

© Copyright by K.U.Leuven

Zonder voorafgaande schriftelijke toestemming van de promotoren en de auteurs is overnemen, kopiëren, gebruiken of realiseren van deze uitgave of gedeelten ervan verboden. Voor aanvragen tot of informatie i.v.m. het overnemen en/of gebruik en/of realisatie van gedeelten uit deze publicatie, wendt u zich tot de K.U.Leuven, Secretariaat Master Biomedische Technologie, Celestijnenlaan 300 C - bus 2419, B-3001 Heverlee (België), tel. 016/327096.

Voorafgaande schriftelijke toestemming van de promotor is eveneens vereist voor het aanwenden van de in dit afstudeerwerk beschreven (originele) methoden, producten, toestellen, schakelingen en programma's voor industrieel nut en voor inzending van deze publicatie ter deelname aan wetenschappelijke prijzen of wedstrijden.

Voorwoord

Deze thesis maakt deel uit van de opleiding tot het bekomen van de graad van Burgerlijk Werktuigkundig- Elektrotechnisch Ingenieur, richting Elektrotechniek - optie Biomedische Technologie: Medische Informatietechnologie aan de Katholieke Universiteit Leuven.

Na bijna tien maanden zoeken in de wereld van de medische beeldanalyse en graven in de veelheid aan segmentatiealgoritmes, zijn we verheugd u, als lezer, dit script te mogen voorstellen. Echter, alvorens u het eigenlijke werk voor te stellen, willen we graag nog enkele mensen bedanken. In de eerste plaats en in het bijzonder, dank aan onze begeleiders Dr. ir. Dirk Loeckx en ir. Bart De Dobbelaer. Ze hebben zich beiden hard ingezet om ons tot het einde te begeleiden. Dankzij hun uitgebreide kennis ter zake, waarbij ze elkaar perfect aanvulden, kon onze thesis worden wat het vandaag geworden is.

Daarnaast willen we ook onze promotor, prof. Dr. ir. Paul Suetens, bedanken omdat hij ons de mogelijkheid geboden heeft ons verder te verdiepen in dit fascinerende onderwerp.

Het was voor ons een groot genoegen te mogen werken in de aangename omgeving van het Medical Imaging Center (MIC), dat dankzij haar unieke ligging, te midden in U.Z. Gasthuisberg, ons de mogelijkheid gaf makkelijk in contact te treden met klinici en zo enige voeling te ontwikkelen met de klinische realiteit van elke dag. Speciaal vermelden we hierbij Walter Coudyzer, die ons tijdens de thesis begeleid heeft en ons voorzag van medische beelden. De klinici maakten het ook mogelijk voor ons om onze algoritmes te ontwikkelen door handmatig de metastasen af te lijnen, hetgeen een tijdrovende bezigheid is. In deze context, willen we graag onze dank uiten, in de eerste plaats aan Dr. Hans Prenen, maar ook aan Dr. Caroline De Vleeschouwer en prof. Dr. Sabine Tejpar. Ook hebben we mogen rekenen op de hulp van enkele andere onderzoekers aan het MIC: prof. Dr. ir. Frederik Maes, prof. Dr. ir. Dirk Vandermeulen en ir. Dieter Seghers.

Ook willen wij onze familie bedanken, die ons steeds gesteund heeft. Ze hebben ons met veel liefde opgevoed met veel aandacht voor doorzettingsvermogen en inzet, zowel op studie- als op sociaal vlak. Daarom van ons beiden: mama, papa, dankjewel. Tot slot vermelden we ook graag onze vrienden: onze collega-studenten burgerlijk ingenieur en voor Dirk, de kotgenoten op Herman Servotte en voor Bert, zijn vriendin Tine.

Toch was het ons een groot plezier om samen, als twee goede vrienden, bijna dagelijks aan deze thesis te bouwen.

Bert & Dirk
Leuven, 27 september 2008

*“It is not the mind of the single individual that provides the sense of certitude,
but the process of communal relationship.”*

Kenneth Gergen, An Invitation to Social Construction, 1999.

Katholieke Universiteit Leuven
Faculteit Ingenieurswetenschappen
Departement Elektrotechniek

DEFINITIEVE EINDWERKTITEL & SAMENVATTING van thesis nr: 3

Namen en voornamen studenten:

SMEETS, Dirk
 STIJNEN, Bert

Richting en optie:

Burgerlijk Werktuigkundig-Elektrotechnisch Ingenieur, Richting Elektrotechniek - optie Biomedische Technologie

Academiejaar:

2007-2008

Nederlandse titel van het eindwerk:

De lever: 3D visualisatie en kwantificatie van metastasen

Engelse vertaling van de titel:

The liver: 3D visualization and quantification of metastases

Korte inhoud:

De kern van deze masterthesis is het kwantificeren van metastasen, uitzaaiingen van kanker, hetgeen in feite neerkomt op de segmentatie van de metastasen. In deze thesis komen verschillende segmentatiemethodes aan bod. Deze zijn gebaseerd op de fotometrische eigenschappen van de metastasen in CT-beelden. Lokale geometrische eigenschappen kunnen gebruikt worden in de segmentatie. Een globaal model echter biedt geen meerwaarde omwille van de willekeurige vorm. De grijswaarden hebben wel een belangrijke functie omdat deze in de metastase gemiddeld verschillen van deze buiten de metastase. Toch is er een belangrijke overlap, hetgeen het probleem vormt voor de eenvoudige segmentatiemethodes zoals *Region Growing* en *Thresholding*. Een ander belangrijk probleem is uitgroei van de methode naar weefsels buiten de lever, die soms gelijkaardige grijswaarden hebben als de metastasen. Ook de belangrijkste Haralick textuurkenmerken zijn onderzocht. Correlatie en variantie zijn het meest discriminerend, doch op kleine schaal onvoldoende onderscheidend.

Een eerste belangrijke methode is *Region Growing*, gebruik makend van een statistische masker. Met behulp van handmatig aangeduide assen in de snede waarin de metastase de grootste oppervlakte heeft, bepaalt het algoritme de verdeling van de grijswaarden in een cirkelvormig of een vierhoekig masker. De gemiddelde waarde en standaarddeviatie worden gebruikt om een boven- en ondergrens vast te leggen voor een *Region Growing* operatie. Een harde beperking in de vorm van een cirkelvormig of een vierhoekig masker moet uitgroei vermijden.

Een tweede belangrijke groep van algoritmes maakt gebruik van een transformatie. De *Spiral Scanning Technique* vergt één zaadpunt in het midden van de tumor, waarna het beeld spiraalvormig wordt afgescand. Een minimaal kostpadalgoritme zorgt voor de segmentatie in het bekomen 2D beeld. De terugtransformatie gebeurt a.d.h.v. een nearest neighbor-interpolatie. Het potentieel voor deze methode in beelden met een lage axiale resolutie is beperkt. Hiervoor biedt *Circle Scanning Technique* een beter alternatief. Hierin plaatst de gebruiker op elke snede een zaadpunt en vindt er voor elke snede een transformatie plaats van een cartesiaans naar een polair assenstelsel. In dit polaire vlak zorgt opnieuw een minimaal kostpadalgoritme voor de segmentatie.

Een derde type van algoritmes zijn de *Pixel Classification*-algoritmes gebaseerd op een statistisch model van zowel grijswaarden als waarden van textuurkenmerken. De kans dat een voxel tot een bepaalde klasse (lever of tumor) behoort, kan uitgezet worden in een beeld.

De meest uitgebreide methode is gebaseerd op een *Level Set*-algoritme. Dit algoritme verandert een initiële level set op basis van een snelheidsbeeld. De initiële level set is het resultaat van een snelle, eerste segmentatie op basis van de *Circle Scanning Technique*. Het snelheidsbeeld is gebaseerd op het beeld bekomen door fuzzy pixel classification met toevoeging van een EM-algoritme om de optimale parameters

van het statistische model te schatten. Dit levert een gemiddelde DSC op van 76,5% als het algoritme wordt toegepast op data waar de lever niet op voorhand gesegmenteerd werd en 78,4% voor een gesegmenteerde lever. Toch is de nauwkeurigheid kleiner als de door ons handmatig afgelijnde tumoren.

Abstract:

The main objective of this master thesis is the quantification of metastases, which is achieved by segmentation. In this thesis several segmentation methods will be discussed. They are all based on the photometric properties of the metastases in CT-images. Local geometric properties can also be used in the segmentation. A global model does not have any added value because of the random shape. The grey values however do have an important function. They differ when compared to the grey values in the healthy liver tissue. Still there is an important overlap between the grey value distributions of liver and metastatic tissue. This is a significant problem when using simple segmentation methods like *Region Growing* and *Thresholding*. Another important problem of these methods is the outgrowth in healthy liver tissue caused by a similarity in grey value with neighboring pixels. Also the most important Haralick texture features are investigated. Our research has shown correlation and variance are the two most discriminant features, although not sufficiently distinctive when applied on small scale.

A first important method is *Region Growing* with a statistical mask. This method consists of manually labeling the two largest orthogonal diameters on the slice with the largest metastatic surface. The algorithm will then calculate the grey value distribution within a circular or quadrangular region determined by the two diameters. The average and standard deviation of the statistical model are used to determine the lower and upper boundary of the Region Growing operation. A hard constraint in the form of a circle or a rectangle is used to prevent outgrowth.

A second important group of algorithm uses a transformation. The *spiral scanning technique* requires a seed point in the center of the tumor, whereupon the image is scanned spirally. From this scan a 2D image is created. A minimal cost path algorithm will calculate the segmentation path in 2D. The back transformation is performed using nearest neighbor-interpolation. The potential for this method in scans with low axial resolution is limited. Therefore the *Circle Scanning Technique* is a better alternative when dealing with low resolution scans. Here a seed point is placed on every slice and a polar transformation is performed per slice. In the polar plane again a minimal cost path algorithm performs the segmentation.

A third type of algorithms is based on *Pixel Classification*-algorithms. For them a statistical model is prepared which contains grey values as well as texture feature values. The probability a pixel belongs to a certain class is shown in an image.

The most elaborate method in this thesis is based on *Level Sets*. The algorithm changes an initial level set by using a velocity field. The initial level set is the result of a quick first segmentation by the Circle Scanning Technique. The velocity field is acquired by performing a fuzzy pixel classification on the original image with addition of a EM algorithm to estimate the optimal parameters for the statistical model. This results in an average DSC of 76,5% when the algorithm is applied to data without a presegmented liver and an average DSC of 78,4% starting from a segmented liver. Still the accuracy remains smaller when compared to the inter-observer variability.

Promotor:

Prof. Dr. ir. P. Suetens

Handtekening promotor(en):

Assessoren:

Prof. Dr. ir. F. Maes

Dr. ir. D. Loeckx

Inhoudsopgave

Voorwoord	i
Inhoudsopgave	iv
Lijst van Acroniemen en Symbolen	vii
1 Inleiding	1
2 Literatuurstudie	2
2.1 De lever	2
2.1.1 Anatomie	2
2.1.2 Fysiologie	2
2.1.3 Pathologie	3
2.2 Kanker	3
2.2.1 Algemeen	3
2.2.2 Metastasen in de lever	4
2.3 Beeldvormingsmodaliteiten	5
2.3.1 Computed Tomography	5
2.3.2 Magnetic Resonance Imaging	6
2.3.3 Positron Emission Tomography	6
2.3.4 Vergelijking beeldvormingsmodaliteiten	6
2.4 Segmentatiemethodes	8
2.4.1 Algemeen	8
2.4.2 Overzicht van de semi- en volautomatische methodes	10
2.4.3 Toepassingen voor segmentatie van metastasen	13
2.5 Kwantificatie van metastasen	17
2.5.1 RECIST	18
2.5.2 Volume	18
2.5.3 Choi-parameters	19
3 Karakteristieken van de metastasen in CT beelden	20
3.1 Dataset	20
3.2 Fotometrische eigenschappen	20
3.2.1 Analyse van de grijswaarden	21
3.2.2 Analyse van de textuur	25
3.3 Geometrische eigenschappen	31
3.3.1 De grootte van de metastase	31
3.3.2 De globale vorm	31

4	Segmentatie van metastasen	34
4.1	Interval thresholding	34
4.2	Statistische pixelclassificatie	34
4.2.1	Gesuperviseerde pixelclassificatie	35
4.2.2	Niet-gesuperviseerd: Expectation Maximization Algoritme	37
4.3	Region Growing	37
4.3.1	Region Growing met één seed per snede in een bepaalde metastase	37
4.3.2	Region Growing met twee seeds per metastase	38
4.3.3	Region Growing met één seed per metastase	38
4.3.4	Region Growing op basis van RECIST en de loodrechte hierop	38
4.3.5	Region Growing gegeven een fuzzy classificatiebeeld	40
4.4	Spiral Scanning Technique	40
4.4.1	Transformatie m.b.v. Spiral Scan	40
4.4.2	Segmentatie m.b.v. Minimal Cost Path (MCP)	41
4.4.3	Terugtransformatie	42
4.4.4	Problemen bij toepassing van het algoritme	42
4.4.5	Verbeteringen van het basisalgoritme	43
4.4.6	Conclusie	45
4.5	Circle Scanning Technique	45
4.5.1	Basisalgoritme	45
4.5.2	Problemen met en verbeteringen aan het algoritme	45
4.5.3	Conclusie	46
4.6	Level Sets met Circle- of Spiral Scanning	46
4.6.1	Kernalgoritme	47
4.6.2	Berekenen van de initiële Level Set	47
4.6.3	Berekenen van het snelheidsbeeld	47
4.7	Besluit	49
5	Technische Validatie	50
5.1	Maatstaven	50
5.1.1	Sensitiviteit en specificiteit	50
5.1.2	Dice Similarity Coefficient en Jaccard Similarity Coefficient	51
5.1.3	Relatieve absolute volumeverschil	51
5.1.4	Afstand tussen oppervlakken	52
5.2	Automatische methodes	52
5.2.1	Interval thresholding	52
5.2.2	Statistische pixelclassificatie	54
5.3	Semi-automatische methodes	55
5.3.1	Region Growing	56
5.3.2	Spiral Scanning Technique	63
5.3.3	Circle Scanning Technique	66
5.3.4	Level Sets met Circle Scan	69
5.4	Kritische evaluatie van de resultaten	74
5.4.1	Overzicht van de resultaten	74
5.4.2	De resultaten in relatie tot andere studies	76
5.4.3	Resultaten in relatie tot de inter-observatorvariabiliteit	76
6	Visualisatie van metastasen	78
6.1	2D Weergave	78
6.2	3D Weergave	79
6.3	Weergave van kwantiteiten	81
6.4	Besluit	81

7 Algemeen besluit	82
7.1 Methodes	82
7.2 Een kritische kijk op de resultaten	83
7.3 Suggesties voor verder werk en onderzoek	83
Bibliografie	85
Lijst van Figuren	89
Lijst van Tabellen	93
A Level Sets	94
A.1 Inleiding	94
A.2 Interface propagatie	94
A.2.1 Grenswaarde formulering	94
A.2.2 Initiële waarde formulering	94
A.2.3 Voordelen	95
A.3 Numerieke methodes	95
A.3.1 Convectie	96
A.3.2 Kromming	96
A.3.3 Tijdsdiscretisatie	96
B Textuuranalyse	97
B.1 Verschillende Technieken voor textuuranalyse	97
B.1.1 Run Length Encoding	97
B.1.2 Co-occurrence matrices	98
B.1.3 Wavelet Gebaseerd	98
B.2 Scatterplots Textuuranalyse	99
C Het opvullen van een vierhoek	105
D MeVisLab	107
D.1 Opbouw en structuur	107
D.2 LiverCAD	108
D.3 Ontwikkelde modules	109
D.3.1 Ontwikkelde Macro Modules	109
D.3.2 Ontwikkelde ML Modules	110
E Validatie	112
E.1 Extra figuren bij gevalideerde methodes	112
E.2 De inter-observatorvariabiliteit: Vergelijking van twee studies	126
F Artikel MICCAI: Liver Tumor Segmentation Challenge	129
F.1 Introduction	129
F.2 Method	130
F.3 Results	134
F.4 Discussion	135
F.5 Conclusions	136

Lijst van Acroniemen en Symbolen

Acroniemen

1D	Eéndimensionaal
2D	Tweedimensionaal
3D	Driedimensionaal
COV	Coefficient of Variation
CT	Computed Tomography
DM	Discriminant Measure
DNA	Desoxyribonucleïnezuur
DSC	Dice Similarity Coefficient
EM	Expectation Maximization
FDG	Fluorodeoxyglucose
FP	False Positives
FPF	False Positive Fraction
FPF'	False Positive Fraction (Geherdefinieerd)
FN	False Negatives
FNF	False Negative Fraction
Gd	Gadolinium
GIST	Gastrointestinal Stromal Tumor
GPU	Graphical Processing Unit
AASD	Average Absolute Surface Distance
GW	Grijswaarden
HU	Hounsfield Unit
ITK	Insight Tool Kit
JSC	Jaccard Similarity Coefficient

KKO	Kleinste Kwadraten Oplossing
kNN	k Nearest Neighbor
MCP	Minimal Cost Path
MDL	Minimum Description Length
MGP	Maximum Gradient Path
MICCAI	Medical Image Computing and Computer-Assisted Intervention
ML	Maximum Likelihood
MPR	Multi Planar Reconstruction
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
MASD	Maximal Absolute Surface Distance
PCA	Principal Component Analysis
PET	Positron Emission Tomography
RAV	Relative Absolute Volume Difference
RAAD	Relative Absolute Area Difference
RECIST	Response Evaluation Criteria In Solids Tumors
RGBA	Rood Groen Blauw Alfa
RLE	Run Length Encoding
RMS	Root Mean Square
ROI	Region Of Interest
SPIO	Super-Paramagnetic Iron Oxide
RMSSD	RMS Surface Distance
TP	True Positives
TPF	True Positive Fraction
TN	True Negatives
TNF	True Negative Fraction
US	Ultrasonic
UV	Ultraviolet
VOI	Volume Of Interest
WHO	World Health Organization criteria

Symbolen

symbool	beschrijving
Γ	contour (2D), oppervlak (3D)
$\phi(\mathbf{x}, t)$	level set functie
φ	elevatiehoek
κ	kromming
π	pi, 3.1415926...
ρ	straal
θ	projectiehoek
μ	gemiddelde
σ	standaarddeviatie
a	scalar a
\mathbf{a}	vector a
$C(i, j)$	co-occurrence matrix
$C_n(i, j)$	genormalizeerde co-occurrence matrix
D	pixelafstand
E_C^{int}	interne kost
E_C^{ex}	externe kost
$I_{cor,k}$	de correlatie van het beeld in voxel k
$I_{var,k}$	de variantie van het beeld in voxel k
I_k	intensiteit van pixel k
I_V	het snelheidsbeeld
\log	neperiaans logaritme
\vec{M}	netto magnetisatie
M_{F1}^t	kenmerkmatrix van de klasse t
P	kans
$p(r, \theta)$	sinogram
Q25	eerste kwartiel
Q50	tweede kwartiel (mediaan)
Q75	derde kwartiel
V	volume
\vec{V}	snelheidsveld
V_N	normale component van \vec{V}
s	standaarddeviatie van een steekproef
\bar{X}	gemiddelde van een steekproef X
\tilde{X}	mediaan van een steekproef X
∇	Nabla-operator
\mathcal{R}	Radontransformatie

Hoofdstuk 1

Inleiding

Mede door verregaande ontwikkelingen in de gezondheidszorg, neemt de gemiddelde levensduur in in België en in de rest van de wereld sinds begin vorige eeuw steeds toe. Hierdoor neemt de aandacht voor een betere kwaliteit in ons gezondheidszorgsysteem toe, met als gevolg dat de levensverwachting stijgt. Doodsoorzaken verschuiven steeds meer van acute naar chronische aandoeningen, mede door het overwinnen van verschillende infectieziekten. Het gewijzigde ziektepatroon weerspiegelt zich vandaag de dag steeds meer in de ziekenhuizen. Patiënten met chronisch hartfalen en kankerpatiënten hebben een belangrijk aandeel in de ziekenhuispopulatie van vandaag. Voor deze laatste groep patiënten is er de laatste jaren veel veranderd, mede dankzij meer en betere technologische middelen. Hierdoor stijgt de kans dat de patiënt deze vreselijke ziekte overwint. Indien hij niet dat geluk heeft, kan de technologie vaak de levensduur verlengen.

Binnen deze technologie neemt de medische beeldvorming als diagnostisch hulpmiddel een belangrijke plaats in. Met behulp van hoog-energetische straling, radiogolven of geluidsgolven kan de clinicus op een niet-invasieve manier in het lichaam kijken en op die manier tot een snellere en betere diagnose te komen. Een belangrijk toepassingsgebied is de opvolging van kanker. Om de overlevingskans van de patiënt beter in de schatten en om de behandeling meer effectief te maken, moet de evolutie van de tumoren in kaart worden gebracht. Hoe meer de tumor in grootte afneemt of hoe lager de dichtheid, des te beter voor de patiënt. Het inschatten van de grootte van tumoren gebeurt vandaag de dag meestal met een eenvoudige unidimensionale meting. Studies hebben uitgewezen dat complexere metingen, zoals het bepalen van het volume, betrouwbaarder zijn, maar ook tijdrovender. Een meer geautomatiseerde meting, hetgeen de kerntaak is van dit werk, kan hiervoor een oplossing bieden. In deze thesis is toegespitst op tumoren in de lever. De lever is een veel voorkomende plaats voor kankergezwellen, meestal metastasen (uitzaaiingen van primaire tumoren). Vandaar dat we een algoritme uitgewerkt hebben voor de kwantificatie van deze metastasen in de lever.

Het probleem van het meten van het volume van tumoren herleidt zich eigenlijk tot een segmentatieprobleem. In hoofdstuk 2 bekijken we de belangrijkste segmentatie-algoritmes, zowel de algemene als de meer specifieke voor tumorsegmentatie in lever. Uit studies blijkt dat dit geen eenvoudige taak is en dat er vooralsnog geen algoritme bestaat dat voldoende nauwkeurig de tumor kan kwantificeren[1]. Het doel van deze thesis is dan ook te zorgen voor een segmentatie die vergelijkbaar of, in het optimale geval, beter presteert dan algoritmes die reeds in studies gepubliceerd werden. Vooraleer tot de meer complexe algoritmes over te gaan, is het noodzakelijk de metastasen in de medische beelden van nabij te bestuderen (hoofdstuk 3). Verschillende segmentatie-algoritmes worden geïmplementeerd en vervolgens gevalideerd a.d.h.v. door klinici handmatig afgelijnde metastasen (hoofdstukken 4 en 5). Tot slot gaat hoofdstuk 6 kort in op het visualiseren van de metastasen in de lever.

Hoofdstuk 2

Literatuurstudie

Dit hoofdstuk kadert de thesis in een breder geheel. Het behandelt zowel enkele medische aspecten alsook technische. Eerst volgt een elementaire beschouwing van de lever in 2.1 en kanker in 2.2, zodat technische aspecten steeds geplaatst kunnen worden in een klinische context. De technische zijde omvat enerzijds het tot stand brengen van medische beelden (2.3) en anderzijds de extractie van informatie uit deze beelden voor klinische doeleinden. Bij dit laatste speelt segmentatie een belangrijke rol. Verschillende segmentatiemethodes zijn beschreven in 2.4, beginnende met enkele categorieën van algemene methodes en gevolgd door specifieke segmentatie-algoritmes.

2.1 De lever

In deze tekst wordt regelmatig verwezen naar bepaalde kenmerken van de lever die van nut zijn voor een beter begrip van de thesisopdracht. Daarom wordt hier een kort overzicht gegeven van de anatomie, fysiologie en enkele veelvoorkomende pathologieën van de lever.

2.1.1 Anatomie

De lever (Lat:*hepar*) is het tweede grootste orgaan uit het menselijk lichaam en gelegen in de buikholte net onder het middenrif. Traditioneel is de lever opgebouwd uit vier lobben gescheiden door ligamenten. Ze zijn benoemd volgens hun ligging in de buikholte: voor, achter, links en rechts. De lobben bestaan op hun beurt uit lobuli die de levercellen bevatten en waarlangs het bloed wordt afgevoerd.

Naast de traditionele anatomie kan de lever ook opgedeeld worden volgens zijn functionele anatomie. Deze functionele opdeling gebeurt op basis van de opsplitsing van de leverader en -slagader in een linker- en rechtertak die ieder een apart deel van de lever bevoeien. Dit levert een opdeling in een linker- en rechterlob op. Vervolgens kan de lever nog verder opgedeeld worden volgens het Couinaud systeem [2].

2.1.2 Fysiologie

De lever is één van de belangrijkste organen uit het menselijk lichaam. Hij speelt een uitgesproken rol in de stofwisseling en de secretie. Vandaar ook dat problemen aan de lever vaak onherroepelijke gevolgen hebben. Bij het falen van de lever worden een aantal belangrijke functies uitgeschakeld. Een eerste belangrijke functie van de lever is de productie van gal door de levercellen, noodzakelijk voor de afbraak van vetten. Daarnaast speelt de lever ook een belangrijke rol in de opbouw en afbraak van koolhydraten via processen als gluconeogenese, glycogenolysis en glycogenese. Een volgende functie van de lever is de vorming en opslag van bepaalde vitamines zoals vitamine B12. Tenslotte is de lever ook verantwoordelijk voor de afbraak van vele gifstoffen (vb. alcohol en bilirubine) en de omzetting van ammoniak naar urea. Bij het falen van de lever leidt een ammoniakvergiftiging immers tot de dood.

2.1.3 Pathologie

De verschillende functies van de lever hierboven aangehaald, maken het duidelijk dat aandoeningen aan de lever zeer ernstige schade kunnen veroorzaken. De verschillende pathologieën die in de lever optreden gaan vaak gepaard met geelzucht veroorzaakt door een verhoogde concentratie bilirubine in het bloed. Bij normale werking breekt de lever het excès aan bilirubine af.

Een veel voorkomende pathologie is levercirrose. Hierbij worden de levercellen vernietigd en vervangen door bindweefsel. De oorzaak hiervan ligt o.a. bij overmatig alcoholgebruik of een virale hepatitis infectie.

Het grootste probleem vormt echter kanker. Dit kan zich voordoen onder twee verschillende vormen: leverkanker en uitzaaiingen of metastasen van de lever. Deze uitzaaiingen zijn meestal het gevolg van andere kankers in het darmstelsel of in de longen. De reden waarom de lever zo gevoelig is voor uitzaaiingen, is haar felle doorbloeding.

2.2 Kanker

Kanker is een aandoening waarbij cellen zich ongecontroleerd vermenigvuldigen. Deze woekerende cellen dringen binnen in het weefsel waar ze veel schade aanbrengen. Bovendien kunnen ze zich verder in het lichaam verspreiden. Dit zijn dan uitzaaiingen of ook wel metastasen genoemd.

2.2.1 Algemeen

Kanker houdt in dat cellen zich ongecontroleerd vermenigvuldigen ten gevolge van defecten in het DNA. Het proces van de mutaties, die aan de basis liggen van het woekeren van de cellen, is wereldwijd onderwerp van onderzoek.

De classificatie van tumoren gebeurt volgens het type cel waaruit het gezwel is opgebouwd. Men onderscheidt 5 typen kankers of kwaadaardige tumoren:

- Carcinomen: opgebouwd uit epitheelcellen en vooral voorkomend bij oudere mensen.
vb. longkanker, darmkanker, borstkanker
- Sarcomen: opgebouwd uit steunweefsel of mesenchymale weefsels.
vb. Ewing-saroom
- Kwaadaardige lymfomen: vindt zijn oorsprong in de lymfocyten.
vb. ziekte van Hodgkin
- Blastomen: kanker afkomstig uit weefsel in ontwikkeling.
vb. netvlieskanker, nierkanker
- Kiemceltumoren: ontstaat bij de geslachts- of kiemcellen.
vb. Teelbalkanker

De meest voorkomende kankers worden weergegeven in tabel 2.1:

Tabel 2.1: Meest voorkomende kankers [3]

Mannen	Vrouwen
Prostaatkanker (29,4 %)	Borstkanker (30,4 %)
Longkanker (15,8 %)	Longkanker (13,3 %)
Colorectale kanker (11,2 %)	Colorectale kanker (11,7 %)
Blaaskanker (6,5 %)	Baarmoederhalskanker (5,7 %)
non-Hodgkin lymfoom (4,1 %)	non-Hodgkin lymfoom (3,9 %)

De symptomen waarmee de patiënt aanvankelijk geconfronteerd wordt, zijn vaak een ongewone zwelling of knobbeltjes, ongewone bloedingen, vermoeidheid en een aanhoudende pijn. Dit gaat vaak gepaard met gewichtsverlies, verminderde eetlust en excessief zweten.

Hoewel de symptomen een grote aanwijzing zijn, kan de precieze vorm van kanker pas vastgesteld worden na een meer uitgebreide diagnose. Hierbij wordt de patiënt aan een aantal medische tests onderworpen waaronder een bloedtest, een punctie in het gewel, een Röntgen-, een CT- en/of een PET-scan. Vervolgens kunnen nog andere testen worden uitgevoerd specifiek voor het type kanker. Om een definitieve zekerheid te krijgen over de aard van de kanker, dient een biopsie uitgevoerd te worden.

Nadat een vorm van kanker is vastgesteld, gaat men zo vlug mogelijk over naar de behandeling. Deze kan bestaan uit een operatie om de aanwezige tumor te verwijderen. Indien een operatie niet mogelijk is, zijn verscheidene andere opties mogelijk. De belangrijkste zijn chemotherapie en bestraling of radiotherapie. Bij chemotherapie zal men cytotoxische medicijnen gebruiken om de deling bij sneldelende kankercellen tegen te houden. Het nadeel hiervan is dat ook het lichaam fel verzwakt wordt. Bij een bestralingstherapie wordt de tumor gericht bestraald. De radioactieve straling zorgt ervoor dat tumorcellen afsterven.

De oorzaak van kanker dient op verschillende vlakken gezocht te worden. Enerzijds zijn er de verschillende chemische carcinogenen (bijv. sigaretten) die voor mutaties in het DNA kunnen zorgen en de kans op kanker vergroten. Daarnaast is er het gevaar van de ioniserende straling. Een voorbeeld hiervan is de toenemende kans op huidkanker bij blootstelling aan UV-licht. Tenslotte bestaan er ook nog verschillende ziektes die kunnen aanleiding geven tot kanker. Sommige types van kanker zijn erfelijk, terwijl voor andere types er geen significant verband is.

Via het epidemiologisch onderzoek naar kanker kunnen onderzoekers meer en meer verbanden leggen tussen kanker en zijn mogelijke oorzaken. Door de toenemende en steeds ouder wordende bevolking is het aantal kankergevallen de laatste 30 jaar verdubbeld [5]. Kanker is als het ware een welvaartsziekte geworden binnen onze Westerse samenleving. Het zal één van de uitdagingen van de 21^{ste} eeuw worden om hiervoor oplossingen te vinden. Met deze thesis hopen wij in ieder geval een bijdrage te leveren.

2.2.2 Metastasen in de lever

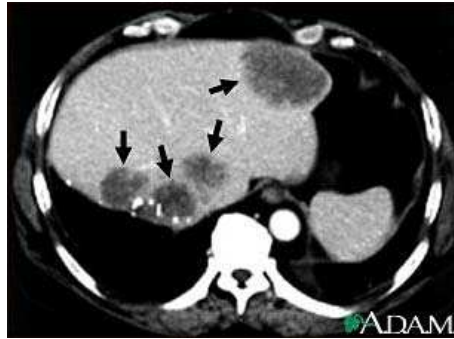
Metastasen komen vaak voor in de laatste stadia van kanker. Via het lymfevatensstelsel en de bloedsomloop kunnen de woekerende cellen zich verspreiden over het lichaam. Zo kunnen er overal in het lichaam metastasen ontstaan. Deze metastasen komen veelvuldig voor in de lever ten gevolge van haar sterke doorbloeding en dit vooral als uitzaaiingen van andere kankers in het maag-darmstelsel of de longen. Ze kunnen zich zeer snel en agressief verspreiden maar soms ook pas jaren na de behandeling van de oorspronkelijke tumor ontdekt worden. De meest voorkomende plaatsen voor metastasen zijn de longen, bot, hersenen en lever [6]. Deze laatste wordt in deze thesis behandeld.

Volgens de “seed and soil-theorie” [7] is er een neiging voor bepaalde tumoren om uit te zaaien naar particuliere organen. Hierdoor zijn levermetastasen vaak het gevolg van darmkanker. Tot 70 % van de patiënten met colorectale kanker ontwikkelen metastasen in de lever. De verklaring hiervoor is dat het voor kankercellen moeilijk is om te overleven buiten hun oorsprongsgebied, zodat de vorming van metastasen vaak gebeurt in gebieden met gelijkaardige karakteristieken. Metastasen zijn theoretisch gezien steeds afkomstig van een primaire tumor, maar in praktijk is deze soms niet bekend.

Sommige mensen met uitgezaaide kanker hebben geen symptomen. Wanneer symptomen van uitgezaaide kanker voorkomen, is de aard en de frequentie van de symptomen afhankelijk van de grootte en de locatie van de metastase. Bijvoorbeeld, kanker die zich verspreidt naar de botten zal waarschijnlijk pijn veroorzaken en kan leiden tot botfracturen. Kanker die zich verspreidt naar de hersenen kan leiden tot verschillende symptomen, waaronder hoofdpijn, epileptische aanvallen en evenwichtsstoornissen. Kortademigheid kan een teken zijn van metastasen in de longen. Abdominale zwelling of geelzucht kunnen aangeven dat kanker zich naar de lever verspreidt. Soms wordt de primaire kanker pas ontdekt nadat de metastase symptomen veroorzaakt [6].

De cellen in een metastase lijken op deze van de primaire tumor. Eens het kankerweefsel onderzocht is onder de microscoop om het celtype te bepalen, kan een arts gewoonlijk vertellen of dat celtype normaliter gevonden wordt in het deel van het lichaam waar het weefselmonster

genomen werd [6]. Alvorens een monster te nemen, wordt zeer vaak gebruik gemaakt van medische beelden om de metastasen te detecteren. Een voorbeeld hiervan is Figuur 2.1.



Figuur 2.1: Voorbeeld van levermetastasen, opgenomen met CT [8]

Ondanks uitgebreid onderzoek naar diagnostische performantie van CT, MRI en PET voor de detectie van colorectale levermetastasen is de optimale beeldvormingsstrategie nog niet gedefinieerd [1]. Of de kanker lokaal woekert, of zich verspreidt naar andere plaatsen, heeft een grote invloed op de behandeling en de overlevingskans. Als de kanker uitzaait, daalt de overlevingskans van de patiënt.

Wanneer er metastasen aanwezig zijn, kan de behandeling bestaan uit radiochirurgie, chemotherapie, bestraling, biologische therapie, hormoontherapie, chirurgische ingreep, laser-immunotherapie, of een combinatie van deze. De keuze hangt gewoonlijk af van het type primaire tumor, de grootte en de locatie van de metastasen, de leeftijd van de patiënt en zijn algemene gezondheidstoestand, en eventuele voorafgaande behandelingen. Jammer genoeg zijn de mogelijkheden van behandeling, die op dit moment beschikbaar zijn, zelden in staat om een uitgezaaide kanker te genezen [6].

2.3 Beeldvormingsmodaliteiten

Er zijn verschillende beeldvormingsmodaliteiten beschikbaar die in staat zijn levermetastasen in beeld te brengen. Eerst volgt hiervan een korte bespreking, waarna ze met elkaar vergeleken worden.

2.3.1 Computed Tomography

Het in beeld brengen van de lever met eventuele uitzaaiingen kan gebeuren met behulp van Computed Tomography (CT). Bij CT wordt de patiënt bestraald met X-stralen. Voor het volledige begrip van deze thesis is een basiskennis rond CT-beelden noodzakelijk. In CT-beelden wordt aan de hand van een dunne X-stralenbundel de attenuatie in een doorsnede opgemeten. Door te roteren over alle hoeken kan men met behulp van de Radontransformatie de attenuatie in ieder punt berekenen:

$$p(r, \theta) = \mathcal{R}\{f(x, y)\} = \int_{-\infty}^{\infty} f(r \cdot \cos\theta - s \cdot \sin\theta, r \cdot \sin\theta + s \cdot \cos\theta) ds \quad (2.1)$$

Vervolgens kan men deze operatie uitvoeren voor iedere snede om op die manier de gewenste doorsnede van een bepaald lichaamsdeel of orgaan te bekomen.

Bij de gewone (sequentiële) CT-beelden roteert de X-stralenbuis 180 graden rond de patiënt waarbij parallelle stralen worden uitgezonden. Sequentieel worden de verschillende snedes afgegaan om uiteindelijk het volledige gewenste volume op te meten. De afstand tussen de verschillende opeenvolgende snedes wordt bepaald door de effectieve snededikte. Meestal wordt als afstand

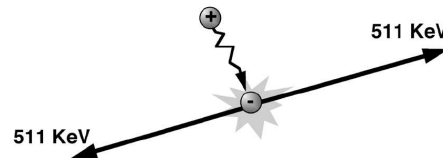
hier de helft van de snededikte gekozen. De huidige CT-scanners gebruiken echter spiral CT of helical CT. Dit houdt in dat de patiënt doorheen de gantry transleert terwijl de X-stralenbuis ronddraait. Het voordeel hiervan is dat de scantijd aanzienlijk verlaagd kan worden wat een lagere dosis en minder bewegingsartefacten als gevolg heeft [9, 10].

2.3.2 Magnetic Resonance Imaging

Een andere mogelijkheid om metastasen te visualiseren is met behulp van Magnetic Resonance Imaging (MRI). MRI meet in feite een magnetische eigenschap van het weefsel. Door het weefsel in een groot magnetisch veld te leggen gaan magnetische momenten, die het gevolg zijn van het spinnen van de elektronen, zich richten. Dit zorgt voor een netto magnetisatie \vec{M} in de richting van het grote aangelegde magnetische veld. De reden hiervoor is dat deze richting energetisch het meest interessant is (meeste spin-up toestanden). Met radiofrequente golven wordt de magnetisatie verstoord. Met een α° -puls wordt \vec{M} omgeklapt over een hoek α . Er zijn twee relaxatiefenomenen die de magnetisatie terug in zijn oorspronkelijke toestand brengen. Dit is echter weefselafhankelijk en vormt zo de basis voor MRI [9].

2.3.3 Positron Emission Tomography

Positron Emission Tomography (PET) behoort tot de nucleaire geneeskunde en is dus gebaseerd op radio-actief verval. Belangrijk bij PET is dat een proton getransformeerd wordt in een neutron en een positron. Dit gebeurt op de plaatsen in het lichaam waar de radioactieve stof is ingespoten. Het positron raakt na een zeer korte tijd en binnen een paar millimeter van zijn oorsprongsgebied een electron en annihileert. Er vindt een omzetting plaats van de massa van de twee deeltjes naar energie, in de vorm van twee fotonen die in de tegenovergestelde richting uitgezonden worden, zoals weergegeven in Figuur 2.2.



Figuur 2.2: Annihilatie van een electron en een positron

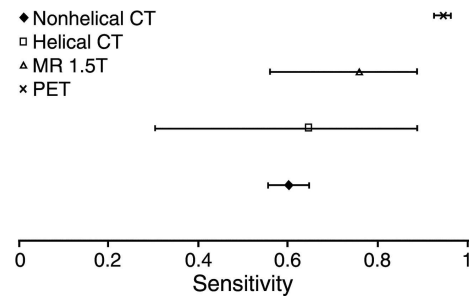
Elk foton heeft een energie van 511 keV. Een elektronisch coïncidentiecircuit zorgt voor de detectie van de fotonen. De reconstructie van het oorspronkelijke beeld gebeurt met iteratieve reconstructie gebaseerd op een Expectation Maximization (EM) algoritme [9].

2.3.4 Vergelijking beeldvormingsmodaliteiten

Uit studies blijkt dat er een kwalitatief verschil is tussen de verschillende beeldvormingsmodaliteiten wat betreft de detectie van metastasen in de lever en wat betreft het aflijnen van de metastasen in de lever. Algemeen kunnen we stellen dat PET hiervoor de meest geschikte beeldvormingsmodaliteit is.

Detectie van metastasen

Een eerste vergelijking tussen de verschillende beeldvormingsmodaliteiten betreft de detectie van metastasen. De vergelijking leert ons in welke beeldvormingsmodaliteit het percentage van aantal gedetecteerde metastasen het grootst is. Het resultaat is weergegeven in Figuur 2.3.



Figuur 2.3: Vergelijking beeldvormingsmodaliteiten per patiënt [1]

De gebruikte maat is de *sensitiviteit* = $\frac{TP}{TP+FN}$, waarbij ook het betrouwbaarheidsinterval is weergegeven. De True Positives (TP) zijn de correct gedetecteerde metastasen, de False Negatives (FN) zijn de niet gedetecteerde metastasen.

Bipat *et al.* [1] gaat uit van 61 studies tussen januari 1990 en december 2003, die aan specifieke criteria voldoen. In Tabel 2.2 staat een overzicht van deze studie en de studie van Kinkel *et al.* [11].

Tabel 2.2: Vergelijking Beeldmodaliteiten: sensitiviteit [1, 11]

Beeldvormingsmodaliteit	Bipat <i>et al.</i> (in %)	Kinkel <i>et al.</i> (in %)
US	-	55
CT	-	72
non-helicoïdale CT	60,2	-
helicoïdale CT	64,7	-
MRI	75,8	76
FDG PET	84,6	90

Hieruit blijkt dat Fluorodeoxyglucose (FDG) PET de hoogste kans op detectie geeft. MRI is gemiddeld gezien iets beter dan CT, maar dit is dan vooral dankzij de MRI met versterking van bijvoorbeeld Gadolinium (Gd).

Aflijning van de metastasen

Naast het al dan niet correct detecteren van de metastasen werd in Bipat *et al.* ook gekeken naar de metastasen zelf, m.a.w. hoe goed de metastasen afgelijnd werden. De gebruikte maat is opnieuw sensitiviteit die de verhouding weergeeft tussen het deel van de metastase dat binnen de aflijning ligt en de totale tumor. De resultaten zijn weergegeven in Figuur 2.4 waarbij tussen haakjes het betrouwbaarheidsinterval is gegeven.

Modality and Subgroup	Sensitivity Estimate (%)*
Nonhelical CT, overall	52.3 (52.1, 52.5)
Helical CT	
Overall	63.8 (54.4, 72.2)
Section thickness of 5 mm	68.2 (50.5, 81.9)
Section thickness of >5 mm	69.1 (59.8, 77.1)
Amount of iodine of ≤45 g	61.4 (43.5, 76.6)
Amount of iodine of >45 g	64.0 (55.1, 72.0)
Two phases (arterial and portal phases)	65.7 (56.8, 73.7)
One phase (portal phase only)	71.4 (57.7, 82.1)
1.0-T MR imaging, overall	66.1 (65.9, 66.3)
1.5-T MR imaging	
Overall	64.4 (57.8, 70.5)
Nonenhanced MR imaging	59.8 (49.0, 69.7)
Gadolinium-enhanced MR imaging	78.2 (63.0, 88.3)
SPIO-enhanced MR imaging	73.2 (62.3, 81.9)
FDG PET, overall	75.9 (61.1, 86.3)

Figuur 2.4: Vergelijking tussen verschillende beeldvormingsmodaliteiten per metastase [1]

Ook hier geeft FDG PET het beste resultaat van de verschillende modaliteiten. MRI, met Gd versterkt geeft nog een beter resultaat. Omwille van zijn toegankelijkheid en snelheid is in deze thesis gekozen om te vertrekken van CT.

2.4 Segmentatiemethodes

In de hedendaagse geneeskunde zijn diagnostische beeldvormingstechnieken onmisbaar geworden. Via beeldvormingsmodaliteiten als CT, MRI en andere is het mogelijk het menselijk lichaam op een niet-invasieve manier te bestuderen. Door het steeds toenemende aantal medische beelden, is het voor de arts niet meer mogelijk alle eigenschappen en abnormaliteiten manueel te detecteren. De ontwikkeling van visualisatietools en -algoritmes vereenvoudigen het werk aanzienlijk.

Eén van de belangrijkste vooruitgangen op dit gebied vormen de segmentatiealgoritmes geschikt voor het aflijnen en de volumebepaling van anatomische structuren en pathologische verschijnselen zoals tumoren en metastasen. Deze algoritmen in het bijzonder hebben een sterke vooruitgang betekend in het radiologische onderzoek. Aangezien in deze thesis de ontwikkeling van een segmentatiealgoritme vereist is, volgt hier een overzicht van enkele mogelijk nuttige algoritmes. Eerst plaatsen we deze segmentatie-algoritmes in een breder kader.

2.4.1 Algemeen

In computervisie slaat segmentatie op het proces dat een digitaal beeld opdeelt in verschillende gebieden. Dit wordt in de medische wereld vaak gebruikt om de éne anatomische structuur te onderscheiden van de andere. De segmentatie kan manueel gebeuren of met tussenkomst van een computer.

Manuele segmentatie

Bij manuele segmentatie wordt door een observator, meestal een clinicus, een gesloten kromme getrokken om op die manier het onderscheid te maken tussen metastase (gelegen in de gesloten kromme) en de rest van de lever (buiten de kromme). De observator dient op elke snede zo'n kromme te trekken. Dit maakt het manueel aflijnen van metastasen een tijdrovende bezigheid, hetgeen het grootste nadeel is aan handmatige segmentatie. Daarnaast heeft manuele segmentatie ook het grote nadeel dat het afhankelijk is van:

- veranderingen in helderheid- en contrastinstellingen
- inter- en intra-observatorvariabiliteit
- externe lichtcondities

Het stelt bovendien sterke eisen aan hardware, beeldscherm en software [9].

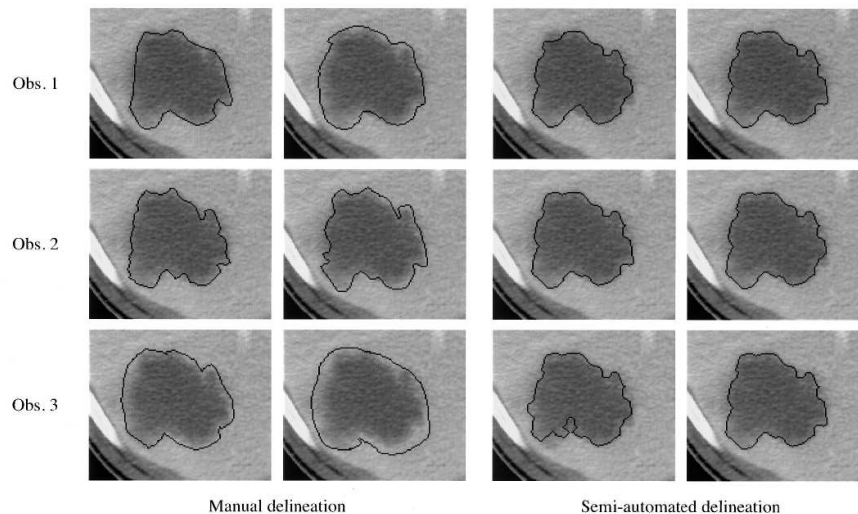
In [12] wordt de reproduceerbaarheid nagegaan voor verschillende maten voor de kwantificatie van de grootte van levermetastasen. De combinatie van inter- en intra-observatorvariabiliteit van het volume zorgt voor een Coefficient of Variation (COV)¹ van $\pm 8\%$. De COV is groter voor kleinere tumoren. Voor uni- en bidimensionele maten is deze COV ook groter.

Semi-automatische en automatische segmentatie

Om het nadeel van inter- en intra-observatorvariabiliteit te vermijden is een meer geautomatiseerde segmentatie noodzakelijk. Dergelijke segmentatie is bovendien minder tijdrovend (voor de klinici) en onafhankelijk van de instellingen voor contrast en helderheid. Voor het bepalen van het volume is het belangrijk dat dit op een consistente manier gebeurt om zo de volumes van de metastasen op verschillende ogenblikken te kunnen vergelijken. In dat opzicht is het belangrijk kleine variaties te hebben op de fout op de volumemeting. Het is dus beter om systematisch bijvoorbeeld 10% fout te meten met heel kleine variaties dan gemiddeld juist te meten maar met grote variaties. Vaak is geldig dat semi-automatische methodes beter presteren dan volautomatische methodes. Deze laatste vereisen meer rekenkracht en vragen dikwijls nog manuele correcties [13]. Daarnaast vereisen volautomatische methodes a priori-kennis van vormen of intensiteiten, hetgeen moeilijk kan zijn door de complexiteit en de diversiteit van anatomische en pathologische objecten [14].

Vergelijkende studie

[15] vergelijkt manuele en semi-automatische aflijningen van levertumoren op CT beelden wat betreft efficiëntie, kwaliteit en intra- en interobservator variabiliteit. Om met een redelijke mate van zekerheid te kunnen concluderen dat een gemeten volumeverandering inderdaad een werkelijk effect reflecteert, moet de meting vergeleken worden met de variantie van de meting. Deze variantie kan gezien worden als de som van de variabiliteit tussen de onderzoeken en de inter- of intra-observatorvariabiliteit. Deze laatste hangt af van de subjectieve interpretatie van de grootte van tumor en van de variabiliteit van de aflijningsmethode. Dit wordt nagegaan door een semi-automatische methode 2.4.3 te vergelijken met een manuele methode. Drie radiologen lijnen de levertumoren tweemaal af voor elke methode met een tussentijd van ongeveer een week.



Figuur 2.5: Inter- en intra-observatorvariabiliteit: vergelijking manuele met semi-automatische aflijning [15]

¹ $COV = \frac{\sigma}{\mu}$, met σ de standaarddeviatie en μ het gemiddelde.

Uit het onderzoek blijkt dat de oppervlakken bekomen door de semi-automatische aflijningen significant kleiner (11%) zijn dan diegene bekomen door de handmatige aflijningen. De inter- en intra-observatorvariabiliteit is ongeveer driemaal lager met de semi-automatische methode. De intra-observatorvariabiliteit, gemeten a.d.h.v. Relative Absolute Area Difference (RAAD)(zie 5.1.3), voor de manuele aflijningen bedroeg gemiddeld 10%, voor de semi-automatische gemiddeld 4%. De inter-observatorvariabiliteit, gemeten a.d.h.v. de COV, bedraagt respectievelijk 10,3% en 3,3%. De efficiëntie werd subjectief als beter beoordeeld, evenals de kwaliteit (beter in 73% van de gevallen voor de semi-automatische aflijning).

2.4.2 Overzicht van de semi- en volautomatische methodes

Het type segmentatiemethode varieert sterk voor verschillende toepassingen binnen de medische beeldvorming. Segmentatie van de witte materie in de hersenen vereist een ander type algoritme dan bijvoorbeeld de segmentatie van de lever. Daarnaast heeft ook iedere beeldvormingsmodaliteit zijn voor- en nadelen. Daarom bestaat er ook niet één algoritme dat goede resultaten geeft voor alle beelden. De ontwikkeling van een eigen algoritme, al dan niet gebaseerd op bestaande methodes, is noodzakelijk om de best mogelijke performantie te bekomen.

De onderstaande paragrafen geven een overzicht van de verschillende technieken die van nut kunnen zijn voor de segmentatie van medische beelden [16]. Er bestaan verschillende manieren om deze methodes op te delen in subcategorieën. Deze opdeling is arbitrair aangezien meestal gewerkt wordt met een combinatie van verschillende technieken in het uiteindelijke segmentatiealgoritme. Van elke techniek zullen het algemene werkingsprincipe en de verschillende voor- en nadelen besproken worden.

