi: 10.1056/NEJMoa1200384. Epub 2012 Mar 26. PMID: 22443479.
40. Kotter E, Pinto Dos Santos D. Ethik und künstliche Intelligenz [Ethics and artificial intelligence]. *Radiologie (Heidelb).* 2024 Jun;64(6):498-502. German. doi: 10.1007/s00117-024-01286-0. Epub 2024 Mar 18. PMID: 38499692
41. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol.* 2017 May 30;69(21):2657-2664. doi: 10.1016/j.jacc.2017.03.571. PMID: 28545640.
42. Küster D, Schultz T. Künstliche Intelligenz und Ethik im Gesundheitswesen – Spagat oder Symbiose? [Artificial intelligence and ethics in healthcare-balancing act or symbiosis?]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2023 Feb;66(2):176-183. German. doi: 10.1007/s00103-022-03653-5. Epub 2023 Jan 17. PMID: 36650296; PMCID: PMC9892090.
43. Kwon JM, Lee SY, Jeon KH, Lee Y, Kim KH, Park J, Oh BH, Lee MM. Deep Learning-Based Algorithm for Detecting Aortic Stenosis Using Electrocardiography. *J Am Heart Assoc.* 2020 Apr 7;9(7):e014717. doi: 10.1161/JAHA.119.014717. Epub 2020 Mar 21. PMID: 32200712; PMCID: PMC7428650.
44. Larentzakis A, Lygeros N. Artificial intelligence (AI) in medicine as a strategic valuable tool. *Pan Afr Med J.* 2021 Feb 17;38:184. doi: 10.11604/pamj.2021.38.184.28197. PMID: 33995790; PMCID: PMC8106796.
45. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015 May 28;521(7553):436-44. doi: 10.1038/nature14539. PMID: 26017442.
46. Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG, Tuzcu EM, Webb JG, Fontana GP, Makkar RR, Brown DL, Block PC, Guyton RA, Pichard AD, Bavaria JE, Herrmann HC, Douglas PS, Petersen JL, Akin JJ, Anderson WN, Wang D, Pocock S; PARTNER Trial Investigators. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. *N Engl J Med.* 2010 Oct 21;363(17):1597-607. doi: 10.1056/NEJMoa1008232. Epub 2010 Sep 22. PMID: 20961243.

47. Lindman BR, Clavel MA, Mathieu P, Iung B, Lancellotti P, Otto CM, Pibarot P. Calcific aortic stenosis. *Nat Rev Dis Primers*. 2016 Mar 3;2:16006. doi: 10.1038/nrdp.2016.6. PMID: 27188578; PMCID: PMC5127286.
48. Makimoto H, Shiraga T, Kohlmann B, Magnisali CE, Gerguri S, Motoyama N, Clasen L, Bejinariu A, Klein K, Makimoto A, Jung C, Westenfeld R, Zeus T, Kelm M. Efficient screening for severe aortic valve stenosis using understandable artificial intelligence: a prospective diagnostic accuracy study. *Eur Heart J Digit Health*. 2022 May 16;3(2):141-152. doi: 10.1093/ehjdh/ztac029. PMID: 36713014; PMCID: PMC9707975.
49. Makimoto H, Höckmann M, Lin T, Glöckner D, Gerguri S, Clasen L, Schmidt J, Assadi-Schmidt A, Bejinariu A, Müller P, Angendohr S, Babady M, Brinkmeyer C, Makimoto A, Kelm M. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction. *Sci Rep*. 2020 May 21;10(1):8445. doi: 10.1038/s41598-020-65105-x. PMID: 32439873; PMCID: PMC7242480.
50. Maknickas V, Maknickas A. Recognition of normal-abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients. *Physiol Meas*. 2017 Jul 31;38(8):1671-1684. doi: 10.1088/1361-6579/aa7841. PMID: 28594638.
51. Mangieri A, Montalto C, Pagnesi M, Lanzillo G, Demir O, Testa L, Colombo A, Latib A. TAVI and Post Procedural Cardiac Conduction Abnormalities. *Front Cardiovasc Med*. 2018 Jul 3;5:85. doi: 10.3389/fcvm.2018.00085. PMID: 30018969; PMCID: PMC6038729.
52. Marquis-Gravel G, Redfors B, Leon MB, G  n  reux P. Medical Treatment of Aortic Stenosis. *Circulation*. 2016 Nov 29;134(22):1766-1784. doi: 10.1161/CIRCULATIONAHA.116.023997. PMID: 27895025.
53. McGee S. Etiology and diagnosis of systolic murmurs in adults. *Am J Med*. 2010 Oct;123(10):913-921.e1. doi: 10.1016/j.amjmed.2010.04.027. PMID: 20920693.
54. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019 Apr;28(2):73-81. doi: 10.1080/13645706.2019.1575882. Epub 2019 Feb 27. PMID: 30810430.
55. Mirolo A, Viart G, Durand E, Savour   A, Godin B, Auquier N, Raiti  re O, Eltchaninoff H, Anselme F. Pacemaker memory in post-TAVI patients: Who should benefit from permanent pacemaker implantation? *Pacing Clin Electrophysiol*. 2018 Sep;41(9):1178-1184. doi: 10.1111/pace.13422. Epub 2018 Jul 16. PMID: 29931693.

