

**Weiterentwicklung und prospektive klinische Validierung
eines lernenden künstlichen neuronalen Netzwerks
zur automatischen Detektion und Segmentierung
von fokalen kortikalen Dysplasien**

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Medizin (Dr. med.)

an der

Medizinischen Fakultät der Universität Hamburg

vorgelegt von

Vicky Chanra

aus

Kabul

2025

Betreuer:in / Gutachter:in der Dissertation: PD Dr. Thomas Sauvigny

Gutachter:in der Dissertation: Prof. Dr. Stefano Panzeri

Vorsitz der Prüfungskommission: Prof. Dr. Stefano Panzeri

Mitglied der Prüfungskommission: Prof. Dr. Patrick Czorlich

Mitglied der Prüfungskommission: PD Dr. Mathias Gelderblom

Datum der mündlichen Prüfung: 04.11.2025

Inhaltsverzeichnis

	<u>Seite</u>
1. Darstellung der Publikation	3
2. Artikel	27
3. Zusammenfassung	48
4. Literaturverzeichnis	49
5. Abkürzungsverzeichnis	54
6. Abbildungsverzeichnis	55
7. Tabellenverzeichnis	56
8. Erklärung des Eigenanteils	57
9. Eidesstattliche Versicherung	58
10. Danksagung	59

1. Darstellung der Publikation

1.1 Einleitung

Epilepsie ist eine neurologische Störung, bei der das Gehirn eine anhaltende Neigung zu spontanen epileptischen Anfällen hat (Fischer et al., 2005). Gemäß der Definition der International League Against Epilepsy (ILAE) liegt eine Epilepsie vor, wenn (1) mindestens zwei unprovizierte Anfälle mit einem Abstand von mehr als 24 Stunden aufgetreten sind, (2) nach einem unprovizierten Anfall das Risiko für einen weiteren Anfall in den nächsten 10 Jahren über 60 % liegt oder (3) die Diagnose eines Epilepsiesyndroms gestellt werden kann (Fischer et al., 2014). Weltweit sind rund 50 Millionen Menschen betroffen (World Health Organization, 2019), von denen etwa 30 bis 40 % pharmakoresistent sind (Chen et al. 2018), d. h., dass trotz Einnahme von zwei ausreichend hoch dosierten und verträglichen anfallssupprimierenden Medikamenten (ASM) als Mono- oder Kombinationstherapie keine Anfallsfreiheit erreicht wird (Kwan et al. 2010). Eine Anfallspersistenz geht mit einer reduzierten Lebensqualität (Ridsdale et al., 2017), einem erhöhten Risiko für psychische Erkrankungen (Tsigebrhan et al., 2023), Stigmatisierung, sozioökonomischen Nachteilen (Wiebe, 2006) sowie einer erhöhten Mortalität durch Verletzungen, Suizid, Status epilepticus und SUDEP (Sudden Unexpected Death in Epilepsy) einher (Thurman et al., 2017).

Fokale kortikale Dysplasien (FCD) bilden die häufigste Ursache pharmakoresistenter Epilepsien im Kindesalter und die dritthäufigste Ursache bei Erwachsenen nach Hippocampussklerose und Tumoren (Blümcke et al., 2017). Diese angeborenen Fehlbildungen der Hirnrinde entstehen durch Störungen in der neuronalen Proliferation, Migration und Differenzierung, was zu einer Desorganisation der kortikalen Architektur und abnormen Neuronenstrukturen führt (Palmini et al., 2004). Nach der aktuellen ILAE-Klassifikation von 2022 (Najm et al., 2022) werden FCD in drei Haupttypen unterteilt (s. [Tabelle 1](#)). Die vollständige operative Resektion bietet die beste Chance auf langfristige Anfallsfreiheit und eine Verringerung von Komplikationen, weshalb eine frühzeitige und präzise Detektion prognostisch entscheidend ist (Willard et al., 2022). Zudem ist eine exakte Identifizierung der Läsion essenziell für die Planung und Implantation intrakranieller EEG-Elektroden.

Die Erkennung von FCD basiert primär auf der visuellen Analyse hochauflösender zerebraler MRT-Bilder unter Anwendung des „Harness“-Protokolls (Bernasconi et al., 2019). Da die Bildgebungsmerkmale von FCD oft subtil sind und fließend in das umliegende Gewebe übergehen (s. [Abbildung 1](#)), können sie bei der visuellen Analyse leicht übersehen werden, was zu einer fehlerhaften Einstufung der Epilepsie als „MRT-negativ“ führen kann (Wang et al., 2016). Trotz technischer Fortschritte bleiben weiterhin 30 bis 78 % der FCD-Fälle bei der visuellen MRT-Analyse unentdeckt (Urbach et al., 2022). Der fehlende Nachweis einer Läsion geht mit einem schlechteren postoperativen Outcome einher,

weshalb bei MRT-negativen Patienten seltener ein epilepsiechirurgischer Eingriff in Betracht gezogen wird (Spencer et al., 2005).

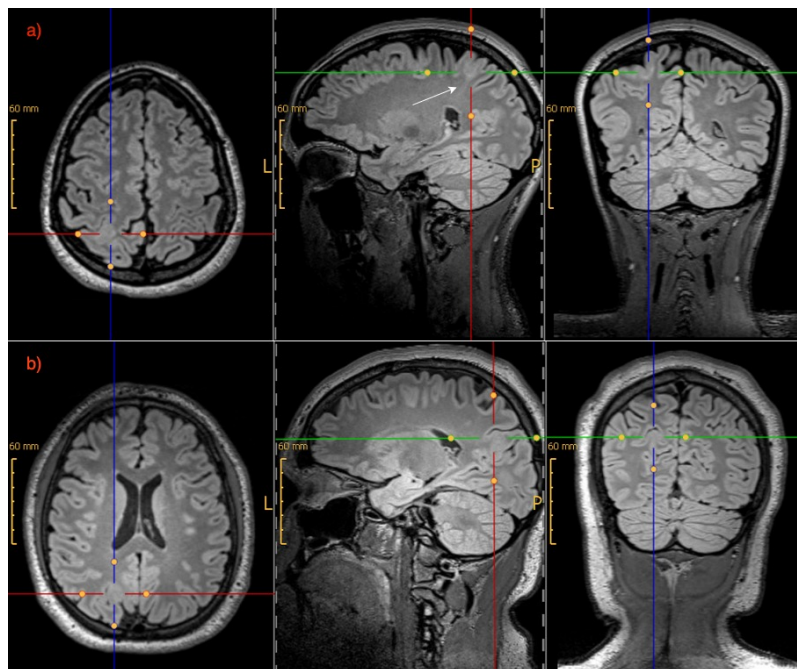


Abbildung 1: Charakteristische MRT-Merkmale von fokalen kortikalen Dysplasien (FCD)

Die Fadenkreuze markieren die FCD auf FLAIR-gewichteten 3T-Bildern, die im Rahmen der präoperativen Diagnostik erstellt wurden.

- MRT einer histologisch bestätigten FCD Typ IIb im rechten Parietallappen. Die charakteristischen MRT-Merkmale von FCD Typ II umfassen eine lokale Verdickung der Hirnrinde, eine unscharfe Mark-Rinden-Grenze, eine abnorme kortikale Gyrierung sowie ein fokales erhöhtes T2-Signal. Ein spezifisches Merkmal ist das „Transmantle“-Zeichen (weißer Pfeil): ein kegelförmiges hyperintenses T2-Signal, das sich von der kortikalen Läsion bis zum Ventrikel erstreckt (Urbach et al., 2022; Barkovich und Kuzniecky, 1996).
- MRT einer histologisch bestätigten FCD Typ Ia im rechten Parietallappen. FCD Typ I sind visuell oft schwer abgrenzbar, können jedoch subtile Hinweise wie ein moderat erhöhtes T2/FLAIR-Signal und ein vermindertes T1-Signal in der weißen Substanz, eine unscharfe Mark-Rinden-Grenze sowie eine Hypoplasie der betroffenen Region aufweisen (Marcelis et al., 2022).

Tabelle 1: Dreistufige ILAE-Klassifikation der fokalen kortikalen Dysplasien (Najm et al., 2022)

Typ	Beschreibung	Untertyp
FCD Typ I	FCD mit abnormer kortikaler Laminierung	Ia: Abnorme radiale kortikale Laminierung Ib: Abnorme tangentielle kortikale Laminierung Ic: Abnorme radiale und tangentielle kortikale Laminierung
FCD Typ II	FCD mit abnormer kortikaler Laminierung und zellulären Anomalien	IIa: Dysmorphie Neurone IIb: Dysmorphie Neurone und Ballonzellen
FCD Typ III	FCD mit abnormer kortikaler Laminierung, die mit einer anderen primären Läsion assoziiert ist	IIIa: Mit einer Hippocampussklerose assoziiert IIIb: An einen glialen oder glioneuronalen Tumor angrenzend IIIc: An eine vaskuläre Malformation angrenzend IIId: An andere früh erworbene Läsionen angrenzend, z. B. Trauma, ischämische Schädigung, Enzephalitis

Der Begriff „Fokale kortikale Dysplasie“ wurde erstmals 1971 von Taylor et al. eingeführt, die bei zehn operierten Patienten mit pharmakoresistenter Epilepsie systematisch charakteristische Veränderungen im Neokortex dokumentierten. Bereits 1957 hatte Crome bei drei Patienten mit schwer behandelbarer Epilepsie ähnliche mikroskopische Befunde beschrieben, darunter „große, unregelmäßig geformte Nervenzellen mit dicken, gewundenen Fortsätzen“, „dysmorphe Neurone“ und „Riesennervenzellen“.

Digitale Bildnachbearbeitungsmethoden (Postprocessing) unterstützen die Erkennung von FCD, indem sie quantitative und qualitative Ansätze zur Extraktion von Bildmerkmalen kombinieren. Das „Morphometric Analysis Program“ (MAP18, Version 2018) (Huppertz et al., 2008, 2005) ist eine halb-automatisierte Bildnachbearbeitungsmethode auf Grundlage der voxelbasierten Morphometrie (Ashburner et al., 2000) und fester Bestandteil der präoperativen Diagnostik in über 60 Epilepsiezentren weltweit (David et al., 2021). Trotz der breiten klinischen Anwendung weist die Methode mehrere Einschränkungen auf, die den universellen Einsatz erschweren. Dazu zählen die aufwendige Implementierung einer MRT-Scanner-spezifischen Normdatenbank, was die Übertragbarkeit auf andere Scannerprotokolle oder Populationen einschränkt (alternativ kann eine unspezifische MRT-Normdatenbank mit einer wesentlich geringeren statistischen Spezifität verwendet werden), sowie die zeitintensive visuelle Auswertung, die aufgrund häufiger falsch-positiver Befunde Expertenwissen erfordert.

Convolutional Neural Networks (CNN) sind eine Klasse künstlicher neuronaler Netzwerke, die sich besonders für die effiziente Verarbeitung von Daten mit Gitterstrukturen wie beispielsweise Bilder eignen und eine automatisierte und untersucherabhängige Auswertung von MRTs ermöglichen. Durch ihre Fähigkeit, räumliche Hierarchien in Daten zu erfassen, Merkmale automatisch zu extrahieren und Translationsinvarianz zu gewährleisten, eignen sich CNN besonders für Aufgaben wie Objekterkennung, Klassifikation und Segmentierung (LeCun et al., 2015).

Ein zentrales Problem im maschinellen Lernen ist das „katastrophale Vergessen“ (McCloskey und Cohen, 1989), bei dem ein Modell zuvor erlerntes Wissen verliert, wenn es auf neuen Daten trainiert wird. Dieses Phänomen resultiert aus dem Spannungsfeld zwischen Lernplastizität und Gedächtnisstabilität: Eine höhere Anpassungsfähigkeit führt zu einer Schwächung der Gedächtnisstabilität und umgekehrt. Die Wiederverwendung aller bisherigen Daten kann diesem Problem entgegenwirken, erfordert jedoch erheblichen Speicher- und Rechenaufwand.

Kontinuierliches Lernen (Kirkpatrick et al., 2017) ist ein Ansatz im maschinellen Lernen, bei dem ein Modell fortlaufend aus neuen Daten lernt, ohne zuvor erlerntes Wissen zu verlieren. Es ermöglicht ein ressourcenschonenderes Lernen, das ohne umfassendes Neutraining auskommt und dadurch Rechenzeit, Speicherbedarf sowie die Abhängigkeit von alten Daten verringert (Wang et al., 2024). Dies verbessert sowohl die Effizienz als auch die Anpassungsfähigkeit und Anwendbarkeit der Modelle.

House et al. (2021) entwickelten ein CNN zur FCD-Erkennung und Segmentation, das ein U-Net (Ronneberger et al., 2015) mit einem zusätzlichen Variational-Decoder-Zweig (Myronenko, 2018) kombiniert. Das CNN wurde mit einem 3T-MRT-Datensatz des Epilepsie-Zentrums Hamburg trainiert, der auf einem 3T-Siemens-Skyra-Scanner (Magnetom Syngo MR D11) erstellt wurde. Der Trainingsdatensatz umfasste insgesamt 201 FCD-MRTs, 100 normale MRTs und 50 pathologische Nicht-FCD-

MRTs. Die Leistungsbewertung erfolgte durch Kreuzvalidierung mit vier verschiedenen Datenkonfigurationen. Das leistungsstärkste Modell, im Folgenden als „Basismodell“ bezeichnet, wurde anschließend prospektiv an einem unabhängigen Datensatz mit 100 MRTs, darunter neun FCD-MRTs, validiert. Dabei zeigte das Modell eine Sensitivität von 77,8 % und eine Spezifität von 5,5 %. Aufgrund der niedrigen Spezifität war das Modell bisher höchstens für ein „Pre-Screening“ von MRTs, aber nicht für den automatisierten klinischen Einsatz geeignet.

In dieser prospektiven Studie wurde das von House et al. (2021) entwickelte CNN durch daten-gesteuertes Training und die Implementierung von kontinuierlichem Lernen weiterentwickelt. Seine Leistungsfähigkeit wurde anhand eines klinischen Datensatzes evaluiert. Ziel war es, ein Modell zu schaffen, das für den klinischen Einsatz geeignet ist.

1.2 Material und Methoden

1.2.1 Datenerhebung

Die Studie wurde prospektiv in einem zehnmonatigen Zeitraum zwischen März und Dezember 2022 am Epilepsie-Zentrum Hamburg durchgeführt. Eingeschlossen wurden alle Patienten des Hamburger Epilepsie-Zentrums, bei denen hochauflösende cerebrale 3D-Datensätze (T1, T2, FLAIR) an einem 3T-Siemens-Skyra-Scanner (Magnetom Syngo MR D11) an der Klinik für Neuroradiologie des Universitätsklinikums Hamburg-Eppendorf angefertigt wurden. Um die Eignung des Algorithmus für die klinische Praxis sicherzustellen, wurden MRTs mit suboptimaler Bildqualität absichtlich nicht aus der Studie ausgeschlossen.

Bei allen Teilnehmern wurden zusätzlich eine Anfallsanamnese erhoben, eine neurologische Untersuchung durchgeführt und Video-EEGs aufgezeichnet.

Die MRTs wurden von einem Assistenzarzt und einem Facharzt der Klinik für Neuroradiologie am Universitätsklinikum Hamburg-Eppendorf visuell analysiert.

Der 3D-Datensatz wurde im DICOM-Format exportiert und mit DICOM Anonymizer Pro (v2.0.15, NeoLogica s.r.l.) anonymisiert.

Die morphometrische MRT-Analyse wurde auf anonymisierten 3D T1- und T2-Sequenzen mit coregistrierten 3D FLAIR-Sequenzen mittels MAP18 (Huppertz et al. 2008, 2005) vom Erstautor durchgeführt und im Konsensverfahren vom Erst- und Letztautor ausgewertet.

Die abschließenden Bewertungen der MRTs berücksichtigten die Anfallssemiologie, Video-EEG-Daten sowie die Mitbeurteilung durch weitere Epileptologen (vor allem Chefarzt und Oberärzte) des Hamburger Epilepsie-Zentrums.

Die Datenerhebung wurde fortgeführt, bis insgesamt 30 FCDs identifiziert waren. Fokale Polymikrogyrien (PMG) (n = 4) wurden ebenfalls in die FCD-Kohorte aufgenommen, da ihre Bildmerkmale denen von FCD ähneln und sie für die prächirurgische Bewertung bei Patienten mit fokaler pharmakoresistenter Epilepsie von gleicher Bedeutung sind. Zur Vereinfachung werden FCD und PMG im Folgenden gemeinsam als „FCD“ bezeichnet. Die regionale Verteilung und die anatomischen Charakteristika der identifizierten FCDs sind in [Tabelle 3](#) dargestellt.

Als Kontrollgruppe wurden 270 randomisierte MRTs verwendet, darunter 150 normale MRTs und 120 pathologische Nicht-FCD-MRTs mit verschiedenen Pathologien (s. [Tabelle 2](#)). Die demografischen Merkmale des gesamten Datensatzes sind in [Tabelle 4](#) zusammengefasst.

Tabelle 2: Zusammensetzung der pathologischen Nicht-FCD-MRTs

Pathologie	Prozentualer Anteil [%]
Hippocampusklerose	7,5
DNET/Gangliogliom/Gangliozytom	7,5
Glioblastom/Astrozytom	5
Hamartom	2,5
Kavernom	7,5
Arteriovenöse Malformation	2,5
DVA	5
Zerebraler Infarkt	7,5
Zerebrale Hämorrhagie	2,5
Arachnoidalzyste	7,5
Porencephalie	2,5
Moderate Mikroangiopathie	2,5
Schwere Mikroangiopathie	2,5
Multiple Sklerose-Läsion	2,5
Enzephalitis	15
Leukodystrophie	5
Posttraumatischer Defekt	7,5
Postoperativer Defekt	7,5

Abkürzungen: DNET – Dysembryoplastischer Neuroepithelialer Tumor;
DVA – Developmental Venous Anomaly

Tabelle 3: Regionale Verteilung und anatomische Merkmale der FCD-Kohorte

		Teildatensatz 1 (n = 10 FCD)	Teildatensatz 2 (n = 10 FCD)	Teildatensatz 3 (n = 10 FCD)	Gesamt
Größe (Volumen)	Klein (< 1,5 ml)	4	2	0	6 (20,0 %)
	Mittel (1,5 – 7,5 ml)	3	4	3	10 (33,3 %)
	Groß (> 7,5 ml)	3	4	7	14 (46,7 %)
Hemisphäre	Rechts	6	8	7	21 (70,0 %)
	Links	4	2	3	9 (30,0 %)
Transmantle- Zeichen	Ja	4	2	4	10 (33,3 %)
	Nein	6	8	6	20 (66,7 %)
Fokale PMG		1	2	1	4 (13,3 %)
Lokalisation	Frontal	6	3	1	10 (33,3 %)
	Temporal	1	0	0	1 (3,3 %)
	Insula	1	3	1	5 (16,7 %)
	Parietal	2	4	6	12 (40,0 %)
	Okzipital	0	0	2	2 (6,7 %)

Abkürzungen: PMG – Polymikrogyrie

Tabelle 4: Demografische Merkmale des Datensatzes

Gesamt		FCD-MRTs (jeweils n = 10)	Pathologische Nicht-FCD-MRTs (jeweils n = 40)	Normale MRTs (jeweils n = 50)
Teildatensatz 1 (n = 100 MRTs)	62,0 % weiblich, Altersspanne: 15 bis 82 J, M + SD: 34,6 ± 17,9 J	40,0 % weiblich, Altersspanne: 18 bis 43 J, M + SD: 25,3 ± 7,4 J	57,5 % weiblich, Altersspanne: 17 bis 80 J, M + SD: 38,9 ± 19,5 J	70,0 % weiblich, Altersspanne: 15 bis 82 J, M + SD: 33,0 ± 17,0 J
Teildatensatz 2 (n = 100 MRTs)	59,0 % weiblich, Altersspanne: 14 bis 78 J, M + SD: 38,1 ± 16,0 J	50,0 % weiblich, Altersspanne: 18 bis 54 J, M + SD: 30,1 ± 11,9 J	55,0 % weiblich, Altersspanne: 14 bis 78 J, M + SD: 41,2 ± 17,8 J	64,0 % weiblich, Altersspanne: 17 bis 73 J, M + SD: 37,2 ± 14,5 J
Teildatensatz 3 (n = 100 MRTs)	51,0 % weiblich, Altersspanne: 14 bis 78 J, M + SD: 31,2 ± 14,2 J	20,0 % weiblich, Altersspanne: 18 bis 48 J, M + SD: 29,0 ± 9,6 J	47,5 % weiblich, Altersspanne: 14 bis 78 J, M + SD: 35,3 ± 17,2 J	60,0 % weiblich, Altersspanne: 17 bis 68 J, M + SD: 28,4 ± 11,2 J
Gesamt		36,7 % weiblich, Altersspanne: 17 bis 54 J, M + SD: 28,1 ± 10,1 J	53,3 % weiblich, Altersspanne: 14 bis 80 J, M + SD: 38,4 ± 18,4 J	64,7 % weiblich, Altersspanne: 15 bis 82 J, M + SD: 32,8 ± 14,9 J

Abkürzungen: J – Jahre; M – Mittelwert; SD – Standardabweichung

1.2.2 Digitale Nachbearbeitung (Postprocessing) nach MAP18

MAP18 (Huppertz et al., 2008, 2005) verwendet ein automatisiertes MATLAB-Skript (MathWorks, USA) in Kombination mit Algorithmen der Open-Source-Software „Statistical Parametric Mapping“ (SPM12, Functional Imaging Laboratory, UCL Queen Square Institute of Neurology, London, UK), um die individuelle Gehirnanatomie mit einer Normdatenbank zu vergleichen (s. [Abbildung 2](#)). Hierfür werden hochauflösende T1- und T2-gewichtete cerebrale MRT-Volumendatensätze in den standardisierten anatomischen Raum des Montreal Neurological Institute (MNI-Raum) überführt, um interindividuelle Vergleiche zu ermöglichen. Der „Unified Segmentation“-Algorithmus segmentiert das Gehirn in graue Substanz, weiße Substanz und Liquor cerebrospinalis, wobei Bildintensitätsinhomogenitäten korrigiert werden. Auf Basis der Mittelwerte und Standardabweichungen der Voxelintensitäten in grauer und weißer Substanz werden individuelle Schwellenwerte berechnet, um relevante Regionen hervorzuheben. Subkortikale und infratentorielle Regionen werden mithilfe vordefinierter Masken ausgeblendet. Das resultierende Binärbild wird durch einen dreidimensionalen Faltungsfilter geglättet, um zusammenhängende Strukturen hervorzuheben und Rauschen zu reduzieren. Anschließend wird die Verteilung der grauen und weißen Substanz voxelweise analysiert und mit der Normdatenbank verglichen. Aus dieser Analyse entstehen drei spezifische 3D-Merkmalkarten, die charakteristische Merkmale kortikaler Dysplasien visualisieren. Diese Karten stellen die Ergebnisse als z-Score-Karten dar, wobei Regionen mit hohen z-Scores hervorgehoben werden, da sie auf morphologische Anomalien hinweisen. Dabei hebt das „Junction image“ unscharfe Übergänge zwischen grauer und weißer Substanz hervor, das „Thickness image“ markiert Bereiche mit abnormer kortikaler Dicke und das „Extension image“ zeigt eine untypische Ausdehnung der grauen Substanz in die weiße Substanz.

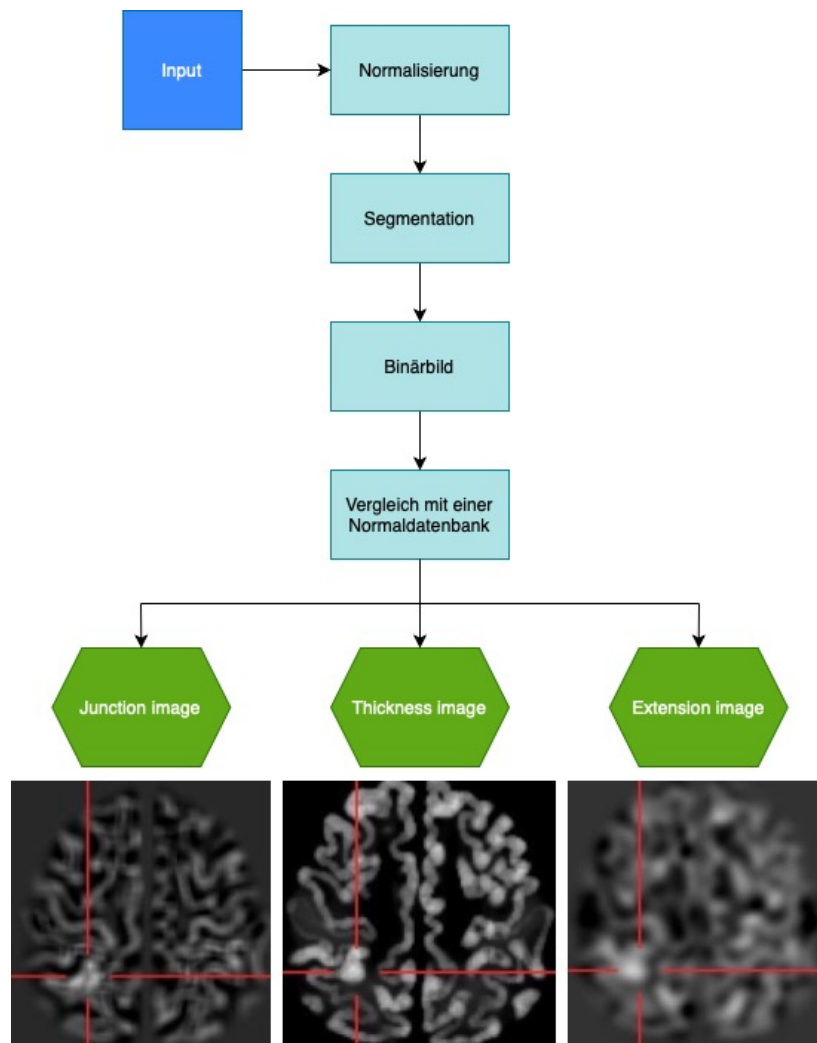


Abbildung 2: Überblick über die Funktionsweise von MAP18

Unten in der Abbildung sind exemplarisch die axialen z-Score-Karten von T1-gewichteten Sequenzen für die in [Abbildung 1a](#) beschriebene FCD IIb dargestellt.

1.2.3 Convolutional Neural Network

Die Architektur eines typischen CNN besteht aus mehreren Blöcken von Convolutional layers und Pooling layers, gefolgt von einem oder mehreren Fully connected layers (s. [Abbildung 3](#)). Die Eingabedaten durchlaufen schrittweise die einzelnen Schichten des Netzwerks, was als Vorwärtspropagation bezeichnet wird. Jede Schicht erzeugt dabei neue Informationen in Form von Merkmalskarten (Feature maps).

Die Convolutional layers extrahieren Merkmale aus den Eingabedaten mithilfe kleiner Matrizen (Kernels), deren Größe und Anzahl als Hyperparameter definiert sind. Mehrere Kernels erzeugen dabei verschiedene Merkmalskarten. Die Ergebnisse dieser linearen Faltungsoperationen werden durch eine nichtlineare Aktivierungsfunktion wie die Rectified Linear Unit (ReLU; berechnet die Funktion $f(x) = \max(0, x)$) transformiert, um komplexe Beziehungen in den Daten zu modellieren.

Pooling layers reduzieren anschließend die räumliche Dimension der Merkmalskarten, indem sie z. B. den größten Wert in einem Bereich auswählen (Max pooling) oder den Durchschnitt berechnen (Global averaging pooling). Dies führt zu einer Verringerung der Anzahl der lernbaren Parameter, reduziert die Rechenkomplexität und erhöht die Robustheit des Modells gegenüber kleinen Verzerrungen.

Am Ende des letzten Convolutional layers oder Pooling layers werden die Merkmalskarten in eine flache Datenstruktur umgewandelt (Flattening) und mit einer oder mehreren Fully connected layers verbunden. Dabei wird jede Eingabe mit jeder Ausgabe durch lernfähige Gewichtungen (Weights) verknüpft. Fully connected layers werden häufig von nichtlinearen Aktivierungsfunktionen wie ReLu begleitet, um komplexe Zusammenhänge zwischen den Merkmalen zu erfassen. Der letzte Fully connected layer trifft die endgültige Entscheidung, z. B. die Berechnung von Wahrscheinlichkeiten für spezifische Klassen bei Klassifizierungsaufgaben.

Beim Training eines CNN werden die lernbaren Parameter so angepasst, dass die Differenz zwischen den Modellvorhersagen und den tatsächlichen Labels (Ground-Truth) minimiert wird. Die Verlustfunktion misst hierfür diese Differenz. Mithilfe von Rückwärtspropagation und Gradientenabstieg werden die lernbaren Parameter iterativ optimiert, um den Verlust zu reduzieren (Yamashita et al., 2018).