Thresholding

De techniek van het thresholden stelt een drempelwaarde in voor de intensiteit. Op basis van deze waarde wordt een scheiding tussen de beeldintensiteiten gecreëerd. Dit levert een binair beeld op waarbij alle waarden kleiner dan de drempelwaarde eenzelfde intensiteit krijgen toegemeten. Alle waarden groter dan de drempelwaarde krijgen een andere intensiteit toegewezen.

Het nadeel is dat de spatiale kenmerken van het beeld niet in rekening worden gebracht. Dit maakt deze segmentatiemethode gevoelig voor ruis en inhomogeniteiten binnen de structuur. Dankzij de sterke eenvoud is thresholding ondanks de vele nadelen toch een veel gebruikte techniek binnen de segmentatie van medische beelden. Vaak is het een eerste stap in een keten van verschillende beeldverwerkingsoperaties.

Region Growing

Bij region growing is eveneens een intensiteitsgebaseerde methode. De eenvoudigste vorm van region growing is met behulp van een zaadpunt. Hierbij wordt een punt geselecteerd binnen een gewenste regio en op basis van de beeldintensiteit in en rond dit punt wordt dit uitgebreid tot een gebied. De grootte van het gesegmenteerde gebied wordt bepaald door de intervalgrootte van de intensiteiten die men meeneemt. Het nadeel bij deze methode is dat er geen rekening wordt gehouden met de spatiale informatie van de structuur die men wilt aflijnen. Dit maakt region growing gevoelig voor ruis (gaten in het gesegmenteerde gebied) en het partieel volume effect (verbindingen tussen afzonderlijke structuren).

Verwant met deze methode zijn de split- en merge-algoritmes. Bij splitalgoritmes vertrekt men van het ganse beeld. Vervolgens splitst het algoritme het beeld zodanig op dat enkel nog gebieden overblijven met een homogene intensiteit. Region merging methodes werken omgekeerd. Zij gaan het beeld oversegmenteren en vervolgens op basis van a priori kennis, bepaalde gebieden opnieuw samenvoegen.

Classifiers

Classificatiemethodes gebruiken patroonherkenning op basis van gekende informatie over de textuur of de intensiteit van het beeld om het beeld op de delen in verschillende klassen van weefsels. Om een juiste opdeling te krijgen baseert het algoritme zich op trainingsdata. Dit is ook meteen één van de nadelen van deze methode. De gebruiker dient zelf deze trainingsdata op te geven en men spreekt daarom van een gesuperviseerde methode.

De voordelen van classificatiemethodes zijn de computationele efficiëntie van het algoritme, en het voordeel dat men op een eenvoudige manier het beeld kan opdelen in verschillende klassen. Het nadeel is dat maar weinig rekening wordt gehouden met de vorm van de structuur. Voor ingewikkelde vormen zoals bijvoorbeeld de witte materie in de hersenen, kan dit echter een voordeel zijn.

Enkele voorbeelden van classificatiemethodes zijn de k Nearest Neighbor (k NN) classifier, Parzen window en het Maximum Likelihood (ML) algoritme. Bij het k NN-algoritme worden de pixels geïdentificeerd volgens het naaste buren principe. Een pixel wordt geplaatst in die klasse waarin hij het meeste voorkomt volgens de k dichtste trainingsdata. Bij Parzen Window daarentegen wordt gekeken naar alle pixels binnen een bepaald venster en kiest men de klasse die het meeste voorkomt binnen deze range. Het ML-algoritme is statistisch geïnspireerd.

Een ander type van classificatiemethodes zijn de Atlas Guided methodes. Bij Atlas Guided methodes vertrekt men van een atlas van de anatomische structuur die men wilt segmenteren. Deze methodes kunnen daarom vergeleken worden met de classificatiemethodes met dit verschil dat men hier gebruik maakt van de spatiale informatie om een beeld in een bepaalde klasse op te delen. De anatomische structuren in de atlas worden geregistreerd op het beeld. Men zoekt de beste match. Dit principe noemt men atlas warping.

Clustering

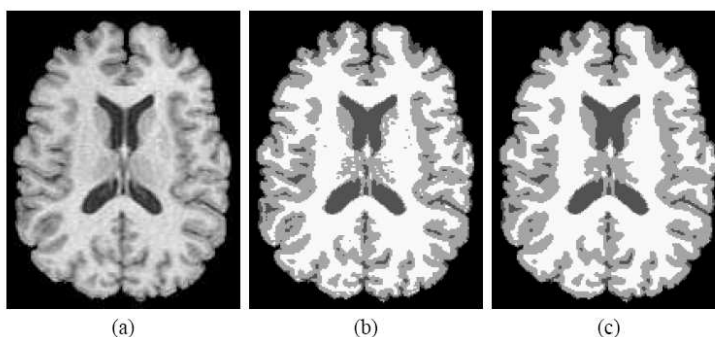
Clusteringmethodes zijn gelijkaardig aan de eerder vernoemde classificatiemethodes. Hier is echter dit verschil dat bij clustering geen trainingsdata wordt gebruikt. De voor- en nadelen zijn dezelfde als voor classifiers, met dit verschil dat de manuele trainingsstap wordt uitgeschakeld. Dit geeft aanleiding tot niet-gesuperviseerde methodes zoals het EM-, K-means- en fuzzy c-means algoritme. Bij het EM-algoritme vertrekt men van een schatting van het gemiddelde en de covariantie en worden de pixels volgens het ML-algoritme geïdentificeerd. Vervolgens wordt iteratief een nieuwe schatting gemaakt en een nieuwe classificatie. Het K-means algoritme berekent voor iedere klasse de gemiddelde intensiteit en plaatst op basis hiervan een pixel in de dichtstbijzijnde klasse. Dit wordt vervolgens iteratief geüpdatet. Het c-means algoritme is een veralgemeende versie van het K-means algoritme.

Als hulpmiddel voor deze segmentatiemethodes (bijv. K-means) kan een Markov Random Field (MRF) gebruikt worden. Dit is een statistisch model dat gebruik maakt van het spatiale kenmerk dat naburige pixels vaak tot dezelfde klasse behoren. Het toepassen van de MRF-methode heeft als gevolg dat de ruis in een beeld sterk gereduceerd wordt. Dit is vooral nuttig in MRI-beelden. Een voorbeeld van de werking van MRFs is weergegeven in Figuur 2.6.

Vervormbare modellen

Vervormbare modellen worden opgebouwd uit parametrische curven of oppervlakken die vervormen onder invloed van interne en externe krachten die werken op de curve. Hoeveel deze krachten doorwegen ten opzichte van elkaar wordt bepaald door een kostfunctie. Bij vervormbare modellen vertrekt men meestal van een eenvoudige meetkundige vorm die bovenop de structuur wordt geplaatst. Onder invloed van de kostfunctie levert dit dan een optimale segmentatie op.

Werken met vervormbare modellen heeft verschillende voordelen. Er ontstaat altijd een gesloten oppervlak. De curve kan gemakkelijk gesmooth worden voor visualisatiedoeleinden en het algoritme zelf kan zo robuust of gevoelig gemaakt worden afhankelijk van hoe accuraat men wil segmenteren. Nadeel is dat de initialisatie nog altijd manueel dient te gebeuren en dat vaak een groot aantal parameters dienen ingesteld te worden.

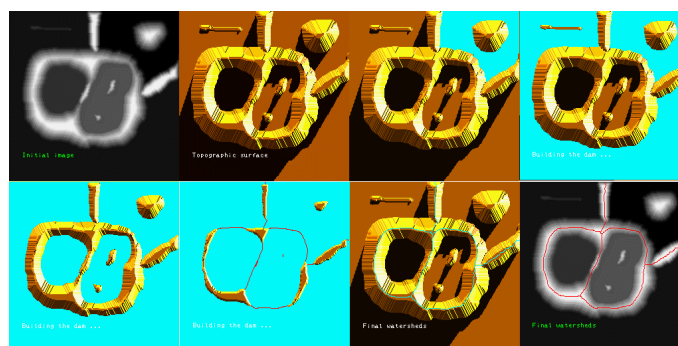


Figuur 2.6: Segmentatie : (a) origineel beeld, (b) segmentatie met behulp van K-means algoritme, (c) segmentatie met behulp van het K-means algoritme en het gebruik van MRF [16].

Waterscheiding

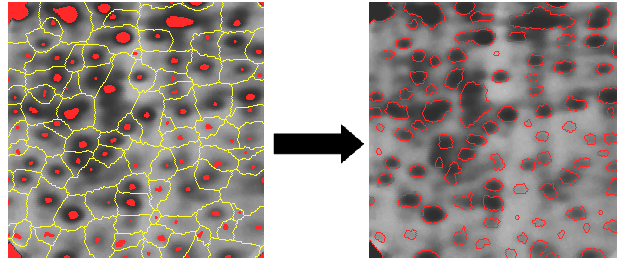
Waterscheiding is een methode die gebruikt wordt om het gebied in homogene delen op te delen. Meestal levert dit een overgesegmenteerd beeld op. Er is een postprocessing stap nodig om de oversegmentatie te compenseren en bij elkaar horende deeltjes opnieuw samen te voegen.

De opdeling in homogene gebieden gebeurt op basis van de topologie van het beeld. Elk grijswaardenbeeld kan beschouwd worden als een topografisch oppervlak. Indien men dit oppervlak laat onderlopen, zullen geleidelijk aan enkel op de hoogste plaatsen een scheidingslijn overblijven, de waterscheidingslijn. Een voorbeeld van een dergelijke waterscheiding is weergegeven in Figuur 2.7. We zien hier dat de verschillende bassins 1 voor 1 worden opgevuld en bijgevuld tot uiteindelijk alleen de waterscheiding op 1 lijn overblijft.



Figuur 2.7: De verschillende fases in het waterscheidingsalgoritme [17]: links boven het beeld voor segmentatie, rechts onder het beeld met segmentatie. De tussenliggende beelden geven het vollopen van het bassin aan.

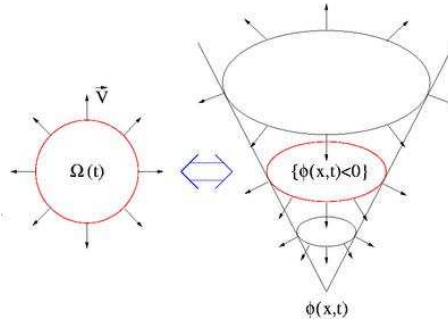
In het geval van een grijswaardenbeeld wordt de gradiënt genomen van het beeld om de topologie weer te geven. Een ideale werking van het algoritme zou de perfecte scheiding opleveren. Echter tengevolge van ruis en lokale onregelmatigheden treedt een oversegmentatie op. Een oplossing hiervoor is het gebruik van markers zodanig dat een marker gecontroleerde waterscheiding ontstaat. Het principe van markers bestaat eruit dat de gebruiker zelf bepaalde plaatsen aanduidt waar de overstrooming dient te starten. Dit levert een sterk verbeterde segmentatie op, zoals aangetoond in Figuur 2.8



Figuur 2.8: Marker gecontroleerd waterscheidingsalgoritme [17]

Level Sets

De methode van de level sets is een numerieke methode voor het volgen van de evolutie van structuren in een arbitraire dimensie, bijv. contouren of oppervlakken. In plaats van het rechtstreeks manipuleren van de structuur, is de structuur ingebed als een zero level set van een hoger dimensionale functie, de level set functie $\phi(x, t)$ genaamd. Dit is een impliciete functie. Dat wil zeggen dat $\phi(x, t)$ negatief is binnen de structuur en positief erbuiten. Een deelverzameling van de impliciete functies zijn de “signed distance” functies. Dit zijn impliciete functie waarvoor de absolute waarde de afstand tot de structuur (contour of oppervlak) voorstelt.



Figuur 2.9: Het concept van een zero set in een level set [18]

De level set functie $\phi(x, t)$ kan vervormen onder invloed van een snelheidsveld \vec{V} . Dit snelheidsveld is gedefinieerd in elk punt van de impliciete functie. Figuur 2.9 stelt de vervorming van een cirkelvormige contour voor onder invloed van een snelheidsveld met richting de normale op de contour [19]. Het rechterdeel van de figuur stelt verschillende tijdstappen voor. De verandering van ϕ is gegeven door de partiële differentiaalvergelijking [20]:

$$\frac{\partial \phi}{\partial t} + V_N |\nabla \phi| = 0$$

met V_N de normale component van \vec{V} op het oppervlak. Omdat alleen de verandering van de zero level set interessant is, is het niet nodig de nieuwe level set volledig te bepalen. Lokale level set methodes definiëren ϕ enkel in de buurt van de zero level set hetgeen rekenkundig minder intensief is. Op elk ogenblik kan de ontwikkelde structuur uit de zero level set gehaald worden: $\Gamma(\mathbf{x}, t) = \phi(\mathbf{x}, t) = 0$. Het belangrijkste voordeel van het gebruik van level sets is dat complexe vormen gemodelleerd kunnen worden. Een uitgebreide beschrijving van level set methodes is terug te vinden in Bijlage A.

2.4.3 Toepassingen voor segmentatie van metastasen

In de literatuur zijn verschillende artikels terug te vinden die de segmentatie van metastasen beschrijven of gebruikt kunnen worden voor de segmentatie van metastasen. Methodes voor wa-

terscheiding [21], region growing, vervormbare modellen (actieve contouren [21]) worden uitvoerig beschreven. Ondanks deze vele publicaties blijft de segmentatie van metastasen een uitdaging.

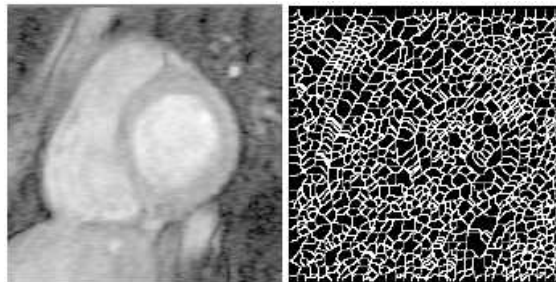
Waterscheiding - Paintbrush algoritme

In onderstaande paragraaf zijn een aantal artikels beschreven die een toepassing van het waterscheidingsalgoritme voor medische beeldsegmentatie beschrijven. In [13] en [14] wordt een methode voorgesteld waarin een object kan ingekleurd worden aan de hand van een intelligent penseel. Dit maakt het mogelijk een strategie te schrijven om de segmentatie van anatomische objecten te vereenvoudigen.

Vooreerst wordt het beeld opgedeeld in primitieve objecten. Deze worden bekomen door waterscheidingsgebieden met dezelfde intensiteit te verbinden aangezien deze met grote waarschijnlijkheid tot hetzelfde object behoren. Dit biedt een grotere computationele efficiëntie omdat men met gebieden werkt in plaats van individuele pixels. De opdeling in gebieden gebeurt onafhankelijk van de segmentatie waardoor het algoritme voor een waaier aan verschillende toepassingen kan aangewend worden. Men moet bijgevolg enkel nog die gebieden selecteren (inkleuren) waarvan men vermoedt dat ze tot het object behoren.

Bepaling van de primitieve gebieden

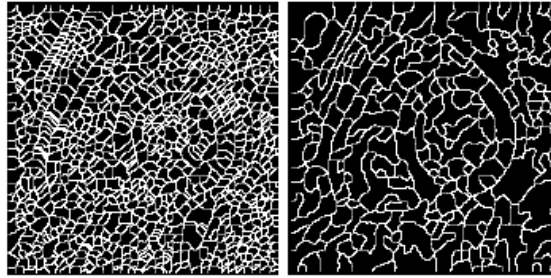
De waterscheiding van het gradiëntbeeld is een ideaal vertrekpunt om de randen in een beeld te detecteren. De echte waterscheidingstransformatie is echter te rekenintensief. Daarom wordt in deze paper pseudo-waterscheidingsgebieden geconstrueerd, gebruik makend van Maximum Gradient Paths (MGPs) [22]. Het beeld wordt gesmooth en in de omgeving van iedere pixel wordt de minimale gradiënt bepaald. Het algoritme stopt als er geen gradiënt in de omgeving wordt gevonden die kleiner is dan deze in de voorgaande pixel. Deze laatste pixel duidt dan de minimale gradiënt aan in een gebied. Iedere pixel krijgt dan het label van het minimale gradiënt waarin hij eindigt. Uiteindelijk bekomt men hieruit een overgesegmenteerd beeld bestaande uit de gewenste primitieve gebieden. Een voorbeeld is weergegeven in Figuur 2.10 [14].



Figuur 2.10: MR doorsnede van het hart (l) en waterscheiding van het gradiëntbeeld (r)

Samenvoeging van primitieve gebieden

De oversegmentatie van het waterscheidingsalgoritme kan gereduceerd worden door gebieden samen te voegen. De intensiteiten binnen 1 object verschillen vaak niet veel en kunnen samengevoegd worden zonder dat men informatie over de rand verliest. Daarnaast kan men ook de gebieden samenvoegen die een grote deel van hun rand gemeenschappelijk hebben. Deze criteria kunnen ingevoerd worden met behulp van het Minimum Description Length (MDL) principe. De optimale verdeling van het beeld is diegene waarvoor het totaal aantal bits om de beelddata te beschrijven in een bepaalde descriptieve taal minimaal is. Dit komt hier overeen met het aantal bits om de intensiteiten te coderen en het aantal bits om de grenzen tussen regio's te coderen. Het proces van samenvoegen gebeurt iteratief, waarbij in iedere stap de naburige gebieden die de grootste positieve description length gain opleveren, samengenomen worden. Dit proces wordt herhaald zolang er gebieden met positieve description length gain kunnen samengevoegd worden. Een voorbeeld is getoond in Figuur 2.11



Figuur 2.11: waterscheiding van het gradiëntbeeld (l) en gebieden na samenvoeging (r) [14]

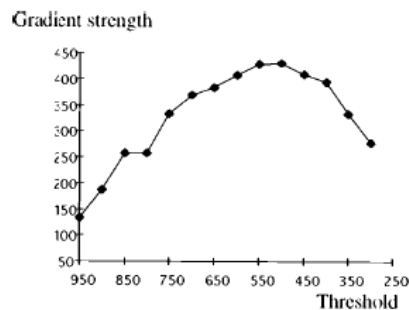
Besluit

Opdelen van het beeld in primitieve gebieden laat een snelle en robuuste computergesteunde segmentatie van anatomische objecten toe. Dit algoritme vraagt echter nog relatief veel gebruikersinteractie. Ook zal het segmenteren van verschillende objecten verschillende strategieën vereisen afhankelijk van de vorm- en intensiteitsverschillen.

2D multi-criteria segmentatie van longmetastasen

In [23] wordt een segmentatiemethode beschreven gebaseerd op een region growing binnen een vooraf geselecteerde Region Of Interest (ROI). Hierbij zijn enkele constraints opgelegd om uitgroei te beperken. De vormbeperking voorkomt dat andere structuren met een gelijkaardige densiteit, mee gesegmenteerd worden.

Allereerst wordt het beeld binnen de ROI gethreshold zodanig dat een binair beeld ontstaat. In verschillende stappen wordt de threshold verlaagd. Voor iedere threshold worden de gradiënt in de normale richting en de compactheid (bijv. cirkel als meest compacte vorm) van de kandidaat metastase bepaald. Op basis van de resultaten voor deze factoren wordt de optimale threshold geselecteerd. Een voorbeeld van de gradiëntsterkte in functie van de threshold is weergegeven in Figuur 2.12.



Figuur 2.12: Gemiddelde gradiëntsterkte van de contour van de metastase voor variërende threshold. De optimale is degene waarvoor de gradiënt maximaal is [23].

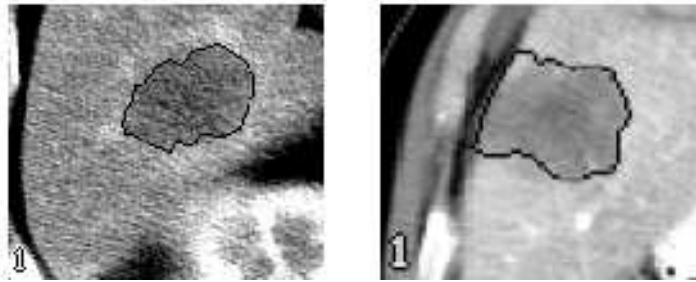
De compactheid die in rekening wordt genomen maakt het mogelijk om de longmetastasen te scheiden van omliggende bloedvaten met dezelfde intensiteit. Wanneer de vorm sterk verandert (bijv. bloedvat wordt mee gesegmenteerd), zal de compactheid sterk dalen. Dit bepaalt welke threshold moet genomen worden.

Waterscheidings- en actieve contour-algoritmes

Het principe van waterscheiding werd uitgelegd in 2.4.2. Dit deel gaat verder in op de toepassing van waterscheiding specifiek voor de segmentatie van levermetastasen. In [21] wordt een studie uitgewerkt over de segmentatie van levermetastasen op basis van waterscheiding. In deze studie zijn

levermetastasen van vijf verschillende patiënten opgenomen (in totaal 36 snedes). Iedere segmentatie werd tweemaal uitgevoerd om de reproduceerbaarheid van de methode na te gaan. Het artikel vergelijkt de resultaten van waterscheiding met een manuele aflijning. Uit de evaluatie van de resultaten blijkt dat het waterscheidingsalgoritme in vele gevallen grote segmentatiefouten oplevert. Slechts in 17 van de 36 snedes is het algoritme succesvol. Slecht 2 van de 10 aanwezige metastasen zijn volledig juist gesegmenteerd. De meeste fouten treden op indien de tumor aan de rand van de lever gelegen is. Daarnaast heeft het Tweedimensionaal (2D)-waterscheidingsalgoritme ook nog het nadeel dat het veel gebruikersinteractie vereist. Voor iedere snede moet de gebruiker een inwendige en uitwendige ROI aanduiden. Waterscheiding is bovendien een stuk trager dan andere methodes zoals region growing. Vooral op grote beelden zoals de hier gebruikte Driedimensionaal (3D) CT beelden vormt dit een probleem.

Naast waterscheiding beschrijft artikel [21] ook de resultaten van een Active Contour Algoritme voor de segmentatie van levermetastasen. Men vertrekt van een contour gebaseerd op een eenvoudige geometrische figuur die vervolgens vervormd wordt naar de randen van de metastasen. Dit gebeurt eveneens door een kostenfunctie. Vervolgens wordt ook het dual-scale active contour algoritme besproken. Dit algoritme heeft als belangrijkste wijziging dat men op twee schalen gaat werken. Eerst wordt op grote schaal het volledige beeld gesmooth. Vervolgens wordt het active contour algoritme op kleine schaal toegepast. Op deze manier worden lokale minima gereduceerd en de segmentatiefout verkleind. Een voorbeeld van het resultaat is terug te vinden in figuur 2.13.



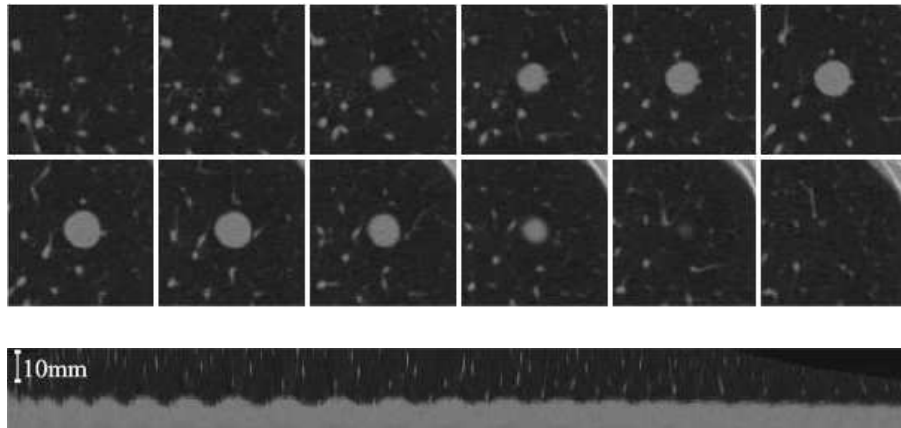
Figuur 2.13: Resultaten van dual scale active contour algoritme op de segmentatie van levermetastasen [21]

Het artikel vermeldt enkel de resultaten van de dual scale active contour methode. Deze geeft in 35 op 36 snedes een bevredigend resultaat. Vooral metastasen aan de rand van de lever gaven segmentatiefouten. Het reproduceren van de segmentatie gaf verwaarloosbare verschillen. Het nadeel van deze methode is echter dat veel gebruikersinteractie vereist is. Op iedere snede moet een geometrische figuur getekend worden.

Spiral Scanning Technique voor segmentatie van longmetastasen

Dit artikel beschrijft een segmentatiemethode ontwikkeld voor de detectie van longmetastasen [24]. In dit artikel wordt gebruik gemaakt van een spiral scanning techniek om over te gaan van een 3D Volume Of Interest (VOI) naar een 2D beeldruimte. Het beeld wordt vervolgens in 2D gesegmenteerd m.b.v een dynamisch programmeringsalgoritme. Vervolgens wordt het resultaat teruggetransformeerd naar 3D. Deze techniek heeft als voordeel dat men de 3D informatie van het beeld behoudt en tegelijkertijd de berekeningen relatief eenvoudig in 2D kan uitvoeren.

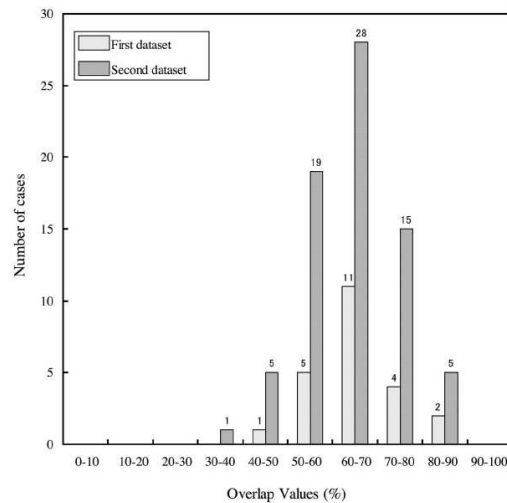
Om over te gaan van de 3D ruimte naar de 2D ruimte wordt de 3D sferische VOI beschreven door een aantal radiale lijnen waarbij de pixels op vaste afstanden liggen op de radiale lijn. Uit deze radiale lijnen wordt een spiraal gevormd door de elevatiehoek voor ieder punt te laten variëren. Vervolgens wordt deze spiraal geresampled om alle punten evenredig te verdelen over de spiraal. Door deze transformatie wordt de segmentatie aanzienlijk vereenvoudigd. In Figuur 2.14 wordt een voorbeeld gegeven van deze transformatie toegepast op een 3D CT beeld van een longmetastase.



Figuur 2.14: 3D snedes van longmetastasen(bovenaan) en transformatie naar 2D(onderaan) [24]

Het 3D beeld is getransformeerd naar een 2D curve. Via dynamisch programmeren wordt iteratief de optimale segmentatierand berekend op basis van de kostenfunctie om van de vorige naar de volgende stap te gaan. De kostenfunctie is opgebouwd uit een interne en een externe kost. De interne kost bepaalt de smoothness van de rand zowel in longitudinale als laterale richting. De externe kost geeft de randsterkte weer. Door een gewogen som van de interne en externe kosten te minimaliseren, bekomt men de optimale segmentatierand. Ten slotte moet het beeld opnieuw gereconstrueerd worden in de 3D beeldruimte.

De resultaten van deze methode worden gevalideerd aan de hand van de 3D volume-overlapping met een manuele segmentatie. De resultaten voor beide datasets worden weergegeven in Figuur 2.15.



Figuur 2.15: Histogram van overlapwaarden voor de metastasen in de 2 datasets [24]

Voor de eerste dataset wordt een gemiddelde overlap bereikt van 66%, voor de tweede een overlap van 64%.

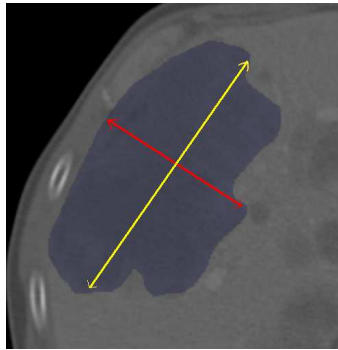
2.5 Kwantificatie van metastasen

De verandering van grootte van een metastase is de belangrijkste parameter om het resultaat van een experimentele behandeling op te volgen en om eventueel de behandeling bij te sturen. Oncolo-

gen en artsen op radiotherapie kunnen op basis hiervan bijvoorbeeld inschatten of een bepaalde chemotherapie goed aanslaat of niet. Daarom is het van groot belang de grootteverandering kwantitatief te kunnen bepalen. Hiervoor is een accurate bepaling van de grootte noodzakelijk. De volgende paragraaf beschrijft hoe dit de dag van vandaag gebeurt. Sectie 2.5.2 legt een alternatief uit waarvan de uitwerking de kerntaak van deze thesis zal zijn.

2.5.1 RECIST

Tot nu toe gebeurt de meting van deze grootteverandering vrij primitief, namelijk met behulp van een uni- of bidimensioneel criterium. Oorspronkelijk was de standaard een bidimensionele meting, maar studies hebben aangetoond dat een unidimensionale meting haast even goede resultaten geeft. In de praktijk komt dit alles neer op meting van de grootste diameter op een axiale snede, ook wel Response Evaluation Criteria In Solids Tumors (RECIST) genoemd. Sinds 2000 wordt RECIST als standaard gebruikt. Men spreekt van een volledige respons op een behandeling als de tumor volledig is verdwenen. Een gedeeltelijke respons is een daling van de som van de langste diameters van 30% of meer. Men spreekt van een progressieve ziekte als de som van de langste diameters stijgt met 20% of meer. Daartussen is de ziekte stabiel [25]. Een voorbeeld van een bidimensionale meting is weergegeven in Figuur 2.16. Het product van RECIST met de langste diameter van de metastase loodrecht op RECIST, noemt men de World Health Organization criteria (WHO).



Figuur 2.16: Meting van RECIST (geel) en zijn grootste orthogonale diameter (rood)

Deze eenvoudige metingen zijn niet volledig correct. Een metastase heeft namelijk geen eenvoudige vorm die met één of twee parameters kan beschreven worden. De enige reden waarom deze maatstaven gebruikt worden, is hun eenvoud. Het manueel aflijnen van alle metastasen is erg tijdrovend en een simpele RECIST-meting geeft reeds een goede indicatie van de toe- of afname van de tumorgrootte. In het verdere verloop van de thesis kan RECIST zowel slaan op het meten van de grootste diameter als op het lijnstuk zelf.

2.5.2 Volume

Een alternatief voor het bepalen van de grootteverandering van de metastasen is door gebruik te maken van volumemetingen op verschillende tijdstippen. Intuïtief voelt men makkelijk aan dat het verschil in volume van een metastase op verschillende tijdsogenblikken een betere maat is voor het kwantificeren van de grootteverandering dan door gebruik te maken van uni- of bidimensionele criteria. Ook uit verschillende studies blijkt dat het volume een betere maat is om metastasen te karakteriseren dan RECIST. Hierna volgen enkele studies die een vergelijking maken tussen uni- of bidimensionale metingen en volumemetingen.

In een studie [26] uit 2002 werden de effecten van een behandeling van levermetastasen bij borstkankerpatiënten geëvalueerd. Dit gebeurde enerzijds aan de hand van uni- en bidimensionale metingen en anderzijds aan de hand van een manueel afgelijnd volume. Uit deze studie met

37 testpatiënten blijkt dat evaluatie van de resultaten uit uni- en bidimensionale metingen in respectievelijk 12 en 13 gevallen verschilt met de evaluatie bij volumemetingen.

Een gelijkaardige studie [27] toegepast op longmetastasen gaf een discordantie tussen bidimensionale en volumemetingen in 7 op 30 gevallen. Tussen unidimensionale en volumemetingen was een discordantie in 1 op 30 gevallen. Alhoewel niet alle studies sluitende resultaten geven, is het duidelijk dat volumemetingen weldegelijk meer precieze en beter reproduceerbare resultaten geven. Indien zij geautomatiseerd worden, kan dit een waardevolle aanvulling zijn voor de evaluatie van een behandeling.

Ook in [12] wordt gewezen op de voordelen van een driedimensionale meting. Ten eerste geeft dergelijke meting een objectieve maat voor het absolute volume van een tumor. Ten tweede maakt zo'n meting de evaluatie van veranderingen in de tijd van de craniocaudale diameter mogelijk. Ze kan bovendien beter overweg met onregelmatige massa's. Nadeel is echter dat de derde dimensie van de tumor moeilijker beschikbaar is.

In [28] uit 2007 worden 1D- en 2D-metingen van tumoren in de borst, buik en bekken vergeleken met volumemetingen voor de evaluatie van de verandering in tumorgrootte. Hieruit blijkt dat 2D tumormetingen, gebruik makend van de WHO, beter de verandering van het tumorvolume weergeven dan 1D metingen, gebruik makend van de RECIST en bovendien de ruis in de meting beter onderdrukken.

Het meten van het volume van een metastase, hetgeen de kern van de opdracht is, vereist een segmentatie van die metastase, bij voorkeur met zo weinig mogelijk gebruikersinteractie. Eens dit gebeurd is, is het eenvoudig om het volume te bepalen door het aantal voxels in de segmentatie te vermenigvuldigen met de voxelgrootte. Het probleem van volumemetingen komt dus in essentie neer op het zo goed mogelijk segmenteren van de metastasen. Belangrijk hierbij is de segmentatie ook te valideren a.d.h.v. handmatige aflijningen.

2.5.3 Choi-parameters

De laatste jaren hebben twee ontwikkelingen RECIST in vraag gesteld [25].

1. Beeldvormingstechnieken laten meer en meer toe om kleinere veranderingen in tumorgrootte te detecteren. Deze beeldvormingstechnieken kunnen ook veranderingen in de dichtheid van de tumor weergeven, hetgeen ook een belangrijke indicator voor de therapeutische respons blijkt te zijn zelfs wanneer de tumor niet krimpt.
2. Er komen steeds meer geneesmiddelen ter bestrijding van kanker op de markt. Sommige inhiberen de celtgroei (cytostatica), andere, zoals de conventionele chemotherapie, doden kankercellen (cytotoxische geneesmiddelen). Cytostatica zorgen er dus niet voor dat de tumor inkrimpt, zodat hiervoor RECIST geen goede maat kan zijn.

In [29] stelt Haesun Choi nieuwe criteria voor om de tumorrespons te meten. In plaats van de verandering van langste diameter te vergelijken op verschillende tijdsogenblikken, gaat Choi zowel het tumorvolume als de dichtheid van de tumor na. De dichtheid kan gemeten worden op contrast-versterkte CT-scans. Er is sprake van een (gedeeltelijke) respons vanaf een afname van 10% in het tumorvolume *of* een afname van 15% in de tumordichtheid op CT-beelden.

In een klinische studie [30] wordt aangetoond dat de Choi responscriteria gevoeliger en preciezer zijn voor het meten van de respons van Gastrointestinal Stromal Tumors (GISTs) behandeld met imatinib. Dit type behandeling zorgt er namelijk vooral voor dat de dichtheid van de tumor daalt.

Om de Choi criteria te meten moet dus zowel het volume als de gemiddelde dichtheid van de tumor gekend zijn. Hiervoor is een segmentatie vereist. Uit de gesegmenteerde metastasen kan makkelijk het volume bepaald worden, maar ook de gemiddelde dichtheid door de segmentatie te gebruiken als masker op het originele beeld. Optellen van de grijswaarden binnen dit masker en delen door het aantal voxels in het masker levert de gemiddelde intensiteit van de tumor.

Hoofdstuk 3

Karakteristieken van de metastasen in CT beelden

Vooraleer met de segmentatie van de metastasen te beginnen, is het belangrijk meer te weten te komen over specifieke kenmerken van metastasen in de beelden. Op die manier is het mogelijk deze kenmerken te gebruiken in het segmentatie-algoritme om tot een zo efficiënt en correct mogelijke segmentatie te komen. Eerst volgt de bespreking van de gebruikte data, gevolgd door de analyse van de fotometrie (3.2) en de geometrie (3.3).

3.1 Dataset

Voor deze thesis maken we gebruik van 11 trainingsdatasets afkomstig van twee patiënten en opgenomen in het UZ Leuven. Deze datasets bevatten in totaal 31 metastasen die door clinici¹ manueel afgelijnd werden. Daarnaast werden op de 31 metastasen ook RECIST en de loodrechte op RECIST aangeduid. De grootte van deze metastasen varieert tussen 4,39 ml en 775,90 ml met een gemiddelde van 146,50 ml. De beelden zijn gegenereerd in een CT-scanner, met de “Siemens CT Volume Zoom - 2000” (4 slice), de “Siemens Somatom Plus 4” (single slice) of de “GE HiSpeed CTe” (single slice). De beelden werden gereconstrueerd met 4096 grijswaarden (-1024 tot 3072 HU). Iedere slice bevat 512×512 voxels met een voxelgrootte van 0,68 mm tot 0,82 mm. De datasets bestaan uit 37 tot 64 sneden met een snededikte van 5 mm tot 7 mm.

Het protocol waarmee de beeldopname gebeurde is een CT abdomen drie-fasenprotocol. De eerste fase is zonder contraststof. De tweede, arteriële fase start wanneer de trigger in de aorta 140 HU bereikt heeft. De contraststof bevindt zich dan voornamelijk in de slagaders. De derde en laatste fase is de veneuze fase. De contraststof is dan vooral aanwezig in de aders. Afwijkingen van dit protocol gebeuren frequent aangezien de meting patiëntafhankelijk is. De gebruikte trainingsdata zijn allemaal opgenomen tijdens de veneuze fase. De gebruikte contraststof is op basis van jodium (30 tot 40 g per inspuiting).

3.2 Fotometrische eigenschappen

Het onderzoeken van fotometrische eigenschappen van metastasen en van de lever kan nuttige informatie opleveren voor bepaalde segmentatiemethodes, zoals statistische pixelclassificatie. In de komende paragrafen gaat de aandacht vooral uit naar de intensiteitswaarden in het beeld en de textuur.

¹Dr. Hans Prenen, Dr. Caroline De Vleeschouwer en prof. Dr. Sabine Tejpar

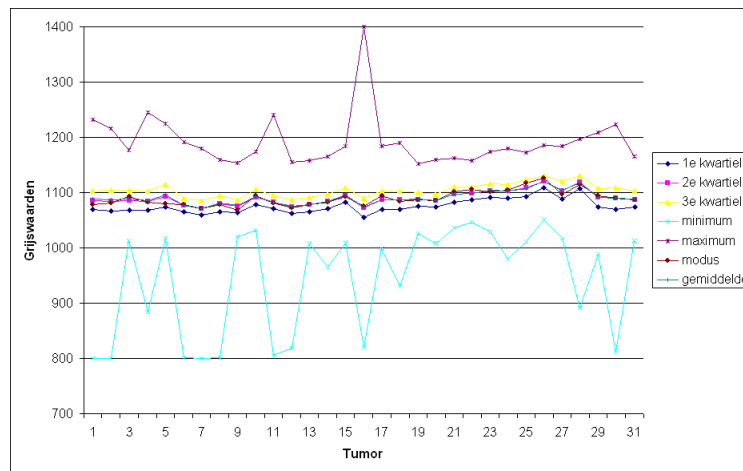
3.2.1 Analyse van de grijswaarden

Meestal wordt de beeldintensiteit uitgedrukt in Hounsfield Unit (HU). In het verdere verloop van de thesis gebruiken we echter Grijswaarden (GW) waarbij geldt dat: $GW=HU+1024$.

Grijswaarden per metastase

Een eerste onderzoek betreft de grijswaarden. Dit is nuttig voor functies die op de één of de andere manier gebruik maken van *thresholding*. Hierbij is de waarde van de threshold bepalend voor het opdelen van het beeld.

Opdat de informatie over grijswaarden zinvol zou zijn, is een onderzoek nodig van verschillende metastasen. Figuur 3.1 geeft de statistieken van de grijswaardenverdeling van de metastasen. Opvallend hierbij is dat de meeste grijswaarden geconcentreerd zijn in een klein interval rond 1100. Dit sluit het gebruik van een thresholding tijdens segmentatie dus zeker nog niet uit. Ook de gemiddelde standaarddeviatie van de tumoren heeft een kleine waarde, nl. 22,5 (t.o.v. verwachtingswaarde 1090), wat dus duidelijk maakt dat de grijswaarden binnen een metastasen weinig verschillen.



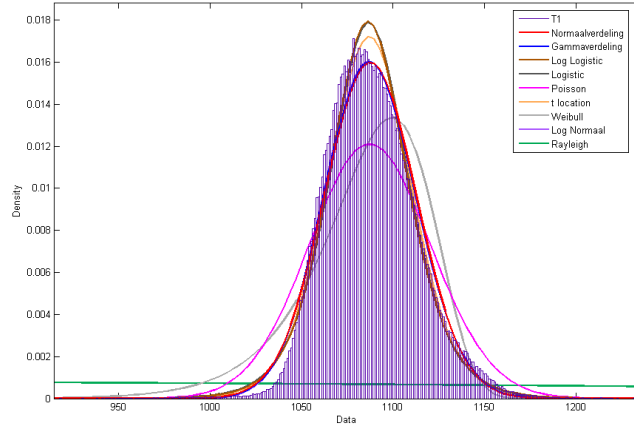
Figuur 3.1: Statistieken van de grijswaardenverdeling

De aanname dat de grijswaarden normaal verdeeld zijn, is een goede aanname. Dit blijkt uit de statistieken. Enerzijds moet in een normale verdeling de mediaan gelijk zijn aan het gemiddelde. Voor de 31 testmetastasen wijkt de mediaan maximaal 0,28% af van het gemiddelde. Anderzijds moet het verschil van Q_{25} en Q_{75} theoretisch gelijk zijn aan $1,36 \sigma$. In werkelijkheid is dit verschil gemiddeld gelijk aan $1,30 \sigma$.

Het verschil tussen de waarde van het eerste en het derde kwartiel blijft vrij constant over de verschillende tumoren, ook al varieert de waarde van het eerste en derde kwartiel. Dit wil dus zeggen dat als het eerste kwartiel stijgt de kans dat dan ook het derde kwartiel stijgt groter wordt, wat statistisch neerkomt op een kleine standaarddeviatie. De gemiddelde waarde van deze standaarddeviatie over alle tumoren is ongeveer half zo groot als het verschil tussen de minimale en de maximale gemiddelde waarde van alle tumoren. Dit betekent dat het nuttiger is om gebruik te maken van een relatieve threshold dan van een absolute threshold.

Nader onderzoek buigt zich over de verdeling van de grijswaarden met de bedoeling te controleren of de grijswaarden daadwerkelijk normaal verdeeld zijn.

Op Figuur 3.2 is een histogram getekend van de grijswaarden van één tumor. Daarnaast staan de kansdichtheidsfuncties van enkele belangrijke verdelingen met elk hun beste fit. Hier blijkt dat het moeilijk te zeggen of de grijswaarden van de tumor verdeeld zijn volgens één van de verdelingen.



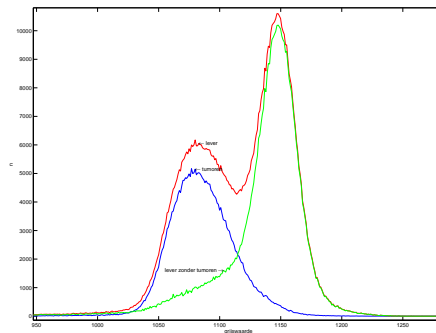
Figuur 3.2: Voorbeeld van de verdeling van de grijswaarden in 1 metastase

Er is een systematische afwijking vast te stellen tussen de grijswaardenverdeling van een metastase en de normale verdeling. In een metastase zijn er namelijk meer uitlopers voor grote waarden dan voor kleine waarden: de rechterstaart is groter dan de linkerstaart. De verklaring hiervoor ligt in het niet lineair partieel volume-effect. Hierdoor verschuiven de grijswaarden van de metastase in de richting van de grijswaarden van het omliggende weefsel, in dit geval de lever zelf. Het gevolg is dus dat de metastasen een langere staart hebben in de richting van de grijswaarden van de rest van de lever. Dit is de reden dat de lognormaal-verdeling beter past aan het histogram van de metastase.

De bekomen informatie omtrent grijswaarden kan nuttig gebruikt worden door na te gaan of de verdeling van de voxels in een (semi-)automatische afijning van de metastase voldoet aan een normale verdeling met een standaarddeviatie van de grootte-orde van de hier bekomen standaarddeviaties.

Grijswaarden van de volledige lever

De veronderstelling is dus dat de grijswaarden van de verschillende metastasen ongeveer normaal verdeeld zijn. Dit is niet zo voor de volledige lever, waarin metastasen aanwezig zijn. Een voorbeeld van een grijswaardenverdeling van een pathologische lever is gegeven in Figuur 3.3.



Figuur 3.3: Voorbeeld van de verdeling van de grijswaarden in de lever

Op deze figuur zijn duidelijk twee pieken te onderscheiden, waarvan de eerste piek zich rond

de gemiddelde waarde van de grijswaarden van de metastasen bevindt. De figuur geeft ook de grijswaardenverdeling van de tumoren in die lever en de grijswaardenverdeling van de lever zonder tumoren weer. Hierbij moet wel opgemerkt worden dat de lever zonder tumoren wel degelijk niet afgelijnde tumoren kan bevatten. Enkel de vijf grootste tumoren worden afgelijnd door de clinici. We kunnen dus concluderen dat de eerste piek een belangrijke component heeft die te wijten is aan de tumoren, de tweede piek voornamelijk aan leverweefsel.

Over een gezonde lever hebben we geen gegevens. We vermoeden dat daarin de grijswaarden wel normaal verdeeld zijn.

Grijswaarden van alle metastasen van alle patiënten

Het is ook mogelijk uitspraken te doen over de verdeling van de grijswaarden, niet binnen één metastase, maar over alle metastasen van alle patiënten. Hiervoor bekijken we twee benaderingen. De eerste benadering bekijkt de metastasen apart, de tweede benadering bekijkt niet de metastasen, maar de grijswaarden van de tumoren onafhankelijk van welke tumor ze afkomstig zijn. De eerste benadering gebruikt de veronderstelling dat de metastasen onderling onafhankelijk zijn, de tweede benadering gebruikt deze veronderstelling niet. Bovendien kan de tweede benadering als controle dienen voor de eerste.

- *Metastasen als geheel bekeken*

De eerste benadering bekijkt dus de metastasen als geheel en probeert uitspraken te doen over de grijswaardenverdeling in alle metastasen. Hiervoor gebruiken we de centrale limietstelling.

De centrale limietstelling of stelling van Lindeberg-Levy geeft aan dat de som van een groot aantal onderling onafhankelijke en gelijk verdeelde stochastische variabelen met eindige variantie bij benadering een normaalverdeling heeft. De variabelen zelf behoeven daarvoor geen normaalverdeling te hebben.

Hierdoor veronderstellen we dus dat de verschillende metastasen onderling onafhankelijk zijn, ook al is dit waarschijnlijk niet voor metastasen binnen eenzelfde lever, en dat de grijswaardenverdeling dezelfde is voor de verschillende metastasen.

In de tweede benadering worden de metastasen niet onderling onafhankelijk verondersteld. Aangezien er slechts 31 afgelijnde metastasen ter beschikking zijn, is het aantal stochastische variabelen beperkt. De *student t-verdeling* geeft een benadering voor de normale verdeling in het geval van een klein aantal variabelen [31]. Gegeven N onafhankelijke metingen X_i , beschouw

$$t \equiv \frac{\bar{X} - \mu}{s/\sqrt{N}}, \quad (3.1)$$

waar μ het populatiegemiddelde is, \bar{X} het steekproef gemiddelde ($\bar{X} = 1089,8$), en s de schatter voor de populatie standaarddeviatie, gedefiniëerd door

$$s^2 \equiv \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$$

De waarde hiervan voor de 31 metastasen is $s = 12,0$. Dan is de kansdichtheidsfunctie voor t :

$$f(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)} (1+t^2)^{-(\nu+1)/2} \quad (3.2)$$

met ν het aantal vrijheidsgraden ($\nu = n - 1$) en Γ de gamma-functie. Op basis hiervan geldt dat de kans dat t binnen het volgende interval 90% is:

$$P(-A < t < A) = 0,90 \Rightarrow A = 1,697$$

Gebruik makend van vergelijking 3.1 kunnen we zeggen dat de kans dat het werkelijke populatiegemiddelde, hier de werkelijke gemiddelde grijswaarde van alle metastasen van alle patiënten, 90% is als het binnen het volgende interval ligt:

$$P(1086, 2 < \mu < 1093, 5) = 0,90 \Rightarrow \mu \in [1086, 2; 1093, 3] \quad (3.3)$$

Analoog geldt:

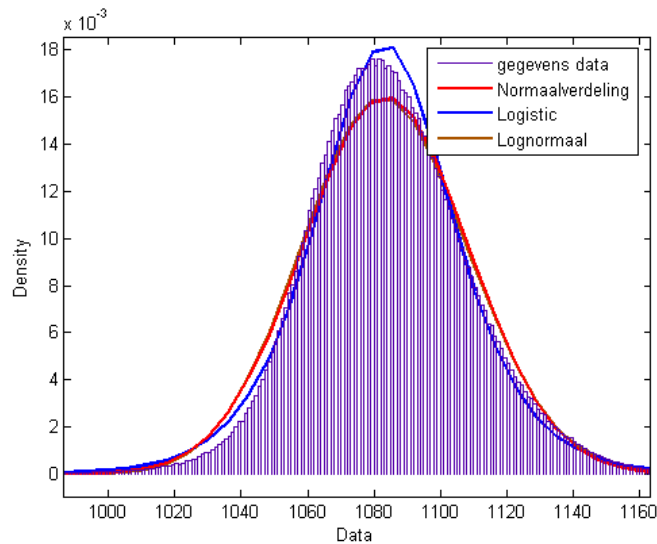
$$\begin{cases} P(1085, 4 < \mu < 1094, 2) = 0,95 \\ P(1083, 9 < \mu < 1095, 8) = 0,99 \\ P(1082, 0 < \mu < 1097, 7) = 0,999 \end{cases}$$

Het kenmerk dat de gemiddelde grijswaarde van eender welke metastase met grote zekerheid binnen een redelijk smal interval ligt, kan gebruikt worden tijdens segmentatie. Doordat de gemiddelde grijswaarde binnen een klein interval ligt, kan ook een absolute threshold bekoerende resultaten geven.

Hierbij dient wel opgemerkt te worden dat de gemiddelde grijswaarde binnen een metastase afhangt van het type kanker. Ook primaire tumoren kunnen verschillen van deze grijswaarden. In die zin duidt het smalle interval waarbinnen de gemiddelde grijswaarde zich bevindt, erop dat op basis van deze beelden niet het onderscheid kan gemaakt worden tussen verschillende types kanker.

- *Grijswaarden bekeken*

Een andere mogelijke benadering van het onderzoek naar grijswaarden is de grijswaarden te bekijken onafhankelijk van welke metastase ze afkomstig zijn. Figuur 3.4 geeft het histogram van alle grijswaarden, samen met enkele belangrijke verdelingen.