56. Montinari MR, Minelli S. The first 200 years of cardiac auscultation and future perspectives. *J Multidiscip Healthc.* 2019 Mar 6;12:183-189. doi: 10.2147/JMDH.S193904. PMID: 30881010; PMCID: PMC6408918.
57. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. *Lancet.* 2006 Sep 16;368(9540):1005-11. doi: 10.1016/S0140-6736(06)69208-8. PMID: 16980116.
58. Olszowska M. Pathogenesis and pathophysiology of aortic valve stenosis in adults. *Pol Arch Med Wewn.* 2011 Nov;121(11):409-13. English, Polish. PMID: 22129836
59. Otto CM, Nishimura RA, Bonow RO, Carabello BA, Erwin JP 3rd, Gentile F, Jneid H, Krieger EV, Mack M, McLeod C, O'Gara PT, Rigolin VH, Sundt TM 3rd, Thompson A, Toly C. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation.* 2021 Feb 2;143(5):e35-e71. doi: 10.1161/CIR.0000000000000932. Epub 2020 Dec 17. Erratum in: *Circulation.* 2021 Feb 2;143(5):e228. Erratum in: *Circulation.* 2021 Mar 9;143(10):e784. PMID: 33332149.
60. Pedriali G, Morciano G, Patergnani S, Cimaglia P, Morelli C, Mikus E, Ferrari R, Gasbarro V, Giorgi C, Wieckowski MR, Pinton P. Aortic Valve Stenosis and Mitochondrial Dysfunctions: Clinical and Molecular Perspectives. *Int J Mol Sci.* 2020 Jul 11;21(14):4899. doi: 10.3390/ijms21144899. PMID: 32664529; PMCID: PMC7402290.
61. Pianykh OS, Langs G, Dewey M, Enzmann DR, Herold CJ, Schoenberg SO, Brink JA. Continuous Learning AI in Radiology: Implementation Principles and Early Applications. *Radiology.* 2020 Oct;297(1):6-14. doi: 10.1148/radiol.2020200038. Epub 2020 Aug 25. PMID: 32840473.
62. Radanliev P, Santos O, Brandon-Jones A, Joinson A. Ethics and responsible AI deployment. *Front Artif Intell.* 2024 Mar 27;7:1377011. doi: 10.3389/frai.2024.1377011. PMID: 38601110; PMCID: PMC11004481.
63. Rajput FA, Zeltser R. Aortic Valve Replacement. 2023 May 29. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. PMID: 30725821.