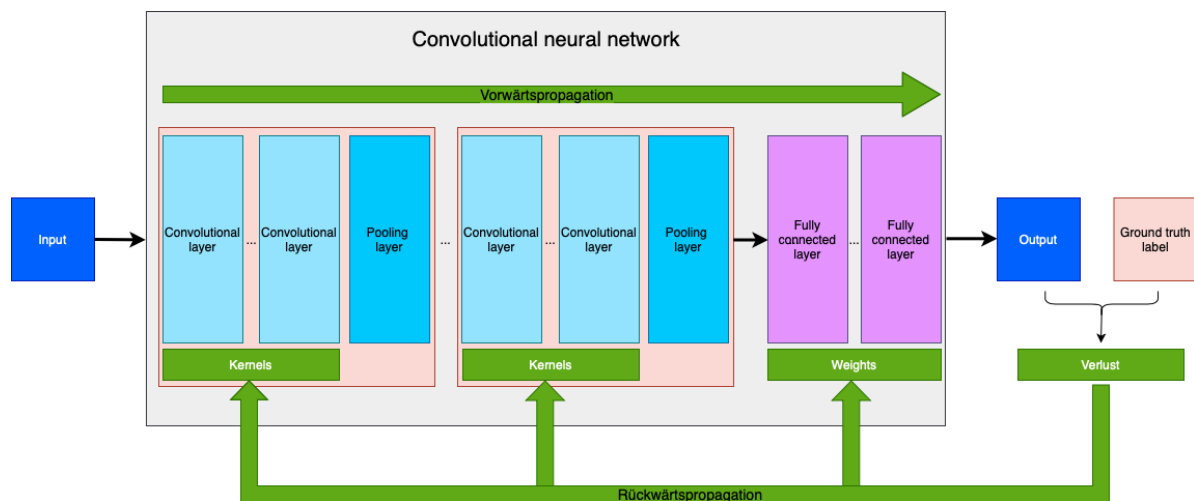


Abbildung 3: Überblick über die allgemeine Funktionsweise und Architektur eines CNN (Yamashita et al., 2018)

Das von House et al. (2021) entwickelte neuronale Netzwerk basiert auf einer Variante der Encoder-Decoder-Architektur, dem U-Net, das für seine gute Leistung bei einer Vielzahl von medizinischen Segmentierungsaufgaben bekannt ist (Ronneberger et al., 2015). Ein zusätzlicher Variational Decoder (Myronenko, 2018) dient ausschließlich der Regularisierung, um die Generalisierungsfähigkeit des Modells zu erhöhen und Überanpassungen an die Trainingsdaten zu vermeiden. Die Implementierung des CNN erfolgte mithilfe des Open-Source-Frameworks PyTorch (Paszke et al., 2019). Die

Netzwerkarchitektur des CNN ist in [Abbildung 4](#) grafisch dargestellt und wird ausführlich in Chanra et al. (2024), [Section 2.5](#), beschrieben.

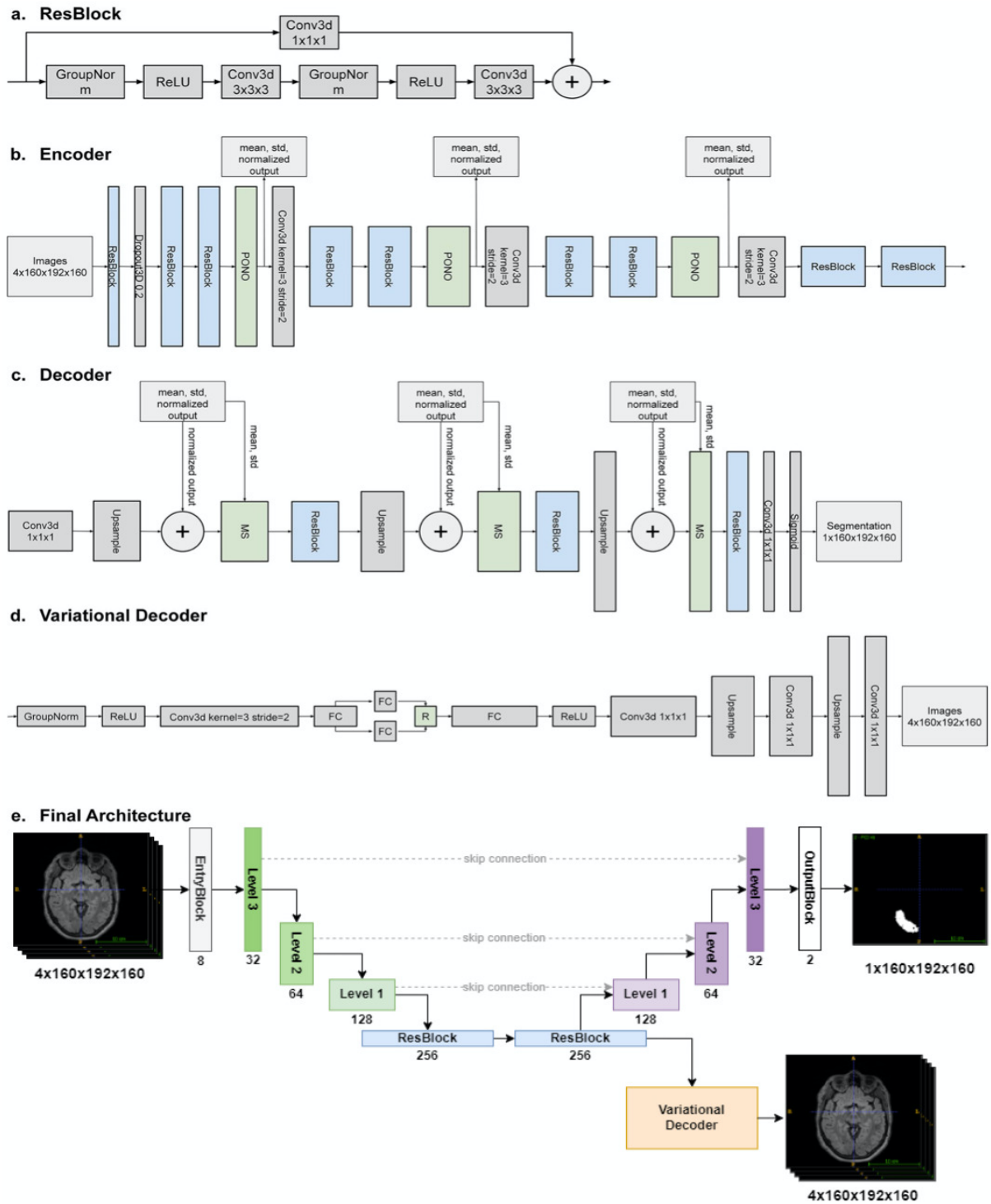


Abbildung 4: Netzwerkarchitektur (House et al., 2021; Chanra et al., 2024)

1.2.4 Trainings- und Auswertungsstrategie

Volumetrische FCD-Läsionskarten wurden manuell vom Erstautor voxelweise in axialer, koronarer und sagittaler Orientierung an T1- und FLAIR-Scans mit ITK-SNAP (Version 3.8.0) erstellt (s. [Abbildung 5](#)) und anschließend vom Letztautor validiert. Diese 3D-Läsionskarten dienten als Referenzdaten (Ground-Truth) und wurden zusammen mit anonymisierten DICOM-Daten für das Training des CNN verwendet.

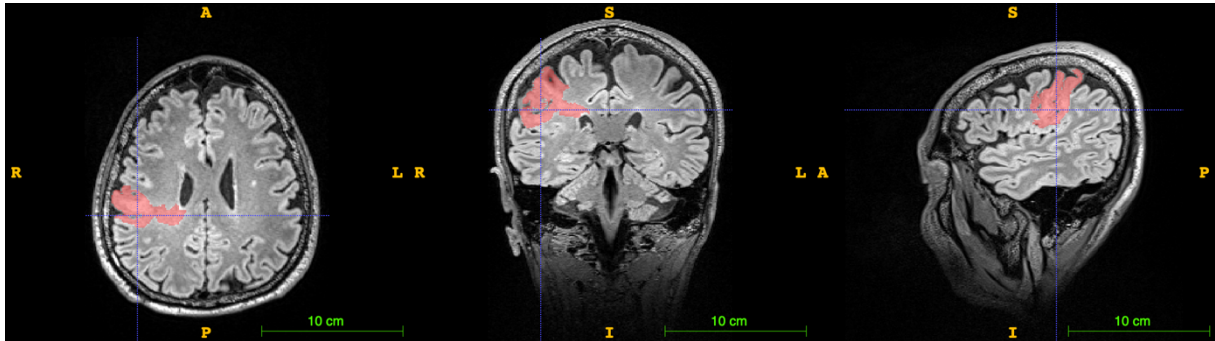


Abbildung 5: Volumetrische FCD-Läsionskarten

Beispielhafte Darstellung einer volumetrischen FCD-Läsionskarte auf einer FLAIR-Sequenz. Die Segmentierungsergebnisse der drei CNN-Modelle (Basismodell, OWD, CLOND) für diese FCD sind in Chanra et al. (2024), [Figure 5a](#), dargestellt.

Der Datensatz wurde in drei Teildatensätze („Datensatz 1–3“) mit jeweils 100 MRT-Bildern aufgeteilt, darunter 10 FCD-Fälle, 50 normale Fälle und 40 pathologische Nicht-FCD-Fälle. Das von House et al. (2021) entwickelte CNN wurde als „Basismodell“ verwendet. Der zugehörige Trainingsdatensatz wurde als „Basismodell-Datensatz“ bezeichnet.

Es wurde eine dreiphasige Trainings- und Auswertungsstrategie entwickelt, um klassisches Training und kontinuierliches Training zu vergleichen. Das OWD-Modell („Old model with whole dataset“) wurde mit dem gesamten Datensatz nach einem klassischen Trainingsansatz trainiert, während beim CLOND-Modell („Continual learning with old model and new dataset“) ausschließlich neue Daten mittels kontinuierlicher Lernstrategie verwendet wurden. Der Trainings- und Auswertungsprozess ist ausführlich in [Tabelle 5](#) beschrieben.

In jeder Trainingsphase wurden beide Modelle jeweils über 100 Epochen trainiert. Hierbei wurde der Adam Optimizer mit einem Weight Decay von 10^{-5} verwendet. Die anfängliche Lernrate betrug 10^{-4} und nahm während des Trainings schrittweise gemäß der folgenden Formel ab:

$$\alpha = \alpha_0 \cdot \left(1 - \frac{e}{N_e}\right)^{0.9},$$

e bezeichnet die aktuelle Epoche und N_e die Gesamtanzahl der Epochen ($N_e = 100$). Für das Training wurde eine Batch-Größe von 2 verwendet.

Tabelle 5: Dreiphasige Trainings- und Auswertungsstrategie

Phase	Schritt	Beschreibung
1. Phase	Analyse	Der erste Datensatz („Datensatz 1“) wurde vom Basismodell analysiert.
	Training	a) Klassisches Training: Das Basismodell wurde mit dem gesamten verfügbaren Datensatz („Basismodell-Datensatz“ und „Datensatz 1“) trainiert. Das resultierende Modell wurde als „OWD“ bezeichnet. b) Kontinuierliches Lernen: Das Basismodell wurde ausschließlich mit den neu hinzugekommenen Daten („Datensatz 1“) trainiert. Das resultierende Modell wurde als „CLOND“ bezeichnet.
2. Phase	Analyse	Der zweite Datensatz („Datensatz 2“) wurde von allen drei Modellen (Basismodell, OWD, CLOND) analysiert.
	Training	a) Klassisches Training: Das Basismodell wurde mit dem gesamten bisherigen Datensatz („Basismodell-Datensatz“, „Datensatz 1“ und „Datensatz 2“) weitertrainiert. Das resultierende Modell wurde als „Modifiziertes OWD“ bezeichnet. b) Kontinuierliches Lernen: Das CLOND-Modell wurde ausschließlich mit „Datensatz 2“ weitertrainiert. Das resultierende Modell wurde als „Modifiziertes CLOND“ bezeichnet.
3. Phase	Analyse	Der dritte Datensatz („Datensatz 3“) wurde vom Basismodell sowie vom modifizierten OWD- und CLOND-Modell ausgewertet.

Abkürzungen: CLOND – „Continual learning with old model and new dataset“; OWD – „Old model with whole dataset“

1.2.5 Kontinuierliches Lernen und Elastic Weight Consolidation

Beim Training eines CNN werden die Filter (Kernels) und Gewichtungen (Weights) angepasst, um die Differenz zwischen den Vorhersagen und den tatsächlichen Labels zu minimieren. Hierzu berechnet das Netzwerk den Fehler mithilfe der Verlustfunktion und optimiert die Parameter über den Rückwärtspropagation-Algorithmus.

Abbildung 6 illustriert die Problematik des sequenziellen Lernens in neuronalen Netzwerken. Bei fehlender Regularisierung führen Anpassungen der Parameter für eine neue Aufgabe häufig zu einem Verlust der Leistung bei der vorherigen Aufgabe, was zum „katastrophalen Vergessen“ (schwarzer Pfeil) führt. Eine zu starke Regularisierung hingegen kann das Lernen der neuen Aufgabe einschränken und möglicherweise auch die Leistung der ersten Aufgabe beeinträchtigen (blauer Pfeil). Eine flexible Regularisierung mittels kontinuierlicher Lernmethoden ermöglicht es, die Leistung beider Aufgaben zu bewahren, indem die Parameter in einem kontrollierten Bereich angepasst werden (roter Pfeil).

Methoden des kontinuierlichen Lernens lassen sich konzeptionell in fünf Gruppen unterteilen (s. Tabelle 6). Regularisierungsbasierte Methoden bieten den Vorteil, dass sie mit geringem Implementierungsaufwand flexibel in bestehende Modelle integriert werden können, ohne alte Daten vorhalten

zu müssen. Ein prominentes Beispiel für diesen Ansatz ist die Elastic Weight Consolidation (EWC) (Aich, 2021), die entscheidende Modellparameter vor starker Veränderung schützt, um bestehendes Wissen zu bewahren.

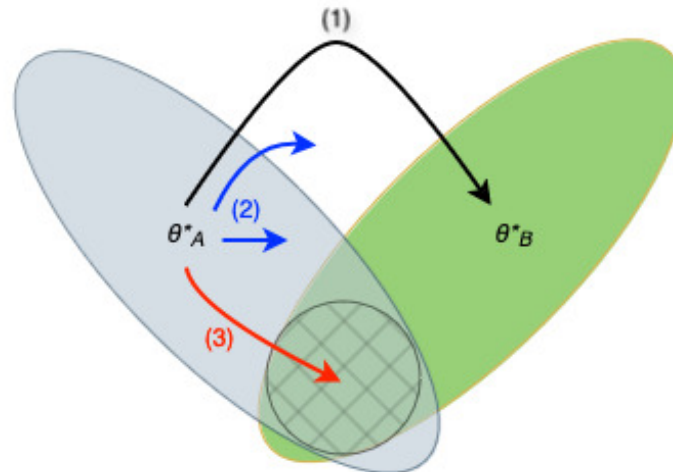


Abbildung 6: Sequenzielles Lernen und Elastic Weight Consolidation

Die Abbildung illustriert die Problematik des sequenziellen Lernens zweier Aufgaben (A und B) in einem neuronalen Netzwerk. Beim Training eines neuronalen Netzwerks werden die Parameter θ optimiert, um den Fehler bei der jeweiligen Aufgabe zu minimieren (optimales $\theta = \theta^*$). θ^*_A (grauer Bereich) und θ^*_B (grüner Bereich) repräsentieren die Parameterbereiche, die zu einer guten Leistung bei den Aufgaben A bzw. B führen. Nach dem Lernen der Aufgabe A befinden sich die Parameter bei θ^*_A .

- (1) Schwarzer Pfeil: Ohne Regularisierung des Trainings auf Aufgabe B verschieben sich die Parameter von θ^*_A zu θ^*_B , wodurch die Leistung für Aufgabe A vollständig verloren geht („katastrophales Vergessen“).
- (2) Blauer Pfeil: Bei übermäßiger Regularisierung wird die Anpassung der Parameter an θ^*_B stark eingeschränkt, was das Lernen von Aufgabe B verhindert und möglicherweise auch die Leistung für Aufgabe A beeinträchtigt.
- (3) Roter Pfeil: Eine flexible Regularisierung durch EWC ermöglicht dem Netzwerk die optimalen Parameter $\theta^*_{A,B}$ (schraffierter Bereich) zu erlernen, die sowohl die Leistung für Aufgabe A bewahren als auch eine effektive Anpassung an Aufgabe B erlauben (Kirkpatrick et al., 2017; Aich, 2021).

Tabelle 6: Methoden des kontinuierlichen Lernens (Wang et al., 2024)

Methoden	Beschreibung
Regularisierungsbasierter Ansatz	Hinzufügen von Regularisierungsbedingungen, um Parameter (Gewichtsregularisierung) oder Verhaltensweisen (Funktionsregularisierung) des alten Modells zu bewahren
Wiederholungsbasierter Ansatz	Annäherung und Wiederherstellung alter Daten, z.B. durch Speicherung von Trainingsbeispielen (experience replay), Training eines generativen Modells (generative replay) oder Speicherung von Prototypen oder statistischer Informationen zur Wiederherstellung alter Merkmale (feature replay)
Optimierungsbasierter Ansatz	Anpassung des Lernprozesses (Optimierungsstrategie), um Interferenzen zwischen den Aufgaben zu minimieren, z.B. durch spezielle Lernverfahren wie Metalernen oder die Steuerung der Parameteränderungen mittels Gradientenprojektion
Repräsentationsbasierter Ansatz	Lernen allgemeiner Muster oder Strukturen in den Daten (Repräsentationen), z. B. durch Vortraining und Wiederverwendung dieser Grundlagen für neue Aufgaben
Architekturbasierter Ansatz	Dynamische Anpassung der Netzwerkarchitektur für neue Aufgaben, z. B. durch Zuweisung spezifischer Parameter für jede Aufgabe (Parameterzuweisung), Hinzufügen aufgabenspezifischer Module (modularer Netzwerkansatz) oder Zerlegung in gemeinsame und aufgabenspezifische Komponenten (Modellzerlegung)

EWC betrachtet das Training neuronaler Netzwerke aus einer probabilistischen Perspektive. Die Bayes'sche Regel kombiniert die Prior $\log p(\theta|\mathcal{D}_A)$, die das Wissen für Aufgabe A bewahrt, mit der Likelihood $\log p(\mathcal{D}_B|\theta)$, die die Anforderungen der Parameter θ für die neue Aufgabe B berücksichtigt. Daraus ergibt sich die Posterior-Wahrscheinlichkeit, die das alte Wissen mit den neuen Anforderungen integriert:

$$\log p(\theta|\mathcal{D}) = \log p(\mathcal{D}_B|\theta) + \log p(\theta|\mathcal{D}_A) - \log p(\mathcal{D}_B),$$

$\mathcal{D} = \{\mathcal{D}_A, \mathcal{D}_B\}$ beschreibt den kombinierten Datensatz aller Aufgaben. Der Normalisierungsfaktor $-\log p(\mathcal{D}_B)$ skaliert die Verteilung.

Die Fisher-Informationsmatrix identifiziert die für Aufgabe A relevanten Parameter, indem sie die Sensitivität der Log-Likelihood gegenüber diesen Parametern quantifiziert. Basierend auf diesen Berechnungen schützt die Verlustfunktion wichtige Parameter durch einen Regularisierungsterm. Die Verlustfunktion lautet:

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i (\theta_i - \theta_{A,i}^*)^2,$$

$\mathcal{L}_B(\theta)$ beschreibt den Verlust für Aufgabe B, λ legt die relative Gewichtung der alten Aufgabe im Vergleich zur neuen fest und i indiziert die einzelnen Parameter.

Der Regularisierungsterm führt eine quadratische Strafe ein, die Abweichungen von den optimalen Werten für Aufgabe A bestraft. Weniger wichtige Parameter bleiben hingegen flexibel und können sich an die Anforderungen der neuen Aufgabe B anpassen (Kirkpatrick et al., 2017; Aich, 2021).

1.2.6 Auswertung

Eine FCD wurde anhand typischer neuroradiologischer Merkmale in T1- oder FLAIR-Bildern diagnostiziert (Urbach et al., 2022; Barkovich und Kuzniecky, 1996). Die manuell erstellten FCD-Läsionskarten dienten als Referenzgrundlage (Ground-Truth-Mask, GTM) für die Auswertung.

Die Bewertung der Modelle erfolgte auf Fall-, Läsions- und Voxel Ebene:

1. Klassifikationskriterien:

- Fallebene:
 - True Positive (TP): Eine FCD wurde korrekt erkannt und der Dice-Koeffizient beträgt $\geq 0,1$.
 - False Negative (FN): Eine FCD wurde nicht erkannt oder der Dice-Koeffizient beträgt $< 0,1$.
 - True Negative (TN): Kontrollgruppen-MRTs, in denen keine Cluster detektiert wurden.

- False Positive (FP): Kontrollgruppen-MRTs, in denen Cluster registriert wurden.
- Läsions- und Voxel Ebene:
 - True Positive (TP): Cluster bzw. Voxel überlappen mit der GTM in FCD-MRTs.
 - False Positive (FP): Cluster bzw. Voxel in Kontrollgruppen-MRTs oder ohne Überlappung mit der GTM in FCD-MRTs

2. Metriken:

- Fallebene:

$$\text{Sensitivität} = \frac{TP}{TP + FN}$$

$$\text{Spezifität} = \frac{TN}{TN + FP}$$

$$\text{Genauigkeit} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Balancierte Genauigkeit} = \frac{\text{Sensitivität} + \text{Spezifität}}{2}$$

$$\text{Präzision auf Fallebene} = \frac{TP}{TP + FP}$$

- Läsionsebene:

$$\text{Präzision} = \frac{TP \text{ cluster}}{TP \text{ cluster} + FP \text{ cluster}}$$

- Voxel Ebene:

$$\text{Sensitivität} = \frac{TP \text{ voxel}}{GTM}$$

$$\text{Präzision} = \frac{TP \text{ voxel}}{TP \text{ voxel} + FP \text{ voxel}}$$

Die Leistung der FCD-Segmentierung wurde anhand des Dice-Koeffizienten bewertet:

$$\text{Dice-Koeffizient} = \frac{2 * |A \cap B|}{|A| + |B|},$$

|A| beschreibt das Volumen der Modellvorhersage und |B| das Volumen der FCD-Läsionskarte.

Zusätzlich wurden Receiver-Operating-Characteristic-(ROC-)Kurven erstellt und die Fläche unter der Kurve (AUC) berechnet. Der Modellvergleich erfolgte mittels Intersection-Union-Test, wobei ein Modell als überlegen galt, wenn es eine höhere Spezifität und eine Sensitivität innerhalb einer Toleranzgrenze von 10 % aufwies.

Die statistischen Analysen wurden mit Microsoft Excel (Version 16.69.1) und R (Version 4.2.3, R Foundation for Statistical Computing, Wien, Österreich) durchgeführt. Die Erstellung der ROC-Kurven erfolgte mithilfe von XLSTAT statistical and data analysis solutions (Version 2023.1.3., Lumivero, New York, USA).

1.3 Ergebnisse

Die Ergebnisse sind in [Tabelle 7](#) und [Abbildung 7](#) zusammengefasst.

Nach der ersten Trainingsphase erreichten alle drei Modelle auf Fallebene eine Sensitivität von 80 % (95%-KI: 55,2 % – 100 %). Nach der zweiten Trainingsphase lag die Sensitivität des Basismodells und von CLOND bei 90 % (95%-KI: 71,4 % – 100 %), während OWD 100 % erreichte.

Die Spezifität verbesserte sich bei OWD und CLOND stetig im Verlauf des Trainingsprozesses. In der finalen Auswertung erzielte CLOND mit 70 % (95%-KI: 60,5 % – 79,5 %) die höchste Spezifität.

Hinsichtlich der Genauigkeit, der ausgewogenen Genauigkeit und der Präzision auf Fall-, Läsions- und Voxel Ebene übertraf OWD das Basismodell, während CLOND in allen diesen Metriken die besten Ergebnisse erzielte.

Die Segmentierungsleistung verbesserte sich ebenfalls durch das Training. Das Basismodell erzielte einen mittleren Dice-Koeffizienten von 0,46 (95%-KI: 0,26 – 0,65), während CLOND mit einem mittleren Dice-Koeffizienten 0,56 (95%-KI: 0,24 – 0,88) in der abschließenden Auswertung die beste Segmentierungsleistung zeigte.

Tabelle 7: Auswertung der Modellleistung auf Voxel-, Läsions- und Fallebene

Modell	Phase (jeweils n=100)	TP Voxel	FP Voxel	TP Cluster	FP Cluster	TP Fälle	FN Fälle	TN Fälle	FP Fälle
Basismodell	1	63.416	546.493	13	241	8	2	0	90
	2	40.446	452.841	8	219	8	2	2	88
	3	88.873	473.330	9	216	9	1	5	85
OWD	2	62.934	333.689	12	127	8	2	19	71
	3	110.435	148.135	10	63	10	0	51	39
CLOND	2	71.324	249.794	10	87	8	2	38	52
	3	120.765	118.976	10	41	9	1	63	27

Modell	Phase (jeweils n=100)	Sensitivität (%)	Spezifität (%)	Genauigkeit (%)	Balancierte Genauigkeit (%)	Dice	SV (%)	PV (%)	PL (%)	PF (%)	AUC	FP Cluster pro MRT
Basismodell	1	80,0	0,0	8,0	40,0	0,45	22,4	10,4	5,1	8,2	0,87	2,41
	2	80,0	2,2	10,0	41,1	0,43	15,3	8,2	3,5	8,3	0,87	2,19
	3	90,0	5,6	14,0	47,8	0,50	33,3	15,8	4,0	9,6	0,97	2,16
OWD	2	80,0	21,1	27,0	50,6	0,50	23,7	15,9	8,6	10,1	0,93	1,27
	3	100	56,7	61,0	78,4	0,52	41,4	42,7	13,7	20,4	0,96	0,63
CLOND	2	80,0	42,2	46,0	61,1	0,55	26,9	22,2	10,3	13,3	0,94	0,87
	3	90,0	70,0	72,0	80,0	0,56	45,3	50,4	19,6	25,0	0,98	0,41

Abkürzungen: AUC – Fläche unter der Kurve (Area under the curve); Dice – Dice-Koeffizient; FN – False Negative; FP – False Positive; PF – Präzision auf Fallebene; PL – Präzision auf Läsionsebene; PV – Präzision auf Voxel Ebene; SV – Sensitivität auf Voxel Ebene; TN – True Negative; TP – True Positive

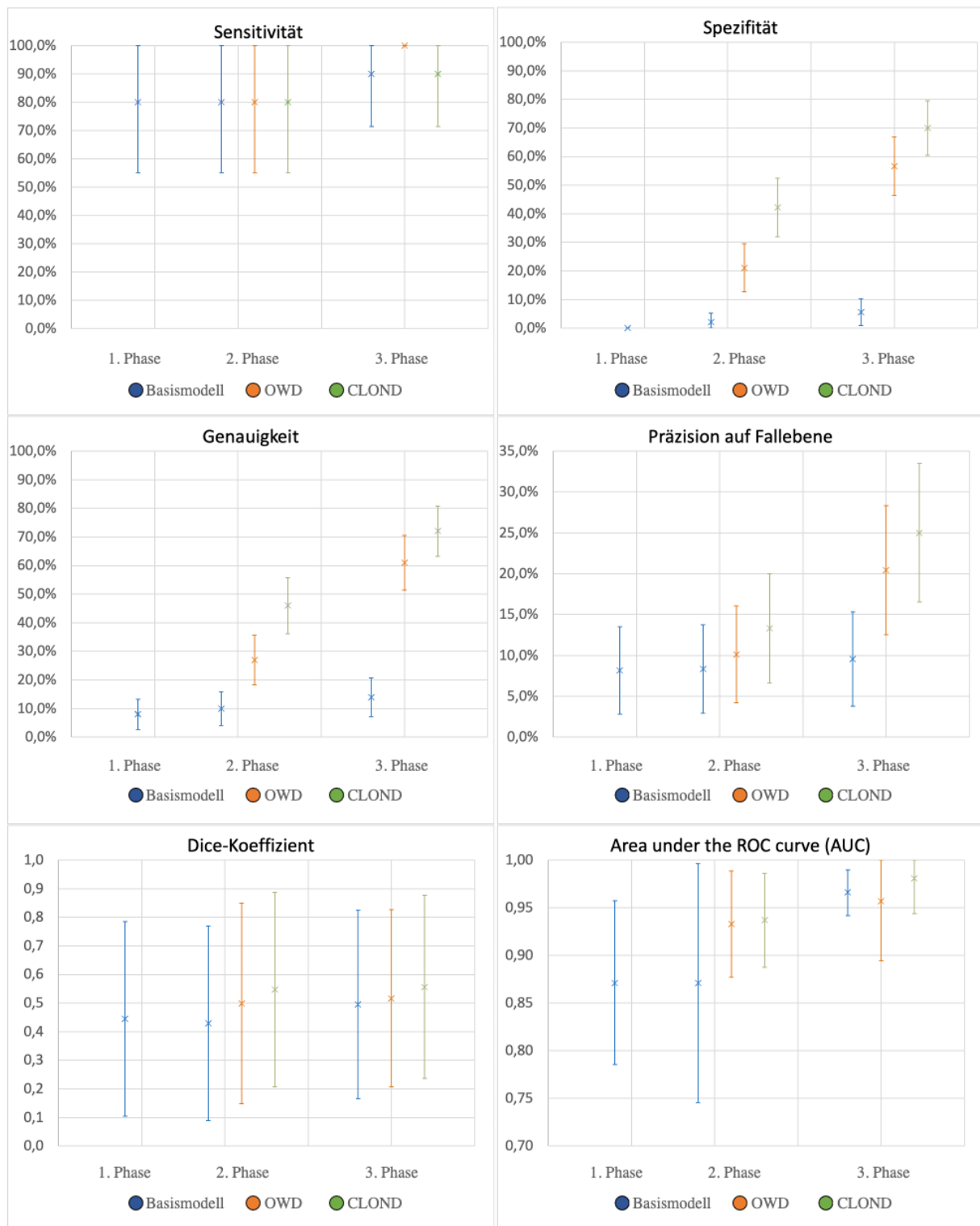


Abbildung 7: Leistungsmetriken der CNN-Modelle auf Fallebene mit 95%-Wald-Konfidenzintervallen

Das Basismodell identifizierte kumulativ 676 FP-Cluster in 300 MRTs (durchschnittlich 2,3 FP-Cluster pro MRT, SD = 1,2). Durch das Training konnte die Anzahl der FP-Cluster signifikant reduziert werden. In der abschließenden Auswertung erzielte CLOND mit durchschnittlich 0,4 FP-Clustern pro MRT (SD = 0,79) das beste Ergebnis (s. [Abbildung 8](#)).

