Aus der Klinik und Poliklinik für Psychosomatik und Psychotherapeutische Medizin

Direktor: Prof. Dr. Dr. Wolfgang Schneider

Sektion für Gerotopsychosomatik und dementielle Erkrankungen

Leiter: Prof. Dr. Stefan Teipel

In Kooperation mit dem Deutschen Zentrum für Neurodegenerative Erkrankungen e.V.

Automatisierte Erkennung manifester und prodromaler Alzheimer-Krankheit mittels multizentrisch akquirierter, multimodaler Bilddaten

Kumulative Dissertationsschrift

zur

Erlangung des akademischen Grades

Doktor der Medizinwissenschaften

der Universitätsmedizin Rostock

vorgelegt von

Martin Dyrba, geb. am 5. August 1985 in Zwenkau

wohnhaft in Rostock

Rostock, den 18. Mai 2015

urn:nbn:de:gbv:28-diss2015-0235-7

Gutachter:

Prof. Dr. Stefan Teipel
Klinik und Polyklinik für Psychosomatik und Psychotherapeutische Medizin, Universitätsmedizin Rostock
und Deutsches Zentrum für Neurodegenerative Erkrankungen e.V., Standort Rostock/ Greifswald

Prof. Dr. Thomas Kirste
Lehrstuhl Mobile Multimediale Informationssysteme, Fakultät für Informatik und Elektrotechnik, Universität Rostock

Prof. Dr. Martin Lotze
Fachbereich Funktionelle Bildgebung, Institut für Diagnostische Radiologie und Neuroradiologie, Universitätsmedizin Greifswald

Eingereicht am 18. Mai 2015
Verteidigt am 16. Dezember 2015

Inhaltsverzeichnis

Zusammenfassung 1
1 Einleitung 3
1.1 Klinische und pathologische Charakterisierung der Alzheimer-Krankheit 3
1.2 Bedeutung der neuronalen Konnektivität bei der Alzheimer-Krankheit 4
1.3 Bildgebung bei der Alzheimer-Krankheit 5
1.4 Bedeutung multizentrischer Stichproben für die Untersuchung der Alzheimer-Krankheit 7
1.5 Die European DTI Study on Dementia 8
1.6 Maschinelle Lernverfahren zur Analyse multimodaler multizentrischer Daten 9
1.7 Untersuchungsziele und Hypothesen 9
2 Methodik 11
2.1 Studienkohorten 11
2.2 Bilddatenvorverarbeitung 12
2.3 Verarbeitung mittels maschineller Lernverfahren 15
2.3.1 Validierungstechniken 15
2.3.2 Merkmalsextraktion 17
2.3.3 Varianzreduktion 17
2.3.4 Merkmalsauswahl 20
2.3.5 Lernverfahren 21
2.3.6 Evaluation 23
2.3.7 Permutationstests 24
2.3.8 Visualisierung der informativsten Merkmale 25
3 Ergebnisse 27
3.1 Umsetzung der Datenverarbeitung 27
3.2 Leistung der Lernverfahren 27
4 Diskussion 35
4.1 Diskussion der Bilddatenvorverarbeitung und Datenaufbereitung 35
4.2 Bewertung der Leistung der Lernverfahren 35
4.3 Beurteilung der Varianzreduktion 36
4.4 Ausblick 37
5 Publikationen 39

6 Literaturverzeichnis 41
Selbstständigkeitserklärung 49
Lebenslauf und wissenschaftlicher Werdegang 51

Abkürzungsverzeichnis

AD	Alzheimer-Krankheit
AUC	Fläche unter der Kurve (Area under the receiver operating characteristic curve)
CSF	Cerebrospinalflüssigkeit
CT	Computertomographie
DMN	Default Mode Network
DTI	Diffusions-Tensor-Bildgebung
EDSD	European DTI Study on Dementia
FA	Fraktionelle Anisotropie
FDG-PET	Fluordesoxyglukose Positronen-Emissions-Tomographie
fMRT	Funktionelle Magnetresonanztomographie
GM	Graue Substanz
MCI	Leichte kognitive Störung
MD	Mittlerer Diffusionskoeffizient
MK-SVM	Multikernel Support Vector Machine
MO	Modus der Anisotropie
MRT	Magnetresonanztomographie
NB	Naive Bayes
PCA	Hauptkomponentenanalyse
PET	Positronen-Emissions-Tomographie
Ruhe-fMRT	Funktionelle MRT im Ruhezustand
SVM	Support Vector Machine
SWPL	Kürzeste gewichtete Pfadlänge
WEKA	Waikato Environment for Knowledge Analysis
WLCC	Gewichteter lokaler Clustering-Koeffizient
WM	Weiße Substanz

Abbildungsverzeichnis

Tabelle 1. Demographische Charakterisierung der Studienkohorten 12

Abbildung 1. Akquisition und Analyse mittels Diffusions-Tensor-Bildgebung..... 6
Abbildung 2. Charakterisierung von neuen Biomarkern mit dem Ziel der Implementierung in die klinische Praxis..... 8
Abbildung 3. Teilnehmende Zentren der European DTI Study on Dementia 8
Abbildung 4. Vorverarbeitungsschritte zur räumlichen Normalisierung der Bilddaten..... 13
Abbildung 5. Illustration der Datenverarbeitung mittels maschineller Lernverfahren 16
Abbildung 6. Berechnung der Kennwerte *gewichteter lokaler Clustering-Koeffizient* (WLCC) und *kürzeste gewichtete Pfadlänge* (SWPL) aus der funktionellen Konnektivitätsmatrix. 18
Abbildung 7. Illustration des Support Vector Machine (SVM) Verfahrens..... 22
Abbildung 8. Bestimmung der *Fläche unter der Kurve* (AUC) durch Variation der Entscheidungsebene des SVM-Modells 24
Abbildung 9. Erkennungsraten für die Unterscheidung von Alzheimer-Patienten und Kontrollprobanden 30
Abbildung 10. Erkennungsraten für die Unterscheidung von Alzheimer-Patienten und Kontrollprobanden (Scanner-basierte Kreuzvalidierung) 30
Abbildung 11. Erkennungsraten für die Unterscheidung von Patienten mit leichter kognitiver Störung mit positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor (MCI-Aβ42^{+}) und gesunden Kontrollprobanden 31
Abbildung 12. Erkennungsraten für die Unterscheidung von Patienten mit leichter kognitiver Störung mit negativem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor (MCI-Aβ42^{-}) von Patienten mit positivem Befund (MCI-Aβ42^{+})..... 31
Abbildung 13. Erkennungsleistung für die Unterscheidung von Alzheimer-Patienten und gesunden Kontrollprobanden (monozentrisch) 32
Abbildung 14. SVM Sensitivitätskarten für die von Patienten mit Alzheimer-Demenz und Kontrollprobanden (monozentrisch)..... 32
Abbildung 15. SVM Sensitivitätskarten für die fraktionelle Anisotropie und die mittlere Diffusivität..... 33

Danksagung

Eine wissenschaftliche Arbeit ist nie das Werk einer einzelnen Person. Deshalb möchte ich mich an dieser Stelle bei den zahlreichen Unterstützern bedanken.

Zuerst gilt mein Dank meinen beiden Doktorvätern, Herrn Prof. Dr. Stefan Teipel und Herrn Prof. Dr. Thomas Kirste, für ihre sachkundige und wertvolle Unterstützung meiner Arbeit, ihren kritischen Hinweisen, den konstruktiven Diskussionen und für ihre Geduld. Weiterhin gilt mein Dank meinen Kollegen und Freunden von der Universitätsmedizin, aus dem DZNE und dem Institut für Informatik, ohne deren Rat und Tat, Motivation und Hilfestellung diese Arbeit nicht möglich gewesen wäre. Zuletzt gilt mein größter Dank meiner Familie, meiner Frau Anica und meiner Tochter Elisabeth, für ihre Unterstützung und ihren Beistand, ihre Liebe und Fürsorge in allen Höhen und Tiefen.

Zusammenfassung

Alzheimer-Patienten zeigen bereits früh im Verlauf der Krankheit Veränderungen der strukturellen Integrität der weißen Substanz. Die Diffusions-Tensor-Bildgebung (DTI) erlaubt die Beurteilung der Fasertraktintegrität *in vivo*, wobei die multizentrische Erhebung und Anwendung dieser Bildgebungstechnik als diagnostischer Marker für die Alzheimer-Krankheit bisher noch nicht systematisch untersucht wurden. Multivariate maschinelle Lernverfahren, bei denen mehrere Merkmale gemeinsam ausgewertet werden, zeigen sich sensibel für verteilte Intensitätsveränderungen und wurden bisher erfolgreich in zahlreichen monozentrischen Bildgebungsstudien eingesetzt.

Ziele dieser Dissertation waren: erstens, die Charakterisierung multizentrischer DTI als Bildgebungsmarker für die Alzheimer-Krankheit im Demenzstadium sowie im Prodromalstadium der leichten kognitiven Störung. Zweitens, die Untersuchung geeigneter Verfahren zur Reduktion von Störeffekten bedingt durch die multizentrische Akquisition, um damit die Trennleistung des Markers zu verbessern. Und drittens, die multimodale Auswertung mehrerer Bildgebungstechniken zur Evaluation, ob sich dadurch die Erkennungsleistung weiter steigern lässt.

Im Rahmen dieser Dissertation wurden zwei Ansätze zur Verarbeitung von medizinischen Bilddaten mit multivariaten maschinellen Lernverfahren erarbeitet und umgesetzt. Die statistischen Verfahren Hauptkomponentenanalyse und lineare Regression wurden zur Modellierung und Reduktion der Störvarianz eingesetzt und deren Nutzen verglichen. Die Leistung der Verfahren wurde mit dem Kreuzvalidierungsverfahren ermittelt.

Das multivariate Lernverfahren Support Vector Machine erzielte bei der Erkennung einer manifesten Alzheimer-Demenz mit Hilfe von multizentrischen DTI-Daten eine Erkennungsrate von 85%, wobei dagegen mit Hilfe von anatomischer Magnetresonanztomographie eine signifikant höhere Erkennungsrate von 92% erzielt wurde. Die Erkennung von Patienten im Prodromalstadium der Alzheimer-Krankheit sowie die Abgrenzung von Patienten mit leichter kognitiver Störung und positivem Befund für Neurodestruktionsmarker im Liquor von Patienten mit negativem Befund waren am besten anhand von DTI-Daten möglich (77% bzw. 68% Erkennungsrate). Die Varianzreduktion mittels linearer Regression steigerten die Erkennungsraten um durchschnittlich 4% im Vergleich zur Verwendung der unkorrigierten Originaldaten. Demgegenüber steigerte die

Varianzreduktion basierend auf der Hauptkomponentenanalyse die Erkennungsraten um durchschnittlich 2%, wobei diese Methode eine deutlich höhere Variabilität der Ergebnisse aufwies. Die multimodale Auswertung von DTI, anatomischer Magnetresonanztomographie und funktioneller Magnetresonanztomographie im Ruhezustand erzielte keine Verbesserung der Erkennungsleistung im Vergleich zur jeweils besten Einzelmodalität.

Die DTI erwies sich als vielversprechender Bildgebungsmarker für die Alzheimer-Krankheit, der insbesondere als diagnostisches Instrument zur Frühdiagnose im Prodromalstadium der Krankheit geeignet wäre oder als Kennwert in Interventionsstudien genutzt werden könnte. Folgestudien könnten Assoziationen bzw. die Dynamik zwischen verschiedenen Bildgebungsmarkern untersuchen, um weitere Erkenntnisse über die bei der Alzheimer-Krankheit ablaufenden neurobiologischen Prozesse zu erhalten.

1 Einleitung

1.1 Klinische und pathologische Charakterisierung der Alzheimer-Krankheit

Die Alzheimer-Krankheit ist eine Erkrankung des Gehirns mit einem schleichenden, aber stetig fortschreitenden Degenerationsprozess [1 S.204]. Sie ist die häufigste neurologische Ursache für eine Demenz bei Personen über 60 Jahren [2, 3] und entwickelt sich über mehrere Jahre, wobei bereits einige Jahre vor den ersten klinischen Anzeichen charakteristische neuropathologische und neurochemische Veränderungen nachweisbar sind [3, 4]. Im Verlauf der Krankheit kommt es zunehmend zu einer Störung höherer kognitiver Funktionen, die unter anderem Gedächtnisleistung, Orientierung, Denken, Auffassung und Lernfähigkeit umfassen [3]. Ebenso treten Veränderungen im Sozialverhalten und in der emotionalen Kontrolle auf [3]. Klinisch wird die Alzheimer-Krankheit in drei Phasen eingeteilt [3]: 1. eine viele Jahre dauernde *präklinische Phase*, die rein biologisch definiert ist; 2. ein ca. zwei Jahre dauerndes klinisches Vorstadium mit manifester, aber noch prädemenzieller Symptomatik, *leichte kognitive Störung* bzw. *mild cognitive impairment*; und 3. das Stadium der manifesten *Alzheimer-Demenz*.

Histopathologisch ist die Alzheimer-Krankheit durch sogenannte „senile Plaques“ gekennzeichnet, extrazelluläre Ablagerungen von aggregierten Amyloid-β-Peptiden [3, 4]. Des Weiteren ist die Krankheit durch intraneuronale fibrilläre Aggregate des τ-Proteins charakterisiert, die als „neurofibrilläre Bündel“ bezeichnet werden [3, 4]. Die Ätiologie der Erkrankung ist derzeit noch nicht abschließend geklärt [3]. Das am weitesten diskutierte Modell für die Pathogenese der Krankheit ist die Amyloid-Kaskaden-Hypothese [5, 6]: Demnach sei die übermäßige Bildung von Amyloid-β durch die enzymatische Spaltung des Amyloid-Precursor-Proteins durch die β- und γ-Sekretase sowie die Aggregation des Amyloid-β in Form kleinerer Oligomere und die extrazelluläre Ablagerung größerer Plaques der primäre ätiologische Faktor der Alzheimer-Krankheit. Das Amyloid-β wirkt neurotoxisch durch Induzierung von oxidativem Stress, erhöhtem Einstrom von Calcium-Ionen und Schädigung der Mitochondrien [7]. Ferner führen diese Effekte zu Veränderungen der Phosphorkinaseaktivität und zur Hyperphosphorylierung des τ-Proteins [8]. In Folge dessen würden sich die τ-Proteine von den Mikrotuboli ablösen und als neurofibrilläre Bündel aggregieren [4]. Letztendlich führe dies zum Zusammenbruch des Cytoskeletts, zur Störung des Zellmetabolismus und final zum Zelltod des Neurons [4].

Als primärer biochemischer Marker für die Alzheimer-Diagnostik wird derzeit die Liquordiagnostik verwendet [9]. Schon mehrere Jahre vor den ersten Symptomen einer kognitiven Beeinträchtigung sind in der Cerebrospinalflüssigkeit für die Alzheimer-Erkrankung typische Veränderungen nachweisbar: Dazu zählen die Reduktion von Amyloid-β_{1-42}, die Erhöhung von hyperphosphoryliertem τ-Protein und Gesamt-τ-Protein sowie die Erhöhung des Quotienten aus Amyloid-β_{1-42} und Amyloid-β_{1-40} [9]. Mit einer Sensitivität und Spezifität von über 80% kann mit diesen Markern bereits im Prodromalstadium der leichten kognitiven Störung die Hypothese der Alzheimer-Krankheit als Ursache bestärkt oder wiederlegt werden [10, 11].

Die Amyloid-Kaskaden-Hypothese ist in den letzten Jahren stark in die Kritik geraten, insbesondere nachdem pharmakologische Studien mit dem Fokus auf der Bereinigung von Amyloid-Ablagerungen keinen signifikanten Effekt bezüglich der Kognition sowie der τ-Pathologie zeigten [5, 6]. Diesbezüglich werden derzeit ergänzende oder alternative Modelle diskutiert, die eine Vielzahl primär und sekundär relevanter Faktoren bezüglich der Alzheimer-Erkrankung beinhalten [12, 13].

1.2 Bedeutung der neuronalen Konnektivität bei der Alzheimer-Krankheit

Die Ausbreitung seniler Plaques und neurofibrillärer Bündel und die damit einhergehende neuronale Schädigung sind im Kortex nicht gleichverteilt, sondern regional verschieden. Braak und Braak begründeten ein Stufenmodell für die Ausbreitung von senile Plaques und neurofibrillärer Bündel [14], das z.B. durch Delacourte et al. erweitert wurde [15]. Die Ausbreitung der neurofibrillären Bündel und die damit assoziierte neuronale Schädigung beginnt in entorhinalen Regionen, gefolgt von limbischen Arealen des medialen Temporallappens und schließlich neokortikalen Assoziationsarealen [14, 15]. Die sensomotorischen Areale sowie das Cerebellum sind erst in den Endstadien der Erkrankung betroffen. Die ersten Kognitionsstörungen treten auf, wenn sich die Degeneration im limbischen Allokortex ausgebreitet hat, während eine klinisch manifeste Demenz schließlich mit der Schädigung neokortikaler Gebiete einhergeht [15].

Pathologische Befunde zeigen, dass bereits sehr früh im Verlauf der Alzheimer-Krankheit die Fasertrakte der weißen Substanz geschädigt werden [16]. Die Schädigung der Fasertrakte der weißen Substanz wird dabei als frühere Veränderung bei der Alzheimer-Krankheit

angenommen, während die zerebrale Atrophie als Ausdruck des Untergangs einer Vielzahl von Nervenzellen erst später im Verlauf der Krankheit einsetzt [17].

1.3 Bildgebung bei der Alzheimer-Krankheit

Die bildgebenden Verfahren, vor allem die Magnetresonanztomographie (MRT) und Computertomographie (CT), haben im Kontext der Alzheimer-Krankheit eine besondere Bedeutung. Die T_1-gewichtete MRT ermöglichen es, die Hirnstruktur bzw. eine Reduktion des Hirnvolumens (Atrophie) in vivo zu beurteilen [18]. Die Diffusions-Tensor-Bildgebung (DTI) dagegen misst die Bewegungsfreiheit (=Diffusion) von Wassermolekülen im Gewebe. Diese wird anhand eines Ellipsoidmodells repräsentiert, das durch die Richtung und Größe der drei Hauptachsen charakterisiert ist (Abbildung 1). Da die Diffusionsfähigkeit und Diffusionsrichtung von der Mikrostruktur des Gewebes abhängt, z.B. von Zellmembranen oder Mikrotuboli, lassen sich mittels DTI Rückschlüsse auf die Integrität der Fasertrakte der weißen Substanz ziehen [19]. Die Fluorodesoxyglukose-Positronen-Emissions-Tomographie (FDG-PET) [20] bemisst den Glukosemetabolismus. Patienten mit einer prodromalen Alzheimer-Erkrankung oder einer Alzheimer-Demenz zeigen einen regional spezifischen Hypometabolismus [21, 22]. Weiterhin wurden für die Alzheimer-Krankheit spezielle Tracer-Substanzen entwickelt, die mittels Positronen-Emissions-Tomographie (PET) das Vorhandensein und die lokale Ausbreitung von Amyloid-β abbilden können [23]. Eine weitere Bildgebungstechnik ist die funktionelle Magnetresonanztomographie (fMRT), bei der die unterschiedlichen magnetischen Eigenschaften von oxygeniertem und desoxygeniertem Blut genutzt werden, um über den Mechanismus der neurovaskulären Kopplung Rückschlüsse auf die neuronale Aktivität in den verschiedenen Hirnregionen zu ziehen [24, 25]. Bei verschiedenen kognitiven Aufgaben zeigen sich verteilt spezifische Hirnregionen aktiv, die ein synchrones Zeitsignal aufweisen. Diese werden deshalb als *intrinsische funktionelle Konnektivitätsnetzwerke* bezeichnet [26, 27]. Die bei der funktionellen Magnetresonanztomographie im Ruhezustand des Probanden (Ruhe-fMRT) synchron aktiven Regionen, darunter Cluster im posterioren Gyrus cinguli und Precuneus, im medialen Frontallappen und beidseitig in den inferioren Parietallappen, werden *Default Mode Network* (DMN) genannt [28, 29]. Im Verlauf der Alzheimer-Krankheit verändert sich die Synchronizität bzw. die funktionelle Konnektivität dieses Netzwerkes [30-37]. Methodisch werden diese Veränderungen häufig mithilfe von Graphenmodellen untersucht [36, 38-43].