Figuur 3.4: Voorbeeld van de verdeling van de grijswaarden in de lever

De grijswaarden zijn bij benadering normaal verdeeld, met een gemiddelde van $\mu = 1083,58$ en een standaarddeviatie van $\sigma = 24,93$. Een wiskundige beschrijving van de kansdichtheidsfunctie van een normale verdeling is gegeven in volgende vergelijking:

$$f(I_k | \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(I_k - \mu)^2}{2\sigma^2}\right) \quad (3.4)$$

Ook hier weer is een afwijking zichtbaar, te wijten aan het niet lineair partieel volume-effect.

Zoals Figuur 3.4 aantoont, sluit de logistische verdeling beter aan bij de grijswaardenverdeling dan de normale verdeling. De kansdichtheidsfunctie van een logistische verdeling is [32]:

$$f(I_k|\mu, s) = \frac{\exp(\frac{I_k - \mu}{s})}{s(1 + \exp(\frac{I_k - \mu}{s}))^2} \quad (3.5)$$

Hierin is $s = \frac{\sigma\sqrt{3}}{\pi}$.

Op dit ogenblik is het niet mogelijk een fysische interpretatie te geven aan deze logistische distributie in het kader van medische beelden.

Besluiten

Het onderzoek naar de grijswaarden levert volgende besluiten op:

1. De grijswaarden binnen een metastase zijn bij benadering normaal verdeeld. De systematische afwijking is te wijten aan het niet lineair partieel volume effect.
2. De standaarddeviatie van de grijswaarden binnen de tumoren is klein, zodat een relatieve threshold betere resultaten zal geven dan een absolute threshold.
3. De grijswaarden van de metastasen bevinden zich in een ander interval dan de grijswaarden van de rest van de lever met een overlap die de segmentatie een stuk moeilijker maakt.
4. De gemiddelde grijswaarde van alle mogelijke levermetastasen van een bepaald type ligt met grote zekerheid in een smal interval.

3.2.2 Analyse van de textuur

Naast de grijswaarde van de metastase, vormt de textuur een tweede distinctief kenmerk van de lever en de metastasen. Textuur is een maat voor de variatie van de intensiteit van een beeldregio en kwantificeert eigenschappen zoals de gladheid, grofheid, homogeniteit en regelmatigheid van een oppervlak [33]. Textuuranalyse is van belang bij de classificatie van weefsels van verschillende organen. De vorm van het orgaan is vaak niet consistent en varieert snel doorheen de verschillende snedes van een CT-scan. Daarnaast is er ook een grote overlap in grijswaarden tussen verschillende organen. Binnen het weefsel van één orgaan kan men echter wel verwachten dat bepaalde aspecten van de textuur consistent zijn [34]. We gaan hier van de veronderstelling uit dat tumor- en leverweefsel op een verschillende manier zijn opgebouwd en dat bijgevolg op basis van de verschillende textuurkenmerken een onderscheid kan gemaakt worden.

In de literatuur worden drie technieken besproken om texturen te analyseren: Run Length Encoding (RLE), co-occurrence matrices en een wavelet-gebaseerde textuuranalyse. Meer details hierover zijn terug te vinden in Bijlage B, deel B.1.

Op één van deze technieken, namelijk textuuranalyse op basis van co-occurrence matrices, gaat een eerste paragraaf dieper in. Vervolgens wordt de gemiddelde textuur van metastase- en leverweefsel besproken om de meest onderscheidende textuurkenmerken te selecteren voor de textuuranalyse per voxel.

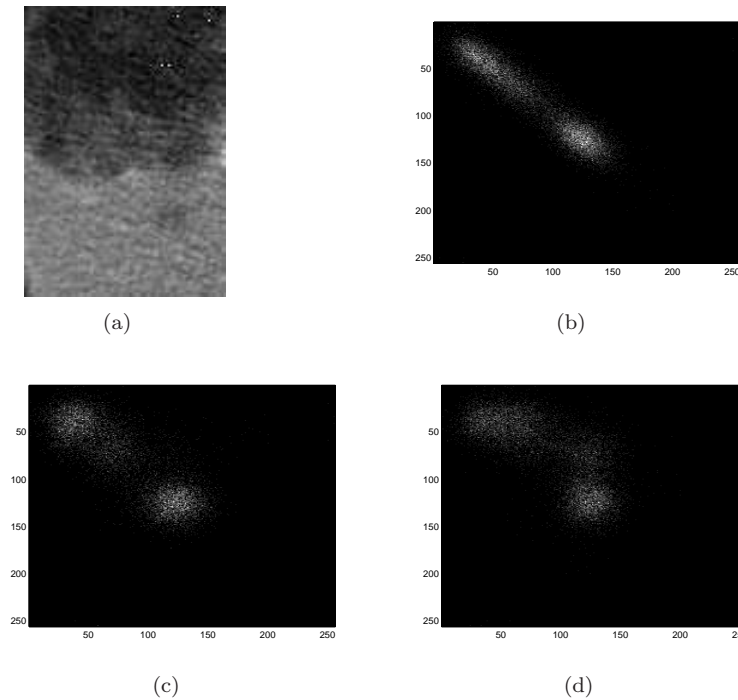
Textuuranalyse op basis van de co-occurrence matrix

Gebruikmakend van de methode beschreven in sectie B.1 is een algoritme ontwikkeld om de co-occurrence matrix te bepalen.

Om de verwerking van het beeld te versnellen en het aantal mogelijke combinaties te verlagen, worden de grijswaarden eerst in bins van 16 grijswaarden gestopt. Dit komt overeen met een downsampling van 4096 naar 256 grijswaarden. Dit verhoogt de computationele efficiëntie van het algoritme. Vervolgens gebruikt men formule 3.6 om de co-occurrence matrix te berekenen.

$$C(i, j) = \sum_{x=1}^m \sum_{y=1}^n \sum_{z=1}^o \begin{cases} 1, & \text{als } I(x, y, z) = i \text{ en } I(x + \Delta x, y + \Delta y, z + \Delta z) = j \\ 0, & \text{elders} \end{cases} \quad (3.6)$$

Hierin is I het gegeven CT-beeld en $\Delta x, \Delta y, \Delta z$ de verplaatsingen in de x,y en z-richting om in de buurpixel terecht te komen. Deze verplaatsingen zijn afhankelijk van de richtingen waarin men de matrix berekent en de afstand tussen de referentie-pixel en de buur-pixel. Een voorbeeld is weergegeven in Figuur 3.5. Hier berekenen we de co-occurrence matrix van een deelbeeld dat zowel leverweefsel (licht) als tumorweefsel (donker) bevat. De resultaten zijn verschillend afhankelijk van de richting en de pixelafstand waarover we berekenen. Indien bijvoorbeeld pixelafstand 1 gekozen wordt zoals in Figuur 3.5(b) krijgen we twee dichte wolken. Indien de pixelafstand 20 wordt gekozen zullen de twee wolken veel meer gespreid zijn. Dit is te zien in Figuur 3.5(c). Ook de richting heeft een invloed op de vorm van de co-occurrencematrix. In Figuur 3.5(c) is de richting horizontaal gekozen. In Figuur 3.5(d) echter is de richting verticaal gekozen. Hierdoor ontstaat een derde wolk die overeenstemt met de overgangen tussen lever- en tumorweefsel. Algemeen wordt aangenomen dat de grootte van de verplaatsingen zodanig wordt gekozen dat ze in de grootte-orde liggen van de patronen of texturen die men zoekt [35].



Figuur 3.5: Van een sample(a) met lever- en tumorweefsel wordt de co-occurrencematrix berekend voor een pixelafstand van respectievelijk 1(b) en 20(c) in de verticale richting. In (d) gebeurt dezelfde berekening in de horizontale richting.

In dit geval zijn de gekozen afstanden $D=1,2,4,6$ en 8 pixels. De gekozen richtingen zijn weergegeven in Tabel 3.1. Hierbij is θ de hoek gemeten in het XY-vlak vertrekkende bij 0 in de positieve x-richting en ϕ de hoek gemeten boven en onder het XY-vlak. Vervolgens wordt de co-occurrence matrix genormaliseerd:

$$C_n(i, j) = \frac{C(i, j)}{\sum_{k, l} C(k, l)} \quad \forall i, j$$

De grootte van het beeld waarop men de berekeningen uitvoert, speelt in dat geval geen rol. Nadat het algoritme de co-occurrence matrix bepaald heeft, worden de textuurkenmerken uit Tabel 3.2 berekend.

Vervolgens kan men een aantal statistieken berekenen van de co-occurrence matrix $C(i, j)$ die kunnen verschillen naargelang de textuur. Deze zijn weergegeven in Tabel 3.2.

Tabel 3.1: Verplaatsingsvectoren van de co-occurrence matrix

Richting (θ, ϕ)	Verplaatsingsvector
(0°, 45°)	(D,0,D)
(0°, 90°)	(D,0,0)
(0°, 135°)	(D,0,-D)
(45°, 45°)	(D,D,D)
(45°, 90°)	(D,D,0)
(45°, 135°)	(D,D,-D)
(90°, 45°)	(0,D,D)
(90°, 90°)	(0,D,0)
(90°, 135°)	(0,D,-D)
(135°, 45°)	(-D,D,D)
(135°, 90°)	(-D,D,0)
(135°, 135°)	(-D,D,-D)
(-, 0°)	(0,0,D)

Tabel 3.2: Statistieken uit de co-occurrence matrix

Statistiek	Formule	Wat wordt gemeten?
Entropie	$-\sum_i^M \sum_j^N C(i, j) \cdot \log P[i, j]$	Maat voor de willekeurigheid van de grijswaardenverdeling.
Energie	$\sum_i^M \sum_j^N C^2(i, j)$	Maat voor het aantal terugkerende paren
Contrast	$\sum_i^M \sum_j^N (i - j)^2 C(i, j)$	Maat voor het lokale contrast in het beeld
Homogeniteit	$\sum_i^M \sum_j^N \frac{C(i, j)}{1 + i - j }$	Maat voor de lokale homogeniteit van het pixelpaar
Gemiddelde	$\frac{1}{2} \sum_i^M \sum_j^N (iC(i, j) + jC(i, j))$	Gemiddelde van alle grijswaardenniveaus in het beeld
Variantie	$\frac{1}{2} \sum_i^M \sum_j^N ((i - \mu)^2 C(i, j) + (j - \mu)^2 C(i, j))$	Geeft een idee van de spreiding van de grijswaarden
Correlatie	$\sum_i^M \sum_j^N \frac{(i - \mu)(j - \mu)C(i, j)}{\sigma_i^2 \sigma_j^2}$	Geeft de correlatie weer tussen de pixels in het pixelpaar
Maximale waarschijnlijkheid	$Max_{i,j}^{M,N} C(i, j)$	Het meest voorkomende resultaat in de pixelparen
Invers Verschil Moment	$\sum_i^M \sum_j^N \frac{C(i, j)}{ i - j ^k} \quad i \neq j$	Maat voor de gladheid van het beeld
Neiging tot clustering	$\sum_i^M \sum_j^N (i + j - 2\mu)^k C(i, j)$	Meet hoe sterk de pixels gegroepeerd zijn

Dit zijn een deel van de Haralick textuurkenmerken. Aan de hand van een analyse worden de meest significante kenmerkvectoren bekomen. De evaluatie van de meerwaarde van de verschillende kenmerkvectoren gebeurt op basis van kenmerkselectie [36]. Kenmerkselectie kan gebeuren volgens twee benaderingen:

- Subset Selectie
- Projectie: Principale Componenten Analyse (PCA)

Bij Subset Selectie wordt op basis van een aantal discriminante maten besloten welke kenmerken beter en welke minder goed onderscheid maken tussen verschillende klassen. Principal Component Analysis (PCA) is een methode om hoogdimensionale gegevens om te zetten in laagdimensionale gegevens via de bepaling van een kleinere deelruimte die goed bij de data aansluit [37]. Hier is gekozen voor de eerste benadering.

Omdat hier lever en tumoren data wordt gebruikt, kiest men logischerwijs voor twee klassen(1,t). Dit komt neer op twee kenmerkmatrices M_{F1}^t, M_{F2}^l die de verschillende kenmerken voor de verschillende datasets bevatten. In de analyse maken we geen gebruik van de combinatie van verschillende kenmerken (buiten een gemiddelde over de verschillende richtingen). Vervolgens definiëren we de drie Discriminant Measures (DMs) die zullen gebruikt worden:

1. Criterium van Fisher: $DM(t_i, l_i) = \frac{(\text{mean}(t_i) - \text{mean}(l_i))^2}{\text{var}(t_i) + \text{var}(l_i)}$
2. Euclidische Afstand: $DM(t_i, l_i) = \|t_i - l_i\|$
3. Relatieve Entropie: $DM(t_i, l_i) = t_i \log \frac{t_i}{l_i}$

Deze maten worden voor iedere dataset en iedere kenmerkvector berekend. Op basis hiervan kan men vervolgens een deelruimte selecteren die de meest significante kenmerkvectoren bundelt. Het besproken algoritme wordt uitgevoerd op de twee trainingspatiënten. Ideaal worden de textuurkenmerken voor elke voxel berekend. Aangezien het enkel mogelijk is de textuur te berekenen voor een gebied (of volume) en niet voor één enkele voxel, moet voor elke voxel een regio rond die voxel geselecteerd worden. In die regio dienen dan de verschillende textuurkenmerken berekend te worden en te worden opgeslagen in de voxel waarrond de regio zich bevindt. De textuur in elke voxel berekenen is echter enorm rekenintensief en tijdrovend. De oplossing is het maken van een preselectie op basis van de gemiddelde textuurwaarden voor een ganse metastase en voor de ganse lever (zonder metastasen). Een minimale vereiste is dat de gemiddeldes voor metastase en lever voldoende ver uit mekaar liggen opdat het textuurkenmerk algemeen toepasbaar zou zijn. Uitzondering zijn die textuurkenmerken met weinig verschillende waarde voor metastase en lever, maar met een duidelijk verschillende waarde voor de rand van de metastasen.

Een toepassing van deze textuurkenmerken gebeurt bij statistische pixelclassificatie 4.2. Hierbij worden de minder significante kenmerken buiten beschouwing gelaten.

Gemiddelde textuurkenmerken voor metastase- en leverweefsel

De gemiddelde textuurkenmerken worden afzonderlijk berekend voor het leverweefsel en het tumorweefsel binnen de verschillende datasets. In eerste instantie wordt de co-occurrence matrix berekend voor alle referentie-pixels waar de buur-pixels binnen hetzelfde weefsel vallen. Indien dit niet zo is, worden ze niet opgenomen in de co-occurrence matrix. De kenmerkvectoren zijn de tien geselecteerde Haralick textuurkenmerken in iedere richting (dertien mogelijkheden) en iedere afstand (vijf mogelijkheden). Hieraan wordt ook nog het gemiddelde over alle richtingen toegevoegd. Dit geeft een totaal van 700 kenmerkvectoren voor iedere dataset. Hiervan worden vervolgens de DMs berekend. Als vertrekpunt nemen we eerst het gewogen gemiddelde en de gewogen variantie over alle datasets. Met deze informatie kunnen de drie DMs berekend worden. De bekomen data wordt vervolgens verwerkt.

Resultaten

Een eerste zicht op de resultaten van de DMs geeft weer dat de richting waarin de co-occurrence matrix wordt berekend geen systematische invloed heeft. Daarom zal abstractie gemaakt worden van de richting. Bij overgangen tussen lever- en tumorweefsel kan de richting echter wel een rol spelen. Wanneer men een classificatie wil uitvoeren, neemt men daarom best het gemiddelde over alle richtingen of tenminste een gemiddelde in de horizontale en verticale richting van de snede. De afstand D heeft wel een invloed op bepaalde textuurkenmerken. Het verschil in contrast tussen lever- en tumorweefsel vergroot als de co-occurrencematrix over een grotere pixelafstand wordt berekend. Daarnaast nemen maten als correlatie en homogeniteit af voor grotere afstanden. Alle maten vertonen een toenemende standaarddeviatie naarmate de pixelafstand toeneemt. Dit komt overeen met een grotere spreiding binnen de co-occurrence matrix getoond in Figuur 3.5(c). Voor een visuele weergave hiervan verwijzen we naar de scatterplots in Bijlage B.2.

In Tabel 3.3 is een samenvatting weergegeven van de resultaten bekomen uit de DMs. Voor iedere DM is een rangschikking opgesteld. Daarnaast is ook een rangschikking toegevoegd op basis van een combinatie van de drie maten. Dit gebeurde op basis van een sortering volgens de gemiddelde plaats in de rangschikking over de drie maten heen. In de tabel wordt een onderscheid gemaakt tussen de textuurkenmerken die volgens de verschillende maten voldoende discriminerend zijn (vetgedrukt) en degenen die onvoldoende discriminerend zijn (grijs). Bij de combinatiekolom werd een kenmerk onvoldoende discriminerend beschouwd als het voor twee van de drie maten onvoldoende onderscheid maakt tussen lever- en tumorweefsel. De meeste textuurkenmerken scoren consistent doorheen de maten. Kenmerken als de variantie, het contrast en de correlatie scoren over het algemeen vrij goed. Entropie en gemiddelde scoren eerder matig. Homogeniteit, energie, neiging tot clustering en maximale waarschijnlijkheid scoren slecht wat betreft onderscheid tussen lever- en tumorweefsel. Tussen de verschillende DMs treedt toch enig variatie op. Dit heeft verschillende oorzaken. Voor het Fisher criterium veronderstelt men een normaalverdeling van de

data. Door de aanwezigheid van andere weefsels (bloedvaten) in de lever zal de data echter licht scheefverdeeld zijn. In dit geval is de Fisher maat niet helemaal correct. Daarnaast meet men met de euclidische afstand een verschil in absolute waarde, terwijl de relatieve entropie een verhouding meet. Hierdoor zal ook het resultaat verschillen.

Tabel 3.3: Samenvatting resultaten van discriminante maten

Top 10	Relatieve Entropie	Euclidische afstand	Fisher Maat	Combinatie
1.	Variantie	Variantie	Correlatie	Variantie
2.	Contrast	Contrast	Variantie	Correlatie
3.	Gemiddelde	Correlatie	Gemiddelde	Contrast
4.	Entropie	Max. Wsl.	Entropie	Gemiddelde
5.	Correlatie	Gemiddelde	Inv. Verschil Mom.	Entropie
6.	Energie	Energie	Contrast	Inv. Verschil Mom.
7.	Inv. Verschil Mom.	Entropie	Homogeniteit	Max. Wsl.
8.	Homogeniteit	Inv. Verschil Mom.	Max. Wsl.	Energie
9.	Max. Wsl.	Homogeniteit	Neig. Tot Clust.	Homogeniteit
10.	Neig. Tot Clust.	Neig. Tot Clust.	Energie	Neig. Tot Clust.

Op basis van de resultaten in Tabel 3.3 kunnen we besluiten dat de drie beste kenmerkvectoren de variantie, het contrast en de correlatie zijn. De pixelafstand van de co-occurrencematrix heeft ook een invloed op de resultaten. Grotere afstanden geven gemiddeld gezien een iets groter onderscheid, maar ook de variantie neemt toe. [35] geeft aan dat de afstand best in de grootte-orde van het textuurelement gekozen wordt. Voor metastasen is hierover geen informatie gevonden. Men kan dit best experimenteel nagaan. Hierbij moet echter wel rekening gehouden worden met de beperkte axiale resolutie (snededikte=7mm). Indien men 3D werkt, kan de pixelafstand niet te groot gekozen worden.

Textuuranalyse per voxel

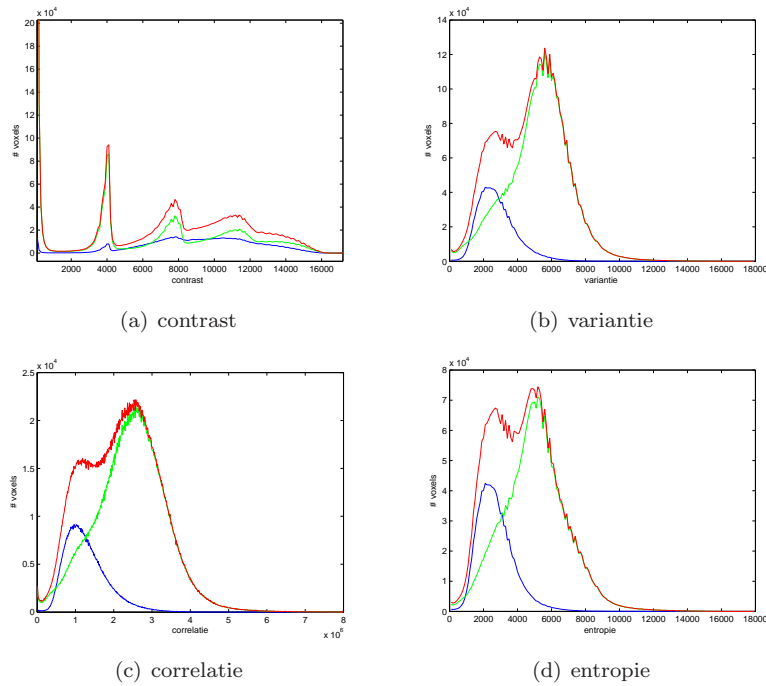
Tot hiertoe zijn de berekende resultaten gemiddelde waarden van textuurkenmerken over een geheel gebied. Dit biedt geen garanties bij het toepassen van eventuele lokale textuurkenmerken. Kenmerken die gemiddeld echter al slecht scoren, zullen lokaal niet beter doen. Een kenmerk waar op basis van een gemiddelde waarde geen onderscheid gemaakt kan worden tussen lever en metastase, is ook niet geschikt om voor onderscheid te zorgen op pixelniveau. Variantie, correlatie, contrast, gemiddelde en entropie zijn bijgevolg de meest interessante kenmerken om lokaal te bestuderen. Omdat het textuurkenmerk “gemiddeld” neerkomt op het meten van de gemiddelde intensiteit in een bepaalde regio laten we deze uit de verdere analyse. Buiten het verlies aan spatiale informatie is dit immers equivalent met het toepassen van een uitmiddelfilter op het beeld.

Aangezien het berekenen van de textuur voor elke voxel enorm rekenintensief en tijdrovend is, blijft de analyse beperkt voor kleine gebiedjes rond elke voxel, waardoor textuurberekeningen over grote afstanden praktisch niet haalbaar zijn, ook met het oog op later gebruik tijdens segmentatiemethodes. De beschouwde afstand is dan ook één voxel.

Om niet te veel informatie te verliezen tijdens het converteren van 4096 naar 128 grijswaarden, wordt eerst een grijswaardentransformatie uitgevoerd met als centrum 1107,5 en breedte 256 (zie 4.2.1). Uit de grijswaardenanalyse 3.2.1 blijkt immers dat bijna alle lever- en metastaseweefsel een grijswaarde heeft binnen dat interval. Rond elke voxel wordt een gebied geselecteerd van $3 \times 3 \times 1$ waarbinnen de textuurkenmerken berekend worden. Zo kan opnieuw een verdeling opgesteld worden voor alle metastasen en alle levers van de datasets samen.

Resultaten

Het resultaat is weergegeven in Figuur 3.6, waarbij de blauwe curve de verdeling van het textuurkenmerk in alle metastasen voorstelt, de groene curve de verdeling in de lever zonder metastasen en de rode curve de som van beide verdelingen. Het is meteen duidelijk dat het contrast lokaal geen discriminerend karakter heeft. Dit was ook al duidelijk bij de analyse van de gemiddelde textuurkenmerken, waar het gemiddelde contrast van metastase ongeveer gelijk is aan het gemiddelde contrast van de lever voor een afstand één. Contrast is waarschijnlijk wel discriminerend bij



Figuur 3.6: Verdeling van (a)contrast, (b)variantie, (c)correlatie, (d)entropie

grotere afstanden.

De overige drie kenmerken zijn op het eerste zicht wel onderscheidend. De eigenschappen van de verdeling (het gemiddelde μ en de standaarddeviatie σ) zijn weergegeven in Tabel 3.4.

Tabel 3.4: Eigenschappen van de verdeling van de textuurkenmerken

	$\mu_{contrast}$	$\sigma_{contrast}$	$\mu_{variantie}$	$\sigma_{variantie}$	$\mu_{correlatie}$	$\sigma_{correlatie}$	$\mu_{entropie}$	$\sigma_{entropie}$
lever	$8,26 \cdot 10^3$	$3,65 \cdot 10^3$	$5,27 \cdot 10^3$	$1,81 \cdot 10^3$	$2,45 \cdot 10^6$	$9,27 \cdot 10^5$	$5,05 \cdot 10^3$	$1,86 \cdot 10^3$
metastase	$9,36 \cdot 10^3$	$2,82 \cdot 10^3$	$2,81 \cdot 10^3$	$1,14 \cdot 10^3$	$1,28 \cdot 10^6$	$5,65 \cdot 10^5$	$2,73 \cdot 10^3$	$1,09 \cdot 10^3$

Deze eigenschappen worden gebruikt om een statistisch model op te stellen van de verdeling van de textuurkenmerken. Op basis van Figuur 3.6 lijkt het een goede benadering om een normale verdeling te veronderstellen. Deze is mathematisch weergegeven in formule 3.4. Aangezien ook hier de staarten te groot zijn voor een normale verdeling, is ook de logistische verdeling (formule 3.5) geen slechte benadering.

Om te bestuderen welke de maat is die het beste onderscheid maakt tussen lever en metastase wordt de Fisher maat berekend en uitgezet in Tabel 3.5.

Tabel 3.5: Textuurkenmerken volgens hun onderscheidend karakter op voxelniveau

	Fisher maat
variantie	1,3239
entropie	1,1668
correlatie	1,1435
contrast	0,0573
grijswaarde	4,4076

Ter referentie is ook de DM voor het intensiteitsbeeld toegevoegd. Geen enkele van de textuurkenmerken kan weefsel beter onderscheiden dan de oorspronkelijke grijswaarde. Een mogelijke

verklaring hiervoor is dat de lokale textuurkenmerken enkel voor een afstand één berekend werden. Er kan ook opgemerkt worden dat entropie relatief t.o.v. de andere maten beter scoort dan bij de analyse van de gemiddelde textuurkenmerken. De verklaring hiervoor is dat er slechts één patiënt in de analyse werd opgenomen.

Door de analyse per dataset uit te voeren, kan onderzocht worden of binnen een dataset er wel een beter onderscheid is tussen lever en metastase. Dit blijkt niet zo te zijn. De maximale waarde voor de Fisher maat is slecht 2,0648.

Besluiten

Het vertrekpunt om de textuur van tumor- en leverweefsel te onderzoeken, is de handmatige aflijning. De textuuranalyse gaat uit van het gebruik van een co-occurrencematrix die berekend wordt uit een beeld met minder grijswaarden. Deze matrix kan in verschillende richtingen en over verschillende pixelafstanden bepaald worden. Uit de co-occurrencematrix kunnen de Haralick textuurkenmerken berekend worden. Deze kenmerken kunnen verschillen in waarde naargelang de textuur van het weefsel. In eerste instantie gebeurt een analyse waarbij regio's van een bepaald type weefsel vergeleken worden om zo tot een preselectie te komen voor een analyse van de textuurkenmerken in de pixels. Deze analyse levert een top drie van meest discriminante kenmerken op: de variantie, het contrast en de correlatie. De richting waarin de co-occurrencematrix berekend wordt speelt weinig tot geen rol. Lever- en tumorweefsel zijn bijgevolg isotroop. Het verder onderzoeken van de textuurkenmerken levert echter weinig resultaat op. Met een voxelafstand één in een kernel van $3 \times 3 \times 1$, is geen enkel kenmerk meer onderscheidend dan de grijswaarden.

3.3 Geometrische eigenschappen

Naast de fotometrische eigenschappen, is het zinvol ook de geometrie van de metastasen te bestuderen opdat die mogelijk een meerwaarde kan bieden tijdens de (semi-)automatische segmentatie. We maken hierbij onderscheid tussen de grootte, de globale vorm en lokale geometrische eigenschappen.

3.3.1 De grootte van de metastase

Het eerste facet dat de geometrie van de metastase karakteriseert, is de grootte. Het volume is de beste manier om de grootte van de tumor te kwantificeren (2.5.2). Het gemiddelde tumorvolume hiervan is 146,5 ml, maar het volume vertoont grote variaties hetgeen de grote standaarddeviatie van 225,8 ml verklaart. Deze grote variaties zijn vooral te wijten aan een grote metastase in de lever van één patiënt en sluiten een modelgebaseerde segmentatie zonder scalering uit.

3.3.2 De globale vorm

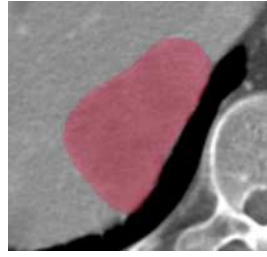
Een visuele inspectie van de metastasen leert dat deze soms willekeurige vormen hebben. Toch lijken de kleinere metastasen erg bolvormig. Het is ook zichtbaar dat de randen van de lever ervoor zorgen dat de groei van de metastase in de richting van de rand afgeremd wordt, waardoor de metastase een meer langwerpige vorm kent, zoals weergegeven op Figuur 3.7.

Een meer kwantitatieve analyse van de vorm kan gebeuren op basis van RECIST en WHO of op basis van een PCA van de handmatige segmentaties.

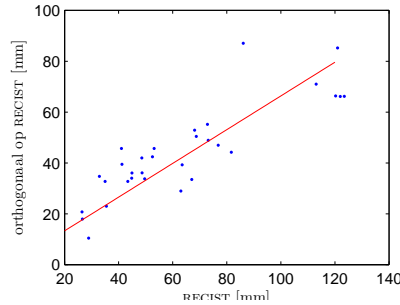
Het verband tussen RECIST en de orthogonale op de RECIST is zichtbaar in Figuur 3.8. Er is, zoals te verwachten, een sterke correlatie tussen beide maten. Het lineair verband tussen beide maten kan gevonden worden door het zoeken van de Kleinste Kwadraten Oplossing (KKO):

$$KKO = (x^T x)^{-1} x^T y$$

Dit levert de rechte $y = 0,6636x$ op. Punten boven deze rechte komen overeen met metastasen die iets ronder zijn dan gemiddeld, punten onder deze rechte zijn meer langwerpige tumoren.



Figuur 3.7: Voorbeeld van een langwerpige metastase t.g.v. de rand van de lever



Figuur 3.8: De loodrechte op RECIST in functie van RECIST

In 2.5.2 worden enkele studies aangehaald die aantonen dat het volume een betere maat is dan RECIST voor het karakteriseren van de metastase. Daarom wordt het volume vergeleken met RECIST en WHO. Dit doen we door de metastase te modelleren als een ellipsoïde, gelijkaardig als in [28]. We veronderstellen immers dat de metastasen er min of meer uitzien als een ellipsoïde. De wiskundige beschrijving van een ellipsoïde is:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

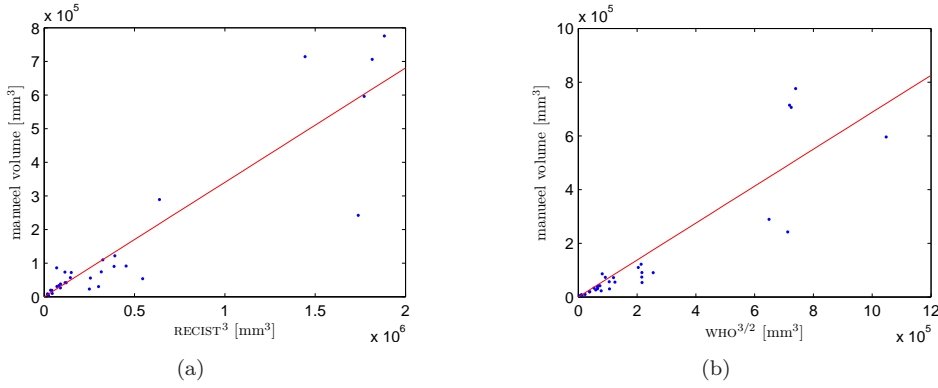
Hierin zijn a, b en c de halve lengtes van de verschillende assen. Het volume van een ellipsoïde is dan:

$$V_{ellipsoïde} = \frac{4}{3}\pi abc \tag{3.7}$$

Om het volume te bepalen aan de hand van RECIST stellen we $a = b = c = \frac{RECIST}{2}$, om het te bepalen aan de hand van de WHO stellen we $abc = (\sqrt{WHO})^3$. Er is dus een lineair verband tussen volume en RECIST tot de derde macht en tussen volume en $WHO^{3/2}$ te verwachten. Dit is uitgezet in Figuur 3.9(a) en Figuur 3.9(b).

De punten dicht bij de rechte gelegen in Figuur 3.9(a), komen overeen met metastasen waarvoor een verband tussen volume en RECIST tot de derde macht bij benadering geldig is. Voor deze metastasen kan RECIST dus als maat gebruikt worden voor het bepalen van grootteveranderingen, aangezien er een verband is tussen grootte van de metastase en de lengte van RECIST. De punten ver van de rechte gelegen in Figuur 3.9(a), komen overeen met metastasen waarvoor een verband tussen volume en RECIST tot de derde absoluut niet geldig is. Voor deze metastasen is RECIST dus een slechte maat. Het verband tussen $WHO^{3/2}$ en het manuele volume benadert iets beter een rechte, zoals ook aangetoond in [28], maar is ook onvoldoende. We concluderen hieruit dat RECIST en WHO onvoldoende maten zijn om de grootteveranderingen van de metastasen te kunnen kwantificeren. Een beter alternatief is deze kwantificatie uit te voeren a.d.h.v. volumemetingen. Hiervoor is echter de meting van het volume op verschillende tijdstippen noodzakelijk.

Een gelijkaardige, maar meer geautomatiseerde en dus nauwkeurigere analyse gebeurt met behulp van een PCA-analyse. Deze berekent de lengte van de hoofdassen van de handmatige



Figuur 3.9: Het manueel volume van de metastase in functie van (a) RECIST³ en (b) WHO^{3/2}

segmentaties van de metastasen. Ook is het mogelijk de richting van deze hoofdasen te berekenen. Met deze gegevens is het mogelijk om de bolvormigheid van de tumoren te bepalen. Deze is gegeven door [38]:

$$\Psi = \frac{\pi^{\frac{1}{3}}(6V_{obj})^{\frac{2}{3}}}{A_{obj}} \quad (3.8)$$

met V_{obj} het volume van het object en A_{obj} de oppervlakte van het objectoppervlak. Het volume van een ellipsoïde is gegeven in (3.7). Het oppervlak van een ellipsoïde is echter wiskundig niet te berekenen, vandaar dat we een benaderende formule gebruiken:

$$A_{obj} \approx 4\pi \left(\frac{a^p b^p + a^p c^p + b^p c^p}{3} \right)^{\frac{1}{p}}$$

Hierin is de relatieve fout t.o.v. het werkelijke oppervlak maximaal 1,178% voor $p=1,6$ (formule van David W. Cantrell [39]).

De gemiddelde bolvormigheid van het ellipsoïde-model is 95,12%, de minimale 87,16% en de maximale 99,17%. Ter referentie, de bolvormigheid van een bol is 100%, van een kubus 80,6% en van een octaëder 84,6%. De globale vorm van de metastasen (a.d.h.v. het ellipsoïde-model) is dus vrij bolvormig.

Aangezien de studie van de vorm tot doel heeft informatie te vergaren om een model op te stellen dat bruikbaar is voor segmentatie, is het belangrijk de bruikbaarheid van het ellipsoïde-model na te gaan. Dit doen we door het volume van het ellipsoïde-model te vergelijken met het werkelijke volume van de metastase. Deze vergelijking leert ons dat er een gemiddelde afwijking is van 30,8% met een maximum van $\pm 90\%$. De afwijking is in bijna alle gevallen positief (groter in het model). De afwijkingen zijn behoorlijk grote waarden, hetgeen een segmentatie die zuiver gebaseerd is op een geometrisch model, bijna onmogelijk maakt. Dit strookt ook met onze intuïtie. Een mogelijke toepassing van een ellipsoïde-model zou de localisatie van de metastasen kunnen zijn.

Oppervlakedetails

Bij het grondig bekijken van de tumoren valt op dat de metastasen een licht onregelmatig oppervlak hebben. Mathematisch kan dit nagegaan worden door de oppervlakte van het oppervlak van het ellipsoïde-model te vergelijken met het werkelijke oppervlak. Het oppervlak van het ellipsoïde-model is 10,7% groter in het ellipsoïde-model. De karakteristieke lengte is dan 5,2% groter ($\sqrt{1,107}=1,052$). Het volume gemiddeld 30% groter is. De karakteristieke lengte is dan 5,2% groter ($\sqrt{1,306}=1,093$). Dit wil zeggen dat het werkelijke oppervlak in verhouding een beetje groter is dan het werkelijke volume wat wijst op een licht onregelmatig oppervlak. Scherpe uitsteeksels komen weinig voor. Deze kennis kan gebruikt worden om een beperking op te leggen aan de lokale kromming van de afijning.

Hoofdstuk 4

Segmentatie van metastasen

In dit hoofdstuk volgen een aantal methodes die geïmplementeerd werden in *MeVisLab* [40] met als doel de metastasen in de lever te segmenteren. Binnen *MeVisLab* is het mogelijk gebruik te maken van een groot aantal algemene segmentatiemethodes uit de Insight Tool Kit (ITK) [41].

In de methodes die volgen is het belangrijk onderscheid te maken tussen methodes die de metastasen afzonderlijk kunnen segmenteren (beschreven in 4.3, 4.4, 4.5 en 4.6) en de methodes die dit niet kunnen (beschreven in 4.1 en 4.2).

Daarnaast onderscheiden de methodes zich ook in de mate waarin gebruikersinteractie nodig is. Een laatste onderscheid dat van belang is, is de nood aan een afgelijnde lever.

4.1 Interval thresholding

De eenvoudigste manier van segmenteren is door gebruik te maken van thresholds. Hiermee gebeurt een selectie van een bepaald interval van grijswaarden. Dit noemen we interval thresholding.

Uit onderzoek blijkt dat de grijswaarden van de metastasen zich in een ander, echter wel overlappend, interval bevinden dan de grijswaarden van de rest van de lever en dat de gemiddelde grijswaarde zich in een smal interval bevindt (zie sectie 3.2.1). De interval thresholding vindt plaats op een gesegmenteerde lever met de originele grijswaarden. Deze methode is dus automatisch, gegeven een gesegmenteerde lever.

Na de interval thresholding operatie gebeuren er nog operaties om de losse voxels weg te eroderen en om gaten te vullen met beschikbare morfologische operatoren. Een groot nadeel van deze methode is dat het de verschillende metastasen niet afzonderlijk segmenteert en het op die manier niet mogelijk is het volume van één metastase op te volgen. Interval thresholding kan wel een globaal beeld geven van de toe- of afname van tumorweefsel in de lever, hetgeen ook nuttig kan zijn voor het opvolgen van patiënten. Deze automatische methode moet toegepast worden op een afgelijnde lever, aangezien er weefsel buiten de lever aanwezig is met ongeveer dezelfde grijswaarden als de metastasen.

4.2 Statistische pixelclassificatie

Pixelclassificatie kent aan elke pixel (voxel) een kans toe dat hij tot een bepaalde klasse behoort. We gebruiken twee klassen: de klasse van pixels die tot de lever behoren en de klasse van de pixels die tot een metastase behoren. Dit vraagt een gesegmenteerde lever met de originele grijswaarden. Om een kans toe te kunnen kennen aan een pixel is een statistisch model nodig. De parameters van het model kunnen geleerd worden van een testdataset of kunnen geschat worden uit het beeld zelf [9].

4.2.1 Gesuperviseerde pixelclassificatie

Het onderzoek van de fotometrische eigenschappen (3.2) van lever en metastasen levert voldoende informatie op om de parameters van een statistisch model op te stellen voor de lever en de metastasen. Het gaat dus om statistische pixelclassificatie met *supervised learning*. Een eerste algoritme gebruikt enkel intensiteitswaarden om aan pixels een kans toe te kennen, een tweede algoritme gebruikt daarnaast ook nog textuurkenmerken. Aangezien het berekenen van textuurkenmerken redelijk tijdsintensief is, kan het eerste algoritme het beste alternatief zijn als rekentijd belangrijk is.

Classificatie enkel op basis van intensiteiten

Basisalgoritme

Zoals aangetoond in 3.2.1 is het aanvaardbaar om een Gaussiaanse verdeling te veronderstellen voor de grijswaarden van zowel de lever als de metastasen:

$$p(I_k|c_j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(I_k - \mu)^2}{2\sigma^2}\right) \quad (4.1)$$

met I_k de grijswaarde van voxel k , μ en σ de gemiddelde grijswaarde en de standaarddeviatie van de klasse. Als alternatief voor de verdeling van de grijswaarden in de metastase is een logistische verdeling mogelijk:

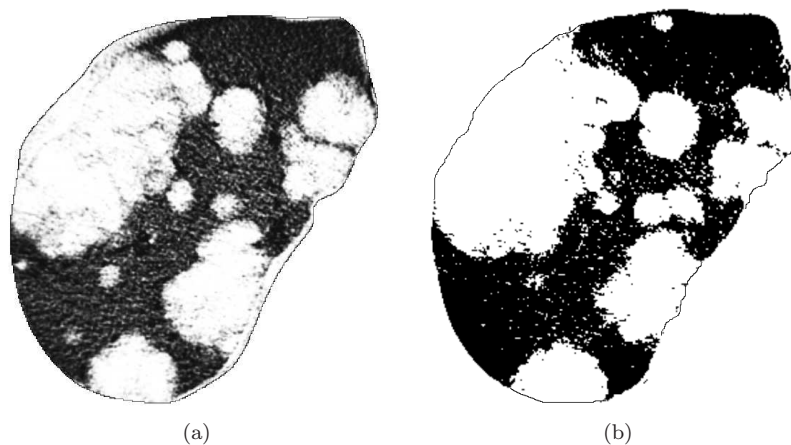
$$p(I_k|c_j) = \frac{\exp\left(\frac{I_k - \mu}{s}\right)}{s(1 + \exp\left(\frac{I_k - \mu}{s}\right))^2} \quad (4.2)$$

met $s = \frac{\sigma\sqrt{3}}{\pi}$. Door toepassing van de regel van Bayes kan een kans toegekend worden aan een voxel dat hij tot een bepaalde klasse behoort:

$$p(c_j|I_k) = \frac{p(I_k|c_j)p(c_j)}{\sum_{i=1}^M p(I_k|c_i)p(c_i)} \quad (4.3)$$

Hierin is $p(c_j)$ de a priori waarschijnlijkheid dat voxel k tot klasse c_j behoort. Deze nemen we $p(c_{\text{lever}}) = p(c_{\text{metastase}}) = \frac{1}{2}$ [9], aangezien op voorhand niet gekend is hoeveel metastaseweefsel er zich in de lever bevindt en dit te fel varieert van dataset tot dataset om aan te leren.

Het is mogelijk om de kans dat een voxel tot een bepaalde klasse behoort weer te geven in een beeld (fuzzy classification), maar om het beeld te segmenteren moet er voor elke voxel een beslissing genomen worden tot welke klasse hij behoort (binary classification). Figuur 4.1(a) illustreert fuzzy classification, Figuur 4.1(b) de binaire classificatie.



Figuur 4.1: Pixelclassificatie: fuzzy (a) en binair (b)

Voor de verdeling van de metastasen gebruikt het algoritme een logistische distributie, voor de verdeling van de lever een normaalverdeling. Tabel 4.1 geeft de nodige parameters van de distributies.

Tabel 4.1: μ en σ van verdeling van lever- en metastaseweefsel

	μ	σ
lever	1160,24	26,68
metastase	1083,58	24,93

Het gebruik van een normale verdeling i.p.v. een logistische verdeling voor metastaseweefsel geeft nagenoeg hetzelfde resultaat.

Problemen met en verbeteringen aan het algoritme

Het grootste probleem is dat er binnen de afijning van de lever meer dan twee types weefsel aanwezig zijn. Naast het leverweefsel (hepatocieten) en metastasen zijn er ook nog bloedvaten, holtes, ... Hierbij zijn vooral die weefsels problematisch die door het algoritme als metastase aangeduid worden, maar het eigenlijk niet zijn. Dit zijn vooral holtes die een veel lagere intensiteit hebben dan de metastasen.

Om dit oplossen, voegen we een beperking toe. Pixels met intensiteiten beneden een bepaalde drempelwaarde hebben steeds een kans nul om tot de klasse van de metastasen te behoren. Er is eveneens een bovengrens toegevoegd. De onder- en bovengrens zijn resp. 980 en 1235. Dit lost het probleem deels op. Omwille van het partiel volume effect zijn er bij de overgang tussen de holtes en leverweefsel voxels met een intensiteit die eerder karakteristiek is voor metastaseweefsel.

Classificatie op basis van intensiteiten en textuurkenmerken

Toevoegen van textuurkenmerken in het statistische model maakt het vorige algoritme betrouwbaarder. Deel 3.2.2 bespreekt de textuur van de lever en de metastasen. Er werd onderzocht welke onderscheidmakende textuurkenmerken zijn en welke niet. Voor de onderscheidmakende kenmerken is de analyse gelijkaardig aan deze van de intensiteiten.

Basialgoritme

Om de textuur te berekenen mogen er maar 256 verschillende grijswaarden in het beeld zijn, omwille van snelheid en stabiliteit van het algoritme dat de textuur berekent (3.2.2). Vandaar dat eerst een grijswaardentransformatie wordt toegepast: een level-windowing operatie met als centrum (level) 1107,5 en breedte (window) 256 grijswaarden. Door dit te doen gebeurt een selectie van het grijswaardeninterval van de lever en de metastasen.

Op het getransformeerde beeld berekent een textuurfilter verschillende textuurkenmerken. Voor details van de implementatie: zie 3.2.2. Uit 3.2.2 blijkt dat correlatie en variantie de meest onderscheid makende kenmerken zijn. Het statistische model gebruikt een normale verdeling voor elk van deze kenmerken. Ook de intensiteiten worden mee opgenomen in het model. Opnieuw gebruiken we een normaalverdeling.

Om de uiteindelijke kans te bepalen dat een bepaalde voxel tot de klasse van de metastase behoort combineren we:

$$\begin{aligned}
 p(c_{met}|I_k) &= \frac{p(I_k|c_{met})p(c_{met})}{\sum_{i=1}^M p(I_k|c_i)p(c_i)} \\
 p(c_{met}|I_{cor,k}) &= \frac{p(I_{cor,k}|c_{met})p(c_{met})}{\sum_{i=1}^M p(I_{cor,k}|c_i)p(c_i)} \\
 p(c_{met}|I_{var,k}) &= \frac{p(I_{var,k}|c_{met})p(c_{met})}{\sum_{i=1}^M p(I_{var,k}|c_i)p(c_i)}
 \end{aligned}$$

met $I_{cor,k}$ en $I_{var,k}$ respectievelijk de correlatie en de variantie van voxel k .

4.2.2 Niet-gesuperviseerd: Expectation Maximization Algoritme

Het hierboven beschreven pixelclassificatie-algoritme vereist een trainingsfase om tot de parameters van het statistische model te komen. Wanneer deze parameters verschillen van dataset tot dataset kan het nodig zijn om de parameters te schatten uit het beeld zelf. Dit wordt *unsupervised learning* genoemd. Een mogelijke benadering hiervoor is het EM-algoritme. Het voordeel van niet-gesuperviseerde pixelclassificatie is dat per tumor het statistisch model geoptimaliseerd wordt. Uit 3.2.1 blijkt immers dat een relatieve threshold interessanter is dan een absolute threshold.

Classificatie enkel op basis van intensiteiten

Basisalgoritme

Het algoritme gebruikt hetzelfde statistische model als bij *supervised learning*, gegeven door vergelijking 4.1. De parameters $\{\mu_j, \sigma_j\}$ worden bij elke iteratie aangepast.

$$\begin{aligned}\mu_j^{nieuw} &= \frac{\sum_k p(c_j|I_k)^{oud} I_k}{\sum_k p(c_j|I_k)^{oud}} \\ \sigma_j^{nieuw} &= \frac{\sum_k p(c_j|I_k)^{oud} (I_k - \mu_j^{nieuw})^2}{\sum_k p(c_j|I_k)^{oud}}\end{aligned}$$

Ook wordt er bij elke iteratie een nieuwe a priori waarschijnlijkheid berekend voor de volgende iteratie:

$$p(c_j)^{nieuw} = \frac{1}{N} \sum_{k=1}^N p(c_j|I_k)^{oud}$$

met N het aantal pixels. Deze a priori waarschijnlijkheid convergeert ook naar de fractie dat een bepaalde klasse voorkomt.

Problemen met en verbeteringen aan het algoritme

Ook dit algoritme geeft problemen bij holtes tussen de leverlobben, aangezien deze een lage intensiteit hebben. Vandaar dat dezelfde beperking wordt ingevoerd als bij *supervised learning*. Om een oneindige lus te vermijden wordt een maximum aantal iteraties ingevoerd. Een mogelijke verbetering zou zijn om hiernaast ook nog een ondergrens vast te leggen voor het verschil tussen twee opeenvolgende stappen van de totale waarschijnlijkheid van het beeld. Wanneer dit verschil de ondergrens (bijv. 1%) bereikt heeft, zal er geen nieuwe iteratiestap meer volgen.

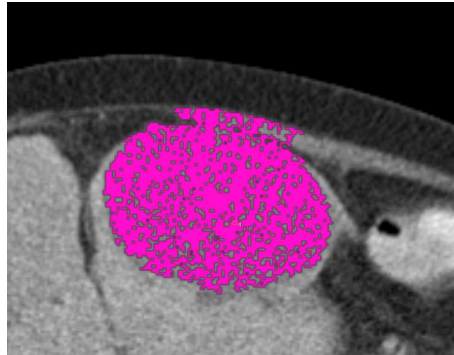
4.3 Region Growing

Voor alle beschreven Region Growing-methodes is een interactie van de gebruiker nodig wat wil zeggen dat al deze methodes semi-automatisch zijn. De motivatie voor de implementatie van Region Growing is een handboek van Suri, Wilson en Laxminarayan [42], pp 196-197. Hierin wordt vermeld dat op basis van initiële resultaten van een vergelijkend onderzoek tussen segmentatiemethodes uit ITK voor de afijning van tumoren in de pancreas, Region Growing de beste segmentatie geeft. We maken voor al deze methodes gebruik van een beschikbaar basisalgoritme voor region growing.

4.3.1 Region Growing met één seed per snede in een bepaalde metastase

Deze methode is de klassieke 2D-Region Growing-methode. De gebruiker dient een seed te plaatsen op elke snede waar er zich een deel van de metastase bevindt. De gemiddelde grijswaarde van alle punten waar een seed aanwezig is, is bepalend voor het interval waarbinnen de grijswaarden moeten liggen om tot het gesegmenteerde gebied te mogen behoren. De methode deelt het beeld op in verschillende gebieden door aangrenzende pixels (voxels) met grijswaarden binnen het interval

samen te nemen. Deze methode vergt dus nog vrij veel interactie van de gebruiker. Het resultaat hiervan is zichtbaar in Figuur 4.2.



Figuur 4.2: Een metastase na 2D region growing per snede

De postprocessing zorgt voor de opvulling van kleine gaatjes en voor het verwijderen van staarten. In de meeste gevallen zijn deze verkeerd gesegmenteerd weefsel. Het voordeel van deze methode is dat ze eenvoudig en snel is. Het grote nadeel is dat uitgroei bijna onvermijdelijk is, zeker wanneer het verschil in grijswaarde tussen metastase en omliggend weefsel klein is.