64. Reardon MJ, Van Mieghem NM, Popma JJ, Kleiman NS, Søndergaard L, Mumtaz M, Adams DH, Deeb GM, Maini B, Gada H, Chetcuti S, Gleason T, Heiser J, Lange R, Merhi W, Oh JK, Olsen PS, Piazza N, Williams M, Windecker S, Yakubov SJ, Grube E, Makkar R, Lee JS, Conte J, Vang E, Nguyen H, Chang Y, Mugglin AS, Serruys PW, Kappetein AP; SURTAVI Investigators. Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients. *N Engl J Med*. 2017 Apr 6;376(14):1321-1331. doi: 10.1056/NEJMoa1700456. Epub 2017 Mar 17. PMID: 28304219.
65. Ristić-Andelkov A, Gligić B, Damnjanović M, Milovanović J, Matunović R, Obradović S. Prognoza i terapija stenoze aortne valvule [Aortic valve stenosis]. *Vojnosanit Pregl*. 2003 Mar-Apr;60(2):167-74. Serbian. doi: 10.2298/vsp0302167r. PMID: 12852159.
66. Roldan CA, Shively BK, Crawford MH. Value of the cardiovascular physical examination for detecting valvular heart disease in asymptomatic subjects. *Am J Cardiol*. 1996 Jun 15;77(15):1327-31. doi: 10.1016/s0002-9149(96)00200-7. PMID: 8677874.
67. Salazar SA, Borrero JL, Harris DM. On systolic murmurs and cardiovascular physiological maneuvers. *Adv Physiol Educ*. 2012 Dec;36(4):251-6. doi: 10.1152/advan.00128.2011. PMID: 23209004.
68. Scherk, J., Pöchlacker-Tröscher, G., & Wagner, K. (2017). Künstliche Intelligenz-Artificial Intelligence. A Report commissioned by the German Federal Ministry of Transport and Infrastructure.
69. Shah KS, Xu H, Matsouaka RA, Bhatt DL, Heidenreich PA, Hernandez AF, Devore AD, Yancy CW, Fonarow GC. Heart Failure With Preserved, Borderline, and Reduced Ejection Fraction: 5-Year Outcomes. *J Am Coll Cardiol*. 2017 Nov 14;70(20):2476-2486. doi: 10.1016/j.jacc.2017.08.074. Epub 2017 Nov 12. PMID: 29141781.
70. Schütze, B., Schlieter, H. Künstliche Intelligenz. *Radiologe* 59, 1091–1096 (2019). <https://doi.org/10.1007/s00117-019-00599-9>
71. Siontis GCM, Overtchouk P, Cahill TJ, Modine T, Prendergast B, Praz F, Pilgrim T, Petrinic T, Nikolakopoulou A, Salanti G, Søndergaard L, Verma S, Jüni P, Windecker S. Transcatheter aortic valve implantation vs. surgical aortic valve replacement for treatment of symptomatic severe aortic stenosis: an updated meta-analysis. *Eur Heart J*. 2019 Oct 7;40(38):3143-3153. doi: 10.1093/eurheartj/ehz275. PMID: 31329852.
72. Stassen J, Ewe SH, Pio SM, Pibarot P, Redfors B, Leipsic J, Genereux P, Van Mieghem NM, Kuneman JH, Makkar R, Hahn RT, Playford D, Marsan NA, Delgado V, Ben-Yehuda O, Leon MB, Bax JJ. Managing Patients With Moderate Aortic Stenosis. *JACC Cardiovasc Imaging*. 2023 Jun;16(6):837-855. doi: 10.1016/j.jcmg.2022.12.013. Epub 2023 Feb 8. PMID: 36881428.