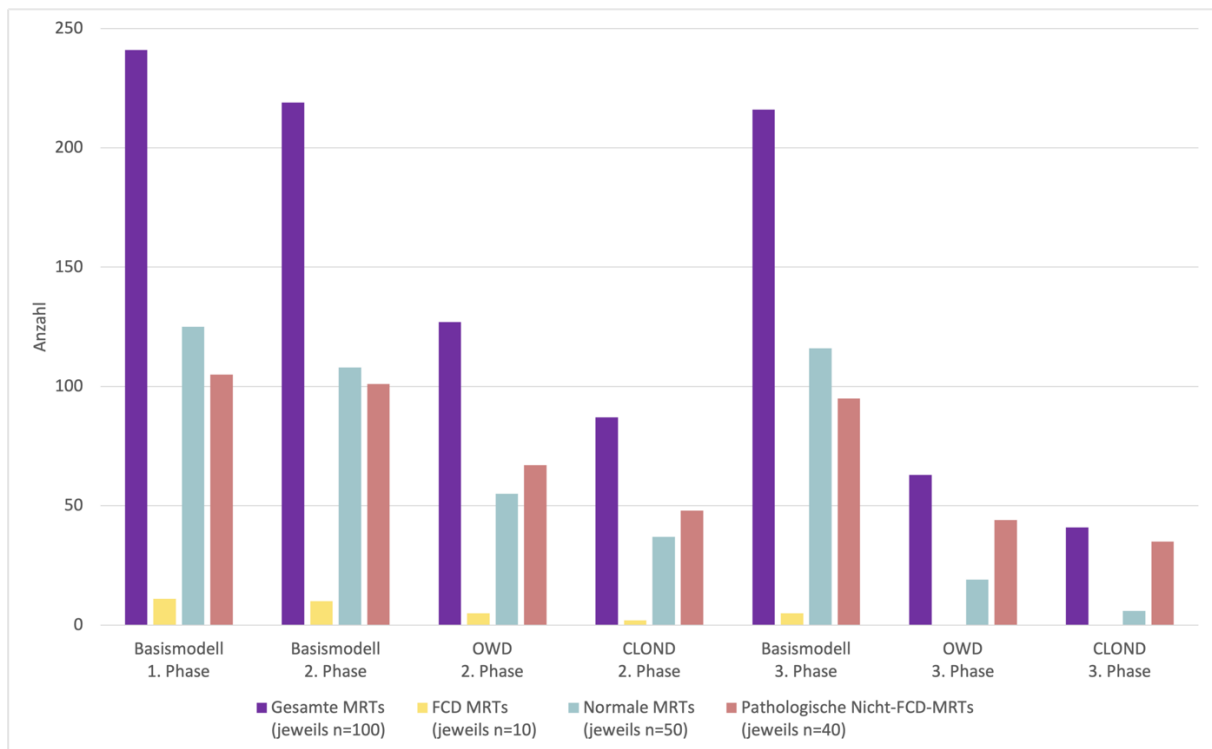


Abbildung 8: Anzahl der falsch-positiven Cluster

Der Intersection-Union-Test bestätigte die Überlegenheit von CLOND gegenüber OWD und dem Basismodell; zudem war OWD dem Basismodell überlegen (s. [Abbildung 9](#)).

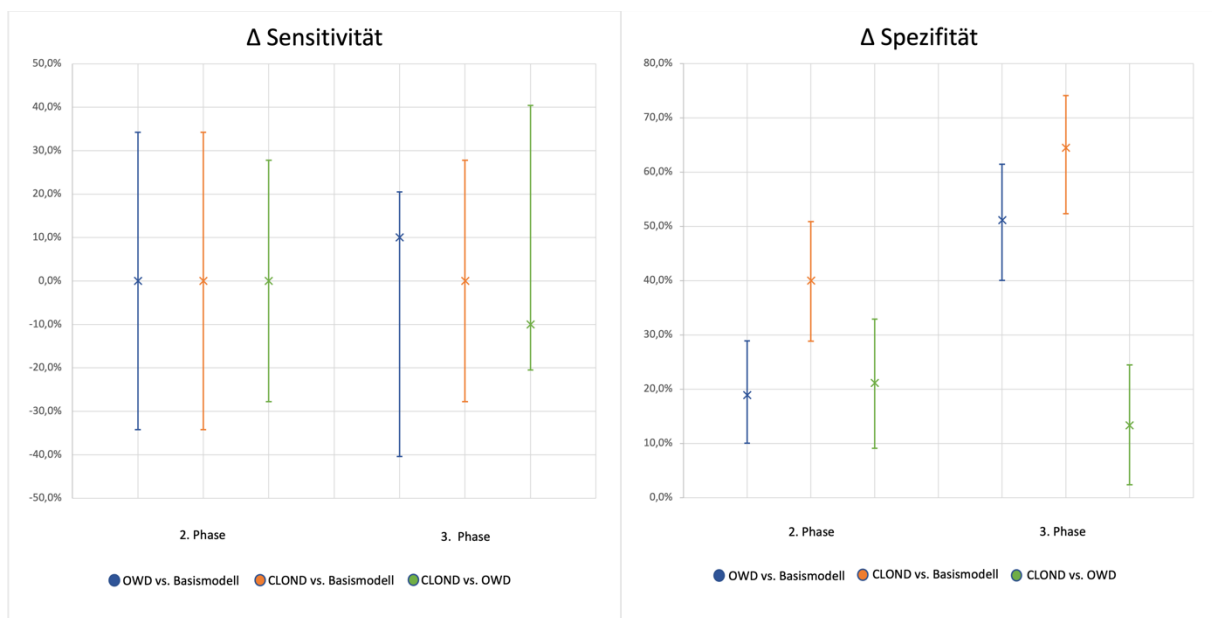


Abbildung 9: Unterschiede in Sensitivität und Spezifität zwischen den Modellen mit 95%-asymptotischen Konfidenzintervallen nach Tango

In der ROC-Analyse erzielten alle Modelle konsistent hohe AUC-Werte. Das Basismodell zeigte in den drei Auswertungsphasen AUC-Werte zwischen 0,87 und 0,97. In der finalen Analyse erreichte CLOND mit einer AUC von 0,98 (95%-KI: 0,94–1) die beste Gesamtleistung (s. [Abbildung 10](#)).

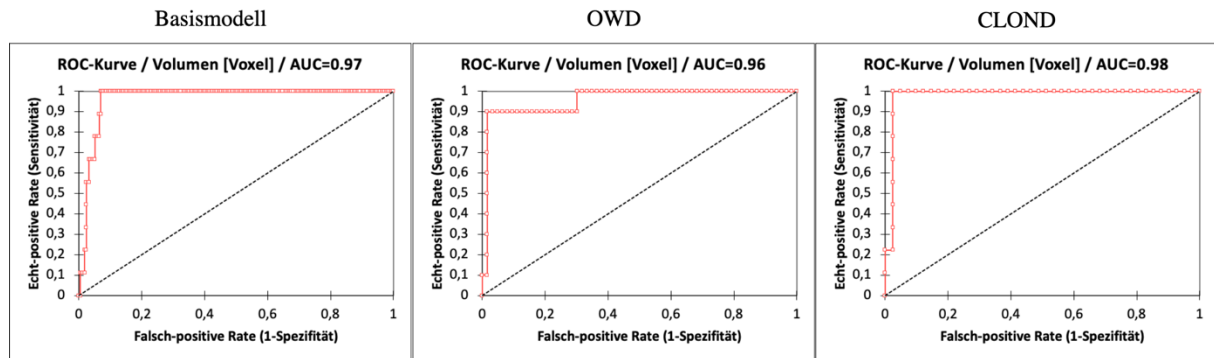


Abbildung 10: ROC-Kurven der drei Modelle in der finalen Auswertungsphase

1.4 Diskussion

In dieser Studie wurde das Basismodell mittels des kontinuierlichen Lernansatzes (CLOND-Modell) und des klassischen Ansatzes (OWD-Modell) mit einem prospektiv erhobenen MRT-Datensatz aus der klinischen Routine trainiert und vergleichend ausgewertet.

Das Basismodell erzielte über alle Auswertungsphasen hinweg stabile Ergebnisse und zeigte sich robust gegenüber Datenvariationen, was wichtig für den klinischen Einsatz ist. Beide Trainingsmethoden führten zu einer Verbesserung der Leistungsmetriken, ohne die Sensitivität zu beeinträchtigen. Das CLOND-Modell übertraf das OWD-Modell und erzielte eine höhere Spezifität (70,0 %, $p < 0,05$), reduzierte FP-Cluster auf 0,4 pro MRT ($p < 0,05$) und verbesserte die FCD-Segmentierung (mittlerer Dice-Koeffizient 0,56). Diese Ergebnisse wurden bei klinischen Datenverteilungen erreicht, bei denen 90 % der MRTs aus Nicht-FCD-Daten bestanden. Zudem erzielte das CLOND-Modell diese Verbesserungen bei geringeren Anforderungen an Datensatzgröße, Rechenleistung und Speicherbedarf, was seine Effizienz und Anwendbarkeit im klinischen Alltag hervorhebt.

Durch das Training wurden FP-Cluster in infratentoriellen Regionen vollständig beseitigt und das initial überproportionale Auftreten von FP-Clustern in frontalen und zentralen Regionen ging ebenfalls deutlich zurück. FP-Cluster traten vor allem bei pathologischen Nicht-FCD-MRTs auf, da deren komplexe Bildmerkmale teilweise denen von FCDs ähneln und die Differenzierung erschweren. Dennoch reduzierte das Training FP-Cluster auch in dieser Gruppe, was für die Anpassungsfähigkeit des Modells spricht. Da die einzelnen Pathologien derzeit nur 2,5 bis 7,5 % des Datensatzes ausmachen, könnte eine Erweiterung des Trainings um zusätzliche Pathologien die Ergebnisse weiter verbessern.

Ein Cut-off-Wert von 1,8 ml für das FCD-Volumen hätte die Spezifität von CLOND auf 85,4 % (95%-KI: 71,1 % – 93,4 %) steigern können, ohne die Sensitivität zu beeinträchtigen. Angesichts der therapeutischen Relevanz kleiner Läsionen, der bisher unbekannten Mindestgröße von FCDs und der Tatsache, dass unser Datensatz vier FCDs $\leq 0,5$ ml enthielt, wurde jedoch auf die Einführung eines willkürlichen Schwellenwerts verzichtet.

Ein direkter Vergleich unseres CNN mit anderen Modellen ist aufgrund methodischer Unterschiede, wie heterogener Datensätze, Ground-Truth-Definitionen und Evaluierungsmetriken, schwierig. Viele Studien basieren zudem auf retrospektiven Daten, verwenden ausschließlich normale MRTs als Kontrollgruppe, setzen strenge Ausschlusskriterien und berichten ausschließlich über Kreuzvalidierungsergebnisse. Dies kann zu einer Überschätzung der Leistungsmetriken führen. Um diese Limitationen zu adressieren und die klinische Realität besser abzubilden, wurde unser CNN mit einem prospektiven Datensatz aus der klinischen Routine bewertet. Dieser Datensatz umfasste MRTs mit variabler Bildqualität und heterogenen Pathologien und wurde ohne Ausschlusskriterien erstellt. Trotz der eingeschränkten Vergleichbarkeit erzielt unser Modell vielversprechende Ergebnisse. [Tabelle 8](#) bietet eine vergleichende Übersicht über moderne Verfahren zur automatischen FCD-Detektion.

Dennoch weist die Studie einige Einschränkungen auf, wobei die bedeutendste das Fehlen einer histopathologischen Validierung der FCD-Diagnosen ist. Da keiner der Patienten der Studienpopulation bisher operiert wurde oder ein intrakranielles EEG erhalten hat, bleibt die Klassifizierung der Läsionen ungesichert. Zwar wurden die Diagnosen von erfahrenen Epileptologen anhand der Anfallssemiotik, EEG-Daten, neuroradiologischen Befunde und morphometrischen Analysen gestellt und alle 20 histologisch bestätigten FCDs im Trainingsdatensatz des Basismodells korrekt identifiziert (House et al., 2021). Dennoch besteht die Möglichkeit einer fehlerhaften Klassifikation der detektierten FCDs. Zudem könnten vorhandene FCDs im aktuellen MRT-Datensatz unentdeckt geblieben sein und einige als FP eingestufte Cluster könnten echten FCDs entsprechen.

Eine retrospektive Analyse pharmakoresistenter Patienten mit fokaler Epilepsie des Epilepsiezentrums Hamburg, die zwischen Januar 2014 und April 2023 an der Klinik für Neurochirurgie des Universitätsklinikums Hamburg-Eppendorf operiert wurden (s. [Tabelle 9](#)), ergab, dass die histopathologische Diagnose in 48 von 68 Fällen mit der präoperativen Verdachtsdiagnose einer kortikalen Fehlbildung unter Anwendung der genannten Methoden übereinstimmte. In 10 Fällen, in denen präoperativ eine FCD Typ IIIa vermutet wurde, konnte lediglich eine Hippocampusklerose bestätigt werden. In 4 der 20 histopathologisch negativen Fällen zeigte sich trotz unauffälliger Histopathologie eine postoperative Verbesserung der Anfallsfrequenz. Die histopathologische Untersuchung gilt zwar als Goldstandard zur Diagnose einer FCD, ist jedoch stark von der Qualität der Biopsieproben abhängig. Bei kleinen oder fragmentierten Gewebeproben können charakteristische Merkmale leicht übersehen

werden (Walger et al., 2023), sodass auch die Ergebnisse der histopathologischen Untersuchung keine vollständige Diagnosesicherheit gewährleisten. Zusammenfassend bieten die in der Studie angewandten Methoden, trotz fehlender histopathologischer Bestätigung, eine solide Grundlage für die FCD-Diagnose.

Eine weitere Einschränkung stellt die ausschließliche Nutzung monozentrischer MRT-Daten eines einzelnen Scanners dar. Dies kann zu einer Überschätzung der Modellleistung führen (Varoquaux, 2018). Zudem birgt es das Risiko einer Überanpassung (Overfitting), bei der das Modell spezifische Muster der Trainingsdaten lernt, die nicht auf unbekannte Daten übertragbar sind. Obwohl die Ergebnisse im unabhängigen Testdatensatz derzeit keine Hinweise auf Überanpassung zeigen, könnte die begrenzte Vielfalt der Trainingsdaten die Generalisierbarkeit des Modells einschränken.

Das Modell zeigt zudem Schwächen bei der Erkennung kleiner FCDs: In der Studie wurde lediglich eine von vier FCDs $\leq 0,5$ ml korrekt identifiziert. Diese Limitation ist vermutlich auf die geringe Anzahl kleiner FCDs im Trainingsdatensatz zurückzuführen. Darüber hinaus könnten Bildrauschen oder Artefakte subtile Merkmale überdeckt und die Erkennung zusätzlich erschwert haben.

Eine mögliche Lösung für diese Einschränkungen besteht in der Erweiterung des Trainingsdatensatzes durch die Integration von Daten aus verschiedenen Zentren, Scannern und Feldstärken sowie zusätzlichen Fällen kleiner FCDs. Allerdings wird dies durch die begrenzte Verfügbarkeit solcher Daten erschwert. Zudem ist die manuelle Annotation von FCDs zur Erstellung von Ground-Truth-Daten zeit- und kostenintensiv. Dies ist jedoch entscheidend für die Qualität der Trainingsdaten und somit für die Modellleistung entsprechend dem Grundsatz „Garbage in, garbage out“.

Eine Kombination verschiedener Ansätze könnte dazu beitragen, die Vielfalt der Trainingsdaten zu erhöhen und die Generalisierbarkeit des Modells zu verbessern. Datenaugmentation würde eine Erweiterung des Trainingsdatensatzes durch verschiedene Transformationen der vorhandenen Daten ermöglichen. Generative Adversarial Networks (GANs) bieten die Möglichkeit, synthetische MRT-Daten zu erzeugen, die verschiedene Scanner-Eigenschaften, Feldstärken und subtile Läsionen simulieren. CycleGANs könnten ferner genutzt werden, um Scannerunterschiede zu harmonisieren (Goceri, 2023). Pseudolabels erlauben die Nutzung unannotierter MRT-Daten, indem sie vorläufige Labels durch das bestehende Modell generieren (Mao et al., 2022). Die Integration von 7T-MRT-Daten (Chen et al, 2021), moderner Sequenzen wie MP2RAGE (Demerath et al., 2022) oder FLAWS (Sun et al., 2021) sowie die Kombination multimodaler Bildgebungsverfahren, beispielsweise 18F-Fluorodeoxyglucose-(FDG-)Positron-Emissions-Tomographie (PET) (Zheng et al., 2024), könnten die Läsionsidentifikation zusätzlich verbessern.

Um die Präzision neurochirurgischer Eingriffe zu erhöhen und die Planung invasiver EEG-Elektrodenimplantationen zu optimieren, könnte die Integration der generierten Läsionskarten in intraoperative

Navigationssysteme eine vielversprechende Strategie sein (Wellmer et al., 2010). Die Erweiterung des Modells auf die Detektion anderer epileptogener Läsionen, wie beispielsweise Heterotopien oder gar vaskuläre Pathologien, stellt eine weitere interessante Möglichkeit dar. Die Kombination dieser Techniken und die Weiterentwicklung des Modells bieten vielversprechende Perspektiven, deren Umsetzung jedoch weiterer Forschung bedarf.

1.5 Schlussfolgerungen

Durch datengesteuertes Training mittels eines umfangreichen 3T-MRT-Datensatzes aus der klinischen Routine und der Implementierung von kontinuierlichem Lernen konnte das CNN signifikant verbessert werden, insbesondere hinsichtlich der zuvor eingeschränkten Spezifität. Der kontinuierliche Lernansatz übertraf den klassischen Ansatz und ermöglichte eine fortlaufende Anpassung des Modells an neue Daten, ohne zuvor erlerntes Wissen zu verlieren – ein entscheidender Vorteil in klinischen Umgebungen mit variablen Datenbedingungen. Trotz Herausforderungen wie der fehlenden histopathologischen Validierung und der begrenzten Vielfalt der Trainingsdaten zeigt das Modell großes Potenzial für den klinischen Einsatz. Das CNN ermöglicht eine vollautomatische FCD-Detektion und ist dadurch für eine breitere klinische Anwendung geeignet. Eine verbesserte FCD-Detektion könnte die Planung und Durchführung neurochirurgischer Eingriffe erheblich optimieren und die Möglichkeiten der Epilepsiechirurgie bei pharmakoresistenten Patienten effektiver nutzen. Zukünftige Studien sollten darauf abzielen, die Vielfalt der Trainingsdaten zu erweitern, moderne Bildgebungsmethoden zu integrieren und das Modell auf andere epileptogene Läsionen auszuweiten.

Tabelle 8: Gegenüberstellung aktueller Methoden zur automatisierten FCD-Detektion

Studie	Studiendesign	Studienzentren	MRT-Scanner	MRT-Sequenzen	Trainingskohorte	Testkohorte	Sensitivität	Spezifität
Chanra et al. (2024)	Prospektiv	Monozentrisch	Siemens Skyra 3T	3D T1 & FLAIR	210 FCD-MRTs, 11 fokale PMG-MRTs, 200 normale MRTs, 130 pathologische MRTs	10 FCD-MRTs, 50 normale MRTs, 40 pathologische MRTs	90%	70%
Zheng et al. (2024)	Retrospektiv	Monozentrisch	3T (nicht spezifiziert)	3D T1, zusätzlich FDG-PET Aufnahmen	62 FCD-Fälle	20 FCD-Fälle	90%	-
Spitzer et al. (2022)	Retrospektiv	Multizentrisch (22 Zentren)	Siemens, GE, Philips 1,5T/3T	3D T1	278 FCD-Fälle, 180 gesunde Kontrollen	260 FCD-Fälle, 193 gesunde Kontrollen	59%	54%
David et al. (2021)	Retrospektiv	Multizentrisch	Training: 13 Scanner, Auswertung: 1 Scanner	3D T1	113 FCD-Fälle, 362 gesunde Kontrollen	60 FCD-Fälle, 70 gesunde Kontrollen	81%	84,3%
Demerath et al. (2021)	Retrospektiv	Monozentrisch	Siemens Magnetom Prisma 3T	3D T1 (MPRAGE & MP2RAGE)	Verwendung des ANN aus David et al. (2021)	32 FCD-Fälle, 32 gesunde Kontrollen	MPRAGE: 82,4% MP2RAGE: 97,1%	MPRAGE: 62,5% MP2RAGE: 34,4%
Ganji et al. (2021)	Retrospektiv	Monozentrisch	Philips Achieva 1,5T, Siemens 1,5T/3T	3D T1 & FLAIR	30 FCD-Fälle, 28 gesunde Kontrollen	Aufteilung des Datensatzes: 70% fürs Training, 15% für die Validierung, 15% fürs Testen	96,7%	100%
Gill et al. (2021)	Retrospektiv	Multizentrisch (9 Zentren)	Siemens, GE, Philips 1,5T/3T	3D T1 & FLAIR	148 FCD-Fälle	23 FCD-Fälle, 42 gesunde Kontrollen, 90 Kontrollen mit TLE und HS	83%	89%
House et al. (2021)	Prospektiv	Monozentrisch	Siemens Skyra 3T	3D T1 & FLAIR	194 FCD-MRT, 7 fokale PMG-MRTs, 100 normale MRTs, 50 pathologische MRTs	9 FCD-MRTs, 91 Kontroll-MRTs (normale und pathologische MRTs)	77,8%	5,5%
Feng et al. (2020)	Retrospektiv	Monozentrisch	Keine Angabe	3D FLAIR	34 MRTs von 19 FCD-Fällen, hiervon 7 FLAIR-positiv, 12 FLAIR-negativ		83,3%*	99,0%*
Wang et al. (2020)	Retrospektiv	Monozentrisch	Siemens 1,5T (FCD-Fälle), Philips 3T (gesunde Kontrollen), GE 1,5T (TLE)	3D T1	10 FCD-Fälle, 10 gesunde Kontrollen, 10 TLE		90%*	85%*
Bijay Dev et al. (2019)	Retrospektiv	Monozentrisch	Siemens 1,5T, GE 3T	3D FLAIR	43 FCD-Fälle	Aufteilung des Datensatzes: 80% für Training und Validierung, 20% fürs Testen	82,5%*	-
Mo et al. (2019)	Retrospektiv	Monozentrisch	3T (nicht spezifiziert)	3D T1 & FLAIR zusätzlich FDG-PET Aufnahmen	40 FCD-Fälle, 30 Fälle mit HS oder Epidermoidzyste	Aufteilung des Datensatzes: 70% fürs Training, 15% für die Validierung, 15% fürs Testen	70%*	69,9%*
Jin et al. (2018)	Retrospektiv	Multizentrisch (3 Zentren)	GE Discovery, Siemens Verio, Siemens Trio (alle 3 T)	3D T1	61 FCD-Fälle, 120 normale MRTs für die Normdatenbank, zusätzlich 35 gesunde Kontrollen und 15 Fälle mit HS zur Berechnung der Spezifität		73,7%*	90,0%*
Adler et al. (2017)	Retrospektiv	Monozentrisch	Siemens Avanto 1,5 T	3D T1 & FLAIR	22 FCD-Fälle, 28 gesunde Kontrollen		73%*	

Abkürzungen: HS – Hippocampussklerose; PMG – Polymikrogyrie; TLE – Temporallappenepilepsie
 Die mit einem Sternchen (*) markierten Werte stammen aus Kreuzvalidierungsergebnissen.

Tabelle 9: Vergleich von MAP18- und histopathologischen Befunden bei der Diagnose kortikaler Fehlbildungen der neuronalen Entwicklung (Malformations of Cortical Development, MCD) am Hamburger Epilepsie-Zentrum (2014–2023)

Fall	Alter zum OP-Zeitpunkt [Jahre]	Geschlecht	MAP18 Befund	Histopathologischer Befund
1	28	w	FCD links temporal + HS	FCD Typ IIIa + HS ILAE Typ 3
2	5	m	FCD rechts frontal	FCD Typ I
3	30	m	FCD links temporal + HS	FCD Typ IIIa + HS
4	13	w	FCD links temporo-parietal	FCD Typ IIb
5	13	m	FCD links frontal	FCD Typ IIb
6	30	w	FCD rechts frontal	FCD Typ IIb
7	46	w	FCD links temporal + HS	FCD Typ IIIa + HS
8	36	w	FCD rechts frontal	FCD Typ IIb
9	27	m	FCD rechts frontal	FCD Typ Ic
10	45	w	FCD rechts temporal	FCD Typ IIIa + HS ILAE Typ 2
11	28	m	FCD rechts frontal	FCD Typ Ic
12	20	w	FCD links temporal	FCD Typ IIIa + HS ILAE Typ 2
13	40	m	FCD rechts temporal + HS	FCD Typ IIIa + HS
14	41	w	FCD rechts frontal	FCD Typ IIb
15	22	w	FCD links temporal	FCD Typ IIIa + HS
16	56	m	FCD rechts frontal	FCD Typ IIa
17	54	m	FCD rechts parietal	FCD Typ IIb
18	25	m	FCD links frontal	FCD Typ IIb
19	20	w	FCD links frontal	FCD Typ IIb
20	60	w	FCD links frontal	FCD Typ IIb
21	19	m	FCD links frontal	FCD Typ IIb
22	24	w	FCD links temporal	FCD Typ IIIa + HS ILAE Typ 2
23	52	w	FCD links temporal + HS	FCD Typ IIIa + HS ILAE Typ 1
24	21	w	FCD rechts parietal	FCD Typ IIa
25	34	m	FCD links temporal + HS	FCD Typ IIIa + HS ILAE Typ 1
26	26	m	FCD links frontal	FCD Typ IIb
27	62	m	FCD rechts temporal	FCD Typ IIb
28	26	m	FCD links temporal	FCD Typ IIb
29	45	m	FCD rechts frontal	MOGHE
30	24	m	FCD links frontal	Verstreute dysmorphe Neurone
31	33	m	FCD rechts temporal	FCD Typ I
32	24	m	FCD rechts temporal	FCD Typ IIa
33	12	m	FCD links frontal	FCD Typ IIb
34	62	m	FCD links temporal + HS	FCD Typ IIIa + HS ILAE Typ 1
35	32	m	FCD links frontal	FCD Typ IIb
36	48	m	FCD rechts frontal	FCD Typ IIb
37	49	w	FCD links frontal	FCD Typ IIb
38	7	w	FCD links occipital	FCD Typ IIb
39	16	m	FCD links parietal	Komplexe kortikale Fehlbildung
40	26	m	FCD links temporal	FCD Typ IIb
41	23	m	FCD IIIa links temporal + HS	FCD Typ IIIa + HS ILAE Typ 1
42	35	m	FCD links frontal	FCD Typ IIb
43	29	w	FCD links frontal	FCD Typ IIb
44	45	w	FCD links frontal	FCD Typ IIa
45	17	m	FCD links temporo-occipital	FCD Typ IIa
46	24	w	FCD rechts frontal	FCD Typ IIa
47	28	w	FCD rechts parietal	FCD Typ IIb
48	17	m	FCD rechts parietal	FCD Typ Ia
49	48	w	FCD rechts temporal	Niedriggradiger glioneuraler Tumor
50	54	w	FCD IIIa rechts temporal + HS	HS, keine FCD
51	15	w	FCD IIIa links temporal + HS	HS, keine FCD
52	32	m	FCD IIIa rechts temporal + HS	HS ILAE Typ 1, keine FCD
53	38	w	FCD IIIa rechts temporal + HS	HS ILAE Typ 1, keine FCD
54	17	m	FCD IIIa links temporal + HS	HS ILAE Typ 1, keine FCD
55	28	m	FCD IIIa links temporal + HS	HS ILAE Typ 2, keine FCD
56	53	m	FCD IIIa rechts temporal + HS	HS ILAE Typ 3, keine FCD
57	50	m	FCD IIIa rechts temporal + HS	HS ILAE Typ 1, keine FCD

58	25	w	FCD IIIa rechts temporal + HS	HS, keine FCD
59	22	w	FCD IIIa rechts temporal + HS	HS ILAE Typ 1, keine FCD
60	21	w	FCD rechts temporal	Unauffällige Histopathologie
61	30	m	FCD links temporal	Unauffällige Histopathologie
62	23	w	FCD links frontal	Unauffällige Histopathologie
63	34	w	FCD rechts temporal	Unauffällige Histopathologie
64	37	m	FCD links temporal	Unauffällige Histopathologie
65	45	w	FCD links temporal	Reaktive Veränderungen
66	30	m	FCD links frontal	Reaktive Veränderungen
67	25	w	FCD links > rechts central	Fehlt
68	22	m	FCD links central	Fehlt
69	29	m	keine FCD	FCD Typ IIIa + HS
70	56	m	keine FCD, rechts HS	FCD Typ IIIa + HS ILAE Typ 1
71	30	m	keine FCD	FCD Typ IIIa + HS ILAE Typ 1
72	56	m	keine FCD, links HS	FCD Typ IIIa + HS ILAE Typ 2
73	36	w	keine FCD, links HS	FCD Typ IIIa + HS ILAE Typ 2
74	34	w	keine FCD, rechts HS	FCD Typ IIIa + HS ILAE Typ 1
75	38	m	keine FCD, links HS	FCD Typ IIIa + HS ILAE Typ 1
76	54	m	keine FCD	FCD Typ IIIa + HS ILAE Typ 1
77	50	w	keine FCD	FCD Typ IIIa + HS ILAE Typ 1
78	52	m	keine FCD, links HS	FCD Typ IIIa/b + HS ILAE Typ 1
79	6	w	keine FCD	FCD Typ IIb
80	55	w	keine FCD, links HS	MOGHE + HS

Abkürzungen: FCD – Fokale kortikale Dysplasie; HS – Hippocampusklerose, m – männlich; MCD – Malformation of Cortical Development; MOGHE - Mild malformation of cortical development with oligodendroglial hyperplasia in epilepsy; w – weiblich

Zwischen Januar 2014 und April 2023 wurde bei 68 Patienten (Fälle 1 bis 68) am Epilepsie-Zentrum Hamburg präoperativ mittels MAP18 auf cerebralen 3T-MRTs der Verdacht auf eine MCD gestellt (20 weiblich; Durchschnittsalter $\pm$ Standardabweichung: $31 \pm 14,3$ Jahre; Altersspanne: 5 – 62 Jahre).