4.3.2 Region Growing met twee seeds per metastase

Een tweede methode maakt gebruik van slechts twee seeds per metastase. Deze dienen door de gebruiker geplaatst te worden op de bovenste en de onderste snede waar de te segmenteren metastase nog zichtbaar is. Het algoritme bekomt de tussenliggende seeds d.m.v. lineaire interpolatie. Zo bekomt het programma één seed per snede. Het vervolg is volledig gelijkaardig aan de vorige paragraaf 4.3.1. Deze methode vergt minder gebruikersinteractie, maar blijkt niet veel slechter te presteren dan Region Growing met één seed per snede. Bij zeer willekeurige vormen is het echter mogelijk dat de geïnterpoleerde seed minder goed in het midden van of zelfs buiten de snede ligt.

4.3.3 Region Growing met één seed per metastase

Deze methode gebruikt een 3D-region growing, in tegenstelling tot de twee vorige methodes. Hierdoor moet de gebruiker maar één seed plaatsen. Het grote gevaar bij 3D-region growing is dat het gesegmenteerde gebied snel uit de tumor groeit doordat er een grote kans bestaat dat het gebied ergens in de ruimte raakt aan een ander gebied met een gelijkaardige intensiteit. Het vastleggen van het interval waarbinnen de grijswaarden van de voxels moeten liggen, gebeurt opnieuw a.d.h.v. de intensiteit van de seed: $[I_{seed} - interval, I_{seed} + interval]$, waarbij *interval* een parameter is uitgedrukt in % van de maximale in het beeld aanwezige grijswaarde.

Het toevoegen van de beperking dat het gebied slechts mag groeien naar de volgende voxel als alle omliggende voxels binnen het toegelaten grijswaardeninterval liggen vermijdt zware uitgroei tijdens de segmentatie.

4.3.4 Region Growing op basis van RECIST en de loodrechte hierop

Het belangrijkste nadeel van de drie vorige methodes is uitgroei van de segmentatie uit de metastase. Deze is met de vorige methodes onvermijdelijk indien verschillende metastasen aan mekaar grenzen. Een volgende methode is gebaseerd op het handmatig trekken van twee lijnen: RECIST en de loodrechte hierop. Het tekenen van deze twee lijnen is iets wat in de dagelijkse klinische praktijk standaard gebruikt wordt (zie 2.5.1).

De methode werkt in drie belangrijke stappen:

1. **Statistische stap:** In deze stap verzamelt het algoritme statistische informatie over de grijswaardenverdeling in die welbepaalde metastase. Dit gebeurt door een masker te creëren waar we van kunnen veronderstellen op basis van de getekende lijnen dat dit masker met grote waarschijnlijkheid in de metastase ligt. Over de vorm van het masker zijn er twee mogelijkheden geïmplementeerd:

- Een schijfvormig masker, waarvan de straal de minimale straal is van een eindpunt van een lijn tot het snijpunt van beide lijnen en waarvan het centrum samenvalt met het snijpunt van de twee lijnen.
- Een vierhoek door de vier eindpunten van de twee lijnen. Het tekenen van de vierhoek gebeurt door het opvullen van twee driehoeken gebruik makend van een scanline-algoritme. Details van de implementatie zijn terug te vinden in Bijlage C.

Vervolgens bepaalt het algoritme de gemiddelde grijswaarde van het masker en de standaarddeviatie van de grijswaardenverdeling in het masker die de basis vormen van een normaal verdeeld statistisch model.

2. **Region Growing stap:** Dit gebeurt met een standaard 2D region growing-algoritme, met als zaadpunt het centrum van de twee lijnen en met als interval:

$$\begin{cases} thr_{upper} = \bar{I}_k + \alpha \cdot s \\ thr_{lower} = \bar{I}_k - \beta \cdot s \end{cases}$$

met \bar{I}_k de gemiddelde grijswaarde in het masker en s de standaarddeviatie. Door een normaalverdeling te veronderstellen voor de grijswaardenverdeling (zie 3.2.1) wil dit dus zeggen dat tussen de twee drempelwaarden zich 99% van grijswaarden bevinden voor $\alpha = \beta = 3$.

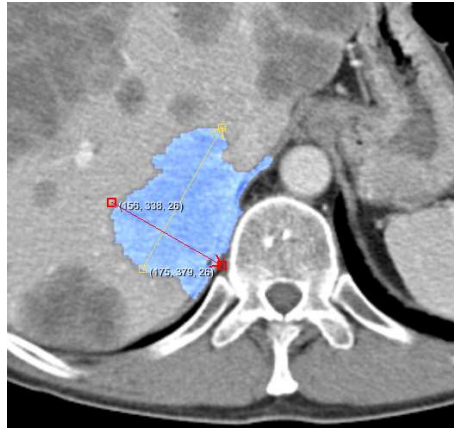
3. **Propagatie:** Er wordt een nieuw zaadpunt gevonden door het zwaartepunt te bepalen voor de gesegmenteerde snede. De beste resultaten worden gevonden door dit zaadpunt naar boven en beneden te schuiven zolang die waarde binnen het opgelegde grijswaardeninterval ligt. Het gebruikte interval blijft hetzelfde voor alle snedes.

Om te vermijden dat de segmentatie uit de metastase groeit zijn er beperkingen gedefinieerd in grootte. Deze beperkingen kunnen opnieuw twee vormen aannemen:

- Cirkelvormig: de grootste straal van een eindpunt van een lijn tot het snijpunt van beide lijnen. Aangezien RECIST de grootste afmeting van de tumor is in 2D is deze cirkel steeds groter dan de metastase.
- Trapezoïdaal: de zijden van het parallellogram worden gevormd door evenwijdigen aan RECIST en de lijn (ongeveer) loodrecht daarop. Ook deze bijna rechthoek is steeds groter dan de tumor. Het tekenen van dit parallellogram gebeurt eveneens met het in Bijlage C besproken algoritme.

Een voorbeeld van segmentatie met de methode die gebruikt maakt van vierhoekig masker en een trapezoïdale beperking is gegeven in Figuur 4.3. De statistieken van de grijswaarden in het vierhoekig masker bepalen de boven- en ondergrens voor segmentatie. In dit voorbeeld is de benedengrens 1024,19 en de bovengrens 1117,32. De beperking in de vorm van een parallellogram belet dat er uitgroei plaatsvindt. Dit is zichtbaar aan de rechterrands waar er bij gewone region growing uitgroei zou zijn.

De voordelen van deze methode zijn dat er geen uitgroei mogelijk is binnen één snede door de harde beperking en dat het statistisch model het optimale interval specifiek voor iedere metastase kiest. Bovendien vraagt deze methode relatief weinig gebruikersinteractie. De gebruiker dient enkel RECIST en de loodrechte hierop te tekenen. Het grote pijnpunt ligt in de propagatie. Als door het verschuiven van zaadpunt het nieuwe zaadpunt per toeval buiten het opgelegde interval terecht komt, zal het algoritme soms ten onrechte de propagatie stoppen. Aan de andere kant is het ook mogelijk dat de propagatie te ver doorgaat omdat het gepropageerde zaadpunt in een andere structuur is terecht gekomen met gelijkaardige grijswaarden, bijv. andere metastasen of bepaalds weefsel buiten de lever.



Figuur 4.3: Een metastase na 2D region growing per snede

4.3.5 Region Growing gegeven een fuzzy classificatiebeeld

Een alternatief voor de methodes beschreven in 4.3.1, 4.3.2 en 4.3.3, zijn methodes die vertrekken van een beeld bekomen door fuzzy pixelclassificatie. De ingang van het 2D region growing algoritme is de uitgang van het fuzzy pixelclassificatie-algoritme, beschreven in 4.2 waarbij de kans dat een pixel tot de klasse van de metastasen behoort, wordt uitgezet in een beeld. Op deze manier is het mogelijk om de verschillende metastasen afzonderlijk te segmenteren. Een alternatief is om etikettering (labeling) toe te passen op een beeld bekomen door binaire pixelclassificatie.

4.4 Spiral Scanning Technique

De algemene beschrijving van de Spiral Scanning Technique is terug te vinden in Sectie 2.4.3, gebaseerd op [24]. De manier waarop wij dit geïmplementeerd hebben wordt hieronder in detail besproken. De implementatie gebeurt in drie grote stappen. Eerst vindt een transformatie van de 3D- naar 2D-beeldruimte plaats m.b.v. een spiraalscan. Daarna volgt de segmentatie in 2D d.m.v. een berekening van het minimaal kostpad. De derde en laatste stap is de terugtransformatie naar de 3D-beeldruimte.

4.4.1 Transformatie m.b.v. Spiral Scan

Een 3D beeld wordt spiraalvormig gescand en getransformeerd naar 2D met in het centrum van de spiraal een handmatig aangeduide seed. De wiskundige beschrijving van de spiraal op een boloppervlak is een parametrisatie, gegeven door:

$$x = x_C + \rho \sin(t) \cos(2Nt) \quad (4.4)$$

$$y = y_C + \rho \sin(t) \sin(2Nt) \quad (4.5)$$

$$z = z_C + \rho \cos(t) \quad (4.6)$$

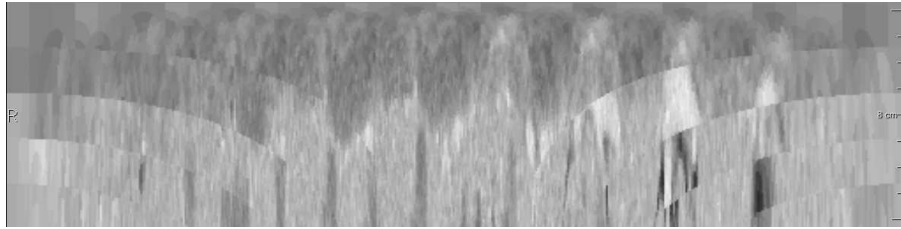
Hierin is $t \in [0, \pi]$ de variabele parameter, N het aantal omwentelingen dat de spiraal maakt en ρ de straal van de bol. Het aantal gescande stralen is belangrijk om een voldoende resolutie te krijgen in het getransformeerde beeld. We kiezen $2N$ punten op de spiraal per omwenteling, met $N = 16$. Zo zijn er dus in totaal $2N^2 = 512$ te scannen lijnen. Dit is dan ook het aantal pixels van het getransformeerde beeld in de horizontale richting k_1 . Op één scanlijn kiezen we $M = 128$ monsters, dit is de verticale resolutie na transformatie. Het verband tussen de parameter t en k_1 ,

en het verband tussen ρ en k_2 , de verticale richting, is gegeven door:

$$t = \frac{\pi k_1}{2N^2} \quad (4.7)$$

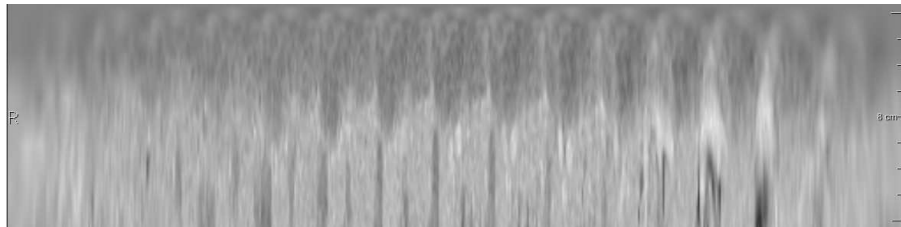
$$\rho = \frac{\rho_{max} k_2}{M} \quad (4.8)$$

Hierin is M het aantal samples per scanlijn en ρ_{max} de handmatig ingestelde maximale straal waarbinnen de metastase dient te liggen. De ingangen van het transformatie-algoritme zijn het inputbeeld, het middelpunt van de tumor en een straal, waarbinnen de tumor zeker moet liggen. De opvulling van het outputbeeld gebeurt kolom per kolom, waarbij voor elke voxel de door afronding overeenkomstige voxel van het inputbeeld opgevraagd wordt. De berekening van het overeenkomstige punt in het inputbeeld, van het punt (k_1, k_2) gebeurt door combinatie van vergelijking (F.1) t.e.m. (4.8). Het resultaat van de transformatie is te zien in Figuur 4.4.



Figuur 4.4: Transformatie volgens een spiraal scan

Omdat het ingangsbeeld bestaat uit 5 – 7 mm snedes, is het resultaat van mindere kwaliteit, vooral aan de zijkanten van het beeld. De lijn die gescand wordt snijdt vijf verschillende z-snedes, die duidelijk te onderscheiden zijn. Dit zorgt mogelijk voor onnauwkeurigheden in de berekening van het minimale kostpad, doordat er een voorkeur voor het pad zal zijn volgens het scheidingsoppervlak tussen de z-snedes. Door de waarde van een punt te bepalen op basis van interpolatie van omliggende voxels, verdwijnen de overgangen te wijten aan de lage z-resolutie (Figuur 4.5). Door de lage z-resolutie is er echter een grote uitsmering aan de randen.



Figuur 4.5: Transformatie volgens een spiraal scan

4.4.2 Segmentatie m.b.v. Minimal Cost Path (MCP)

Het is de bedoeling het 2D beeld te segmenteren door in het beeld een kromme te trekken van links naar rechts. Dit gebeurt door het MCP te bepalen door dynamisch programmeren. De segmentatie bestaat uit twee grote stappen, die voorafgegaan worden door een preprocessing-operatie van het beeld, namelijk het berekenen van de afgeleide in de verticale richting. Dit is de richting van een toenemende straal.

1. Berekenen van kostmatrix

De lokale kost is gedefinieerd als de gewogen som van twee componenten: de interne en de externe kost. De interne kost meet de smoothness en kan geschreven worden als:

$$E_C^{int}(i) = w_{int} \cdot |y_i - y_{i-1}| \quad (4.9)$$

Hoe groter de stap in de verticale richting is om in een punt in de volgende kolom te belanden, des te groter de interne kost. De gewichtsfactor w_{int} weegt de interne kost. Het vergroten van deze gewichtsfactor levert een zachter verlopend pad op.

De externe kost is de kost gerelateerd aan de sterkte van de gradiënt. Hoe sterker de gradiënt in een bepaald punt des te lager de kost.

$$E_C^{ext}(i) = w_{ext}(|\nabla I|_{max} - |\nabla I(x_i, y_i)|) \quad (4.10)$$

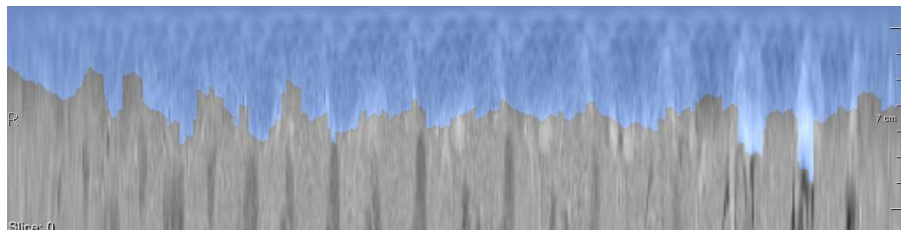
De eerste kolom van de kostmatrix bevat enkel een externe kost. Voor de verdere kolommen moet de minimale kost in elk punt gezocht worden, dit is het minimum van de verzameling die de sommen van de interne en externe kost vertrekkende van elke punt in de vorige kolom bevat. Op deze manier gebeurt de opbouw van de hele kostmatrix.

2. Segmentatie via scheiding d.m.v. het minimale kosttraject

De volgende stap na het berekenen van de kostmatrix is het bepalen van het MCP. Dit gebeurt door het minimum te nemen in de laatste kolom van de kostmatrix en vervolgens door terugrekenen het minimale pad te bepalen. Dit kan gemakkelijk door tijdens het berekenen van de kostmatrix in een andere matrix het pad bij te houden om in een zeker punt te geraken. Deze padmatrix bevat de verticale positie van de vorige kolom horende bij de minimale kost voor een zeker punt.

Alle pixels met een y-coördinaat kleiner dan de y-coördinaat van het minimale kostpad voor een bepaalde x-waarde, krijgen de waarde 1, de rest de waarde 0.

Het resultaat van de segmentatie m.b.v. MCP-berekening is te zien op Figuur 4.6.



Figuur 4.6: Segmentatie na bepaling MCP

4.4.3 Terugtransformatie

Na segmentatie van het beeld in 2D, dient het beeld teruggetransformeerd te worden naar 3D. Voor alle punten in het 3D beeld gelegen binnen de maximale straal ρ_{max} gebeurt er een conversie naar polaire coördinaten (φ, θ, ρ) . φ , de elevatiehoek, bepaalt in welk deel van het 2D beeld gezocht moet worden. Als φ gekend is, weet men ook dat het dichtstbijzijnde punt op de spiraal met straal ρ gelegen is tussen $\varphi - \frac{\pi}{N}$ en $\varphi + \frac{\pi}{N}$. A.d.h.v. vergelijking (4.7) en met $t = \varphi$, ligt het interval vast waarbinnen de juiste waarde van k_1 moet liggen. In dat interval, zoekt een algoritme naar de dichtstbijzijnde sample op de spiraal. Door opnieuw gebruik te maken van (4.7) is het berekenen van t mogelijk. Deze waarde van t wordt ingevuld in (F.1), (F.3) en (F.3) om op die manier een punt $P(x, y, z)$ op de spiraal te vinden. Tot slot berekent het algoritme de euclidische afstand tussen dit punt P en het te bepalen punt in het uitgangsbild. Uiteindelijk krijgt het te berekenen punt de waarde van het dichtstbijzijnde punt op de spiraal. Dit is het nearest neighbor principe.

4.4.4 Problemen bij toepassing van het algoritme

Omliggende structuren rond de tumor

Door structuren die rond de tumor liggen en binnen de maximale straal ρ_{max} kunnen gradiënten ontstaan die groter zijn dan de gradiënt t.g.v. de rand van de tumor. Hierdoor zal het MCP naar de structuur buiten de tumor toetrokken worden.

Een eenvoudige, maar zeker geen volledige oplossing is een extra kost toekennen aan de buitenkant van de spiraal, waar de kans op dergelijke structuren groter is. Dit doen we door daar de kost te vermenigvuldigen met een factor 1,25. Wanneer een structuur met een grote gradiënt aanwezig is, zal de externe kost overal in het beeld groter zijn ($E_C^{ext}(i) = |\nabla I|_{max} - |\nabla I(x_i, y_i)|$). Een vermenigvuldiging met een factor zorgt dan voor een grotere extra kost aan de buitenkant dan in de situatie wanneer er zich geen structuur bevindt (en de externe kost in het beeld dus lager is).

Verminderde nauwkeurigheid bij toenemende straal

Bij toenemende straal zullen opeenvolgende samples op de spiraal verder uit elkaar gelegen zijn. Dit heeft twee belangrijke gevolgen:

- Diep in de metastase is er een oversampling, terwijl aan de buitenzijde relatief weinig monsters zijn. Echter aan de buitenzijde is het belangrijk voldoende samples te hebben opdat de segmentatie nauwkeurig zou zijn.
- De Spiral Scanning Technique zal beter resultaten opleveren bij kleine metastasen dan bij grotere indien het totaal aantal samples hetzelfde blijft. Het opvoeren van het aantal samples gaat gepaard met een significante toename van de rekencomplexiteit.

Een oplossing hiervoor zou het zijn om naast een maximale straal ook een minimale straal aan te duiden. De gebruiker moet er dan voor zorgen dat de scheiding tussen metastase en omliggend weefsel steeds tussen beide stralen ligt.

Problemen bij slechte plaatsing van de seed

De werking van het algoritme is optimaal wanneer de seed perfect in het midden van de metastase geplaatst wordt. In praktijk is het echter niet altijd even gemakkelijk om dit midden aan te duiden.

Onduidelijke rand van de metastasen

Waar de rand van de metastasen onduidelijk is, is er ook nauwelijks een verhoging van de gradiënt vast te stellen. Hierop toepassen van het MCP-algoritme geeft een onbetrouwbare scheiding.

4.4.5 Verbeteringen van het basisalgoritme

Equidistante scanlijnen

Het basisalgoritme maakt gebruik van eenzelfde aantal scanlijnen per omwenteling. Dit heeft tot gevolg dat aan de polen, waar de lengte van de spiraal over één omwenteling kleiner is dan op de evenaar, de scanlijnen dichter bij elkaar liggen, zoals zichtbaar op Figuur 4.7.

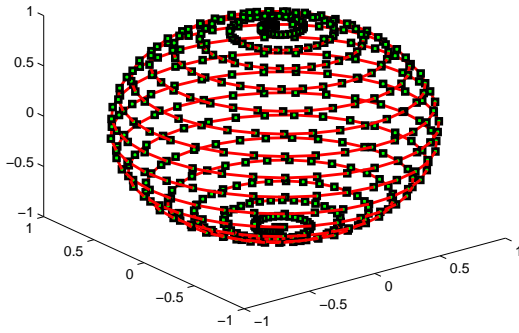
Aangezien bij equidistante bemonstering de spiraal dezelfde is, is ook de parametrisatie dezelfde. Echter het verband tussen t en k_1 is wel gewijzigd:

$$t = \frac{\pi}{2} + \arcsin\left(\frac{2k_1}{2N^2} - 1\right)$$

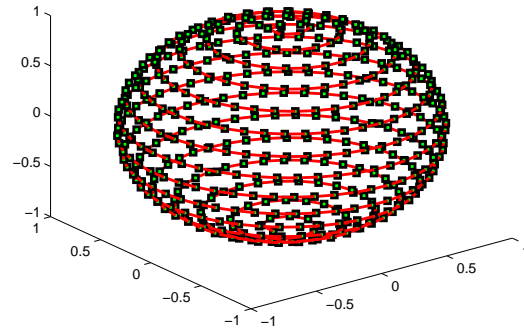
Dit zorgt ervoor dat de transformatie m.b.v. Spiral Scan en de terugtransformatie lichtjes wijzigen.

Aanpassing berekening Minimal Cost Path

In het basisalgoritme gebeurt de berekening van het minimaal kostpad op basis van twee kosttermen, één die de gradiënt in weging neemt en een andere die de smoothness in weging neemt. De segmentatie is dus uiteindelijk gradiënt gebaseerd. Dit heeft als risico dat onduidelijk randen voor slechte segmentaties zorgen. Vandaar dat extra kosttermen in rekening worden gebracht die rekening houden met de intensiteit. We voegen twee termen toe:



Figuur 4.7: Niet equidistante bemonstering



Figuur 4.8: Equidistante bemonstering

- De binnenkant van de metastase wordt door de Spiral Scan-transformatie getransformeerd naar de bovenzijde van het beeld. De pixels bovenaan het beeld zouden dus een waarde in de buurt van de verwachtingswaarde van weefsel uit een metastase moeten hebben. De grijswaarden binnen een metastase kan normaal verdeeld verondersteld worden (zie sectie 3.2.1). Het is dus niet moeilijk om een kansdichtheidsfunctie hiervan op te stellen (3.4):

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(I_k - \mu)^2}{2\sigma^2}\right)$$

Hoe groter de kans om metastase te zijn, des te kleiner de kost moet zijn, vandaar de keuze van

$$\frac{1}{\sigma\sqrt{2\pi}} (1 - \exp(-\frac{(I_k - \mu)^2}{2\sigma^2}))$$

als kost voor alle pixels boven de scheidingslijn. De kost in een punt is de accumulatie van de kosten over alle pixels boven dat punt.

- Voor punten die gelegen zijn beneden in het getransformeerde beeld, d.w.z. punten het verst verwijderd van het middelpunt van de metastase, is het nuttig een kost toe te kennen aan pixels die een hoge waarschijnlijkheid hebben metastase te zijn. Hoe groter de kans om metastase te zijn, des te groter de kost moet zijn, vandaar de keuze als kost van

$$\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(I_k - \mu)^2}{2\sigma^2})$$

Versterking van de gradiënt aan de randen van de metastasen

Om deels het probleem op te lossen dat hoge gradiënten van omliggende structuren het minimale pad naar zich toe trekken, wordt de gradiënt van de randen van de metastasen versterkt. Dit gebeurt door gebruik te maken van één van volgende mogelijkheden:

- Interval thresholding: om het beeld binair te segmenteren. De gradiënt van dit binaire beeld wordt bij de gradiënt van het beeld opgeteld.
- Binary classification: hier wordt het binaire beeld geleverd door een binair pixelclassificatie-algoritme
- Fuzzy classification: in plaats van een gradiënt van een binaire segmentatie bij een de beeld-gradiënt op te tellen kan ook de gradiënt genomen worden van het resultaat van een fuzzy pixelclassificatie.

4.4.6 Conclusie

De segmentatie op basis van de Spiral Scanning Technique heeft zowel voor- als nadelen. Het belangrijkste voordeel is dat op deze manier uitgroei onmogelijk is. Daarnaast is er relatief weinig gebruikersinteractie nodig, nl. twee muisklikken, één in het midden van de metastase en één voor het aangeven van de maximale straal.

De nadelen zijn de mindere nauwkeurigheid, vooral bij datasets met dikke snedes en bij grote metastasen. Ook trekken hoge gradiënten van omliggende structuren het minimaal kostpad naar zich toe.

4.5 Circle Scanning Technique

De Circle Scanning Technique is onze variant van de Spiral Scanning Technique gericht op segmentatie in datasets met een lage z-resolutie. Het is een methode waarbij de segmentatie snede per snede gebeurt.

4.5.1 Basisalgoritme

De Circle Scanning Technique maakt gebruik van een transformatie van elke snede van een metastase van de 3D-data. De eerste stap is de plaatsing van een seed binnen de metastase op elke snede. Daarna vindt het radiaal scannen van het beeld plaats om zo een transformatie van een cartesiaans assenstelsel naar een polair assenstelsel te bekomen. De uiteinden van de scanlijnen liggen dus op cirkels in het oorspronkelijke beeld. De wiskundige beschrijving van een cirkel is een eenvoudige parametrisatie, gegeven door:

$$\begin{aligned}x &= x_C + \rho \cos t \\y &= y_C + \rho \sin t \\z &= z_C\end{aligned}$$

Hierin is $t \in [0, 2\pi]$ de variabele parameter. De transformatie vindt plaats door ρ te laten variëren van 0 tot r_{max} voor elke waarde van t om op die manier een beeld te verkrijgen i.f.v. t en ρ .

Op het getransformeerde beeld gebeurt de berekening van het MCP op dezelfde manier als in 4.4.2. Het enige verschil is dat het om een kleiner beeld gaat, waardoor de minimale kostpadberekening sneller verloopt. De terugtransformatie is eenvoudiger dan in 4.4.3. Voor elke waarde in het xy-vlak berekent het algoritme ρ , t_1 en t_2 met

$$\begin{aligned}\rho &= \sqrt{(x - x_C)^2 + (y - y_C)^2} \\t_1 &= \arccos\left(\frac{x - x_C}{\rho}\right) \\t_2 &= \arcsin\left(\frac{x - x_C}{\rho}\right).\end{aligned}$$

Hierin geldt dat $t_1 \in [0, \pi]$ en $t_2 \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Op basis van t_1 en t_2 is het mogelijk een eenduidige waarde voor t te bepalen. Aangezien ρ en t meestal geen gehele getallen zijn, is een interpolatie nodig tussen de vier omliggende waarden. Aangezien het om een binair beeld gaat, kan ook het nearest-neighborprincipe of zelfs eenvoudige afronding toegepast worden, aangezien bij interpolatie opnieuw afronding nodig is om een binair beeld te bekomen.

Wanneer de metastase in een snede gesegmenteerd is, wordt deze snede toegevoegd aan de snedes waar de segmentatie reeds op gebeurde. Uiteindelijk bekomt het algoritme een 3D-blok met daarin de gesegmenteerde metastase.

4.5.2 Problemen met en verbeteringen aan het algoritme

De problemen die opduiken bij toepassing van het algoritme zijn veeleer dezelfde als bij Spiral Scanning Technique. Sterke gradiënten van omliggende structuren trekken het MCP naar zich toe.

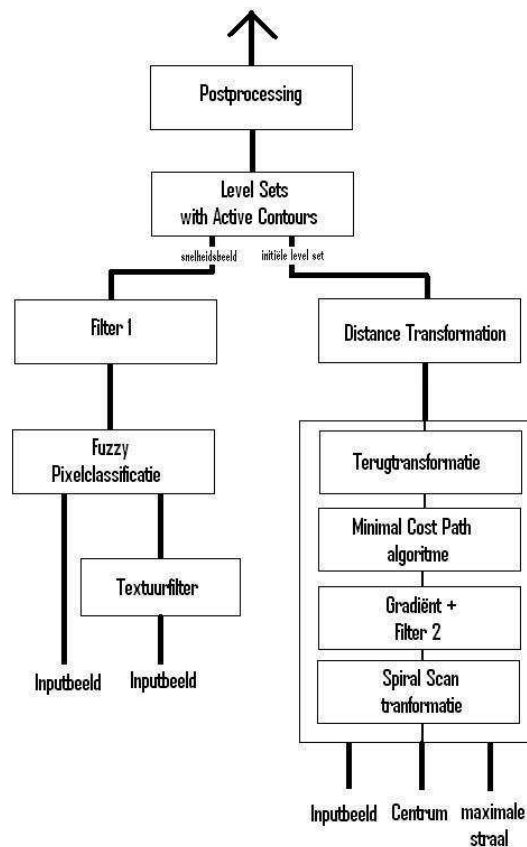
Net als bij de Spiral Scanning Technique is het belangrijk het zaadpunt vrij goed in het midden van de metastase te plaatsen, hetgeen in 2D veel gemakkelijker gaat dan in 3D, zodat dit in feite geen probleem is.

4.5.3 Conclusie

De Circle Scanning Technique is zeer verwant met de Spiral Scanning Technique en heeft ook dezelfde voor- en nadelen. Deze methode vergt meer gebruikersinteractie maar geeft betere resultaten voor datasets met grote snededikte.

4.6 Level Sets met Circle- of Spiral Scanning

De algemene beschrijving van Level Sets is terug te vinden in Sectie 2.4.2. Een belangrijke toepassing van level set-algoritmes is vormdetectie en -herkenning. Een initiële segmentatie wordt hierbij vervormd naar de grenzen van een object, hier de metastasen in de lever. De implementatie van het volledige algoritme op basis van Level Sets kan verduidelijkt worden a.d.h.v. Figuur 4.9.



Figuur 4.9: Blokdiagramma van de Level Set methode

De kern van het Level Sets met Spiral Scanning-algoritme verandert een initiële Level Set, weergegeven door de rechtertak in Figuur 4.9, a.d.h.v. een snelheidsbeeld, weergegeven door de linkertak. We bespreken deze aspecten in de komende paragrafen meer in detail.

4.6.1 Kernalgoritme

Het kernalgoritme is in Figuur 4.9 aangegeven met het “Level Sets met actieve contouren”-blok. Dit is een beschikbaar algoritme uit ITK. Het transformeert een initiële segmentatie onder invloed van een snelheidsbeeld. De initiële segmentatie wordt voorgesteld als een impliciete functie, waarbij elk punt de kortste afstand weergeeft tot het gesegmenteerde oppervlak. Deze impliciete functie, de initiële level set functie ϕ , evolueert door oplossing van volgende partiële differentiaalvergelijking:

$$\frac{d}{dt}\phi = -\alpha\mathbf{A}(\mathbf{x}) \cdot \nabla\phi - \beta P(\mathbf{x})|\nabla\phi| + \gamma Z(\mathbf{x})\kappa|\nabla\phi| \quad (4.11)$$

waar \mathbf{A} een advectieterm is, P een propagatie- (of expansie-)term en Z een term voor de spatiale aanpassing van κ , de kromming. De scalaire constanten α , β en γ wegen de relatieve invloed van elk van deze termen op de beweging van het oppervlak. De propagatieterm en de krommingsterm zijn bepaald door het snelheidsbeeld: $P(\mathbf{x}) = Z(\mathbf{x}) = I_V(\mathbf{x})$. De propagatieterm $P(\mathbf{x})$ zorgt ervoor dat de zero level set snel naar buiten beweegt daar waar het snelheidsveld hoge waarden heeft en die weinig (uitwaarts) beweegt waar het snelheidsbeeld lage waarden heeft. Het effect van $\gamma Z(\mathbf{x})\kappa|\nabla\phi|$ op de evolutie van de level set is enerzijds afhankelijk van $\kappa|\nabla\phi|$, anderszijds van $Z(\mathbf{x}) = I_V(\mathbf{x})$. Het toepassen van de Laplace operator op de level set, $\kappa|\nabla\phi| = \nabla \cdot \nabla\phi$ (zie Bijlage A), meet de kromming van de level set. Waar deze groot is, zal de level set snel inwaarts bewegen indien het snelheidsveld op die plaats een hoge waarde heeft. Het kernalgoritme houdt echter alleen rekening met de gemiddelde kromming, zodat het snelheidsveld bepaalt hoe groot de inwaartse beweging is. De advectieterm wordt geconstrueerd uit de negatieve gradiënt van het snelheidsbeeld: $\mathbf{A}(\mathbf{x}) = -\nabla I_V(\mathbf{x})$. Daar waar deze gradiënt negatief is, zal de component van deze steilste daling volgens de richting van de steilste helling van de level een expansie van de level set veroorzaken. Waar de gradiënt van het snelheidsbeeld positief is, zal dit voor een inkrimping zorgen. De grootte van de inkrimping en van de expansie is afhankelijk van de grootte van de helling in de level set. Het effect van de verschillende termen is dus afhankelijk van het snelheidsbeeld en zal in 4.6.3 verder verklaard worden. Bijlage A geeft verdere details betreft numerieke implementatie in A.3.

4.6.2 Berekenen van de initiële Level Set

Vooraleer het kernalgoritme toe te kunnen passen, is een eerste, snelle segmentatie nodig. Hoe beter deze initiële segmentatie is, des te groter de kans op goede uiteindelijke segmentatie en des te sneller het kernalgoritme naar de uiteindelijke segmentatie convergeert. Het berekenen van de initiële Level Set gebeurt in twee stappen: een primaire segmentatie, gevolgd door afstandstransformatie.

Voor de primaire segmentatie kan elk van de vorige algoritmes dat de metastases afzonderlijk kan segmenteren gebruikt worden (zie 4.3, 4.4 en 4.5). Voor datasets met een hoge axiale resolutie gaat de voorkeur uit naar de Spiral Scanning Technique (4.4), met een lage axiale resolutie naar Circle Scanning Technique (4.5). Deze primaire segmentatie levert een binair beeld op. Dit binaire beeld wordt nu vertaald naar een afstandsfunctie. Dit gebeurt met een afstandstransformatie. Die berekent van elke voxel in de data de afstand tot het oppervlak van de segmentatie. In dit geval gebruiken we de euclidische afstand tussen de voxel en de dichtstbijzijnde punt dat op het oppervlak gelegen is. Om een geschikte initiële level set te verkrijgen moeten de afstanden tot het oppervlak binnen de segmentatie negatief zijn om op die manier een impliciete functie te bekomen.

4.6.3 Berekenen van het snelheidsbeeld

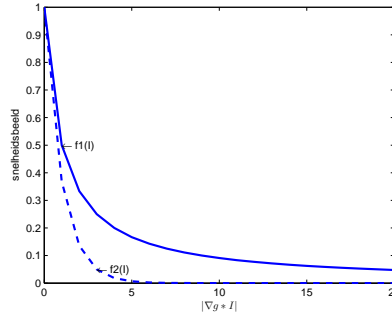
Het snelheidsbeeld zorgt ervoor dat de initiële level set in een bepaalde richting gestuurd wordt. Dit beeld moet getransformeerd worden in een snelheidsbeeld (Filter 1 in Figuur 4.9). Er zijn

verschillende transformaties mogelijk.

$$I_V(\mathbf{x}) = f_1(I) = \frac{1}{1 + |\nabla g * I|}$$

$$I_V(\mathbf{x}) = f_2(I) = \exp(-|\nabla g * I|)$$

Hierin is $g(\vec{r}) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-r^2}{2\sigma^2}\right)$ de gediscrèteerde Gaussiaanse functie [9]. De beide transformaties worden geïllustreerd in Figuur 4.10 met op de x-as de ingangswaarden, hier het fuzzy pixelclassificatiebeeld, en op de verticale as de uitgangswaarden van de filter.



Figuur 4.10: Twee filters die vaak in de literatuur gebruikt worden

In de veronderstelling dat de gradiënt van het beeld I groot is aan het metastase-oppervlak wordt een snelheidsbeeld bekomen dat hoge waarden heeft op plaatsen waar de level set veel mag veranderen en lage waarden rond het metastase-oppervlak. Dit heeft als gevolg dat de propagatieterm de zero level set steeds naar buiten zal bewegen, maar minder rond het metastase-oppervlak. De krommingsterm zal steeds voor inkrimping zorgen en des te meer wanneer de zero level set zich verder van het metastase-oppervlak bevindt. De advectieterm zal de zero level set naar het metastase-oppervlak toetrekken. Hij werkt dus samen met de propagatieterm als de zero level set binnen de metastase dichtbij het oppervlak ligt en is op die manier de belangrijkste drijvende kracht omdat de propagatie dichtbij het oppervlak klein is. Anderzijds werkt de advectieterm de propagatie tegen wanneer de zero level set buiten het metastase-oppervlak ligt. Het is dan belangrijk dat de advectieterm groter is dan de propagatieterm. Beide filters hebben twee nadelen:

- Bij grote waarden van $|\nabla g * I|$ worden grote intervallen afgebeeld op kleine intervallen. Een duidelijke rand geeft een groot verschil in $|\nabla g * I|$, maar een veel kleiner verschil in het snelheidsbeeld. Kleine verschillen bij lage waarden van $|\nabla g * I|$ zoals ruis, leiden tot grote verschillen in het snelheidsbeeld. Dit is afgeleid worden uit Figuur 4.10.
- De veronderstelling van een duidelijke rand van de metastasen is meestal niet van toepassing. Bovendien wordt het verschil in gradiënt nog verkleind door de filters. Op die manier is de advectieterm onvoldoende sterk om de zero level set naar de metastaserand toe te trekken.

Om aan deze tekortkomingen tegemoet te komen, stellen we twee alternatieven voor. Ze maken beide gebruik van fuzzy statistische pixelclassificatie (zie 4.2) i.p.v. het oorspronkelijke grijswaardenbeeld. Een eerste voorstel is opnieuw f_1 , maar toegepast op een fuzzy pixelclassificatiebeeld:

$$I_V(\mathbf{x}) = f_1(p(c_{metastase}|I_k)) = \frac{1}{1 + |\nabla g * p(c_{metastase}|I_k)|}$$

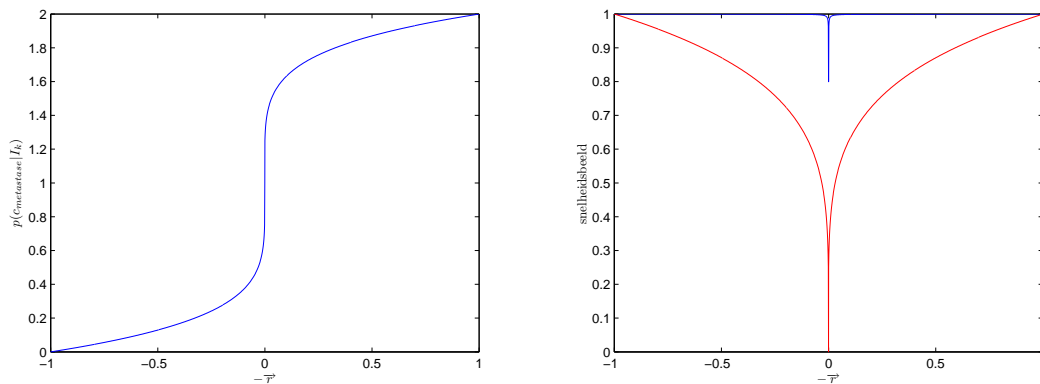
Dit geeft duidelijk betere resultaten omdat door de pixelclassificatie de rand tussen metastase en lever duidelijker wordt. Om de ruis te onderdrukken maar de randen te behouden, wordt een anisotrope diffusie-operator toegepast op het fuzzy pixelclassificatiebeeld. Deze filter f_1 behoudt

echter nog altijd het nadeel dat grote verschillen in gradiënt van het pixelclassificatiebeeld leiden tot kleine verschillen in het snelheidsbeeld.

Volgend filter komt aan dit probleem tegemoet:

$$I_V(\mathbf{x}) = f_3(p(c_{metastase}|I_k)) = |2p(c_{metastase}|I_k) - 1|$$

Ook hier wordt eerst een anisotrope diffusie-operator toegepast. De vergelijking van beide filters voor een theoretisch verloop van het fuzzy pixelclassificatiebeeld over de rand is geïllustreerd in Figuur 4.11 en Figuur 4.12. Als de gradiënt in het pixelclassificatiebeeld niet extreem groot is



Figuur 4.11: Een typisch theoretisch verloop van Figuur 4.12: Resultaat van het snelheidsbeeld de fuzzy pixelclassificatie overheen een rand van voor f_1 (blauw) en f_2 (rood) een metastase

(zoals in de reële beelden), is duidelijk dat het snelheidsbeeld een diepere en bredere kloof vertoont voor filter f_3 . De breedte is vooral belangrijk voor de advectieterm die zo een segmentatie, die verder van de rand van de metastase gelegen is, beter naar het metastaseoppervlak kan toetrekken. Doordat de kloof dieper is, zal de kans kleiner zijn dat de propagatieterm de zero level set weg van de rand duwt. Dit betekent ook dat de goede werking sterker afhankelijk is van de kwaliteit van de fuzzy pixelclassificatie. Indien deze incorrecte randen bevat, zal de level set daar gemakkelijker blijven plakken. Uit testen blijkt dat het gebruik van f_3 betere resultaten geeft.

4.7 Besluit

Dit hoofdstuk besprak verschillende segmentatiemethodes voornamelijk gebaseerd op de fotometrische eigenschappen van de lever en de metastasen. De grijswaarden zijn belangrijke eigenschappen van de voxels om lever en metastase van elkaar te onderscheiden. Alle vormen van thresholding maken hier rechtstreeks gebruik van, de vormen van pixelclassificatie door middel van een statistische model. Textuurkenmerken leveren een beperkte meerwaarde door de grote rekentijd vereist voor hen te berekenen. Impliciet zorgen vormbeperkingen ervoor dat ook geometrische eigenschappen in de methodes zijn opgenomen. Dit gebeurt door het gebruik van een masker en de connectiviteitsvoorwaarde bij region growing, de externe kost bij Spiral- en Circle Scanning Technique en de krommingsterm bij level sets. In het volgende hoofdstuk volgen de resultaten van de verschillende methodes.

Hoofdstuk 5

Technische Validatie

In dit hoofdstuk volgt de validatie van de eerder besproken segmentatiemethodes. Dit maakt het mogelijk om objectief de verschillende methodes met elkaar te vergelijken. Eerst volgt een korte bespreken van mogelijke maatstaven om de methodes te vergelijken. Daarna volgt de validatie van de methodes. Ten slotte worden de bekomen resultaten grondig besproken.

Naast de trainingsdata die gebruikt worden om de methode te optimaliseren worden ook test-data gebruikt. De testdata zijn afkomstig van een studie voor de opvolging van tumoren in de lever. In totaal bestaat het bestand uit 28 verschillende datasets van 14 patiënten, opgenomen in het UZ Leuven. Dit komt neer op een totaal van 72 tumoren. Daarvan worden 19 tumoren buiten beschouwing gelaten tijdens de validatie. In totaal blijven er dus 53 tumoren over waarop kan getest worden. De data bevatten de aflijningen van de metastasen en informatie over RECIST en de loodrechte op RECIST. Het gebruikte protocol is hetzelfde als voor de trainingsdata (zie 3.1).

5.1 Maatstaven

Er zijn verschillende maatstaven om de segmentatie door een algoritme te toetsen aan een handmatige aflijning. Deze sectie bespreekt de meest courante op basis van artikels uit de literatuur.

5.1.1 Sensitiviteit en specificiteit

In de meeste standaardwerken gebruikt men als maatstaf de sensitiviteit, gedefinieerd als:

$$\text{sensitiviteit} = \frac{TP}{TP + FN}, \quad (5.1)$$

met

- True Positives (TP): dit zijn de voxels die de segmentatiemethode als tumor aanduidt en die dit in werkelijkheid ook zijn.
- True Negatives (TN): dit zijn de voxels die de segmentatiemethode niet als tumor aanduidt en die dit in werkelijkheid ook niet zijn.
- False Positives (FP): dit zijn de voxels die de segmentatiemethode als tumor aanduidt maar die dit in werkelijkheid niet zijn.
- False Negatives (FN): dit zijn de voxels die de segmentatiemethode niet als tumor aanduidt maar die dit in werkelijkheid wel zijn.

Op basis hiervan bepaalt men fracties¹:

$$\begin{cases} TPF = \frac{TP}{TP+FN} = \text{sensitiviteit} \\ TNF = \frac{TN}{TN+FP} = \text{specificiteit} \\ FPF = \frac{FP}{TN+FP} = 1 - \text{specificiteit} \\ FNF = \frac{FN}{TP+FN} = 1 - \text{sensitiviteit} \end{cases}$$

Aangezien het moeilijk is om de TN te bepalen, aangezien dit afhankelijk is van het beeld, voegen we een extra fractie toe:

$$FPF' = \frac{FP}{TP + FP}$$

De FPF' is een maat voor het te groot zijn van de segmentatie, FNF voor het te klein zijn. Deze maten moeten steeds samen bekeken worden en zijn in het optimale geval nul. In de meta-analyse in [1] is de gemiddelde FNF voor verschillende methodes om levermetastasen af te lijnen gelijk aan 36,2% met een betrouwbaarheidsinterval [27,8%;45,6%].

5.1.2 Dice Similarity Coefficient en Jaccard Similarity Coefficient

Een andere mogelijk maatstaf is de Dice Similarity Coefficient (DSC), ook Dice's coefficient genoemd. De DSC is nuttig voor het meten van spatiale overlap tussen twee regio's A en B [43]:

$$DSC(A, B) = \frac{2(A \cap B)}{A + B} \quad (5.2)$$

Indien er geen overlap is, is de DSC gelijk aan 0, bij volledige overlap 1. De DSC beboet zowel de FP als de FN. Het is ook mogelijk de DSC uit te drukken als functie van TP, FP en FN:

$$DSC(TP, FP, FN) = \frac{2TP}{2TP + FP + FN}$$

De Jaccard Similarity Coefficient (JSC) of Jaccard index, wordt gegeven door [44, 45]:

$$JSC(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (5.3)$$

Het is ook mogelijk de JSC uit te drukken als functie van TP, FP en FN:

$$JSC(TP, FP, FN) = \frac{TP}{TP + FP + FN}$$

Ook de JSC is in het optimale geval 1. Ze is zeer gelijkaardig aan de DSC, maar wordt hier vermeld omdat ze in de literatuur vaak terugkomt als "volume overlap".

5.1.3 Relatieve absolute volumeverschil

Een maatstaf voor het verschil in volume tussen manuele segmentatie (A) en de (semi-)automatische segmentatie (B), uitgedrukt als percentage van het manuele volume is het Relative Absolute Volume Difference (RAV). De definitie is:

$$RAV(A, B) = \frac{|A - B|}{A} \quad (5.4)$$

In het beste geval is de RAV dus nul. Deze maat kan ook geschreven worden in functie van TP, FP en FN:

$$RAV(TP, FP, FN) = \frac{|FN - FP|}{TP + FN}$$

Opmerking:

De variant van RAV in 2D is de RAAD.

¹True Positive Fraction (TPF), True Negative Fraction (TNF), False Positive Fraction (FPF), False Negative Fraction (FNF)

5.1.4 Afstand tussen oppervlakken

Een goede maat om de kwaliteit van de (semi-)automatische segmentatie te meten is door van beide segmentaties (manuele en (semi-)automatische) een mesh te maken en te kijken hoeveel deze oppervlakken van elkaar verschillen [46]. Elke maat kan gedefinieerd worden met zowel de handmatige aflijning als de (semi-)automatische segmentatie als referentie-aflijning. De vergelijking van de oppervlakken kan op vier manieren.

De eerste is de Average Absolute Surface Distance (AASD). Van zowel de handmatige aflijning als van de (semi-)automatische aflijning wordt een mesh gemaakt. Voor elk punt in de mesh van de referentie-aflijning bepaalt het validatie-algoritme het dichtstbijzijnde punt uit de andere mesh, gebruik makend van de euclidische afstand. Dit gebeurt in wereldcoördinaten om op die manier rekening te houden met de verschillende resoluties in de verschillende scan-richtingen. Een punt dat binnen de mesh ligt, heeft een negatieve afstand. Van al deze afstanden wordt het gemiddelde de AASD genoemd. De maximale afstand is de Maximal Absolute Surface Distance (MASD) en de Root Mean Square (RMS) waarde van de afstanden is de RMS Surface Distance (RMSSD). Dit zijn tweede en derde manier om oppervlakken met elkaar te vergelijken. Verwant met de MASD is de minimale oppervlakte-afstand voor de grootste afstand binnen de mesh. Deze methodes hebben als voordeel dat ze uitgedrukt worden in millimeter waardoor de maat op een concretere manier kan voorgesteld worden. De MASD kan nul zijn wanneer het deel van de segmentatie buiten de referentie-aflijning het deel erbinnen opheft. Vandaar dat de RMSSD een meerwaarde biedt. Uit de MASD kan dan weer afgeleid worden of de aflijning gemiddeld te groot of te klein is. Een nadeel van deze maten is dat ze geen rekening houden met de grootte van de metastase, zodat grote metastase grotere fouten zullen hebben. Deze vier maten zijn in het beste geval 0 mm.

5.2 Automatische methodes

In de validatie maken we onderscheid tussen automatische methodes die een aflijning van de lever vereisen en de semi-automatische methodes die niet noodzakelijk een afgelijnde lever nodig hebben. Daarnaast is er ook een onderscheid tussen het vermogen om de methodes afzonderlijk te segmenteren. De hier besproken automatische methodes kunnen de metastasen niet afzonderlijk segmenteren, waardoor het niet gemakkelijk is om de methodes correct te valideren. Niet alle metastasen in één lever zijn namelijk handmatig afgelijnd, waardoor het niet altijd geldig is om de automatische segmentatie te vergelijken met de unie van handmatige segmentaties. Hierdoor zeggen de meeste validatiematen weinig. Enkel de FNF geeft enigszins een correct beeld van hoe goed de metastasen gesegmenteerd zijn omdat die maat geen gebruik maakt van FP. We bekijken echter ook de DSC omdat de FNF meestal monotoom stijgt of daalt i.f.v. een parameter. Het stijgen van een parameter gaat immers vaak gepaard met het groter (of kleiner) worden van de segmentatie. Hierdoor dalen de FN, metastaseweefsel dat buiten de segmentatie valt. Belangrijke knikken in de stijging of daling van de FNF zijn meestal een indicatie voor een optimale parameter. Het aanwezig zijn van niet afgelijnde tumoren heeft een effect op de DSC. De FP zouden immers opgesplitst kunnen worden in twee groepen. Enerzijds de FP_w , dit zijn de FP die werkelijk leverweefsel voorstellen en binnen de handmatige aflijningen vallen. Anderzijds de FP_v , dit zijn de FP die volgens de handmatige aflijningen geen tumoren zijn, maar toch werkelijk tumorweefsel zijn. De DSC zou dan de volgende vorm krijgen:

$$DSC = \frac{2TP}{2TP + FP_w + FP_v + FN} \quad (5.5)$$

De FP_v mogen in feite niet meegerekend worden in de DSC. Merk op dat deze automatische methodes vertrekken vanuit een gesegmenteerde lever zoals reeds eerder vermeld.