73. Strange G, Stewart S, Celermajer D, Prior D, Scalia GM, Marwick T, Ilton M, Joseph M, Codde J, Playford D; National Echocardiography Database of Australia contributing sites. Poor Long-Term Survival in Patients With Moderate Aortic Stenosis. *J Am Coll Cardiol*. 2019 Oct 15;74(15):1851-1863. doi: 10.1016/j.jacc.2019.08.004. Epub 2019 Sep 3. PMID: 31491546.
74. Thaden JJ, Nkomo VT, Enriquez-Sarano M. The global burden of aortic stenosis. *Prog Cardiovasc Dis*. 2014 May-Jun;56(6):565-71. doi: 10.1016/j.pcad.2014.02.006. Epub 2014 Mar 2. PMID: 24838132.
75. Vahanian A, Alfieri O, Andreotti F, Antunes MJ, Barón-Esquivias G, Baumgartner H, Borger MA, Carrel TP, De Bonis M, Evangelista A, Falk V, Iung B, Lancellotti P, Pierard L, Price S, Schäfers HJ, Schuler G, Stepinska J, Swedberg K, Takkenberg J, Von Oppell UO, Windecker S, Zamorano JL, Zembala M. Guidelines on the management of valvular heart disease (version 2012). *Eur Heart J*. 2012 Oct;33(19):2451-96. doi: 10.1093/eurheartj/ehs109. Epub 2012 Aug 24. PMID: 22922415.
76. Vahanian A, Beyersdorf F, Praz F, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. *Eur Heart J* 2021:ehab395. doi: 10.1093/eurheartj/ehab395.
77. Van Mieghem NM, Reardon MJ, Yakubov SJ, Heiser J, Merhi W, Windecker S, Makkar RR, Cheng W, Robbins M, Fail P, Feinberg E 2nd, Stoler RC, Hebel R, Serruys PW, Popma JJ. Clinical outcomes of TAVI or SAVR in men and women with aortic stenosis at intermediate operative risk: a post hoc analysis of the randomised SURTAVI trial. *EuroIntervention*. 2020 Nov 20;16(10):833-841. doi: 10.4244/EIJ-D-20-00303. PMID: 32715995.
78. Wang R, Pan W, Jin L, Li Y, Geng Y, Gao C, Chen G, Wang H, Ma D, Liao S. Artificial intelligence in reproductive medicine. *Reproduction*. 2019 Oct;158(4):R139-R154. doi: 10.1530/REP-18-0523. PMID: 30970326; PMCID: PMC6733338.
79. Widder JD, Bauersachs J. Therapie der Aortenklappenstenose [Therapy of aortic valve stenosis]. *Internist (Berl)*. 2014 Dec;55(12):1391-2, 1394-6, 1398-9. German. doi: 10.1007/s00108-014-3555-1. PMID: 25373710.
80. Yasaka K, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology*. 2018 Mar;286(3):887-896. doi: 10.1148/radiol.2017170706. Epub 2017 Oct 23. PMID: 29059036.

81. Yi B, Zeng W, Lv L, Hua P. Changing epidemiology of calcific aortic valve disease: 30-year trends of incidence, prevalence, and deaths across 204 countries and territories. *Aging (Albany NY)*. 2021 May 11;13(9):12710-12732. doi: 10.18632/aging.202942. Epub 2021 May 11. PMID: 33973531; PMCID: PMC8148466.
82. Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep Learning in Neuroradiology. *AJNR Am J Neuroradiol*. 2018;39(10):1776-1784. doi:10.3174/ajnr.A5543
83. Zhou Q, Shan J, Ding W, Wang C, Yuan S, Sun F, Li H, Fang B. Cough Recognition Based on Mel-Spectrogram and Convolutional Neural Network. *Front Robot AI*. 2021 May 7;8:580080. doi: 10.3389/frobt.2021.580080. PMID: 34026854; PMCID: PMC8138471.
84. Zippel-Schultz B, Schultz C, Müller-Wieland D, Remppis AB, Stockburger M, Perings C, Helms TM. Künstliche Intelligenz in der Kardiologie : Relevanz, aktuelle Anwendungen und nächste Schritte [Artificial intelligence in cardiology : Relevance, current applications, and future developments]. *Herzschrittmacherther Elektrophysiol*. 2021 Mar;32(1):89-98. German. doi: 10.1007/s00399-020-00735-2. Epub 2021 Jan 15. PMID: 33449234.

7 Anhang

Fragebogen für die eingeschlossenen Patienten

Sheet for UKD AI Project

Enrollment Date: __. __.

Name: _____ **Birthday:** __. __.