Die histopathologische Untersuchung bestätigte bei 48 Patienten (Fälle 1 bis 48) die Diagnose einer MCD. Diese Gruppe umfasste fünf Patienten mit einer FCD Typ I, 28 Patienten mit FCD Typ II, zwölf Patienten mit FCD Typ III, einen Patienten mit einer MOGHE, einen Patienten mit einer „komplexen Fehlbildung“ sowie einen weiteren Patienten mit „verstreuten dysmorphen Neuronen ohne definitive Klassifikation“.

Bei zwei Patienten (Fälle 67 und 68) lag keine histopathologische Untersuchung vor (ein Fall mit Kallosotomie, ein Fall mit fehlender Untersuchung).

In 18 Fällen (Fälle 49 bis 66) konnte der präoperative Verdacht auf eine MCD nicht bestätigt werden. Ein Patient (Fall 49) wurde stattdessen mit einem glioneuronalen Tumor diagnostiziert. Bei zehn Patienten mit einem präoperativen Verdacht auf FCD Typ IIIa (Fälle 50 bis 59) wurde lediglich eine Hippocampusklerose festgestellt.

Trotz unauffälliger Histopathologie wiesen vier Patienten postoperativ eine verbesserte Anfallsfrequenz auf: zwei Patienten wurden nach Engel IA klassifiziert (Fälle 63 und 64), ein Patient nach Engel IIA (Fall 62) und ein weiterer Patient nach Engel IIIA (Fall 61). Lediglich ein Patient (Fall 60) mit unauffälliger Histopathologie zeigte keine Anfallsreduktion (Engel IV).

Bei zwei weiteren Patienten (Fälle 65 und 66) wurden reaktive Veränderungen festgestellt, das postoperative Outcome dieser Fälle ist jedoch nicht bekannt.

Darüber hinaus wurden bei zwölf Patienten, bei denen präoperativ kein bildmorphologischer Verdacht auf eine MCD bestand, histopathologisch MCDs nachgewiesen (Fälle 69 bis 80). Diese Gruppe umfasste zehn Patienten mit FCD Typ III, einen Patienten mit FCD Typ II und einen Patienten mit MOGHE.



Contents lists available at ScienceDirect

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Development and prospective clinical validation of a convolutional neural network for automated detection and segmentation of focal cortical dysplasias

Vicky Chanra^a, Agata Chudzinska^b, Natalia Braniewska^b, Bartosz Silski^b, Brigitte Holst^c, Thomas Sauvigny^d, Stefan Stodieck^a, Sirko Pelzl^b, Patrick M. House^{a,b,e,*}

^a Hamburg Epilepsy Center, Protestant Hospital Alsterdorf, Department of Neurology and Epileptology, Hamburg, Germany

^b theBlue.ai GmbH, Hamburg, Germany

^c University Hospital Hamburg-Eppendorf, Department of Neuroradiology, Hamburg, Germany

^d University Hospital Hamburg-Eppendorf, Department of Neurosurgery, Hamburg, Germany

^e Epileptologicum Hamburg, Specialist's Practice for Epileptology, Hamburg, Germany

ARTICLE INFO

Keywords:

FCD
MRI
Convolutional neural network
Epilepsy
Artificial intelligence
Continual learning

ABSTRACT

Purpose: Focal cortical dysplasias (FCDs) are a leading cause of drug-resistant epilepsy. Early detection and resection of FCDs have favorable prognostic implications for postoperative seizure freedom. Despite advancements in imaging methods, FCD detection remains challenging. House et al. (2021) introduced a convolutional neural network (CNN) for automated FCD detection and segmentation, achieving a sensitivity of 77.8%. However, its clinical applicability was limited due to a low specificity of 5.5%. The objective of this study was to improve the CNN's performance through data-driven training and algorithm optimization, followed by a prospective validation on daily-routine MRIs.

Material and methods: A dataset of 300 3 T MRIs from daily clinical practice, including 3D T1 and FLAIR sequences, was prospectively compiled. The MRIs were visually evaluated by two neuroradiologists and underwent morphometric assessment by two epileptologists. The dataset included 30 FCD cases (11 female, mean age: 28.1 ± 10.1 years) and a control group of 150 normal cases (97 female, mean age: 32.8 ± 14.9 years), along with 120 non-FCD pathological cases (64 female, mean age: 38.4 ± 18.4 years). The dataset was divided into three subsets, each analyzed by the CNN. Subsequently, the CNN underwent a two-phase-training process, incorporating subset MRIs and expert-labeled FCD maps. This training employed both classical and continual learning techniques. The CNN's performance was validated by comparing the baseline model with the trained models at two training levels.

Results: In prospective validation, the best model trained using continual learning achieved a sensitivity of 90.0%, specificity of 70.0%, and accuracy of 72.0%, with an average of 0.41 false positive clusters detected per MRI. For FCD segmentation, an average Dice coefficient of 0.56 was attained. The model's performance improved in each training phase while maintaining a high level of sensitivity. Continual learning outperformed classical learning in this regard.

Conclusions: Our study presents a promising CNN for FCD detection and segmentation, exhibiting both high sensitivity and specificity. Furthermore, the model demonstrates continuous improvement with the inclusion of more clinical MRI data. We consider our CNN a valuable tool for automated, examiner-independent FCD

Abbreviations: AUC, Area under the curve; CLOND, "Continual learning with old model and new dataset", a CNN model trained by a continual learning strategy; CLP, Case-level precision; CNN, Convolutional neural network; Dice, Dice coefficient; ESI, Electric source imaging; EWC, Elastic weight consolidation; FCD, Focal cortical dysplasia; FDG-PET, Fluorodeoxyglucose positron emission tomography; FLAWS, Fluid and white matter suppression; FN, False negative; FP, False positive; GTM, Ground truth mask; LLP, Lesion-level precision; M, Mean; MRI, Magnetic resonance imaging; MSI, Magnetic source imaging; OWD, "Old model with whole dataset", a CNN model trained by classical learning; PMG, Focal polymicrogyria; ROC, Receiver operating characteristic; SD, Standard deviation; SNR, Signal-to-noise ratio; SPECT, Subtraction ictal single photon emission computed tomography; TE, Echo Time; TN, True negative; TP, True positive; TR, Repetition Time; T1, Longitudinal Relaxation Time; VAE, Variational autoencoder; VLP, Voxel-level precision; VLS, Voxel-level sensitivity; 3D-EDGE, 3D Edge-Enhancing Gradient Echo.

* Corresponding author at: Hamburg Epilepsy Center, Protestant Hospital Alsterdorf, Department of Neurology and Epileptology, Hamburg, Germany

E-mail address: house@epileptologicum.de (P.M. House).

<https://doi.org/10.1016/j.epilepsyres.2024.107357>

Received 13 November 2023; Received in revised form 28 February 2024; Accepted 1 April 2024

Available online 3 April 2024

0920-1211/© 2024 Elsevier B.V. All rights reserved.

detection in daily clinical practice, potentially addressing the underutilization of epilepsy surgery in drug-resistant focal epilepsy and thereby improving patient outcomes.

1. Introduction

Focal cortical dysplasias (FCDs) are focal malformations of cortical development first described by Crome (1957) and named by Taylor et al. (1971). They result from abnormal neuronal proliferation, migration, and differentiation and are characterized by architectural neocortex disorganization (FCD type I), dysmorphic neurons (FCD type II), and may co-occur with another epileptogenic lesion (FCD type III) (Najm et al., 2022).

FCDs account for 5–25% of focal epilepsies (Bast et al., 2006), being the primary cause of focal drug-resistant epilepsies in children and ranking as the third most common cause in adult patients (Blumcke et al., 2017). While epilepsy typically begins in childhood (Wiwattana-dittakul et al., 2021), it can also manifest as late as the second to the sixth decade of life (Fauser et al., 2006). Patients generally do not exhibit significant neurological deficits apart from seizures (Guerrini and Barba, 2021). Early and complete resective surgery offers the best chance for long-term seizure freedom (Willard et al., 2022), making FCDs the main indication for epilepsy surgery. Shorter epilepsy duration and younger age at surgery favor better postoperative outcomes (Salemdawod et al., 2022; Lamberink et al., 2020), whereas incomplete resection is associated with a worse prognosis (Willard et al., 2022; Seong et al., 2021). Hyperactivation of the mTOR pathway plays an important role in the pathogenesis of FCD type II (Wang et al., 2023; Blumcke et al., 2021), suggesting that targeting this pathway with mTOR inhibitors might be a future treatment option for drug-resistant epilepsy caused by FCD (Moloney et al., 2021). Hence, early detection of FCDs and accurate assessment of their extent on cerebral imaging are pivotal for optimal treatment outcomes.

Structural magnetic resonance imaging (MRI) serves as the primary modality for assessing suspected FCD cases. FCD type II typically presents with characteristic MRI findings, including a blurred white-gray matter junction, localized cortical thickening, abnormal cortical gyration, and a focal increase in subcortical white matter T2 signal, often tapering towards the underlying ventricle (referred to as the “transmantle sign” in FCD type IIb) (Urbach et al., 2022b, 2002; Barkovich and Kuzniecky, 1996). In contrast, FCD type I is usually not visible on an MRI but may present with moderately increased T2/FLAIR-signal, decreased white matter T1 signal, blurring of the white-gray matter junction, and hypoplasia of the affected region (Marcelis et al., 2022).

Despite improvements in clinical imaging techniques, FCDs remain the most common pathologic finding in MRI-negative cases of drug-resistant focal epilepsy, with missed lesion rates ranging from 30% to 78% (Urbach et al., 2022b). The main challenge lies in the subtlety of MRI abnormalities and the variability of FCDs in terms of location, extent, and shape, usually merging into adjacent tissues or located at the bottom of a sulcus (Ying et al., 2019).

Advanced high-resolution MRI techniques, such as 7 T MRI (Tang et al., 2022; Chen et al., 2021), and improved imaging protocols, e.g., 3D Edge-Enhancing Gradient Echo (3D-EDGE) (Middlebrooks et al., 2020), fluid and white matter suppression (FLAWS) (Chen et al., 2018), MP2RAGE (Demerath et al., 2020) and quantitative MRI analyses (Maiworm et al., 2020), have enhanced the identification of epileptogenic lesions. Complementary multimodality epilepsy imaging techniques, including fluorodeoxyglucose positron emission tomography (FDG-PET) (Yokota et al., 2020), subtraction ictal single photon emission computed tomography (SPECT) (Foiadelli et al., 2020), magnetic source imaging (MSI) (Otsuka et al., 2021; Mohamed et al., 2020) and electric source imaging (ESI) (Wanders et al., 2022), provide additional insights in lesion localization. However, subtle lesions may remain undetected, and the widespread clinical application of these techniques is

constrained by limited availability and the need for specialized expertise (Duncan et al., 2016).

Semi-automated postprocessing methods, including surface-based (SBM) (Dale et al., 1999; Fischl et al., 1999) and voxel-based morphometry (VBM) (Ashburner and Friston, 2000), have significantly contributed to the improvement of FCD detection. The “Morphometric Analysis Program” (v2018; MAP18) (Huppertz et al., 2008, 2005) forms an integral part of presurgical examinations in epilepsy centers worldwide (David et al., 2021) and has independently been validated for its clinical benefit (Wang et al., 2015). Despite its significant diagnostic benefit, this method has several drawbacks. These include the intricate implementation requiring a healthy control group database, concerns regarding its low specificity (David et al., 2021), reliance on time-consuming human evaluation, and the necessity of specialized expertise in neuroradiology. To address these limitations and enable examiner-independent evaluation, several machine learning techniques based on postprocessing methods have been developed in recent years. Notable contributions in this field include the works of Demerath et al. (2022), Ganji et al. (2022), Spitzer et al. (2022), Urbach et al. (2022a), David et al. (2021), Gill et al. (2021), Feng et al. (2020), Wang et al. (2020), Bijay Dev et al. (2019), Mo et al. (2019), Jin et al. (2018), Adler et al. (2017), Ahmed et al. (2015), and Hong et al. (2014).

House et al. (2021) introduced a convolutional neural network (CNN) for automated detection and segmentation of FCDs, trained on an 3 T MRI dataset comprising 201 FCD cases. In a prospective validation on an independent dataset, the CNN exhibited a sensitivity of 77.8%, outperforming visual analysis by experienced neuroradiologists. However, the CNN exhibited a low specificity of 5.5%, requiring additional visual confirmation. Therefore, its clinical application has been limited to a pre-screening tool.

Our study aimed to improve this CNN’s performance by employing data-driven training and optimizing the algorithm. Subsequently, we prospectively validated it using daily routine MRIs to simulate everyday use.

2. Materials & methods

2.1. Data collection

All patients of the Hamburg Epilepsy Center, Germany, who underwent high-resolution cranial 3 T MRIs after March 2022 on a Siemens Skyra scanner at the University Hospital Hamburg-Eppendorf, Germany, were included in the study.

Neurological examinations, seizure history assessments, and video-EEG recordings were performed in all participants.

The MRIs were visually analyzed by two neuroradiologists, including a resident and an experienced senior neuroradiologist, at the University Hospital Hamburg-Eppendorf. Morphometric analysis was conducted in a consensus reading by two epileptologists at the Hamburg Epilepsy Center, including a resident with nine years of experience in clinical neuroimaging and more than one year of full-time training in morphometrical analysis, alongside an experienced senior epileptologist with over 20 years of expertise. Final evaluations considered seizure semiology, clinical and video-EEG data, and visual double checks by experienced epileptologists. These evaluations served as a reference for comparative assessments. Following the positive detection of an FCD, the findings were reviewed and discussed at interdisciplinary case meetings involving epileptologists from the Hamburg Epilepsy Center, as well as neurosurgeons and neuroradiologists from the University Hospital Hamburg-Eppendorf.

Data collection was continued until 30 FCD cases were identified.

Focal polymicrogyria (PMG) cases were intentionally included as they can resemble FCD and are equally important for presurgical assessment in patients with drug-resistant focal epilepsy (FCD/PMG hereinafter referred to as “FCD”). Although the feature maps of the morphometric analysis were not initially developed for the purpose of PMG detection, they have been effective in identifying PMG (House et al., 2021). The control group was composed of 270 randomly selected non-FCD MRIs, including 150 normal MRIs and 120 non-FCD-related pathological conditions (hereinafter referred to as “non-FCD pathological MRIs”).

The identified FCDs had to be clearly recognizable on artifact-free MRIs with sufficiently high signal-to-noise ratio (SNR). However, no criteria for data exclusion were established for the control group MRIs, intentionally preserving MRIs with suboptimal image quality to enable the algorithm’s suitability for practical clinical use.

All patients’ MRIs were fully anonymized. Our study was performed in accordance with the Ethics Committee of the regional Medical Association Hamburg (WF-062/18) and the ethical standards laid down in the Declaration of Helsinki, as well as local and institutional laws.

2.2. MRI processing pipeline

A 3 T Siemens Skyra scanner (Magnetom Syngo MR D11) at the University Hospital Hamburg-Eppendorf was used to acquire high-resolution 3D datasets (T1, T2, FLAIR). Table 1 provides details of the acquisition protocol for 3 T MRI measurements.

The 3D dataset was exported in DICOM format and anonymized using DICOM Anonymizer Pro (NeoLogica s.r.l.).

FLAIR sequences were resampled to match T1 image resolution, ensuring voxel compatibility and preventing upscaling artifacts.

Morphometric MRI analysis applied MATLAB (MathWorks) scripts MAP18 on 3D T1 and T2 sequences (House et al., 2013) with coregistered 3D FLAIR sequences to calculate junction, extension, and thickness images according to the method of Huppertz et al. (2008), (2005).

As described by House et al. (2021), preprocessing involved brain tissue segmentation, morphometric MRI analysis, and final input preparation. T2 sequences were excluded from the training dataset.

For tissue segmentation, we utilized a 2D Unet with a ResNet34 encoder (Howard, 2018) to segment T1 MRIs into three primary brain tissues: cerebrospinal fluid, gray matter and white matter. This segmentation aimed to create a detailed brain mask for skull stripping and to enable tissue segmentation for subsequent preprocessing.

A volumetric FCD lesion map was created through manual voxelwise labeling on T1 and FLAIR scans using ITK-SNAP (version 3.8.0) by an experienced epileptologist and validated by a senior epileptologist.

2.3. Participant demographics and control group characterization

In a 10-month period between March and December 2022, a total of 893 high-resolution cranial 3 T MRIs were conducted. Among these, 30 FCD cases (3.4%), including four focal PMG, were identified using our reference method (cf. Section 2.1) according to typical MRI criteria (11 female, mean age at MRI: 28.1 years $\pm$ 10.1 years, age range: 17 years to 54 years) on MRIs with good quality without artifacts and with

sufficiently high SNR. No additional or questionable FCDs were detected on MRIs of suboptimal quality.

The control group consisted of 270 consecutive clinical-routine MRIs (270 patients, 161 female, mean age at MRI: 35.3 years $\pm$ 16.7 years, age range: 14 years to 82 years), including 150 normal MRIs (150 patients, 97 female, mean age at MRI: 32.8 years $\pm$ 14.9 years, age range: 15 years to 82 years) and 120 non-FCD pathological MRIs (120 patients, 64 female, mean age at MRI: 38.4 years $\pm$ 18.4 years, range: 14 years to 80 years) with the following pathologies: hippocampal sclerosis (7.5%), DNET/ganglioglioma/gangliocytoma (7.5%), glioblastoma/astrocytoma (5.0%), hamartoma (2.5%), cavernoma (7.5%), arteriovenous malformation (2.5%), developmental venous anomaly (5.0%), infarction (7.5%), hemorrhage (2.5%), arachnoid cyst (7.5%), porencephaly (2.5%), medium microangiopathy (2.5%), strong microangiopathy (2.5%), multiple sclerosis lesions (2.5%), encephalitis (15.0%), leukodystrophy (5.0%), post-traumatic defects (7.5%), and postoperative defects (7.5%). The distribution of these pathologies was identical in the three subsets.

Table 2 summarizes the demographic characteristics of the dataset.

Table 2
Demographic characteristics of the dataset.

	Total	FCD MRIs (each n = 10)	non-FCD pathological MRIs (each n = 40)	normal MRIs (each n = 50)
Subset 1 (n = 100 MRIs)	62.0% female, range: 15 years to 82 years, M + SD: 34.6 $\pm$ 17.9 years	40.0% female, range: 18 years to 43 years, M + SD: 25.3 $\pm$ 7.4 years	57.5% female, range: 17 years to 80 years, M + SD: 38.9 $\pm$ 19.5 years	70.0% female, range: 15 years to 82 years, M + SD: 33.0 $\pm$ 17.0 years
Subset 2 (n = 100 MRIs)	59.0% female, range: 14 years to 78 years, M + SD: 38.1 $\pm$ 16.0 years	50.0% female, range: 18 years to 54 years, M + SD: 30.1 $\pm$ 11.9 years	55.0% female, range: 14 years to 78 years, M + SD: 41.2 $\pm$ 17.8 years,	64.0% female, range: 17 years to 73 years, M + SD: 37.2 $\pm$ 14.5 years
Subset 3 (n = 100 MRIs)	51.0% female, range: 14 years to 78 years, M + SD: 31.2 $\pm$ 14.2 years	20.0% female, range: 18 years to 48 years, M + SD: 29.0 $\pm$ 9.6 years	47.5% female, range: 14 years to 78 years, M + SD: 35.3 $\pm$ 17.2 years	60.0% female, range: 17 years to 68 years, M + SD: 28.4 $\pm$ 11.2 years
Total		36.7% female, range: 17 years to 54 years, M + SD: 28.1 $\pm$ 10.1 years	53.3% female, range: 14 years to 80 years, M + SD: 38.4 $\pm$ 18.4 years	64.7% female, range: 15 years to 82 years, M + SD: 32.8 $\pm$ 14.9 years

Abbreviations: M – Mean; SD – Standard deviation.

Table 1
MRI scanner, sequences, and technical parameters.

Scanner, model, field strength	Signal weighting, name of sequence	Dimensionality, orientation, slice thickness (mm)	TR (ms)	TE (ms)	T1 (ms)	DICOM file resolution (pixels)
Siemens Skyra 3 T (Magnetom Syngo MR D11)	3D T1w MPRAGE	3D, sagittal, 0.9 $\times$ 0.9 $\times$ 0.9	1900	2.5	900	256 $\times$ 256 $\times$ 192
	T2w SPACE	3D, sagittal, 0.9 $\times$ 0.9 $\times$ 0.9	3200	409		256 $\times$ 256 $\times$ 192
	3D T2w FLAIR	3D, sagittal, 0.9 $\times$ 0.9 $\times$ 0.9	4700	392	1800	320 $\times$ 320 $\times$ 192
	PDw/T2w	2D, axial, 3	2800	18/90		384 $\times$ 512
	T1w IR	2D, coronal, 2	2500	9.5	500	256 $\times$ 256
	T2w TSE	2D, coronal, 2	8900	83		320 $\times$ 320

Abbreviations: TE – Echo Time; TR – Repetition Time; T1 – Longitudinal Relaxation Time.

2.4. FCD lesion characterization

The 30 FCDs were distributed anatomically as follows: right hemisphere 70.0%, left hemisphere 30.0%, frontal 33.4%, insula 16.7%, temporal 3.3%, parietal 40.0%, and occipital 6.7%. Seven FCDs (23.3%) spread out into one or two adjacent regions. The average lesional volume was $13.8 \text{ ml} \pm 18.7 \text{ ml}$, ranging from 0.19 ml to 94.25 ml. The FCD size was categorized as follows: 20.0% small ($< 1.5 \text{ ml}$), 33.3% medium (1.5 ml – 7.5 ml), 46.7% big ($> 7.5 \text{ ml}$). The neuroradiologists detected 73.3% of the FCDs.

Additional anatomical FCD characteristics are presented in **Table A.1** and **Fig. A.1**.

2.5. Artificial neural network

As described in [House et al. \(2021\)](#), the neural network employs an encoder-decoder architecture variation called U-Net ([Ronneberger et al., 2015](#)), with an auxiliary variational decoder branch for regularization ([Myronenko, 2018](#)).

The CNN was implemented using PyTorch ([Paszke et al., 2019](#)). The final architecture of the CNN is presented in **Fig. 1**.

2.5.1. Network input

The network input comprises four channels: T1 with N4 Bias Field Correction, resampled FLAIR, and two auxiliary channels consisting of ‘junction images’ and ‘extension images’ based on [Huppertz et al. \(2008, 2005\)](#).

2.5.2. ResBlock

The network’s key component is ResBlock, consisting of two sets of group normalization, 3D convolution, and ReLU activation function, followed by a skip connection.

2.5.3. Encoder

The encoder starts with an initial ResBlock and is followed by the DropBlock regularization method ([Ghiasi et al., 2018](#)). The encoder comprises three subsequent levels, each consisting of two ResBlocks, a Positional Normalization ([Li et al., 2019](#)), and a strided convolution used for downsampling. It concludes with two ResBlocks serving as the bottleneck.

2.5.4. Decoder

The decoder complements the encoder with a single ResBlock for each level. It employs trilinear interpolation for upsampling and

integrates the output with skip connections from the encoder. Moment Shortcut (MS) ([Li et al., 2019](#)) is used for further normalization after a skip connection is added to the upsampled output. The decoder generates a probability segmentation map from the sigmoid activation function.

2.5.5. Variational decoder

The variational decoder serves as a regularization method without affecting output segmentation. It includes three convolutions and two upsampling layers. The bottleneck employs two linear layers to represent means and logarithm of variances. The output is calculated using the classical VAE reparameterization method.

2.5.6. Training of the baseline model

The neural network had been trained on a 3 T MRI dataset including 201 FCD cases, 100 normal cases, and 50 non-FCD pathological cases (referred to as the “baseline dataset”). Evaluation was performed through cross-validation in four data configurations. The best model, referred to as the “baseline model”, was selected for further development.

In [House et al. \(2021\)](#), the baseline model’s performance was prospectively validated on an independent dataset of 100 MRIs, including nine FCD MRIs, from patients with focal epilepsy (46% female, mean age at MRI: 44 years, range: 12 years – 74 years). The validation demonstrated a sensitivity of 77.8% and specificity of 5.5%.

For a detailed description of the CNN’s architecture and the baseline model’s training and validation, please refer to [House et al. \(2021\)](#).

2.6. Training and evaluation of the models

Three datasets (“Dataset 1–3”) were prospectively acquired, each comprising 100 MRIs with 10 FCD cases, 50 normal cases, and 40 non-FCD pathological cases.

In the first phase, the baseline model analyzed the first dataset (“Dataset 1”). Following this, two different training strategies were applied. In the first approach, a classical training method was employed by training the baseline model with the entire dataset (“Baseline dataset” and “Dataset 1”), with the resulting model designated the “Old model with whole dataset” (OWD). The second approach employed continual learning by training the baseline model exclusively on new data (“Dataset 1”), with the resulting model designated the “Continual learning with old model and new dataset” (CLOND).

In the second phase, all three models (baseline model, OWD and CLOND) analyzed the second dataset (“Dataset 2”). The baseline model

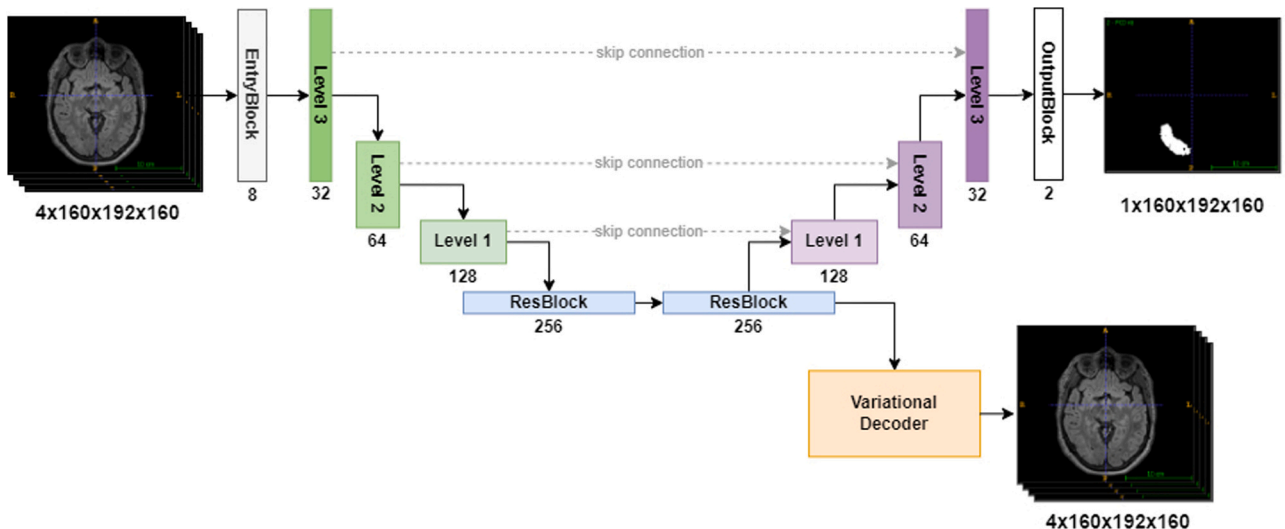


Fig. 1. Neuronal network architecture.

was then trained in a classical manner using the entire dataset ("Baseline dataset", "Dataset 1" and "Dataset 2"), resulting in a modified OWD model. CLOND was further trained using exclusively the second dataset ("Dataset 2") in a continual learning approach, resulting in a modified CLOND model.

In the third phase, all three models (baseline model, modified OWD and modified CLOND) analyzed the third dataset ("Dataset 3").

Fig. 2 illustrates the training and evaluation process.

2.6.1. Training strategies

The first strategy led to the development of OWD using a classical training approach. The training loop was not modified, and the loss function remained the same as in House et al. (2021):

$$L = L_{dice} + 0.1MSE + 0.1L_{KL}$$

Here, MSE represents the mean squared error for the Variational Decoder, and L_{KL} denotes the Kullback-Leibler divergence.

The baseline model served as the starting point for OWD, and the entire available dataset was utilized for training. For a detailed explanation of the loss function, please refer to House et al. (2021).

In contrast, in CLOND the training loop was modified according to the principles and strategies of continual learning, and the Elastic Weight Consolidation (EWC) (Aich, 2021; Kirkpatrick et al., 2017) was incorporated into the loss function as a penalty:

$$L_{CL} = L_{dice} + 0.1MSE + 0.1L_{KL} + \lambda \cdot EWC$$

Here, λ sets the importance of the old task relative to the new one, and EWC serves as penalty employed in continual learning.

Similar to OWD, CLOND started with the baseline model but exclusively utilized the most recent dataset for training.

In each training phase, both models were trained for 100 epochs using the Adam optimizer with weight decay set to 10^{-5} and the learning rate starting at 10^{-4} , gradually decreasing according to the formula:

$$\alpha = \alpha_0 \cdot \left(1 - \frac{e}{N_e}\right)^{0.9}$$

Here, e represents the current epoch, N_e is the total number of epochs (in this case 100), and a batch size of 2 was employed.

2.6.2. Continual learning

Continual learning (Kirkpatrick et al., 2017) addresses the challenge of "catastrophic forgetting" (McCloskey and Cohen, 1989) in machine learning, enabling models to adapt to new tasks without forgetting previously acquired knowledge. In CLOND, we incorporated Domain-Incremental Learning (van de Ven et al., 2022) and EWC as fundamental techniques.

EWC operates as a regularization technique, optimizing model parameters for multiple tasks by regulating the model's adaptation to new tasks during training. This is achieved by introducing a quadratic penalty to the loss function. This penalty prioritizes the preservation of crucial model weights learned during the initial task, preventing significant changes to these essential weights, while allowing other parameters to adapt to the new task.

The quadratic penalty in EWC is calculated using the Fisher Information Matrix, which quantifies the sensitivity of the loss with respect to each model parameter. This information determines the weight's contribution to the performance on the previously learned task. The penalty term is then added to the loss function, with a scaling factor (λ) controlling its influence. The scaling factor (λ) regulates the balance between preserving the knowledge of the initial task and adapting to the new task.

In our model, EWC ensures the retention of knowledge gained from FCD segmentation while the model undergoes additional training for the discrimination task against non-FCD images.