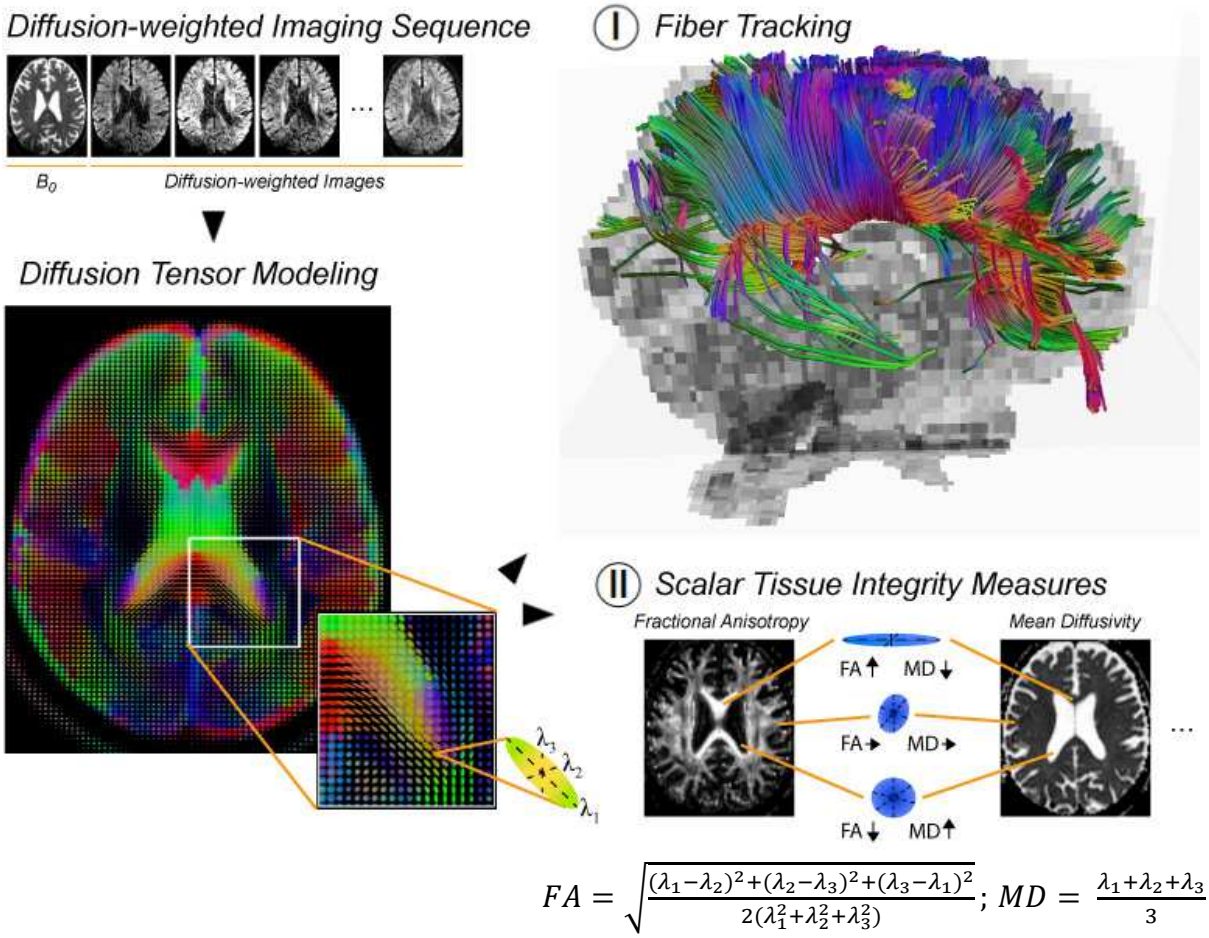

Abbildung 1. Akquisition und Analyse mittels Diffusions-Tensor-Bildgebung

Die diffusionsgewichtete Akquisitionssequenz besteht aus einer diffusionsungewichteten Aufnahme (sog. B_0-Bild) und mehreren diffusionsgewichteten Aufnahmen, die die Bewegungsfreiheit der Wassermoleküle in verschiedene Raumrichtungen erfassen (oben links). Aus diesen Daten wird anhand des Tensor-Modells die Diffusionsfähigkeit berechnet, repräsentiert als Ellipsoid, der durch die drei Hauptachsen λ_1, λ_2, λ_3 charakterisiert ist (unten links). Die Länge der Hauptachsen λ_1, λ_2, λ_3 entspricht hierbei der Bewegungsfreiheit der Wassermoleküle entlang der jeweiligen Richtung. Beim sogenannten Fiber-Tracking (I) nutzen Algorithmen die Form und Richtung des Ellipsoids, um den wahrscheinlichsten Verlauf von Nervenfaserbahnen zu rekonstruieren. Skalare Messwerte für die Gewebeintegrität (II), z.B. die fraktionelle Anisotropie (FA) und mittlere Diffusivität (MD), nutzen dagegen rein die Form des Ellipsoids. Je nach Gewebetyp weisen diese Kennwerte charakteristische Größen auf: Bei vielen, parallel orientierten Nervenfasern, z.B. im Corpus callosum, ist der Ellipsoid eher zigarrenförmig, so dass die FA ihre höchsten Werte erreicht bzw. die MD ihre niedrigsten. Dem entgegen ist im Liquor die Diffusion in alle Richtungen gleich groß, so dass der Ellipsoid hier eher kugelförmig ist, FA klein bzw. MD groß. In der grauen Substanz oder in Bereichen mit kreuzenden Fasern zeigen beide Kennwerte eher mittelgroße Werte, die stark von der jeweiligen Konfiguration der einzelnen Hauptachsen des Ellipsoids abhängen (unten links).

Während für die Bildgebungstechniken MRT, FDG-PET und Amyloid-PET die zeitlichen und regionalen Schädigungs- bzw. Ablagerungsmuster im Verlauf der Alzheimer-Krankheit in zahlreichen Studien charakterisiert wurden [17, 44], waren zu Beginn dieser Promotion (2011) vergleichbare, multizentrische Ergebnisse für die DTI noch nicht verfügbar. Da insbesondere durch Vorstudien bekannt war, dass multizentrische DTI-Daten im Vergleich zu T_1-gewichteten MRT-Daten eine etwa 50% höhere Variabilität aufweisen [45, 46], ist die Untersuchung der Eignung multizentrischer DTI als Bildgebungsmarker für die Alzheimer-Krankheit von besonderer Relevanz.

1.4 Bedeutung multizentrischer Stichproben für die Untersuchung der Alzheimer-Krankheit

Die Charakterisierung von neuen Biomarkern bzw. Bildgebungsmarkern mit dem Ziel der Implementierung in die klinische Praxis kann in drei Schritte gegliedert werden (Abbildung 2) [47]:

1. Die Untersuchung der technischen Eigenschaften einer Methodik bzw. eines Verfahrens.
2. Die Bestimmung des diagnostischen Wertes des Markers, ob und wie gut die Unterscheidung von Patienten mit einer bestimmten Krankheit von gesunden Kontrollprobanden anhand einer monozentrischen Stichprobe möglich ist. Dies erlaubt eine erste Abschätzung der Sensitivität und Spezifität.
3. a) Die Anwendung in **multizentrischen Kohorten**, zur Validierung der Ergebnisse vorheriger monozentrischer Studien, zur Identifikation von Konfundierungsfaktoren bei der multizentrischen Datenerhebung bzw. Evaluation des Einflusses auf die Zielgrößen. Diese Stufe beinhaltet ebenfalls die Anwendung eines Verfahrens in multizentrischen klinischen Studien.
3. b) Studien in repräsentativen Stichproben der Population zur Bewertung der Effektivität des Markers im realen Versorgungssystem. Dies umfasst die Translation des Verfahrens in die Regelversorgung, die Bewertung der Sensitivität und Spezifität des Markers anhand der populationsbasierten Stichprobe, sowie die Evaluierung, welche primären und sekundären Effekte dieser Marker auf das Versorgungssystem zeigt.

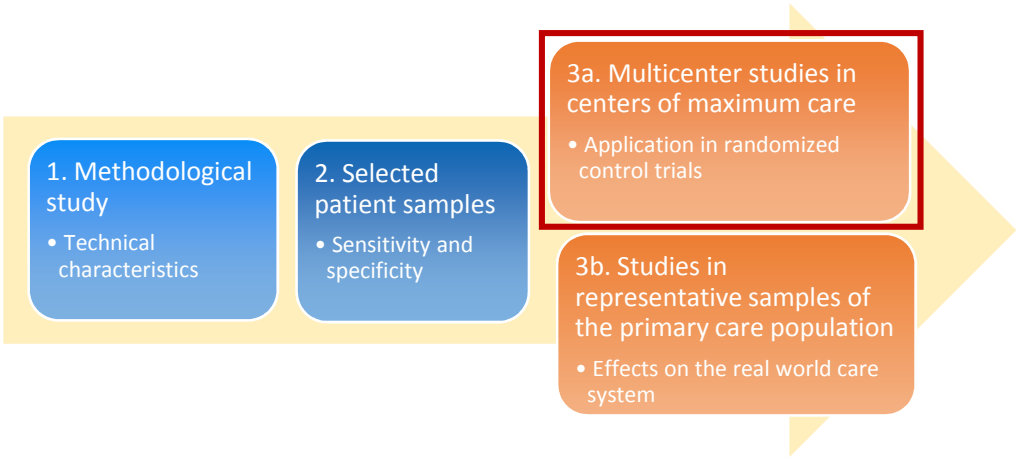

Abbildung 2. Charakterisierung von neuen Biomarkern mit dem Ziel der Implementierung in die klinische Praxis

1.5 Die European DTI Study on Dementia

Für die Untersuchung der Eignung und Stabilität der DTI als Bildgebungsmarker bei der Alzheimer-Krankheit wurde die *European DTI Study on Dementia* (EDSD) initiiert. Ziel der EDSD ist die Schaffung einer großen, multizentrischen Bilddatenbank, um den Einfluss der multizentrischen Datenakquisition im Vergleich zur monozentrischen Datenerhebung zu evaluieren oder eine geeignete Methoden zur Repräsentation und Reduktion von Zentrumseffekten zu untersuchen. Dazu wurde initial eine Studie mit einem physikalischen Phantom und einer gesunden Probandin durchgeführt, die an zwölf verschiedenen MRT-Scannern in Europa gescannt wurden [45]. Schrittweise wurden MRT- und DTI-Scans von Alzheimer-Patienten und gesunden Kontrollen [48] bzw. von Patienten mit leichter kognitiver Störung und gesunden Kontrollen [49] ergänzt. Die teilnehmenden Zentren (siehe Abbildung 3) haben dazu bereits vorhandene Daten beigetragen, so dass die Datenbank zügig einen ausreichend großen Datenbestand erreichte und zur Untersuchung der verschiedenen Fragestellungen genutzt werden konnte, z.B. die Bewertung des Einflusses der multizentrischen Akquisition auf automatisiertes Fiber-Tracking (siehe Abbildung 1 I auf S.6) [50] oder die Bewertung des Einflusses des Apolipoprotein E4 auf die Integrität der Fasertrakte der weißen Substanz [51].

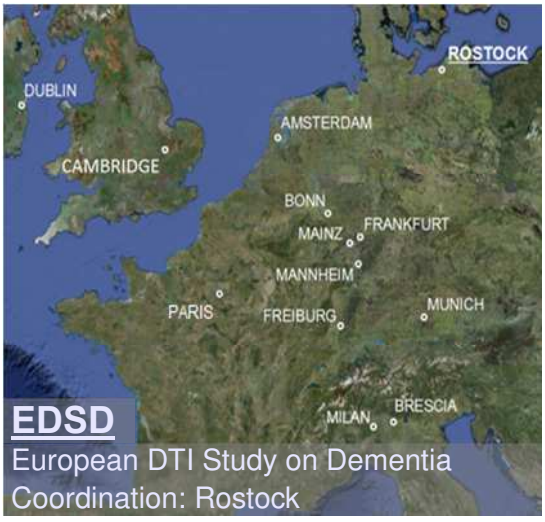

Abbildung 3. Teilnehmende Zentren der European DTI Study on Dementia

1.6 Maschinelle Lernverfahren zur Analyse multimodaler multizentrischer Daten

Maschinelle Lernverfahren sind in der Lage, automatisiert Muster in Daten zu identifizieren und in Form einer Abbildungsfunktion verschiedenen Werten zuzuordnen [52 S.16ff]. Besonders die multivariaten Verfahren, bei denen zahlreiche Merkmale kombiniert ausgewertet werden, zeigen sich sensibel für verteilte Gewebe- bzw. Intensitätsveränderungen, wie sie beispielsweise im Verlauf der Alzheimer-Krankheit auftreten. Damit erscheinen solche Verfahren gut für die Verwendung in diagnostischen Assistenzsystemen geeignet. Das multivariate Verfahren *Support Vector Machine* (SVM) [53] wurden in den vergangenen Jahren in zahlreichen monozentrischen Studien für die automatisierte Erkennung von Alzheimer-Demenz oder leichter kognitiver Störung mit Hilfe des Volumens der grauen Substanz verwendet [54-58] oder auf Basis von DTI-Kennwerten [59-64]. Dabei konnten Erkennungsraten von ca. 90% für die Gruppentrennung von Patienten mit Alzheimer-Demenz versus gesunden Kontrollprobanden erzielt werden [57, 58, 60, 64], 80-90% für die Trennung von Patienten mit leichter kognitiver Störung versus gesunden Probanden [57, 58, 62-64] sowie 70% für die Prädiktion der Konversion von leichter kognitiver Störung zur Alzheimer-Demenz versus Patienten, deren leichte kognitive Störung sich über den Beobachtungszeitraum von ca. 18 Monaten nicht veränderte [57, 58].
Zusätzlich zur Verwendung einzelner Bildgebungsmodalitäten wurden ebenfalls Studien unter Verwendung mehrerer Bildgebungstechniken durchgeführt. Die Verwendung von metabolischer und anatomischer Bildgebung mittels FDG-PET und T_1-gewichteter MRT zeigte eine leichte Verbesserung der Erkennungsrate um 3-6% Prozent für die Gruppentrennung von Patienten mit Alzheimer-Demenz oder leichter kognitiver Störung versus Kontrollprobanden gegenüber der Verwendung nur der T_1-gewichteten MRT [65-67]. Die kombinierte Verwendung funktioneller und struktureller Konnektivitätsmaße mittels Ruhe-fMRT und DTI steigerte die Gruppentrennung von Patienten mit leichter kognitiver Störung versus Kontrollprobanden um 7% im Vergleich zur Verwendung nur der DTI-Daten [62].

1.7 Untersuchungsziele und Hypothesen

Zur Untersuchung der automatisierten Erkennung der Alzheimer-Krankheit mittels multizentrischer, multimodaler Bilddaten wurden folgende Studienziele formuliert:

I. **Charakterisierung multizentrischer DTI-Daten als Biomarker für die Alzheimer-Krankheit mittels maschineller Lernverfahren (monomodal):**

Die Anwendung von multivariaten maschinellen Lernverfahren auf die multizentrisch akquirierten Bilddaten zur Bestimmung der Gruppentrennung. Dazu soll untersucht werden, inwiefern die Erkennungsrate für die Alzheimer-Krankheit mittels multizentrischer DTI-Daten in der im vorherigen Abschnitt für monozentrische Studien genannten Größenordnung erreicht werden können. Die Vorergebnisse der etwa 50% höheren Variabilität von multizentrischen DTI-Daten im Vergleich zu T_1-gewichteten MRT-Daten [45, 46] lassen eine deutlich reduzierte Erkennungsrate von multizentrischer DTI im Vergleich zu multizentrischer MRT oder monozentrischer DTI erwarten.

II. **Untersuchung und Anwendung von statistischen Verfahren zur Modellierung und Reduktion der Inter-Scanner-Variabilität:**

Die Bestätigung der oben genannten Hypothese würde es erforderlich machen, geeignete Verfahren für die Modellierung und Reduktion des Einflusses der Inter-Scanner-Variabilität sowie von demographischen Störvariablen (z.B. Alter) einzusetzen. Dazu sollen auf Basis theoretischer Überlegungen zur Verteilungsform der Störeffekte geeignete Verfahren ausgewählt, in die Datenaufbereitung integriert und deren Nutzen evaluiert werden. Die höhere Variabilität von multizentrischen DTI-Daten im Vergleich zu T_1-gewichteten MRT-Daten lässt eine signifikante Verbesserung der Erkennungsrate unter der Verwendung bereinigter multizentrischer DTI-Daten im Vergleich zur Verwendung unkorrigierter multizentrischer DTI-Daten erwarten.

III. **Untersuchung und Anwendung von maschinellen Lernverfahren unter Verwendung multimodal Bilddaten zur Erkennung der Alzheimer-Krankheit:**

Die in Abschnitt 1.6 (S.9) genannten Studien legen nahe, dass sich die Erkennungsrate bei der kombinierten Analyse mehrerer Bildgebungsmodalitäten im Vergleich zur Einzelanalyse verbessert. Dazu sollen geeignete Verfahren zur Integration mehrerer Bildgebungsmodalitäten ausgewählt und evaluiert werden. Es soll untersucht werden, ob die kombinierte Auswertung unter Zunahme anderer Bildgebungsverfahren den diagnostischen Nutzen der DTI signifikant steigert, insbesondere bei der Erkennung der Alzheimer-Krankheit im Prodromalstadium.

2 Methodik

2.1 Studienkohorten

Zur Untersuchung der oben genannten Fragestellungen wurden drei Studienkohorten genutzt:

Studie 1: In der ersten Studie wurde die diagnostische Güte der automatisierten Erkennung der Alzheimer-Krankheit im Demenzstadium mittels DTI- und MRT-Daten bestimmt. Dazu wurden 137 Patienten mit Alzheimer-Demenz und 143 gesunden Kontrollprobanden aus der Datenbank der EDSD ausgewählt. Die demographische Charakteristik der Stichprobe ist in der ersten Doppelspalte in Tabelle 1 angegeben.

Studie 2: In der zweiten Studie wurde die diagnostische Güte der automatisierten Erkennung der Alzheimer-Krankheit im Prodromalstadium auf Basis von DTI- und MRT-Daten bestimmt. Dazu wurden aus der Datenbank der EDSD zwei Teilstichproben ausgewählt: (a) 25 Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor versus 25 gesunden Kontrollprobanden bzw. (b) 35 Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor versus 35 Patienten mit leichter kognitiver Störung und negativem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor. Die Differenz beider Teilstichproben begründet sich darin, dass nicht alle EDSD-Zentren Daten von Kontrollprobanden zur Verfügung gestellt haben. Die Übersicht der demographischen Informationen ist in der zweiten und dritten Doppelspalte in Tabelle 1 angegeben.

Studie 3: In der dritten Studie wurden ausschließlich Daten des Deutschen Zentrums für Neurodegenerative Erkrankungen vom Standort Rostock/Greifswald genutzt. Zur kombinierten Untersuchung von DTI-, MRT-, und Ruhe-fMRT-Daten wurden 28 Patienten einer leichten bis mittelgradigen Alzheimer-Demenz mit positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor und 25 gesunde Kontrollprobanden ausgewählt. Die demographischen Informationen sind in der letzten Doppelspalte in Tabelle 1 angegeben.

In den ersten beiden Studien wurde die Charakterisierung der DTI als Marker für die Alzheimer-Krankheit im Demenz- und Prodromalstadium untersucht (Untersuchungsziel I), Methoden zur Varianzreduktion evaluiert, vor allem zur Reduktion des Störeffekts durch die multizentrische Akquisition (Untersuchungsziel II), sowie die kombinierte Verwendung der

zwei Bildgebungsmodalitäten DTI und MRT ausgewertet (Untersuchungsziel III). In der dritten Studie wurde darauf aufbauend die multimodale Auswertung um die Ruhe-fMRT erweitert, um die Kombination der drei Bildgebungsmodalitäten DTI, MRT und Ruhe-fMRT als Bildgebungsmarker für die Alzheimer-Krankheit zu bewerten (Untersuchungsziel III).

Alle Probanden bzw. deren gesetzlicher Vertreter haben schriftlich ihr Einverständnis zur Studienteilnahme erklärt. Alle Probanden wurden im Vorfeld auf signifikante neurologische, psychiatrische oder medizinische Ausschlusskriterien hin überprüft. Für die gesunden Kontrollprobanden galten folgende Bedingungen: Abwesenheit kognitiver Beschwerden sowie alters- und ausbildungskorrigiertes Testergebnis in der neuropsychologischen Testbatterie *Consortium to Establish a Registry of Alzheimer's Disease* (CERAD) [68] innerhalb einer Standardabweichung.

Tabelle 1. Demographische Charakterisierung der Studienkohorten

	Studie 1		***Studie 2a***		***Studie 2b***		***Studie 3***	
	AD	**HC**	**MCI-Aβ42$^+$**	**HC**	**MCI-Aβ42$^-$**	**MCI-Aβ42$^+$**	**AD**	**HC**
N (Frauen)	137 (59)	143 (72)	25 (9)	25 (11)	35 (15)	35 (13)	28 (14)	25 (13)
Alter[1] (SD)	73 (8)	69 (6)	68 (6)	69 (5)	70 (7)	69 (6)	72 (7)	73 (6)
Ausbildungs-dauer[1] (SD)	10 (3)	13 (4)	12 (3)	13 (3)	12 (3)	12 (3)	13 (3)	13 (2)
MMST (SD)	21 (5)*	29 (1)	26 (2)*	29 (1)	27 (2)	26 (2)	24 (3)*	28 (1)
Scanner	9		3		5		1	

* Statistisch signifikanter Unterschied zwischen AD und HC bzw. MCI und HC, Mann-Whitney-Test, $p<0.05$;

[1] Einheit: Jahre; Abkürzungen: AD – Patienten mit Alzheimer-Demenz, HC – gesunde Kontrollprobanden, MCI – Patienten mit leichter kognitiver Störung, MCI-Aβ42$^-$ bzw. MCI-Aβ42$^+$ – Patienten mit negativem bzw. positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor, N – Anzahl, SD – Standardabweichung, MMST – Mini-Mental-Status-Test

2.2 Bilddatenvorverarbeitung

Im Vorfeld wurden alle Scans visuell überprüft, um Messfehler oder Artefakte auszuschließen. Bevor die Bilddaten statistisch bzw. mittels maschineller Lernverfahren ausgewertet werden konnten, mussten sie entsprechend aufbereitet werden. Dazu durchliefen sie nach dem Protokoll von [45, 48] eine Reihe von Vorverarbeitungsschritten, die mit den Programmen *FMRIB Software Library* (*FSL*, Version 5.0.4, Oxford Center for Functional MRI of the Brain, Oxford, UK) [69], *Statistical Parametric Mapping 8* (*SPM8*, Release 4290, Wellcome Trust Center for Neuroimaging, London, UK) [70], der SPM8 Toolbox *Voxel-Based Morphometry 8* (*VBM8*, Release 433, Structural Brain Mapping Group,

Kliniken für Neurologie und Psychiatrie, Universität Jena) [71] und *Data Processing Assistant for Resting-State fMRI* (*DPARSF*, Version 2.2, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China) [72] ausgeführt wurden. Die Verarbeitungsschritte sind in Abbildung 4 illustriert.