Checklist: *all following checkpoints should be “yes”

Age 20- years old ☐ yes ☐ no

Informed Consent ☐ yes ☐ no

Please check if the following criteria are “no”

Pacemaker ☐ yes ☐ no

ICD ☐ yes ☐ no

CRT ☐ yes (Date : __. __. __) ☐ no

EKG parameters:

Rhythm: ☐ sinus rhythm ☐ atrial fibrillation ☐ other rhythm

Heart Rate:

P wave duration:

PQ interval:

QRS duration:

☐ incomplete RBBB ☐ incomplete LBBB ☐ LAHB ☐ LPHB ☐ RBBB

QT interval:

QRS axis: _____ grade

Clinical Investigation:

Heart Sound:

I sound:

II sound:

III sound:

IV sound:

NYHA classification ☐ I ☐ II ☐ III ☐ IV

ICD: implantierbarer Cardioverter-Defibrillator, CRT: *cardiac resynchronization therapy*, EKG: Elektrokardiogramm, RBBB: *right bundle branch block*, LAHB: linksanteriorer Hemiblock, LPHB: linksposteriorer Hemiblock, LBBB: *left bundle branch block*, NYHA: *New York Heart Association*

Detailed Information 1:

Past Histories:

Hypertension ☐ yes ☐ no
Diabetes Mellitus ☐ yes ☐ no
Dyslipidemia ☐ yes ☐ no
Hyperuricemia ☐ yes ☐ no
Dialysis ☐ yes ☐ no
Myocardial infarction ☐ yes (Date : __. __. __) ☐ no
Notes on myocardial infarction:
History of PCI ☐ yes ☐ no
Notes on PCI:
History of CABG ☐ yes ☐ no
Notes on CABG:
History of valve operation ☐ yes ☐ no
Notes on Valve Operation:

Echocardiography: date : __. __.
LVDd __mm ; LVDs __mm ; LVEF __% ; IVSd __mm ; LVPWd __mm
LAD __mm; LA area __cm²
RV dysfunction ☐ yes ☐ no
Mitral regurgitation ☐ no ☐ I ☐ II ☐ III ; Mitral stenosis ☐ no ☐ I ☐ II ☐ III
Aortic regurgitation ☐ no ☐ I ☐ II ☐ III ; Aortic stenosis ☐ no ☐ I ☐ II ☐ III
Mitral valve replacement ☐ yes (date : __. __. __) ☐ no
Aortic valve replacement ☐ yes (date : __. __. __) ☐ no
Notes on TTE :

PCI: perkutane Koronarintervention, CABG: *coronary artery bypass graft*, LVDd: linksventrikulärer enddiastolischer Durchmesser, LVDs: linksventrikulärer endsystolischer Durchmesser, LVEF: linksventrikuläre Ejektionsfraktion, IVSd: interventrikuläre Septumdicke (diastolisch), LVPWd: linksventrikuläre posterior Wanddicke (diastolisch), LAD: linksatrialer Durchmesser, LA: linkes Atrium, RV: rechter Ventrikel; TTE: transthorakale Echokardiographie

Detailed Information 2:

Medications:

Calcium antagonist ☐ yes (name_____; dose_____) ☐ no

ACE ☐ yes (name_____; dose_____) ☐ no

ARB ☐ yes (name_____; dose_____) ☐ no

β -blocker ☐ yes (name_____; dose_____) ☐ no

Mineralocorticoid receptor antagonist ☐ yes (name_____; dose_____) ☐ no

ARNI ☐ yes (name_____; dose_____) ☐ no

Statin ☐ yes (name_____; dose_____) ☐ no

PCSK9 ☐ yes (name_____; dose_____) ☐ no

α -blocker ☐ yes (name_____; dose_____) ☐ no

Amiodarone ☐ yes (dose_____since___. ___. ___) ☐ no

Sotalol ☐ yes (dose_____since___. ___. ___) ☐ no

Ajmalin ☐ yes (dose_____since___. ___. ___) ☐ no

Mexiletine ☐ yes (dose_____since___. ___. ___) ☐ no

Flecainid ☐ yes (dose_____since___. ___. ___) ☐ no

Other medications:

1. name_____; dose

2. name_____; dose

3. name_____; dose

ACE: *Angiotensin-converting* Enzym, ARB: Angiotensinrezeptorblocker, ARNI: Angiotensin-Rezeptor-Neprilysin-Inhibitor, PCSK9: Proproteinkonvertase Subtilisin Kexin Typ 9