Domain-Incremental Learning enables the algorithm to learn diverse tasks without requiring explicit identification. In our context, the algorithm views a new dataset as a distinct domain, facilitating efficient learning of FCD segmentation and discrimination of non-FCD cases.

This adaptability, achieved by EWC and Domain-Incremental learning, is essential for reducing false positive results in the discrimination task against non-FCD images. Moreover, the model demonstrates enhanced adaptability to novel tasks and domains, especially in scenarios where data is incrementally received, eliminating the need for simultaneous access to the complete dataset.

For a detailed description of these methods, please refer to Kirkpatrick et al. (2017) and van de Ven et al. (2022).

2.7. Image evaluation and performance metrics

An FCD was diagnosed based on identifiable neuroradiological features in either T1 or FLAIR images, i.e., abnormal gyration, blurring of the gray-white matter junction, abnormal cortical thickness, and hyperintensity in FLAIR-weighted images (Urbach et al., 2022b; Bar-kovich and Kuzniecky, 1996).

Expert-labeled FCD maps served as the ground truth mask (GTM) for evaluation.

FCD detection performance was evaluated at the case level using the following metrics:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Balanced accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2}$$

$$\text{Case-level precision (CLP)} = \frac{TP}{TP + FP}$$

FCD MRIs were labeled as true positive (TP) if the Dice coefficient was ≥ 0.1 . Conversely, FCD MRIs were classified as false negative (FN) if no clusters were registered or if clusters overlapped with the GTM with a Dice coefficient < 0.1 . Control group MRIs were labeled as true negative (TN) when no clusters were detected, and as false positive (FP) when clusters were found.

Voxel and lesion levels evaluations were performed using the following metrics:

$$\text{Voxel-level sensitivity (VLS)} = \frac{TP_{\text{voxel}}}{\text{ground truth mask}}$$

$$\text{Voxel-level precision (VLP)} = \frac{TP_{\text{voxel}}}{TP_{\text{voxel}} + FP_{\text{voxel}}}$$

$$\text{Lesion-level precision (LLP)} = \frac{TP_{\text{cluster}}}{TP_{\text{cluster}} + FP_{\text{cluster}}}$$

Voxels and clusters overlapping with the GTM in FCD MRIs were labeled as TP, while voxels and clusters in the control group and non-overlapping voxels and clusters in FCD MRIs were labeled as FP.

The Dice coefficient was used to evaluate FCD segmentation performance by measuring spatial overlap with the GTM at the voxel level:

$$\text{Dice} = \frac{2 * |A \cap B|}{|A| + |B|}$$

Here, $|A|$ represents the volume of the predicted FCD and $|B|$ represents the volume of the expert-labeled FCD maps.

Receiver operating characteristic (ROC) curve analysis was

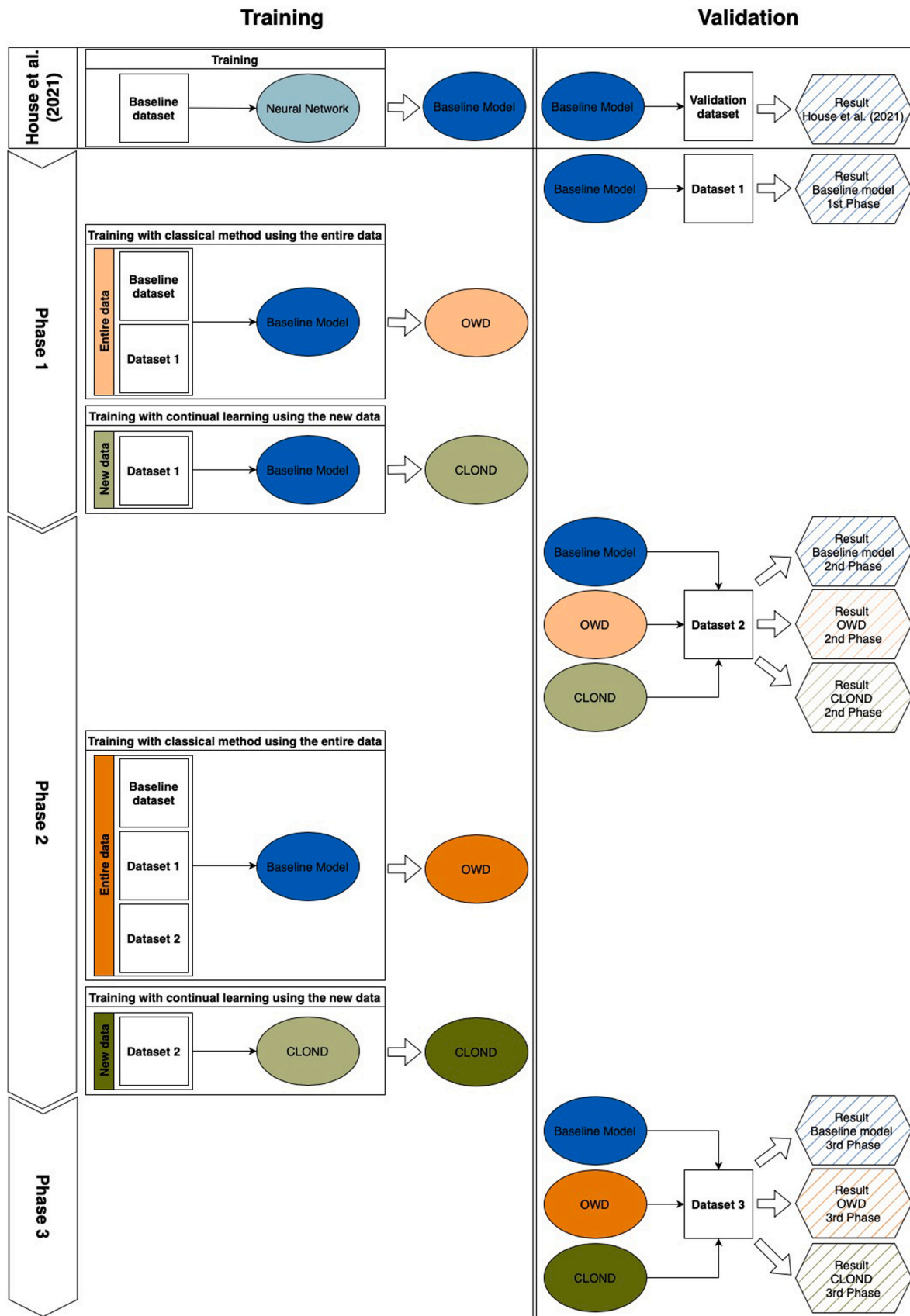


Fig. 2. Training and validation process. The datasets were composed as follows: -“Baseline dataset”: 201 FCD MRIs, 100 normal MRIs, 50 non-FCD pathological MRIs. - “Validation dataset”: 9 FCD MRIs, 91 control group MRIs. - “Dataset 1”, “Dataset 2” and “Dataset 3”: Each includes 10 FCD MRIs, 50 normal MRIs, and 40 non-FCD pathological MRIs. * Results by [House et al. \(2021\)](#): Sensitivity 77.8%, Specificity 5.5%.

conducted and the area under the curve (AUC) was calculated for all models considering TP and FP clusters.

A comparison between the models was performed using an intersection-union test. A model was considered superior if it demonstrated superior specificity and non-inferior sensitivity within a $\leq 10\%$ limit.

Statistical analysis was conducted using Microsoft Excel (Version 16.69.1) and R (version 4.2.3., R Foundation for Statistical Computing, Vienna, Austria). ROC curves were generated using XLSTAT statistical and data analysis solutions (version 2023.1.3., Lumivero, New York, USA).

3. Results

The results of the prospective validation are summarized in Table 3 and Fig. 3. Table A.2 presents additional results of the subgroup analysis.

3.1. Detection performance

After the first round of training, all three models demonstrated a sensitivity of 80.0% (95% CI: 55.2% – 100%) by correctly identifying 8 out of 10 FCDs. After the second round of training, the baseline model and CLOND achieved a sensitivity of 90.0% (95% CI: 71.4% – 100%) by detecting 9 out of 10 FCDs, while OWD achieved a sensitivity of 100% by detecting all 10 FCDs. Accordingly, no relevant differences in sensitivity were found among the three models.

The specificity progressively increased in both trained models across the training phases. OWD demonstrated higher specificity compared to the baseline model, and CLOND exhibited higher specificity compared to both OWD and the baseline model, with CLOND reaching a specificity of 70.0% (95% CI: 60.5% – 79.5%) in the last evaluation phase. Cochran's Q test indicated relevant differences in specificity between the models (Cochran's Q test: second phase $\chi^2(2) = 41.4$, $p < .001$; third phase: $\chi^2(2) = 82.7$, $p < .001$). Post hoc McNemar test confirmed relevant differences for all pairwise comparisons ($p < .05$).

OWD outperformed the baseline model in terms of accuracy, balanced accuracy, precision at voxel level, lesion level and case level. CLOND exceeded both models in these metrics. Additionally, the intersection union test confirmed CLOND's superiority over OWD and the baseline model, whereas OWD outperformed the baseline model (cf. Fig. A.2).

Table 3
Results of the prospective validation at voxel, lesion, and case level.

Model	Phase (each n=100)	TP voxel	FP voxel	TP cluster	FP cluster	TP cases	FN cases	TN cases	FP cases
Baseline	1	63,416	546,493	13	241	8	2	0	90
	2	40,446	452,841	8	219	8	2	2	88
	3	88,873	473,330	9	216	9	1	5	85
OWD	2	62,934	333,689	12	127	8	2	19	71
	3	110,435	148,135	10	63	10	0	51	39
CLOND	2	71,324	249,794	10	87	8	2	38	52
	3	120,765	118,976	10	41	9	1	63	27

Model	Phase (each n=100)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Balanced accuracy (%)	Dice	VLS (%)	VLP (%)	LLP (%)	CLP (%)	AUC	FP cluster per MRI
Baseline	1	80.0	0.0	8.0	40.0	0.45	22.4	10.4	5.1	8.2	0.87	2.41
	2	80.0	2.2	10.0	41.1	0.43	15.3	8.2	3.5	8.3	0.87	2.19
	3	90.0	5.6	14.0	47.8	0.50	33.3	15.8	4.0	9.6	0.97	2.16
OWD	2	80.0	21.1	27.0	50.6	0.50	23.7	15.9	8.6	10.1	0.93	1.27
	3	100	56.7	61.0	78.4	0.52	41.4	42.7	13.7	20.4	0.96	0.63
CLOND	2	80.0	42.2	46.0	61.1	0.55	26.9	22.2	10.3	13.3	0.94	0.87
	3	90.0	70.0	72.0	80.0	0.56	45.3	50.4	19.6	25.0	0.98	0.41

Abbreviations: AUC – Area under the curve; CLP – Case-level precision; Dice – Dice coefficient; FN – False negative; FP – False positive; LLP – Lesion-level precision; TN – True negative, TP – True positive; VLP – Voxel-level precision; VLS – Voxel-level sensitivity.

3.2. False positive clusters

The baseline model detected 676 FP clusters in 300 MRIs throughout the three phases, averaging 2.3 FP clusters per MRI (SD 1.2) and an average FP cluster volume of $1.1 \text{ ml} \pm 1.8 \text{ ml}$.

Both OWD and CLOND demonstrated a reduction of the number of FP clusters per phase, with CLOND showing fewer FP clusters compared to OWD. After the second round of training, CLOND achieved the lowest average of 0.4 FP clusters per MRI (SD 0.79) (cf. Fig. 4).

A Friedman test indicated relevant differences in the number of FP clusters among the three models in the second and third phase (second phase: $\chi^2(df = 2) = 65.0$, $p < .001$; third phase $\chi^2(df = 2) = 127.4$, $p < .001$). Multiple post-hoc Dunn-Bonferroni tests confirmed differences between the models for all paired comparisons in the second phase, and between the baseline model and OWD, as well as the baseline model and CLOND in the third phase ($p < .05$). In the third phase, CLOND demonstrated fewer FP clusters than OWD ($p = .48$), particularly in normal MRIs ($p = .02$).

All models had fewer average FP clusters in FCD MRIs compared to control group MRIs. However, in the baseline model, there were no relevant differences in the number of FP clusters within the control group between normal MRIs (2.3 FP cluster per MRI) and non-FCD pathological MRIs (2.5 FP clusters per MRI). In the final evaluation phase, both OWD and CLOND exhibited the highest proportion of FP clusters in non-FCD pathological MRIs (OWD 69.8%, CLOND 85.4%), whereas a lower proportion was observed in normal MRIs (OWD 30.2%, CLOND 14.6%), and none in FCD MRIs.

FP cluster sizes did not differ relevantly between the models (one-way ANOVA, $F(6, 987) = [0.47]$, $p = .83$) (cf. Fig. A.4), with more than 80% of clusters $< 1.5 \text{ ml}$ (cf. Fig. A.5).

The number of FP clusters are displayed in Fig. 4.

Fig. A.3 provides additional information on the anatomical distribution of the FP clusters.

3.3. Segmentation performance

The baseline model achieved an average Dice of 0.46 (95% CI: 0.26 – 0.65). The segmentation performance improved throughout the training phases, with CLOND achieving the highest average Dice of 0.56 in the final evaluation phase.

A Friedman test indicated relevant differences in segmentation performance among the models (second phase $\chi^2(df = 2) = 10.13$, $p < .01$; third phase $\chi^2(df = 2) = 6.51$, $p < .05$). Multiple post-hoc Dunn-Bonferroni tests confirmed differences between the baseline model and CLOND in the second phase ($p < .05$), while no relevant differences were

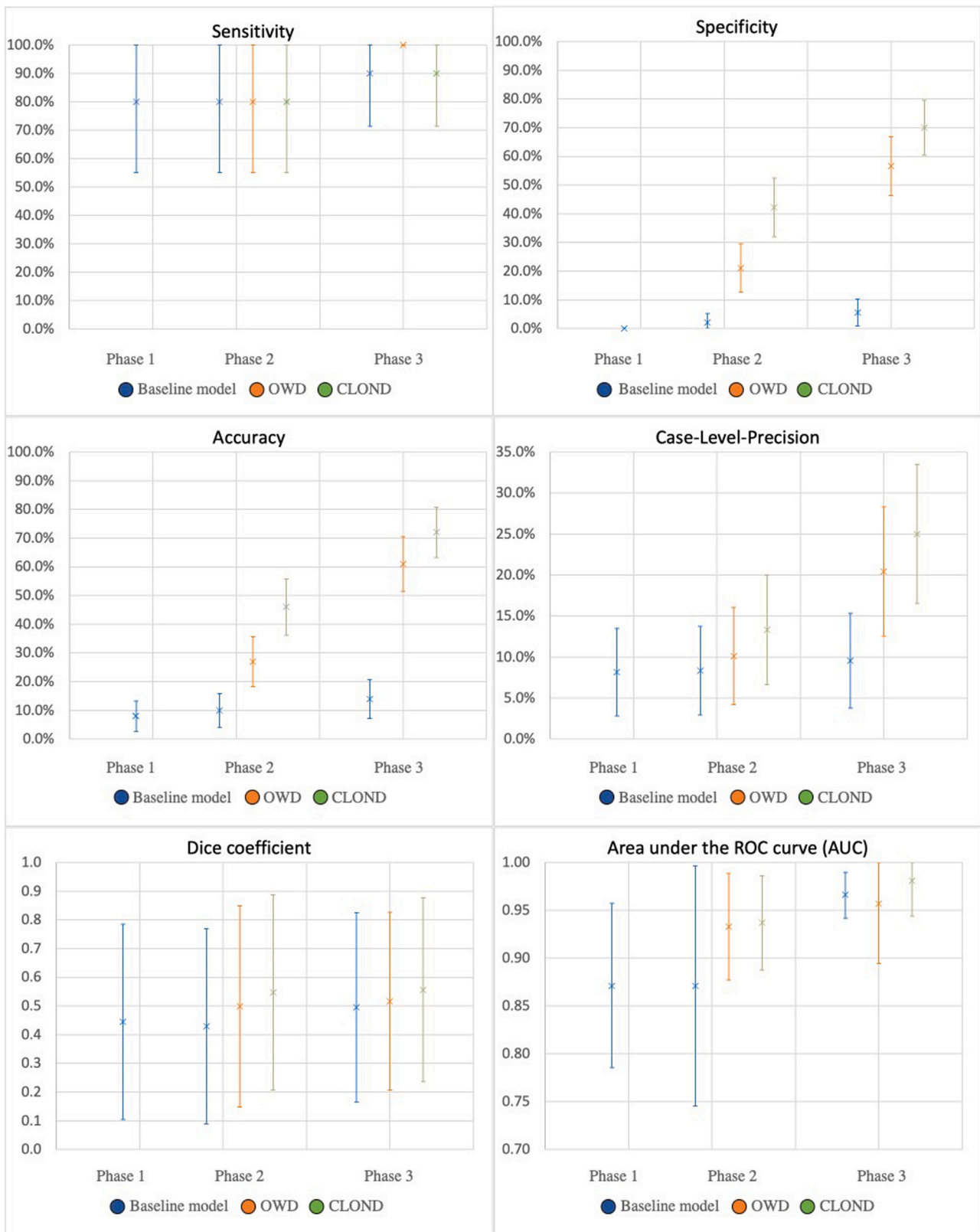


Fig. 3. Performance metrics of the CNN models at the case level with 95% Wald confidence intervals.

found in the other comparisons.

The segmentation performance of the three models is exemplified in Fig. 5, while Fig. 5b displays the segmentation of a small FCD.

The Dice coefficient for every FCD is presented in Fig. A.6.

3.4. Receiver operating characteristic analysis

The baseline model consistently performed well with AUC values ranging from 0.87 to 0.97. Both OWD and CLOND exhibited higher AUC than the baseline model, while CLOND exhibited the highest AUC of

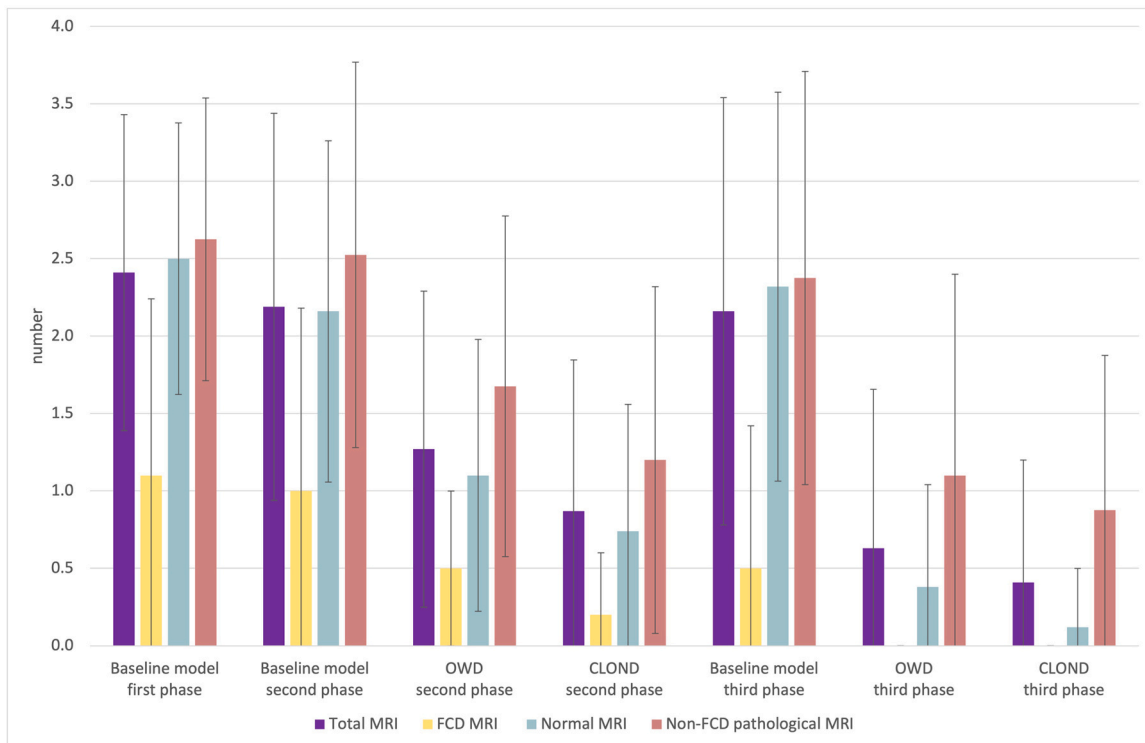


Fig. 4. Average number of false positive clusters per MRI with standard deviation.

0.98 in the final evaluation phase. However, there were no significant differences in AUC between the models.

The volumes of TP clusters were significantly larger than the volumes of FP clusters (t-test, $p < .01$). In the second phase, cut-off values of 1.4 ml for OWD and 1.8 ml for CLOND were determined to achieve the best possible differentiation between TP and FP clusters while maintaining sensitivity. Applying these cut-off values in the final phase would have resulted in an improvement of specificity to 82.5% (95% CI: 71.1% – 90.1%) with a decrease in sensitivity to 90.0% (95% CI: 57.1% – 100%) for OWD, and an improvement of specificity to 85.4% (95% CI: 71.1% – 93.4%) with the same sensitivity for CLOND.

Fig. A.7 displays the ROC curves for each model.

4. Discussion

In this study, we adopted the CNN model introduced by House et al. (2021) as our baseline model and trained it employing two distinct training strategies: a classical approach utilizing the entire MRI dataset (OWD), and a continual learning approach using exclusively new MRIs (CLOND). We evaluated the performance of FCD detection and segmentation for each training method. Our dataset comprised 300 anonymized 3 T MRIs from routine clinical practice, including 30 FCD cases and various non-FCD pathologies.

Both trained models demonstrated improvements compared to the baseline model while maintaining high sensitivity. Training significantly improved specificity, reaching up to 70.0% ($p < .05$), and reduced the occurrence of FP clusters to 0.4 per MRI ($p < .05$) in CLOND. Furthermore, the trained models demonstrated improved precision in segmenting FCD lesions with CLOND achieving an average Dice of 0.56.

Performance improvements were particularly evident in normal and FCD MRIs. Following training, FP clusters primarily occurred in non-FCD pathological MRIs. This observation can be attributed to the complexity and heterogeneity of these pathologies which may exhibit imaging features resembling FCDs, thus posing challenges for accurate differentiation. Nevertheless, the training process also reduced the occurrence of FP clusters both at case level and lesion level within this

cohort. Expanding the training dataset to encompass additional pathologies, each representing a minor proportion (2.5–7.5%) of the current training dataset, is expected to enhance the model's ability to distinguish FCDs from other pathologies.

Although the number of FP clusters per MRI decreased with training, their size remained consistent. Notably, over 80% of these clusters were small (< 1.5 ml), and more than 50% were very small (< 0.5 ml), unlikely to be mistaken for FCDs upon subsequent visual analysis.

In the ROC analysis, a cut-off value of 1.8 ml was determined for CLOND to differentiate TP from FP clusters. The application of this cut-off value increased specificity to 85.4% (95% CI 71.1–93.4%) while maintaining sensitivity. However, considering the presence of four FCDs in our dataset measuring 0.5 ml or less and the as-yet-unknown minimum size of FCDs, we decided not to limit our CNN to an arbitrarily chosen lesion-size threshold. In FCD detection, prioritizing sensitivity over specificity is crucial to prevent the potential risk of overlooking small FCDs.

As training progressed, FP clusters in infratentorial regions were eliminated, and the disproportionate occurrence of FP clusters in frontal and central regions declined (cf. Fig. A.3). These improvements were achieved without imposing any constraints on the CNN, indicating the potential for further enhancement through continued training.

CLOND outperformed OWD, demonstrating the superiority of continual learning in our setting. CLOND effectively leveraged prior knowledge to accelerate learning on new tasks while preventing catastrophic forgetting, leading to improved performance. Additionally, it allowed a robust understanding of FCD features, even when predominantly trained with non-FCD MRIs (90%), reflecting the clinical data distribution. Furthermore, continual learning requires smaller training datasets, facilitating accelerated training and reducing computational and data storage requirements.

Gill et al. (2021) reported a deep-learning algorithm achieving the highest sensitivity (83%) and specificity (89%) in detecting FCDs to date. Their multicenter study focused on retrospective data from histologically confirmed FCD type II cases and comprised a control group of healthy individuals and individuals with temporal lobe epilepsy and

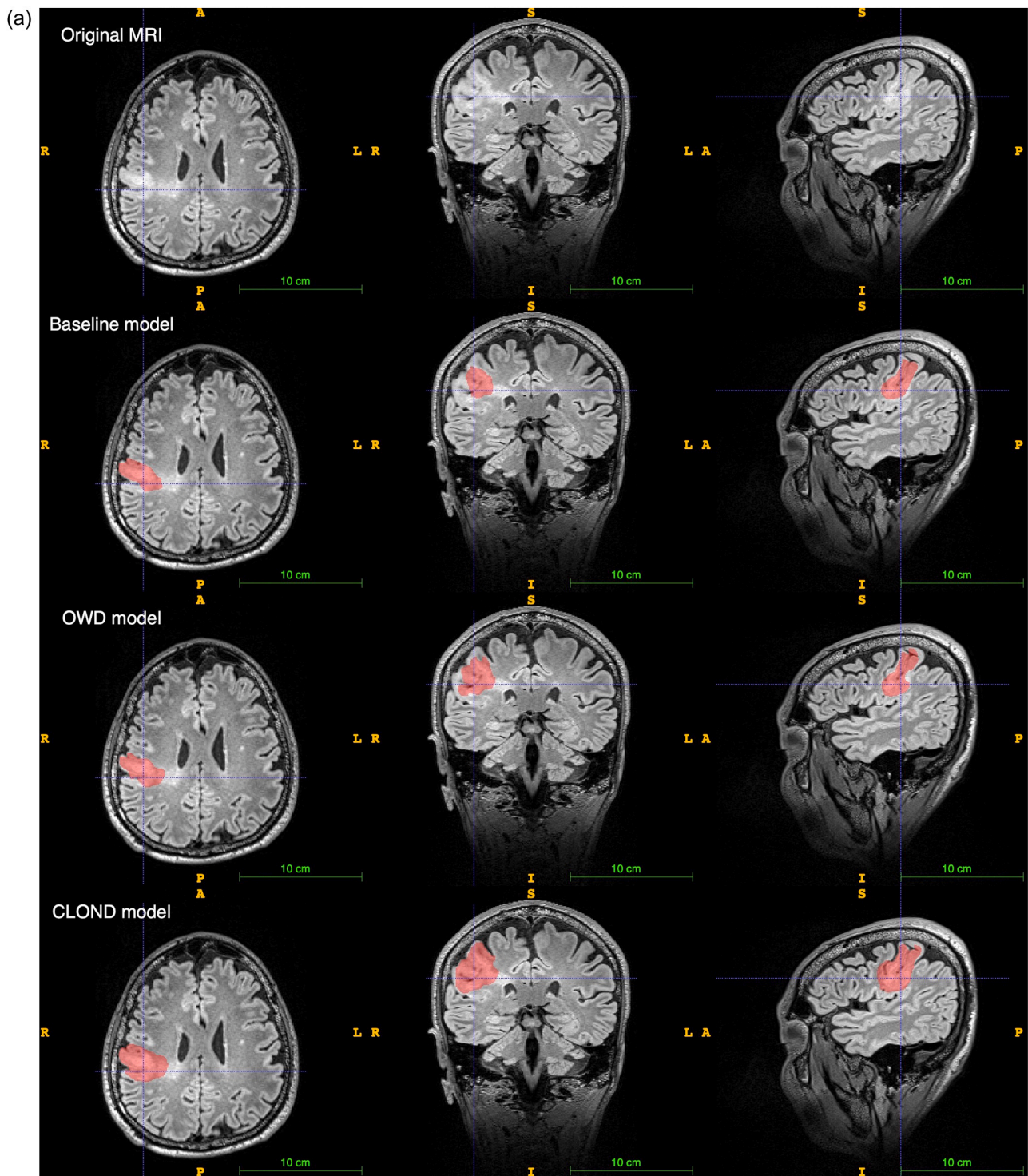


Fig. 5. a. Segmentation results of all three CNN models (baseline model, OWD, CLOND) in the final evaluation phase on an exemplary FCD. **b.** Segmentation results for a small FCD ($V = 0.51$ ml) in the first evaluation phase. Additionally, an infratentorial false positive cluster is found in the right cerebellum. Remarkably, infratentorial clusters were no longer detected after training.

hippocampal sclerosis. In contrast, the dataset of our baseline model includes 20 different histopathologically confirmed FCD types (I, II and III) (House et al., 2021). Moreover, our evaluation dataset is prospectively obtained from routine clinical practice, without the application of exclusion criteria, and includes a variety of non-FCD pathologies and low-quality scans. This approach ensures that our results are more representative of clinical reality. Our model achieved a higher

sensitivity of 90.0% for FCD detection with fewer FP clusters but demonstrated a lower overall specificity of 70.0% when considering all control group MRIs. Remarkably, when considering only normal MRIs as control group, a common practice in AI-based FCD detection studies (Demerath et al., 2022; Ganji et al., 2022; Spitzer et al., 2022; Urbach et al., 2022a; David et al., 2021), our model exhibited a comparable specificity of 90.0%.