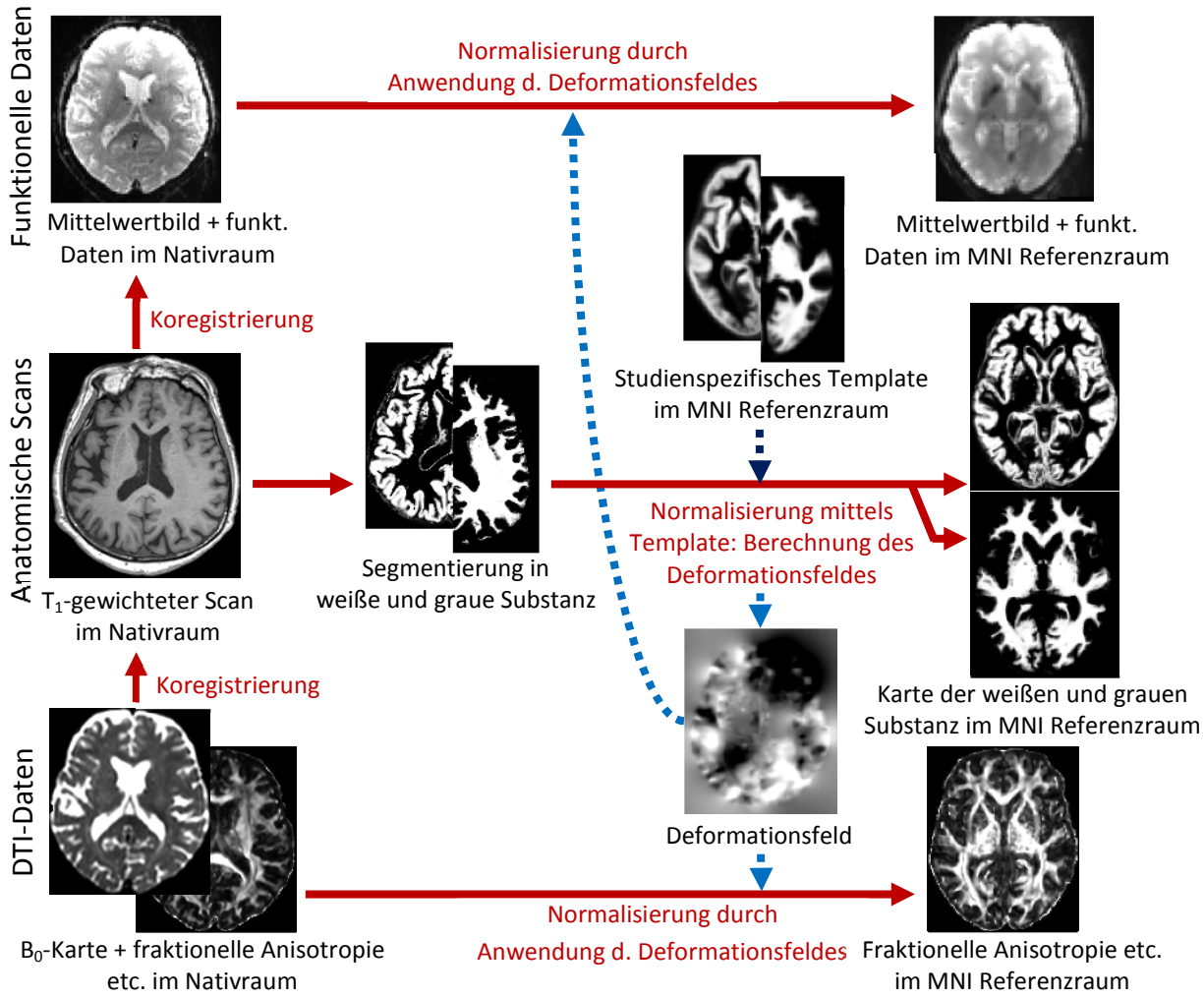

Abbildung 4. Vorverarbeitungsschritte zur räumlichen Normalisierung der Bilddaten

Für die T_1-gewichteten anatomischen Scans wurde mittels Template im MNI Referenzraum das Deformationsfeld zur räumlichen Normalisierung berechnet. Anhand des Deformationsfeldes konnten die funktionellen oder DTI-Daten ebenfalls in den MNI Referenzraum transformiert werden. Das inverse Deformationsfeld (ohne Abb.) konnte genutzt werden, um Bilddaten vom MNI Referenzraum zurück in den Nativraum des Probanden zu transformieren.

Von den **anatomischen T_1-gewichteten MRT-Scans** wurden mit VBM8 die Bereiche der grauen Substanz, weißen Substanz und Cerebrospinalflüssigkeit extrahiert (*Segmentierung*). Dann wurden diese Segmente mit dem Verfahren *Diffeomorphic Anatomical Registration Through Exponentiated Lie* (DARTEL) Algebra [73] und einem studienspezifischen Template

in das *Montreal Neurological Institute* (MNI) Referenzkoordinatensystem gebracht (*Normalisierung*). Dazu berechnet der Algorithmus eine nichtlineare Projektion von den Originalscans im Nativraum des Probanden auf das Template-Gehirn, so dass danach die Gehirne aller Probanden etwa die gleiche Form haben. Der Algorithmus bewahrt dazu die lokal unterschiedlichen Proportionen bzw. Volumina der einzelnen Gyri und subkortikalen Strukturen, so dass die Hirne anschließend voxelweise verglichen werden können. Dazu wurden abschließend die Karten der grauen Substanz und der weißen Substanz mit einem 8 mm breiten Gaußkernel geglättet (*Smoothing*).

Die **DTI-Daten** wurden zuerst mit der *Diffusion Toolbox* von FSL vorverarbeitet, um Verzerrungen bedingt durch Wirbelströme oder Kopfbewegung zu korrigieren, Messwiederholungen zur Verbesserung des Signal-Rausch-Verhältnisses zu mitteln, das Diffusions-Tensor-Modell auf die Rohdaten anzuwenden und die skalaren Kennwerte *fraktionelle Anisotropie*, *mittlere Diffusivität* und *Modus der Anisotropie* [74] zu berechnen (siehe dazu Abbildung 1, S.6). Das diffusionsungewichtete B_0-Bild wurde anschließend auf das T_1-gewichtete MRT koregistriert und diese Transformation für die skalaren DTI-Kennwertkarten übernommen. Die aus der Normalisierung des T_1-gewichteten MRTs gewonnenen Deformationsfelder wurden dann auf die skalaren DTI-Kennwertkarten angewendet, um diese ebenfalls in den MNI-Standardraum zu bringen. Alle DTI-Karten wurden danach maskiert, so dass nur Bereiche innerhalb der weißen Substanz für die nachfolgenden Analysen genutzt wurden. Abschließend wurden die DTI-Karten ebenfalls mit einem 8 mm breiten Gaußkernel geglättet.

Für die Vorverarbeitung der **Ruhe-fMRT-Daten** wurde das Protokoll von [62, 75, 76] verwendet. Die ersten sechs Messzeitpunkte (ca. 15 Sekunden) wurden verworfen, um Artefakte bedingt durch den Aufbau des Magnetgradientenfeldes zu vermeiden. Dann wurde der Effekt von unterschiedlichen Aufnahmezeitpunkten der einzelnen Schichten innerhalb des drei Sekunden dauernden Aufnahmezyklus korrigiert (sog. *Slice-Time Correction*) und die Einzelbilder auf das Mittelwertbild koregistriert, um leichte Kopfbewegung zu kompensieren. Anschließend wurden die Ruhe-fMRT-Daten mittels Regressionsmodell und den Kovariaten Kopfbewegung, gemitteltes Zeitsignal im Gesamthirn, Liquor und Bereich der weißen Substanz bereinigt. Entsprechend der Vorarbeit von Wee et al. [62] wurden mittels Bandpassfilter für jedes Voxel die Zeitsignale innerhalb der fünf Frequenzbänder 0.025–0.03929 Hertz, 0.03929–0.05357 Hz, 0.05357–0.06786 Hz, 0.06786–0.08214 Hz und

0.08214–0.1 Hz extrahiert. Abschließend wurde das funktionelle Mittelwertbild genutzt, um die funktionellen Zeitreihenbilder auf das T_1-gewichtete MRT zu koregistrieren. Für die dritte Studie wurden die Daten im Gegensatz zu den ersten beiden Studien nicht voxelweise ausgewertet, sondern entsprechend der Vorarbeit von Wee et al. [62] regionsbasiert. Dazu wurden die inversen Deformationsfelder der T_1-gewichteten MRT genutzt, um die Atlanten mit den spezifischen Hirnregionen (siehe Abschnitt 2.3.2 weiter unten) aus dem MNI-Referenzraum zurück in den Nativbildraum der einzelnen Probanden zu projizieren.

2.3 Verarbeitung mittels maschineller Lernverfahren

Für die Verarbeitung mittels statistischer oder maschineller Lernverfahren müssen die vorverarbeiteten Bilddaten eingelesen, aufbereitet und dem Lernverfahren zugeführt werden. Diese Schritte umfassen die Anwendung von geeigneten Validierungsmethoden, die die Generalisierbarkeit der Ergebnisse sicherstellen, die Anwendung von Verfahren zur Reduktion der Inter-Scanner-Variabilität, die Auswahl von Merkmalen, die einen Beitrag zur Gruppenunterscheidung liefern, das Trainieren des Klassifikatormodells mithilfe von Trainingsdaten und abschließend die Anwendung des Klassifikatormodells auf Testdaten zur Ermittlung der Leistungskennwerte (Abbildung 5).

2.3.1 Validierungstechniken

Zum Trainieren der maschinellen Lernverfahren und der anschließenden Bestimmung der Leistungskennwerte kam das stratifizierte Kreuzvalidierungsverfahren zum Einsatz. Dabei wurden initial alle Probanden zufällig in verschiedene Gruppen eingeteilt, diese aber stratifiziert, so dass jede Gruppe in etwa die gleiche Anzahl an Patienten und gesunden Probanden aufwies. Die Schritte der Merkmalsselektion, Varianzreduktion und Training des Lernalgorithmus wurden mit allen Probanden außer denen in der Gruppe *i* durchgeführt. Die zurückgehaltenen Probanden aus der Gruppe *i* wurden anschließend zur Bestimmung der Leistungskennwerte genutzt. Dieses Verfahren wurde wiederholt, bis jede Gruppe einmal zur Evaluierung genutzt wurde. Somit wurde die Generalisierbarkeit der Klassifikationsleistung sichergestellt und eine Überanpassung des gelernten Modells an die Trainingsdaten vermieden. In der ersten und zweiten Studie wurde eine zehnmal wiederholte zehnfache Kreuzvalidierung durchgeführt. Nach Kohavi [77] sei die stratifizierte *k*-fache Kreuzvalidierung der Leave-One-Out Kreuzvalidierung oder dem Bootstrap-Ansatz wenn möglich vorzuziehen, da diese Methodik im Vergleich den geringsten systematischen Fehler

und die kleinste Variabilität der Leistungskennwerte aufweise. Für die erste Studie wurde neben der Kreuzvalidierung mit den Daten aller MRT-Scanner (*Pooling*) zusätzlich auch eine Kreuzvalidierung mit den Daten aufgeteilt nach den Scannern durchgeführt. Bei diesem Ansatz wurden die Klassifikationsmodelle mit den Daten anderer MRT-Scanner trainiert und die Leistung anhand der Daten des dem Modell vorenthaltenen Scanners evaluiert. In der dritten Studie wurde anknüpfend an die Vorarbeit von Wee et al. [62] die Leave-One-Out Kreuzvalidierung verwendet, um eine bessere Vergleichbarkeit der Ergebnisse zu ermöglichen.

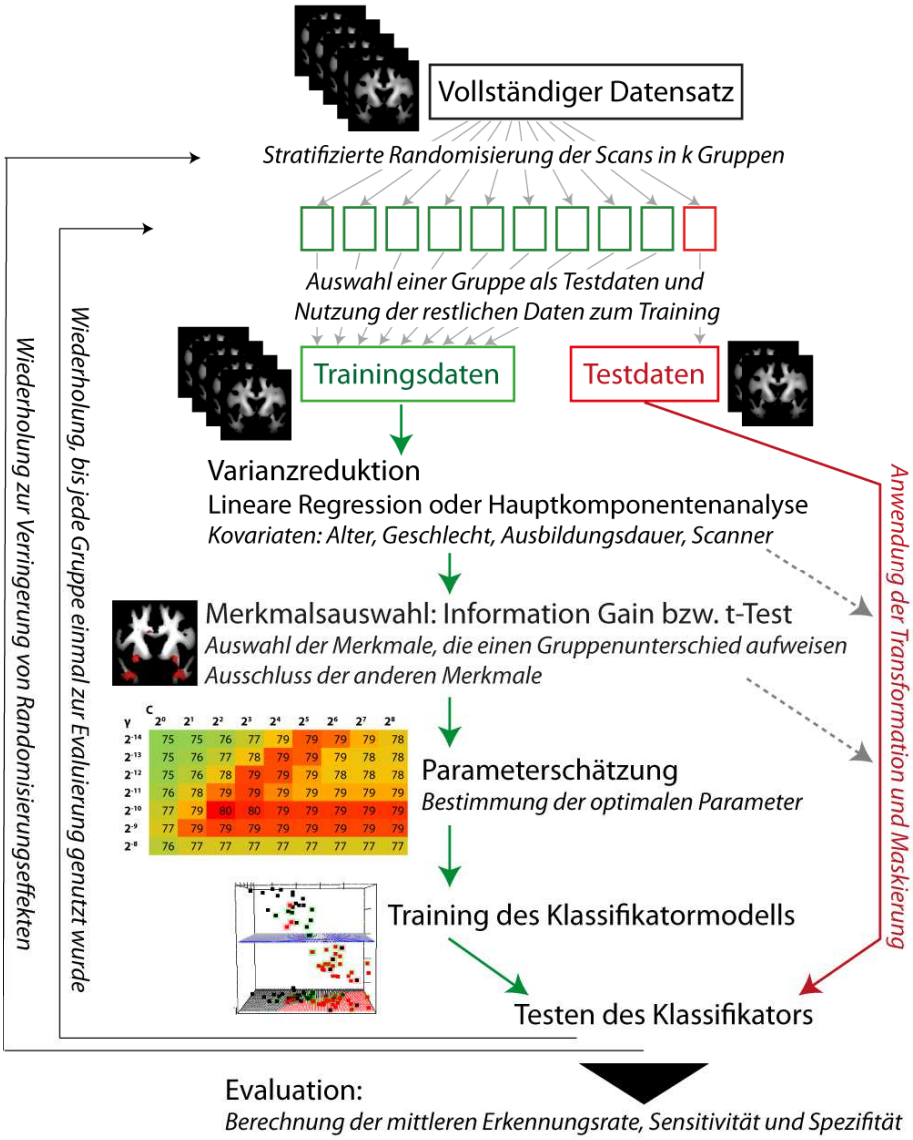

Abbildung 5. Illustration der Datenverarbeitung mittels maschineller Lernverfahren

Die Schritte Varianzreduktion, Auswahl von Merkmalen, die einen Beitrag zur Gruppenunterscheidung liefern, Parameterschätzung und Training des Klassifikatormodells wurden innerhalb der k-fachen Kreuzvalidierung rein anhand der Trainingsdaten ausgeführt. Abschließend wurden die Leistungskennwerte des Klassifikatormodells anhand der Testdaten evaluiert.

2.3.2 Merkmalsextraktion

Für die **ersten beiden Studien** wurden die DTI-Daten und volumetrischen MRT-Daten voxelbasiert ausgewertet, d.h. jedes Voxel wurde als eigenes Merkmal betrachtet. Die Analysen wurden für jede der Modalitäten fraktionelle Anisotropie, mittlere Diffusivität, Modus der Anisotropie (nur zweite Studie), Volumen der weißen Substanz und Volumen der grauen Substanz separat durchgeführt und anschließend noch einmal multimodal für die Kombination aus allen wiederholt.

Für die **dritte Studie** wurde anknüpfend an die Vorarbeit von Wee et al. [62] mit Hilfe verschiedener Atlasse die Kennwerte spezifischer Hirnregionen extrahiert und zur weiteren Analyse verwendet. Für die Ruhe-fMRT wurde anhand eines funktionell definierten Atlasses [76] die Korrelation des fMRT-Zeitsignals zwischen den verschiedenen Hirnregionen bestimmt, die sogenannte *funktionelle Konnektivität*. Basierend auf der Korrelationsmatrix wurden die graphentheoretischen Kennwerte *gewichteter lokaler Clustering-Koeffizient* (WLCC) [78] und *kürzeste gewichtete Pfadlänge* (SWPL) [36, 42] berechnet. Diese beiden Kennwerte bemessen die Vernetzungsstärke einzelner Regionen untereinander bzw. die Effizienz des Netzwerks, Informationen von Region *a* zu einer Region *b* zu übermitteln. Die Berechnung der beiden Kennwerte ist in Abbildung 6 illustriert. Für die DTI-Daten wurden mit Hilfe eines DTI-Atlasses [79] die Mittelwerte der fraktionellen Anisotropie, der mittleren Diffusivität und des Modus der Anisotropie für zahlreiche Hauptfasertrakte des Gehirns extrahiert. Für die anatomischen T_1-gewichteten MRT-Daten wurden atlasbasiert das Volumen der grauen Substanz zahlreicher kortikaler und subkortikaler anatomischer Strukturen [80] bestimmt.

2.3.3 Varianzreduktion

Zur Reduktion der Störeffekte der multizentrischen Bildakquisition und zur Verringerung des Effekts bedingt durch die Kovariaten Alter, Geschlecht und Ausbildungsniveau wurden zwei Verfahren zur Varianzreduktion implementiert und deren Einfluss auf die Erkennungsrate der maschinellen Lernverfahren evaluiert. Vorherige Studien haben gezeigt, dass die demographischen Kovariaten *Alter*, *Geschlecht* und *Ausbildungsniveau* einen signifikanten Einfluss auf das Volumen der grauen Substanz bzw. die DTI-Kennwerte haben [81-89]. Zur Repräsentation der Störeffekte der multizentrischen Bildakquisition wurde eine Dummycodierung des jeweiligen Scanners und der Magnetfeldstärke verwendet.

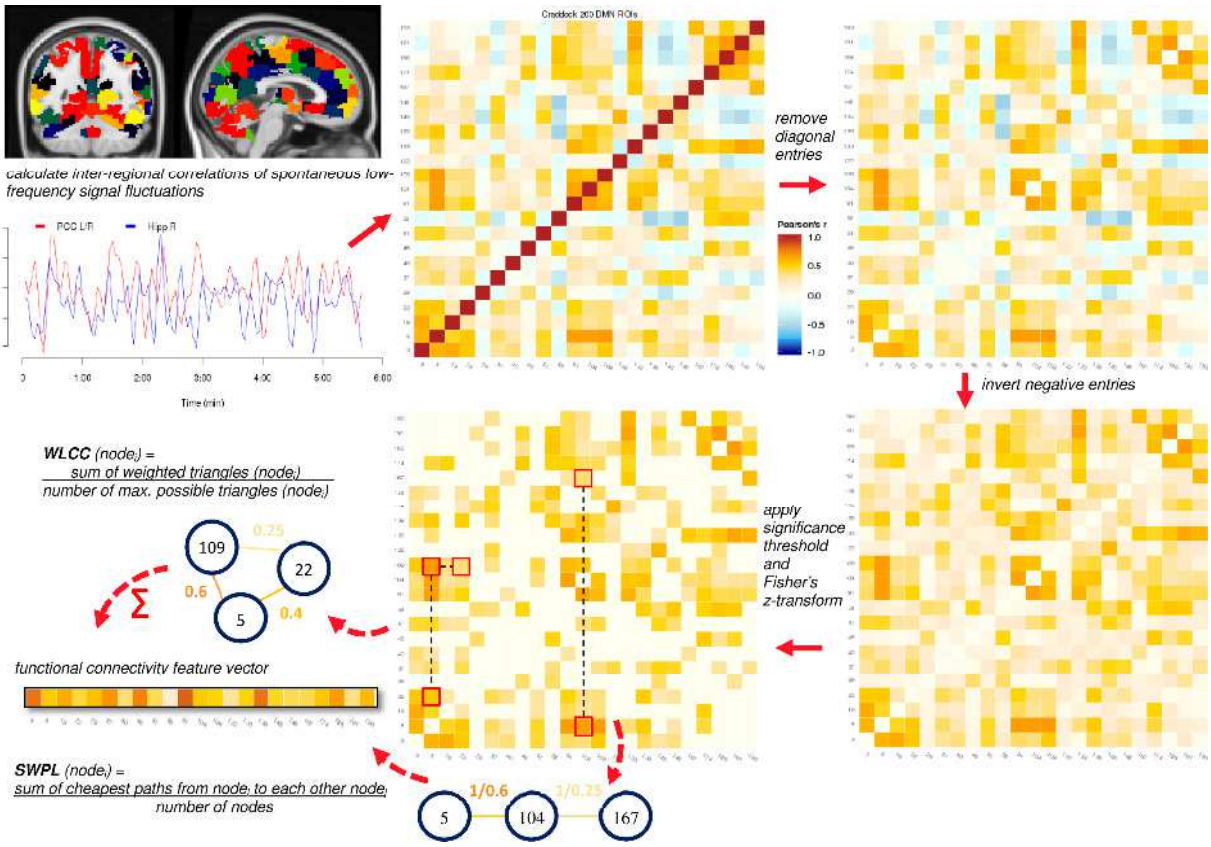

Abbildung 6. Berechnung der Kennwerte *gewichteter lokaler Clustering-Koeffizient* (WLCC) und *kürzeste gewichtete Pfadlänge* (SWPL) aus der funktionellen Konnektivitätsmatrix

Der Ablauf der Berechnung ist hier schematisch dargestellt: 1. Berechnung der Konnektivitätsmatrix, die die Korrelation des Zeitsignals zweier Regionen enthält; 2. Entfernen der Diagonale und Invertierung von negativen Korrelationskoeffizienten; 3. Anwendung eines Schwellwerts zum Entfernen statistisch nicht signifikanter Korrelationen; 4. Anwendung der Fisher z-Transformation zur Normalisierung; 5. Berechnung der graphentheoretischen Kennwerte. Zu Illustrationszwecken enthält die abgebildete Konnektivitätsmatrix nur einen Ausschnitt der insgesamt 200 Cluster aus dem Craddock-Atlas, die für die Analyse verwendet wurden.