Tabelle 11: Datensatzvorbereitung aufgeteilt nach den Schweregraden einer begleitenden Mitralklappeninsuffizienz

	Gesammelte Datensätze	Testdatensätze	Trainings- und Validierungsdatensätze
Schwere Mitralklappeninsuffizienz	keine AS n=42 leichte AS n=3 moderate AS n=3 schwere AS n=6	keine AS n=12 leichte AS n=1 moderate AS n= 1 schwere AS n=2	keine AS n=30 leichte AS n=2 moderate AS n=2 schwere AS n=4
Moderate Mitralklappeninsuffizienz	keine AS n=64 leichte AS n=6 moderate AS n=12 schwere AS n=18	keine AS n=10 leichte AS n=1 moderate AS n=2 schwere AS n=3	keine AS n=54 leichte AS n=5 moderate AS n=10 schwere AS n=15
Keine/leichte Mitralklappeninsuffizienz	keine AS n=564 leichte AS n=42 moderate AS n=36 schwere AS n=90	keine AS n=40 leichte AS n=7 moderate AS n=6 schwere AS n=15	keine AS n=524 leichte AS n=35 moderate AS n=30 schwere AS n=75

8 Tabellenverzeichnis

Tabelle 1: Einteilung der Schweregrade der Aortenklappenstenose anhand echokardiographisch erhobener Parameter	5
Tabelle 2: Charakteristika der Patienten der Trainings-, Validierungs- und Testgruppe	18
Tabelle 3: Charakteristika der exkludierten Patienten	20
Tabelle 4: Charakteristika der Patienten der klinischen Validierung	22
Tabelle 5: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten von 2RSB.....	36
Tabelle 6: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten des Erb-Punktes	38
Tabelle 7: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten des Apex.....	40
Tabelle 8: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L1 Modell.....	54
Tabelle 9: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L2 Modell.....	59
Tabelle 10: Fokusscoreanalyse pro Herzschlag im Vergleich zwischen der schweren und nicht-schweren Aortenklappenstenose beim L3 Modell.....	62
Tabelle 11: Datenvorbereitung aufgeteilt nach den Schweregraden einer begleitenden Mitralklappeninsuffizienz	93

Abbildungsverzeichnis

Abb. 1: Visualisierung der Tonspuren	25
Abb. 2: Flowchart der rekrutierten Patienten	27
Abb. 3: Beispiel eines einfachen <i>convolutional neural network</i>	29
Abb. 4: Beispiel eines kombinierten <i>convolutional neural network</i>	30
Abb. 5: Rekrutierung neuer Datensätze im Rahmen der klinischen Validierung	33
Abb. 6: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten von 2RSB	37
Abb. 7: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten des Erb-Punktes	39
Abb. 8: Vergleich der Leistungsdaten der verschiedenen <i>convolutional neural networks</i> anhand der Auskultationsdaten des Apex	41
Abb. 9: Vergleich der Ergebnisse der <i>convolutional neural networks</i> anhand des F1-Wertes	42
Abb. 10: Vergleich der Ergebnisse der <i>convolutional neural networks</i> anhand der Treffsicherheit	42
Abb. 11: Vergleich der Ergebnisse der <i>convolutional neural networks</i> anhand der <i>area under the curve</i>	43
Abb. 12: Vergleich der Treffsicherheit der Modelle L1-L3 in Bezug auf den Schweregrad der Aortenklappenstenose	44
Abb. 13: Vergleich der Sensitivität und der Spezifität der Modelle L1-3.....	45
Abb. 14: Vergleich der Sensitivität und Spezifität der zusammengesetzten Modelle E1-E3	46
Abb. 15: Vergleich der Treffsicherheit der Modelle E1-E3 in Bezug auf den Schweregrad der Aortenklappenstenose	47
Abb. 16: Vergleich zwischen den Modellen L1-L3 und E1-E3 in Hinblick auf die Treffsicherheit, F1 Wert, Sensitivität, Spezifität und <i>area under the curve</i> (AUC)	48
Abb. 17: Vergleich der Modelle L1, L2, L3 und E1 in Hinblick auf die Treffsicherheit, die Sensitivität, die Spezifität und den F1 Wert bei der klinischen Validierung.....	49
Abb. 18: Vergleich der erreichten Treffsicherheit des E1 Modells bei den verschiedenen Schweregraden der Aortenklappenstenose	50
Abb. 19: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei korrekt erkannter schwerer AS beim L1 Modell	53