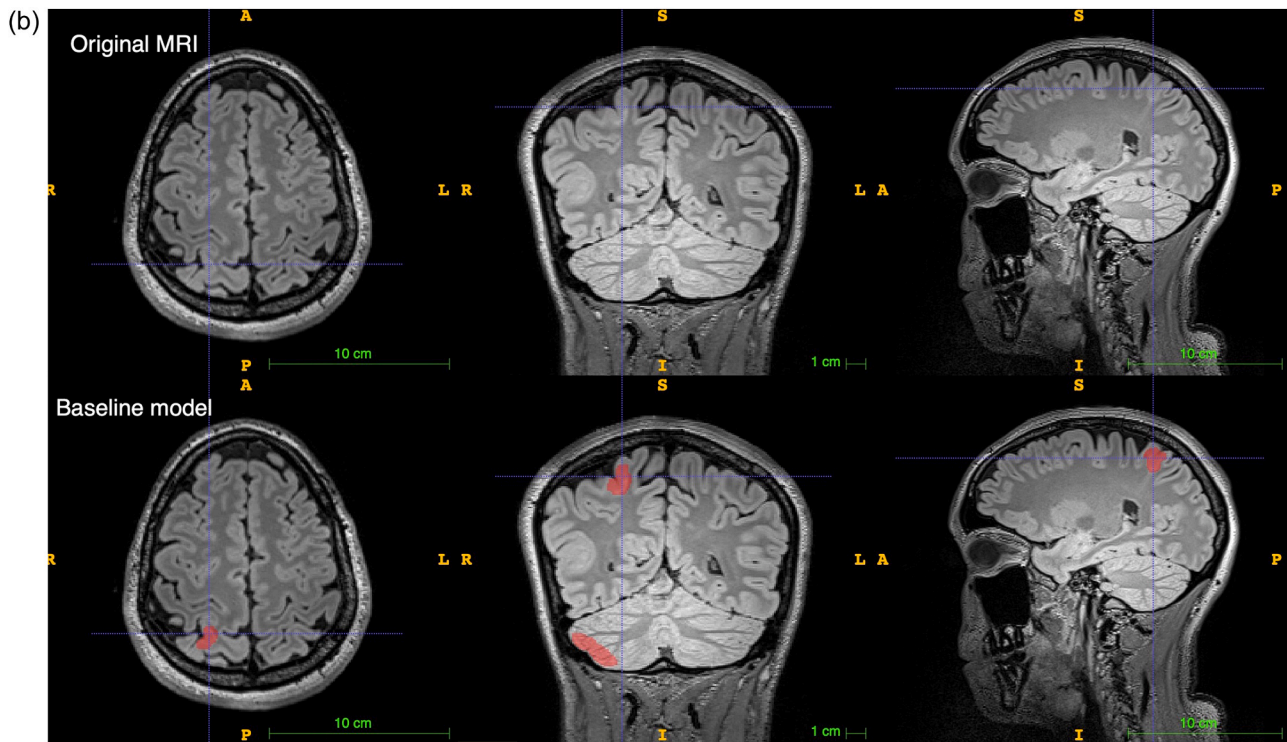


Fig. 5. (continued).

However, comparing our results with other studies is challenging due to variability in training and evaluation datasets with heterogeneous FCDs and control groups, ground truth definitions, exclusion criteria, study designs, and reported evaluation metrics. Many studies rely on retrospective data (Spitzer et al., 2022; David et al., 2021; Gill et al., 2021; Bijay Dev et al., 2019; Mo et al., 2019; Adler et al., 2017; Ahmed et al., 2015; Hong et al., 2014), report cross-validation results (Ganji et al., 2022; Wang et al., 2020; Jin et al., 2018; Mo et al., 2019; Adler et al., 2017; Ahmed et al., 2015; Hong et al., 2014), utilize healthy participants as control groups (Demerath et al., 2022; Ganji et al., 2022; Spitzer et al., 2022; Urbach et al., 2022a; David et al., 2021; Adler et al., 2017; Ahmed et al., 2015), and exclude poor quality MRIs (Spitzer et al., 2022; David et al., 2021; Mo et al., 2019; Jin et al., 2018; Adler et al., 2017), potentially leading to an overestimation of performance parameters and lacking alignment with clinical reality. Hence, we strongly support Walger et al.'s (2023) proposal to develop a public benchmark dataset to standardize AI model evaluations and improve cross-study comparability. In our efforts to enhance comparability, we have provided a detailed description of our model's performance parameters.

One limitation of our algorithm is its reliance on MRI data from a single scanner operating at a specific field strength, potentially limiting its broader clinical applicability. Single-scanner data may also lead to an overestimation of performance metrics (Varoquaux, 2018). Incorporating data from multiple scanners and field strengths for training and evaluation would provide a more comprehensive assessment of the model's performance and its applicability in real-world scenarios. However, leveraging continual learning with Domain-Incremental Learning offers promising potential to address these constraints, opening possibilities for further research in training our CNN with new data from different scanners, field strengths, and sequences.

Despite the large and diverse MRI dataset, selection bias may be present in our study population as all our data were obtained from a single specialized epilepsy center. These limitations impact the generalizability of our results to broader clinical contexts. These issues have recently been addressed by multi-center studies (Spitzer et al., 2022; Gill et al., 2021).

Furthermore, our CNN currently struggles to detect small FCDs and may miss subtle features, particularly in the presence of image noises or artifacts. It successfully detected only one out of four FCDs sized 0.5 ml or smaller. We attribute this restriction to the limited number of small FCDs in the training dataset which points to a need for additional training with such cases to enhance detection.

Although our prospective validation did not demonstrate significant improvements in sensitivity through training, it is important to consider that the study was underpowered to detect potential effects here due to the high baseline sensitivity and the low number of ten FCD cases per evaluation phase.

A downside of our CNN is the requirement to create manually labeled lesion maps for its training. While this meticulous approach ensures accurate training data, it also demands time and effort. Fortunately, a significant part of this workload is alleviated by the CNN's segmentation capability.

The main limitation of this study is the lack of histopathologic confirmation as no patient in the study population has yet undergone surgery or invasive EEG. This may raise concerns regarding the definitive classification of lesions. However, our reference method involved comprehensive assessments by experienced experts, considering neuroradiologic findings, morphometric analyses, seizure semiology, and EEG data. Over the past 14 years, this approach successfully detected over 800 FCD cases at our epilepsy center. In the last nine years, the histopathological diagnosis aligned with the preoperative morphometric suspicion in 48 out of 68 cases. Additionally, 12 cases were histopathologically diagnosed with cortical dysplasia despite the absence of visual morphologic suspicion. Out of the 20 non-positive cases, two cases lacked histopathology, and in 10 cases with suspected FCD type IIIa, only hippocampal sclerosis could be confirmed. Notably, four cases showed clinical improvement after surgery despite unremarkable histopathology. Reactive changes were observed in two patients, but the outcomes of their recent surgeries are still unknown. However, neuropathological detection depends on well-preserved anatomic specimens and characteristic features may be missed in small or highly fragmented biopsy specimens (Walger et al., 2023). Yet, the reference method may

have missed existing FCDs, and some clusters labeled as FP could potentially represent genuine FCDs. Nevertheless, our reference method demonstrates reasonable accuracy in detecting FCDs and all 20 histologically confirmed dysplastic lesions in the training cohort for the baseline model were correctly identified preoperatively (House et al., 2021).

The sensitivity of neuroradiological visual analysis in our current study, averaging 73.3%, is higher than the reported sensitivity of 33.3% in House et al. (2021). While the exact reasons for this discrepancy remain unclear, the enhanced expertise developed through regular interdisciplinary case discussions at our center may have played a role in this considerable improvement. Nonetheless, the CNN's sensitivity continues to outperform conventional visual neuroradiological analysis.

Recent studies have highlighted the superiority of 7 T MRI for FCD detection (Chen et al., 2021), offering improved SNR and gray-white matter differentiation. The use of advanced sequences like MP2RAGE (Demerath et al., 2022, 2020) and FLAWS (Sun et al., 2021), combined with postprocessing methods, has also improved lesion identification. Future implementation of these techniques into our CNN may enhance its performance. Additional training on stronger field strengths and sequences may not be necessary (Demerath et al., 2022). However, this remains subject for further studies.

To support translating these promising results into clinical practice and enabling independent validation, user-friendly software is currently being developed.

5. Conclusions

This study confirmed the superior sensitivity of our CNN in detecting FCDs compared to conventional neuroradiological analysis. Continual learning and data-driven training significantly improved the CNN's performance, addressing the previous model's main limitation of specificity. Our method exhibits robust performance achieving both a high sensitivity of 90.0% and a specificity of 70.0% on a large MRI dataset comprising 3D T1 and FLAIR sequences from daily clinical practice. The evaluation is fully automated and does not require specialized expertise making it suitable for easy FCD detection in any clinical setting, including outpatient clinics and hospitals without specialized epilepsy departments. Our algorithm has the potential to address the underutilization of epilepsy surgery in patients with drug-resistant focal epilepsy. Overall, our CNN offers a practical and efficient solution for FCD detection, continuously enhancing its performance through ongoing training with real-world data.

Ethics approval statement

All patients' MRIs were fully anonymized and thus not traceable in

Appendix

Color is crucial for the interpretation of the figures. In case of print publication, please ensure that all figures are reproduced in color to maintain accuracy and clarity.

reverse. Our study was performed in accordance with the Ethics Committee of the regional Medical Association Hamburg (WF-062/18) and the ethical standards laid down in the Declaration of Helsinki, as well as local and institutional laws. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Thomas Sauvigny: Writing – review & editing, Supervision. **Brigitte Holst:** Writing – review & editing, Resources. **Agata Chudzinska:** Software. **Vicky Chanra:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Bartosz Silski:** Software, Methodology. **Natalia Braniewska:** Writing – review & editing, Visualization, Software, Methodology. **Patrick M. House:** Writing – review & editing, Supervision, Project administration, Data curation, Conceptualization. **Sirko Pelzl:** Resources. **Stefan Stodieck:** Resources.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used DeepL Translator (<https://www.deepl.com/translator>) for language translation, improved clarity and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of Competing Interest

None of the authors has any conflict of interest to disclose.

Acknowledgements

We wish to thank Angela McOrmond, Juliane House-Edmondson, Maria Stark, Gerhard Schön and Mathias Trabs for constructive criticism, as well as Irene Lorenzi, Olga Simova, Michael Lanz, Berthold Voges and Anja Herzer for some of the MRI data.

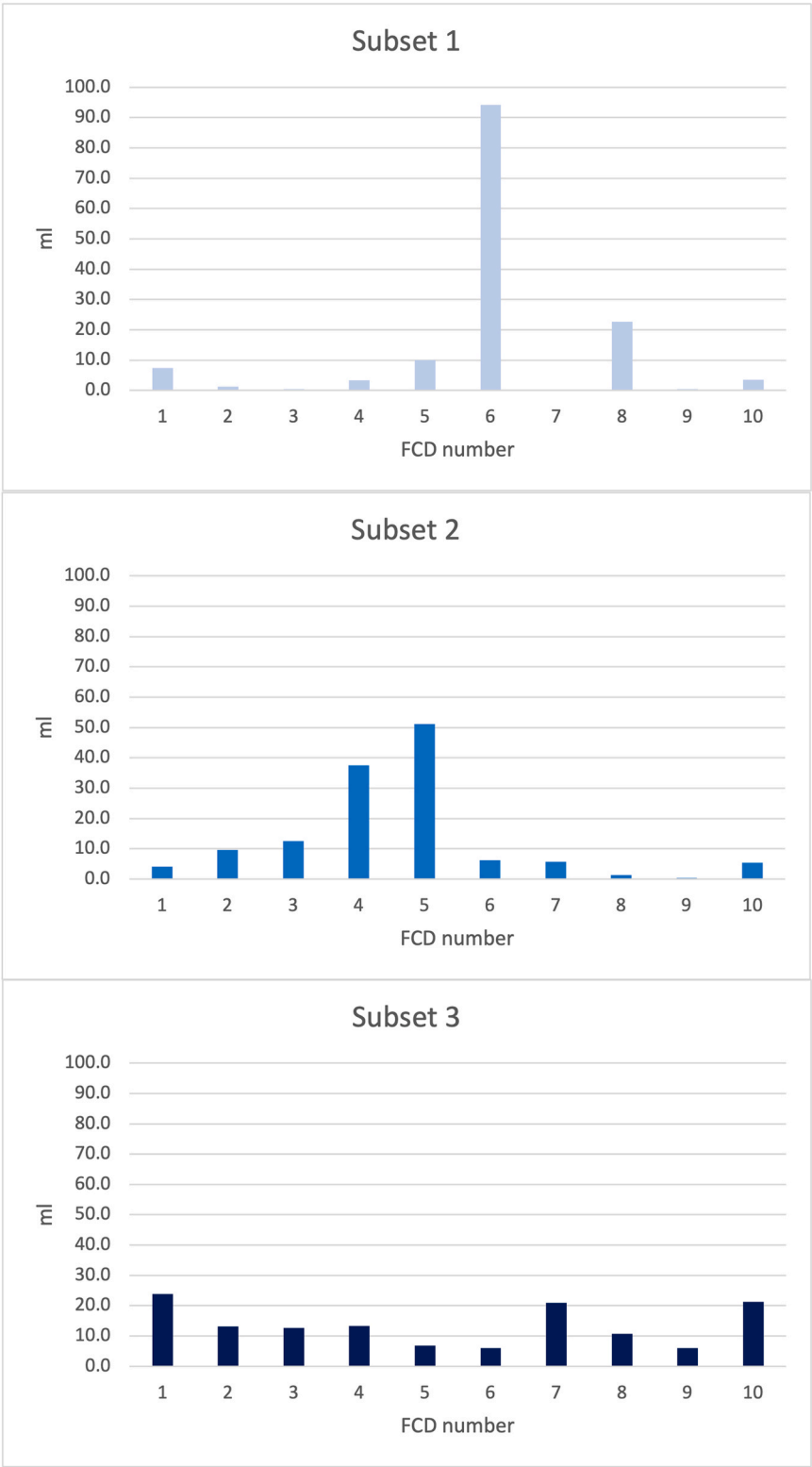


Figure A.1. FCD sizes according to expert-labeled maps (in ml).

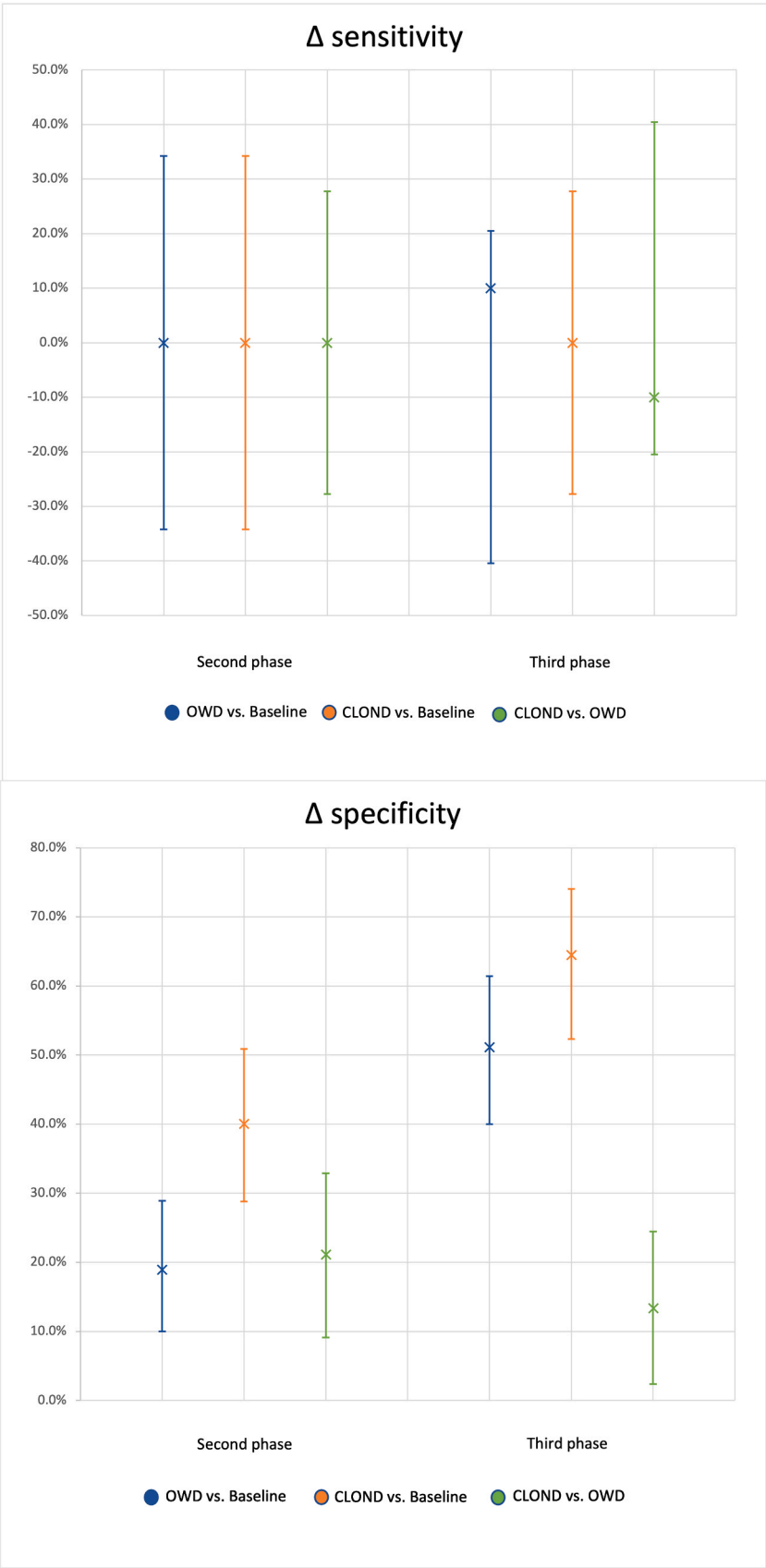


Figure A.2. Differences in sensitivity and specificity between the three CNN models with 95% Tango's asymptotic confidence intervals.

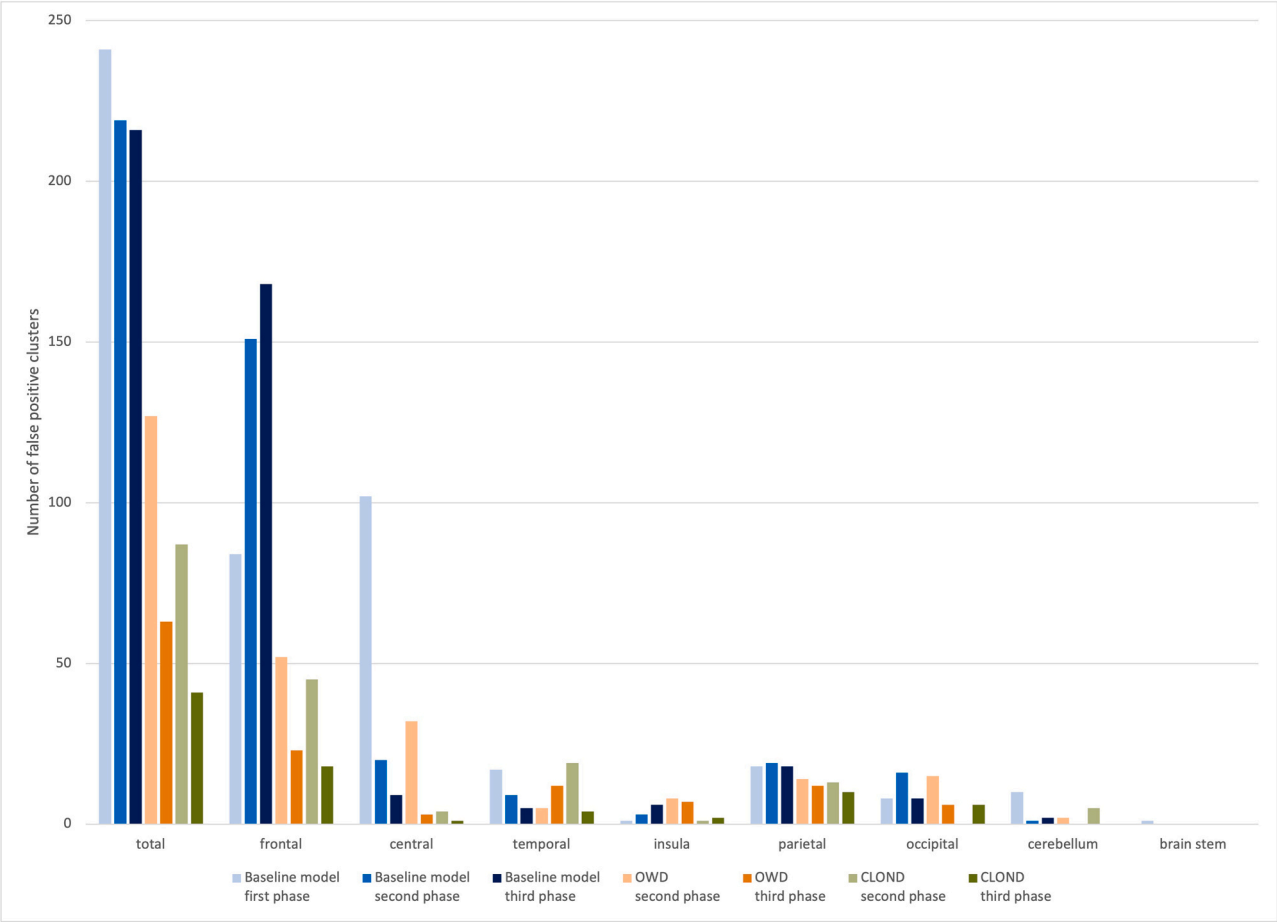


Figure A.3. Anatomical distribution of the false positive clusters.

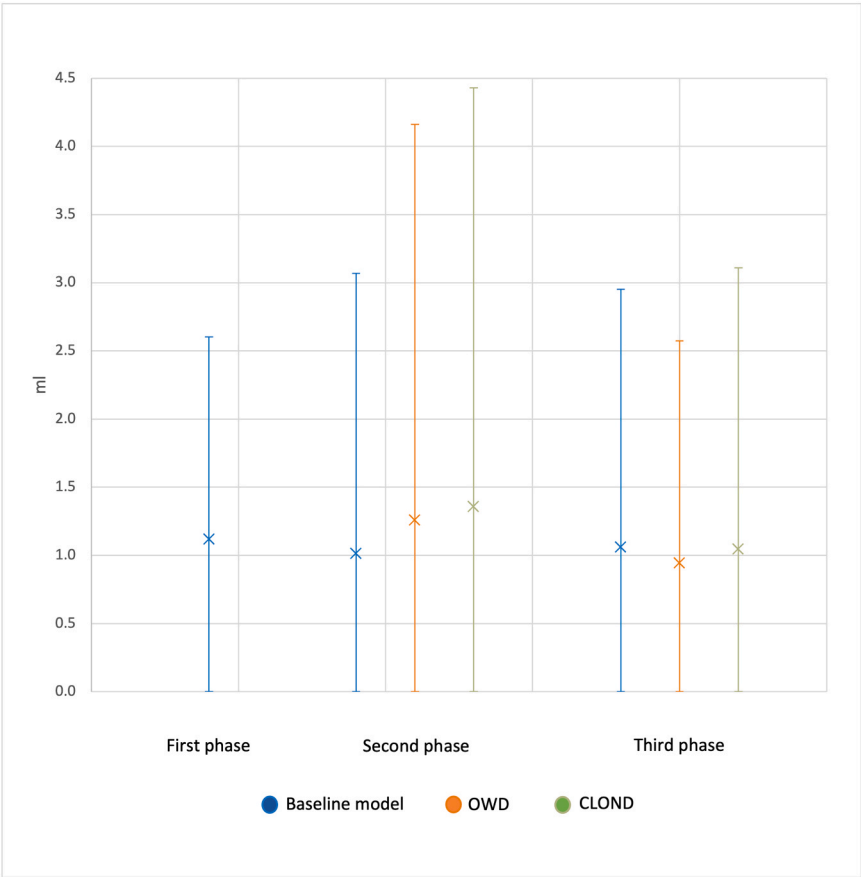


Figure A.4. Average false positive cluster sizes with standard deviation (in ml).

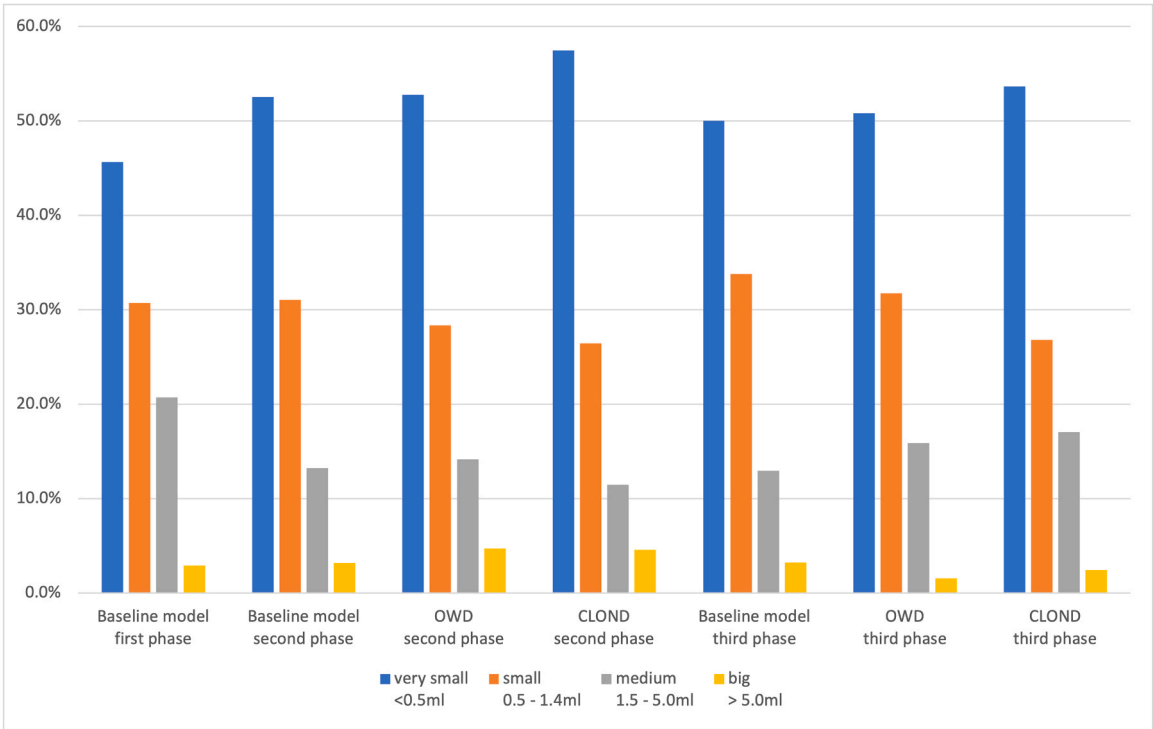


Figure A.5. Size distribution of false positive clusters.

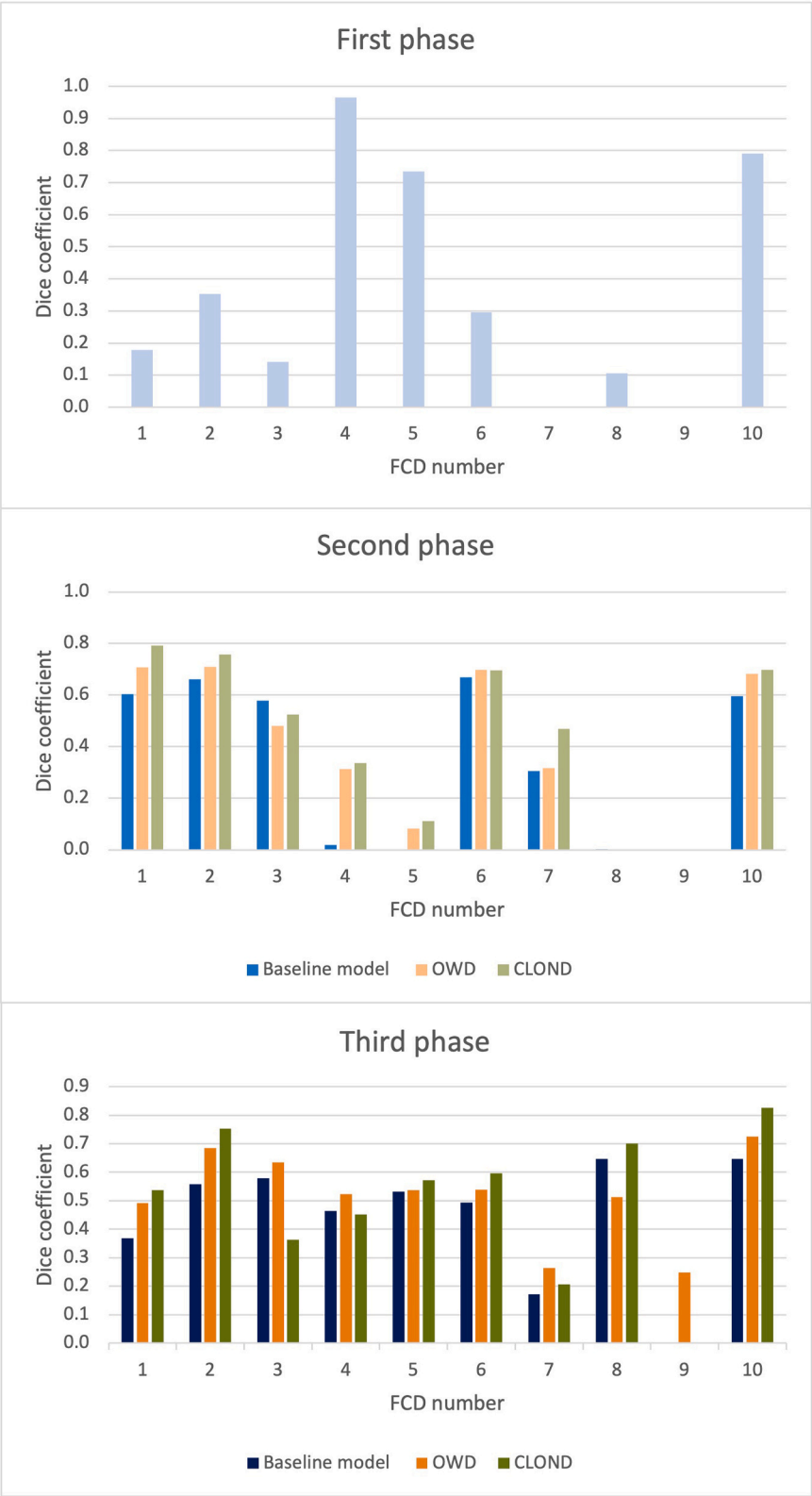


Figure A.6. FCD segmentation performance for each FCD (Dice coefficient).