Dieser Ansatz wurde erforderlich, da jedes Zentrum ein anderes DTI-Scanprotokoll verwendet hatte und sich die verwendeten Scannertypen größtenteils unterschieden. Frühere Untersuchungen der Intra-Scanner- und Inter-Scanner-Variabilität bei DTI und anatomischer MRT haben gezeigt, dass sogar bei gleichem Scannertyp und Scanprotokoll signifikante, systematische Unterschiede der Variabilität zwischen zwei Scannern auftreten [45, 46, 48, 56]. Diese Studien haben ebenfalls gezeigt, dass die Variablen *Scanner* und *Magnetfeldstärke* im Vergleich mit anderen Parametern den größten Anteil der Varianz aufklären. Auf die Hinzunahme weiterer Variablen in das statistische Modell wurde verzichtet, um die Komplexität der Modelle sowie die Gefahr der Überschätzung von Modellparametern zu verringern.

In Anlehnung an das statistische Standardverfahren der Faktorenanalyse [90 S.28-33, 91 Kap.23] wurde eine Bereinigung auf Basis der **Hauptkomponentenanalyse** (PCA) durchgeführt. Dieser Methode liegt die Annahme zugrunde, dass die Effekte bedingt durch die Alzheimer-Krankheit und die Störeffekte bedingt durch die multizentrische Bildakquisition linear unabhängig voneinander sind [48]. Dieser Annahme folgend würden diese Effekte unter Anwendung der Hauptkomponentenanalyse in verschiedene Varianzkomponenten aufgeteilt und der Einfluss der Störeffekte dann gezielt eliminiert werden können. Hierzu wurde mittels der Hauptkomponentenanalyse die Bilddaten $\boldsymbol{X}_{Training} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)^T$ in ihre Varianzkomponenten $\boldsymbol{V}_{Training} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ zerlegt und eine Projektion ermittelt, die die Originaldaten in den Vektorraum überführt, der durch die Hauptkomponenten aufgespannt wird: $\boldsymbol{Y}_{Training} = \boldsymbol{V}_{Training}{}^T \boldsymbol{X}_{Training}$. Die Werte im Hauptkomponentenbezugssystem $\boldsymbol{Y}_{Training} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)^T$ wurden dann mit den o.g. Kovariaten korreliert, um diejenigen Varianzkomponenten zu identifizieren, die am ehesten mit den einzelnen Kovariaten übereinstimmen. Das heißt, wenn der Korrelationskoeffizient einen bestimmten Schwellwert überschritt, wurde die entsprechende Komponente $\boldsymbol{v}_i$ als Störvarianz interpretiert und entfernt, d.h. mit einem Nullvektor $\boldsymbol{v}_0$ ersetzt: $\boldsymbol{V}_{partial} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_0, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n)$. Anschließend wurden die Daten unter Ausschluss der Störkomponenten zurück in den Originalbildraum projiziert und weiterverarbeitet: $\boldsymbol{X}_{Training,cleaned} = \boldsymbol{V}_{partial} \boldsymbol{Y}_{Training}$. Entsprechend wurde diese Datenbereinigung auch auf die Testdaten $\boldsymbol{X}_{Test}$ angewandt: $\boldsymbol{Y}_{Test} = \boldsymbol{V}_{Training}{}^T \boldsymbol{X}_{Test}$ und $\boldsymbol{X}_{Test,cleaned} = \boldsymbol{V}_{partial} \boldsymbol{Y}_{Test}$. Als Schwellwerte zur Identifikation von Hauptkomponenten mit hohem Anteil an Störvarianz wurden empirisch die Korrelationskoeffizienten $|r| \geq 0.6$, $|r| \geq 0.5$ und $|r| \geq 0.4$ ermittelt und der Einfluss auf die Erkennungsleistung der Lernverfahren verglichen.

Als zweite Methode wurde die Varianzreduktion mittels **linearer Regression** gewählt [52 Kap.3, 91 Kap.11], ein statistisches Standardverfahren zur Bemessung des linearen Zusammenhangs von unabhängigen Variablen auf eine abhängige Zielgröße. Hierbei wurde für jedes Voxel der Beitrag der Kovariaten und der Diagnose geschätzt (nur Trainingsdaten): $x_{ij} = \beta_{0j} + \beta_{1j}\, age_i + \beta_{2j}\, gender_i + \beta_{3j}\, edu_i + \beta_{4j}\, center_{1i} + \ldots + \beta_{5j}\, center_{ki} + \beta_{6j}\, diagnosis_i + \varepsilon_{ij}$, mit der Intensität x_{ij} des Voxels j von Proband i, den Regressionsparametern $\beta_{0j} \ldots \beta_{6j}$ für Voxel j, den Kovariaten $age_i \ldots diagnosis_i$ für Proband i und dem Fehlerterm ε_{ij} für Voxel j und Proband i. Im zweiten Schritt wurde das Residuum e_{ij} aus der Differenz des vom Regressionsmodell ohne Verwendung der Diagnoseinformation vorhergesagten Wertes und

der gemessenen Intensität berechnet (Trainings- und Testdaten): $e_{ij} = x_{ij} - (\beta_{0j} + \beta_{1j}\, age_i + \beta_{2j}\, gender_i + \beta_{3j}\, edu_i + \beta_{4j}\, center_{1i} + \ldots + \beta_{5j}\, center_{ki})$. Damit setzt sich das Residuum e_{ij} vereinfacht aus der Diagnose und dem Fehlerterm zusammen: $e_{ij} = \beta_{6j}\, diagnosis_i + \varepsilon_{ij}$. Da die Diagnoseinformation der Testdaten an dieser Stelle nicht berücksichtigt werden soll, wurde die erstgenannte Berechnungsformel genutzt. Die Residuen e_{ij} wurden dann statt der originalen Voxelwerte für die weitere Verarbeitung verwendet. Regressionsmodelle mit höherer Ordnung oder Gauß-Prozess Regression [92, Kap.2.2] wurden nicht verwendet. Hauptgrund dafür war die hohe Anzahl der zu schätzenden Modellparameter im Vergleich zur Stichprobengröße, die möglicherweise eine starke Überanpassung des Modells an die Trainingsdaten verursacht hätte. Vorstudien haben gezeigt, dass zum Beispiel zweistufige Regressionsverfahren oder eine metaanalytische Modellierung des Zentrumseffekts keinen signifikanten Einfluss auf die erzielten Erkennungsraten haben [48]. Neuere Studienergebnisse zeigen in Übereinstimmung dazu, dass Gauß-Prozess-Regression keine Verbesserung der Erkennungsrate im Vergleich zu weniger elaborierten Regressionsverfahren ermöglicht [93].

Als Varianzreduktionsverfahren bei der Scanner-basierten Kreuzvalidierung (Studie 1) wurde eine einfache **Mittelwertkorrektur** durchgeführt: Von den gesunden Probanden, die im selben Scanner gescannt wurden, wurde das Mittelwertbild berechnet und anschließend von jedem Scan dieses Scanners subtrahiert: $x_{ij}^* = x_{ij} - \bar{x}_{sj}$, mit der Intensität x_{ij} des Voxels j von Proband i und dem Mittelwert $\bar{x}_{sj}$ des Voxels j für die gesunden Probanden, die im gleichen Scanner s wie Proband i gescannt wurden. Dieses Verfahren vermeidet, dass die Testdaten bei der Modellschätzung inkludiert werden müssen.

2.3.4 Merkmalsauswahl

Ein Großteil der Scans enthält Voxel, die in der Vorverarbeitung maskiert wurden oder deren Intensität keine Unterschiede im Vergleich zwischen den diagnostischen Gruppen aufweisen. Solche Voxel wurden als uninformativ bezüglich der Gruppenunterscheidung betrachtet und vor dem Trainieren des Lernverfahrens entfernt.

In den **ersten beiden Studien** wurde hierzu, aufbauend auf der Arbeit von Plant et al. [55], das informationstheoretische Maß *Information Gain* (IG) [94 Kap.4.3] genutzt, das in der Software-Bibliothek *Waikato Environment for Knowledge Analysis* (WEKA) (Machine Learning Group, University of Waikato, Hamilton, New Zealand) [94] implementiert ist. Der

Algorithmus kommt unter anderem beim Training von Entscheidungsbäumen zum Einsatz und unterteilt die Werte eines Merkmals in mehrere, möglichst klassenreine Intervalle. Anschließend wird auf Basis der Entropie der Informationsgehalt dieses Merkmals bestimmt, d.h. eine Kennzahl zwischen null und eins, die angibt, wie gut dieses Merkmal zur Gruppenunterscheidung geeignet ist[1]. Zur Weiterverarbeitung wurden alle Merkmale (Voxel) ohne Information bezüglich der Gruppentrennung, d.h. mit *IG* = 0, entfernt.

In der **dritten Studie** wurde der rechnerisch weniger komplexe Zweistichproben-*t*-Test genutzt. Die Bestimmung des am besten geeigneten Schwellwerts für den Mittelwertunterschied *t* wurde im Schritt der Parametersuche integriert (siehe nachfolgender Abschnitt).

2.3.5 Lernverfahren

Zur automatisierten Gruppenunterscheidung wurde der Algorithmus *Support Vector Machine* (SVM) [53] genutzt, der bereits in zahlreichen Bildgebungsstudien verwendet wurde [54, 55, 57, 59-63]. Die Funktionsweise des SVM-Algorithmus ist in Abbildung 7 skizziert. Das multivariate Verfahren berechnet mit Hilfe einer sogenannten Kernel-Funktion eine Matrix, die die Ähnlichkeit der Merkmalsvektoren aller Probanden zueinander repräsentiert. Hierzu wurde der nichtlineare Gauß-Kernel $k(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2)$ genutzt. Auf Basis der Kernel-Matrix bestimmt der Algorithmus die sogenannten *Support*-Vektoren, d.h. diejenigen Probanden, die ähnlich zu sein scheinen, aber einer unterschiedlichen Klasse (=Diagnose) angehören. Weiter bestimmt der Algorithmus Gewichte (Faktoren) für die Support-Vektoren, um damit eine nichtlineare Projektion der Ursprungsdaten in einen höherdimensionalen Raum zu beschreiben (siehe Abbildung 7 A und B). Anhand dieser Projektion werden die Daten klassifiziert: Punkte oberhalb der Entscheidungsebene, d.h. mit einem Entscheidungswert mit positivem Vorzeichen, werden zur einen Klasse zugeordnet; Punkte unterhalb der Entscheidungsebene, d.h. mit negativem Entscheidungswert, werden zur anderen Klasse zugeordnet (siehe Abbildung 7 C und D). Durch die nichtlineare Projektion ergibt sich eine kurvenförmige Entscheidungsgrenze. Der Vorteil des Gauß-Kernels gegenüber der linearen Kernelfunktion besteht darin, dass, je nach Parametrisierung des Kernels mit dem Parameter *γ*, die Entscheidungsgrenze von sehr

[1] Details und die formale mathematische Definition des *Information Gain* Kriteriums sind im Anhang aufgeführt: Dyrba et al., 2013, Abschnitt *Feature Selection* (S.5).

kurvenförmig bis hin zu einer linearen Form justiert werden kann. Zur Anwendung des SVM-Modells auf neue Daten wird die Kernel-Funktion für diese und die Support-Vektoren berechnet. Anschließend wird die Projektionsfunktion ausgewertet und anhand des Vorzeichens des Ergebnisses die geschätzte Klassenzugehörigkeit bestimmt. Für den SVM-Algorithmus müssen zwei Parameter festgelegt werden: Die Kosten oder Strafe C für fehlklassifizierte Trainingsdaten und die Breite des Gauß-Kernels γ. Um die optimalen Parameter zu bestimmen, wurde eine Gittersuche für jede Kombination der beiden Parameter im Bereich $C = 2^{-3}, 2^{-2}, \ldots, 2^{8}$ und $\gamma = 2^{-14}, 2^{-13}, \ldots, 2^{-1}$ durchgeführt. Dieser Bereich wurde empirisch aus den Daten bestimmt.

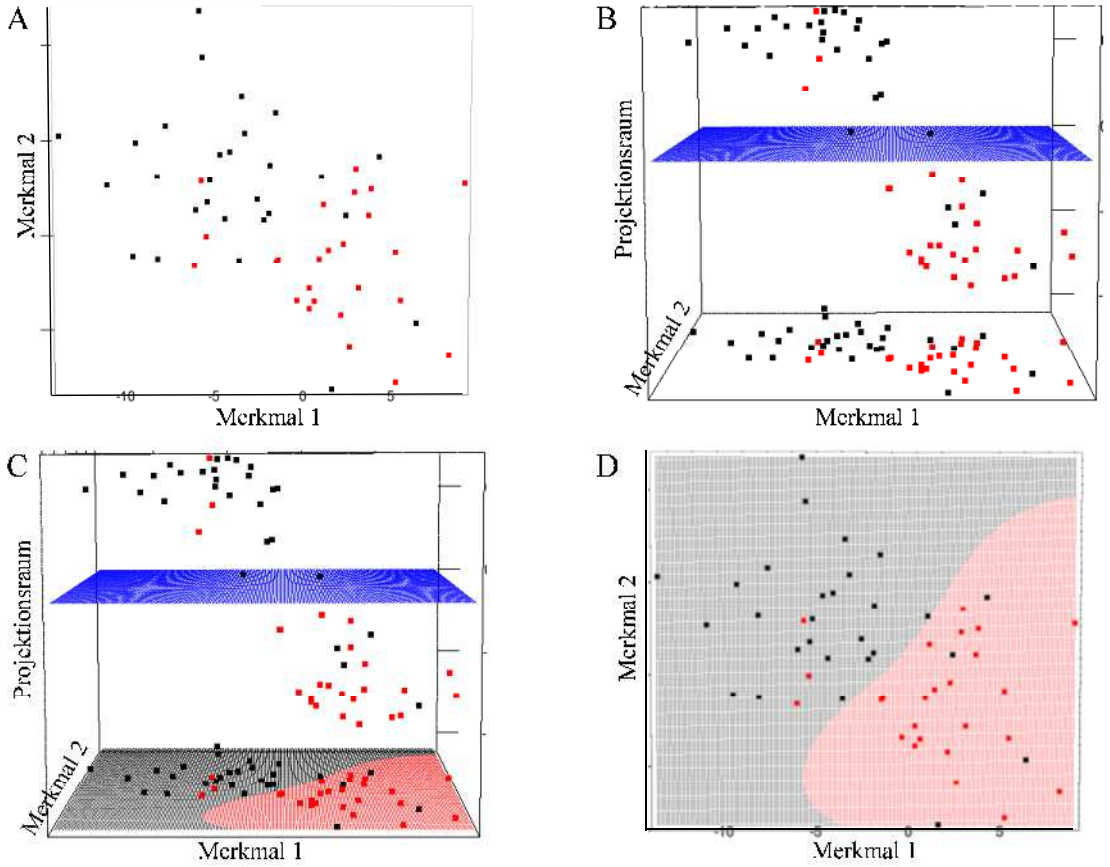

Abbildung 7. Illustration des Support Vector Machine (SVM) Verfahrens

A: Streudiagramm für zwei Merkmale und die zwei Gruppen rot und schwarz. **B:** Der SVM-Algorithmus bestimmt eine nichtlineare Projektion der Daten in einen höherdimensionalen Raum (oben), in dem die Daten mittels linearer Entscheidungsebene (blau) klassifiziert werden. Durch die nichtlineare Projektion folgt in diesem Fall die Anordnung der Datenpunkte einer Sigmoid-Form. Im unteren Bereich sind zum Vergleich die Ausgangsdaten abgebildet. **C** und **D:** Durch die nichtlineare Projektion ergibt sich eine nichtlineare, kurvenförmige Entscheidungsgrenze für die Trennung zwischen den beiden Gruppen rot und schwarz.

Das Höhenprofil der Erkennungsrate bildete im Allgemeinen ein V-förmiges Muster, da die höchsten Erkennungsraten bei einem konstanten *γ* oder einem konstanten Verhältnis von *C* und *γ* erreicht wurden (siehe Abbildung 5 auf S.16, *Parameterschätzung*). Um eine Überanpassung der beiden Parameter an die Trainingsdaten zu verhindern, wurde eine interne Kreuzvalidierung genutzt: Die SVM-Modelle wurden mit einem Parameterpaar parametrisiert, an einer Untermenge der Trainingsdaten trainiert und anhand einer anderen Untermenge der Trainingsdaten die Erkennungsleistung bestimmt. Aus diesen Ergebnissen wurde das Parameterpaar mit der besten Erkennungsleistung ausgewählt und abschließend für das Trainieren und Evaluieren der Klassifikationsleistung genutzt. Für die Parametersuche des optimalen IG- oder t-Schwellwerts zur Merkmalsauswahl (zweite bzw. dritte Studie) wurde die Gittersuche um die entsprechenden Parameter erweitert und äquivalent durchgeführt. Da die Parametersuche nur die Trainingsdaten nutzte, aber nicht die Evaluierungsdaten, wurde sichergestellt, dass die final bestimmten Leistungskennwerte auch tatsächlich einer objektiven Schätzung entsprachen.

Für die Verarbeitung von multimodalen Daten wurde der SVM-Algorithmus zur *Multikernel Support Vector Machine* (MK-SVM) erweitert [62, 95]. Zuerst wird dabei die Kernelmatrix separat für jede Datenmodalität *m* berechnet $k^{(m)}(x_i^{(m)}, x_j^{(m)}) = \exp(-\gamma_m \parallel x_i^{(m)} - x_j^{(m)} \parallel^2)$ und im zweiten Schritt der sogenannte *Mixed Kernel* als gewichtete Summe der Einzelkernelmatrizen $k(x_i, x_j) = \sum_{m \in \mathrm{M}} \beta_m \, k^{(m)}(x_i^{(m)}, x_j^{(m)})$. Die Gewichte $\beta_m \in \{0, 0.5, 0.75, 1, 1.25, 1.5\}$ wurden wahlweise a priori festgelegt, so dass sie der Güte der Erkennungsleistung der Einzelmodalitäten entsprachen, oder alternativ im Rahmen der Gittersuche bestimmt.

Zusätzlich zum SVM-Algorithmus wurde in der ersten Studie auch der *Naive Bayes* (NB) Klassifikator verwendet. Im Vergleich zum multivariaten SVM-Verfahren arbeitet der NB-Ansatz univariat, d.h. jedes Merkmal wird separat ausgewertet. Das Verfahren ist unter der Voraussetzung der statistischen Unabhängigkeit der Merkmale optimal bezüglich des statistischen Fehlers. Plant et al. [55] zeigten, dass das Verfahren ebenfalls gute Ergebnisse liefert, auch wenn die Annahme der statistischen Unabhängigkeit verletzt wird.

2.3.6 Evaluation

Zur Evaluation der maschinellen Lernverfahren wurden die Mittelwerte der diagnostischen Genauigkeit, Sensitivität und Spezifität berechnet. Als Streumaß wurde das 2,5 und 97,5

Perzentil bestimmt, um das 95%-Konfidenzintervall zu schätzen. In der zweiten Studie war eine Bewertung des Nutzens der Varianzreduktionsverfahren mittels dieser Kennwerte allein nicht möglich, da sie zwischen den Modalitäten stark variierten. Deshalb wurde zusätzlich der Anteil der Support-Vektoren an den Trainingsprobanden ermittelt, um den Nutzen der Varianzreduktion differenzierter bewerten zu können. Dieser Kennwert beschreibt die Komplexität des gelernten Modells [53] und lässt sich damit als indirekte Kennzahl für das Verhältnis von Gruppenunterschied zu Störvarianz nutzen. In der dritten Studie wurde außerdem die *Fläche unter der Kurve* (AUC) bestimmt. Dieser Kennwert beschreibt ergänzend zu den oben genannten Maßen das Fehlklassifikationsverhalten, wenn die Entscheidungsebene des gelernten SVM-Modells nicht fixiert ist, sondern in der Höhe variiert wird (siehe Abbildung 8).