5.2.1 Interval thresholding

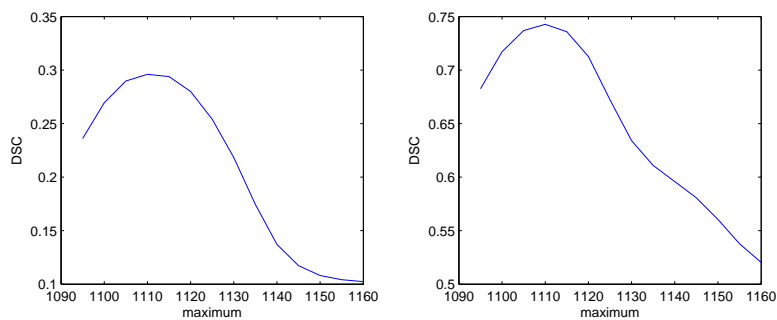
De resultaten van interval thresholding zijn afhankelijk van de boven- en benedenthreshold. De trainingsdata worden gebruikt om de optimale drempelwaardes te bepalen voor het gegeven pro-

toel. Vervolgens worden deze optimale waarden gebruikt om de testdata te valideren.

Trainingsdata

Om de optimale intervalthreshold te bepalen, wordt de ondergrens vastgelegd op 1015. Dit maakt gebruik van de waarden die berekend zijn in Sectie 3.2.1. Hieruit bleek dat de grijswaarden van de metastasen een gemiddelde waarde $\mu = 1084$ en een standaarddeviatie $\sigma = 25$ hebben. Het 99%-betrouwbaarheidsinterval ($\mu \pm 3\sigma$) stelt dat de minimale grijswaarde van de metastase met grote zekerheid boven 1009 ligt. Deze ondergrens is echter minder kritisch aangezien toch met een afgeijnde lever wordt gewerkt en er bijna geen lagere grijswaarden in de lever voorkomen.

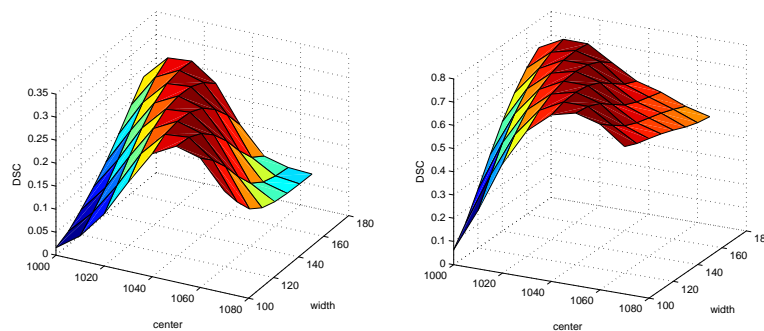
De bovengrens die men kiest, is echter wel kritisch. Om het optimum te bepalen laat men deze variëren van 1095 tot 1160. In Figuur 5.1 is de DSC weergegeven in functie van de bovengrens. Links de gemiddelde waarde voor patiënt 1 en rechts de gemiddelde waarde voor patiënt 2. Zoals uit de figuur blijkt, stemmen beide optimale waarden voor de bovengrens vrij goed overeen. Deze optimale waarde is 1110.



Figuur 5.1: DSC in functie van de maximale threshold voor patiënt 1(links) en patiënt 2(rechts)

Aangezien de FP_v in vergelijking 5.5 toeneemt in functie van de bovendrempelwaarde, betekent dit dat de werkelijke DSC afwijkt van deze weergegeven in Figuur 5.1. Het werkelijke maximum zal bijgevolg iets hoger gelegen zijn.

Gelijkaardig kan men ook het optimale interval bepalen door het centrum en de breedte van het interval te laten variëren. Dit is weergegeven in Figuur 5.2. De DSC is maximaal daar waar de som van het centrum en de helft van de breedte 1110. Het is dus niet zozeer het centrum of de breedte die van belang zijn, maar wel de bovengrens. Deze ligt opnieuw rond 1110.



Figuur 5.2: DSC in functie van de breedte en het centrum van het interval voor patiënt 1(links) en patiënt 2(rechts)

Besluit

Het nut van interval thresholding is beperkt. Dit blijkt onder andere uit de waarden voor de DSC. Het grote nadeel van thresholding is dat men geen informatie over de vorm of de omgeving gebruikt, maar enkel rekening houdt met de intensiteitswaarden. Individuele tumoren kan men niet bestuderen en de resultaten zijn bedrieglijk indien niet alle tumoren zijn afgelijnd.

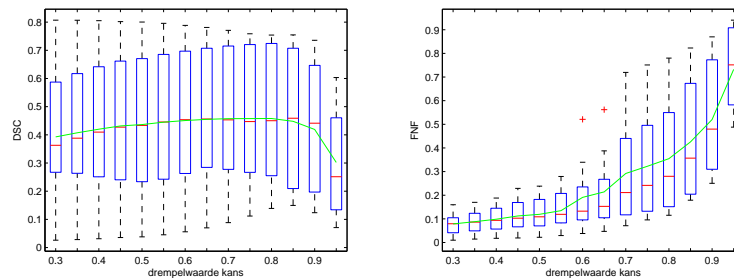
5.2.2 Statistische pixelclassificatie

Een meer gesofisticeerde methode dan thresholding voor automatische segmentatie van metastasen is statistische pixelclassificatie. We kunnen enerzijds onderscheid maken tussen gesuperviseerde en niet-gesuperviseerde methodes en anderzijds tussen een classificatie gebaseerd op intensiteiten en een classificatie gebaseerd op textuur en intensiteiten.

Gesuperviseerde pixelclassificatie

pixelclassificatie op basis van intensiteiten

Theoretisch gezien behoort een pixel tot de klasse van de metastasen als de kans dat de pixel tot de klasse van de metastase behoort groter is dan die van de klasse van het leverweefsel. In het geval van twee klassen moet de kans groter zijn dan 0,5 om tot een klasse te behoren. Omdat het moeilijk is de a priori waarschijnlijkheid ($p(c_{metastase}) = \frac{1}{2}$) aan te leren, is het niet altijd zo dat 0,5 de optimale drempelwaarde is om tot de klasse van de metastasen te behoren. Vandaar is het nuttig deze drempelwaarde een beetje te laten variëren in het interval $[0,30;0,95]$ met stappen van 0,05. We beschouwen hierbij twee validatiematen: de DSC en de FNF. Dit is weergegeven in Figuur 5.3.

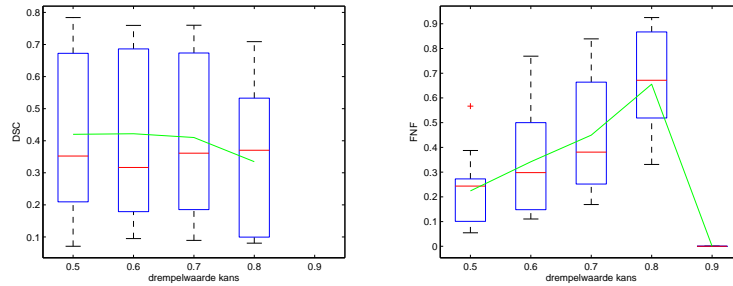


Figuur 5.3: De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten

Bij de DSC zoekt men het maximum, bij de FNF gaat men op zoek naar een grote knik in de curve die wijst op een significante toename van het aantal foutief als niet-tumor gesegmenteerde pixels. De DSC vertoont een maximum bij 0,8. Aangezien de FP_v in formule (5.5) afneemt in functie van de drempelwaarde, betekent dit dat de werkelijke DSC afwijkt van deze weergegeven in Figuur 5.3. Het werkelijke maximum zal bijgevolg iets lager zijn gelegen.

pixelclassificatie op basis van textuur en intensiteiten

Ook voor deze methode laten we de drempelwaarde variëren. Dit om dezelfde reden als voor de vorige paragraaf. Er is gekozen voor een interval van $[0,50;0,90]$ in stappen van 0,1. Het aantal stappen wordt beperkt aangezien de textuuranalyse voorafgaand aan de classificatie vrij rekenintensief is. We beschouwen opnieuw de DSC en de FNF. De resultaten zijn weergegeven in Figuur 5.4.



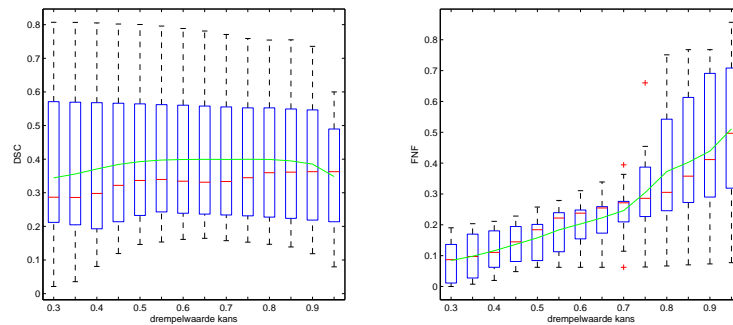
Figuur 5.4: De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten en textuur

Uit de figuur is het moeilijk een globaal optimum te bepalen. Een optimum kan gekozen worden voor de drempelwaarde gelijk aan 0,6. De FNF blijft beperkt en de gemiddelde DSC is optimaal. De resultaten van deze methode zijn iets minder goed dan voor pixelclassificatie op basis van intensiteiten. Textuurinformatie lijkt hier dus geen meerwaarde te bieden.

Niet-gesuperviseerd: Expectation Maximization Algoritme

pixelclassificatie op basis van intensiteiten

Opnieuw stelt zich de keuze waar men de drempelwaarde voor de kans dient te leggen (vanaf welke kans we een pixel als tumorweefsel beschouwen). Het gekozen interval waarover de drempelwaarde varieert is $[0,3;0,95]$ met tussenstappen van 0,05. De resultaten hiervan zijn weergegeven in Figuur 5.5.



Figuur 5.5: De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten, gebruik makend van EM

De EM gebaseerde classificatie geeft iets minder goede resultaten dan gesuperviseerde pixelclassificatie. In ruil hiervoor kan wel de trainingsstap uitgeschakeld worden en is het algoritme ook toepasbaar op beelden met een ander protocol. De DSC bereikt een maximum voor de drempelwaarde van 0,8. Opnieuw geldt de opmerking dat het optimum van de DSC mogelijk naar rechts is verschoven. Wanneer we de FNF beschouwen, zien we een grote knik na 0,7. Bij niet gesuperviseerde pixelclassificatie op basis van EM wordt de a priori waarschijnlijkheid bij elke iteratie bijgesteld, waardoor er theoretisch gezien van uitgegaan kan worden dat 0,5 de beste threshold is.

5.3 Semi-automatische methodes

Naast een korte validatie van de automatische methodes gebeurt ook een meer uitgebreide validatie van de semi-automatische methodes, die hier de metastasen wel afzonderlijk kunnen segmenteren.

Aan het einde van dit hoofdstuk wordt in Tabel 5.1 een overzicht gegeven van de resultaten van alle methodes. In Tabel 5.2 volgt een overzicht van de resultaten voor enkele methodes toegepast op een afgelijnde lever.

5.3.1 Region Growing

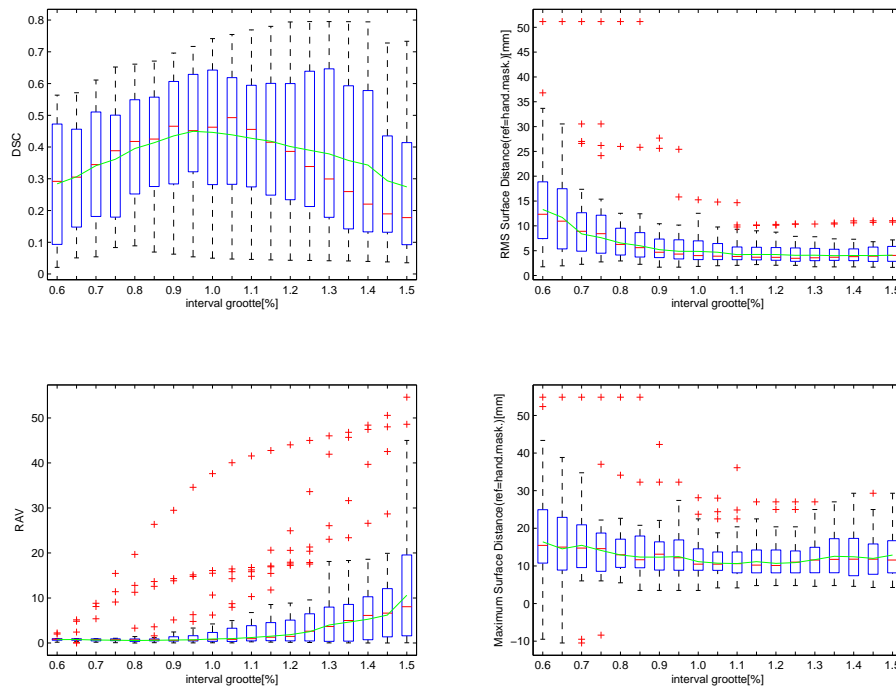
De resultaten van de verschillende region growing methodes worden beïnvloed door de grootte van het interval waarbinnen de intensiteiten van de voxels moeten liggen om tot de segmentatie te kunnen behoren. Men maakt een onderscheid tussen methodes die een trainingsstap vereisen en methodes die het interval bepalen op basis van een statistisch model met RECIST en zijn loodrechte.

Region Growing met één seed per snede in een bepaalde metastase

Bij deze methode is een seed aangeduid op iedere snede en wordt een 2D-Region Growing uitgevoerd. Uit de trainingsstap wordt de optimale intervalgrootte bepaald. De testdata gaan de algemene geldigheid van deze optimale waarde na. De resultaten hiervoor zijn weergegeven in onderstaande paragrafen.

Trainingsdata

Uit visueel onderzoek blijkt dat een goede drempelwaarde voor de intervalgrootte gelegen is rond 1%. Hiermee wordt bedoeld 1% van de maximale grijswaarde die in het beeld voorkomt boven de gemiddelde grijswaarde van de seeds en 1% eronder. Voor de trainingsstap werd gekozen om de validatiematen te berekenen in het interval $[0,6\% - 1,5\%]$ met tussenstappen van 0,05%. Als validatiematen worden de DSC, de RMSSD, de RAV en de MASD gekozen. Aan de hand hiervan wordt de optimale thresholdwaarde voor deze methode bepaald. In Figuur 5.6 worden deze maten weergegeven voor de verschillende intervalgroottes. Dit gebeurt met behulp van een boxplot waarop de mediaan en de interkwartielafstanden zijn weergegeven. De rode kruisjes stellen uitschieters voor en de groene lijn het gemiddelde zonder uitschieters.



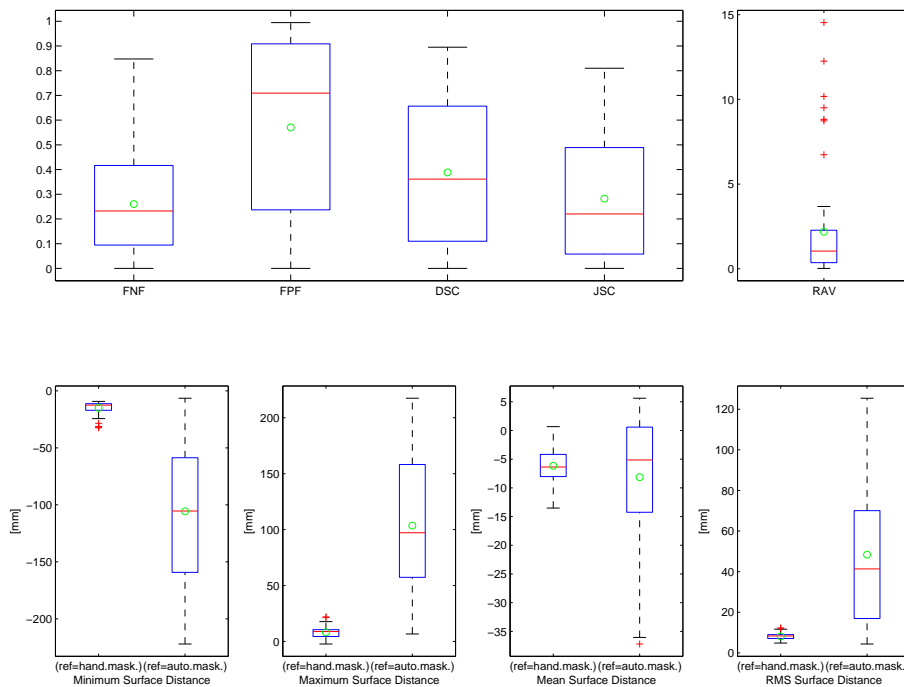
Figuur 5.6: De DSC, RMSSD, RAV en MASD in functie van de intervalgrootte voor Region Growing met één seed per snede

We zien verschillende optimale waarden naargelang de validatiemaat die we beschouwen en naargelang we het gemiddelde of de mediaan beschouwen. We proberen vervolgens op zicht een globaal optimum te zoeken. De DSC heeft een optimum bij 1,05%, de RMSSD bij 1,25%, de RAV bij 0,9% en de MASD bij 1,2%. We kiezen daarom het gemiddelde van de vier: 1,10%. Deze waarde zal gebruikt worden om de testdata te valideren.

De RAV neemt toe met de intervalgrootte. Bij grotere intervallen is de kans op uitgroei immers groter. Opvallend is ook de grootte spreiding van de DSC voor een gegeven intervalgrootte. Dit benadrukt de onbetrouwbaarheid van de region growing-methode. In Bijlage E.1 zijn de andere validatiematen weergegeven. Deze bevestigen het bekomen resultaat.

Testdata

De testdata kunnen gebruikt worden om de consistentie van het algoritme voor andere CT-beelden met een gelijkaardig protocol te onderzoeken. Van de validatiematen is het gemiddelde (uitschieters verwijderd), de mediaan, en de interkwartielafstand voorgesteld in een boxplot. Dit is weergegeven in Figuur 5.7.



Figuur 5.7: Resultaten van de 13 validatiematen voor de testdata van Region Growing met één seed per sneed

Als we de DSC vergelijken voor de testdata en de trainingsdata voor intervalgrootte 1,10% is op de merken dat de testdata duidelijk slechter scoren. De trainingsdata heeft een gemiddelde van 44% t.o.v. 40% voor de testdata. Voor de mediaan is dit respectievelijk 49% en 36%. De interkwartielafstand is ook ongeveer 10% groter voor de testdata. Een verklaring hiervoor zou kunnen zijn dat een ander type CT-scanner werd gebruikt voor een aantal testbeelden met mogelijk een andere optimale intervalgrootte. Bovendien is de uitgroei die optreedt afhankelijk van waar de tumor gelegen is (aan de rand, centraal) en deze beïnvloedt sterk de resultaten. De RMSSD geeft aan dat de automatisch afgelijnde tumoren gemiddeld 8 mm afwijken van de handmatig afgelijnde tumoren voor de testdata. Voor de trainingsdata is dit maar de helft. Bovendien is het opmerkelijk dat er een bijzonder groot verschil is voor de RMSSD afhankelijk welke aflijning als referentie gekozen wordt. Als dit de automatische is, zijn er grote afstanden tussen de verst uitgegroeide voxels en de handmatige aflijning. Dit kan zelfs meer als 12 cm zijn. Als de handmatige aflijning

de referentie is, zijn de minimale afstanden tot de automatische aflijning kleiner.

Besluit

De belangrijkste conclusie die men kan trekken is dat Region Growing met één seed per snede onvoldoende goede resultaten oplevert. De DSC ligt gemiddeld onder de 50% en er is geen restrictie tegen uitgroei in de lever of naar andere metastasen. Dit maakt de methode nogal onbetrouwbaar. Bovendien vereist men nogal wat gebruikers-interactie, zeker bij grote tumoren. Een voordeel is de rekensnelheid van het algoritme.

Region Growing met twee seeds per metastase

Bij deze methode is een seed aangeduid op begin- en eindsnede van de metastase. De andere seeds hiertussen worden automatisch bepaald door lineaire interpolatie. Voor de validatie wordt dezelfde procedure aangewend als bij de vorige methode. Deze is volledig terug te vinden in Bijlage E.

Ook hier kan hetzelfde besluit getrokken worden. De methode presteert ondermaats. De DSC blijft gemiddeld opnieuw onder de 50%. Deze methode heeft wel één voordeel t.o.v. de vorige. Het heeft haast dezelfde efficiëntie terwijl de gebruikersinteractie kan gereduceerd worden tot het plaatsen van twee zaadpunten.

Region Growing met één seed per metastase

Region Growing met één seed per metastase zorgt voor een 3D region growing vanuit het centrum van de metastase. De mate waarin de regio groeit, wordt bepaald door een zelfgekozen grijswaardeninterval. De trainingsdata worden gebruikt om dit interval optimaal te bepalen. Dit is [1010,1115]. Dit wordt vervolgens vergeleken voor de testdata. Dit is verder uitgewerkt in Bijlage E.

Region Growing met één seed per metastase geeft zeer slechte resultaten. Dit heeft een aantal oorzaken. Doordat de neighborhoodfilter alle burens in rekening neemt wordt uitgroei dan wel beperkt, maar een te streng interval zorgt er in een aantal gevallen voor dat helemaal geen groei optreedt en het algoritme faalt. De segmentatie blijft dan beperkt tot de oorspronkelijke seed. Vooral bij de testdata is dit een probleem, mogelijk omdat hier oudere beelden tussenzitten die nogal wat ruis vertonen. Hier kan het algoritme moeilijk mee overweg.

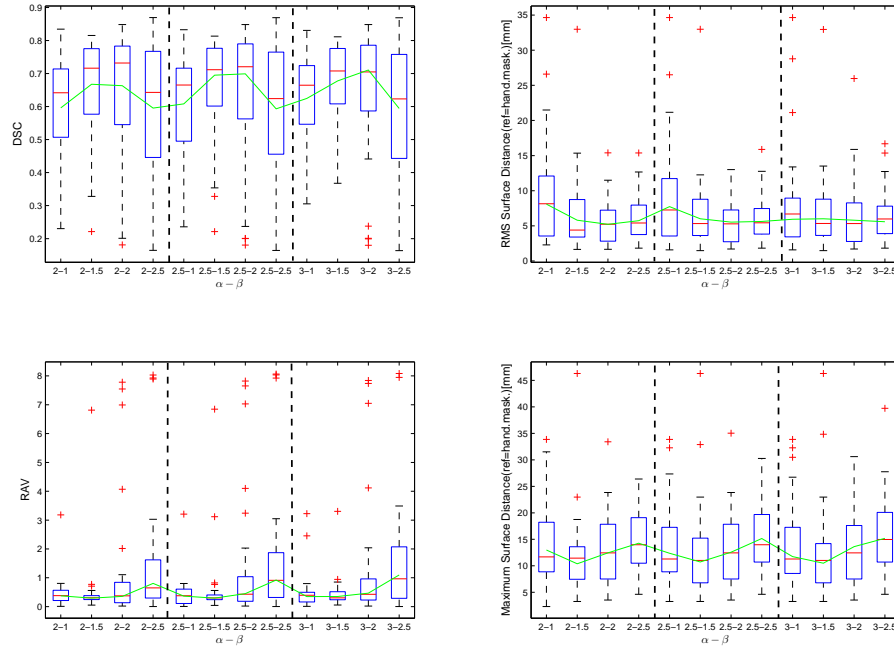
Region Growing op basis van RECIST en de loodrechte hierop (Cirkel)

Region Growing met een cirkel als restrictie maakt gebruik van een statistisch model. Het grijswaardeninterval waarbinnen gegroeid wordt, is hier berekend op basis van het gemiddelde en de standaarddeviatie van het statistisch model. Bijgevolg moeten in het interval $[\bar{T}_k - \alpha \cdot s, \bar{T}_k + \beta \cdot s]$, α en β zodanig gekozen worden dat de automatische segmentatie de handmatige zo goed mogelijk benadert.

Trainingsdata

De trainingsdata worden gebruikt om na te gaan wat de beste waarde is voor α en β . α varieert van 2 tot 3 in stappen van 0,5. β varieert van 1 tot 2,5 in stappen van 0,5. De resultaten zijn weergegeven in Figuur 5.8.

Uit de grafieken blijkt dat vooral de bovengrens, bepaald door β , kritisch is. Eénmaal de standaarddeviatie bij het gemiddelde optellen is duidelijk te weinig. Twee en een halve keer de standaarddeviatie is dan weer duidelijk te veel. De beste waarden worden verkregen voor $\alpha = 2$ en $\beta = 1,5$. Dit levert een gemiddelde DSC op van 67% en een mediaan van 72%. De RAV wijst op een foutvolume van 30%. Een groot deel van deze verbetering t.o.v. voorgaande methodes kan verklaard worden door restrictie tegen uitgroei. Daarnaast bepaalt het statistisch model de optimale threshold voor iedere tumor. Dit zorgt ook voor een aanzienlijke verbetering van de segmentatie. Resultaten voor de andere validatiematen zijn terug te vinden in Bijlage E.1.



Figuur 5.8: De validatiematen i.f.v. α en β voor Region Growing met een statistisch model en een cirkel als restrictie

Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.9.

De resultaten van de trainingsdata en de testdata verschillen aanzienlijk. De DSC heeft een gemiddelde waarde van 67% voor de trainingsdata tegenover 59,3% voor de testdata. De andere validatiematen bevestigen deze resultaten. De verklaring hiervoor kan gezocht worden in het feit dat een aantal tumoren zeer slechte resultaten geven. Dit verslechtert het gemiddelde aanzienlijk en vergroot de spreiding voor de verschillende validatiematen. Een mogelijke oorzaak hiervoor zijn bijv. handmatige aflijningen waar verschillende tumoren zijn samengenomen. Hier groeit de region growing niet altijd uit naar de verschillende tumoren.

Besluit

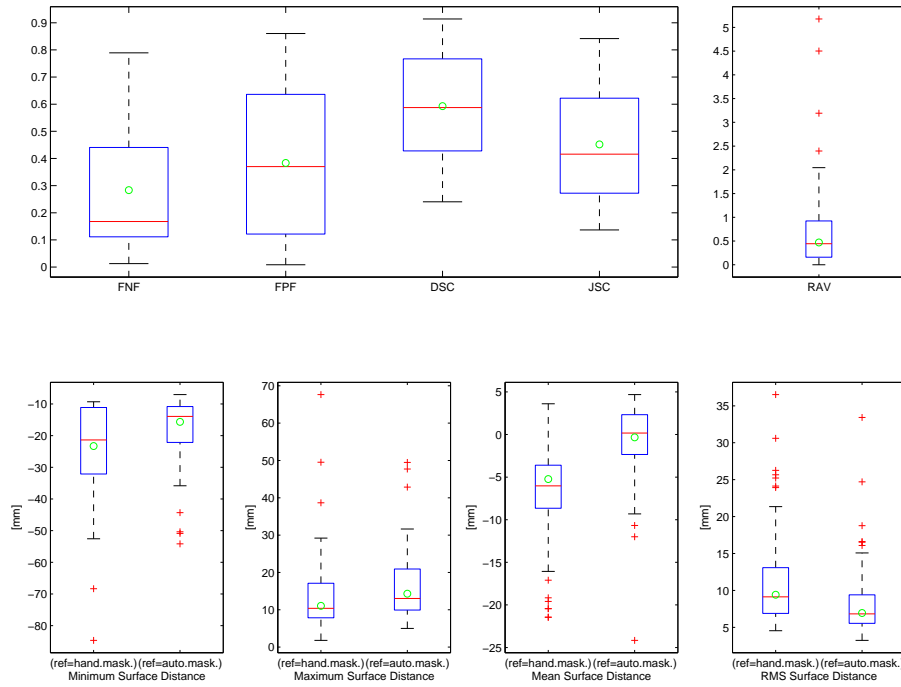
Region Growing op basis van RECIST en de loodrechte hierop met cirkelvormig masker levert aanvaardbare resultaten op. De DSC schommelt rond de 70%. Toch is het resultaat niet optimaal. Naast intensiteiten wordt ook de globale vorm van de metastase impliciet in rekening gebracht door het masker.

Region Growing op basis van RECIST en de loodrechte hierop (Rechthoek)

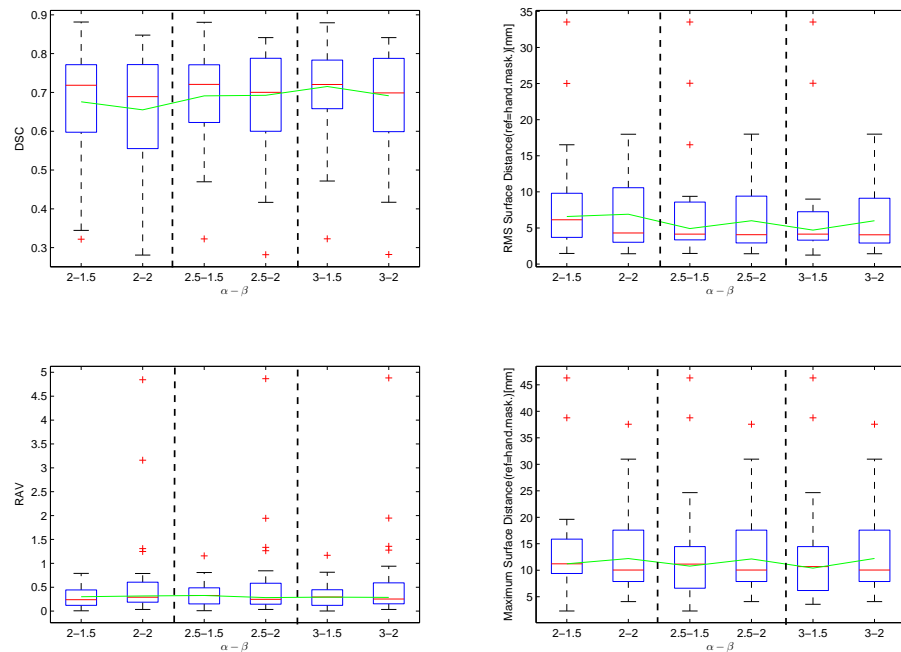
Region Growing met een rechthoek als restrictie maakt eveneens gebruik van een statistisch model. Opnieuw gaan we op zoek naar de optimale waarde voor α om het grijswaardeninterval $[\overline{I}_k - \alpha.s, \overline{I}_k + \beta.s]$ in te stellen.

Trainingsdata

Het statistisch model voor deze methode is op dezelfde manier opgebouwd als de vorige. Bijgevolg zullen de waarden voor α en β dezelfde trends veroorzaken. Aangezien uit de vorige methode is gebleken dat een aantal waarden duidelijk minder presteren, zullen deze buiten beschouwing worden gelaten. α varieert tussen 2 en 3 met stappen van 0,5. Voor β worden enkel de waarden 1,5 en 2 beschouwd. De resultaten zijn weergegeven in Figuur 5.10



Figuur 5.9: Resultaten van de 13 validatiematen voor de testdata van Region Growing met een cirkel als restrictie, $\alpha = 2$, $\beta = 1,5$



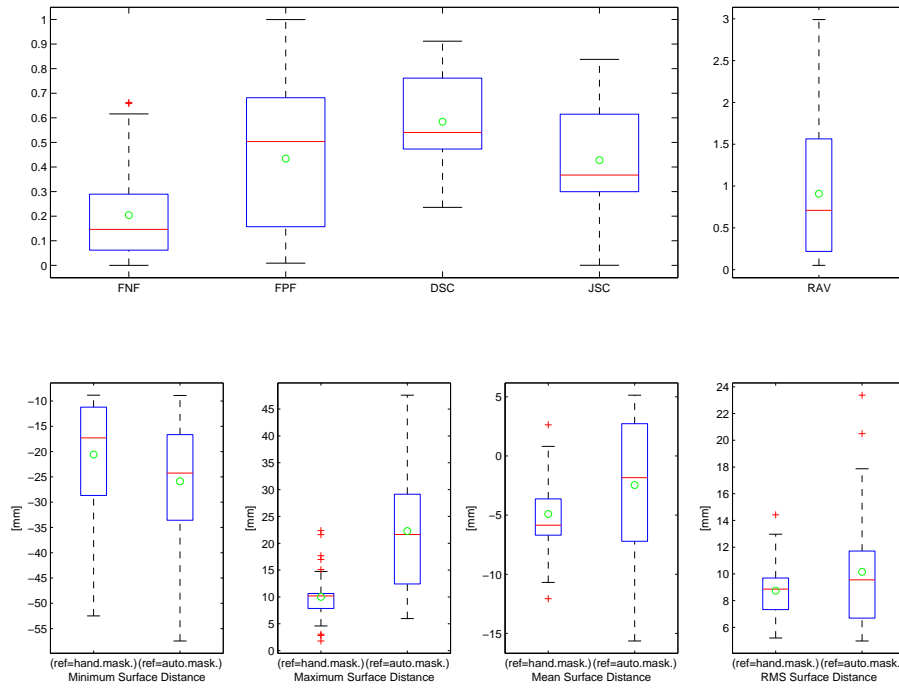
Figuur 5.10: De validatiematen i.f.v. α en β voor Region Growing met een statistisch model en een rechthoek als restrictie

De figuur geeft aan dat $\beta = 1,5$ iets betere resultaten geeft. $\beta = 2$ zorgt voor een grotere spreiding aangezien meer uitgroei zal optreden naarmate het interval groter wordt. Voor DSC, RMSSD

en RAV geven de waarden $\alpha=3$ en $\beta=1,5$ de kleinste spreiding en het beste gemiddelde. We kiezen daarom ook voor deze waarden om de testdata te valideren.

Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.11.



Figuur 5.11: Resultaten van de 13 validatiematen voor de testdata van Region Growing met een rechthoek als restrictie, $\alpha = 3$, $\beta = 1,5$

Het eerste wat valt op te merken is dat ook hier een vrij grote discrepantie bestaat tussen de resultaten van de trainingsdata en de testdata. Deze laatste presteren duidelijk minder goed. De DSC hebben een gemiddelde van 71,5% voor de trainingsdata en 58,4% voor de testdata. De andere validatiematen bevestigen dit fenomeen. Dit kan verklaard worden door een beperkt aantal datasets die mindere resultaten geven. Dit wordt bevestigd door de grotere spreiding van de verschillende validatiematen bij de testdata. Oorzaken hiervan zijn enerzijds handmatige aflijningen waar 2 of meerdere tumoren samengenomen zijn en waarbij de region growing methode maar naar een beperkt aantal van deze tumoren uitgroeit.

Besluit

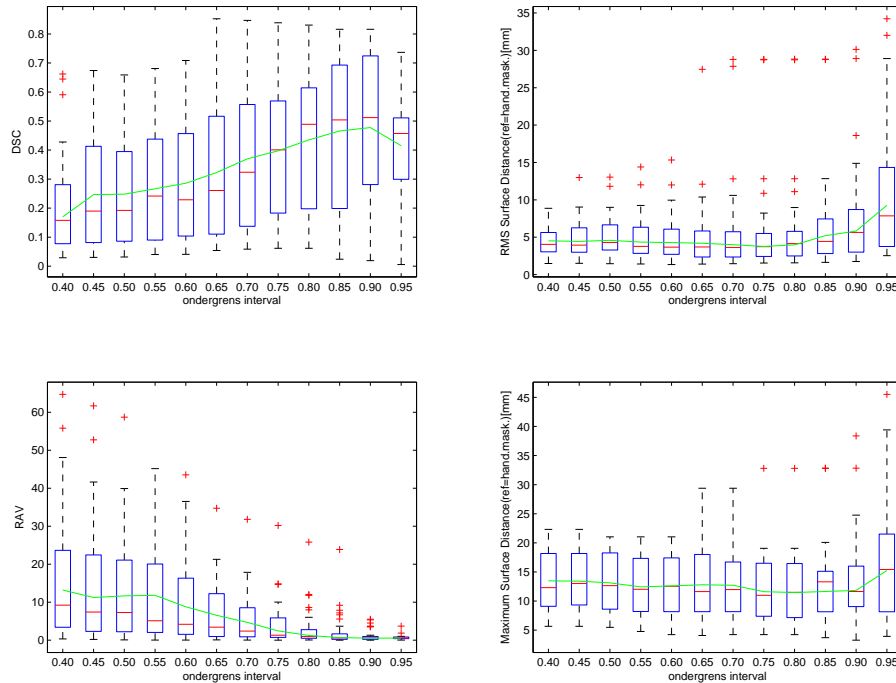
Region Growing op basis van RECIST en de loodrechte hierop (Rechthoek) geeft aanzienlijk betere resultaten dan de eenvoudige region growing methodes. De gemiddelde DSC voor de trainingsdata ligt tegen de 70%. Net als bij de vorige methode moeten we echter besluiten dat de methodes beperkt zijn. Uitgroei, al zij het beperkt, treedt nog altijd op in de axiale richting. Vooral tumoren aan de rand van lever vertonen dit probleem. Daarnaast vertoont het algoritme ook problemen wanneer meerdere tumoren samen gesegmenteerd worden.

Region Growing gegeven een fuzzy classificatiebeeld

Naast Region Growing gebruik makend van intensiteiten, kunnen we ook een region growing uitvoeren vertrekkende van een pixelclassificatie methode. De bovengrens van het interval ligt vast op 1. De ondergrens dient optimaal gekozen te worden.

Trainingsdata

Via de trainingsdata wordt de optimale ondergrens gezocht. We laten deze ondergrens variëren van 0,4 tot 0,95. De resultaten hiervoor zijn weergegeven in Figuur 5.12.



Figuur 5.12: De validatiematen i.f.v. ondergrens voor Region Growing gegeven een fuzzy classificatiebeeld

De RMSSD en de MASD geven een minimale waarde bij respectievelijk 0,75 en 0,80. De DSC en de RAV neigen eerder naar 0,90. We kiezen daarom de tussenliggende suboptimale waarde 0,85 om het algoritme te valideren voor de testdata.

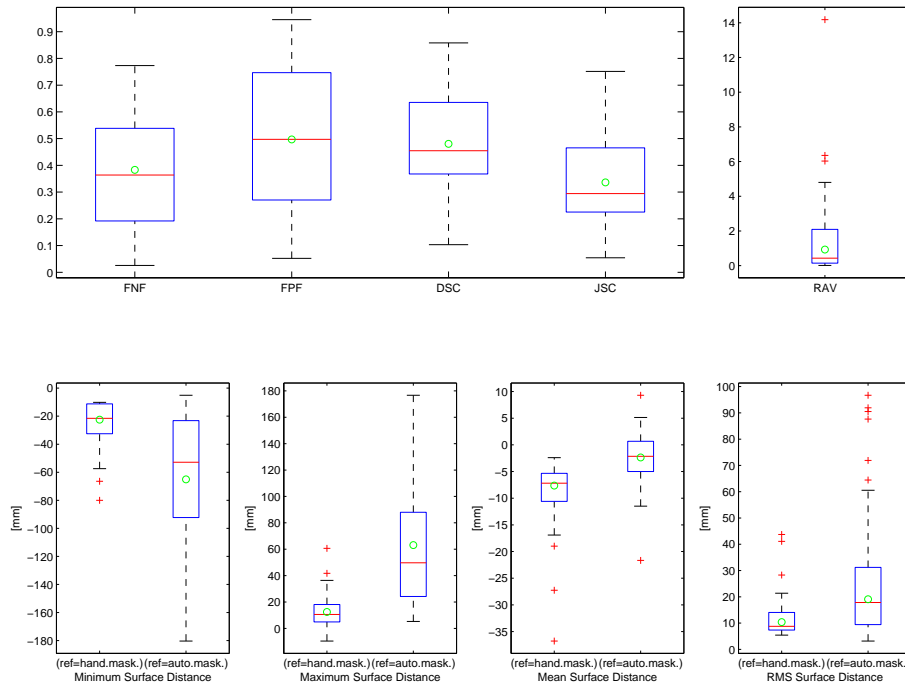
Testdata

Figuur 5.13 geeft de resultaten van de testdata toegepast op de methode voor ondergrens 0,85.

De resultaten van de testdata komen vrij goed overeen met de resultaten van de trainingsdata. De gemiddelde DSC is 46,5% voor de trainingsdata en 48% voor de testdata. Als we dit vergelijken met region growing met één seed per snede, is dit 1,5% beter voor de trainingsdata en maar liefst 9% beter voor de testdata. De pixelclassificatie levert dus weldegelijk een meerwaarde vergeleken met een gewone region growing op basis van intensiteiten. Het combineren van region growing op basis van pixelclassificatie met een cirkelvormig of vierhoekig masker om de statistieken van het fuzzy pixelclassificatiebeeld te bepalen wordt buiten beschouwing gelaten omdat het berekenen van statistieken op een beeld dat al statistieken weergeeft intuïtief geen meerwaarde lijkt te bieden. Toch zou de vormbeperking de uitgroei beter beperken.

Besluit

Door eerst een kans toe te kennen aan een voxel dat hij tot de klasse van de metastasen behoort en dit weer te geven in een beeld om er vervolgens region growing op toe te passen verbetert de segmentatie t.o.v. een region growing operatie zonder pixelclassificatie. Toch geeft deze methode nog steeds geen bevredigende resultaten.



Figuur 5.13: Resultaten van de 13 validatiematen voor de testdata van Region Growing gegeven een fuzzy classificatiebeeld, ondergrens = 0,85

5.3.2 Spiral Scanning Technique

In Spiral Scanning technique wordt een 3D beeld getransformeerd naar een 2D beeld. Dit wordt gesegmenteerd door het minimaal kostenpad te bepalen. Hiervoor moet men de som van de interne en externe kosten minimaliseren. De verhouding van de gewichten van deze twee kosten w_{int}/w_{ext} wordt via de trainingsdata geoptimaliseerd. We maken een onderscheid voor resultaten van een segmentatie op het volledige beeld en op een afgelijnde lever.

Volledig beeld

Trainingsdata

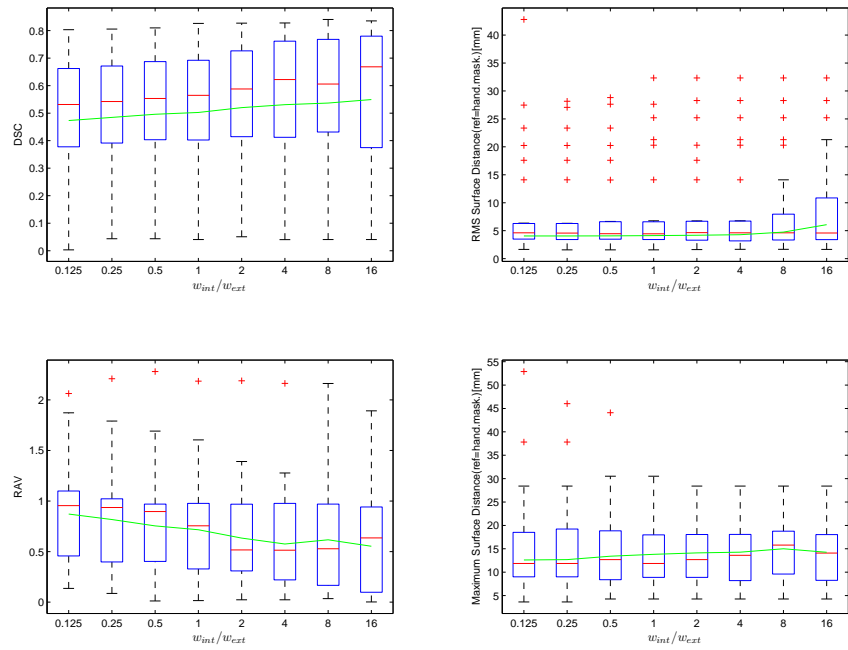
We laten de verhouding van de gewichten variëren volgens een exponentiële reeks namelijk 2^x voor x gaande van -3 tot 4 in stappen van 1. Voor $w_{int}/w_{ext} < 1$ weegt de externe kost zwaarder door. Voor $w_{int}/w_{ext} > 1$ weegt de interne kost zwaarder door. De resultaten hiervan zijn weergegeven in Figuur 5.14.

Spiral Scanning presteert minder goed dan de statische region growing methodes. Dit is echter verwacht aangezien de trainingsdata een lage axiale resolutie hebben en de 3D informatie in de beelden beperkt is. Naarmate w_{int}/w_{ext} groter wordt, nemen de DSC en de RAV licht toe. De RMSSD en de MASD nemen echter toe. Voor grote waarden van w_{int}/w_{ext} vertonen de data echter een grotere spreiding. We kiezen daarom voor $w_{int}/w_{ext}=2$ om de testdata te valideren.

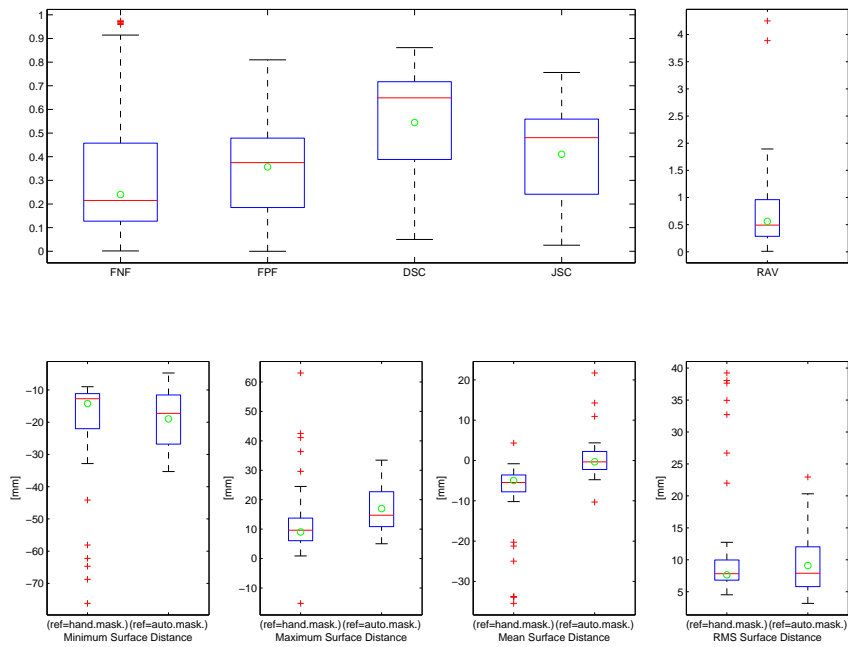
Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.3.2.

De testdata geven iets betere resultaten in vergelijking met de trainingsdata, voornamelijk omdat een aantal uitschieters waar het algoritme faalt, zijn weggelaten. Dit levert een gemiddelde DSC op van 54,5% tegenover 52,0% bij de trainingsdata. Door de outliers te verwijderen geven de testdata een realistischer beeld van de resultaten van de methode. Indien we de MASD vergelijken met zijn equivalent (MASD met het automatisch masker als referentie) in Bijlage E.1, merken we op voor Spiral Scanning deze waarden van dezelfde grootteorde zijn. Dit wijst erop dat de uitgroei



Figuur 5.14: De validatiematen i.f.v. w_{int}/w_{ext} voor de Spiral Scanning methode



Figuur 5.15: Resultaten van de 13 validatiematen voor de testdata van Spiral Scanning

beperkt blijft (gemiddeld een maximale afstand kleiner dan 17 mm). Dit is één van de grote voordelen van Spiral Scanning.

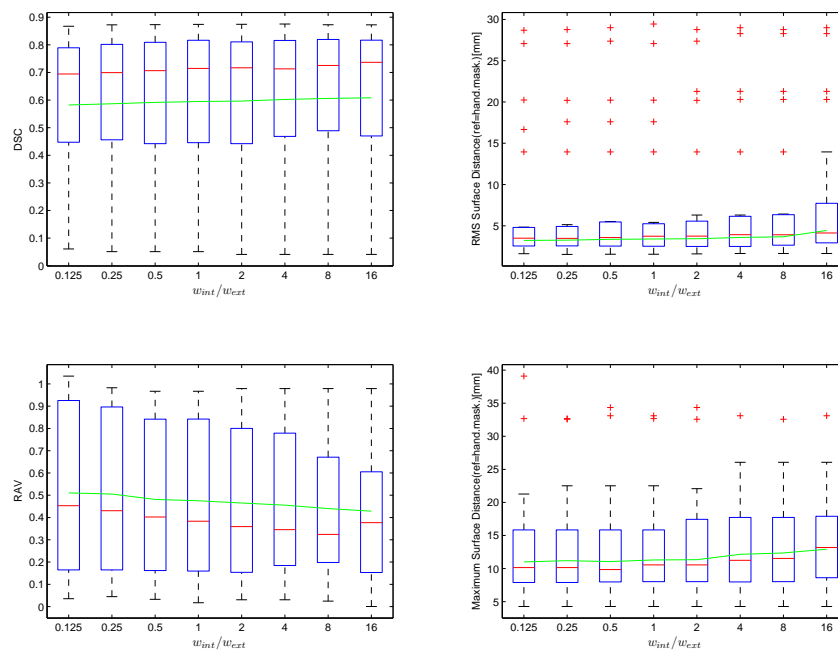
Besluit

Omwille van de grote snededikte geeft de Spiral Scanning Technique onvoldoende resultaten. Doordat een maximale straal moet meegegeven worden aan het algoritme is uitgroei beperkt. De gebruiker moet slechts tweemaal klikken per tumor. De enige andere methode met zo weinig gebruikersinteractie (region growing met één zaadpunt per metastase) is duidelijk veel minder performant.

Afgelijnde lever

Trainingsdata

We laten de verhouding van de gewichten variëren volgens een exponentiële reeks namelijk 2^x voor x gaande van -3 tot 4 in stappen van 1. De resultaten hiervan zijn weergegeven in Figuur 5.16.



Figuur 5.16: De validatiematen i.f.v. w_{int}/w_{ext} voor de Spiral Scanning methode met een afgelijnde lever

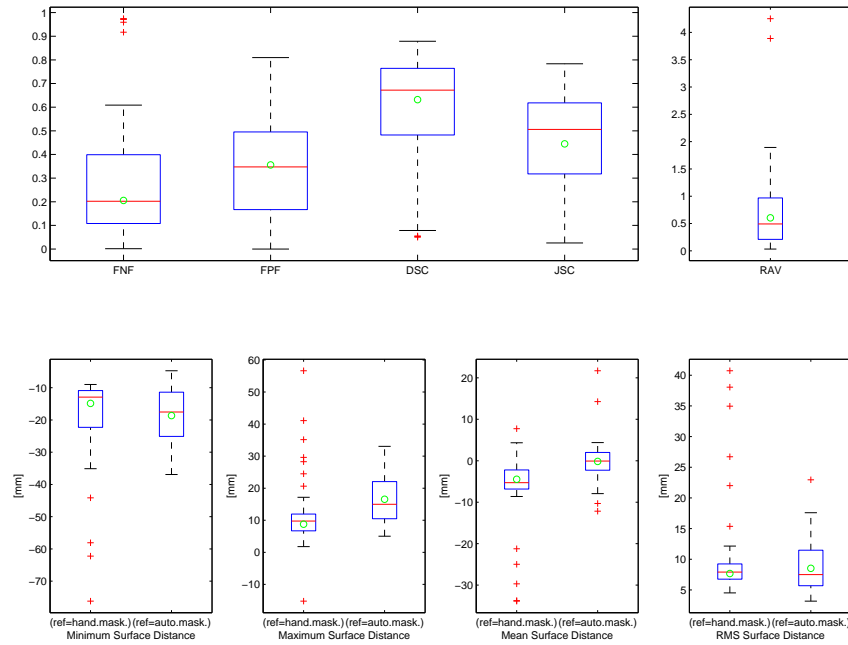
Ook hier is het niet eenvoudig om een optimale waarde te vinden voor w_{int}/w_{ext} . Voor de vier maten lijkt 2 wel een goed compromis te zijn. Deze waarde testen we op de testdata.

Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.17. De resultaten voor een afgelijnde lever zijn beduidend beter (bijna 10% in DSC) dan zonder afgelijnde lever.

Besluit

Zelfs met een afgelijnde lever, presteert de methode onvoldoende om in de klinische praktijk te kunnen gebruikt worden op data met een grote snededikte.



Figuur 5.17: Resultaten van de 13 validatiematen voor de testdata van Spiral Scanning op een afgelijnde lever, $w_{int}/w_{ext} = 2$

5.3.3 Circle Scanning Technique

De Circle Scanning Technique transformeert het beeld naar polaire coördinaten om vervolgens een minimaal kostpad algoritme toe te passen. In sectie 4.5 is een externe en interne kost gedefinieerd voor dit kostenpad. De verhouding van de gewichten van deze twee kosten w_{int}/w_{ext} wordt via de trainingsdata geoptimaliseerd. We maken een onderscheid voor resultaten van een segmentatie op het volledige beeld en op een afgelijnde lever.

Volledig beeld

Trainingsdata

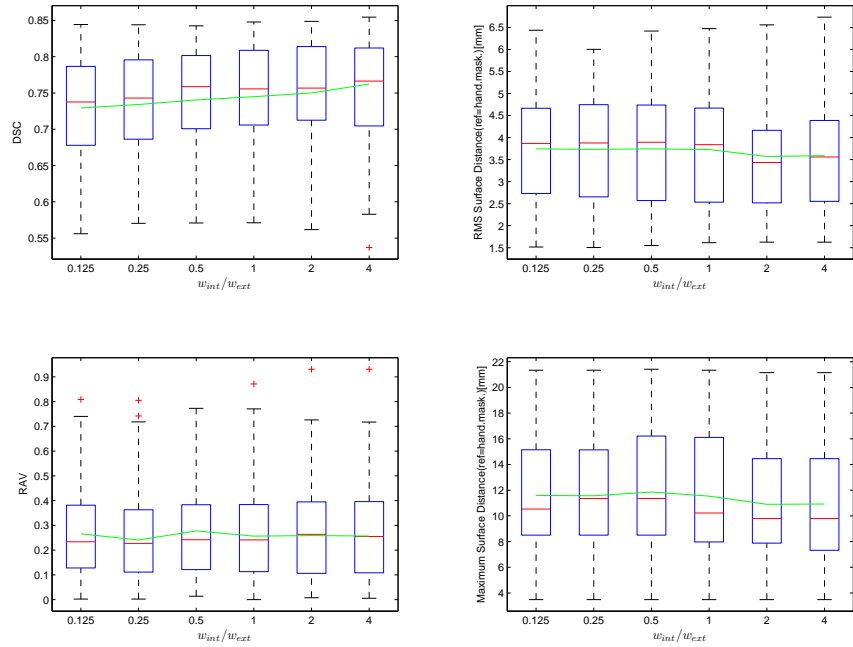
We laten de verhouding van de gewichten variëren volgens een exponentiële reeks namelijk 2^x voor x gaande van -3 tot 4 in stappen van 1. Voor $w_{int}/w_{ext} < 1$ weegt de externe kost zwaarder door, Voor $w_{int}/w_{ext} > 1$ weegt de interne kost zwaarder door. De resultaten hiervan zijn weergegeven in Figuur 5.18.

Indien we deze resultaten van naderbij beschouwen merken we enkele zaken op. De methodes geven significant betere resultaten dan alle voorgaande methodes. De resultaten zijn beter naarmate we een hoger gewicht toekennen aan de interne kost. Dit wil zeggen dat sprongen in minimaal kostpad worden onderdrukt. Als optimale verhouding kiezen we voor $w_{int}/w_{ext} = 2$. Voortgaand op de figuur is dit een goede waarde. Een te grote smoothing heeft immers nadelige effecten voor tumoren met een onregelmatige vorm. De resultaten voor de overige validatiematen zijn weergegeven in Bijlage E.1.

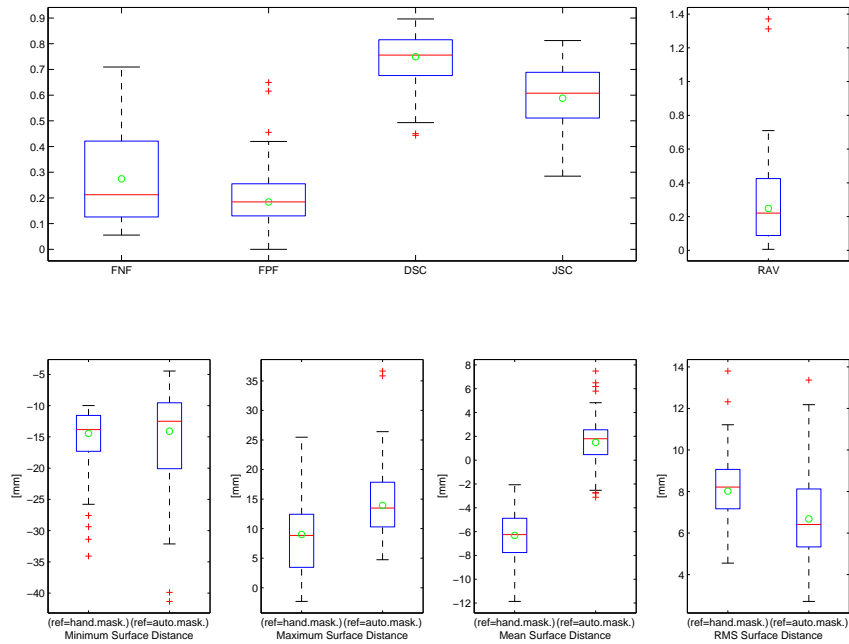
Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.19.

De testdata presteren iets beter dan de trainingsdata. De gemiddelde DSC bedraagt voor beide 75,0%. Voor de mediaan is dit respectievelijk 75,6% en 75,7%. De RAV, RMSSD en MASD vertonen dezelfde trends. De RAV geeft met 24,8% veel betere resultaten dan de tot nu besproken methodes. Bovendien presteert de methode presteert vrij consistent. Er is ook een kleinere

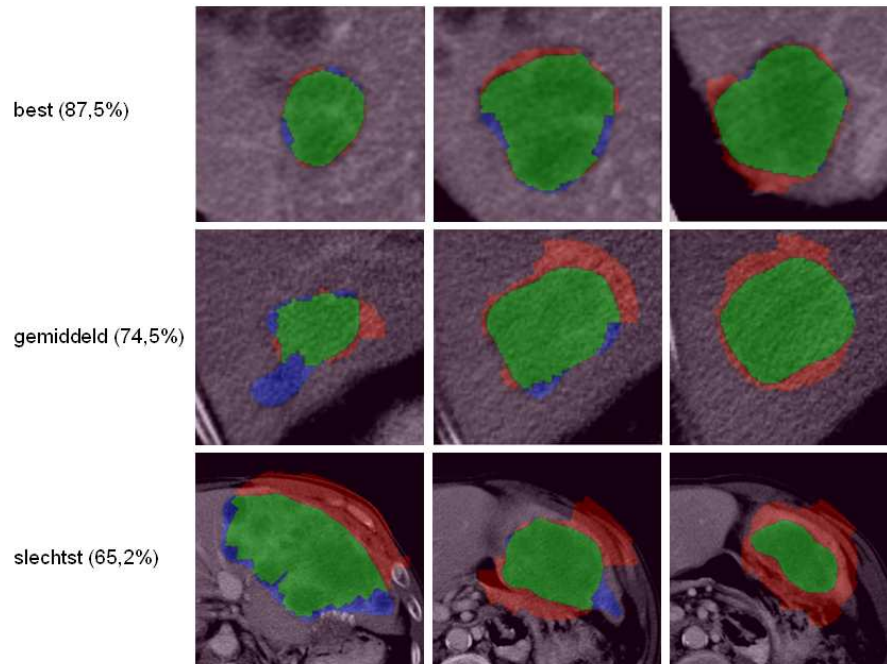


Figuur 5.18: De validatiematen i.f.v. w_{int}/w_{ext} voor de Circle Scanning methode



Figuur 5.19: Resultaten van de 13 validatiematen voor de testdata van Circle Scanning, $w_{int}/w_{ext} = 2$

spreiding op te merken tegenover voorgaande region growing methodes. Dit wijst erop dat de methode betrouwbaarder is. Het is immers belangrijk dat het algoritme voor zoveel tumoren goed werkt. Deze consistentie is misschien voor RAV nog belangrijker. Het belangrijk doel van deze thesis is immers het automatiseren van de volumemetingen van de metastasen. In Figuur 5.20 zijn een aantal beelden getoond voor de beste, gemiddelde en slechtste prestatie van het algoritme.



Figuur 5.20: De beste, gemiddelde en slechtste prestatie van het Circle Scanning algoritme op de testdata

Het is duidelijk te aanschouwen waar zich nog de problemen van het algoritme bevinden. Bij de gemiddelde prestatie zien we dat de methode faalt daar waar zich een grote uitstulping in de tumor bevindt. Dit zou eventueel kunnen verbeterd worden door de interne kost te verhogen t.o.v. de externe kost. Dit zal elders echter slechtere resultaten geven. Mogelijk kan hiervoor dus nog een extra variabele parameter toegevoegd worden. Een ander probleem dat optreedt bij tumoren aan de rand, is dat de gradiënt van de lichaamswand zo groot is dat het algoritme deze wand zal verkiezen boven de eigenlijke rand van de tumor. Dit kan opgelost worden door te vertrekken van een afgelijnde lever.

Besluit

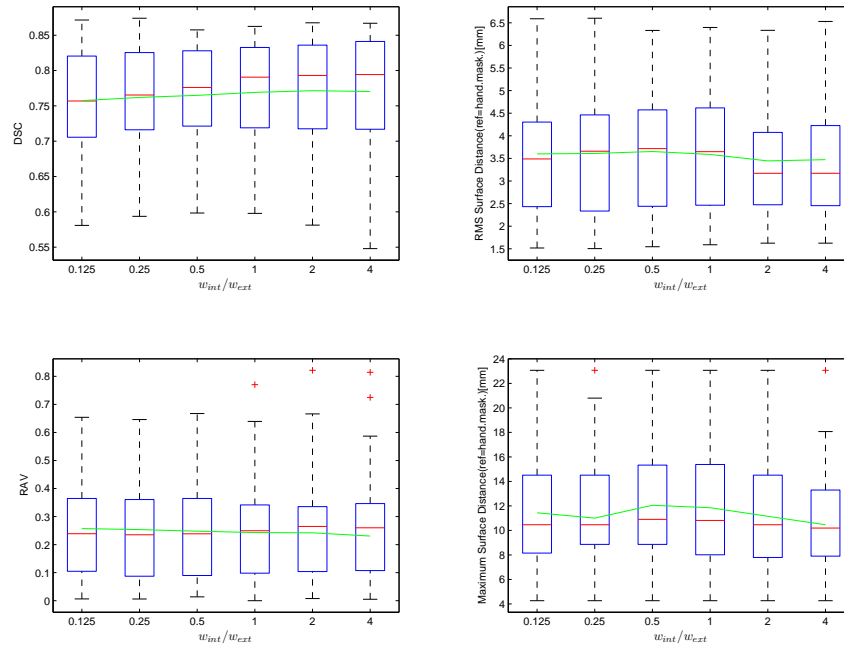
De Circle Scanning methode presteert beduidend beter dan de region growing methodes. De gemiddelde DSC bevindt zich rond 75%. Dit kan verklaart worden door het feit dat deze methode ook gebruik maakt van de gradiënt van het beeld samen met de gradiënt van een binaire interval thresholding segmentatie. De smoothness beperking voorkomt te grote sprongen in het minimaal kostpad waardoor uitgroei efficiënt kan vermeden worden. Daarnaast presteert Circle Scanning ook beter dan Spiral Scanning. Door in 2D te werken, vermijdt het algoritme de problemen die Spiral Scanning heeft met dikke snedes (5 tot 7 mm).

Afgelijnde lever

Trainingsdata

De Circle Scan methode wordt ook getraind ditmaal vertrekkende van een afgelijnde lever. Het is immers mogelijk dat de resultaten van de methode beïnvloed worden door uitgroei buiten de lever.

Voor de verhouding w_{int}/w_{ext} worden dezelfde waarden gekozen. Figuur 5.21 toont de resultaten.



Figuur 5.21: De validatiematen i.f.v. w_{int}/w_{ext} voor de Circle Scanning methode met een afgeleijde lever

Uit de figuur kan afgeleid worden dat zowel voor de DSC als voor de RMSSD $w_{int}/w_{ext}=2$ de optimale waarde is. De RAV en de MASD geven 4 als optimale waarde.

Testdata

De resultaten voor de testdata zijn weergegeven in Figuur 5.22.

De gemiddelde DSC voor de testdata is 75,9%. De mediaan is 78,8%. Voor de trainingsdata was dit respectievelijk 77,1% en 79,3%. De testdata presteren dus bijna even goed als de trainingsdata. Wanneer we vergelijken met de resultaten van Circle Scanning op het volledige beeld zien we dat de methode iets beter presteert op een afgeleijde lever. De methode gaat steeds op zoek naar de sterkste gradiënt. Indien de tumor aan de rand van de lever is gelegen en hij voldoende groot is, kan het zijn dat de rand van het lichaam als sterkste gradiënt wordt aanzien. Bij een gesegmenteerde lever wordt dit probleem vermeden. Er is echter ook een nadeel. Hetzelfde effect kan er immers voor zorgen dat voor tumoren die niet aan de rand gelegen zijn, het algoritme nu de leverwand als rand van de segmentatie gaat kiezen.

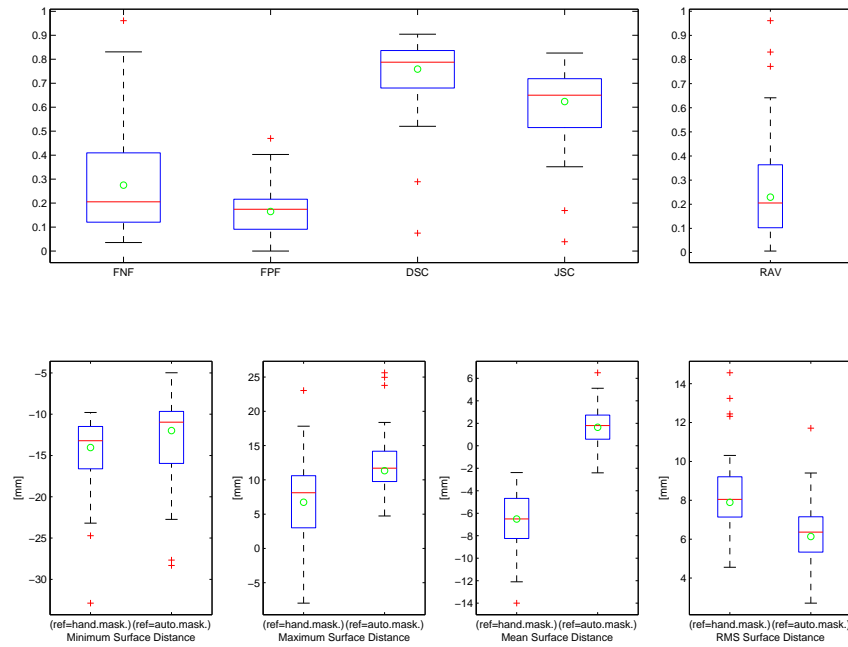
5.3.4 Level Sets met Circle Scan

Ook voor deze Level Sets met Circle Scan wordt het algoritme enerzijds toegepast op het volledige beeld en anderzijds op een gesegmenteerde lever.

Volledig Beeld

Trainingsdata

Deze methode bevat verschillende parameters waarvoor een optimale waarde moet gezocht worden. De gehanteerde strategie is de volgende. Op basis van snelle, voornamelijk kwalitatieve testen, worden de meest invloedrijke parameters bepaald. Er zijn immers parameters die voor



Figuur 5.22: Resultaten van de 13 validatiematen voor de testdata van Circle Scanning op een afgelijnde lever, $w_{int}/w_{ext} = 2$

deze toepassing weinig invloed hebben op het resultaat. De belangrijkste parameters zijn de wegingsfactoren van de advectieterm, de krommingsterm en propagatieterm in vergelijking 4.11 en de wegingsfactor van de interne t.o.v. de externe kost (w_{int}/w_{ext}) in het Circle Scanning Technique-algoritme dat voor de primaire segmentatie zorgt. Deze laatste factor is echter al afgesteld in 5.3.3: $w_{int}/w_{ext} = 2$. Vervolgens wordt een eerste parameter geoptimaliseerd, waarbij de overige parameters constant worden gehouden. In dit geval start de optimalisatie met de wegingsfactor van advectieterm. De wegingsfactoren voor de propagatieterm en de krommingsterm worden gelijkgesteld aan resp. 0,01 en 0,1. De wegingsfactor loopt van 10^{-2} tot 10^2 , waarbij de macht van 10 elke stap met 1 verhoogd wordt. Hieruit blijkt dat de optimale waarde voor de advectieterm gelijk is aan 0,1. Deze waarde wordt vastgelegd. De volgende stap is om de wegingsfactor voor de krommingsterm te optimaliseren. Deze is best 0,1. De laatste invloedrijke term waarvoor nog een optimale waarde moet gevonden worden is de propagatieterm. Deze heeft geen uitgesproken optimum, vandaar dat we de huidige waarde van 0,01 behouden. De figuren voor de optimalisatie van de wegingsfactor van de advectie-, krommings- en propagatieterm zijn terug te vinden in Bijlage E.

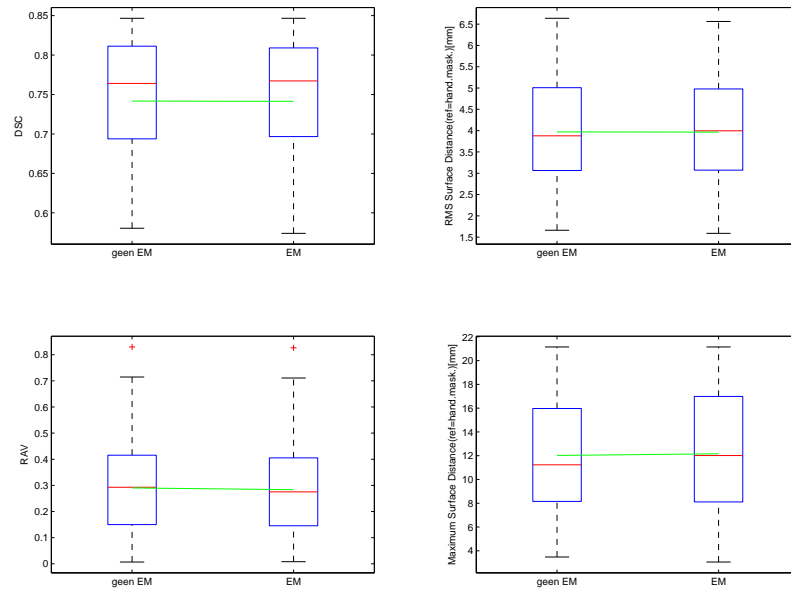
Figuur 5.23 toont de resultaten van de trainingsdata voor de geoptimaliseerde wegingsfactoren. Daarnaast zijn ook de resultaten voor het level sets-algoritme dat gebruik maakt van gesuperviseerde pixelclassificatie gebaseerd op het EM-algoritme. De gekozen parameters zijn voor beide dezelfde (wegingsfactor van de advectie-, krommings- en propagatieterm zijn resp. 0,01, 0,1 en 0,01).

De resultaten voor niet gesuperviseerde pixelclassificatie zijn nagenoeg hetzelfde als voor gesuperviseerde hetgeen duidelijk is in de DSC van respectievelijk 74,2% en 74,1%.

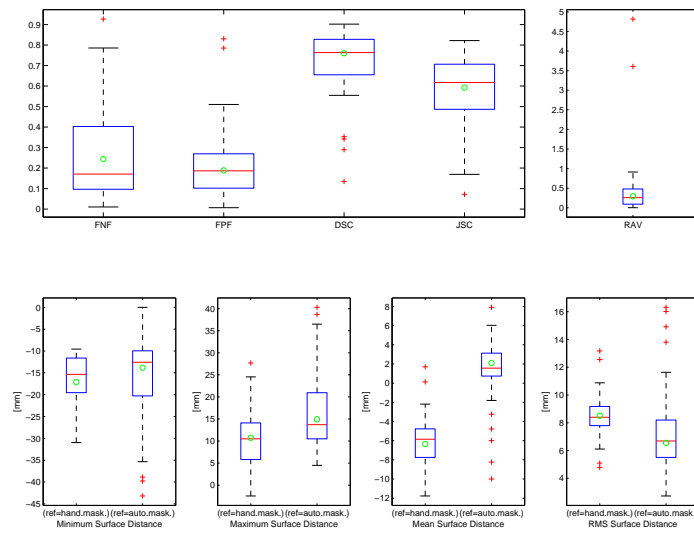
Testdata

De resultaten voor de testdata voor de Level Sets met Circle Scan methode, gebruik makend van gesuperviseerde pixelclassificatie zijn gegeven in Figuur 5.24. Diegene voor Level Sets met Circle Scan methode zonder gesuperviseerde pixelclassificatie in Figuur 5.25.

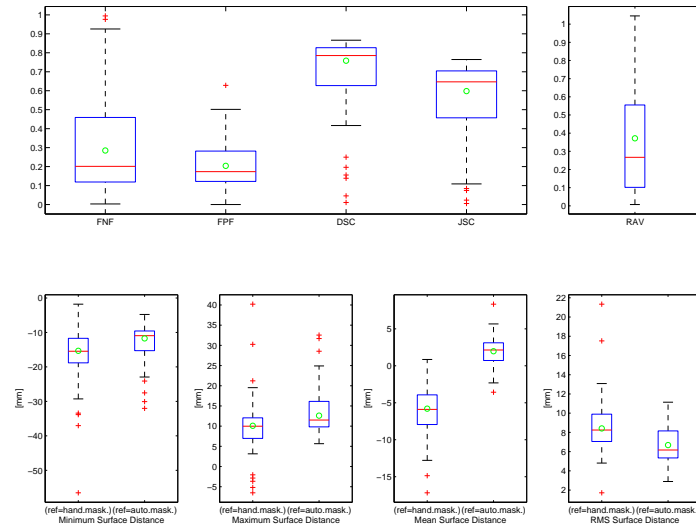
De gemiddelde DSC voor de testdata is 76,5% en 76,0% resp. zonder en met supervised learning. De mediaan is 77,7% en 76,4%. Deze waarden zijn beter dan voor de trainingsdata. Dit wijst erop



Figuur 5.23: De validatiematen voor een level set methode met en zonder supervised learning



Figuur 5.24: Resultaten van de 13 validatiematen voor de testdata van Level Sets met supervised learning



Figuur 5.25: Resultaten van de 13 validatiematen voor de testdata van Level Sets zonder supervised learning

dat dit algoritme beter overweg kan met nieuwe data. Voor de level set-methode zonder supervised learning is dit nog meer waar, aangezien het zichzelf leert voor nieuwe data. In Figuur 5.26 zijn de resultaten weergegeven voor de beste, gemiddelde en slechtste prestatie van het Level Sets algoritme met EM.

In de figuur zijn een aantal redenen terug te vinden waarom het algoritme soms minder goed presteert. Bij de gemiddelde prestatie is bijvoorbeeld te zien dat het algoritme faalt wanneer één metastase is afgelijnd met meerdere aflijningen. Dit is niet het gevolg van de Level Set maar van de initialisatie met Circle Scanning. Dit laatste algoritme is niet geïmplementeerd voor meerdere seeds per slice. Een andere fout die optreedt, is wanneer de tumor een kern heeft met een hogere grijswaardenintensiteit. Dit veroorzaakt een hoge gradiënt die bovendien versterkt is doordat er zich aan de ander kant van de rand intensiteiten bevinden die typisch zijn voor metastasen. Hierdoor is deze gradiënt groter dan de gradiënt aan de werkelijke rand van de metastase. Hierdoor kan de Level Set ook niet verder propageren.

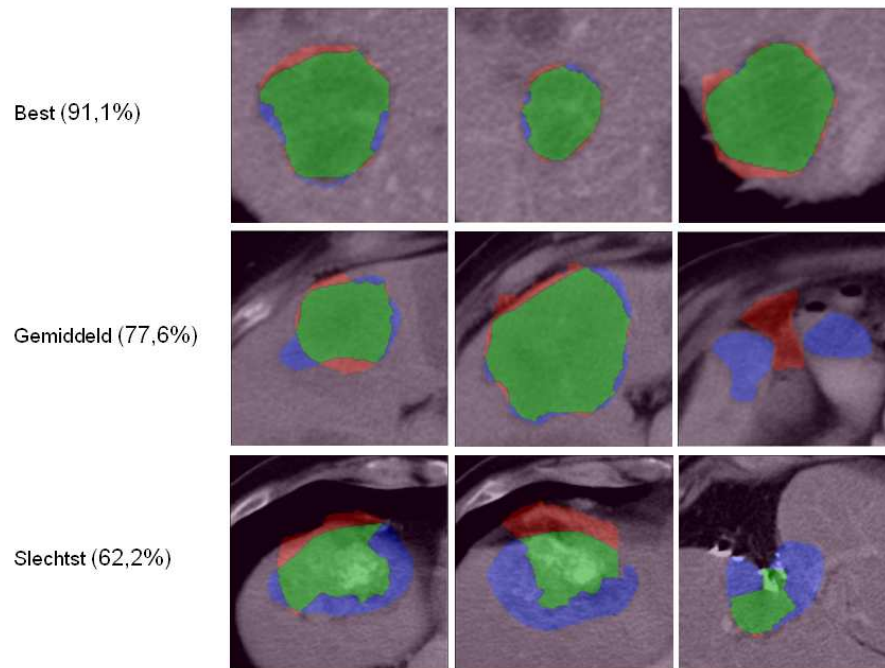
Besluit

Het afstellen van de parameters voor deze methode is belangrijk voor het resultaat. Dit kan gezien worden als een nadeel aangezien het afstellen van de parameters zeer tijdrovend is. Bij goed afgestelde parameters geeft het algoritme bevredigende resultaten. Voor nieuwe data zorgt het EM-algoritme voor een verbeterd statistisch model wat een betere segmentatie oplevert.

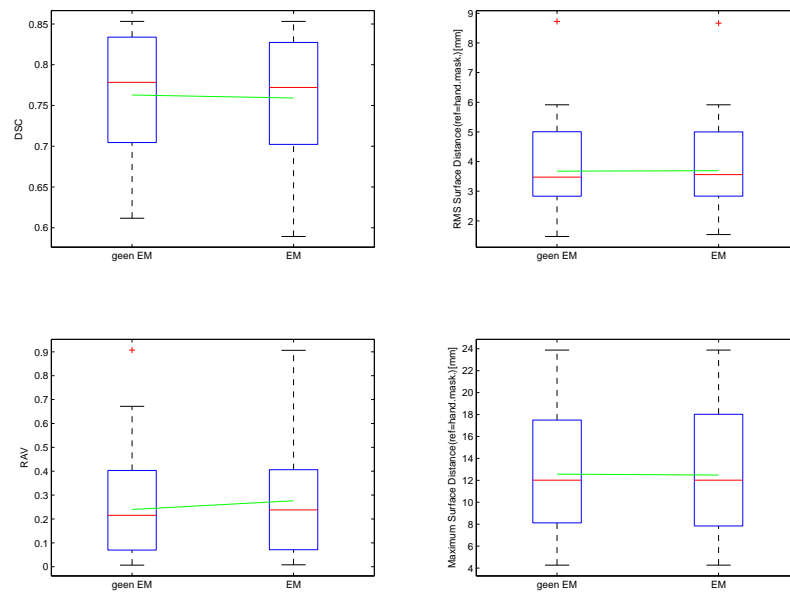
Afgelijnde lever

Trainingsdata

Het afstellen van de wegingsfactoren gebeurt op dezelfde manier als hierboven. De wegingsfactoren voor de advectieterm, de krommingsterm en de propagatie worden geoptimaliseerd en geven als beste waarden resp. 10^{-2} , 10^{-1} en 10^{-2} . Deze waarden worden ook toegepast voor het level set-algoritme dat gebruik maakt van het EM-algoritme. De resultaten van de trainingsdata zijn weergegeven in Figuur 5.27. Het algoritme toepassen op een afgelijnde lever, levert een DSC op die gemiddeld $\pm 2\%$ beter is dan voor een niet afgelijnde lever. De getrainde wegingsfactoren worden gebruikt om de testdata te valideren.



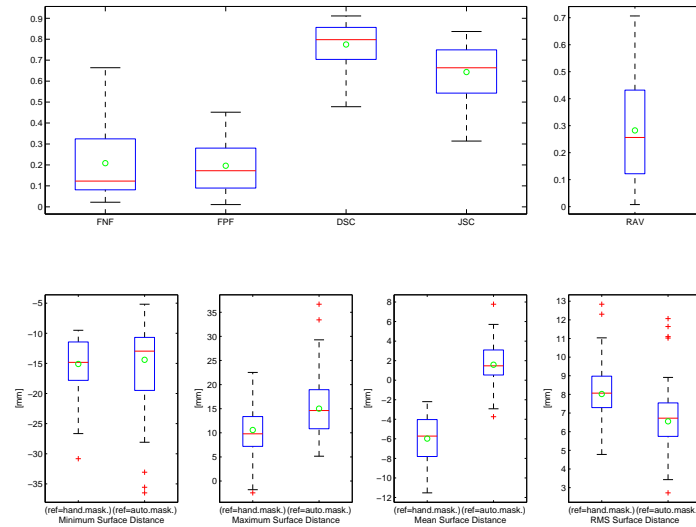
Figuur 5.26: De beste, gemiddelde en slechtste prestatie van het Levels Sets met EM algoritme op de testdata



Figuur 5.27: De validatiematen i.f.v. voor de trainingsdata met en zonder supervised learning (afgelijnde lever)

Testdata

De resultaten voor de testdata voor de Level Sets met Circle Scan methode, gebruik makend van gesuperviseerde pixelclassificatie vertrekkende van een afgelijnde lever zijn gegeven in Figuur 5.28. Diegene voor Level Sets met Circle Scan methode zonder gesuperviseerde pixelclassificatie in Figuur 5.29.



Figuur 5.28: Resultaten van de validatiematen voor de testdata van Level Sets met supervised learning toegepast op een afgelijnde lever

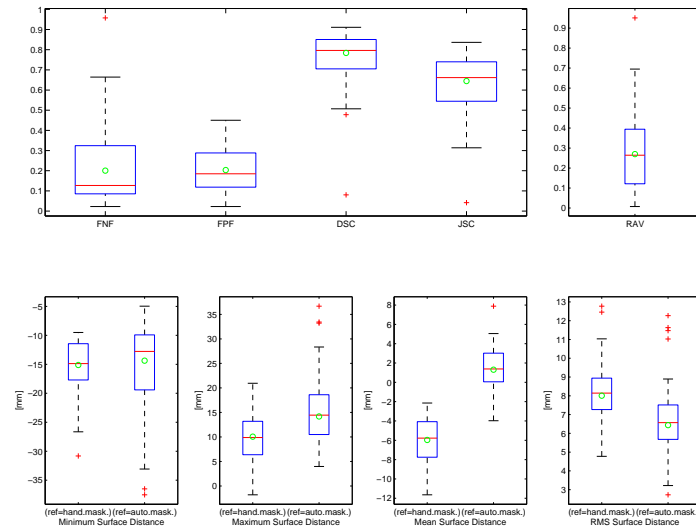
We zien gelijkaardige trends als in de situatie van een niet afgelijnde lever. Opnieuw zijn de resultaten beter dan de trainingsdata: 77,5% voor de testdata t.o.v. 76,3% voor de trainingsdata voor het algoritme met gesuperviseerde pixelclassificatie en 78,4% t.o.v. 75,39% voor unsupervised learning. Het verschil is in het laatste geval dus ook weer groter, wat nogmaals bewijst dat het EM-algoritme het statistisch model specifiek traint.

5.4 Kritische evaluatie van de resultaten

Om de kwaliteit van de resultaten beter in te schatten, is het belangrijk de resultaten te vergelijken enerzijds met gelijkaardige, in de literatuur verschenen studies en anderzijds met de inter-observervariabiliteit. Alvorens een kritische kijk op de resultaten te gooien, wordt een overzicht van de resultaten gegeven. Dit maakt het vergelijken van de methodes onderling eenvoudiger.

5.4.1 Overzicht van de resultaten

Tabel 5.1 geeft de resultaten voor een niet afgelijnde lever weer. In de tabel is voor elke maatstaf de beste waarde voor test- en trainingsdataset vetgedrukt.



Figuur 5.29: Resultaten van de validatiematen voor de testdata van Level Sets zonder supervised learning toegepast op een afgelijnde lever

Tabel 5.1: Resultaten van alle methodes

		DSC[%]		RAV[%]		RMSSD[mm]		MASD[mm]	
		\bar{x}	\tilde{x}	\bar{x}	\tilde{x}	\bar{x}	\tilde{x}	\bar{x}	\tilde{x}
Region Growing seed per snede	training	41,8	47,5	140,6	98,0	4,3	3,6	10,9	10,6
	test	40,0	36,1	249,3	171,0	8,2	8,9	8,1	8,3
Region Growing twee seeds	training	49,3	43,8	99,1	81,4	4,7	3,9	10,7	10,7
	test	39,5	39,6	225,5	218,1	8,0	8,4	8,7	8,8
Region Growing één seed (3D)	training	25,7	25,0	1077	164,2	4,8	4,3	12,1	12,2
	test	22,8	11,0	96,0	99,5	12,8	10,5	13,6	12,4
Statistische Region Growing (○)	training	66,8	71,6	29,6	31,0	5,8	4,4	10,4	11,5
	test	59,3	58,7	47,1	44,3	9,4	9,1	11,1	10,4
Statistische Region Growing (□)	training	71,5	72,0	29,4	29,5	4,7	4,2	10,4	10,7
	test	58,4	54,0	90,9	71,0	8,7	8,9	10,0	10,2
PixelClass. met Reg. Growing	training	46,6	50,4	68,9	57,6	5,2	4,4	11,6	13,3
	test	48,0	45,5	93,1	43,0	10,4	8,8	12,4	10,5
Circle Scanning	training	75,0	75,7	26,4	25,9	3,6	3,4	10,9	9,8
	test	75,0	75,6	24,8	22,0	8,0	8,2	9,0	8,9
Spiral Scanning	training	52,0	58,8	63,3	51,6	4,2	4,7	14,1	12,7
	test	54,5	65,0	56,1	49,2	7,6	7,8	9,0	9,7
Levelsets met Circle Scan	training	74,2	76,4	29,0	29,3	4,0	3,9	12,0	11,2
	test	76,0	76,4	29,5	26,0	8,5	8,4	10,7	10,5
Levelsets met Circle Scan (EM)	training	74,1	76,7	28,3	27,5	4,0	4,0	12,2	12,0
	test	76,5	77,7	27,3	22,6	8,5	8,4	10,8	10,4

De resultaten zullen kort even op een rijtje gezet worden in stijgende volgorde wat betreft performantie. De methode die duidelijk de minste resultaten oplevert, is region growing met één zaadpunt omdat toepassen van deze methode vaak leidt tot ofwel felle uitgroei ofwel een veel te kleine segmentatie. Daarna volgen de region growing-methodes met één seed per snede of twee seeds. Het gebruik van pixelclassificatie zorgt voor een verbetering. Daarna volgt de Spiral Scanning Technique, die een aanzienlijke verbetering t.o.v. de voorgaande algoritmes. Vandaar dat

we kunnen stellen dat de voorgaande methodes eerder primitief zijn terwijl de andere methodes eerder geavanceerd is. De statistische region growing-methodes zijn wat betreft performantie gelijkaardig aan Spiral Scanning Technique. De duidelijk twee beste methodes zijn Circle Scanning Technique en Level Sets met Circle Scan.

Opvallend aan deze ordening is dat de eerder primitieve methodes geen rekening houdt met geometrische eigenschappen terwijl de meer geavanceerde methodes dit wel doen hetzij globaal (statistische region growing), hetzij lokaal (Spiral en Circle Scanning Technique, Level Sets).

Het overzicht van de resultaten voor een afgelijnde lever is weergegeven in Tabel 5.2.

Tabel 5.2: Resultaten van enkele methodes voor een afgelijnde lever

		DSC[%]		RAV[%]		RMSSD[mm]		MASD[mm]	
		\bar{x}	\tilde{x}	\bar{x}	\tilde{x}	\bar{x}	\tilde{x}	\bar{x}	\tilde{x}
Circle Scanning	training	77,1	79,3	24,2	26,5	3,4	3,2	11,1	10,5
	test	75,9	78,8	23,0	20,5	7,9	8,0	6,7	8,1
Spiral Scanning	training	59,7	71,7	45,5	35,9	3,4	3,7	11,3	10,6
	test	63,2	67,2	60,4	49,2	7,7	7,9	8,8	9,7
Levelsets met Circle Scan	training	76,3	77,8	24,0	21,5	3,7	3,5	12,6	12,0
	test	77,5	79,8	28,2	25,6	8,0	8,1	10,6	9,8
Levelsets met Circle Scan (EM)	training	75,9	77,2	27,6	23,8	3,7	3,6	12,5	12,0
	test	78,4	79,7	27,0	26,4	8,0	8,1	10,1	9,9

Een aantal andere validatiematen in tabelvorm zijn terug te vinden in Bijlage E.

5.4.2 De resultaten in relatie tot andere studies

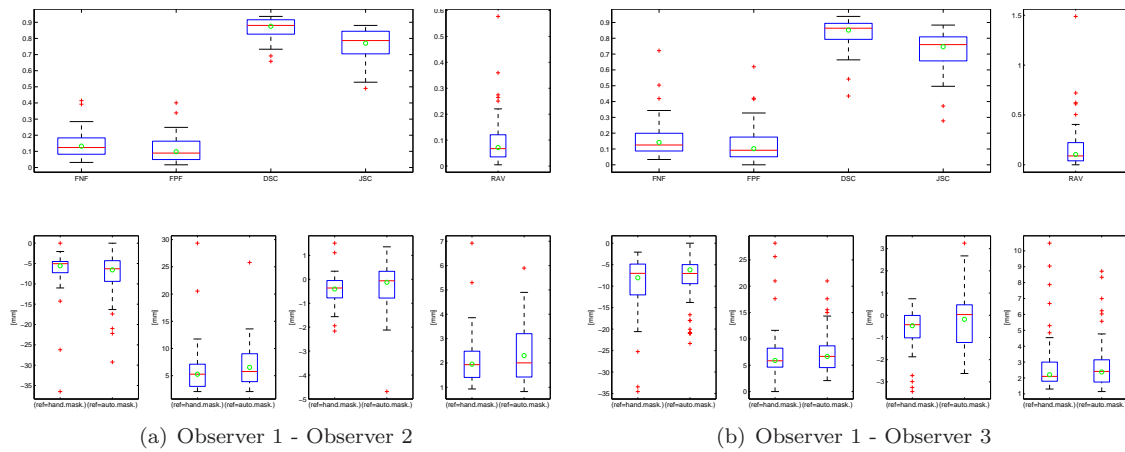
De resultaten vergelijken met in de literatuur gepubliceerde resultaten moet met enige voorzichtigheid gebeuren aangezien de invloed van de verschillende protocols, CT-scanners, snedediktes, hoeveelheden contraststof, patiënten,... moeilijk in kaart te brengen is.

Een eerste studie is [1] waar een meta-analyse besproken wordt over de detectie en segmentatie van colorectale levermetastasen. Voor de segmentatie van metastasen in 39 CT datasets is de gemiddelde sensitiviteit 63,8%. Deze datasets zijn afkomstig uit verschillende studies gebruik makend van een verschillende methodes en verschillende kwaliteiten van beelden. De subcategorie van datasets waarbij minder dan 45 g jodium wordt ingespoten zoals ook de gebruikte data in deze thesis, is de sensitiviteit 61,4%.

De beste resultaten die we in literatuur konden vinden, komen uit [47] waar uitgegaan wordt van een gesegmenteerde lever. De studie werkt met recente data met een snededikte van 2 tot 3 mm pixels van 0,55 tot 0,88 mm² slice. De bekomen DSC was 88,9±4,6%, de FNF 7,2±2,9% en de FPF 12,6±4,2%. De maximale gemiddelde DSC in een afgelijnde lever in deze thesis is 78,4%. Dit minder presteren is deels te wijten aan de data die we gebruikten.

5.4.3 Resultaten in relatie tot de inter-observatorvariabiliteit

Omdat het idee dat wij hebben over een metastase enigszins afwijkt van hoe een metastase beschouwd wordt door een arts, is een studie uitgevoerd naar de inter-observatorvariabiliteit. Ook willen we deze inter-observatorvariabiliteit vergelijken met het presteren van de methodes. Op deze manier kunnen ook de bekomen resultaten van de automatische algoritmes beter gesitueerd worden. Voor de studie zijn de 72 tumoren van de testdata manueel afgelijnd. Observator 1 is een oncoloog. Observator 2 en 3 zijn twee studenten die enige affiniteit hebben met abdominale CT-scans. In Figuur 5.30 vergelijkt men de aflijningen van observator 1 met die van observator 2 en 3.



Figuur 5.30: De resultaten voor de 13 validatiematen van observer 1 vergeleken met observer 2(a) en observer 3(b)

De figuur toont de 13 validatiematen om 2 aflijningen te vergelijken. De gemiddelde DSC wanneer men 1 met 2 vergelijkt is 87,6%, de mediaan is 88,1%. Voor observator 1 versus 3 is dit respectievelijk 85,2% en 86,4%. Dit toont aan dat er nog een aanzienlijk verschil is in de manier van manueel aflijnen. Daarnaast is er ook nog een verschil wanneer de aflijningen van observer 2 en 3 met elkaar vergeleken worden. Dit levert een gemiddelde DSC op van 87,9% en een mediaan van 88,1%. Wanneer we naar de resultaten van de automatische methodes kijken, dient deze inter-observatorvariabiliteit in rekening te worden gebracht. Daarnaast is er ook nog de intra-observatorvariabiliteit. Deze wordt sterk gereduceerd door over te stappen naar een (semi-)automatische methode.

Een vergelijking werd gemaakt met de studie uit [12]. Dit is terug te vinden in Bijlage E.2. Hieruit blijkt eveneens dat de manier waarop observator 1 aflijnt enigszins verschilt van hoe observator 2 en 3 aflijnen. Dit verschil is zoals verwacht groter dan in de literatuur.

Hoofdstuk 6

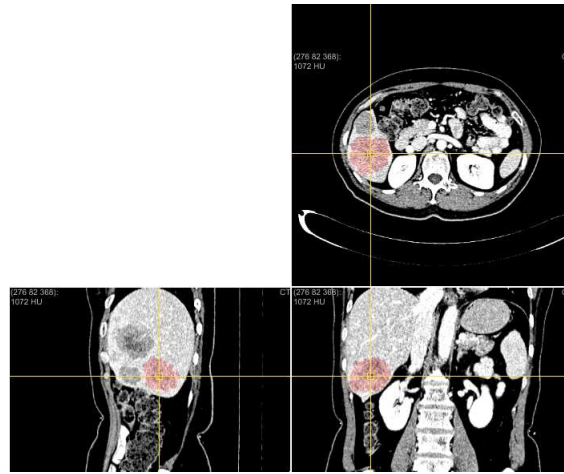
Visualisatie van metastasen

In de medische technologie bestaat er nog altijd een groot verschil tussen de werkwijze van artsen en ingenieurs. Waar artsen vooral geïnteresseerd zijn in kwalitatieve verschillen, zullen ingenieurs steevast werken met kwantitatieve verschillen. Met behulp van visualisatie probeert men deze twee werelden te verzoenen. Visualisatie in zijn meest algemene vorm houdt in dat men een gedachte probeert te vertalen naar een beeld. In dit hoofdstuk zal enerzijds 2D visualisatie en 3D visualisatie kort besproken worden. Daarnaast zal ingegaan worden op de extra kwantitatieve informatie die beeldverwerkingsprogramma's kunnen meegeven aan de clinicus wanneer hij de tumoren bekijkt.

6.1 2D Weergave

Beelden worden traditioneel weergegeven door 2D axiale snedes. Vroeger gebeurde dit op film, tegenwoordig zijn al deze beelden digitaal opgeslagen. Dit geeft de mogelijkheid om te scrollen doorheen alle snedes en op die manier eveneens een ruimtelijke impressie te bekomen. Alles kan weergegeven worden op één scherm wat de nood aan een lichtbak die veel plaats inneemt, reduceert. Uit de standaard viewer kan heel wat informatie gehaald worden. Zo zijn de grijswaardenintensiteit (in HU), patiëntnaam, beeldmodaliteit en andere informatie aangegeven. In sommige gevallen biedt een weergave van eveneens het sagittale en het coronale vlak een meerwaarde. Bijvoorbeeld wanneer men snel het centrum van een tumor wil lokaliseren in het beeld. Hiervoor kan men gebruik maken van een orthoviewer. De orthoviewer zal op basis van het volumebeeld sagittale, coronale en axiale snedes reconstrueren. Met behulp van Multi Planar Reconstruction (MPR) kunnen snedes doorheen het volumebeeld getrokken worden. Dit kan in elke mogelijk richting gebeuren maar meestal kiest men de 3 orthogonale richtingen.

De orthoviewer kan gebruikt worden om een gesegmenteerde metastase weer te geven. Dit gebeurt door het masker van de segmentatie eveneens op te delen in orthogonale snedes en vervolgens bovenop de beelden in de orthoviewer te plaatsen. Op deze manier kan de clinicus in de verschillende vlakken de tumor onderzoeken. Een voorbeeld hiervan is weergegeven in Figuur 6.1.

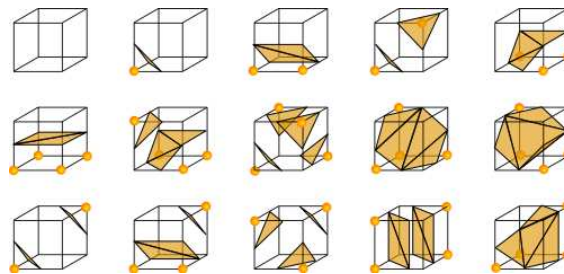


Figuur 6.1: Orthoviewer met een weergave van de metastase in axiaal(rb), sagittaal(lo) en coronaal(ro) vlak

6.2 3D Weergave

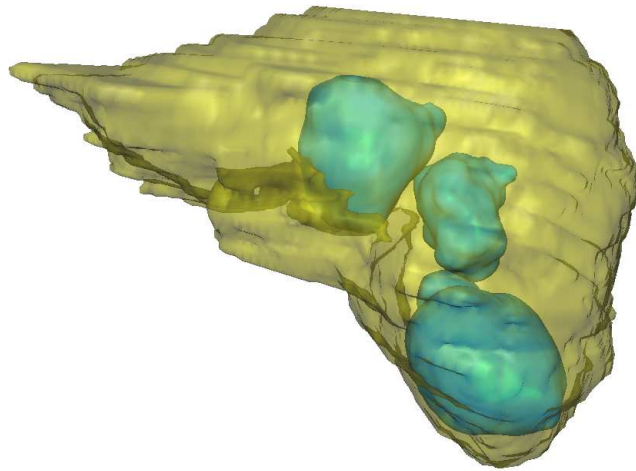
Indien een operatie dient uitgevoerd te worden om primaire tumoren of metastasen in de lever te verwijderen, kan het soms nuttig zijn om naast een 2D ook een 3D weergave van de lever te bekomen. Dit maakt het mogelijk voor de chirurg om de operatie optimaal te plannen. Aan de hand van een 3D beeld kan de chirurg de ideale plaats van incisie bepalen om geen belangrijke bloedvaten te raken en zo weinig mogelijk gezond leverweefsel te beschadigen [48].

Om een 3D beeld te renderen wordt vertrokken van een oppervlak (*mesh*). Een mesh is een verzameling van vertices en vlakken die de vorm van een object in 3D bepalen. Er bestaan twee manieren om 3D medische beeldinformatie te renderen: surface rendering en volume rendering. Bij surface rendering wordt op basis van een binaire segmentatie de verschillende oppervlakken gegenereerd. In medische beeldverwerking stemmen deze oppervlakken overeen met structuren eigen aan het menselijk lichaam (organen, bot, tumoren). De kern bij surface rendering is een goede voorstelling van deze oppervlakken. Dit bekomt men door het *Marching Cubes* algoritme [49] toe te passen op het object. Hierbij gaat men het 3D object opdelen in een raster. Vervolgens wordt voor iedere cel bepaald hoeveel driehoekjes in iedere kubus moeten getekend worden. De 15 verschillende mogelijkheden zijn weergegeven in Figuur 6.2.



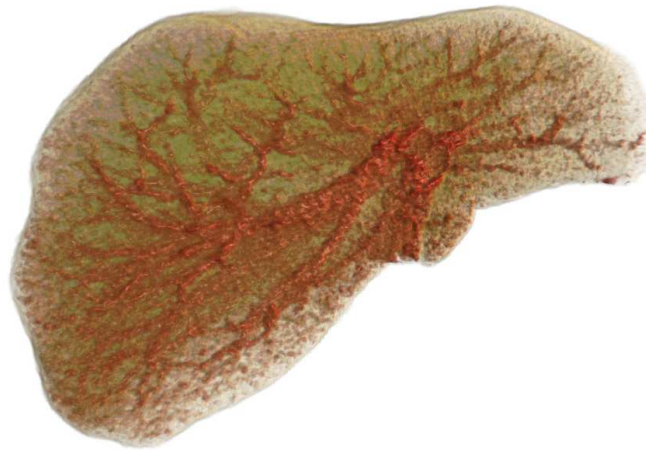
Figuur 6.2: 15 unieke situaties die zich voordoen bij *Marching Cubes*

Surface rendering heeft als voordeel dat de geconstrueerde meshes verder bewerkt kunnen worden. Nadeel is dat enkel de oppervlakken van de objecten gevisualiseerd kunnen worden. Dit maakt het geheel visueel iets minder aantrekkelijk. Bovendien is een harde segmentatie vereist. Het object dient gecreëerd te worden en dit vereist extra tijd. Een voorbeeld van surface rendering is weergegeven in Figuur 6.3.