Abb. 20: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei korrekt erkannter leichter sowie keiner AS beim L1 Modell	55
Abb. 21: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> als Beispiel bei Schwierigkeiten bei der Bestimmung der schweren Aortenklappenstenose	56
Abb. 22: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei einer fälschlicherweise als schwer klassifizierten moderaten Aortenklappenstenose	57
Abb. 23: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei schwererer sowie moderater AS beim L2 Modell.....	58
Abb. 24: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei leichter sowie keiner AS beim L2 Modell	59
Abb. 25: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> bei fälschlicher Einordnung einer schweren als nicht-schwere Aortenklappenstenose beim L2 Modell	60
Abb. 26: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> mit fälschlicher Einordnung einer moderaten als schwere Aortenklappenstenose beim L2 Modell	61
Abb. 27: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> mit korrekter Erkennung der schweren und der moderate AS beim L3 Modell	63
Abb. 28: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> mit korrekter Erkennung der leichten und keiner AS beim L3 Modell	63
Abb. 29: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> mit fälschlicher Erkennung einer moderaten als schwere AS beim L3 Modell.....	64
Abb. 30: Visualisierung der Herztöne mittels Waveform und Mel-Spektrogramm mit <i>Heatmap</i> mit fälschlicher Erkennung einer schweren AS als nicht-schwere AS beim L3 Modell.....	65

Danksagung

Ich danke Herrn Prof. Dr. med. Malte Kelm, Direktor der Klinik für Kardiologie, Pneumologie und Angiologie am Universitätsklinikum Düsseldorf, unter dessen Leitung ich meine Promotionsarbeit anfertigen durfte, für die Überlassung des Dissertationsthemas.

Ein herausragender Dank geht an meinen Doktorvater Herrn Prof. Dr. med. Hisaki Makimoto, ehemaliger leitender Oberarzt der Abteilung für interventionelle Rhythmologie und Devicetherapie an der Klinik für Kardiologie, Pneumologie und Angiologie am Universitätsklinikum Düsseldorf und nun zurückgekehrt nach Japan, für die stetige herzliche und wertschätzende Betreuung, für das offene Ohr und die konstruktiven Verbesserungsvorschläge. Ohne ihn wäre die Promotion nicht möglich gewesen. Vielen, vielen Dank!

Zudem danke ich PD Dr. rer. nat. Sebastian Temme für die Erstellung des Zweitgutachtens sowie die unkomplizierte und wertschätzende Unterstützung am Ende der Promotion. Ein großer Dank gilt außerdem Dr.med. Shqipe Gerguri und Dr. med. Alexandru Bejinariu, welche mir insbesondere zu Beginn des Promotionsvorhabens mit Rat zur Seite standen und mich unterstützten. Zudem danke ich den Mitarbeitenden der kardiologischen Funktionsabteilung, welche mich bei der Datensammlung tatkräftig unterstützt haben, sowie der Abteilung für interventionelle Rhythmologie und Devicetherapie.

Ein besonderer Dank gilt Helen, die sich während der Promotion auch in schwierigen Momenten immer meine Sorgen angehört hat und mich immer wieder aufgebaut hat. Ohne dich wäre die Studienzeit nicht halb so schön gewesen. Zudem auch ein großes Danke an Vincent fürs Motivieren zum Durchhalten, den Glauben an mich und natürlich die kulinarische Unterstützung. Zuletzt möchte ich noch meiner Familie und insbesondere meinen Eltern für die jahrelange Unterstützung in jeglicher Form danken. Ohne euch alle wäre dies nicht möglich gewesen!