Zum Vergleich der Klassifikationsleistung zwischen jeweils zwei Bildmodalitäten wurden Kontingenztabellen verwendet. Von diesen wurde der Jaccard-Ähnlichkeitskoeffizient $J =$ *(Anzahl der mittels beider Modalitäten richtig erkannter Probanden)* / *(Anzahl der von mindestens einer Modalität richtig erkannter Probanden)* bestimmt und der McNemar-X^2-Test zur Prüfung der Unterschiede auf statistische Signifikanz berechnet.

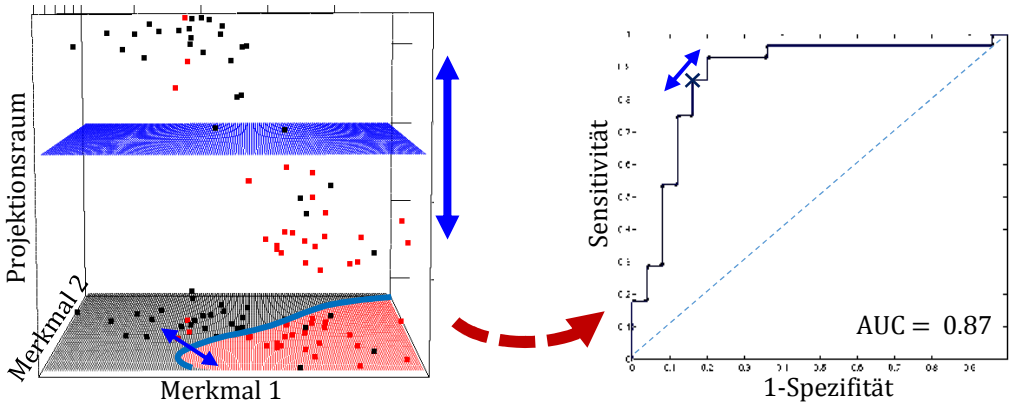

Abbildung 8. Bestimmung der *Fläche unter der Kurve* (AUC) durch Variation der Entscheidungsebene des SVM-Modells

2.3.7 Permutationstests

Bei der zweiten Studie, insbesondere bei der Gruppenunterscheidung von Patienten mit leichter kognitiver Störung und positivem Liquorbefund für Amyloid-β_{1-42} von Patienten mit negativem Liquorbefund erzielten die Lernverfahren relativ niedrige Erkennungsraten von zum Teil 60%. Da die Zufallswahrscheinlichkeit der Gruppenzuordnung ungefähr 50% betrugt, wurde mittels Permutationstests [96, 97] evaluiert, ob die Erkennungsraten

statistisch signifikant über dem Zufallsniveau lagen. Dazu wurde die Diagnoseinformation der Probanden zufällig permutiert und die Schritte Varianzreduktion, Merkmalsauswahl, Modelltraining und Modellevaluation erneut mit der Zufallszuordnung durchgeführt. Dieses Verfahren wurde $k = 100$-mal wiederholt, um die empirische Verteilung der Erkennungsrate für die zufällige Gruppenzuordnung D_Z zu schätzen. Anschließend wurde für jeden der $r = 10$ Ergebniswerte für die Originaldaten D_O (zehnfach wiederholte Kreuzvalidierung) der Anteil an Zufallsergebnissen bestimmt, für den die Zufallserkennungsrate größer als die Originaldatenerkennungsrate war: $p_i = (|acc_Z \in D_Z, acc_i \in D_O: acc_Z > acc_i| + 1) / (k + 1)$, für $i = 1, \ldots, r$. Der Mittelwert der p_i wurde als empirischer p-Wert für die Wahrscheinlichkeit genutzt, dass die Erkennungsraten der Originaldaten ein Zufallsbefund sind [97].

2.3.8 Visualisierung der informativsten Merkmale

Bei der Anwendung multivariater maschineller Lernverfahren als diagnostisches Assistenzsystem für Radiologen oder Neurologen stellt sich die Frage, welche Merkmale am informativsten bezüglich Gruppenunterscheidung sind. Die Transparenz des Verfahrens und der kausale Zusammenhang zwischen den Eingangsdaten und der Entscheidung eines solchen Assistenzsystems sind deshalb von großer Relevanz. Da beim SVM-Algorithmus unter Verwendung des Gauß-Kernels dieser Zusammenhang nicht direkt abgeleitet werden kann, wurde die Methodik der Sensitivitätsanalyse entwickelt [98, 99]. Diese Technik ermittelt heuristisch den Einfluss der Variabilität jedes einzelnen Merkmals auf die Entscheidung des Modells. Mit diesem Verfahren konnte bestimmt werden, welche Voxel oder Regionen den größten Beitrag zur Klassifikationsentscheidung leisteten.

3 Ergebnisse

3.1 Umsetzung der Datenverarbeitung

Zur Datenaufbereitung und –verarbeitung mittels maschineller Lernverfahren wurden zwei Programmbibliotheken entwickelt: Die erste Bibliothek wurde in Java implementiert und lässt sich über die Kommandozeile ausführen. Sie enthält mehrere Komponenten, z.B. zur Zuordnung der Probanden zu den Kreuzvalidierungsgruppen, zum Lesen von 3D-Bilddaten im *NIfTI*-Binärformat, zur Merkmalsauswahl und zum Exportieren der Daten in das *ARFF*-Textformat, das von der Machine-Learning-Bibliothek *WEKA* verwendet wird. Die Machine-Learning-Bibliothek *WEKA* konnte dadurch direkt zur Parametersuche, zum Training und zur Evaluation der Lernverfahren verwendet werden. Die Varianzreduktionsverfahren, die multimodale Auswertung mittels MK-SVM und die Visualisierung wurden in Matlab implementiert und können optional der restlichen Verarbeitung vor- bzw. nachgeschaltet werden. Diese Bibliothek wurde zur Analyse der Bilddaten in den ersten beiden Studien verwendet. Die zweite Programmbibliothek wurde deutlich fokussierter umgesetzt und arbeitet im Gegensatz zu dieser rein auf tabellenförmigen Ausgangsdaten, z.B. Regionsmittelwerten. Sie wurde in Matlab implementiert und beschränkt sich bei den Verarbeitungsschritten auf die Kreuzvalidierung mit eingebetteter Varianzreduktion (regressionsbasiert), Merkmalsselektion (t-Test), Parameteroptimierung und Klassifikation mit dem SVM-Algorithmus mit nichtlinearer Gauß-Kernelfunktion. Diese Bibliothek wurde in der dritten Studie verwendet.

3.2 Leistung der Lernverfahren

Die Erkennungsraten für die Gruppenunterscheidung der multizentrischen Bilddaten sind in Abbildung 9 und Abbildung 10 (Studie 1) sowie Abbildung 11 und Abbildung 12 (Studie 2) angegeben. Die Sensitivität, Spezifität, Anteil der Support-Vektoren sind in den jeweiligen Publikationen im Anhang aufgeführt. Für die Unterscheidung von Alzheimer-Demenzpatienten und gesunden Probanden wurde mittels SVM-Algorithmus eine Erkennungsrate von ca. 85% für die DTI-Kennwerte fraktionelle Anisotropie und mittlere Diffusivität sowie für das Volumen der weißen Substanz erreicht und ca. 90% für das Volumen der grauen Substanz sowie die multimodale Analyse mittels MK-SVM (Abbildung 9). Die Unterscheidung von Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor von gesunden Probanden zeigte eine eher heterogene Leistung

(Abbildung 11): Für die mittlere Diffusivität wurde eine Erkennungsgüte von 77% und für die MK-SVM 74%, für die anderen Modalitäten dagegen signifikant schlechtere Werte von ca. 60% (fraktionelle Anisotropie, mittlere Diffusivität, Volumen der weißen und grauen Substanz; Abbildung 11) ermittelt. Die Gruppentrennung von Patienten mit leichter kognitiver Störung und negativem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor gegenüber von Patienten mit positivem Befund für Amyloid-$\beta_{1\text{-}42}$ erreichte Erkennungsraten von fast 70% für die DTI-Daten (fraktionelle Anisotropie, mittlere Diffusivität, Modus der Anisotropie) sowie die multimodale Analyse und ca. 60% für die volumetrischen Daten (graue oder weiße Substanz; Abbildung 12). Das multivariate Lernverfahren SVM ermöglichte signifikant höhere Erkennungsraten als der univariate Naive Bayes Algorithmus (Differenz ca. 10%, ohne Abbildung; siehe Dyrba et al. 2013, *Table 3* & *Table 5*).

Die besten Leistungen wurden mit der regressionsbasierten Varianzreduktion erzielt, wobei die Unterschiede zur Varianzreduktion mittels Hauptkomponentenanalyse bzw. zu den unkorrigierten Originaltaten größtenteils gering bzw. nicht statistisch signifikant waren (Differenz bis zu 4% für die Unterscheidung von Demenzpatienten und Kontrollen, Abbildung 9; bis zu 8% für die Unterscheidung von Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1\text{-}42}$ von Patienten mit negativem Befund für Amyloid-$\beta_{1\text{-}42}$, Abbildung 11). Die größten Abweichungen zwischen den einzelnen Ansätzen waren bei der Unterscheidung von Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor gegenüber gesunden Probanden zu verzeichnen (Differenz bis zu 18% bei der Erkennungsrate, Abbildung 12). Die SVM-Modelle unter Verwendung der regressionsbasierten Varianzreduktion wiesen mit im Mittel 67% den geringsten Anteil an Support-Vektoren auf (89% für die unkorrigierten Daten, 91% für die Varianzreduktion mittels Hauptkomponentenanalyse; Einzelwerte siehe Dyrba et al. 2015a, *Table 2* & *Table 3*), d.h. unter Verwendung der regressionsbasierten Varianzreduktion wiesen die SVM-Modell die geringste Komplexität auf.

Für die erste Studie wurde neben der Kreuzvalidierung mit den Daten von allen MRT-Scannern (*pooling*) zusätzlich auch eine Kreuzvalidierung mit den Daten aufgeteilt nach dem Scanner durchgeführt. Bei diesem Ansatz wurden Klassifikationsmodelle mit den Daten anderer MRT-Scanner trainiert und die Leistung anhand der Daten eines dem Modell unbekannten Scanners/Zentrums evaluiert. Anhand der unkorrigierten Originaldaten war die Erkennungsleistung für die Alzheimer-Demenz in diesem Fall größtenteils niedriger als bei

der Kreuzvalidierung mit den *gepoolten* Daten (Abbildung 9 und Abbildung 10). Unter der Verwendung der Mittelwertkorrektur zur Varianzreduktion innerhalb der Daten jedes Scanners wurde mit dem SVM-Algorithmus eine ähnlich hohe Erkennungsrate für die volumetrischen Merkmale bzw. eine etwas geringere Erkennungsrate für die DTI-Kennwerte (Differenz ca. 8%) im Vergleich zur Kreuzvalidierung mit den *gepoolten* Daten erreicht (Abbildung 9 und Abbildung 10). Diese Analyse konnte bei der zweiten Studie nicht ausgeführt werden, da hierzu der Stichprobenumfang nicht ausreichte. Für die erste Studie wurden außerdem mehrere Ansätze zur multimodalen Auswertung betrachtet und deren Leistung evaluiert: Kombination der Merkmalsvektoren, Kombination von Klassifikationsergebnissen, z.B. mittels gewichteter Abstimmung, und MK-SVM. Hierbei zeigten alle drei Verfahren ungefähr identische Ergebnisse (ohne Abb., siehe Dyrba et al. 2012, *Table 3* bis *Table 6*).

Die Ergebnisse der dritten Studie sind in Abbildung 13 angegeben. Für die monozentrischen Ruhe-fMRT-Daten erzielte die SVM-Analyse eine Erkennungsrate von 74%, für DTI-Kennwerte 85% und für Volumen der grauen Substanz 81% (Abbildung 13). Die multimodale Auswertung mittels MK-SVM erzielte eine Erkennungsrate von 79% (mit Ruhe-fMRT) bzw. 85% (ohne Ruhe-fMRT).

Mit einem Jaccard-Koeffizient *J* von über 0.74 für die Gruppentrennung von Demenzpatienten oder prodromalen Alzheimer-Patienten von Kontrollprobanden zeigten die Verfahren eine relativ hohe Übereinstimmung der Prädiktion im Vergleich der Modalitäten (die Einzelwerte sind in Dyrba et al. 2012, *Table 2* bzw. Dyrba et al. 2015b, *Table 5* angegeben). Bei der Unterscheidung von Patienten mit leichter kognitiver Störung und positivem bzw. negativem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor sank die Übereinstimmung zwischen den Modalitäten deutlich auf $J = 0.40$ bis $J = 0.73$ (Einzelwerte siehe Dyrba et al. 2015a, Abschnitt *Results*). Bei keiner der drei Studien ermöglichte die multimodale Analyse von DTI, anatomischer MRT und ggf. Ruhe-fMRT eine Verbesserung der Erkennungsleistung von leichter kognitiver Störung oder Alzheimer-Demenz gegenüber der jeweils besten Einzelmodalität.

Die Sensitivitätsanalyse der SVM-Modelle zeigte, dass besonders Veränderungen der Fasertrakte des ventralen und posterioren Cingulums, des Fornix und des Striatums bzw. daran angrenzende Bereiche den größten Einfluss auf die Klassifikationsentscheidung der SVM-Modelle haben (Abbildung 15, Studie 1 und 2, Tabellen mit den MNI-Koordinaten und

anatomischen Regionen im Anhang). Die wichtigsten Bereiche mit atrophischen bzw. funktionellen Veränderungen befanden sich ebenfalls im medialen Temporallappen und im posterioren Gyrus cinguli (Abbildung 14, Studie 3).

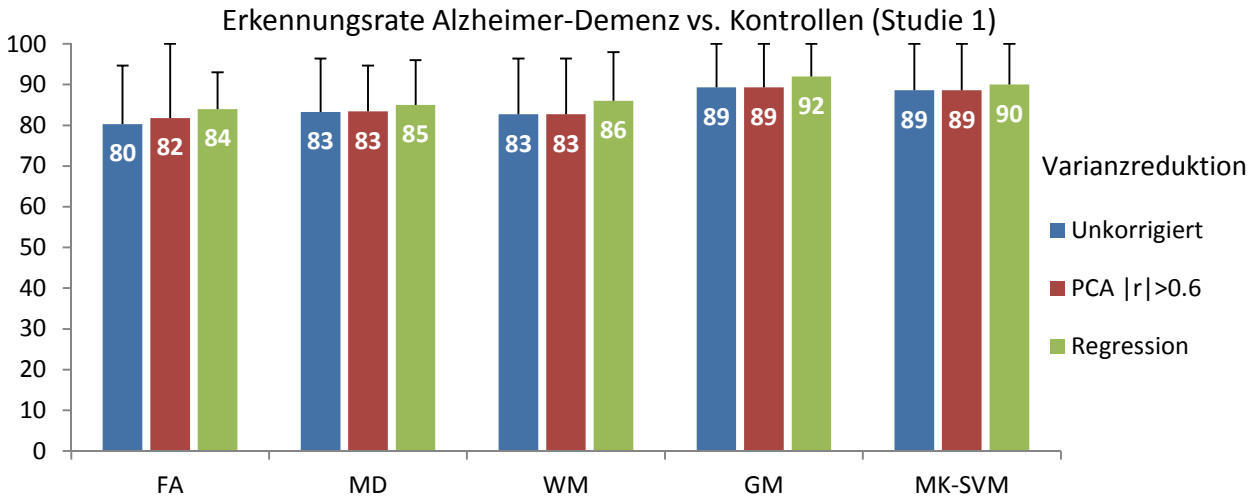

Abbildung 9. Erkennungsraten für die Unterscheidung von Alzheimer-Patienten und Kontrollprobanden

Erkennungsraten für die *gepoolten* Daten der Bildgebungsmodalitäten fraktionelle Anisotropie (FA), mittlere Diffusivität (MD), Volumen der weißen Substanz (WM), Volumen der grauen Substanz (GM) und die multimodale Analyse mittels Multikernel Support Vector Machine (MK-SVM) für die unkorrigierten Originaldaten sowie für die beiden Varianzreduktionsverfahren basieren auf der Hauptkomponentenanalyse (PCA) oder linearer Regression.

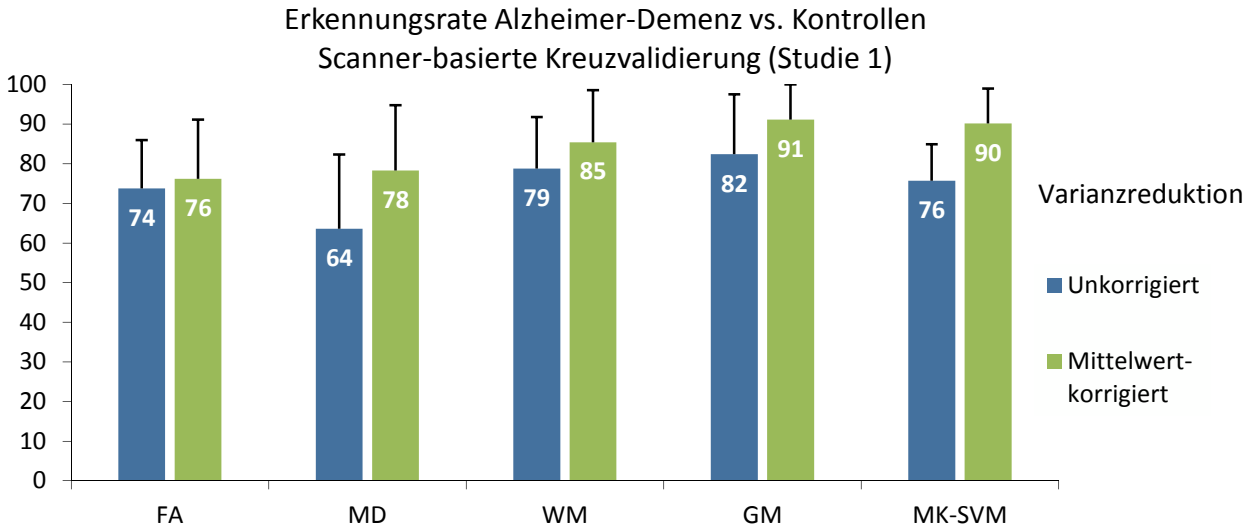

Abbildung 10. Erkennungsraten für die Unterscheidung von Alzheimer-Patienten und Kontrollprobanden (Scanner-basierte Kreuzvalidierung)

Hierbei wurden Daten jeweils eines Scanners als Kreuzvalidierungsfach genutzt, d.h. die Modelle wurden mit den Daten aus allen MRT-Scannern bis auf einen trainiert und die Erkennungsleistung anhand der dem Modell unbekannten Daten evaluiert. Abkürzungen wie oben (Abbildung 9).

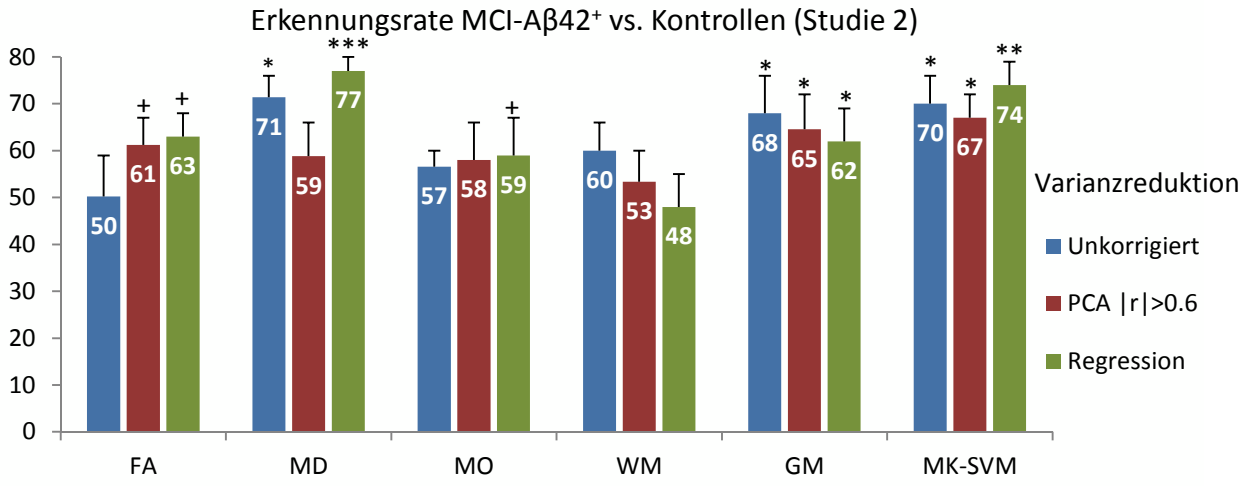

Abbildung 11. Erkennungsraten für die Unterscheidung von Patienten mit leichter kognitiver Störung mit positivem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor (MCI-Aβ42^{+}) und gesunden Kontrollprobanden

Abkürzungen wie oben (Abbildung 9). Die Werte sind signifikant besser als reine Zufallsbefunde: * mit p<0.05, ** mit p<0.01, *** p<0.001, $^{+}$ Trend mit p<0.1, ermittelt durch Permutationstests.