Figuur 6.3: Surface Rendering van de lever met zijn metastasen

Bij volume rendering kiest men een andere benadering. Men kent aan iedere voxel een kleur en een opaciteit toe. Dit gebeurt met een Rood Groen Blauw Alfa (RGBA) transfer functie die in functie van de intensiteit de RGBA-waarde bepaalt in iedere voxel. Vervolgens wordt volume ray casting toegepast om iedere pixel te visualiseren. Dit houdt in dat lichtstralen door het volume worden gestuurd vanuit het camera standpunt waar men kijkt. Afhankelijk van de kleur en de opaciteit zal iedere voxel op een andere manier oplichten. De totale kleur die zichtbaar is in het beeld, wordt bepaald door de som van alle voxels op die lijn. Een ruwe segmentatie kan dienst doen om de volume rendering te versnellen. Zo kunnen anatomische structuren die in de weg zitten verwijderd worden. Verder kan men ook spelen met de radiositeit (ambient, diffuus en/of speculaire belichting). Een voorbeeld van volume rendering is weergegeven in Figuur 6.4



Figuur 6.4: Volume Rendering van de lever met bloedvaten

Volume rendering genereert een ander soort beeld. Dankzij de classificatie op basis van intensiteiten of een ander kenmerk krijgen we nu wel de informatie in iedere voxel van het beeld. We zien nu duidelijk alle weefsels zonder dat er nood is voor een harde segmentatie. Volume rendering vereist ook minder rekenkracht (geen mesh genereren) en laat een real time rendering toe (indien

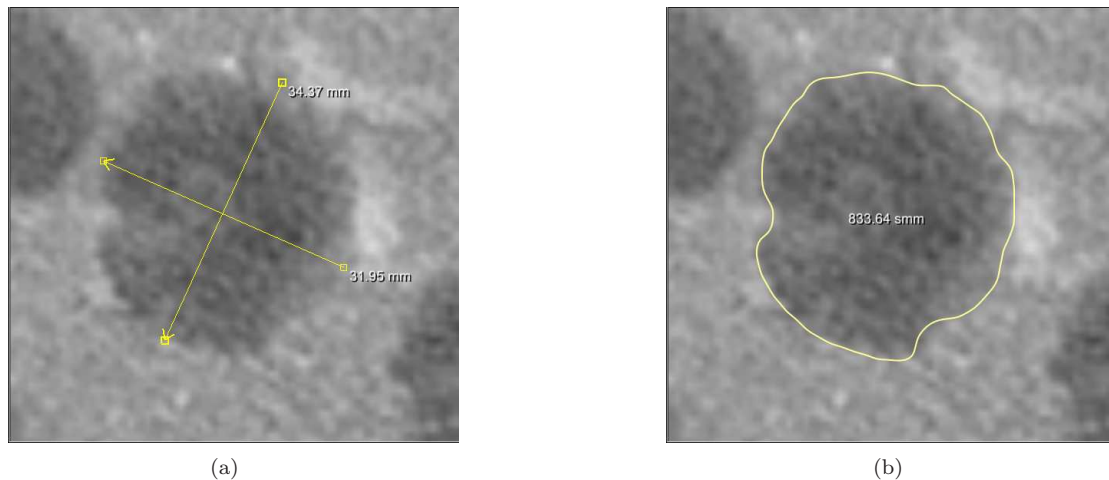
men beschikt over een goede Graphical Processing Unit (GPU)). Het is echter voornamelijk een visueel hulpmiddel. Het is veel moeilijker een volume gerendered beeld verder te bewerken (real time cutting en segmentatie).

6.3 Weergave van kwantiteiten

Vaak is het handig voor de clinicus om over een aantal eenvoudige meetinstrumenten te beschikken. Op die manier kan snel en efficiënt een meting uitgevoerd worden en krijgt men onmiddellijk het resultaat. In deze sectie worden enkele voorbeelden aangehaald van dergelijke metingen.

Een eerste voorbeeld is de meting van RECIST op de metastasen van de lever. Een beeld hiervan is weergegeven in Figuur 6.5(a).

De RECIST en zijn orthogonale worden aangeduid op de figuur en vervolgens wordt de vector-lengte bepaald en mee weergegeven op het beeld. Het grote voordeel hierbij is dat deze meting automatisch de omrekening doet van voxels naar millimeters. Zo kunnen ook beelden met een verschillende resolutie eenvoudig met elkaar vergeleken worden. Een ander voorbeeld is het oppervlak bepalen van een handmatige aflijning. Dit is weergegeven in Figuur 6.5(b).



Figuur 6.5: Meting van de RECIST en zijn loodrechte met weergave van de afmetingen **(a)** en oppervlak van een aflijning weergegeven op het beeld **(b)**

In de figuur is de aflijning van een metastase weergegeven. Tegelijk met deze aflijning wordt ook het oppervlak van de aflijning meegegeven. Zo kan makkelijk nagegaan worden of een metastase in een bepaalde snede gegroeid of gekrompen is.

6.4 Besluit

Visualisatie is een belangrijk concept binnen de medische beeldvorming. Eenvoudige en complexe technieken kunnen het gebruikersgemak voor de klinici aanzienlijk verhogen. 2D weergaven blijven de belangrijkste bron van beeldinformatie voor artsen. Fijne details zijn enkel in 2D zichtbaar. Het is daarom belangrijk om ook voor deze beelden het gebruiksgemak te blijven verbeteren. 3D beelden worden op dit ogenblik nog niet standaard gebruikt, maar deze evolutie kan niet uitblijven. Surface rendering is eenvoudig maar beperkt doordat geen informatie over de inhoud van de anatomische structuren gekend is. Volume rendering heeft meer mogelijkheden op het vlak van visualisatie. De weergave van kwantiteiten ten slotte is een belangrijk aspect. Gebruikers kunnen snel en eenvoudig enkele metingen doen en krijgen onmiddellijk het resultaat.

Hoofdstuk 7

Algemeen besluit

In dit eindwerk “De lever: 3D visualisatie en kwantificatie van de metastasen” staat het kwantificeren van de metastasen centraal. Kwantificatie vereist een segmentatie, bij voorkeur zo betrouwbaar mogelijk en met een minimum aan gebruikersinteractie. Vandaar dat verschillende segmentatiemethodes uitgewerkt werden.

7.1 Methodes

De segmentatiemethodes kunnen onderverdeeld worden in twee categorieën: de automatische en de semi-automatische methodes.

De automatische methodes vereisen een afgelijnde lever en kunnen de metastasen niet afzonderlijk segmenteren. De meest primitieve is *interval thresholding* waar eenvoudigweg twee drempelwaardes bepalen of een voxel tot een metastase behoort of niet. De optimale ondergrens is 1015, de meer kritische optimale bovengrens 1115. Dit blijkt zowel uit de validatie van interval thresholding als uit het uitgevoerde onderzoek naar de verdeling van de grijswaarden. Een meer geavanceerde techniek voor het automatisch segmenteren van de metastasen is *Pixel Classification*. Aan elke voxel wordt een kans toegekend dat deze voxel tot een metastase behoort ofwel op basis van een door een trainingsfase onderbouwd statistische model ofwel op basis van het voor een bepaalde dataset meest waarschijnlijke statistische model. Dit meest waarschijnlijke model wordt berekend met een EM-algoritme. Het pixelclassificatie-algoritme maakt gebruik van een model voor de verdeling van de grijswaarden eventueel aangevuld met Haralick textuurkenmerken (variantie en correlatie). Omwille van computationele efficiëntie, is de berekening van de textuurkenmerken beperkt tot kleine afstanden. Voor die afstanden bieden de textuurkenmerken geen meerwaarde voor het statistisch model. Uit de analyse van de grijswaarden blijkt een normale verdeling met $\{\mu_j, \sigma_j\}$ gegeven in Tabel 4.1 als model een goede benadering, al is ook een logistische verdeling mogelijk. Zoals te verwachten geeft statistische pixelclassificatie betere resultaten dan interval thresholding.

Naast de automatische methodes zijn er de semi-automatische methodes die vaak gebruik maken van één van de vorige methodes. De eerste groep zijn de *Region Growing* methodes. Ze combineren een interval thresholding met een connectiviteitsvereiste. Aangezien de gebruikte data een lage z-resolutie hebben en aangezien er in 3D meer kans is tot uitgroei, levert een 2D Region Growing operatie betere resultaten op. Toevoegen van lokaal statistisch model verbetert de segmentie aanzienlijk. Het algoritme berekent de verdeling van de grijswaarden in een ruitvormig of een cirkelvormig masker en bepaalt op basis daarvan het interval waarbinnen de grijswaarde van de voxel moet liggen om bij de regio toegevoegd te worden. Dit optimale interval legt een bovengrens op van $\bar{I}_k + 1,5s$. Pixelclassificatie kan de metastasen niet afzonderlijk segmenteren, maar door het fuzzy pixelclassificatie-algoritme te combineren met Region Growing kan dit wel. Dit levert een verbetering van $\pm 8\%$ op voor de gemiddelde DSC.

Een tweede belangrijke groep zijn de methodes die gebruik maken van een transformatie. Het

beeld wordt spiraalvormig (3D) of cirkelvormig (2D) gescand. Op het getransformeerde beeld zoekt een minimaal kostpadberekening de beste segmentatie. De laatste stap van dergelijke algoritmes is de terugtransformatie. Door de gradiënt van de mogelijke randen van de metastasen te versterken, zal het minimaal kostpad meer de neiging hebben om de randen van de metastasen te volgen. Dit wordt gedaan door de gradiënt van een binair gesegmenteerd beeld bij de gradiënt van het beeld op te tellen. De binaire segmentatie gebeurt door interval thresholding met grenzen 1033 en 1108. Omdat trainings- en testdata een grote snededikte hebben, werkt het 2D cirkelvormig scannen beduidend beter dan het 3D spiraalvormig scannen, maar vergt ook meer gebruikersinteractie.

De laatste groep segmentatiemethodes zijn gebaseerd op een *Level Set*-methode. Door oplossing van een partiële differentiaalvergelijking verandert een initiële level set onder invloed van een snelheidsveld. De initiële level set is het resultaat van een segmentatie door de algoritmes op basis van een transformatie die daarna geconverteerd werden tot een afstandsfunctie door middel van een afstandstransformatie. Het snelheidsveld wordt gevormd uit een fuzzy pixelclassificatiebeeld door volgende filteroperatie: $I_V = f(I_k) = |2p(c_{metastase}|I_k) - 1|$. Ook andere filters werden getest. Deze methode werd zowel geïmplementeerd voor supervised als unsupervised learning.

7.2 Een kritische kijk op de resultaten

Voor de gegeven trainings- en testdata is duidelijk dat de methode op basis van de 2D-2D transformatie (Circle Scanning Technique) en deze methode als basis voor een level set methode de beste resultaten geven. De dikte van de snedes speelt hier een belangrijke rol doordat hierdoor de methode met de 3D-2D transformatie (Spiral Scanning Technique) mindere resultaten geeft. Er is ook een groot verschil op te merken voor metastasen gelegen in het midden van de lever en metastasen aan de rand van de lever. Deze die aan de rand in de buurt van de borstkas liggen, hebben het grootste risico op een segmentatie van mindere kwaliteit omdat uitgroei uit de lever er nog steeds voorkomt. Met een afgelijnde lever als beginsituatie stelt zich dit probleem niet.

De resultaten vergelijken met in de literatuur gepubliceerde resultaten moet met enige voorzichtigheid gebeuren aangezien de invloed van de verschillend protocols, CT-scanners, patiënten,... moeilijk in kaart te brengen is. Als we deze voorgenoemde facetten buiten beschouwing laten, is de gemiddelde sensitiviteit van 80,0% voor Level Sets met Circle Scan zeker niet slecht. Anderzijds zijn er in de literatuur vergelijkbare studies uitgevoerd die betere resultaten geven. De segmentatie en bijgevolg ook de kwantificatie is de meeste gevallen onvoldoende betrouwbaar voor klinische toepassingen, vooral voor metastasen die slechts weinig verschillen in grijswaarde van de omliggende lever. Bolvormige metastasen met voldoende contrast in vergelijking met de omliggende lever, geven wel goede, accurate resultaten. De gemiddelde nauwkeurigheid ligt echter niet binnen de gemiddelde nauwkeurigheid van de door onszelf afgelijnde metastasen.

7.3 Suggesties voor verder werk en onderzoek

Het grootste potentieel voor beelden met grote snededikte om tot een goede segmentatie te komen, is volgens ons aanwezig in de Level Sets met Circle Scan methode. Door de implementatie van het EM-algoritme is deze methode breder toepasbaar. De waarden voor de wegingsfactoren van de advectie-, krommings- en propagatieterm dienen echter verder geoptimaliseerd te worden om tot een betere segmentatie te komen. In deze thesis werden enkel de grootteordes voor deze wegingsfactoren voor een lokaal optimum gezocht. Met een betere optimalisatiestrategie kunnen waarschijnlijk betere optima gevonden worden.

Een tweede belangrijke mogelijke verbetering aan deze methode is de verbetering van het pixelclassificatie-algoritme. Hiervoor hebben we enkele suggesties:

- Door de textuurkenmerken efficiënter te berekenen, kan ook de textuur voor grotere dimensies bestudeerd worden. Op basis van onze textuuranalyse van grotere gebieden, verwachten we dat bepaalde textuurkenmerken zoals correlatie, contrast en variantie voor de grotere afstanden beter onderscheid maken tussen leverweefsel en metastaseweefsel. De verdeling van

deze textuurkenmerken kan opgenomen worden in het statische model (bijv. een multivariate normale verdeling) voor de pixelclassificatie. Een manier om de schaal van de textuur te bepalen is beschreven in [50].

- Het toevoegen van outlier-detectie aan het niet gesuperviseerde pixelclassificatie-algoritme maakt het algoritme robuuster. Het normaal verdeeld model is immers slechts een benadering van de werkelijkheid en de schatting van de parameters zou niet aangetast mogen worden door de aanwezigheid van metastasen. Het toevoegen van outlier-detectie is uitvoerig beschreven in [51].
- Voor het schatten van de parameters met het EM-algoritme is er een betere stopmaat nodig die rekening houdt met de convergentie van het algoritme. Nu stopt het algoritme wanneer het toegelaten maximaal aantal iteraties bereikt is.

Het lijkt ons ook nuttig het algoritme te testen op recentere data met een betere axiale resolutie. Op die manier kan beter ingeschat worden wat de werkelijke waarde van het algoritme is.

Het uiteindelijke doel van de segmentatie-algoritmes is het meten van volumes en het meten van intensiteitsveranderingen binnen die volumes op verschillende tijdsogenblikken om zo de respons van een behandeling na te gaan en op die manier betere voorspellingen te doen over de evolutie van de ziekte. [29] bespreekt welke minimale veranderingen in grootte of intensiteit nodig zijn om van een klinisch relevante verandering te spreken. Opdat het algoritme in de klinische praktijk kan gebruikt worden, is het bovendien nodig de gebruikersinterface te verbeteren met aandacht voor de vragen van de clinici.

Als eindconclusie stellen we dat verbeteringen aan het algoritme noodzakelijk zijn opdat het algoritme optimaal zou kunnen presteren in de klinische praktijk. Met dit werk is echter een kader geschapen waarbinnen nog verbeteringen mogelijk zijn.

Bibliografie

- [1] S. Bipat, M. S. van Leeuwen, E. F. I. Comans, M. E. J. Pijl, P. M. M. Bossuyt, A. H. Zwinderman, and J. Stoker, “Colorectal liver metastases: CT, MR imaging, and PET for diagnosis meta-analysis,” *Radiology*, pp. 123–131, Augustus 2005.
- [2] G. S. Gazelle and S. Saini, *Hepatobiliary and Pancreatic Radiology: Imaging and Intervention*. Thieme Publishing Group, 1998.
- [3] North American Association of Central Cancer Registries, “Top five most commonly diagnosed cancers in the U.S. by race/ethnic group, 1999-2003,” 2003. laatst bekeken: 18/10/2007.
- [4] Vlaamse Liga tegen Kanker, “Meest voorkomende kankers in Vlaanderen 2000-2001.”
- [5] Elsevier, “Aantal gevallen kanker verdubbeld in 30 jaar,” *Elsevier.nl*, April 2005.
- [6] National Cancer Institute, “Metastatic cancer: Questions and answers.” <http://www.cancer.gov/cancertopics/factsheet/Sites-Types/metastatic>. laatst bekeken: 4/11/2007.
- [7] D. Ribatti, G. Mangialardi, and A. Vacca, “Stephen Paget and the seed and soil theory of metastatic dissemination,” *Clinical and Experimental Medicine*, vol. 6, pp. 145–149, December 2006.
- [8] B. Taragin, “Liver metastases, ct scan.” <http://www.nlm.nih.gov/medlineplus/ency/imagepages/1180.htm>. laatst bekeken: 8/5/2008.
- [9] P. Suetens, *Fundamentals of Medical Imaging*. New York: Cambridge University Press, 2002.
- [10] J. Corthout, D. Smeets, and B. Stijnen, “Automatische nauwkeurige segmentatie van het kaakgewricht uit CT en CBCT data.” 2007.
- [11] K. Kinkel, Y. Lu, M. Both, R. S. Warren, and R. F. Thoeni, “Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): A meta-analysis,” *Radiology*, no. 224, pp. 748–756, 2002.
- [12] L. Van Hoe, E. Van Cutsem, I. Vergote, A. L. Baert, E. Bellon, P. Dupont, and G. Marchal, “Size quantification of liver metastasis in patients undergoing cancer treatment: reproducibility of one-, two-, and three-dimensional measurements determined with spiral CT,” *Radiology*, no. 202, pp. 671–675, 1997.
- [13] F. Maes, D. Vandermeulen, P. Suetens, and G. Marchal, “Automatic image partitioning for generic object segmentation in medical images,” in *Computational Imaging and Vision* (Y. Bizais, C. Barillot, and R. di Paola, eds.), vol. 3, pp. 215–226, XIVth International Conference on Information Processing in Medical Imaging, Kluwer Academic Publishers, 1995.

- [14] F. Maes, D. Vandermeulen, P. Suetens, and G. Marchal, "Computer-aided interactive object delineation using an intelligent paintbrush technique," in *Lecture Notes in Computer Science* (N. Ayache, ed.), vol. 905, pp. 77–83, First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine, Springer, 1995.
- [15] E. Bellon, M. Feron, F. Maes, L. V. Hoe, D. Delaere, F. Haven, S. Sunaert, A. Baert, G. Marchal, and P. Suetens, "Evaluation of manual vs semi-automated delineation of liver lesions on CT images," *European Radiology*, vol. 7, p. 432438, 1997.
- [16] D. Pham, X. Chenyang, and L. Jerry, "A survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, 2000.
- [17] S. Beucher, "Image segmentation and mathematical morphology." <http://cmm.ensmp.fr/~beucher/wtshed.html>. laatst bekeken: 8/5/2008.
- [18] M. Burger and B. Hackl, "Level-set methods." <http://www.sfb013.unilinz.ac.at/index.php?id=f1308-geometric>. laatst bekeken: 8/5/2008.
- [19] J. A. Sethian, *Level Set Methods and Fast Marching Methods*. New York: Cambridge University Press, 1999.
- [20] S. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003.
- [21] P. J. Yim and D. J. Foran, "Volumetry of hepatic metastases in computed tomography using the watershed and active contour algorithms," in *Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems*, IEEE Computer Society, 2003.
- [22] L. D. Griffin, A. C. F. Colchester, and G. P. Robinson, "Scale and segmentation of grey-level images using maximum gradient paths," *Image Vision Comput.*, vol. 10, no. 6, pp. 389–402, 1992.
- [23] B. Zhao and D. Yankelevitz, "Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images," *Medical Physics*, vol. 26(6), pp. 4678–4689, Juni 1999.
- [24] J. Wang, R. Engelmann, *et al.*, "Segmentation of pulmonary nodules in three-dimensional CT images by use of a spiral-scanning technique," *Medical Physics*, vol. 34(12), pp. 4678–4689, December 2007.
- [25] R. S. Tuma, "Sometimes size doesn't matter: Reevaluating RECIST and tumor response rate endpoints," *Journal of the National Cancer Institute*, vol. 98, pp. 1272–1274, September 2006.
- [26] S. Prasad, K. Jhaveri, *et al.*, "CT tumor measurement for therapeutic response assessment: comparison of unidimensional, bidimensional, and volumetric techniques initial observations," *Radiology*, vol. 225(2), pp. 416–9, November 2002.
- [27] L. N. Tran, M. S. Brown, J. G. Goldin, X. Yan, R. C. Pais, M. F. McNitt-Gray, D. Gjertson, S. R. Rogers, and D. R. Aberle, "Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography," *Academic Radiology*, vol. 11(12), pp. 1355–1360, Nov-Dec 2004.
- [28] Y. Hwang, E. Ashton, C. Storgard, S. Wong, E. Malseed, J. Hill, L. Goyal, and J. Evelhoch, "Comparison of 1- and 2-dimensional measurements with volumetric measurements for evaluation of change in tumor size," *J Clin Oncol (Meeting Abstracts)*, vol. 25, no. 18_suppl, pp. 3563–, 2007.

- [29] H. Choi, C. Charnsangavej, S. C. Faria, H. A. Macapinlac, M. A. Burgess, S. R. Patel, L. L. Chen, D. A. Podoloff, and R. S. Benjamin, "Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: Proposal of new computed tomography response criteria," *Journal of Clinical Oncology*, vol. 25, pp. 1753–1759, Mei 2007.
- [30] R. S. Benjamin, H. Choi, H. A. Macapinlac, M. A. Burgess, S. R. Patel, L. L. Chen, D. A. Podoloff, and C. Charnsangavej, "We should desist using recist, at least in gist," *Journal of Clinical Oncology*, vol. 25, pp. 1760–1764, Mei 2007.
- [31] E. Weisstein, "Student's t-distribution." From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/Studentst-Distribution.html>. laatst bekeken: 14/10/2007.
- [32] N. Balakrishnan, *Handbook of the Logistic Distribution*. New York: Marcel Dekker, 1992.
- [33] A. Kurani, D.H.Xu, J.D.Furst, and D.S.Raicu, "Co-occurrence matrices for volumetric data," *Engineering in Medicine and Biology*, Augustus 2004.
- [34] D. Xu, J. Lee, D. Raicu, J. Furst, and D. Channin, "Texture classification of normal tissues in computed tomography," in *The 2005 Annual meeting of the Society for Computer Applications in Radiology*, (Orlando, Florida), Society for Computer Applications in Radiology, Juni 2-5 2005.
- [35] R. Šapina, "Computing textural features based on co-occurrence matrix for infrared images," *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on*, pp. 373–376, 2001.
- [36] P. S. Kostka *et al.*, "Hybrid feature vector extraction in unsupervised learning neural classifier," *Engineering in Medicine and Biology*, vol. 1-4, pp. 5664–5667, September 2005.
- [37] K. Vanden Branden, *Robust methods for high-dimensional data and a theoretical study of depth related estimators*. Thesis-dissertation, K.U. Leuven. Faculteit Wetenschappen. Departement Wiskunde, Mei 2005.
- [38] H. Wadell, "Volume, shape and roundness of quartz particles," *Journal of Geology*, no. 43, pp. 250–280, 1935.
- [39] D. W. Cantrell, "Surface area of an ellipsoid." <http://home.att.net/numericana/answer/ellipsoid.htm#spheroid>, Mei 2004. laatst bekeken: 23/3/2008.
- [40] MeVis Research GmbH, "MeVisLab." <http://www.mevislab.de>.
- [41] L. Ibañez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*. 2005.
- [42] J. S. Suri, D. L. Wilson, and S. Laxminarayan, *Handbook of Biomedical Image Analysis*, vol. II: Segmentation Models Part B. Springer, 2005.
- [43] R. P. Sørli, "Automatic segmentation of liver tumors from MRI images," Master's thesis, University of Oslo, November 2005.
- [44] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des jura," *Bulletin del la Socit Vaudoise des Sciences Naturelles*, vol. 37, pp. 547–579, 1901.
- [45] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining, (First Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2005.
- [46] Medical Image Computing and Computer-Assisted Intervention, "3D segmentation in the clinic: A grand challenge II." <http://grand-challenge2008.bigr.nl/>, 2008.

- [47] L. Massoptier and S. Casciari, "A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans," *European Radiology*, Maart 2008.
- [48] E. K. Fishman, B. S. Kuszyk, D. G. Heath, L. Gao, and B. Cabral, "Surgical planning for liver resection," *Computer*, vol. 29, no. 1, pp. 64–72, 1996.
- [49] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, 1987.
- [50] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1026–1038, Augustus 2002.
- [51] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," Tech. Rep. 0004, KUL/ESAT/PSI, Heverlee, Belgium, Juni 2000.
- [52] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*. No. 153 in Applied Mathematical Sciences, Cambridge University Press, 2003.
- [53] J. S. Suri, S. K. Setarehdan, and S. Singh, *Advanced Algorithmic Approaches to Medical Image Segmentation: state-of-the-art applications in cardiology, neurology, mammography and pathology*. Advances in Pattern Recognition, Springer-Verlag London, 2002.
- [54] D. Xu, A. Kurani, J. Furst, and D. Raicu, "Run-length encoding for volumetric texture," in *The 4th IASTED International Conference on Visualization, Imaging, and Image Processing*, (Marbella Spain), VIIP, September 6-8 2004.
- [55] R. M. Haralick, Dinstein, and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610–621, November 1973.
- [56] M. Hall-Beyer, "The grey level co-occurrence matrix." http://www.fp.ucalgary.ca/mhallbey/the_glcmm.htm.
- [57] L. Semler and J. Furst, "Wavelet-based texture classification of tissues in computed tomography," in *CBMS '05: Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems*, (Washington, DC, USA), pp. 265–270, IEEE Computer Society, 2005.
- [58] M. Wronski, "Triangle fillers." http://www-users.mat.uni.torun.pl/~wrona/3d_tutor/tri_fillers.html. laatst bekeken: 29/11/2007.
- [59] Silicon Graphics Inc, "Open inventor." <http://oss.sgi.com/projects/inventor/>. laatst bekeken: 21/4/2008.
- [60] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, no. 22 (1), p. 6179, 1997.

Lijst van Figuren

2.1	Voorbeeld van levermetastasen, opgenomen met CT [8]	5
2.2	Annihilatie van een electron en een positron	6
2.3	Vergelijking beeldvormingsmodaliteiten per patiënt [1]	7
2.4	Vergelijking tussen verschillende beeldvormingsmodaliteiten per metastase [1]	8
2.5	Inter- en intra-observatorvariabiliteit: vergelijking manuele met semi-automatische aflijning [15]	9
2.6	Segmentatie : (a) origineel beeld, (b) segmentatie met behulp van K-means algoritme, (c) segmentatie met behulp van het K-means algoritme en het gebruik van MRF [16].	12
2.7	De verschillende fases in het waterscheidingsalgoritme [17]: links boven het beeld voor segmentatie, rechts onder het beeld met segmentatie. De tussenliggende beelden geven het vollopen van het bassin aan.	12
2.8	Marker gecontroleerd waterscheidingsalgoritme [17]	13
2.9	Het concept van een zero set in een level set [18]	13
2.10	MR doorsnede van het hart (l) en waterscheiding van het gradiëntbeeld (r)	14
2.11	waterscheiding van het gradiëntbeeld (l) en gebieden na samenvoeging (r) [14]	15
2.12	Gemiddelde gradiëntsterkte van de contour van de metastase voor variërende threshold. De optimale is degene waarvoor de gradiënt maximaal is [23].	15
2.13	Resultaten van dual scale active contour algoritme op de segmentatie van levermetastasen [21]	16
2.14	3D snedes van longmetastasen(bovenaan) en transformatie naar 2D(onderaan) [24]	17
2.15	Histogram van overlapwaarden voor de metastasen in de 2 datasets [24]	17
2.16	Meting van RECIST (geel) en zijn grootste orthogonale diameter (rood)	18
3.1	Statistieken van de grijswaardenverdeling	21
3.2	Voorbeeld van de verdeling van de grijswaarden in 1 metastase	22
3.3	Voorbeeld van de verdeling van de grijswaarden in de lever	22
3.4	Voorbeeld van de verdeling van de grijswaarden in de lever	24
3.5	Van een sample(a) met lever- en tumorweefsel wordt de co-occurrencematrix berekend voor een pixelafstand van respectievelijk 1(b) en 20(c) in de verticale richting. In (d) gebeurt dezelfde berekening in de horizontale richting.	26
3.6	Verdeling van (a)contrast, (b)variantie, (c)correlatie, (d)entropie	30
3.7	Voorbeeld van een langwerpige metastase t.g.v. de rand van de lever	32
3.8	De loodrechte op RECIST in functie van RECIST	32
3.9	Het manueel volume van de metastase in functie van (a) RECIST ³ en (b) WHO ^{3/2}	33
4.1	Pixelclassificatie: fuzzy (a) en binair (b)	35
4.2	Een metastase na 2D region growing per snede	38
4.3	Een metastase na 2D region growing per snede	40
4.4	Transformatie volgens een spiraal scan	41
4.5	Transformatie volgens een spiraal scan	41
4.6	Segmentatie na bepaling MCP	42

4.7	Niet equidistante bemonstering	44
4.8	Equidistante bemonstering	44
4.9	Blokdiagramma van de Level Set methode	46
4.10	Twee filters die vaak in de literatuur gebruikt worden	48
4.11	Een typisch theoretisch verloop van de fuzzy pixelclassificatie overheen een rand van een metastase	49
4.12	Resultaat van het snelheidsbeeld voor f_1 (blauw) en f_2 (rood)	49
5.1	DSC in functie van de maximale threshold voor patiënt 1(links) en patiënt 2(rechts)	53
5.2	DSC in functie van de breedte en het centrum van het interval voor patiënt 1(links) en patiënt 2(rechts)	53
5.3	De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten	54
5.4	De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten en textuur	55
5.5	De DSC en de FNF i.f.v. de drempelwaarde voor pixelclassificatie o.b.v. intensiteiten, gebruik makend van EM	55
5.6	De DSC, RMSSD, RAV en MASD in functie van de intervalgrootte voor Region Growing met één seed per snede	56
5.7	Resultaten van de 13 validatiematen voor de testdata van Region Growing met één seed per snede	57
5.8	De validatiematen i.f.v. α en β voor Region Growing met een statistisch model en een cirkel als restrictie	59
5.9	Resultaten van de 13 validatiematen voor de testdata van Region Growing met een cirkel als restrictie, $\alpha = 2$, $\beta = 1,5$	60
5.10	De validatiematen i.f.v. α en β voor Region Growing met een statistisch model en een rechthoek als restrictie	60
5.11	Resultaten van de 13 validatiematen voor de testdata van Region Growing met een rechthoek als restrictie, $\alpha = 3$, $\beta = 1,5$	61
5.12	De validatiematen i.f.v. ondergrens voor Region Growing gegeven een fuzzy classificatiebeeld	62
5.13	Resultaten van de 13 validatiematen voor de testdata van Region Growing gegeven een fuzzy classificatiebeeld, ondergrens = 0,85	63
5.14	De validatiematen i.f.v. w_{int}/w_{ext} voor de Spiral Scanning methode	64
5.15	Resultaten van de 13 validatiematen voor de testdata van Spiral Scanning	64
5.16	De validatiematen i.f.v. w_{int}/w_{ext} voor de Spiral Scanning methode met een afgelijnde lever	65
5.17	Resultaten van de 13 validatiematen voor de testdata van Spiral Scanning op een afgelijnde lever, $w_{int}/w_{ext} = 2$	66
5.18	De validatiematen i.f.v. w_{int}/w_{ext} voor de Circle Scanning methode	67
5.19	Resultaten van de 13 validatiematen voor de testdata van Circle Scanning, $w_{int}/w_{ext} = 2$	67
5.20	De beste, gemiddelde en slechtste prestatie van het Circle Scanning algoritme op de testdata	68
5.21	De validatiematen i.f.v. w_{int}/w_{ext} voor de Circle Scanning methode met een afgelijnde lever	69
5.22	Resultaten van de 13 validatiematen voor de testdata van Circle Scanning op een afgelijnde lever, $w_{int}/w_{ext} = 2$	70
5.23	De validatiematen voor een level set methode met en zonder supervised learning .	71
5.24	Resultaten van de 13 validatiematen voor de testdata van Level Sets met supervised learning	71
5.25	Resultaten van de 13 validatiematen voor de testdata van Level Sets zonder supervised learning	72
5.26	De beste, gemiddelde en slechtste prestatie van het Levels Sets met EM algoritme op de testdata	73

5.27	De validatiematen i.f.v. voor de trainingsdata met en zonder supervised learning (afgelijnde lever)	73
5.28	Resultaten van de validatiematen voor de testdata van Level Sets met supervised learning toegepast op een afgelijnde lever	74
5.29	Resultaten van de validatiematen voor de testdata van Level Sets zonder supervised learning toegepast op een afgelijnde lever	75
5.30	De resultaten voor de 13 validatiematen van observer 1 vergeleken met observer 2(a) en observer 3(b)	77
6.1	Orthoviewer met een weergave van de metastase in axiaal(rb), sagittaal(lo) en coronaal(ro) vlak	79
6.2	15 unieke situaties die zich voordoen bij <i>Marching Cubes</i>	79
6.3	Surface Rendering van de lever met zijn metastasen	80
6.4	Volume Rendering van de lever met bloedvaten	80
6.5	Meting van de RECIST en zijn loodrechte met weergave van de afmetingen (a) en oppervlak van een aflijning weergegeven op het beeld (b)	81
B.1	Voorbeeld-beeld voor co-occurrence methode	98
B.2	Entropie i.f.v richting en pixelafstand	99
B.3	Energie i.f.v richting en pixelafstand	100
B.4	Contrast i.f.v richting en pixelafstand	100
B.5	Homogeneïteit i.f.v richting en pixelafstand	101
B.6	Gemiddelde i.f.v richting en pixelafstand	101
B.7	Variantie i.f.v richting en pixelafstand	102
B.8	Correlatie i.f.v richting en pixelafstand	102
B.9	Maximale Waarschijnlijkheid i.f.v richting en pixelafstand	103
B.10	Invers Verschil Moment i.f.v richting en pixelafstand	103
B.11	Neiging tot Clustering i.f.v richting en pixelafstand	104
C.1	Mogelijk standen van een driehoek	105
D.1	Voorbeeld van een netwerk van modules in <i>MeVisLab</i>	108
D.2	GUI LiverCAD	108
E.1	Overige validatiematen voor Region Growing met één seed per slice	113
E.2	De DSC,RMSSD,RAV en MASD in functie van de intervalgrootte voor Region Growing met twee seeds	114
E.3	Overige validatiematen voor Region Growing met 2 seeds per metastase	115
E.4	Resultaten van de 13 validatiematen voor de testdata van Region Growing met twee seeds per metastase, interval = 1.05%	116
E.5	De DSC,RMSSD,RAV en MASD in functie van de intervalgrootte voor Region Growing met één seed per metastase	117
E.6	Overige validatiematen voor Region Growing met één seed per metastase	118
E.7	Resultaten van de 13 validatiematen voor de testdata van Region Growing met één seed per metastase, bovengrens = 1115 GW	119
E.8	Overige validatiematen voor Region Growing met een cirkel als restrictie	120
E.9	Overige validatiematen voor Region Growing met een rechthoek als restrictie	121
E.10	Overige validatiematen voor Region Growing van een fuzzy classificatiebeeld	122
E.11	Overige validatiematen voor Spiral Scanning	123
E.12	Overige validatiematen voor Spiral Scanning met een afgelijnde lever	124
E.13	Overige validatiematen voor Circle Scanning	125
E.14	Overige validatiematen voor Circle Scanning met een afgelijnde lever	126
E.15	De validatiematen i.f.v. de wegingsfactor voor de advectieterm	127
E.16	De validatiematen i.f.v. de wegingsfactor voor de krommingsterm	127

E.17 De validatiematen i.f.v. de wegingsfactor voor de propagatieterm	128
E.18 De COV van het volume gegroepeerd per grootte van de tumor. \circ =gemiddelde van de eigen studie, \times =gemiddelde uit de literatuur	128
F.1 block diagram	131
F.2 Original Image	131
F.3 Scan image	131
F.4 Binary segmentation	132
F.5 Speed Image	133
F.6 Example	135

Lijst van Tabellen

2.1	Meest voorkomende kankers [3]	3
2.2	Vergelijking Beeldmodaliteiten: sensitiviteit [1, 11]	7
3.1	Verplaatsingsvectoren van de co-occurrence matrix	27
3.2	Statistieken uit de co-occurrence matrix	27
3.3	Samenvatting resultaten van discriminante maten	29
3.4	Eigenschappen van de verdeling van de textuurkenmerken	30
3.5	Textuurkenmerken volgens hun onderscheidend karakter op voxelniveau	30
4.1	μ en σ van verdeling van lever- en metastaseweefsel	36
5.1	Resultaten van alle methodes	75
5.2	Resultaten van enkele methodes voor een afgelijnde lever	76
B.1	Co-occurrence matrix van beeld in Figuur B.1	98
E.1	Andere validatiematen voor een afgelijnde lever	126
F.1	Results	136
F.2	Results of training set	137
F.3	Results surface distances of training set	137

Bijlage A

Level Sets

A.1 Inleiding

Deze bijlage gaat dieper in op level set methodes en dient als aanvulling op 2.4.2, waar een kort overzicht werd gegeven. In 2.4.2 vindt enkel een korte bespreking van de lokale level set methodes plaats. Hier gebeurt een situering in een groter geheel, nl. interface propagatie [19]. Ook gaan we dieper in op numerieke methodes gebruikt bij level set methodes. De tekst is gebaseerd op [19, 20, 52, 53].

A.2 Interface propagatie

De propagatie van een interface kan geschreven worden als een partiële differentiaalvergelijking. Hierbij zijn er twee perspectieven mogelijk die leiden tot de *Fast Marching Method* en de *Narrow Band Level Set*. Onder invloed van een snelheidsveld \vec{V} evolueert de interface, een curve in twee dimensies of een oppervlak in drie dimensies, in een bepaalde richting. Het snelheidsveld is afhankelijk van verschillende factoren:

- Lokale eigenschappen (lokale geometrische informatie zoals kromming)
- Globale eigenschappen (vorm en positie van de interface)
- Onafhankelijke eigenschappen

Een propagerende interface kan op twee manieren bekeken worden: als grenswaarde formulering en als initiële waarde formulering.

A.2.1 Grenswaarde formulering

Dit wordt beschreven door:

$$|\nabla T|V_N = 1$$

met $T(x, y)$ de aankomsttijd van de interface en V_N de normale component van de snelheid op de interface. Deze moet voor de grenswaarde formulering steeds positief zijn. Dit wil zeggen dat voor het gebruik van *Fast Marching Methods*, die gebaseerd zijn op deze formulering, voor segmentatie het initiële oppervlak kleiner moet zijn dan het uiteindelijke oppervlak. De interface kan steeds gevonden worden door $\Gamma(t) = \{(x, y) | T(x, y) = t\}$.

A.2.2 Initiële waarde formulering

In deze formulering wordt de interface wordt voorgesteld als een impliciete functie. De initiële positie van de interface wordt als een zero level set ingebed in een hoger dimensionale functie ϕ ,

waarvan de evolutie ook de evolutie van de interface beschrijft. De level set vergelijking, ook wel convectie- of advectionvergelijking genoemd, is gegeven door een partiële differentiaalvergelijking:

$$\frac{\partial \phi}{\partial t} + \vec{V} \cdot \nabla \phi = 0 \quad (\text{A.1})$$

met $\phi(x, t = 0)$ ¹. Hierbij mag V_N zowel positief als negatief zijn, wat wil zeggen dat de interface zowel voorwaarts als achterwaarts kan evolueren.

De interface kan steeds teruggevonden worden als: $\Gamma(t) = \{(x, y) | \phi(x, y, t)\}$.

A.2.3 Voordelen

De aanpak van deze formuleringen heeft verschillende voordelen:

- Uitbreidbaarheid naar hogere dimensies
- Natuurlijke manier van omspringen met topologische veranderingen
- Accuraat te benaderen door rekenkundige schema's
- Makkelijk bepalen van intrinsieke geometrische eigenschappen van de interface:
 - normal vector: $\mathbf{n} = \frac{\nabla \phi}{|\nabla \phi|}$ of $\mathbf{n} = \frac{\nabla T}{|\nabla T|}$
 - de kromming: $\kappa = \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|}$ of $\kappa = \nabla \cdot \frac{\nabla T}{|\nabla T|}$
- Beide formuleringen zijn efficiënt om te gebruiken in computationele strategieën die leiden tot *Fast Marching Methods* en *de Narrow Band Level Set Methods*.

A.3 Numerieke methodes

Eens ϕ en \vec{V} gedefiniëerd voor elk punt in een cartesisch rooster, kunnen numerieke methodes toegepast worden om ϕ voorwaarts in de tijd te ontwikkelen en zo de interface bewegen doorheen het rooster. Aangezien vooral lokale level set methodes belangrijk zijn in deze thesis, volgt hier een beschrijving van numerieke algoritmes voor deze benadering.

De level set vergelijking (A.1) bevat een snelheidsveld \mathbf{v} die afhangt van de geometrie van de zero level set en van externe eigenschappen. Een prototype voor \mathbf{v} is:

$$\vec{V} = a(\mathbf{x}) \frac{\nabla \phi}{|\nabla \phi|} - \mu(\mathbf{x}) \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \quad (\text{A.2})$$

De snelheid \mathbf{v} bevat een convectieterm (1^e term) en een krommingsafhankelijke bewegingsterm (2^e term).

(A.2) invullen in (A.1) levert:

$$\frac{\partial \phi}{\partial t} + a(\mathbf{x}) |\nabla \phi| - \mu(\mathbf{x}) |\nabla \phi| \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = 0 \quad (\text{A.3})$$

Deze partiële differentiaalvergelijking geeft de evolutie van de level set functie ϕ . Het implementeren van deze methode als een bruikbare numerieke methode vraagt discretisatie van elk van de termen in (A.3). Voor eindige differenties gebruiken we de volgende operatoren:

$$D_{\pm}^x \phi_{jk} = \pm \frac{(\phi_{j\pm 1, k} - \phi_{jk})}{\Delta}, \quad D_{\pm}^y \phi_{jk} = \pm \frac{(\phi_{j, k\pm 1} - \phi_{jk})}{\Delta}$$

¹In de literatuur wordt ϕ_t vaak gebruikt voor de partiële afgeleide van ϕ naar de tijd

A.3.1 Convectie

De discretisatie van de convectieterm is niet eenvoudig en gebeurt met een Hamilton Jacobi vergelijking met *essentially nonoscillatory* schema HJ ENO of een Hamilton Jacobi vergelijking met *weighted essentially nonoscillatory* schema HJ WENO.

A.3.2 Kromming

De krommingsterm wordt gediscetiseerd door centrale differentie:

$$|\nabla\phi| \approx \sqrt{\frac{1}{2}((D_+^x\phi_{jk})^2 + (D_-^x\phi_{jk})^2 + (D_+^y\phi_{jk})^2 + (D_-^y\phi_{jk})^2)}$$

$$\frac{\partial}{\partial x} \frac{\partial\phi}{|\nabla\phi|} \approx D_-^x \left(\frac{D_+^x\phi_{jk}}{\sqrt{(D_+^x\phi_{jk})^2 + (D_+^y\phi_{jk})^2 + \epsilon}} \right)$$

met $(D^y\phi_{jk})^2 = \frac{1}{4}((D_+^y\phi_{jk})^2 + (D_+^y\phi_{j+1,k})^2 + (D_-^y\phi_{jk})^2 + (D_-^y\phi_{j+1,k})^2)$

A.3.3 Tijdsdiscretisatie

Voor de discretisatie in de tijd kan meestal een lage orde voorwaartse Eulermethode gebruikt worden. Soms is echter een hogere orde temporele discretisatie nodig. De *total variation diminishing (TVD) Runge-Kutta (RK)* methodes verhogen de nauwkeurigheid voor een lijnenmethodebenadering voor de temporele discretisatie. De lijnenmethodebenadering veronderstelt dat de spatiale discretisatie gescheiden kan worden van de temporele discretisatie. Vandaar dat voor de tijdsdiscretisatie de 3^e orde TVD Runge-Kutta methode gebruikt wordt. Stel $\frac{\partial\phi_{jk}}{\partial t} = L_{jk}$ dan is de 1^e orde TVD Runge-Kutta methode:

$$\phi_{jk}^{n+1} = \phi_{jk}^n + \Delta t L_{jk}$$

$$\text{de 2}^e \text{ orde: } \begin{cases} \phi_{jk}^{n+\frac{1}{2}} = \phi_{jk}^n + \Delta t L_{jk}^n \\ \phi_{jk}^{n+1} = \frac{\phi_{jk}^{n+\frac{1}{2}} + \phi_{jk}^n}{2} + \frac{\Delta t}{2} L_{jk}^{n+\frac{1}{2}} \end{cases}$$

$$\text{de 3}^e \text{ orde: } \begin{cases} \phi_{jk}^{n+\frac{1}{3}} = \phi_{jk}^n + \Delta t L_{jk}^n \\ \phi_{jk}^{n+\frac{2}{3}} = \frac{3}{4}\phi_{jk}^n + \frac{1}{4}\phi_{jk}^{n+\frac{1}{3}} + \frac{1}{4}\Delta t L_{jk}^{n+\frac{1}{3}} \\ \phi_{jk}^{n+1} = \frac{1}{3}\phi_{jk}^n + \frac{2}{3}\phi_{jk}^{n+\frac{2}{3}} + \frac{2}{3}\Delta t L_{jk}^{n+\frac{2}{3}} \end{cases} .$$

Bijlage B

Textuuranalyse

B.1 Verschillende Technieken voor textuuranalyse

In de literatuur worden drie technieken besproken om texturen te analyseren: RLE, co-occurrence matrices en een wavelet-gebaseerde textuuranalyse.

B.1.1 Run Length Encoding

Het principe van RLE is vrij eenvoudig. Men gaat een beeld af en men telt het aantal opeenvolgende pixels met dezelfde waarde. Deze techniek wordt vaak gebruikt bij beeldcompressie maar hij kan ook aangewend worden om ruwe en minder ruwe structuren te herkennen in medische beelddata. Door het aantal opeenvolgende pixels te tellen in een specifieke richting met dezelfde grijswaarde kan met fijne en minder fijne structuren detecteren.

Hiervoor wordt een run-length matrix $P(i,j)$ gedefinieerd die het aantal runs van pixels met grijswaarde i en runlengte j weergeeft. Hiervoor wordt meestal het beeld geconverteerd naar minder grijswaarden zodat men voldoende grote runlengtes krijgt. Deze techniek kan dan vervolgens in richtingen worden berekend en hieruit worden vervolgens een aantal textuurbeschrijvende kenmerken gedistilleerd [54]:

- Short Run Emphasis (SRE)
- Long Run Emphasis (LRE)
- high gray-level run emphasis (HGRE)
- low gray-level run emphasis (LGRE)
- short run low gray-level emphasis (SRLGE)
- Short Run High gray-level emphasis (SRHGE)
- long run low gray-level emphasis (LRLGE)
- long run high gray-level emphasis (LRHGE)
- run-length non-uniformity (RLNU)
- grey-level non-uniformity (GLNU)
- Run Percentage (RPC)

B.1.2 Co-occurrence matrices

Een tweede techniek maakt gebruik van co-occurrence matrices (of term-term matrices) om een analyse van de textuur uit te voeren. De theorie van de co-occurrence matrix werd ontwikkeld door Haralick in 1973 [55]. Het is een veelgebruikte benadering in de textuuranalyse. Indien we vertrekken van Figuur B.1 kan de co-occurrence matrix beschouwd worden als de relatie tussen twee naburige pixels. Deze twee pixels noemt men de referentie-pixel en de buur-pixel. In dit



0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figuur B.1: Voorbeeld-beeld voor co-occurrence methode

voorbeeld ligt de buur-pixel rechts van de referentie-pixel. Een venster gaat het volledige beeld af waarbij iedere pixel om beurt de referentie pixel wordt. Voor iedere mogelijke combinatie van grijswaarden telt men het aantal keer dat hij voorkomt (*occurrences*). Dit geeft het resultaat in Tabel B.1. Dit is de co-occurrence matrix over afstand 1 in de oostelijke richting. Dit kan

Tabel B.1: Co-occurrence matrix van beeld in Figuur B.1

	0	1	2	3
0	2	2	1	0
1	0	2	0	0
2	0	0	3	1
3	0	0	0	1

verduidelijkt worden aan de hand van een voorbeeld. Stel men kiest het paar (0,1) met 0 de grijswaarde van de referentie-pixel en 1 de grijswaarde van de buur-pixel. De co-occurrence matrix geeft dan aan dat er twee paren van de combinatie (0,1) voorkomen in Figuur B.1 [56]. Een groot voordeel van deze matrices is dat zij zowel rekening houden met de intensiteiten als de spatiale informatie die zij met zich mee dragen.

B.1.3 Wavelet Gebaseerd

Een derde methode gebruikt wavelets om een textuuranalyse uit te voeren. Een wavelet is een mathematische functie die een signaal of beeld opdeelt in zijn verschillende frequentiegebieden en deze gebieden ieder op een gepaste schaal bestudeert. In [57] wordt een methode beschreven voor textuuranalyse van weefsels in CT-beelden op basis van wavelets. Hierbij kiest men voor drie types wavelets:

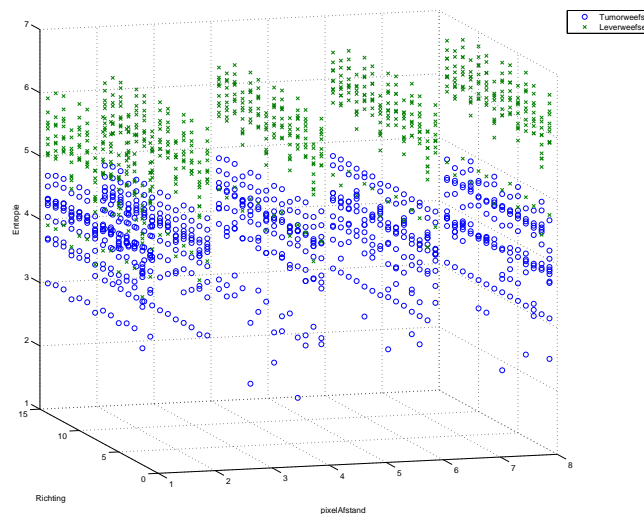
- Haar Wavelet: $\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{elders} \end{cases}$
- Daubechies 4 Wavelet: $c_i = \frac{1-\sqrt{3}}{4\sqrt{2}} \cdot S_{2i} - \frac{3-\sqrt{3}}{4\sqrt{2}} \cdot S_{2i+1} + \frac{3+\sqrt{3}}{4\sqrt{2}} \cdot S_{2i+2} - \frac{1+\sqrt{3}}{4\sqrt{2}} \cdot S_{2i+3}$
- Coiflet Wavelet

De drie waveletfuncties hebben hun voor- en nadelen. Haar wavelets zijn eenvoudig en weinig rekenintensief maar gebruiken geen overlappende vensters in de transformatie waardoor enkel veranderingen tussen naburige pixels worden opgenomen in het model. Daubechies wavelets zijn complexer en meer rekenintensief maar gebruiken wel overlappende vensters waardoor uitgemiddeld wordt over meerdere pixels. Hierdoor krijgt men een vloeiender verloop. Bij de Daubechies 4 wavelet is uitgemiddeld over vier aangrenzende pixels. De Coiflet wavelet is afgeleid van de Daubechies wavelet maar is meer rekenintensief en heeft een grotere overlap in vensters.