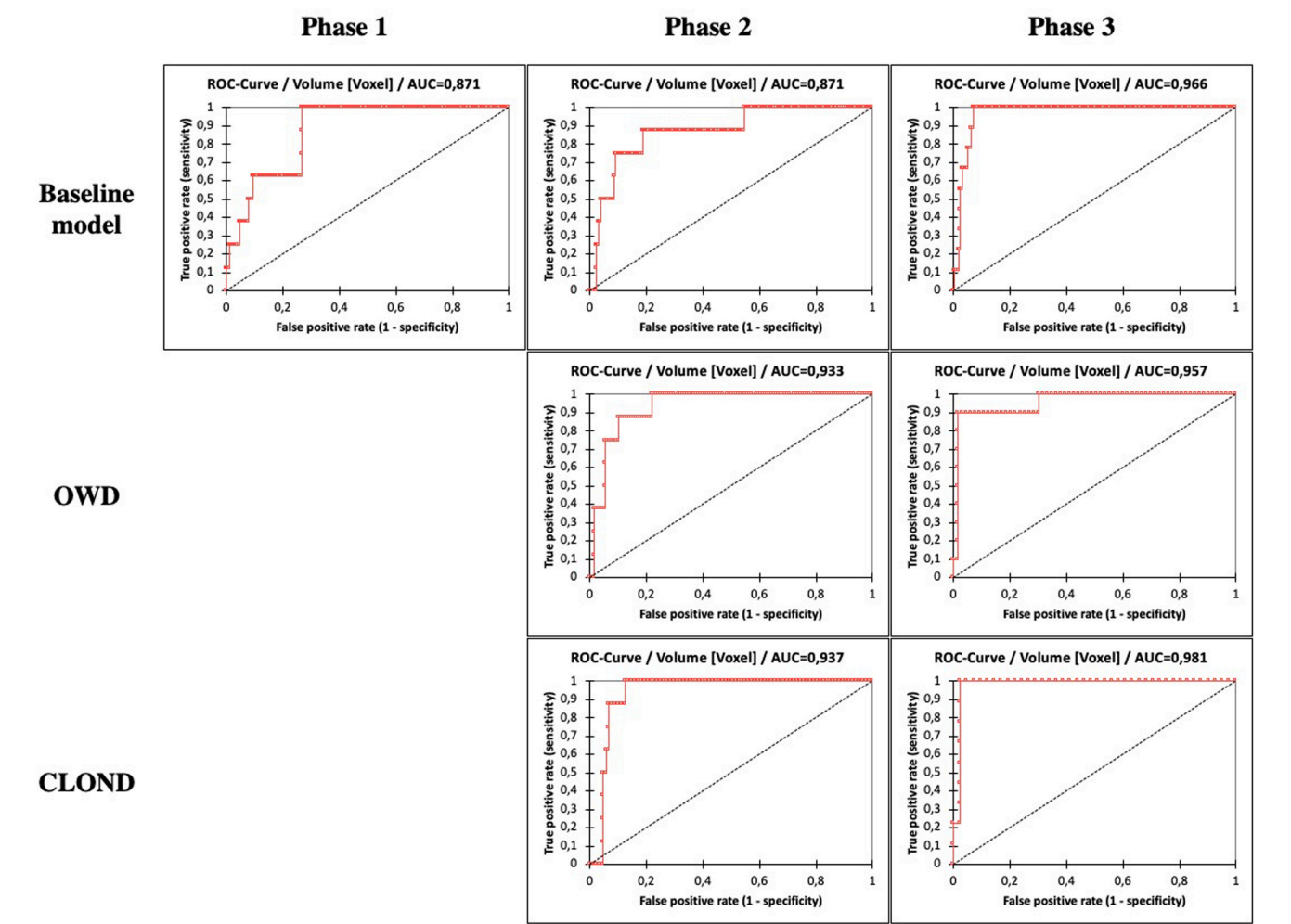


Figure A.7. Receiver operating characteristic (ROC) curves for each CNN model.

Table A.1
Anatomical characteristics of FCDs and focal PMG

		Subset 1 (n = 10 FCD)	Subset 2 (n = 10 FCD)	Subset 3 (n = 10 FCD)	Total
Size*	small	4	2	0	6 (20.0%)
	medium	3	4	3	10 (33.3%)
	big	3	4	7	14 (46.7%)
Hemisphere	right	6	8	7	21 (70.0%)
	left	4	2	3	9 (30.0%)
Transmantle sign	yes	4	2	4	10 (33.3%)
	no	6	8	6	20 (66.7%)
PMG		1	2	1	4 (13.3%)
Location	frontal	6	3	1	10 (33.3%)
	temporal	1	0	0	1 (3.3%)
	insula	1	3	1	5 (16.7%)
	parietal	2	4	6	12 (40.0%)
	occipital	0	0	2	2 (6.7%)
Detected by neuroradiologist		8 / 10	9 / 10	5 / 10	22 / 30 (73.3%)

Abbreviations: PMG – Focal polymicrogyria.
*small: < 1.5 ml; medium: 1.5 – 7.5 ml; big: > 7.5 ml.

Table A.2
Results of the subgroup analysis at the case and lesion level.

Model	Phase (each n = 100 cases)	Subgroup	Total cases	TP cases	FN cases	FP cases	TN cases	FP cluster	FP cluster per MRI	
Baseline model	1	FCD	10	8	2			11	1.1	
		non-FCD	40				40	0	105	2.6
		Normal	50				50	0	125	2.5
	2	FCD	10	8	2			10	1.0	
		non-FCD	40				39	1	101	2.5
		Normal	50				49	1	108	2.2
	3	FCD	10	9	1			5	0.5	
		non-FCD	40				38	2	95	2.4
		Normal	50				47	3	116	2.3
OWD	2	FCD	10	8	2			5	0.5	
		non-FCD	40				35	5	67	1.7
		Normal	50				36	14	55	1.1
	3	FCD	10	10	0			0	0.0	
		non-FCD	40				24	16	44	1.1
		Normal	50				15	35	19	0.4
CLOND	2	FCD	10	8	2			2	0.2	
		non-FCD	40				26	14	48	1.2
		Normal	50				26	24	37	0.7
	3	FCD	10	9	1			0	0.0	
		non-FCD	40				22	18	35	0.9
		Normal	50				5	45	6	0.1

Abbreviations: FN – False negative; FP – False positive; non-FCD – non-FCD pathological MRIs; TN – True negative; TP – True positive.

References

Adler, S., Wagstyl, K., Gunny, R., Ronan, L., Carmichael, D., Cross, J.H., Fletcher, P.C., Baldeweg, T., 2017. Novel surface features for automated detection of focal cortical dysplasias in paediatric epilepsy. *NeuroImage Clin.* 14, 18–27. <https://doi.org/10.1016/j.nicl.2016.12.030>.

Ahmed, B., Brodley, C.E., Blackmon, K.E., Kuzniecky, R., Barash, G., Carlson, C., Quinn, B.T., Doyle, W., French, J., Devinsky, O., Thesen, T., 2015. Cortical feature analysis and machine learning improves detection of “MRI-negative” focal cortical dysplasia. *Epilepsy Behav.* 48, 21–28. <https://doi.org/10.1016/j.yebeh.2015.04.055>.

Aich, A., 2021. Elastic Weight Consolidation (EWC): Nuts and Bolts. arXiv: 2105.04093v1. <https://doi.org/10.48550/arXiv.2105.04093>.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821. <https://doi.org/10.1006/nimg.2000.0582>.

Barkovich, A.J., Kuzniecky, R.I., 1996. Neuroimaging of focal malformations of cortical development. *J. Clin. Neurophysiol.* 13, 481–494. <https://doi.org/10.1097/00004691-199611000-00003>.

Bast, T., Ramantani, G., Seitz, A., Rating, D., 2006. Focal cortical dysplasia: prevalence, clinical presentation and epilepsy in children and adults. *Acta Neurol. Scand.* 113, 72–81. <https://doi.org/10.1111/j.1600-0404.2005.00555.x>.

Bijay Dev, K.M., Jogi, P.S., Niyas, S., Vinayagamani, S., Kesavadas, C., Rajan, J., 2019. Automatic detection and localization of Focal Cortical Dysplasia lesions in MRI using fully convolutional neural network. *Biomed. Signal Process. Control* 52, 218–225. <https://doi.org/10.1016/j.bspc.2019.04.024>.

Blumcke, I., Spreafico, R., Haaker, G., Coras, R., Kobow, K., Bien, C.G., Pfäfflin, M., Elger, C., Widman, G., Schramm, J., Becker, A., Braun, K.P., Leijten, F., Baayen, J.C., Aronica, E., Chassoux, F., Hamer, H., Stefan, H., Rössler, K., Thom, M., Walker, M.C., Sisodiya, S.M., Duncan, J.S., McEvoy, A.W., Pieper, T., Holthausen, H., Kudernatsch, M., Meencke, H.J., Kahane, P., Schulze-Bonhage, A., Zentner, J., Heiland, D.H., Urbach, H., Steinhoff, B.J., Bast, T., Tassi, L., Lo Russo, G., Özkara, C., Oz, B., Krsek, P., Vogelgesang, S., Runge, U., Lerche, H., Weber, Y., Honavar, M., Pimentel, J., Arzimanoglou, A., Ulate-Campos, A., Noachtar, S., Hartl, E., Schijns, O., Guerrini, R., Barba, C., Jacques, T.S., Cross, J.H., Feucht, M., Mühlebner, A., Grunwald, T., Trinka, E., Winkler, P.A., Gil-Nagel, A., Toledano Delgado, R., Mayer, T., Lutz, M., Zountsas, B., Garganis, K., Rosenow, F., Hermesen, A., von Oertzen, T.J., Diepgen, T.L., Avanzini, G., 2017. Histopathological findings in brain tissue obtained during epilepsy surgery. *N. Engl. J. Med.* 377, 1648–1656. <https://doi.org/10.1056/nejmoa1703784>.

Blumcke, I., Budday, S., Poduri, A., Lal, D., Kobow, K., Baulac, S., 2021. Neocortical development and epilepsy: insights from focal cortical dysplasia and brain tumours. *Lancet Neurol.* 20, 943–955. [https://doi.org/10.1016/s1474-4422\(21\)00265-9](https://doi.org/10.1016/s1474-4422(21)00265-9).

Chen, C., Xie, J.-J., Ding, F., Jiang, Y.-S., Jin, B., Wang, Shan, Ding, Y., Li, H., Jiang, B., Zhu, J.-M., Ding, M.-P., Chen, Z., Wu, Z.-Y., Zhang, B.-R., Hsu, Y.-C., Lai, H.-Y., Wang, Shuang, 2021. 7T MRI with post-processing for the presurgical evaluation of pharmacoresistant focal epilepsy. *Ther. Adv. Neurol. Disord.* 14, 17562864211021181 <https://doi.org/10.1177/17562864211021181>.

Chen, X., Qian, T., Kober, T., Zhang, G., Ren, Z., Yu, T., Piao, Y., Chen, N., Li, K., 2018. Gray-matter-specific MR imaging improves the detection of epileptogenic zones in focal cortical dysplasia: A new sequence called fluid and white matter suppression (FLAWS). *NeuroImage Clin.* 20, 388–397. <https://doi.org/10.1016/j.nicl.2018.08.010>.

Crome, L., 1957. Infantile cerebral gliosis with giant nerve cells. *J. Neurol. Neurosurg. Psychiatry* 20, 117–124. <https://doi.org/10.1136/jnnp.20.2.117>.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. *Neuroimage* 9, 179–194. <https://doi.org/10.1006/nimg.1998.0395>.

David, B., Kröll-Seger, J., Schuch, F., Wagner, J., Wellmer, J., Woermann, F., Oehl, B., Van Paesschen, W., Breyer, T., Becker, A., Vatter, H., Hattingen, E., Urbach, H., Weber, B., Surges, R., Elger, C.E., Huppertz, H.-J., Rüber, T., 2021. External validation of automated focal cortical dysplasia detection using morphometric analysis. *Epilepsia* 62, 1005–1021. <https://doi.org/10.1111/epi.16853>.

van de Ven, G.M., Tuytelaars, T., Tolias, A.S., 2022. Three types of incremental learning. *Nat. Mach. Intell.* 4, 1185–1197. <https://doi.org/10.1038/s42256-022-00568-3>.

Demerath, T., Rubensdörfer, L., Schwarzwald, R., Schulze-Bonhage, A., Altenmüller, D.-M., Kaller, C., Kober, T., Huppertz, H.-J., Urbach, H., 2020. Morphometric MRI analysis: Improved detection of focal cortical dysplasia using the MP2RAGE sequence. *AJNR. Am. J. Neuroradiol.* 41, 1009–1014. <https://doi.org/10.3174/ajnr.A6579>.

Demerath, T., Kaller, C.P., Heers, M., Staack, A., Schwarzwald, R., Kober, T., Reisert, M., Schulze-Bonhage, A., Huppertz, H.-J., Urbach, H., 2022. Fully automated detection of focal cortical dysplasia: Comparison of MP2RAGE and MP2RAGE sequences. *Epilepsia* 63, 75–85. <https://doi.org/10.1111/epi.17127>.

Duncan, J.S., Winston, G.P., Koepp, M.J., Ourselin, S., 2016. Brain imaging in the assessment for epilepsy surgery. *Lancet Neurol.* 15, 420–433. [https://doi.org/10.1016/s1474-4422\(15\)00383-x](https://doi.org/10.1016/s1474-4422(15)00383-x).

Fausser, S., Huppertz, H.-J., Bast, T., Strobl, K., Pantazis, G., Altenmueller, D.-M., Feil, B., Rona, S., Kurth, C., Rating, D., Korinthenberg, R., Steinhoff, B.J., Volk, B., Schulze-Bonhage, A., 2006. Clinical characteristics in focal cortical dysplasia: a retrospective evaluation in a series of 120 patients. *Brain* 129, 1907–1916. <https://doi.org/10.1093/brain/awl133>.

Feng, C., Zhao, H., Li, Y., Wen, J., 2020. Automatic localization and segmentation of focal cortical dysplasia in FLAIR-negative patients using a convolutional neural network. *J. Appl. Clin. Med. Phys.* 21, 215–226. <https://doi.org/10.1002/acm2.12985>.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. *Neuroimage* 9, 195–207. <https://doi.org/10.1006/nimg.1998.0396>.

Foiadelli, T., Lagae, L., Goffin, K., Theys, T., De Amici, M., Sacchi, L., Van Loon, J., Savasta, S., Jansen, K., 2020. Subtraction Ictal SPECT coregistered to MRI (SISCOM) as a guide in localizing childhood epilepsy. *Epilepsia Open* 5, 61–72. <https://doi.org/10.1002/epi4.12373>.

Ganji, Z., Aghaee Hakak, M., Zare, H., 2022. Comparison of machine learning methods for the detection of focal cortical dysplasia lesions: decision tree, support vector machine and artificial neural network. *Neurol. Res.* 44, 1142–1149. <https://doi.org/10.1080/01616412.2022.2112381>.

Ghiassi, G., Lin, T.-Y., Le, Q.V., 2018. DropBlock: A regularization method for convolutional networks. arXiv:1810.12890. <https://doi.org/10.48550/arXiv.1810.12890>.

Gill, R.S., Lee, H.-M., Caldairou, B., Hong, S.-J., Barba, C., Deleo, F., D’Incerti, L., Mendes Coelho, V.C., Lenge, M., Semmler, M., Schrader, D.V., Bartolomei, F., Guye, M., Schulze-Bonhage, A., Urbach, H., Cho, K.H., Cendes, F., Guerrini, R., Jackson, G., Hogan, R.E., Bernasconi, N., Bernasconi, A., 2021. Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. *Neurology* 97, e1571–e1582. <https://doi.org/10.1212/WNL.00000000000012698>.

Guerrini, R., Barba, C., 2021. Focal cortical dysplasia: an update on diagnosis and treatment. *Expert Rev. Neurother.* 21, 1213–1224. <https://doi.org/10.1080/14737175.2021.1915135>.

- Hong, S.-J., Kim, H., Schrader, D., Bernasconi, N., Bernhardt, B.C., Bernasconi, A., 2014. Automated detection of cortical dysplasia type II in MRI-negative epilepsy. *Neurology* 83, 48–55. <https://doi.org/10.1212/WNL.0000000000000543>.
- House, P.M., Kopelman, M., Braniewska, N., Silski, B., Chudzinska, A., Holst, B., Sauvigny, T., Martens, T., Stodieck, S., Pelzl, S., 2021. Automated detection and segmentation of focal cortical dysplasias (FCD) with artificial intelligence: Presentation of a novel convolutional neural network and its prospective clinical validation. *Epilepsy Res* 172. <https://doi.org/10.1016/j.eplepsyres.2021.106594>.
- House, P.M., Lanz, M., Holst, B., Martens, T., Stodieck, S., Huppertz, H.-J., 2013. Comparison of morphometric analysis based on T1- and T2-weighted MRI data for visualization of focal cortical dysplasia. *Epilepsy Res* 106, 403–409.
- Howard, J., 2018. Fastai; The fastai deep learning library. Available at: (<https://github.com/fastai/fastai>).
- Huppertz, H.-J., Grimm, C., Fauser, S., Kassubek, J., Mader, I., Hochmuth, A., Spreer, J., Schulze-Bonhage, A., 2005. Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis. *Epilepsy Res* 67, 35–50. <https://doi.org/10.1016/j.eplepsyres.2005.07.009>.
- Huppertz, H.-J., Wellmer, J., Staack, A.M., Altenmüller, D.-M., Urbach, H., Kröll, J., 2008. Voxel-based 3D MRI analysis helps to detect subtle forms of subcortical band heterotopia. *Epilepsia* 49, 772–785. <https://doi.org/10.1111/j.1528-1167.2007.01436.x>.
- Jin, B., Krishnan, B., Adler, S., Wagstyl, K., Hu, W., Jones, S., Najm, I., Alexopoulos, A., Zhang, K., Zhang, J., Ding, M., Wang, S., Pediatric Imaging, Neurocognition, Genetics Study, Wang, Z.I., 2018. Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia* 59, 982–992. <https://doi.org/10.1111/epi.14064>.
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabacka-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., Hadsell, R., 2017. Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. U. S. A.* 114, 3521–3526. <https://doi.org/10.1073/pnas.1611835114>.
- Lamberink, H.J., Otte, W.M., Blümcke, I., Braun, K.P.J., European Epilepsy Brain Bank writing group, study group, European Reference Network EpiCARE, 2020. Seizure outcome and use of antiepileptic drugs after epilepsy surgery according to histopathological diagnosis: a retrospective multicentre cohort study. *Lancet Neurol.* 19, 748–757. [https://doi.org/10.1016/S1474-4422\(20\)30220-9](https://doi.org/10.1016/S1474-4422(20)30220-9).
- Li, B., Wu, F., Wu, Weinberger, K.Q., Belongie, S., 2019. Position. Norm. <https://doi.org/10.48550/arXiv.1907.04312> arXiv:1907.04312.
- Maiworm, M., Nöth, U., Hattingen, E., Steinmetz, H., Knake, S., Rosenow, F., Deichmann, R., Wagner, M., Gracien, R.-M., 2020. Improved visualization of focal cortical dysplasia with surface-based multiparametric quantitative MRI. *Front. Neurosci.* 14, 622. <https://doi.org/10.3389/fnins.2020.00622>.
- Marcelis, S., Vanden Bossche, S., Dekeyser, S., 2022. Not your everyday FCD: Imaging findings of focal cortical dysplasia type 1. *J. Belg. Soc. Radiol.* 106, 39. <https://doi.org/10.5334/jbsr.2710>.
- McCloskey, M., Cohen, N.J., 1989. Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*. Elsevier, pp. 109–165. [https://doi.org/10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8).
- Middlebrooks, E.H., Lin, C., Westerhold, E., Okromelidze, L., Vibhute, P., Grewal, S.S., Gupta, V., 2020. Improved detection of focal cortical dysplasia using a novel 3D imaging sequence: Edge-Enhancing Gradient Echo (3D-EDGE) MRI. *NeuroImage Clin.* 28, 102449. <https://doi.org/10.1016/j.nicl.2020.102449>.
- Mo, J.-J., Zhang, J.-G., Li, W.-L., Chen, C., Zhou, N.-J., Hu, W.-H., Zhang, C., Wang, Y., Wang, X., Liu, C., Zhao, B.-T., Zhou, J.-J., Zhang, K., 2019. Clinical value of machine learning in the automated detection of focal cortical dysplasia using quantitative multimodal surface-based features. *Front. Neurosci.* 12, 1008. <https://doi.org/10.3389/fnins.2018.01008>.
- Mohamed, I.S., Toffa, D.H., Robert, M., Cossette, P., Bérubé, A.-A., Saint-Hilaire, J.-M., Bouthillier, A., Nguyen, D.K., 2020. Utility of magnetic source imaging in nonlesional focal epilepsy: a prospective study. *Neurosurg. Focus* 48, E16. <https://doi.org/10.3171/2020.1.fOCUS19877>.
- Moloney, P.B., Cavalleri, G.L., Delanty, N., 2021. Epilepsy in the mTORopathies: opportunities for precision medicine. *Brain Commun.* 3. <https://doi.org/10.1093/braincomms/fcab222>.
- Myronenko, A., 2018. 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization. arXiv:1810.11654. <https://doi.org/10.48550/arXiv.1810.11654>.
- Najm, I., Lal, D., Alonso Vanegas, M., Cendes, F., Lopes-Cendes, I., Palmieri, A., Paglioli, E., Sarnat, H.B., Walsh, C.A., Wiebe, S., Aronica, E., Baulac, S., Coras, R., Kobow, K., Cross, J.H., Garbelli, R., Holthausen, H., Rössler, K., Thom, M., El-Osta, A., Lee, J.H., Miyata, H., Guerrini, R., Piao, Y.-S., Zhou, D., Blümcke, I., 2022. The ILAE consensus classification of focal cortical dysplasia: An update proposed by an ad hoc task force of the ILAE diagnostic methods commission. *Epilepsia* 63, 1899–1919. <https://doi.org/10.1111/epi.17301>.
- Otsuka, K., Egawa, K., Fujima, N., Kudo, K., Terae, S., Nakajima, M., Ito, T., Yagyu, K., Shiraiishi, H., 2021. Reinterpretation of magnetic resonance imaging findings with magnetoencephalography can improve the accuracy of detecting epileptogenic cortical lesions. *Epilepsy Behav.* 114, 107516. <https://doi.org/10.1016/j.yebeh.2020.107516>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An imperative style, high-performance deep learning library. arXiv [cs.LG]. (<http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>).
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science*. Springer International Publishing, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
- Salemdawod, A., Wach, J., Banat, M., Borger, V., Hamed, M., Haberl, H., Sassen, R., Radbruch, A., Becker, A.J., Vatter, H., Surges, R., Sarikaya-Seiwert, S., 2022. Predictors of postoperative long-term seizure outcome in pediatric patients with focal cortical dysplasia type II at a German tertiary epilepsy center. *J. Neurosurg. Pediatr.* 29, 83–91. <https://doi.org/10.3171/2021.7.PEDS21219>.
- Seong, M.J., Choi, S.J., Joo, E.Y., Shon, Y.-M., Seo, D.-W., Hong, S.B., Hong, S.C., 2021. Surgical outcome and prognostic factors in epilepsy patients with MR-negative focal cortical dysplasia. *PLoS One* 16, e0249929. <https://doi.org/10.1371/journal.pone.0249929>.
- Spitzer, H., Ripart, M., Whitaker, K., D'Arco, F., Mankad, K., Chen, A.A., Napolitano, A., De Palma, L., De Benedictis, A., Folds, S., Humphreys, Z., Zhang, K., Hu, W., Mo, J., Likeman, M., Davies, S., Güttler, C., Lenge, M., Cohen, N.T., Tang, Y., Wang, S., Chari, A., Tisdall, M., Bargallo, N., Conde-Blanco, E., Pariente, J.C., Pascual-Diaz, S., Delgado-Martínez, I., Pérez-Enríquez, C., Lagorio, I., Abela, E., Mullatti, N., O'Muircheartaigh, J., Vecchiato, K., Liu, Y., Caligiuri, M.E., Sinclair, B., Vivash, L., Willard, A., Kandasamy, J., McLellan, A., Sokol, D., Semmelroch, M., Kloster, A.G., Opheim, G., Ribeiro, L., Yasuda, C., Rossi-Espagnet, C., Hamandi, K., Tietze, A., Barba, C., Guerrini, R., Gaillard, W.D., You, X., Wang, I., González-Ortiz, S., Severino, M., Striano, P., Tortora, D., Kälviäinen, R., Gambardella, A., Labate, A., Desmond, P., Lui, E., O'Brien, T., Shetty, J., Jackson, G., Duncan, J.S., Winston, G.P., Pinborg, L.H., Cendes, F., Theis, F.J., Shinohara, R.T., Cross, J.H., Baldeweg, T., Adler, S., Wagstyl, K., 2022. Interpretable surface-based detection of focal cortical dysplasias: a Multi-centre Epilepsy Lesion Detection study. *Brain* 145, 3859–3871. <https://doi.org/10.1093/brain/awac224>.
- Sun, K., Yu, T., Yang, D., Ren, Z., Qiao, L., Ni, D., Wang, X., Zhao, Y., Chen, X., Xiang, J., Chen, N., Gao, R., Yang, K., Lin, Y., Kober, T., Zhang, G., 2021. Fluid and white matter suppression imaging and voxel-based morphometric analysis in conventional magnetic resonance imaging-negative epilepsy. *Front. Neurol.* 12, 651592. <https://doi.org/10.3389/fneur.2021.651592>.
- Tang, Y., Blümcke, I., Su, T.-Y., Choi, J.Y., Krishnan, B., Murakami, H., Alexopoulos, A. V., Najm, I.M., Jones, S.E., Wang, Z.I., 2022. Black line sign in focal cortical dysplasia IIB: A 7T MRI and electroclinicalopathologic study. *Neurology* 99, e616–e626. <https://doi.org/10.1212/wnl.0000000000000702>.
- Taylor, D.C., Falconer, M.A., Bruton, C.J., Corsellis, J.A.N., 1971. Focal dysplasia of the cerebral cortex in epilepsy. *J. Neurol. Neurosurg. Psychiatry* 34, 369–387. <https://doi.org/10.1136/jnnp.34.4.369>.
- Urbach, H., Scheffler, B., Heinrichsmeier, T., von Oertzen, J., Kral, T., Wellmer, J., Schramm, J., Wiestler, O.D., Blümcke, I., 2002. Focal cortical dysplasia of Taylor's balloon cell type: a clinicopathological entity with characteristic neuroimaging and histopathological features, and favorable postsurgical outcome. *Epilepsia* 43, 33–40. <https://doi.org/10.1046/j.1528-1157.2002.38201.x>.
- Urbach, H., Kellner, E., Kremers, N., Blümcke, I., Demerath, T., 2022b. MRI of focal cortical dysplasia. *Neuroradiology* 64, 443–452. <https://doi.org/10.1007/s00234-021-02865-x>.
- Urbach, H., Heers, M., Altenmueller, D.M., 2022a. Within a minute' detection of focal cortical dysplasia. *Neuroradiology* 64, 715–726. <https://doi.org/10.1007/s00234-021-02823-7>.
- Varoquaux, G., 2018. Cross-validation failure: Small sample sizes lead to large error bars. *Neuroimage* 180, 68–77. <https://doi.org/10.1016/j.neuroimage.2017.06.061>.
- Walger, L., Adler, S., Wagstyl, K., Henschel, L., David, B., Borger, V., Hattingen, E., Vatter, H., Elger, C.E., Baldeweg, T., Rosenow, F., Urbach, H., Becker, A., Radbruch, A., Surges, R., Reuter, M., Cendes, F., Wang, Z.I., Huppertz, H.-J., Rübner, T., 2023. Artificial intelligence for the detection of focal cortical dysplasia: Challenges in translating algorithms into clinical practice. *Epilepsia* 64, 1093–1112. <https://doi.org/10.1111/epi.17522>.
- Wanders, A., Garibotto, V., Spinelli, L., Beniczky, S., Vulliemoz, S., Daniel, R.T., Schaller, K., Bartoli, A., Korff, C., Seeck, M., 2022. High density electric source imaging in childhood-onset epilepsy due to focal cortical dysplasia. *Clin. Neurophysiol. Pract.* 7, 245–251. <https://doi.org/10.1016/j.cnp.2022.07.002>.
- Wang, H., Ahmed, S.N., Mandal, M., 2020. Automated detection of focal cortical dysplasia using a deep convolutional neural network. *Comput. Med. Imaging Graph.* 79, 101662. <https://doi.org/10.1016/j.compmedimag.2019.101662>.
- Wang, Y., Yu, T., Blümcke, I., Cai, Y., Sun, K., Gao, R., Wang, Yujiao, Fu, Y., Wang, W., Wang, Yuping, Zhang, G., Piao, Y., 2023. The clinico-pathological characterisation of focal cortical dysplasia type IIB genetically defined by MTOR mosaicism. *Neuropathol. Appl. Neurobiol.* 49. <https://doi.org/10.1111/nan.12874>.
- Wang, Z.I., Jones, S.E., Jaisani, Z., Najm, I.M., Prayson, R.A., Burgess, R.C., Krishnan, B., Ristic, A., Wong, C.H., Bingham, W., Gonzalez-Martinez, J.A., Alexopoulos, A.V., 2015. Voxel-based morphometric magnetic resonance imaging (MRI) postprocessing in MRI-negative epilepsies: MAP in Nonlesional Epilepsies. *Ann. Neurol.* 77, 1060–1075. <https://doi.org/10.1002/ana.24407>.
- Willard, A., Antonic-Baker, A., Chen, Z., O'Brien, T.J., Kwan, P., Perucca, P., 2022. Seizure outcome after surgery for MRI-diagnosed focal cortical dysplasia: A systematic review and meta-analysis. *Neurology* 98, e236–e248. <https://doi.org/10.1212/wnl.00000000000013066>.
- Wiwattanadittakul, N., Suwannachote, S., You, X., Cohen, N.T., Tran, T., Phuachantuck, R., Tsuchida, T.N., Depositario-Cabacar, D.F., Zelleke, T., Schreiber, J.M., Conry, J.A., Kao, A., Bartolini, L., Olutigbo, C., Almira-Suarez, M.I., Havens, K., Whitehead, M.T., Gaillard, W.D., 2021. Spatiotemporal distribution and

- age of seizure onset in a pediatric epilepsy surgery cohort with cortical dysplasia. *Epilepsy Res* 172, 106598. <https://doi.org/10.1016/j.eplepsyres.2021.106598>.
- Ying, Z., Wang, L., Blümcke, I., Bulacio, J., Alexopoulos, A., Jehi, L., Bingaman, W., Gonzalez-Martinez, J., Kobow, K., Niestroj, L.M., Lal, D., Koelble, K., Najm, I., 2019. A comprehensive clinico-pathological and genetic evaluation of bottom-of-sulcus focal cortical dysplasia in patients with difficult-to-localize focal epilepsy. *Epileptic Disord.* 21, 65–77. <https://doi.org/10.1684/epd.2019.1028>.
- Yokota, H., Uetani, H., Tatekawa, H., Hagiwara, A., Morimoto, E., Linetsky, M., Yoo, B., Ellingson, B.M., Salamon, N., 2020. Focal cortical dysplasia imaging discrepancies between MRI and FDG-PET: Unique association with temporal lobe location. *Seizure* 81, 180–185. <https://doi.org/10.1016/j.seizure.2020.08.017>.