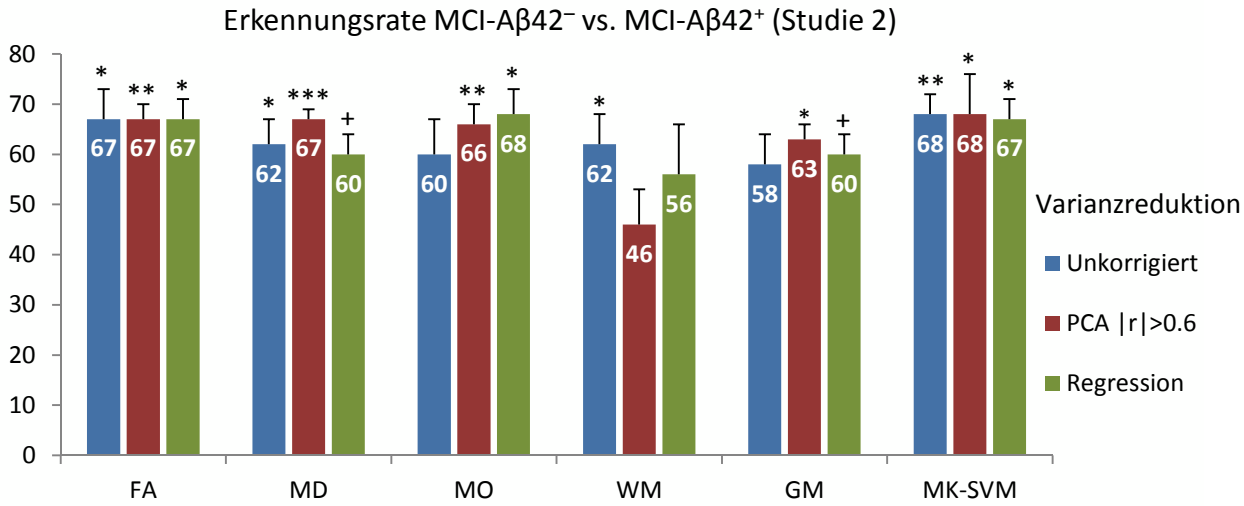

Abbildung 12. Erkennungsraten für die Unterscheidung von Patienten mit leichter kognitiver Störung mit negativem Befund für Amyloid-$\beta_{1\text{-}42}$ im Liquor (MCI-Aβ42^{-}) von Patienten mit positivem Befund (MCI-Aβ42^{+})

Abkürzungen wie oben (Abbildung 9 und Abbildung 11).

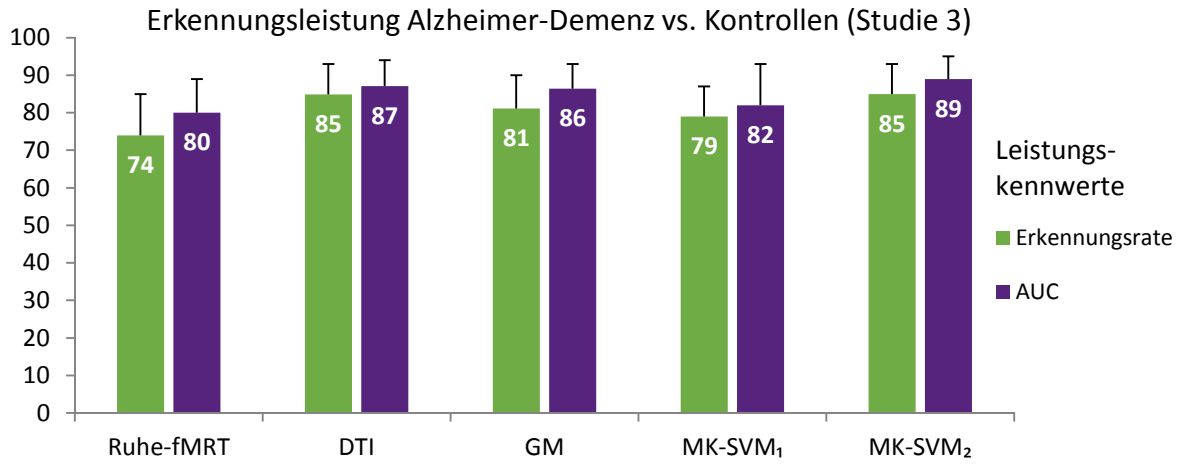

Abbildung 13. Erkennungsleistung für die Unterscheidung von Alzheimer-Patienten und gesunden Kontrollprobanden (monozentrisch)

Erkennungsrate und Fläche unter der Kurve (AUC) für die graphentheoretischen Maße *gewichteter lokaler Clustering-Koeffizient* und *kürzeste gewichtete Pfadlänge* für die Ruhe-fMRT-Daten; für die Kennwerte fraktionelle Anisotropie, Modus der Anisotropie und mittlere Diffusivität der DTI; für das Volumen der grauen Substanz (GM) und für die multimodale Analyse mittels Multikernel Support Vector Machine *mit* ($MK\text{-}SVM_1$) bzw. *ohne* ($MK\text{-}SVM_2$) Verwendung der Ruhe-fMRT-Daten.

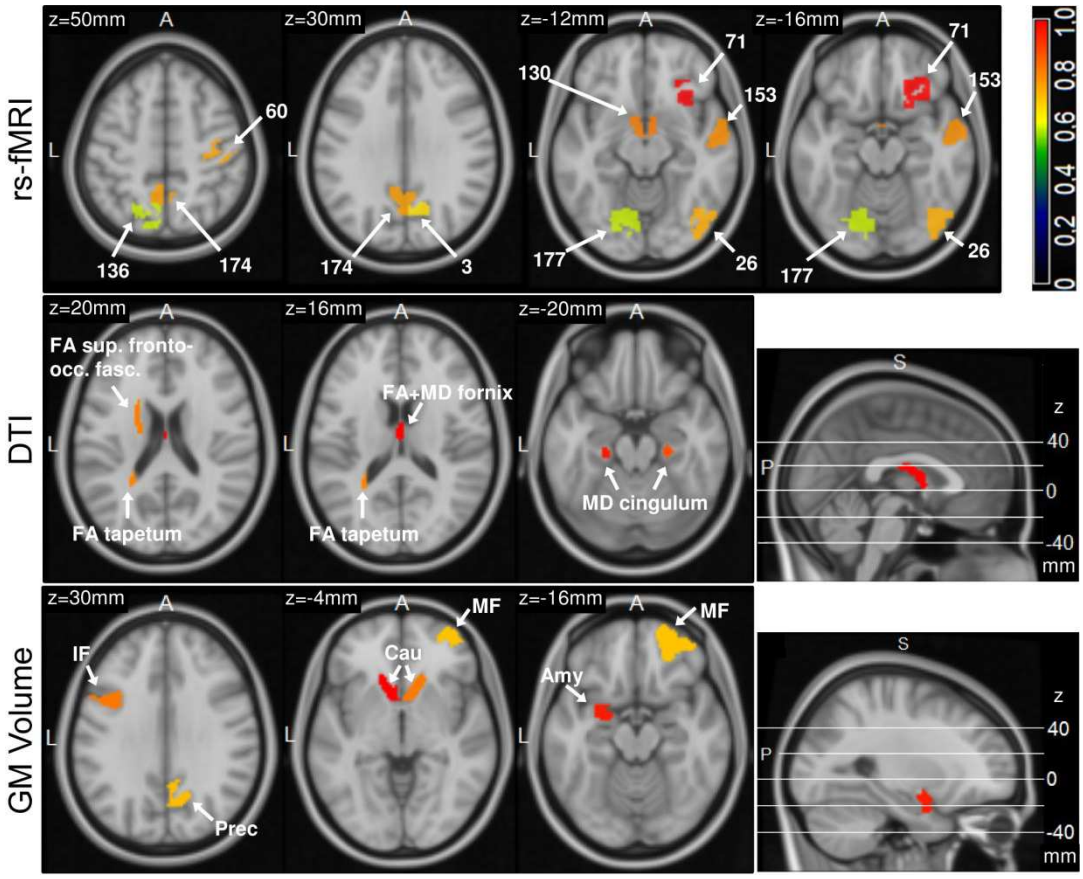

Abbildung 14. SVM Sensitivitätskarten für die von Patienten mit Alzheimer-Demenz und Kontrollprobanden (monozentrisch)

Abkürzungen: rs-fMRI – funktionelle MRT im Ruhezustand, DTI – Diffusions-Tensor-Bildgebung, →

GM Volume – Volumen der grauen Substanz, 71 – insula & gyrus frontalis inferior, 130 – nucleus caudatus, 153 – gyrus temporalis superior & medius, 174 – precuneus, 60 – gyrus praecentralis & postcentralis, 26 – gyrus temporalis inferior & occipitalis, 41 – cerebellum, 3 – precuneus & cuneus, 136 – precuneus, 177 – gyrus lingualis, FA – fraktionelle Anisotropie, MD – mittlere Diffusivität, sup. fronto-occ. fasc. – fasciculus fronto-occipital superior, IF – gyrus frontalis inferior, Prec – precuneus, Cau – nucleus caudatus, MF – gyrus frontalis medius, Amy – amygdala.

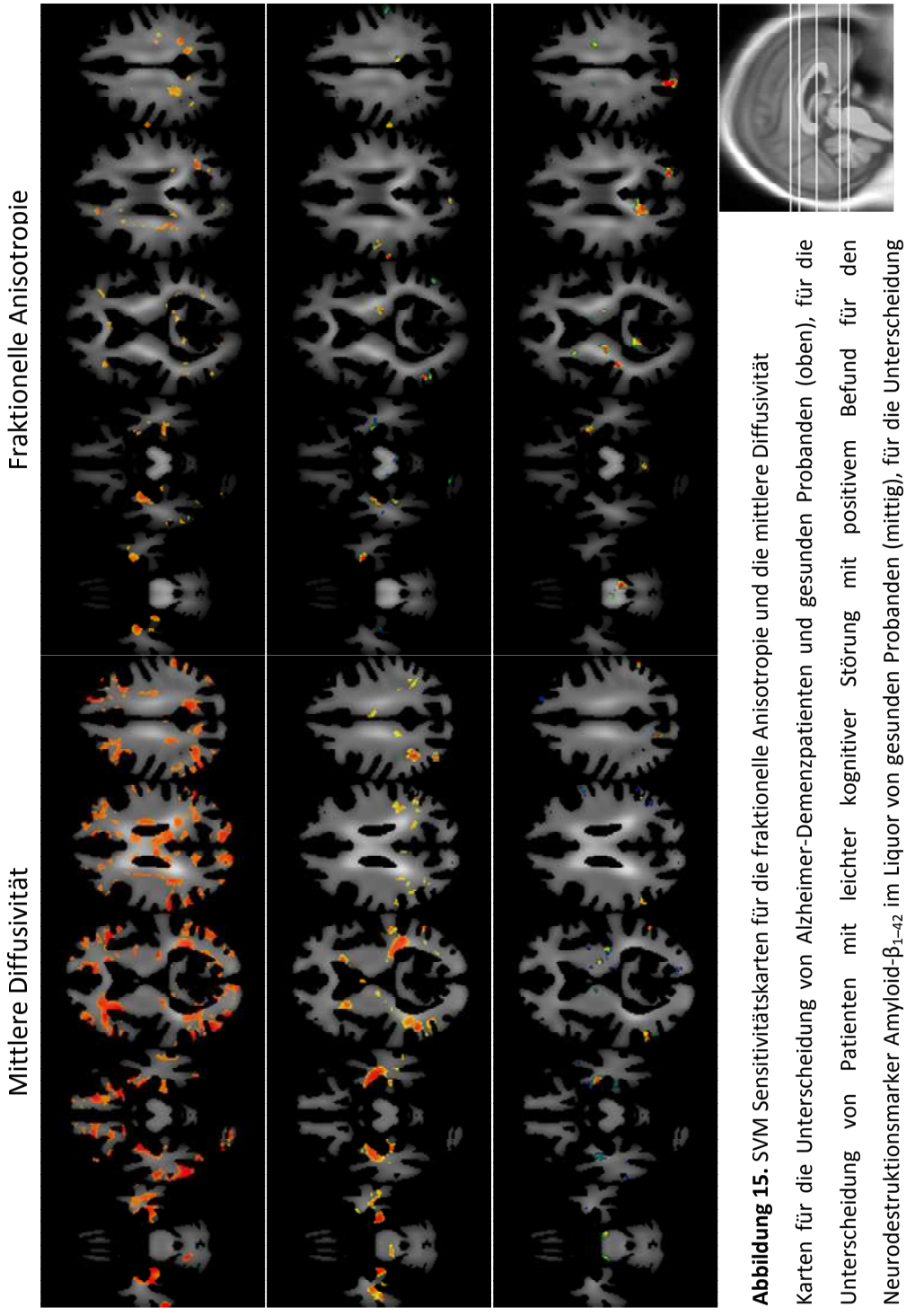

Abbildung 15. SVM Sensitivitätskarten für die fraktionelle Anisotropie und die mittlere Diffusivität
Karten für die Unterscheidung von Alzheimer-Demenzpatienten und gesunden Probanden (oben), für die Unterscheidung von Patienten mit leichter kognitiver Störung mit positivem Befund für den Neurodestruktionsmarker Amyloid-β_{1-42} im Liquor von gesunden Probanden (mittig), für die Unterscheidung von Patienten mit leichter kognitiver Störung mit positivem Befund für den Neurodestruktionsmarker Amyloid-β_{1-42} im Liquor von Patienten mit leichter kognitiver Störung und negativem Befund (unten).
Schichten: z = -26 mm, -16 mm, 8 mm, 26 mm, 34 mm, MNI Referenzsystem.

4 Diskussion

4.1 Diskussion der Bilddatenvorverarbeitung und Datenaufbereitung

Zur Verarbeitung der Bilddaten mit maschinellen Lernverfahren wurden eigene Softwarebibliotheken implementiert. Dazu wurde soweit wie möglich auf vorhandene Bibliotheken zurückgegriffen, z.B. Matlab-Funktionen für die Hauptkomponentenanalyse oder die Regression, WEKA [94] und LibSVM [100] zur Bereitstellung der Lernalgorithmen. Die Erstellung einer eigenen Bibliothek war erforderlich, weil vorhandene Ansätze [101-104] nur rudimentär implementiert [101, 103, 104] oder ineffizient umgesetzt wurden [101, 102, 104]. Ebenso wurde bei der Umsetzung dieser Ansätze nicht auf die Kriterien der Reproduzierbarkeit oder die umfangreiche Dokumentation der Verarbeitungsschritte geachtet, die nach heutigen Maßstäben zur guten wissenschaftlichen Praxis gehören [105]. Die im Rahmen dieser Dissertation erarbeiteten Softwarebibliotheken haben keine grafische Oberfläche und werden über die Kommandozeile bedient. Alle Befehle zur Ein- und Ausgabe wurden in sequenziell aufzurufenden Skriptdateien hinterlegt. Damit wurde zum einen die Reproduzierbarkeit und Dokumentation der Verarbeitung sichergestellt, zum anderen ließ sich dadurch die Verarbeitung der Daten auf hochperformanten Rechencluster auslagern und parallel ausführen, z.B. die Parametersuche und Evaluation der Erkennungsleistung. Weiterhin wurden sämtliche Zwischenergebnisse der Verarbeitung gespeichert, z.B. die Zuordnung der Probanden in die Gruppen zur Evaluation, die selektierten Regionen der Merkmalsauswahl, Ergebnisse der Varianzreduktion, damit mögliche Fehler in der Verarbeitung schnell entdeckt und behoben werden konnten.

4.2 Bewertung der Leistung der Lernverfahren

Bei der Unterscheidung von Alzheimer-Demenzpatienten und gesunden Probanden mit bis zu 85% Erkennungsrate anhand von multizentrischen DTI-Daten erbrachte das Lernverfahren SVM eine leicht geringere Leistung als die in früheren monozentrischen Studien berichteten Ergebnisse in der Größenordnung von etwa 90% [57, 60, 64]. Die Gruppentrennung von bis zu 92% anhand multizentrischer anatomischer MRT-Daten entsprach dagegen den monozentrischen Ergebnissen in der Literatur [54, 55]. Die Scanner-basierte Kreuzvalidierung zeigte erwartungsgemäß eine größere Differenz zwischen den Erkennungsraten von DTI (78%) und MRT (91%). Dieses Ergebnis ist konsistent mit der Vorstudie von Teipel, Reuter [45], die in einer multizentrischen Phantomstudie einen um etwa 50% höheren

Variationskoeffizienten für die DTI erhielten als für das Volumen der grauen Substanz. Bei der Gruppentrennung von Patienten mit leichter kognitiver Störung und positivem Befund für Amyloid-$\beta_{1–42}$ im Liquor und gesunden Probanden lag die Erkennungsrate von bis zu 77% ebenfalls leicht unter den in der Literatur beschriebenen 80-90% [57, 62-64]. Die Erkennungsrate des DTI-Kennwerts *mittlere Diffusivität* war in diesem Fall signifikant höher als die des Volumens der grauen Substanz mit 62% (regressionsbasierte Varianzreduktion), wobei letztere den in der Literatur genannten Ergebnissen von 60-75% entsprach [58, 61, 106]. Die Gruppentrennung von Patienten mit leichter kognitiver Störung und positivem bzw. negativem Befund für Amyloid-$\beta_{1–42}$ im Liquor war mit bis zu 68% ebenso leicht unter den 70-80%, die in der Literatur für monozentrische Stichproben angegeben wurden [63, 64].

Die Kombination von DTI, MRT und ggf. Ruhe-fMRT führte in keiner der Studien zu einer Verbesserung der Leistung im Vergleich zu den Ergebnissen für die jeweils beste Einzelmodalität. In der Literatur sind hierzu bisher nur sehr wenige Studien zu finden: Wee, Yap [62] erreichten für Ruhe-fMRT (79% AUC) und DTI (94%) eine leichte Verbesserung auf 95% AUC für die multimodale Gruppentrennung von Patienten mit leichter kognitiver Störung und Kontrollprobanden mittels MK-SVM. Zhang, Schuff [106] kombinierten DTI (79% AUC) und MRT (67%) und erzielten mittels logistischer Regression eine Leistung von 82% AUC zur Trennung der gleichen Gruppen. Die größte relative Verbesserung von 7% berichteten Cui, Wen [61], jedoch waren in dieser Studie die Leistungskennwerte eher auf einem niedrigen Niveau (DTI: 61% AUC, MRT: 63%, kombiniert: 70%, Patienten mit leichter kognitiver Störung vs. Kontrollen).

Die Sensitivitätsanalyse der SVM-Modelle ermittelte Veränderungen der Fasertrakte des medialen Temporallappens, des Fornix, des Striatums, des Precuneus und des posterioren Cingulums als besonders informativ bezüglich der Gruppentrennung. Genau diese Bereiche sind in der Literatur als diejenigen mit dem stärksten statistischen Gruppenunterschied angegeben worden [84, 107-110].

4.3 Beurteilung der Varianzreduktion

Bei der Unterscheidung von Patienten mit manifester Alzheimer-Demenz und gesunden Probanden gab es eine leichte numerische, aber keine statistisch signifikante Verbesserung der Erkennungsrate durch die Verwendung von Verfahren zur Reduktion der Störvarianz. Bei

der Unterscheidung von Patienten mit leichter kognitiver Störung und positivem bzw. negativem Befund für Amyloid-$\beta_{1\text{-}42}$ zeigte sich ebenfalls nur ein geringer Unterschied. Die größte Differenz zwischen den Verfahren war bei der Gruppentrennung von Patienten mit leichter kognitiver Störung und gesunden Probanden (*gepoolte* Daten) bzw. bei der Gruppentrennung von Demenz-Patienten und Kontrollen (Scanner-basierte Kreuzvalidierung) zu verzeichnen. Insgesamt führte die Varianzreduktion mittels linearer Regression bzw. die daran angelehnte Mittelwertkorrektur zu besseren Ergebnissen, vor allem bei den DTI-Kennwerten, als im Vergleich dazu die Varianzreduktion mittels der Hauptkomponentenanalyse. Die geringere SVM-Modellkomplexität der regressionsbereinigten Daten im Vergleich zu den unkorrigierten bzw. mit der Hauptkomponentenanalyse verarbeiteten Daten weist auf ein besseres Verhältnis von Gruppenunterschied zu Störvarianz unter Verwendung dieses Verfahrens hin. Die unerwartete Abnahme der Erkennungsrate unter Verwendung der Varianzreduktionsverfahren bei den MRT-Daten (Studie 2) könnte möglicherweise durch Interaktionseffekte von Alter, Scanner und Diagnose bzw. einer Überanpassung der statistischen Modelle verursacht worden sein.

4.4 Ausblick

Die bisherigen Ergebnisse zeigten eine gute Erkennungsleistung der Alzheimer-Demenz anhand multizentrischer DTI-Daten, die leicht unter der Erkennungsleistung anhand von multizentrischen volumetrischen Daten lagen. Für die Erkennung von leichter kognitiver Störung bzw. für die Abgrenzung von Patienten mit leichter kognitiver Störung und positivem Befund für den Neurodestruktionsmarker Amyloid-$\beta_{1–42}$ im Liquor von Patienten mit negativem Befund zeigten sich die DTI den volumetrischen Daten überlegen. Damit erwies sich die multizentrische DTI als vielversprechender Bildgebungsmarker für die Alzheimer-Krankheit, der insbesondere als diagnostisches Instrument zur Frühdiagnose im Prodromalstadium der Krankheit geeignet wäre oder als Kennwert in Interventionsstudien genutzt werden könnte. Aufbauend darauf stellt sich jedoch die Frage, welchen prädiktiven Nutzen die multizentrische DTI für die Progression von leichter kognitiver Störung zur Alzheimer-Demenz hat. Diese Fragestellung wurde in einer Folgestudie der EDSD untersucht (Brüggen, Dyrba et al., in Revision): Beim Vergleich von Patienten mit leichter kognitiver Störung, die in einem Zeitfenster von sechs bis 24 Monaten zur Demenz konvertierten,

gegenüber Patienten, die innerhalb von 24 Monaten kognitiv stabil blieben, zeigte sich eine geringere Trennstärke für die DTI als für das Volumen der Hippocampi oder für das Volumen des basalen Vorderhirns. Weitere Studien könnten den Nutzen der multimodalen Auswertung von DTI, volumetrischen Kennwerten und z.B. Glukosemetabolismus für die Prädiktion der Konversion von leichter kognitiver Störung zur Alzheimer-Demenz untersuchen.

5 Publikationen

Die folgenden Publikationen sind Bestandteil dieser Dissertation und im Anhang abgedruckt.

Studie 1:

Dyrba M, Ewers M, Wegrzyn M, Kilimann I, Plant C, Oswald A, Meindl T, Pievani M, Bokde ALW, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group (2013).
Robust Automated Detection of Microstructural White Matter Degeneration in Alzheimer's Disease Using Machine Learning Classification of Multicenter DTI Data.
PLoS ONE 8: e64925. doi: 10.1371/journal.pone.0064925 (IF 4,09)

Dyrba M, Ewers M, Wegrzyn M, Kilimann I, Plant C, Oswald A, Meindl T, Pievani M, Bokde A, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group (2012).
Combining DTI and MRI for the automated detection of Alzheimer's disease using a large European multicenter dataset.
In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L. (Eds.), *Proceedings of the 3rd Workshop on Multimodal Brain Image Analysis*. Springer Berlin/Heidelberg, pp. 18–28. ISBN: 978-3-642-33529-7, doi: 10.1007/978-3-642-33530-3_2

Studie 2:

Dyrba M, Barkhof F, Fellgiebel A, Filippi M, Hausner L, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group (2015a).
Predicting prodromal Alzheimer's disease in subjects with mild cognitive impairment using machine learning classification of multimodal multicenter diffusion-tensor and magnetic resonance imaging data.
Journal of Neuroimaging 25 (5):738–747. doi: 10.1111/jon.12214 (IF 1,82)

Studie 3:

Dyrba M, Grothe M, Kirste T, Teipel SJ (2015b).
Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM.
Human Brain Mapping 36 (6): 2118–2131. doi: 10.1002/hbm.22759 (IF 6,92)

6 Literaturverzeichnis

1. Wallesch C-W, Förstl H. *Demenzen*. 2nd ed: Thieme; 2012. 416 p.
2. Brookmeyer R, Evans DA, Hebert L, Langa KM, Heeringa SG, Plassman BL, Kukull WA (2011). *National estimates of the prevalence of Alzheimer's disease in the United States*. Alzheimer's & Dementia 7(1):61–73. doi: 10.1016/j.jalz.2010.11.007.
3. Schmidtke K, Otto M. *Alzheimer-Demenz*. In: Wallesch C-W, Förstl H, editors. *Demenzen*. 2nd ed: Thieme; 2012. p. 203–227.
4. Wippold FJ, Cairns N, Vo K, Holtzman DM, Morris JC (2008). *Neuropathology for the Neuroradiologist: Plaques and Tangles*. American Journal of Neuroradiology 29(1):18–22. doi: 10.3174/ajnr.A0781.
5. Castello MA, Soriano S (2014). *On the origin of Alzheimer's disease. Trials and tribulations of the amyloid hypothesis*. Ageing research reviews 13:10–12. doi: 10.1016/j.arr.2013.10.001.
6. Drachman DA (2014). *The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease*. Alzheimer's & dementia : the journal of the Alzheimer's Association 10(3):372–380. doi: 10.1016/j.jalz.2013.11.003.
7. Mattson MP (2004). *Pathways towards and away from Alzheimer's disease*. Nature 430(7000):631–639. doi: 10.1038/nature02621.
8. Götz J, Chen F, van Dorpe J, Nitsch RM (2001). *Formation of neurofibrillary tangles in P301l tau transgenic mice induced by Abeta 42 fibrils*. Science 293(5534):1491–1495. doi: 10.1126/science.1062097.
9. Blennow K, Dubois B, Fagan AM, Lewczuk P, de Leon MJ, Hampel H (2015). *Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease*. Alzheimer's & Dementia 11(1):58–69. doi: 10.1016/j.jalz.2014.02.004.
10. Andreasen N, Minthon L, Davidsson P, Vanmechelen E, Vanderstichele H, Winblad B, Blennow K (2001). *Evaluation of CSF-tau and CSF-Aβ42 as Diagnostic Markers for Alzheimer Disease in Clinical Practice*. Archives of Neurology 58(3):373–379. doi: 10.1001/archneur.58.3.373.
11. Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, Herukka S-K, van der Flier WM, Blankenstein MA, Ewers M, Rich K, Kaiser E, Verbeek M, Tsolaki M, Mulugeta E, Rosén E, Aarsland D, Visser PJ, Schröder J, Marcusson J, de Leon M, Hampel H, Scheltens P, Pirttilä T, Wallin A, Jönhagen ME, Minthon L, Winblad B, Blennow K (2009). *CSF Biomarkers and Incipient Alzheimer Disease in Patients With Mild Cognitive Impairment*. Jama 302(4):385–393. doi: 10.1001/jama.2009.1064.
12. Chételat G (2013). *Alzheimer disease: Aβ-independent processes—rethinking preclinical AD*. Nature Reviews Neurology 9(3):123–124. doi: 10.1038/nrneurol.2013.21.
13. Pohanka M (2013). *Alzheimer´s Disease and Oxidative Stress: A Review*. Current Medicinal Chemistry 21(3):356–364. doi: 10.2174/09298673113206660258.
14. Braak H, Braak E (1991). *Neuropathological stageing of Alzheimer-related changes*. Acta Neuropathologica 82:239–259. doi: 10.1007/bf00308809.
15. Delacourte A, David JP, Sergeant N, Buee L, Wattez A, Vermersch P, Ghozali F, Fallet-Bianco C, Pasquier F, Lebert F, Petit H, Di Menza C (1999). *The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease*. Neurology 52(6):1158–1165. doi: 10.1212/wnl.52.6.1158.

16. Brun A, Englund E (1986). *A white matter disorder in dementia of the Alzheimer type: A pathoanatomical study*. Annals of Neurology 19(3):253–262. doi: 10.1002/ana.410190306.
17. Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ (2013). *Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers*. The Lancet Neurology 12(2):207–216. doi: 10.1016/s1474-4422(12)70291-0.
18. Huk WJ, Gademann G (1984). *Magnetic resonance imaging (MRI): method and early clinical experiences in diseases of the central nervous system*. Neurosurgical review 7(4):259–280.
19. Le Bihan D, Turner R, Douek P, Patronas N (1992). *Diffusion MR imaging: clinical applications*. American Journal of Roentgenology 159(3):591–599. doi: 10.2214/ajr.159.3.1503032.
20. Ido T, Wan CN, Casella V, Fowler JS, Wolf AP, Reivich M, Kuhl DE (1978). *Labeled 2-deoxy-D-glucose analogs. 18F-labeled 2-deoxy-2-fluoro-D-glucose, 2-deoxy-2-fluoro-D-mannose and 14C-2-deoxy-2-fluoro-D-glucose*. Journal of Labelled Compounds and Radiopharmaceuticals 14(2):175–183. doi: 10.1002/jlcr.2580140204.
21. Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, Frölich L, Schönknecht P, Ito K, Mielke R, Kalbe E (2002). *Discrimination between Alzheimer Dementia and Controls by Automated Analysis of Multicenter FDG PET*. NeuroImage 17(1):302–316. doi: 10.1006/nimg.2002.1208.
22. Li Y, Rinne JO, Mosconi L, Pirraglia E, Rusinek H, DeSanti S, Kemppainen N, Någren K, Kim B-C, Tsui W, de Leon MJ (2008). *Regional analysis of FDG and PIB-PET images in normal aging, mild cognitive impairment, and Alzheimer's disease*. European Journal of Nuclear Medicine and Molecular Imaging 35(12):2169–2181. doi: 10.1007/s00259-008-0833-y.
23. Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Bergström M, Savitcheva I, Huang G-F, Estrada S, Ausén B, Debnath ML, Barletta J, Price JC, Sandell J, Lopresti BJ, Wall A, Koivisto P, Antoni G, Mathis CA, Långström B (2004). *Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B*. Annals of Neurology 55(3):306–319. doi: 10.1002/ana.20009.
24. Ogawa S, Lee TM, Kay AR, Tank DW (1990). *Brain magnetic resonance imaging with contrast dependent on blood oxygenation*. Proceedings of the National Academy of Sciences 87(24):9868–9872.
25. Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R (1992). *Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation*. Proceedings of the National Academy of Sciences 89(12):5675–5679. doi: 10.1073/pnas.89.12.5675.
26. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995). *Functional connectivity in the motor cortex of resting human brain using echo-planar MRI*. Magn Reson Med 34(4):537–541.
27. Fox MD (2005). *The human brain is intrinsically organized into dynamic, anticorrelated functional networks*. Proceedings of the National Academy of Sciences 102(27):9673–9678. doi: 10.1073/pnas.0504136102.
28. Biswal BB, Kylen JV, Hyde JS (1997). *Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps*. NMR in Biomedicine 10(4/5):165–170. doi: 10.1002/(sici)1099-1492(199706/08)10:4/5<165::aid-nbm454>3.0.co;2-7.

29. Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2001). *Frequencies Contributing to Functional Connectivity in the Cerebral Cortex in "Resting-state" Data*. American Journal of Neuroradiology 22(7):1326–1333.
30. Sorg C, Riedl V, Muhlau M, Calhoun VD, Eichele T, Laer L, Drzezga A, Forstl H, Kurz A, Zimmer C, Wohlschlager AM (2007). *Selective changes of resting-state networks in individuals at risk for Alzheimer's disease*. Proceedings of the National Academy of Sciences 104(47):18760–18765. doi: 10.1073/pnas.0708803104.
31. Supekar K, Menon V, Rubin D, Musen M, Greicius MD, Sporns O (2008). *Network Analysis of Intrinsic Functional Brain Connectivity in Alzheimer's Disease*. PLoS Computational Biology 4(6):e1000100. doi: 10.1371/journal.pcbi.1000100.
32. Sanz-Arigita EJ, Schoonheim MM, Damoiseaux JS, Rombouts SARB, Maris E, Barkhof F, Scheltens P, Stam CJ, S. Breitner JC (2010). *Loss of 'Small-World' Networks in Alzheimer's Disease: Graph Analysis of fMRI Resting-State Functional Connectivity*. PLoS One 5(11):e13788. doi: 10.1371/journal.pone.0013788.
33. Chen G, Ward BD, Xie C, Li W, Wu Z, Jones JL, Franczak M, Antuono P, Li S-J (2011). *Classification of Alzheimer Disease, Mild Cognitive Impairment, and Normal Cognitive Status with Large-Scale Network Analysis Based on Resting-State Functional MR Imaging*. Radiology 259(1):213–221. doi: 10.1148/radiol.10100734.
34. Petrella JR, Sheldon FC, Prince SE, Calhoun VD, Doraiswamy PM (2011). *Default mode network connectivity in stable vs progressive mild cognitive impairment*. Neurology 76(6):511–517. doi: 10.1212/WNL.0b013e31820af94e.
35. Koch W, Teipel S, Mueller S, Benninghoff J, Wagner M, Bokde ALW, Hampel H, Coates U, Reiser M, Meindl T (2012). *Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease*. Neurobiology of Aging 33(3):466–478. doi: 10.1016/j.neurobiolaging.2010.04.013.
36. Brier MR, Thomas JB, Snyder AZ, Benzinger TL, Zhang D, Raichle ME, Holtzman DM, Morris JC, Ances BM (2012). *Loss of Intranetwork and Internetwork Resting State Functional Connections with Alzheimer's Disease Progression*. Journal of Neuroscience 32(26):8890–8899. doi: 10.1523/jneurosci.5698-11.2012.
37. Soldner J, Meindl T, Koch W, Bokde ALW, Reiser MF, Möller H-J, Bürger K, Hampel H, Teipel SJ (2012). *Strukturelle und funktionelle neuronale Konnektivität bei der Alzheimer-Krankheit*. Der Nervenarzt 83(7):878–887. doi: 10.1007/s00115-011-3326-3.
38. Watts DJ, Strogatz SH (1998). *Collective dynamics of 'small-world' networks*. Nature 393(6684):440–442. doi: 10.1038/30918.
39. Stam C, Jones B, Nolte G, Breakspear M, Scheltens P (2006). *Small-World Networks and Functional Connectivity in Alzheimer's Disease*. Cereb Cortex 17(1):92–99. doi: 10.1093/cercor/bhj127.
40. van den Heuvel MP, Stam CJ, Boersma M, Hulshoff Pol HE (2008). *Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain*. NeuroImage 43(3):528–539. doi: 10.1016/j.neuroimage.2008.08.010.
41. Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, Hedden T, Andrews-Hanna JR, Sperling RA, Johnson KA (2009). *Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease*. Journal of Neuroscience 29(6):1860–1873. doi: 10.1523/jneurosci.5062-08.2009.
42. Rubinov M, Sporns O (2010). *Complex network measures of brain connectivity: Uses and interpretations*. NeuroImage 52(3):1059–1069. doi: 10.1016/j.neuroimage.2009.10.003.

43. Zhao X, Liu Y, Wang X, Liu B, Xi Q, Guo Q, Jiang H, Jiang T, Wang P (2012). *Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state FMRI study*. PLoS One 7(3):e33540. doi: 10.1371/journal.pone.0033540.
44. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen PC, Trojanowski JQ, Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ (2010). *Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade*. The Lancet Neurology 9(1):119–128. doi: 10.1016/s1474-4422(09)70299-6.
45. Teipel SJ, Reuter S, Stieltjes B, Acosta-Cabronero J, Ernemann U, Fellgiebel A, Filippi M, Frisoni G, Hentschel F, Jessen F, Klöppel S, Meindl T, Pouwels PJW, Hauenstein K-H, Hampel H (2011). *Multicenter stability of diffusion tensor imaging measures: A European clinical and physical phantom study*. Psychiatry Research: Neuroimaging 194(3):363–371. doi: 10.1016/j.pscychresns.2011.05.012.
46. Vollmar C, O'Muircheartaigh J, Barker GJ, Symms MR, Thompson P, Kumari V, Duncan JS, Richardson MP, Koepp MJ (2010). *Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners*. NeuroImage 51(4):1384–1394. doi: 10.1016/j.neuroimage.2010.03.046.
47. Teipel SJ, Buchert R, Thome J, Hampel H, Pahnke J (2011). *Development of Alzheimer-disease neuroimaging-biomarkers using mouse models with amyloid-precursor protein-transgene expression*. Progress in Neurobiology 95(4):547–556. doi: 10.1016/j.pneurobio.2011.05.004.
48. Teipel SJ, Wegrzyn M, Meindl T, Frisoni G, Bokde ALW, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Hauenstein K, Ewers M, group tEs (2012). *Anatomical MRI and DTI in the diagnosis of Alzheimer's disease: a European Multicenter Study*. Journal of Alzheimer's Disease 31(Suppl 3):S33–47. doi: 10.3233/jad-2012-112118.
49. Teipel SJ, Grothe MJ, Filippi M, Fellgiebel A, Dyrba M, Frisoni GB, Meindl T, Bokde ALW, Hampel H, Klöppel S, Hauenstein K (2014). *Fractional Anisotropy Changes in Alzheimer's Disease Depend on the Underlying Fiber Tract Architecture: A Multiparametric DTI Study using Joint Independent Component Analysis*. Journal of Alzheimer's Disease 41(1):69–83. Epub 27.2.2014. doi: 10.3233/jad-131829.
50. Fischer FU, Scheurich A, Wegrzyn M, Schermuly I, Bokde ALW, Klöppel S, Pouwels PJW, Teipel S, Yakushev I, Fellgiebel A (2012). *Automated tractography of the cingulate bundle in Alzheimer's disease: A multicenter DTI study*. Journal of Magnetic Resonance Imaging 36(1):84–91. doi: 10.1002/jmri.23621.
51. Kljajevic V, Meyer P, Holzmann C, Dyrba M, Kasper E, Bokde ALW, Fellgiebel A, Meindl T, Hampel H, Teipel S (2014). *The ε4 genotype of apolipoprotein E and white matter integrity in Alzheimer's disease*. Alzheimer's & Dementia 10(3):401–404. doi: 10.1016/j.jalz.2013.02.008.
52. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. New York, NY: Springer New York; 2013. 426 p.
53. Cortes C, Vapnik V (1995). *Support-vector networks*. Machine Learning 20:273–297. doi: 10.1007/bf00994018.
54. Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill R, Rohrer JD, Fox NC, Jack CR, Ashburner J, Frackowiak RSJ (2008). *Automatic classification of MR scans in Alzheimer's disease*. Brain 131(3):681–689. doi: 10.1093/brain/awm319.
55. Plant C, Teipel SJ, Oswald A, Böhm C, Meindl T, Mourão-Miranda J, Bokde AW, Hampel H, Ewers M (2010). *Automated detection of brain atrophy patterns based on MRI for the*

prediction of Alzheimer's disease. NeuroImage 50(1):162–174. doi: 10.1016/j.neuroimage.2009.11.046.
56. Abdulkadir A, Mortamet B, Vemuri P, Jack CR, Krueger G, Klöppel S (2011). *Effects of hardware heterogeneity on the performance of SVM Alzheimer's disease classifier*. NeuroImage 58(3):785–792. doi: 10.1016/j.neuroimage.2011.06.029.
57. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert M-O, Chupin M, Benali H, Colliot O (2011). *Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database*. NeuroImage 56(2):766–781. doi: 10.1016/j.neuroimage.2010.06.013.
58. Wee C-Y, Yap P-T, Shen D, for the Alzheimer's Disease Neuroimaging I (2013). *Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns*. Human Brain Mapping 34(12):3411–3425. doi: 10.1002/hbm.22156.
59. Haller S, Nguyen D, Rodriguez C, Emch J, Gold G, Bartsch A, Lovblad KO, Giannakopoulos P (2010). *Individual Prediction of Cognitive Decline in Mild Cognitive Impairment Using Support Vector Machine-Based Analysis of Diffusion Tensor Imaging Data*. Journal of Alzheimer's Disease 22(1):315–327. doi: 10.3233/jad-2010-100840.
60. Graña M, Termenon M, Savio A, Gonzalez-Pinto A, Echeveste J, Pérez JM, Besga A (2011). *Computer Aided Diagnosis system for Alzheimer Disease using brain Diffusion Tensor Imaging features selected by Pearson's correlation*. Neurosci Lett 502(3):225–229. doi: 10.1016/j.neulet.2011.07.049.
61. Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, Luo S, Zhu W, Kochan NA, Reppermund S, Zhuang L, Raamana PR, Liu T, Trollor JN, Wang L, Brodaty H, Sachdev PS (2012). *Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: A combined spatial atrophy and white matter alteration approach*. NeuroImage 59(2):1209–1217. doi: 10.1016/j.neuroimage.2011.08.013.
62. Wee C-Y, Yap P-T, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D (2012). *Identification of MCI individuals using structural and functional connectivity networks*. NeuroImage 59(3):2045–2056. doi: 10.1016/j.neuroimage.2011.10.015.
63. O'Dwyer L, Lamberton F, Bokde ALW, Ewers M, Faluyi YO, Tanner C, Mazoyer B, O'Neill D, Bartley M, Collins DR, Coughlan T, Prvulovic D, Hampel H, Zhan W (2012). *Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment*. PLoS One 7(2):e32441. doi: 10.1371/journal.pone.0032441.
64. Shao J, Myers N, Yang Q, Feng J, Plant C, Böhm C, Förstl H, Kurz A, Zimmer C, Meng C, Riedl V, Wohlschläger A, Sorg C (2012). *Prediction of Alzheimer's disease using individual structural connectivity networks*. Neurobiology of Aging 33(12):2756–2765. doi: 10.1016/j.neurobiolaging.2012.01.017.
65. Zhang D, Wang Y, Zhou L, Yuan H, Shen D (2011). *Multimodal classification of Alzheimer's disease and mild cognitive impairment*. NeuroImage 55(3):856–867. doi: 10.1016/j.neuroimage.2011.01.008.
66. Dukart J, Mueller K, Horstmann A, Barthel H, Möller HE, Villringer A, Sabri O, Schroeter ML (2011). *Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia*. PLoS One 6(3):e18111. doi: 10.1371/journal.pone.0018111.
67. Hinrichs C, Singh V, Xu G, Johnson SC (2011). *Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population*. NeuroImage 55(2):574–589. doi: 10.1016/j.neuroimage.2010.10.081.

68. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C (1989). *The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease*. Neurology 39(9):1159–1165.
69. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, Luca Md, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, Stefano Nd, Brady JM, Matthews PM (2004). *Advances in functional and structural MR image analysis and implementation as FSL*. NeuroImage 23(Suppl 1):S208–219. doi: 10.1016/j.neuroimage.2004.07.051.
70. Friston KJ, Ashburner J, Kiebel S, Nichols T, Penny WD. *Statistical parametric mapping: The analysis of funtional brain images*. First edition ed. Amsterdam; Boston: Elsevier/Academic Press; 2007.
71. Gaser C, Volz H-P, Kiebel S, Riehemann S, Sauer H (1999). *Detecting Structural Changes in Whole Brain Based on Nonlinear Deformations—Application to Schizophrenia Research*. NeuroImage 10(2):107–113. doi: 10.1006/nimg.1999.0458.
72. Chao-Gan Y, Yu-Feng Z (2010). *DPARSF: A MATLAB Toolbox for "Pipeline" Data Analysis of Resting-State fMRI*. Front Syst Neurosci 4(13):1–7. doi: 10.3389/fnsys.2010.00013.
73. Ashburner J (2007). *A fast diffeomorphic image registration algorithm*. NeuroImage 38(1):95–113. doi: 10.1016/j.neuroimage.2007.07.007.
74. Ennis DB, Kindlmann G (2006). *Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images*. Magnetic Resonance in Medicine 55(1):136–146. doi: 10.1002/mrm.20741.
75. Zhou J, Gennatas Efstathios D, Kramer Joel H, Miller Bruce L, Seeley William W (2012). *Predicting Regional Neurodegeneration from the Healthy Brain Functional Connectome*. Neuron 73(6):1216–1227. doi: 10.1016/j.neuron.2012.03.004.
76. Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS (2012). *A whole brain fMRI atlas generated via spatially constrained spectral clustering*. Human Brain Mapping 33(8):1914–1928. doi: 10.1002/hbm.21333.
77. Kohavi R. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. *IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence*1995. p. 1137–1143.
78. Onnela J-P, Saramäki J, Kertész J, Kaski K (2005). *Intensity and coherence of motifs in weighted complex networks*. Physical Review E 71(6):065103.065101–065104. doi: 10.1103/PhysRevE.71.065103.
79. Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K, Hua K, Faria AV, Mahmood A, Woods R, Toga AW, Pike GB, Neto PR, Evans A, Zhang J, Huang H, Miller MI, van Zijl P, Mazziotta J (2008). *Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template*. NeuroImage 40(2):570–582. doi: 10.1016/j.neuroimage.2007.12.035.
80. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002). *Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain*. NeuroImage 15(1):273–289. doi: 10.1006/nimg.2001.0978.
81. Hsu J-L, Leemans A, Bai C-H, Lee C-H, Tsai Y-F, Chiu H-C, Chen W-H (2008). *Gender differences and age-related white matter changes of the human brain: A diffusion tensor imaging study*. NeuroImage 39(2):566–577. doi: 10.1016/j.neuroimage.2007.09.017.
82. Barnes J, Ridgway GR, Bartlett J, Henley SMD, Lehmann M, Hobbs N, Clarkson MJ, MacManus DG, Ourselin S, Fox NC (2010). *Head size, age and gender adjustment in MRI*

studies: a necessary nuisance? NeuroImage 53(4):1244–1255. doi: 10.1016/j.neuroimage.2010.06.025.
83. Hsu J-L, Van Hecke W, Bai C-H, Lee C-H, Tsai Y-F, Chiu H-C, Jaw F-S, Hsu C-Y, Leu J-G, Chen W-H, Leemans A (2010). *Microstructural white matter changes in normal aging: A diffusion tensor imaging study with higher-order polynomial regression models*. NeuroImage 49(1):32–43. doi: 10.1016/j.neuroimage.2009.08.031.
84. Teipel SJ, Meindl T, Wagner M, Stieltjes B, Reuter S, Hauenstein K-H, Filippi M, Ernemann U, Reiser MF, Hampel H (2010). *Longitudinal Changes in Fiber Tract Integrity in Healthy Aging and Mild Cognitive Impairment: A DTI Follow-Up Study*. Journal of Alzheimer's Disease 22(2):507–522. doi: 10.3233/jad-2010-100234.
85. Dukart J, Kherif F, Mueller K, Adaszewski S, Schroeter ML, Frackowiak RSJ, Draganski B (2013). *Generative FDG-PET and MRI Model of Aging and Disease Progression in Alzheimer's Disease*. PLoS Computational Biology 9(4):e1002987. doi: 10.1371/journal.pcbi.1002987.
86. Apostolova LG, Hwang KS, Kohannim O, Avila D, Elashoff D, Jack CR, Shaw L, Trojanowski JQ, Weiner MW, Thompson PM (2014). *ApoE4 effects on automated diagnostic classifiers for mild cognitive impairment and Alzheimer's disease*. NeuroImage: Clinical 4:461–472. doi: 10.1016/j.nicl.2013.12.012.
87. Sexton CE, Walhovd KB, Storsve AB, Tamnes CK, Westlye LT, Johansen-Berg H, Fjell AM (2014). *Accelerated Changes in White Matter Microstructure during Aging: A Longitudinal Diffusion Tensor Imaging Study*. Journal of Neuroscience 34(46):15425–15436. doi: 10.1523/jneurosci.0203-14.2014.
88. Teipel SJ, Lerche M, Kilimann I, O'Brien K, Grothe M, Meyer P, Li X, Sänger P, Hauenstein K (2014). *Decline of fiber tract integrity over the adult age range: A diffusion spectrum imaging study*. Journal of Magnetic Resonance Imaging 40(2):348–359. doi: 10.1002/jmri.24420.
89. Nenonen M, Hakulinen U, Brander A, Ohman J, Dastidar P, Luoto TM (2015). *Possible confounding factors on cerebral diffusion tensor imaging measurements*. Acta Radiologica Open 4(2):1–9. doi: 10.1177/2047981614546795.
90. Hardy MA, Bryman A, editors. *Handbook of Data Analysis*: SAGE Publications Ltd; 2009.
91. Bortz J, Schuster C. *Statistik für Human- und Sozialwissenschaftler*. 7 ed. Berlin [u.a.]: Springer; 2010. 655 p.
92. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*: The MIT Press; 2005. 248 p.
93. Abdulkadir A, Ronneberger O, Tabrizi SJ, Kloppel S (2014). *Reduction of confounding effects with voxel-wise Gaussian process regression in structural MRI*.1–4. doi: 10.1109/prni.2014.6858505.
94. Witten IH, Frank E, Hall MA. *Data mining: Practical machine learning tools and techniques*. 3rd ed. Burlington, MA: Morgan Kaufmann; 2011. 629 p.
95. Sonnenburg S, Rätsch G, Schäfer C, Schölkopf B (2006). *Large Scale Multiple Kernel Learning*. Journal of Machine Learning Research 7:1531–1565.
96. Good PI. *Permutation tests: A practical guide to resampling methods for testing hypotheses*. 2nd ed ed. New York: Springer; 2000. 270 p.
97. Ojala M, Garriga GC (2010). *Permutation Tests for Studying Classifier Performance*. Journal of Machine Learning Research 11:1833–1863.

98. Rasmussen PM, Madsen KH, Lund TE, Hansen LK (2011). *Visualization of nonlinear kernel models in neuroimaging by sensitivity maps*. NeuroImage 55(3):1120–1131. doi: 10.1016/j.neuroimage.2010.12.035.
99. Smith ED, Szidarovszky F, Karnavas WJ, Bahill AT (2007). *Sensitivity Analysis, a Powerful System Validation Technique*. The Open Cybernetics & Systemics Journal 2(1):39–56. doi: 10.2174/1874110x00802010039.
100. Chang C-C, Lin C-J (2011). *LIBSVM: A library for support vector machines*. ACM Transactions on Intelligent Systems and Technology 2(3):1–27. doi: 10.1145/1961189.1961199.
101. Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby JV, Pollmann S (2009). *PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data*. Neuroinformatics 7(1):37–53. doi: 10.1007/s12021-008-9041-y.
102. Böhm C, Oswald A, Sappelt F. *JGrid/FCC: A Tool for Efficient Knowledge Extraction from MRI Scans*. *22nd International Conference on Database and Expert Systems Applications (DEXA)*2011. p. 81–85.
103. Schrouff J, Rosa MJ, Rondina JM, Marquand AF, Chu C, Ashburner J, Phillips C, Richiardi J, Mourão-Miranda J (2013). *PRoNTo: Pattern Recognition for Neuroimaging Toolbox*. Neuroinformatics 11(3):319–337. doi: 10.1007/s12021-013-9178-1.
104. Hebart MN, Görgen K, Haynes J-D (2015). *The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data*. Frontiers in Neuroinformatics 8(88):1–18. doi: 10.3389/fninf.2014.00088.
105. Begley CG, Ioannidis JPA (2014). *Reproducibility in Science: Improving the Standard for Basic and Preclinical Research*. Circulation Research 116(1):116–126. doi: 10.1161/circresaha.114.303819.
106. Zhang Y, Schuff N, Camacho M, Chao LL, Fletcher TP, Yaffe K, Woolley SC, Madison C, Rosen HJ, Miller BL, Weiner MW (2013). *MRI Markers for Mild Cognitive Impairment: Comparisons between White Matter Integrity and Gray Matter Volume Measurements*. PLoS One 8(6):e66367. doi: 10.1371/journal.pone.0066367.
107. Bozzali M (2002). *White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging*. Journal of Neurology, Neurosurgery & Psychiatry 72(6):742–746. doi: 10.1136/jnnp.72.6.742.
108. Teipel SJ, Stahl R, Dietrich O, Schoenberg SO, Perneczky R, Bokde ALW, Reiser MF, Möller H-J, Hampel H (2007). *Multivariate network analysis of fiber tract integrity in Alzheimer's disease*. NeuroImage 34(3):985–995. doi: 10.1016/j.neuroimage.2006.07.047.
109. Agosta F, Pievani M, Sala S, Geroldi C, Galluzzi S, Frisoni GB, Filippi M (2011). *White Matter Damage in Alzheimer Disease and Its Relationship to Gray Matter Atrophy*. Radiology 258(3):853–863. doi: 10.1148/radiol.10101284.
110. Bosch B, Arenaza-Urquijo EM, Rami L, Sala-Llonch R, Junqué C, Solé-Padullés C, Peña-Gómez C, Bargalló N, Molinuevo JL, Bartrés-Faz D (2012). *Multiple DTI index analysis in normal aging, amnestic MCI and AD. Relationship with neuropsychological performance*. Neurobiology of aging 33(1):61–74. doi: 10.1016/j.neurobiolaging.2010.02.004.

Selbstständigkeitserklärung

Ich erkläre hiermit, das ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Informationen und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Die Regeln zur Sicherung guter wissenschaftlicher Praxis wurden beachtet.

Ich versichere, dass ich für die inhaltliche Erstellung der vorliegenden Arbeit nicht die entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen habe. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Rostock, den 18. Mai 2015

Lebenslauf und wissenschaftlicher Werdegang

Martin Dyrba
Kurt-Schumacher-Ring 116
18146 Rostock

Geboren am 5. August 1985 in Zwenkau

Wissenschaftliche Arbeit

Seit April 2011	Wissenschaftlicher Mitarbeiter des Deutschen Zentrums für Neurodegenerative Erkrankungen e.V. (DZNE), Standort Rostock/ Greifswald Aufgabengebiet: Analyse komplexer Diagnose- und Bilddaten mittels statistischer und maschineller Lernverfahren
Seit 2013	Wissenschaftlicher Gutachter für zahlreiche internationale Fachzeitschriften z.B. A*lzheimer's & Dementia*; *PLoS ONE*; *Journal of Neurology, Neurosurgery & Psychiatry (JNNP)*; *British Journal of Psychology*

Ausbildung

Seit April 2011	Doktorand an der medizinischen Fakultät der Universität Rostock in Kooperation mit dem DZNE und der Fakultät für Informatik und Elektrotechnik Doktorväter: Prof. Dr. Stefan Teipel & Prof. Dr. Thomas Kirste Dissertationstitel: *Automatisierte Erkennung manifester und prodromaler Alzheimer-Krankheit mittels multizentrisch akquirierter, multimodaler Bilddaten*
Oktober 2005 – März 2011	Diplomstudium in Informatik, Universität Rostock (Note: 1,6) Diplomarbeit: *Design and Implementation of a SmartLab Controller Based on the Subsumption Architecture*
August 1996 – Juni 2004	Abitur, Elbe-Gymnasium Boizenburg (Note: 1,6)

Zivildienst

Oktober 2004 – Juni 2005	Hausmeistertätigkeit im Jugendfreizeithaus Luna, Boizenburg

Fremdsprachen

Englisch (fließend) und Französisch (Grundkenntnisse)

Rostock, den 18. Mai 2015

Fachpublikationen

2016

Teipel SJ, Grothe M, Zhou J, Sepulcre J, Dyrba M, Sorg C, Babiloni C

Measuring cortical connectivity in Alzheimer's disease as a brain neural network pathology: towards clinical applications.

Journal of the International Neuropsychological Society, Special Issue *Human Brain Connectivity in the Modern Era* (*in press, release scheduled for Feb 2016*). (IF 3,73)

2015

Brüggen K, Dyrba M, Barkhof F, Hausner L, Filippi M, Nestor PJ, Hauenstein K, Klöppel S, Grothe M, Kasper E, Teipel SJ, the EDSD study group

Basal forebrain and hippocampus as predictors of conversion to Alzheimer's disease in patients with Mild Cognitive Impairment – A multicenter DTI and volumetry study.

Journal of Alzheimer's Disease 48(1):197–204. doi: 10.3233/JAD-150063 (IF 4,15)

Dyrba M, Barkhof F, Fellgiebel A, Filippi M, Hausner L, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group

Predicting prodromal Alzheimer's disease in subjects with mild cognitive impairment using machine learning classification of multimodal multicenter diffusion-tensor and magnetic resonance imaging data.

Journal of Neuroimaging 25(5):738–747. doi: 10.1111/jon.12214 (IF 1,82)

Dyrba M, Grothe M, Kirste T, Teipel SJ (2015)

Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM.

Human Brain Mapping 36(6):2118–2131. doi: 10.1002/hbm.22759 (IF 6,92)

2014

Teipel SJ, Grothe MJ, Filippi M, Fellgiebel A, Dyrba M, Frisoni GB, Meindl T, Bokde AL, Hampel H, Klöppel S, Hauenstein K, the EDSD study group (2014)

Fractional anisotropy changes in Alzheimer's disease depend on the underlying fiber tract architecture: a multiparametric DTI study using joint independent component analysis.

Journal of Alzheimer's Disease 41(1):69–83. doi: 10.3233/JAD-131829 (IF 4,17)

Schuster C, Kasper E, Dyrba M, Machts J, Bittner D, Kaufmann J, Mitchell AJ, Benecke R, Teipel SJ, Vielhaber S, Prudlo J (2014)

Cortical thinning and its relation to cognition in amyotrophic lateral sclerosis.

Neurobiology of Aging 35(1):240–246. doi: 10.1016/j.neurobiolaging.2013.07.020 (IF 6,17)

Kljajevic V, Meyer P, Holzmann C, Dyrba M, Kasper E, Bokde AL, Fellgiebel A, Meindl T, Hampel H, Teipel SJ, the EDSD study group (2014)

The ε4 genotype of apolipoprotein E and white matter integrity in Alzheimer's disease.

Alzheimer's & Dementia 10(3):401–404. doi: 10.1016/j.jalz.2013.02.008 (IF 14,48)

2013

Dyrba M, Ewers M, Wegrzyn M, Kilimann I, Plant C, Oswald A, Meindl T, Pievani M, Bokde ALW, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group (2013)

Robust Automated Detection of Microstructural White Matter Degeneration in Alzheimer's Disease Using Machine Learning Classification of Multicenter DTI Data.

PLoS ONE 8(5):e64925. doi: 10.1371/journal.pone.0064925 (IF 4,09)

2012

Dyrba M, Ewers M, Wegrzyn M, Kilimann I, Plant C, Oswald A, Meindl T, Pievani M, Bokde A, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Hauenstein K, Kirste T, Teipel SJ, the EDSD study group (2012)

Combining DTI and MRI for the Automated Detection of Alzheimer's Disease Using a Large European Multicenter Dataset.

In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L. (Eds.), *Proceedings of the 3rd Workshop on Multimodal Brain Image Analysis*. Springer Berlin/Heidelberg, pp. 18–28. ISBN: 978-3-642-33529-7, doi: 10.1007/978-3-642-33530-3_2

2011

Dyrba M (2011)

Design and Implementation of a SmartLab Controller based on the Subsumption Architecture. Institut für Informatik, Universität Rostock, Preprint CS-04-11 (Technical Report/ Diplomarbeit)

Bader S, Dyrba M (2011)

Goalaviour-Based Control of Heterogeneous and Distributed Smart Environments.

In: Proceedings of the 7th International Conference on Intelligent Environments (IE), pp 142–148. ISBN: 978-1-60750-794-9, doi: 10.1109/IE.2011.33

Dyrba M, Nicolay R, Bader S, Kirste T (2011)

Evaluation of two Control Systems for Smart Environments.

In: *Proceedings of the 8th International ICST Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services - 3rd Workshop on Context Systems Design, Evaluation and Optimisation*, Copenhagen, Denmark, pp 1–12.

Präsentationen auf nationalen und internationalen Fachkongressen

2015

Analysis of inter-modal associations and dependencies of regional disease patterns based on multimodal imaging using Markov random fields.
Alzheimer's Association International Conference (AAIC), Washington, DC, USA (Poster)

2014

Prediction of prodromal AD in MCI subjects using multicenter DTI and MRI data and multiple kernels SVM - An EDSD study.
Alzheimer's Association International Conference (AAIC), Kopenhagen, Dänemark (Poster)

Multimodal analysis of functional and structural disconnection in AD using multiple kernels SVM.
Alzheimer's Association International Conference (AAIC), Kopenhagen, Dänemark (Poster)

Kombinierte Erkennung funktioneller und struktureller Diskonnektionsmuster bei der Alzheimer-Krankheit mittels multimodaler MRT und maschineller Lernverfahren.
Jahrestagung der Deutschen Gesellschaft für Klinische Neurophysiologie (DGKN), Berlin (Poster)
Jahrestagung der Deutschen Gesellschaft für Biologische Psychiatrie (DGBP), Aachen (Vortrag)

Machine learning algorithms for multimodal multicenter imaging data – towards an application in clinical studies, Vortrag im Symposium: *Machine learning methods for the individual diagnosis of neuropsychiatric disorders from multimodal imaging and biomarker data.*
Jahrestagung der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN), Berlin

2013

Predicting prodromal Alzheimer's disease in people with mild cognitive impairment using multicenter diffusion-tensor imaging data and machine learning algorithms.
Alzheimer's Association International Conference (AAIC), Boston, MA, USA (Poster)

Multimodal support vector machine for automated detection of functional and structural disconnection in Alzheimer's disease.
Alzheimer's Association International Conference (AAIC), Boston, MA, USA (Poster)

Multizentrische DTI und maschinelle Lernverfahren zu Prädiktion biomarkerbestätigter Prädemenzstadien der AD – Ergebnisse der European DTI Study on Dementia.
Jahrestagung der Gesellschaft für Nervenheilkunde MV, Güstrow (Vortrag)
Jahrestagung der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN), Berlin (Vortrag)

Kombinierte Erkennung funktioneller und struktureller Diskonnektionsmuster bei der Alzheimer-Krankheit mittels multimodaler MRT und maschineller Lernverfahren.
Jahrestagung der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN), Berlin (Vortrag)

2012

Prädemenzdiagnostik der Alzheimer Krankheit mittels maschineller Lernverfahren und multimodaler MRT – Ergebnisse der European DTI Study on Dementia.

Jahrestagung der Gesellschaft für Nervenheilkunde MV, Greifswald (Vortrag)

Jahrestagung der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN), Berlin (Vortrag)

Combining DTI and MRI for the automated detection of Alzheimer's disease using a large European multicenter dataset.

Workshop on Multimodal Brain Image Analysis, Nizza, Frankreich (Poster & Vortag)

Automatic detection of Alzheimer's disease in multicenter DTI data

Alzheimer's Association International Conference (AAIC), Vancouver, BC, Kanada (Poster)

2011

Automatisierte Erkennung von Morbus Alzheimer mittels Diffusions-Tensor-Bildgebung.

Jahrestagung der GMDS, Mainz (Vortrag)

Jahrestagung der Gesellschaft für Nervenheilkunde MV, Schwerin (Vortrag)

Jahrestagung der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN), Berlin (Vortrag)

Goalaviour-Based Control of Heterogeneous and Distributed Smart Environments,

7th International Conference on Intelligent Environments, Nottingham, UK (Vortrag)

Betreute Abschlussarbeiten

Lehmann D (2012)

Vergleich von Lern- und Klassifikationsansätzen für medizinische 3D-Bilddaten.

Diplomarbeit Wirtschaftsinformatik, Fakultät für Informatik und Elektrotechnik, Universität Rostock