3. Zusammenfassung

Deutsch:

Fokale kortikale Dysplasien (FCD) sind die häufigste Ursache pharmakoresistenter Epilepsien im Kindesalter und die dritthäufigste bei Erwachsenen. Eine vollständige Resektion der Läsion bietet die beste Chance auf langfristige Anfallsfreiheit. Trotz Fortschritten in der Bildgebung bleibt die FCD-Erkennung eine klinische Herausforderung. Methoden des maschinellen Lernens eröffnen vielversprechende Möglichkeiten für eine automatisierte Detektion. House et al. (2021) entwickelten ein Convolutional Neural Network (CNN) zur FCD-Detektion und Segmentierung, dessen klinische Anwendbarkeit aufgrund der geringen Spezifität von 5,5 % begrenzt war. Im Rahmen einer prospektiven Studie wurde dieses CNN mithilfe von 300 3T-MRTs aus der klinischen Routine, darunter 30 FCD-Fälle, weiterentwickelt. Das Training erfolgte unter Anwendung klassischer und kontinuierlicher Lernmethoden. Das mit kontinuierlichem Lernen trainierte Modell erzielte eine verbesserte Leistung mit einer Sensitivität von 90 %, einer Spezifität von 70 % und einer Genauigkeit von 72 %, bei durchschnittlich 0,41 falsch-positiven Clustern pro MRT und einem mittleren Dice-Koeffizienten von 0,56 für die FCD-Segmentierung. Darüber hinaus zeigte das Modell stetige Leistungsverbesserungen und übertraf das klassisch trainierte Modell. Damit stellt das weiterentwickelte CNN eine effiziente und vielversprechende Lösung für die automatisierte Detektion und Segmentierung von FCDs im klinischen Alltag dar.

English:

Focal cortical dysplasias (FCD) are the leading cause of pharmaco-resistant epilepsy in children and the third most common cause in adults. Complete resection of the lesion offers the best chance for long-term seizure freedom. Despite advances in imaging techniques, FCD detection remains a clinical challenge. Machine learning methods present promising opportunities for automated detection. House et al. (2021) developed a convolutional neural network (CNN) for FCD detection and segmentation, but its clinical applicability was limited by a low specificity of 5.5%. In a prospective study, this CNN was further developed using 300 3T-MRIs from routine clinical practice, including 30 FCD cases. Training was conducted using both classical and continual learning methods. The model trained with continual learning demonstrated improved performance, achieving a sensitivity of 90%, a specificity of 70%, and an accuracy of 72%, with an average of 0.41 false-positive clusters per MRI and a mean Dice coefficient of 0.56 for FCD segmentation. Additionally, the model demonstrated steady performance improvements and outperformed the classically trained model. The optimized CNN thus presents an efficient and promising solution for the automated detection and segmentation of FCD in clinical practice.

4. Literaturverzeichnis

- Adler S, Wagstyl K, Gunny R, Ronan L, Carmichael D, Cross JH, Fletcher PC, Baldeweg T (2017) Novel surface features for automated detection of focal cortical dysplasias in paediatric epilepsy. *NeuroImage Clin*, 14:18–27.
- Aich A (2021) Elastic Weight Consolidation (EWC): Nuts and bolts. arXiv:2105.04093. [Online im Internet] URL: <https://doi.org/10.48550/arXiv.2105.04093> [Stand: 30.12.2024, 16:30].
- Ashburner J, Friston KJ (2000) Voxel-based morphometry--the methods. *Neuroimage*, 11(6 Pt 1):805–821.
- Barkovich AJ, Kuzniecky RI (1996) Neuroimaging of focal malformations of cortical development. *J Clin Neurophysiol*, 13(6):481–494.
- Bernasconi A, Cendes F, Theodore WH, Gill RS, Koepp MJ, Hogan RE, Jackson GD, Federico P, Labate A, Vaudano AE, Blümcke I, Ryvlin P, Bernasconi N (2019) Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: A consensus report from the International League Against Epilepsy Neuroimaging Task Force. *Epilepsia*, 60(6):1054–1068.
- Bijay Dev KM, Jogi PS, Niyas S, Vinayagamani S, Kesavadas C, Rajan J (2019) Automatic detection and localization of Focal Cortical Dysplasia lesions in MRI using fully convolutional neural network. *Biomed Signal Process Control*, 52:218–225.
- Blumcke I, Spreafico R, Haaker G, Coras R, Kobow K, Bien CG, Pfäfflin M, Elger C, Widman G, Schramm J, Becker A, Braun KP, Leijten F, Baayen JC, Aronica E, Chassoux F, Hamer H, Stefan H, Rössler K, Thom M, Walker MC, Sisodiya SM, Duncan JS, McEvoy AW, Pieper T, Holthausen H, Kudernatsch M, Meencke HJ, Kahane P, Schulze-Bonhage A, Zentner J, Heiland DH, Urbach H, Steinhoff BJ, Bast T, Tassi L, Lo Russo G, Özkara C, Oz B, Krsek P, Vogelgesang S, Runge U, Lerche H, Weber Y, Honavar M, Pimentel J, Arzimanoglou A, Ulate-Campos A, Noachtar S, Hartl E, Schijns O, Guerrini R, Barba C, Jacques TS, Cross JH, Feucht M, Mühlebner A, Grunwald T, Trinka E, Winkler PA, Gil-Nagel A, Toledano Delgado R, Mayer T, Lutz M, Zountsas B, Garganis K, Rosenow F, Hermsen A, von Oertzen TJ, Diepgen TL, Avanzini G, EEBB Consortium (2017) Histopathological findings in brain tissue obtained during Epilepsy Surgery. *N Engl J Med*, 377(17):1648–1656.
- Chen C, Xie J-J, Ding F, Jiang Y-S, Jin B, Wang S, Ding Y, Li H, Jiang B, Zhu J-M, Ding M-P, Chen Z, Wu Z-Y, Zhang B-R, Hsu Y-C, Lai H-Y, Wang S (2021) 7T MRI with post-processing for the presurgical evaluation of pharmacoresistant focal epilepsy. *Ther Adv Neurol Disord*, 14:17562864211021181.
- Chen Z, Brodie MJ, Liew D, Kwan P (2018) Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: A 30-year longitudinal cohort study. *JAMA Neurol*, 75(3):279–286.
- Crome L (1957) Infantile cerebral gliosis with giant nerve cells. *J Neurol Neurosurg Psychiatry*, 20(2):117–124.
- David B, Kröll-Seger J, Schuch F, Wagner J, Wellmer J, Woermann F, Oehl B, Van Paesschen W, Breyer T, Becker A, Vatter H, Hattingen E, Urbach H, Weber B, Surges R, Elger CE, Huppertz H-J,

- Rüber T (2021) External validation of automated focal cortical dysplasia detection using morphometric analysis. *Epilepsia*, 62(4):1005–1021.
- Demerath T, Kaller CP, Heers M, Staack A, Schwarzwald R, Kober T, Reisert M, Schulze-Bonhage A, Huppertz H-J, Urbach H (2022) Fully automated detection of focal cortical dysplasia: Comparison of MPRAGE and MP2RAGE sequences. *Epilepsia*, 63(1):75–85.
- Feng C, Zhao H, Li Y, Wen J (2020) Automatic localization and segmentation of focal cortical dysplasia in FLAIR-negative patients using a convolutional neural network. *J Appl Clin Med Phys*, 21(9):215–226.
- Fisher RS, Acevedo C, Arzimanoglou A, Bogacz A, Cross JH, Elger CE, Engel J Jr, Forsgren L, French JA, Glynn M, Hesdorffer DC, Lee BI, Mathern GW, Moshé SL, Perucca E, Scheffer IE, Tomson T, Watanabe M, Wiebe S (2014) ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482.
- Fisher RS, van Emde Boas W, Blume W, Elger C, Genton P, Lee P, Engel J Jr (2005) Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–472.
- Ganji Z, Aghaee Hakak M, Zare H. (2022) Comparison of machine learning methods for the detection of focal cortical dysplasia lesions: decision tree, support vector machine and artificial neural network. *Neurol Res*, 44(12):1142–1149.
- Gill RS, Lee H-M, Caldaïrou B, Hong S-J, Barba C, Deleo F, D’Incerti L, Mendes Coelho VC, Lenge M, Semmelroch M, Schrader DV, Bartolomei F, Guye M, Schulze-Bonhage A, Urbach H, Cho KH, Cendes F, Guerrini R, Jackson G, Hogan RE, Bernasconi N, Bernasconi A (2021) Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. *Neurology*, 97(16).
- Goceri E (2023) Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev*:1–45.
- House PM, Kopelyan M, Braniewska N, Silski B, Chudzinska A, Holst B, Sauvigny T, Martens T, Stodieck S, Pelzl S (2021) Automated detection and segmentation of focal cortical dysplasias (FCDs) with artificial intelligence: Presentation of a novel convolutional neural network and its prospective clinical validation. *Epilepsy Res*, 172:106594.
- Huppertz H-J, Grimm C, Fauser S, Kassubek J, Mader I, Hochmuth A, Spreer J, Schulze-Bonhage A (2005) Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis. *Epilepsy Res*, 67(1–2):35–50.
- Huppertz H-J, Wellmer J, Staack AM, Altenmüller D-M, Urbach H, Kröll J (2008) Voxel-based 3D MRI analysis helps to detect subtle forms of subcortical band heterotopia. *Epilepsia*, 49(5):772–785.
- Jin B, Krishnan B, Adler S, Wagstyl K, Hu W, Jones S, Najm I, Alexopoulos A, Zhang K, Zhang J, Ding M, Wang S, the Pediatric Imaging, Neurocognition, and Genetics Study, Wang ZI (2018) Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia*, 59(5):982–992.

- Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, Milan K, Quan J, Ramalho T, Grabska-Barwinska A, Hassabis D, Clopath C, Kumaran D, Hadsell R (2017) Overcoming catastrophic forgetting in neural networks. *Proc Natl Acad Sci USA*, 114(13):3521–3526.
- Kwan P, Arzimanoglou A, Berg AT, Brodie MJ, Allen Hauser W, Mathern G, Moshé SL, Perucca E, Wiebe S, French J (2009) Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia*, 51(6):1069–1077.
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature*, 521(7553):436–444.
- Mao J, Yin X, Zhang G, Chen B, Chang Y, Chen W, Yu J, Wang Y (2022) Pseudo-labeling generative adversarial networks for medical image classification. *Comput Biol Med*, 147(105729):105729.
- Marcelis S, Vanden Bossche S, Dekeyser S (2022) Not your everyday FCD: Imaging findings of focal cortical dysplasia type 1. *J Belg Soc Radiol*, 106(1):39.
- McCloskey M, Cohen NJ (1989) Catastrophic interference in connectionist networks: The sequential learning problem. In: *Psychology of Learning and Motivation*. Elsevier, S. 109–165.
- Mo J-J, Zhang J-G, Li W-L, Chen C, Zhou N-J, Hu W-H, Zhang C, Wang Y, Wang X, Liu C, Zhao B-T, Zhou J-J, Zhang K (2019) Clinical value of machine learning in the automated detection of focal cortical dysplasia using quantitative multimodal surface-based features. *Front Neurosci*, 12.
- Myronenko A (2018) 3D MRI brain tumor segmentation using autoencoder regularization. *arXiv:1810.11654*. [Online in Internet] URL: <https://doi.org/10.48550/arXiv.1810.11654> [Stand: 30.12.2024, 16:30].
- Najm I, Lal D, Alonso Vanegas M, Cendes F, Lopes-Cendes I, Palmini A, Paglioli E, Sarnat HB, Walsh CA, Wiebe S, Aronica E, Baulac S, Coras R, Kobow K, Cross JH, Garbelli R, Holthausen H, Rössler K, Thom M, El-Osta A, Lee JH, Miyata H, Guerrini R, Piao Y-S, Zhou D, Blümcke I (2022) The ILAE consensus classification of focal cortical dysplasia: An update proposed by an ad hoc task force of the ILAE diagnostic methods commission. *Epilepsia*, 63(8):1899–1919.
- Palmini A, Najm I, Avanzini G, Babb T, Guerrini R, Foldvary-Schaefer N, Jackson G, Lüders HO, Prayson R, Spreafico R, Vinters HV (2004) Terminology and classification of the cortical dysplasias. *Neurology*, 62(6 Suppl 3):S2-8.
- Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Köpf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S (2019) PyTorch: An imperative style, high-performance deep learning library. *arXiv:1912.01703*. [Online in Internet] URL: <https://doi.org/10.48550/arXiv.1912.01703> [Stand: 30.12.2024, 16:30].
- Ridsdale L, Wojewodka G, Robinson E, Landau S, Noble A, Taylor S, Richardson M, Baker G, Goldstein LH, SMILE Team (2017) Characteristics associated with quality of life among people with drug-resistant epilepsy. *J Neurol*, 264(6):1174–1184.
- Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Lecture Notes in Computer Science*. Springer International Publishing, Cham, S. 234–241.

- Spencer SS, Berg AT, Vickrey BG, Sperling MR, Bazil CW, Shinnar S, Langfitt JT, Walczak TS, Pacia SV, Multicenter Study of Epilepsy Surgery (2005) Predicting long-term seizure outcome after resective epilepsy surgery: the multicenter study. *Neurology*, 65(6):912–918.
- Spitzer H, Ripart M, Whitaker K, D’Arco F, Mankad K, Chen AA, Napolitano A, De Palma L, De Benedictis A, Foldes S, Humphreys Z, Zhang K, Hu W, Mo J, Likeman M, Davies S, Güttler C, Lenge M, Cohen NT, Tang Y, Wang S, Chari A, Tisdall M, Bargallo N, Conde-Blanco E, Pariente JC, Pascual-Díaz S, Delgado-Martínez I, Pérez-Enríquez C, Lagorio I, Abela E, Mullatti N, O’Muircheartaigh J, Vecchiato K, Liu Y, Caligiuri ME, Sinclair B, Vivash L, Willard A, Kandasamy J, McLellan A, Sokol D, Semmelroch M, Kloster AG, Opheim G, Ribeiro L, Yasuda C, Rossi-Espagnet C, Hamandi K, Tietze A, Barba C, Guerrini R, Gaillard WD, You X, Wang I, González-Ortiz S, Severino M, Striano P, Tortora D, Kälviäinen R, Gambardella A, Labate A, Desmond P, Lui E, O’Brien T, Shetty J, Jackson G, Duncan JS, Winston GP, Pinborg LH, Cendes F, Theis FJ, Shinohara RT, Cross JH, Baldeweg T, Adler S, Wagstyl K (2022) Interpretable surface-based detection of focal cortical dysplasias: a Multi-centre Epilepsy Lesion Detection study. *Brain*, 145(11):3859–3871.
- Sun K, Yu T, Yang D, Ren Z, Qiao L, Ni D, Wang X, Zhao Y, Chen X, Xiang J, Chen N, Gao R, Yang K, Lin Y, Kober T, Zhang G (2021) Fluid and white matter suppression imaging and voxel-based morphometric analysis in conventional magnetic resonance imaging-negative epilepsy. *Front Neurol*, 12:651592.
- Taylor DC, Falconer MA, Bruton CJ, Corsellis JA (1971) Focal dysplasia of the cerebral cortex in epilepsy. *J Neurol Neurosurg Psychiatry*, 34(4):369–387.
- Thurman DJ, Logroscino G, Beghi E, Hauser WA, Hesdorffer DC, Newton CR, Scorza FA, Sander JW, Tomson T, Epidemiology Commission of the International League Against Epilepsy (2017) The burden of premature mortality of epilepsy in high-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia*, 58(1):17–26.
- Tsigebrhan R, Derese A, Kariuki SM, Fekadu A, Medhin G, Newton CR, Prince MJ, Hanlon C (2023) Co-morbid mental health conditions in people with epilepsy and association with quality of life in low- and middle-income countries: a systematic review and meta-analysis. *Health Qual Life Outcomes*, 21(1):5.
- Urbach H, Kellner E, Kremers N, Blümcke I, Demerath T (2022) MRI of focal cortical dysplasia. *Neuroradiology*, 64(3):443–452.
- Varoquaux G (2018) Cross-validation failure: Small sample sizes lead to large error bars. *Neuroimage*, 180(Pt A):68–77.
- Walger L, Adler S, Wagstyl K, Henschel L, David B, Borger V, Hattingen E, Vatter H, Elger CE, Baldeweg T, Rosenow F, Urbach H, Becker A, Radbruch A, Surges R, Reuter M, Cendes F, Wang ZI, Huppertz H-J, Rüber T (2023) Artificial intelligence for the detection of focal cortical dysplasia: Challenges in translating algorithms into clinical practice. *Epilepsia*, 64(5):1093–1112.
- Wang H, Ahmed SN, Mandal M (2020) Automated detection of focal cortical dysplasia using a deep convolutional neural network. *Comput Med Imaging Graph*, 79(101662):101662.
- Wang I, Alexopoulos A (2016) MRI postprocessing in presurgical evaluation. *Curr Opin Neurol*, 29(2):168–174.

- Wang L, Zhang X, Su H, Zhu J (2024) A comprehensive survey of continual learning: Theory, method and application. *IEEE Trans Pattern Anal Mach Intell*, 46(8):5362–5383.
- Wellmer J, Parpaley Y, von Lehe M, Huppertz H-J (2010) Integrating magnetic resonance imaging postprocessing results into neuronavigation for electrode implantation and resection of subtle focal cortical dysplasia in previously cryptogenic epilepsy. *Neurosurgery*, 66(1):187–94; discussion 194-5.
- Wiebe S (2006) Burden of intractable epilepsy. *Adv Neurol*, 97:1–4.
- Willard A, Antonic-Baker A, Chen Z, O’Brien TJ, Kwan P, Perucca P (2022) Seizure outcome after surgery for MRI-diagnosed focal cortical dysplasia: A systematic review and meta-analysis. *Neurology*, 98(3):e236–e248.
- World Health Organization (2019) Epilepsy: a public health imperative. [Online im Internet] URL: <https://iris.who.int/handle/10665/325293> [Stand: 30.12.2024, 16:30].
- Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. *Insights Imaging*, 9(4):611–629.
- Zheng R, Chen R, Chen C, Yang Y, Ge Y, Ye L, Miao P, Jin B, Li H, Zhu J, Wang S, Huang K (2024) Automated detection of focal cortical dysplasia based on magnetic resonance imaging and positron emission tomography. *Seizure*, 117:126–132.

5. Abkürzungsverzeichnis

ASM	Anfallssupprimierende Medikamente
AUC	Area Under the Curve
CLOND	„Continual learning with old model and new dataset“ ist ein CNN-Modell, welches mittels einer kontinuierlichen Lernstrategie trainiert wurde
CNN	Convolutional Neural Network
DNET	Dysembryoplastischer Neuroepithelialer Tumor
DVA	Developmental Venous Anomaly
EWC	Elastic Weight Consolidation
FCD	Fokale kortikale Dysplasie
FN	False Negative
FP	False Positive
GANs	Generative Adversarial Networks
GTM	Ground-Truth-Mask
HS	Hippocampussklerose
ILAE	International League Against Epilepsy
KI	Konfidenzintervall
MAP18	Morphometric Analysis Program v2018
MCD	Malformations of cortical development
MNI	Montreal Neurological Institute
MOGHE	Mild malformation of cortical development with oligodendroglial hyperplasia in epilepsy
MRT	Magnetresonanztomographie
OWD	„Old model with whole dataset“ ist ein CNN-Modell, welches mittels klassischer Lernstrategie trainiert wurde
FDG-PET	18F-Fluorodeoxyglucose Positron-Emissions-Tomography
PMG	Polymikrogyrie
ResBlock	Residual Block
ROC	Receiver Operating Characteristic
SD	Standardabweichung
SPM	Statistical Parametric Mapping
SUDEP	Sudden Unexpected Death in Epilepsy
TLE	Temporallappenepilepsie
TN	True Negative
TP	True Positive

6. Abbildungsverzeichnis

	<u>Seite</u>
Abbildung 1: Charakteristische MRT-Merkmale von fokalen kortikalen Dysplasien	4
Abbildung 2: Überblick über die Funktionsweise von MAP18	9
Abbildung 3: Überblick über die allgemeine Funktionsweise und Architektur eines CNN	9
Abbildung 4: Netzwerkarchitektur	11
Abbildung 5: Volumetrische FCD-Läsionskarten	12
Abbildung 6: Sequenzielles Lernen und Elastic Weight Consolidation	13
Abbildung 7: Leistungsmetriken der CNN-Modelle auf Fallebene mit 95%-Wald-Konfidenzintervallen	17
Abbildung 8: Anzahl der falsch-positiven Cluster	18
Abbildung 9: Unterschiede in Sensitivität und Spezifität zwischen den Modellen mit 95 %-asymptotischen Konfidenzintervallen nach Tango	18
Abbildung 10: ROC-Kurven der drei Modelle in der finalen Auswertungsphase	19
 Figure 1: Neuronal network architecture	 30
Figure 2: Training and validation process	32
Figure 3: Performance metrics of the CNN models at the case level with 95% Wald confidence intervals	34
Figure 4: Average number of false positive clusters per MRI with standard deviation	35
Figure 5a: Segmentation results of all three CNN models (baseline model, OWD model, CLOND model) in the final evaluation phase on an exemplary FCD	36
Figure 5b: Segmentation results for a small FCD ($V = 0.51\text{ml}$) in the first evaluation phase	37
 Figure A.1: FCD sizes according to expert-labeled maps (in ml)	 39
Figure A.2: Differences in sensitivity and specificity between the three CNN models with 95% Tango's asymptotic confidence intervals	40
Figure A.3: Anatomical distribution of the false positive clusters	41
Figure A.4: Average false positive cluster sizes with standard deviation (in ml)	42
Figure A.5: Size distribution of false positive clusters	42
Figure A.6: FCD segmentation performance for each FCD (Dice coefficient)	43
Figure A.7: Receiver operating characteristic (ROC) curves for each CNN model	44

7. Tabellenverzeichnis

	<u>Seite</u>
Tabelle 1: Dreistufige ILAE-Klassifikation der fokalen kortikalen Dysplasien	4
Tabelle 2: Zusammensetzung der pathologischen Nicht-FCD-MRTs	7
Tabelle 3: Regionale Verteilung und anatomische Merkmale der FCD-Kohorte	7
Tabelle 4: Demografische Merkmale des Datensatzes	8
Tabelle 5: Dreiphasige Trainings- und Auswertungsstrategie	13
Tabelle 6: Methoden des kontinuierlichen Lernens	14
Tabelle 7: Auswertung der Modellleistung auf Voxel-, Läsions- und Fallebene	17
Tabelle 8: Gegenüberstellung aktueller Methoden zur automatisierten FCD-Detektion	24
Tabelle 9: Vergleich von MAP18- und histopathologischen Befunden bei der Diagnose kortikaler Fehlbildungen der neuronalen Entwicklung (Malformations of Cortical Development, MCD) am Hamburger Epilepsie-Zentrum (2014–2023)	25
Table 1: MRI scanner, sequences, and technical parameters	29
Table 2: Demographic characteristics of the dataset	29
Table 3: Results of the prospective validation at voxel, lesion, and case level	33
Table A.1: Anatomical characteristics of FCDs and focal PMG	44
Table A.2: Results of the subgroup analysis at the case and lesion level	45

8. Erklärung des Eigenanteils

Die Arbeit wurde in der Klinik für Neurochirurgie am Universitätsklinikum Hamburg-Eppendorf sowie im Epilepsie-Zentrum Hamburg unter der Leitung von Priv.-Doz. Dr. med. Thomas Sauvigny und der Betreuung von Dr. med. Patrick M. House durchgeführt.

Mein Eigenanteil an der Publikation umfasst die Entwicklung der Forschungsfrage und die Konzeption der Studie, die ich in enger Zusammenarbeit mit Dr. House erstellt habe. Ich war verantwortlich für die Durchführung und Auswertung der MAP18-Analyse, die anschließend durch Dr. House validiert wurde. Die Datenerhebung, die Erstellung der dreidimensionalen Läsionskarten, die Experimente sowie die statistische Auswertung der Daten wurden von mir eigenständig durchgeführt. Sowohl das Manuskript als auch sämtliche Grafiken und Tabellen habe ich in vollständiger Eigenverantwortung erstellt. Eine Ausnahme bildet „Figure 1“, die in Zusammenarbeit mit Natalia Braniewska entstanden ist. Während des Peer-Review-Prozesses koordinierte ich die Kommunikation mit den Co-Autorinnen und Co-Autoren, integrierte deren Feedback und überarbeitete das Manuskript entsprechend.

Die Beiträge der Co-Autorinnen und Co-Autoren umfassten folgende Bereiche:

- Dr. med. Patrick M. House war hauptverantwortlich für die Projektadministration sowie die Überprüfung und Überarbeitung des Manuskripts
- Agata Chudzinska, Bartosz Silski und Natalia Braniewska übernahmen die Programmierung und Softwareentwicklung
- Priv.-Doz. Dr. med. Thomas Sauvigny übernahm die Supervision des Projekts
- Dr. med. Brigitte Holst, Sirko Pelzl und Dr. med. Stefan Stodieck stellten die notwendigen Ressourcen bereit
- Natalia Braniewska, Priv.-Doz. Dr. med. Thomas Sauvigny und Dr. med. Brigitte Holst unterstützten zusätzlich die Überarbeitung des Manuskripts

Die vorliegende Einschätzung über die erbrachte Eigenleistung erfolgte einvernehmlich mit allen an dem Artikel beteiligten Co-Autorinnen und Co-Autoren.

9. Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Das gilt insbesondere auch für alle Informationen aus Internetquellen.

Soweit beim Verfassen der Dissertation KI-basierte Tools („Chatbots“) verwendet wurden, versichere ich ausdrücklich, den daraus generierten Anteil deutlich kenntlich gemacht zu haben. Die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“ aus September 2023 wurde dabei beachtet.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich damit einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Datum

Unterschrift

10. Danksagung

Mein besonderer Dank gilt meinem Betreuer Dr. med. Patrick M. House, der mir das Thema anvertraut und mich während des gesamten Projekts mit fachlicher Expertise, konstruktivem Feedback und stetiger Unterstützung begleitet hat. Seine hervorragende Betreuung war von unschätzbarem Wert. An dieser Stelle möchte ich mich auch für seine exzellente Ausbildung in der Epileptologie und der EEG-Befundung bedanken.

Ebenso danke ich Herrn Priv.-Doz. Dr. med. Thomas Sauvigny, Leiter der Epilepsiechirurgie am Universitätsklinikum Hamburg-Eppendorf, für die Möglichkeit, diese Arbeit unter seiner Leitung durchzuführen.

Ein herzlicher Dank geht an Agata Chudzinska, Bartosz Silski und Natalia Braniewska (theBlue.ai GmbH) für ihre engagierte Arbeit bei der Programmierung und Softwareentwicklung, die essenziell für den Erfolg dieser Studie war.

Mein Dank gilt allen externen Partnern und Institutionen, deren Unterstützung diese Arbeit ermöglicht hat, insbesondere Dr. med. Stefan Stodieck (Chefarzt, Epilepsie-Zentrum Hamburg), Dr. med. Brigitte Holst (Funktionsüberärztin, Klinik für Neuroradiologie, Universitätsklinikum Hamburg-Eppendorf) und Sirko Pelzl (CEO, theBlue.ai GmbH) für die Bereitstellung der notwendigen Ressourcen. Maria Stark und Gerhard Schön (Institut für Medizinische Biometrie und Epidemiologie, Universitätsklinikum Hamburg-Eppendorf) danke ich für die statistische Beratung. Angela McOrmond und Prof. Dr. Dr. h.c. Juliane House-Edmondson danke ich für das Korrekturlesen der englischsprachigen Publikation sowie Dr. Hermann Eisele für das Korrekturlesen der deutschsprachigen Publikationsdarstellung. Ebenso danke ich Prof. Dr. Mathias Trabs (Fakultät für Mathematik, Karlsruher Institut für Mathematik) für seine konstruktiven Anmerkungen.

Bei meiner Familie, meinen Freunden sowie meinen Kolleginnen und Kollegen bedanke ich mich herzlich für ihre Unterstützung und ihre Motivation während des Projekts.

Besonders bedanken möchte ich mich bei meiner Ehefrau Neha Chanra für ihren unermüdlichen Rückhalt und ihr Verständnis, da ich ihr während der Erstellung dieser Arbeit weniger Zeit widmen konnte.

Mein tiefster Dank gebührt meinen Eltern Kamlesh und Jagdish Chanra für ihre bedingungslose Liebe und ihren unerschütterlichen Glauben an mich. Diese Dissertation widme ich ihnen.

Nicht zuletzt danke ich allen Teilnehmerinnen und Teilnehmern der Studie, die diese Arbeit möglich gemacht haben.