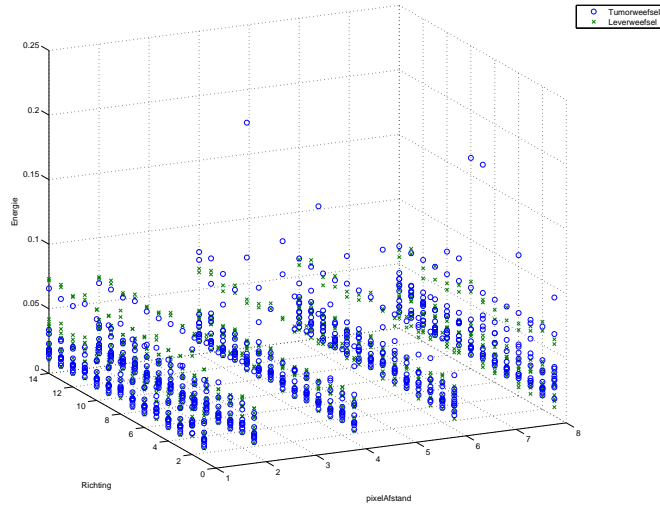
In een volgende stap worden een aantal textuurbeschrijvende kenmerken bepaald. Hiervoor voert men één van de drie voorgaande wavelets uit, waarbij men het aantal resolutieniveau's (opdeling in frequentiegebieden) gelijkstelt aan drie en de detail coëfficiëntmatrices berekent. De detail coëfficiëntmatrices geven informatie over de details die men op de verschillende resolutieniveaus kan onderscheiden in de horizontale, verticale en diagonale richting. Vervolgens worden de co-occurrence matrices van iedere detail coëfficiëntmatrix bepaald zoals beschreven in de vorige paragraaf. De textuurbeschrijvende kenmerken van Haralick worden hieruit berekend. Daarnaast berekent men ook het gemiddelde en de standaarddeviatie van de detail coëfficiëntmatrices als extra textuurbeschrijvende kenmerken. Dit levert een totaal van 396 (2D) kenmerken op die verder kunnen gereduceerd worden tot de meest significante kenmerken om de rekentijd te verlagen.

B.2 Scatterplots Textuuranalyse

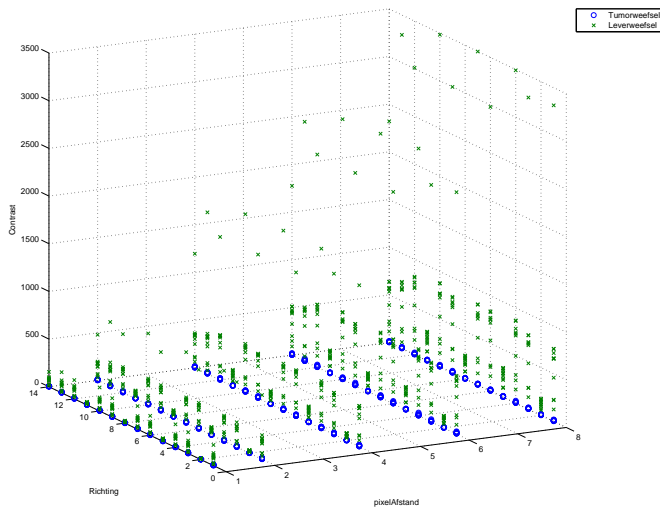
In deze bijlage staan de verschillende scatterplots van de textuurkenmerken van Tabel 3.2. Ieder textuurkenmerk wordt uitgezet in functie van de richting waarin de co-occurrencematrix wordt berekend en in functie van de pixelafstand tussen referentie- en buurpixel [56]. De gebruikte beelddata zijn de 12 datasets van de 2 trainingspatiënten. Er wordt een onderscheid gemaakt tussen leverweefsel (x) en tumorweefsel (o)



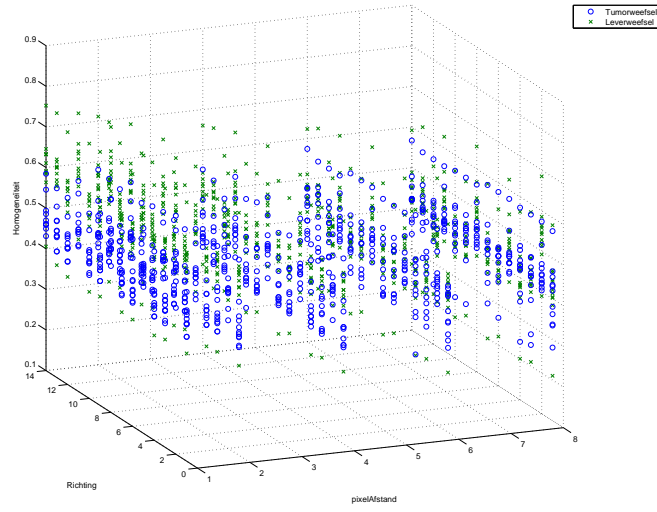
Figuur B.2: Entropie i.f.v richting en pixelafstand



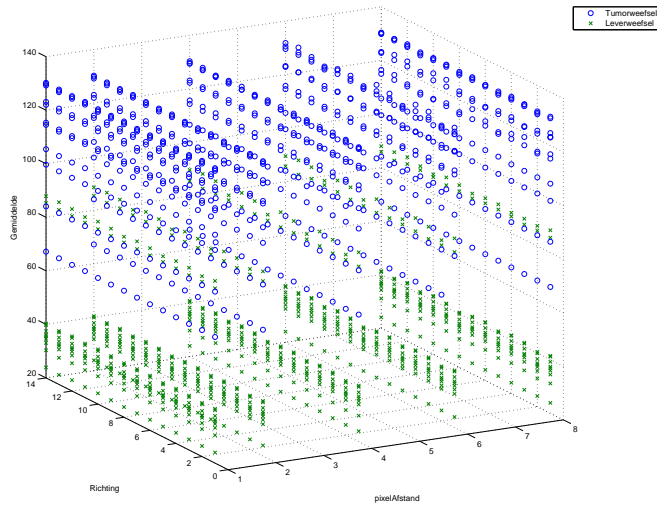
Figuur B.3: Energie i.f.v richting en pixelafstand



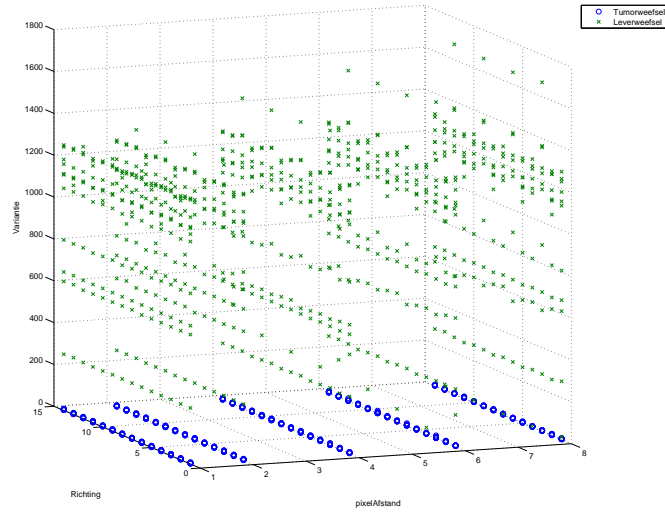
Figuur B.4: Contrast i.f.v richting en pixelafstand



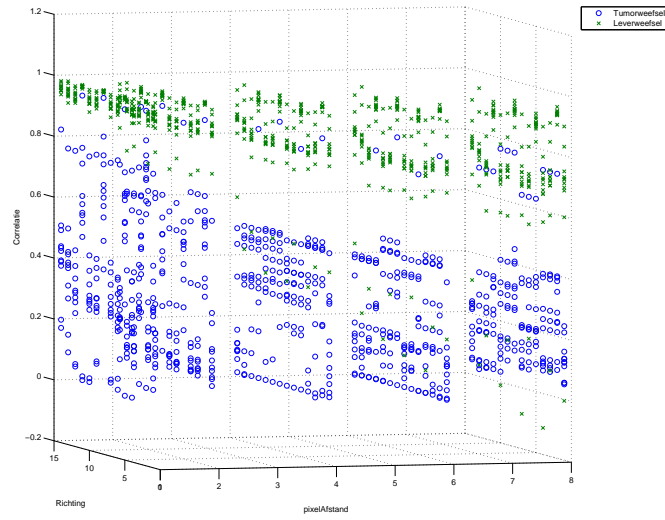
Figuur B.5: Homogeneiteit i.f.v richting en pixelafstand



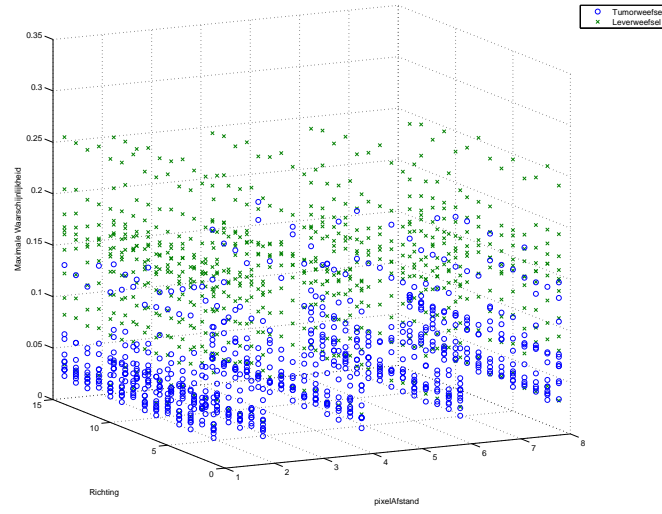
Figuur B.6: Gemiddelde i.f.v richting en pixelafstand



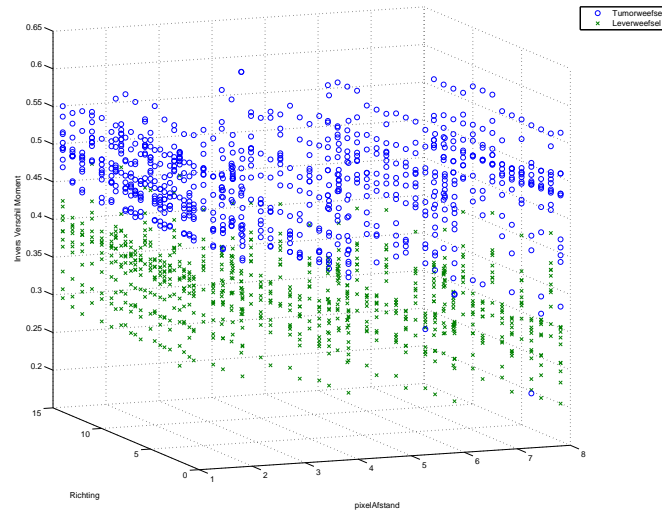
Figuur B.7: Variantie i.f.v richting en pixelafstand



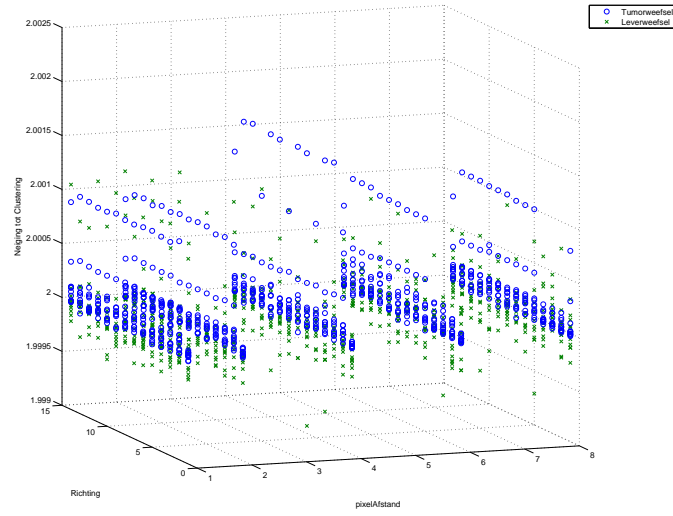
Figuur B.8: Correlatie i.f.v richting en pixelafstand



Figuur B.9: Maximale Waarschijnlijkheid i.f.v richting en pixelafstand



Figuur B.10: Invers Verschil Moment i.f.v richting en pixelafstand



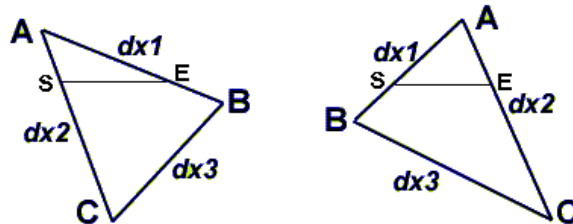
Figuur B.11: Neiging tot Clustering i.f.v richting en pixelafstand

Bijlage C

Het opvullen van een vierhoek

Inleiding

Het opvullen van een willekeurige vierhoek gebeurt door de vierhoek op te delen in 2 driehoeken door het trekken van een diagonaal.



Figuur C.1: Mogelijk standen van een driehoek

Belangrijk is dat de punten van de driehoek op een consistente manier gedefiniëerd zijn. Dit wil zeggen zoals aangegeven op Figuur C.1.

Dit gebeurt op volgende manier:

```
if (y1<=y2 && y2<=y3 ) {_fill_triangle(x1,y1,x2,y2,x3,y3,z,totImg);}
else if (y1<=y3 && y3<=y2) {_fill_triangle(x1,y1,x3,y3,x2,y2,z,totImg);}
else if (y2<=y1 && y1<=y3) {_fill_triangle(x2,y2,x1,y1,x3,y3,z,totImg);}
else if (y2<=y3 && y3<=y1) {_fill_triangle(x2,y2,x3,y3,x1,y1,z,totImg);}
else if (y3<=y2 && y2<=y1) {_fill_triangle(x3,y3,x2,y2,x1,y1,z,totImg);}
else if (y3<=y1 && y1<=y2) {_fill_triangle(x3,y3,x1,y1,x2,y2,z,totImg);}
else {std::cout << "Fout met coördinaten" << std::endl;}
```

Het vullen van een driehoek

Het algoritme voor het vullen van de driehoeken ziet er als volgt uit [58]:

```
float dx1;
float dx2;
float dx3;
if (y2-y1 > 0) {dx1=(x2-x1)/(y2-y1);}
else {dx1=0;}
if (y3-y1 > 0) {dx2=(x3-x1)/(y3-y1);}
else {dx2=0;}
```

```

if (y3-y2 > 0) {dx3=(x3-x2)/(y3-y2);}
else {dx3=0;}
float sx = x1;
float ex = x1;
float sy = y1;
float ey = y1;
if(dx1 > dx2) {
    for(;sy<=y2;sy++,ey++,sx+=dx2,ex+=dx1){
        _qscan(sx,ex,sy,z,totImg);
    }
    //E=B;

    for(;sy<=y3;sy++,ey++,sx+=dx2,ex+=dx3){
        _qscan(sx,ex,sy,z,totImg);
    }
} else {
    for(;sy<=y2;sy++,ey++,sx+=dx1,ex+=dx2){
        _qscan(sx,ex,sy,z,totImg);
    }
    for(;sy<=y3;sy++,ey++,sx+=dx3,ex+=dx2){
        _qscan(sx,ex,sy,z,totImg);
    }
}
}

```

Hierin gebeurt met het `_qscan`-commando een scan van de een lijn, waarbij de pixels op de correcte waarde worden gezet.

Het Scanline-algoritme

Het algoritme loopt elke lijn af en verandert de grijswaarde daar waar nodig.

```

Vector pos((int)sx,(int)sy,(int)z,0,0,0);
totImg.setCursorPos(pos);
for (int i = (int)sx; i <= (int)ex; i++){
    totImg.setCursorVoxel(1);
    totImg.moveFX();}

```

Bijlage D

MeVisLab

[40] Bij het ontwikkelen van de verschillende segmentatie- en validatiealgoritmes wordt gebruik gemaakt van de programmeeromgeving *MeVisLab*. *MeVisLab* is een research en prototyping platform ontwikkeld door MeVis Research GmbH in Bremen, Duitsland. Het is de opvolger van het beeldverwerkingspakket ILAB4. Bij *MeVisLab* ligt de focus op medische beeldvorming en -verwerking. Het laat een snelle integratie en testen van nieuwe algoritmes toe en bevat een heel aantal standaard beeldverwerkingsmodules. Daarnaast bevat het ook meer geavanceerde tools voor medische registratie, segmentatie, volumetrie en kwantitatieve morfologische en functionele analyse. Onderstaande lijst geeft een overzicht van de belangrijkste attributen van *MeVisLab*:

- Basis beeldverwerkingsalgoritmes en geavanceerde medische beeldverwerkingsmodules
- Een flexibele 2D en 3D visualisatie met interactietools
- Een uitbreidbare C++ beeldverwerkingsbibliotheek
- Object-georiënteerde GUI definitie en scripting
- Ondersteuning voor DICOM-formaat en PACS integratie
- Intuïtieve gebruikersinterface
- Genereren van filmpjes en beelden voor presentatiedoeleinden

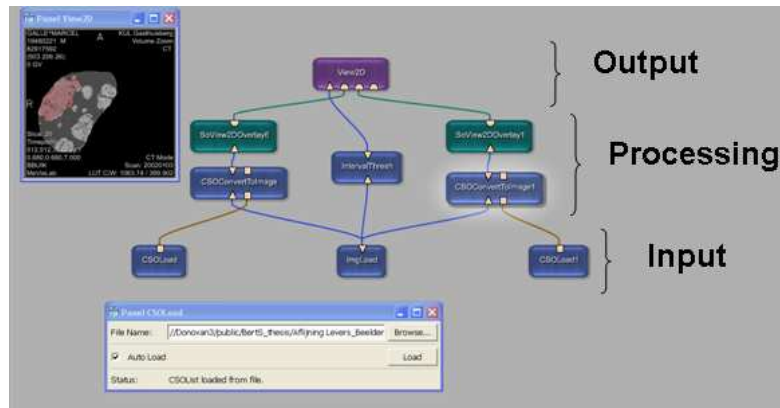
D.1 Opbouw en structuur

MeVisLab is een omgeving voor visueel programmeren. Een visuele programmeertaal (VPL) is een programmeertaal die de gebruiker toelaat programma's te ontwikkelen door de manipulatie van grafische elementen in plaats van textueel te werken. De meeste visuele programmeertalen zijn gebaseerd op het principe van blokken en pijlen die de flow van het programma specificeren. Dit is ook het geval bij *MeVisLab*. Andere voorbeelden van visuele programmeertalen zijn Simulink en G(LabVIEW).

In *MeVisLab* worden individuele bouwblokken of *modules* voor beeldverwerking of visualisatie aan elkaar verbonden om complexe netwerken te vormen. Eigen algoritmes kunnen eenvoudig in C++ geïmplementeerd worden. Voor de grafische gebruikersinterface wordt gebruik gemaakt van een eigen Module Definition Language. Javascript of Pythoncomponenten kunnen toegevoegd worden zowel op het netwerk niveau als op het niveau van de gebruikersinterface.

Een voorbeeld van een netwerk is weergegeven in Figuur D.1. Hierin zijn ook de verschillende types modules weergegeven. De paarse modules zijn Macro Modules. Dit houdt in dat zij sub-netwerken van andere modules bevatten. Ze vereisen geen C++ codering en hebben een eigen gebruikersinterface. De blauwe modules zijn ML Modules. Deze bevatten de eigenlijke beeldverwerkingsalgoritmes en worden in C++ geschreven. De groene modules zijn Inventor Modules.

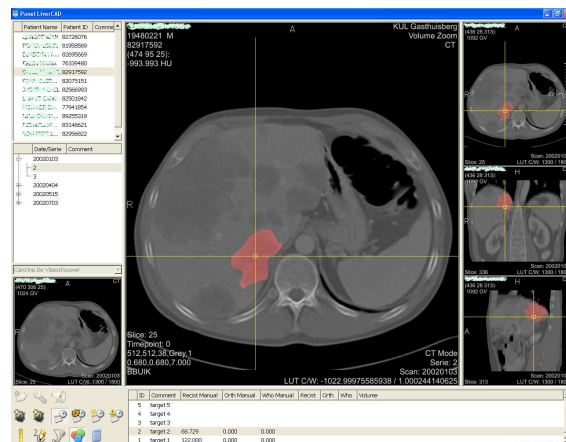
Deze modules worden specifiek voor visualisatie gebruikt. Ze ondersteunen Open InventorTM [59], een C++ object-georiënteerde API ontwikkeld door Silicon Graphics Inc, en laten een snelle visualisatie toe.



Figuur D.1: Voorbeeld van een netwerk van modules in MeVisLab

D.2 LiverCAD

Het oorspronkelijke LiverCAD-programma is een MeVisLabapplicatie ontwikkeld door ir. Bart De Dobbelaer. Het programma dient aanvankelijk als grafische hulptool voor de afdeling Radiologie. Met behulp van de applicatie kan de geneesheer/radioloog de levermetastasen manueel aflijnen en bepaalde afmetingen zoals RECIST en de loodrechte op RECIST aanduiden. Deze afmetingen maken een schatting mogelijk van de tumorgrootte. Hierop kan de oncoloog zich baseren om na te gaan of een behandeling aanslaat bij een patiënt. Op Figuur D.2 is de grafische interface weergegeven van LiverCAD.



Figuur D.2: GUI LiverCAD

Binnen deze thesis bestaat de opdracht uit het automatiseren van de segmentatie aan de hand van één of meerdere methodes. Hierbij kan al dan niet gebruik gemaakt worden van RECIST.

D.3 Ontwikkelde modules

Tijdens het verloop van deze thesis zijn een aantal Macro- en MLModules ontwikkeld om de segmentatiemethodes de implementeren in *MeVisLab* en om de validatie deskundig uit te voeren. In onderstaande secties wordt een kort overzicht gegeven van de ontwikkelde modules samen met hun functie.

D.3.1 Ontwikkelde Macro Modules

CircleScanningTechnique1 voert heen- en terugtransformatie van de Circle Scanning Methode uit, waarbij er bij de minimaal kostpadberekening geen extra kost wordt toegekend aan intensiteiten. De methode vergt een ingangsbeeld, een lijst met punten (XMarkerList) die de zaadpunten voorstellen en bevat een veld waar de maximale straal kan ingesteld worden.

CircleScanningTechnique2 voert heen- en terugtransformatie van de Circle Scanning Methode uit, waarbij er bij de minimaal kostpadberekening wel extra kost wordt toegekend aan intensiteiten.

CompareDelineation vergelijkt 2 handmatige aflijningen met elkaar, gebruikt om interobservatorvariabiliteit op te meten.

LevelSetsWithCircleScan voert een segmentatie uit op basis van een level set methode waarbij de initiële level set geleverd wordt door een Circle Scanning Technique segmentatie. Voor het snelheidsveld zijn er verschillende mogelijkheden, allen op basis van statistische pixelclassificatie. Deze methode vergt de zelfde ingangen als CircleScanningTechnique1.

LiverCADValidation valideert de testdata, GUI gebaseerd op LiverCAD

LiverCADMICCAIChallenge voert de segmentatie en de validatie uit voor de data van de MICCAI Challenge.

PixelClassificationWithRegionGrowing voert een 2D region growing methode uit op een beeld bekomen door statistische pixelclassificatie. Deze methode vergt een ingangsbeeld en een lijst van zaadpunten (in een XMarkerList).

PropagateAndRegionGrowing propageert een zaadpunt naar de volgende snede en voert er een region growing operatie op uit.

RegionGrowingMeth1 is een 2D region growing-methode met twee zaadpunten waar de tusseliggende zaadpunten door lineaire interpolatie gevonden worden.

RegionGrowingMeth2 is een 2D region growing-methode met één zaadpunt per snede.

RegionGrowingMeth3 is een 3D region growing-methode met één zaadpunt.

RegionGrowingMeth4 is een 2D region growing-methode met een trapezoïdaal masker. Ze vereist een ingangsbeeld en twee vectoren (RECIST en de orthogonale hierop) in wereldcoördinaten.

RegionGrowingMeth5b is een 2D region growing-methode met een cirkelvormig masker. Ze vereist een ingangsbeeld en twee vectoren (RECIST en de orthogonale hierop) in wereldcoördinaten.

SpiralScanMacro voert de een transformatie van 3D naar 2D en terug uit, waarbij in de 2D beeldruimte een segmentatie plaatsvindt d.m.v. een minimaal kostpadberekening. De methode vergt een ingangsbeeld, een lijst met één punt (XMarkerList) die het zaadpunt voorstelt en bevat een veld waar de maximale straal kan ingesteld worden.

Validate valideert een methode door de binaire segmentatie te vergelijken met de een binair masker dat gevormd wordt uit de handmatige aflijningen.

D.3.2 Ontwikkelde ML Modules

CalculateCenter berekent het zwaartepunt van een beeld (rekening houdend met de beeldintensiteiten). In het geval van een binair gesegmenteerd beeld is dit zwaartepunt het middelpunt van de segmentatie.

CalculateCenterOnEachSlice berekent op elke snede het zwaartepunt.

CircleBackTransformation berekent volgens met lineaire interpolatie het de segmentatie in het oorspronkelijke beeld. Deze functie verwacht een segmentatie in een 2D polair assenstelsel en transformeert deze segmentatie terug naar een 2D cartesisch assenstelsel. (zie 4.5)

CircleScan transformeert een beeld van een cartesische naar een polair assenstelsel door de beeldruimte vanuit een centrum cirkelvormig af te scannen. (zie 4.5)

Cooccurrence berekent de cooccurrence-matrix op basis van een beeld dat maximaal 256 intensiteiten mag hebben. Het is mogelijk één van de 13 mogelijke richtingen kiezen en om een afstand te kiezen waarover de cooccurrence-matrix berekend wordt.

ExtractDiscregion berekent op basis van RECIST en de loodrechte erop (2 vectoren in voxelcoördinaten) twee schijfvormige maskers. Het eerste masker is binair en bevindt zich volledig in de snede van RECIST en de loodrechte daarop. Het is de kleinste cirkel met het snijpunt van RECIST en de ortho als middelpunt en straal het dichtstbijzijnde eindpunt van één van de vectoren. Het tweede masker strekt zich uit over alle snedes en is schijfvormig volgens de grootste cirkel (van middelpunt tot verst gelegen eindpunt van de vector).

ExtractFeature berekent op basis van een intensiteitenbeeld het contrast, de correlatie, de variantie en de entropie voor elke voxel door eerst de cooccurrence matrix te berekenen voor een $3 \times 3 \times 1$ kernel en daarvan de voorgenoemde Haralick textuurkenmerken te berekenen.

ExtractFeatures berekent de 10 Haralick Features gegeven in Tabel 3.2 op basis van het beeld dat door de Cooccurrence module berekend wordt.

ExtractQuadrangleRegion is volledig gelijkaardig aan de vorige methode maar met een trapezoïdaal masker als kleinste masker door alle eindpunten en een ruitvormig grootste masker met zijden evenwijdig aan RECIST en de orthogonale hierop. Het vullen van de vierhoek gebeurt met een algoritme beschreven in Bijlage C.

InsertSubImg voegt een subimage (deelbeeld) opnieuw in het oorspronkelijke beeld.

MCPwithIntesityCost is een uitbreiding op MinimumCostPath waarbij ook de intensiteiten voor het beeld mede de kost bepalen.

MinimumCostPath is een dynamisch programmeringsalgoritme dat het minimale kostpad berekend op basis van een gradiënt beeld.

PixelClassificationBasedOnIntensitiesEM is een niet gesuperviseerd statistisch pixelclassificatie-algoritme enkel op basis van intensiteiten waarbij het schatten van de modelparameters gebeurt met een Expectation Maximization algoritme.

PixelClassifier is een gesuperviseerd statistisch pixelclassificatie-algoritme op basis van intensiteiten en textuurkenmerken.

PixelClassifierBasedOnIntensities is een gesuperviseerd statistisch pixelclassificatie-algoritme enkel op basis van intensiteiten.

PixelClassifierEM is een niet gesuperviseerd statistisch pixelclassificatie-algoritme op basis van intensiteiten en textuurkenmerken waarbij het schatten van de modelparameters gebeurt met een Expectation Maximization algoritme.

SpiralBackTransformation is gelijkaardig aan CircleBackTransformation. Hier wordt voor elke voxel van het uitgangsbeeld met het nearest neighbor principe de dichtstbijzijnde voxel in het ingangsbeeld gezocht.

SpiralBackTransformationEq doet hetzelfde als SpiralBackTransformation maar voor een SpiralScanEq met equidistante scanlijnen.

SpiralScan scant spiraalvormig (parametrisatie: zie tekst) het beeld af vanuit punt.

SpiralScanB doet hetzelfde als SpiralScan maar gebruikt interpolatie i.p.v. afronding.

SpiralScanEq doet hetzelfde als SpiralScanB maar gebruikt equidistante scanlijnen.

Bijlage E

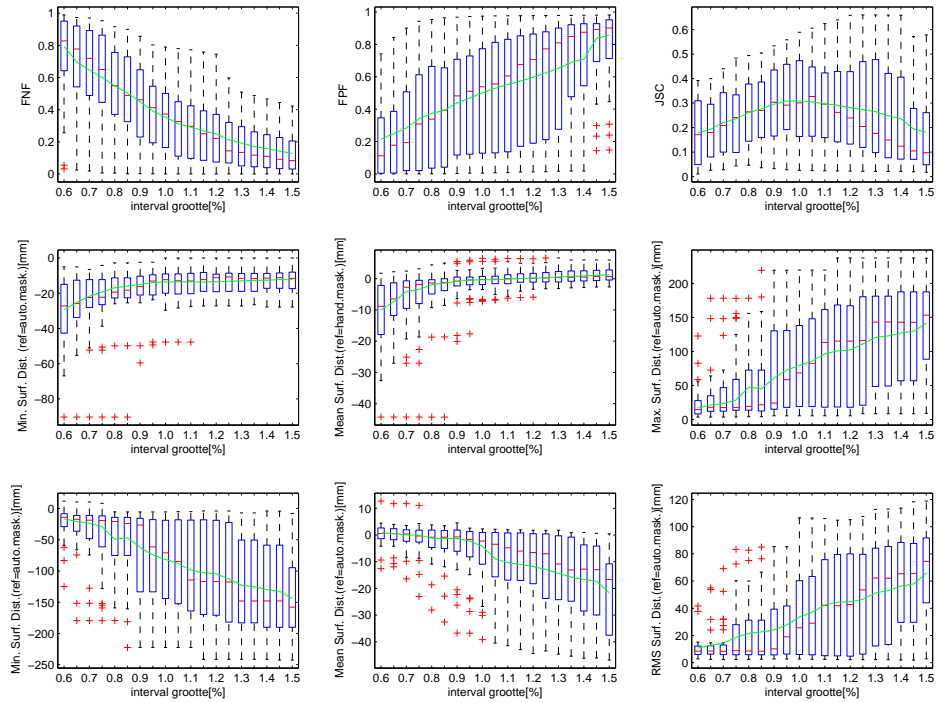
Validatie

E.1 Extra figuren bij gevalideerde methodes

In onderstaande figuren zijn grafieken weergegeven die de overige validatiematen bevatten die niet in de thesistekst vermeld worden. Deze zijn gerangschikt per methode. De overige validatiematen zijn:

- FNF
- FPF
- JSC
- Minimal Symmetric Absolute Surface Distance met de handmatige afijning als referentie
- AASD met de handmatige afijning als referentie
- MASD met de automatische afijning als referentie
- Minimal Symmetric Absolute Surface Distance met de automatische afijning als referentie
- AASD met de automatische afijning als referentie
- RMSSD met de automatische afijning als referentie

Region Growing met één seed per snede

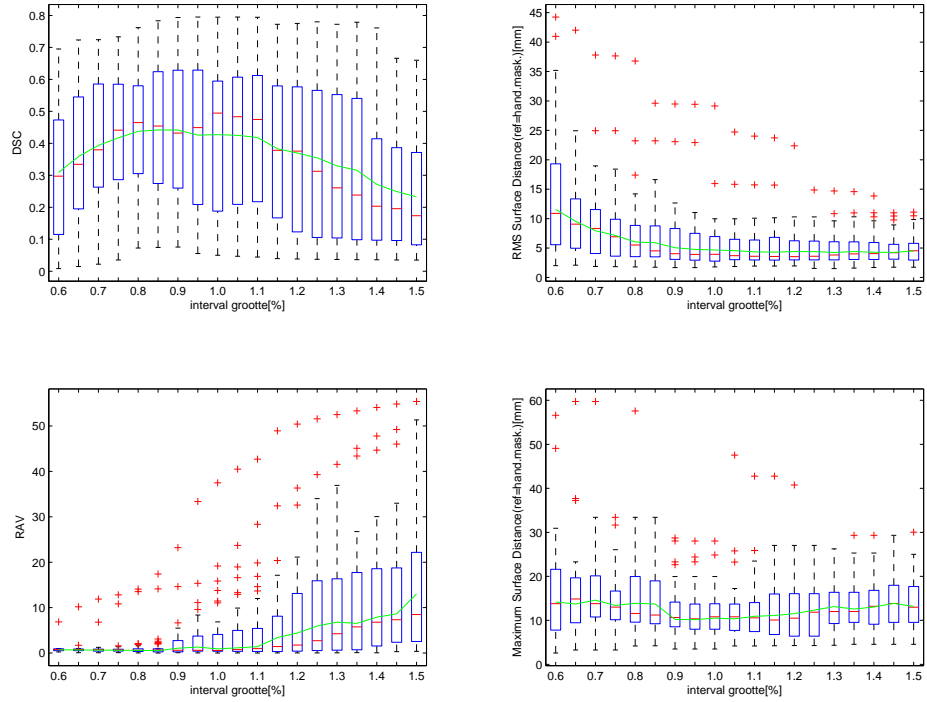


Figuur E.1: Overige validatiematen voor Region Growing met één seed per slice

Region Growing met twee seeds per metastase

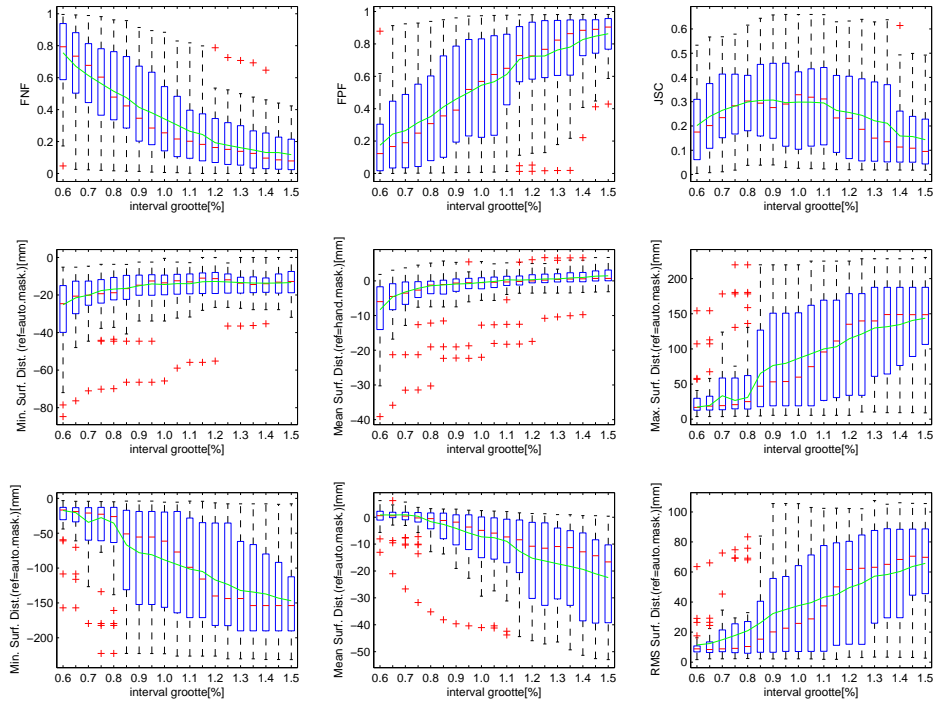
Trainingsdata

Als testwaarden om de optimale intervalgrootte te bepalen, wordt opnieuw het interval $[0,6\%–1,5\%]$ gekozen met tussenstappen $0,05\%$. Er is gekozen om dezelfde validatiematen als in de vorige paragraaf te beschouwen: de DSC, de RMSSD, de RAV en de MASD. In Figuur E.2 zijn deze 4 maten weergegeven in functie van de intervalgrootte. Als optimale waarden krijgen we hier: DSC bij 1% , RMSSD bij $1,2\%$, RAV bij $0,8\%$ en MASD bij $1,15\%$. Het globale optimum om de testdata te valideren, is dan $1,05\%$.



Figuur E.2: De DSC, RMSSD, RAV en MASD in functie van de intervalgrootte voor Region Growing met twee seeds

In Figuur E.2 is op te merken dat de resultaten meer outliers bevatten dan in Figuur 5.6 wat erop wijst dat minder seeds enerzijds het gebruikersgemak verhogen, maar anderzijds een verminderde prestatie teweegbrengen. Verder kunnen dezelfde opmerkingen gemaakt worden over de validatiematen als bij de voorbije methode. De overige validatiematen zijn terug te vinden in Figuur E.3

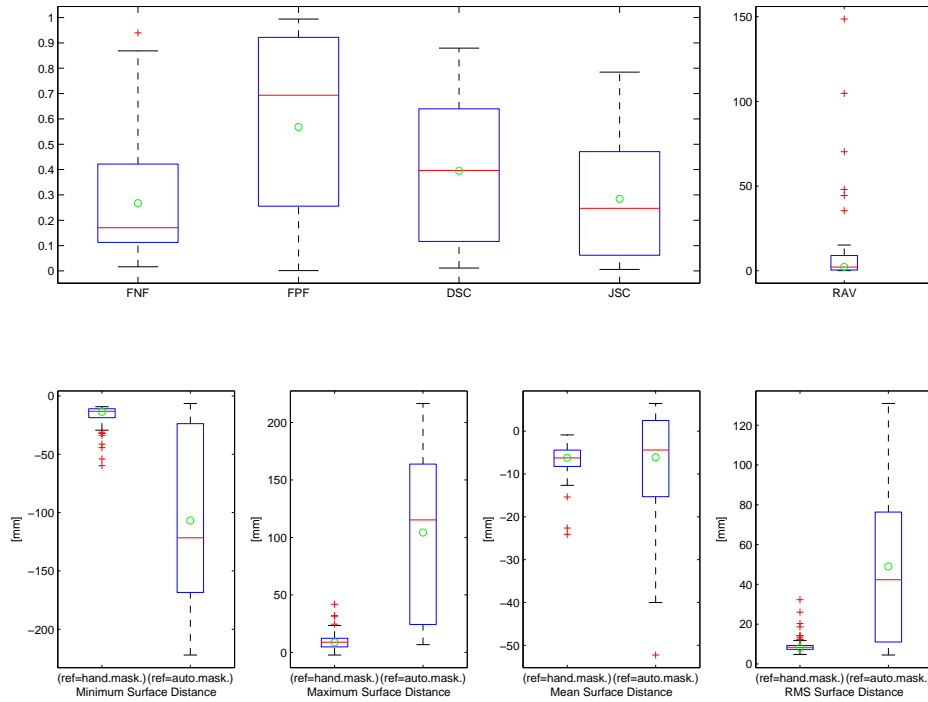


Figuur E.3: Overige validatiematen voor Region Growing met 2 seeds per metastase

Testdata

De resultaten voor de testdata zijn weergegeven in Figuur E.4. Ook hier zijn de verschillen tussen de trainingsdata en de testdata beperkt. De DSC heeft voor de trainingsdata 42% tegenover 39% bij de testdata. Voor de mediaan is dit respectievelijk 47% en 39%. De verschillen zijn klein maar de grote interkwartielafstand wijst erop dat de methode erg onregelmatig presteert.

Ook de overige maten geven ongeveer dezelfde resultaten als de vorige methode. Dit is natuurlijk niet onverwacht aangezien we hetzelfde algoritme aanwenden, enkel nu zijn de tussenliggende seeds automatisch gegenereerd.

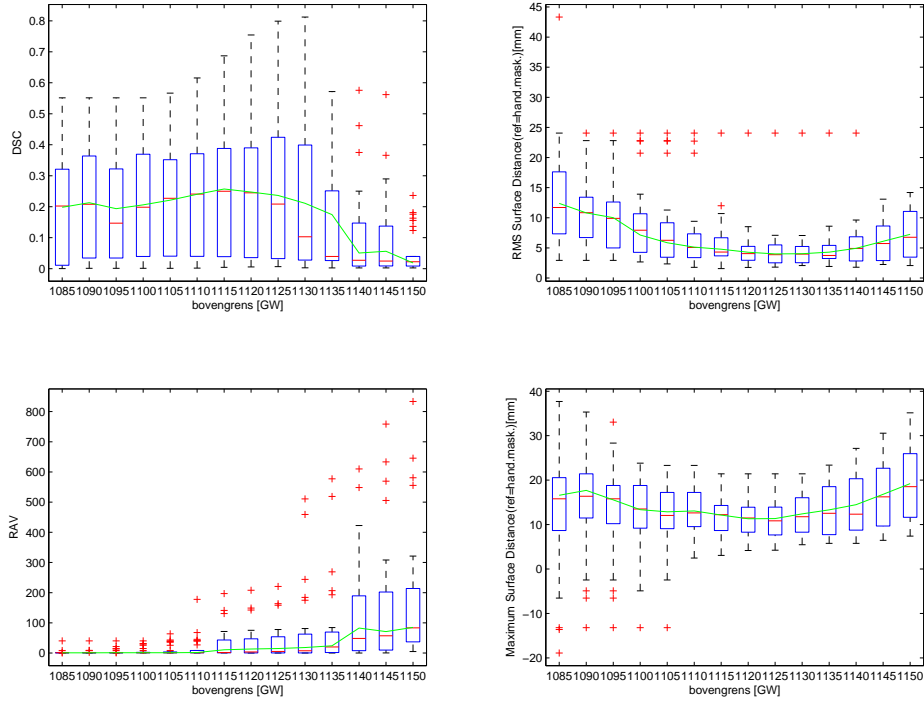


Figuur E.4: Resultaten van de 13 validatiematen voor de testdata van Region Growing met twee seeds per metastase, interval = 1.05%

Region Growing met één seed per metastase

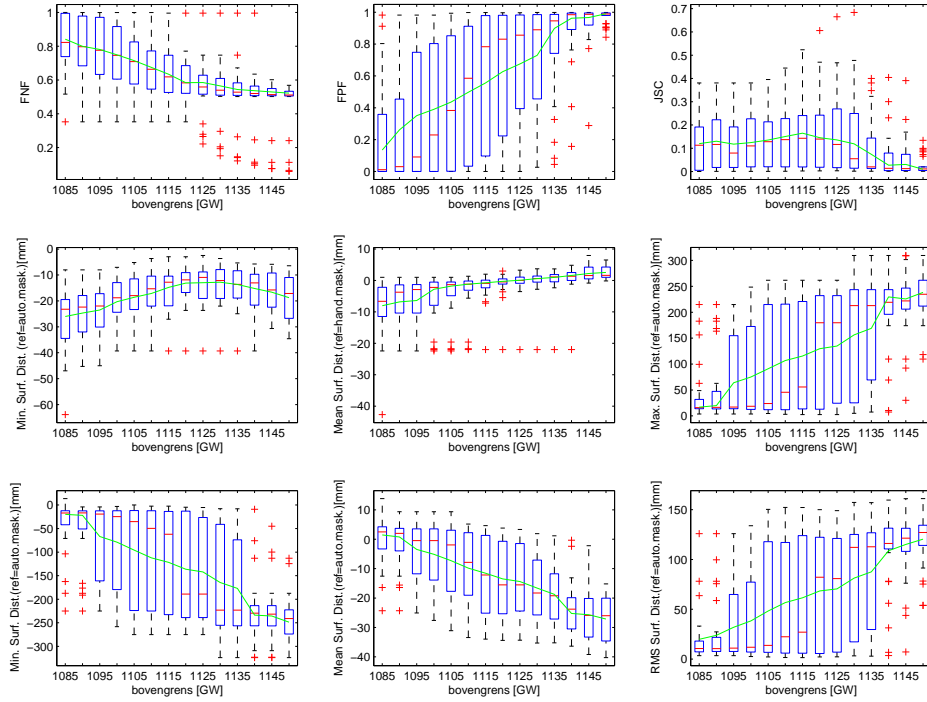
Trainingsdata

De grootte van het grijswaarde interval wordt anders bepaald dan in de voorgaande methodes. Aangezien enkel uitgroei zal optreden indien alle omliggende pixels ook binnen het grijswaarden-interval liggen, wordt de ondergrens van het interval vastgelegd op 1010 GW. Deze waarde is zodanig gekozen dat initieel altijd groei zal optreden binnen de metastase zonder dat te veel ander weefsel wordt mee gesegmenteerd (trade-off). De bovengrens laten we variëren van 1085 GW tot 1150 GW in stappen van 5 GW. Het resultaat van de validatie is weergegeven in Figuur E.5. De verschillende validatiematen zijn weergegeven in boxplots voor de verschillende waarden van de bovengrens.



Figuur E.5: De DSC, RMSSD, RAV en MASD in functie van de intervalgrootte voor Region Growing met één seed per metastase

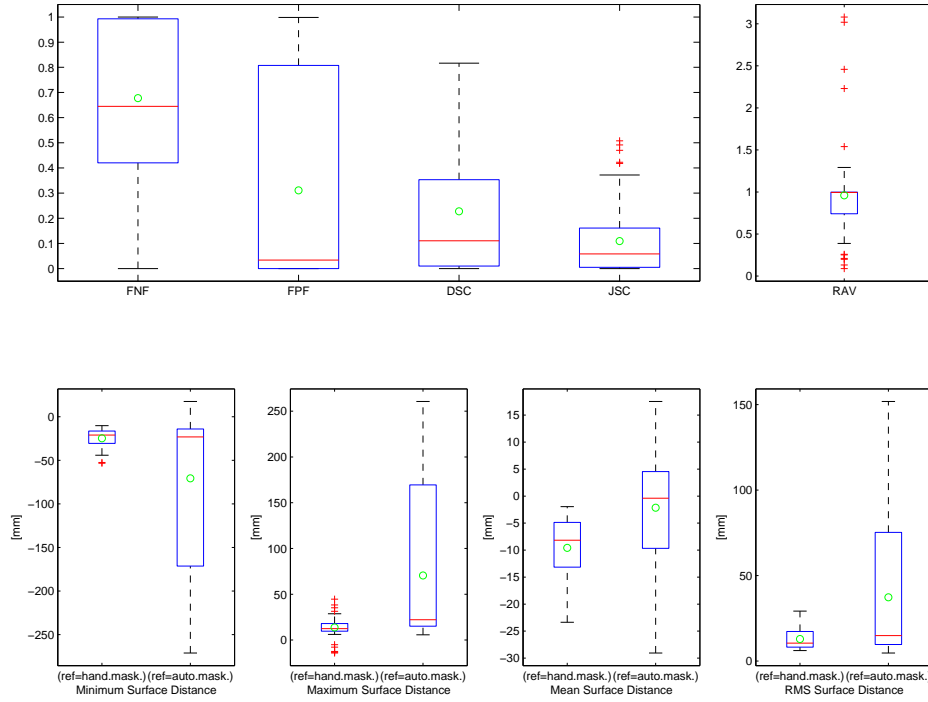
De resultaten zijn beduidend slechter dan bij de voorgaande methodes. Voor een 3D-region growing is de kans op uitgroei groter en dit wordt ook gereflecteerd in de resultaten. We zien echter toch duidelijk optima. Voor de DSC is de optimale bovengrens 1115 GW, voor de RMSSD 1125 GW, voor de RAV 1010 GW en voor de MASD 1020GW. We kiezen daarom als optimale waarde voor de bovengrens 1115 GW. De overige maten zijn gegeven in Figuur E.3.



Figuur E.6: Overige validatiematen voor Region Growing met één seed per metastase

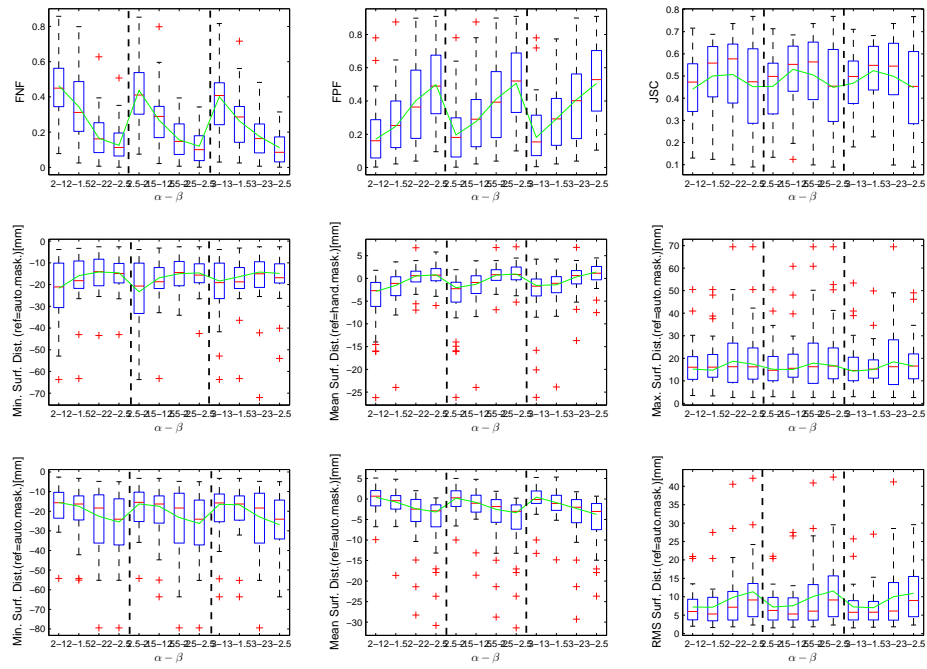
Testdata

De resultaten voor de testdata zijn weergegeven in Figuur E.7. De DSC ligt gemiddeld rond 10%. Het verschil met de trainingsdata (DSC gemiddeld 25%) is toch wel significant.



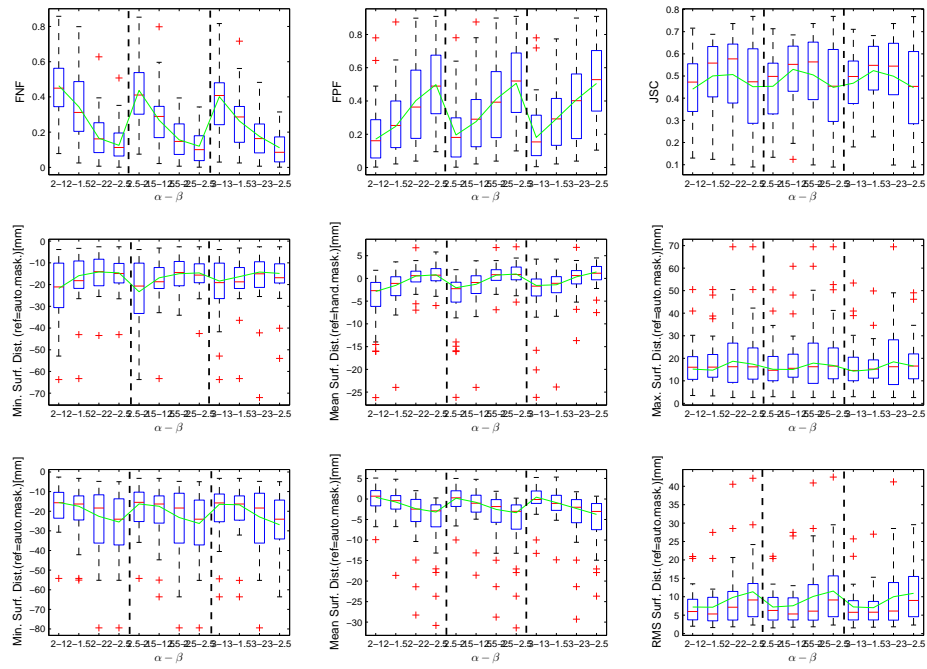
Figuur E.7: Resultaten van de 13 validatiematen voor de testdata van Region Growing met één seed per metastase, bovengrens = 1115 GW

Region Growing op basis van RECIST en de loodrechte hierop (Cirkel)



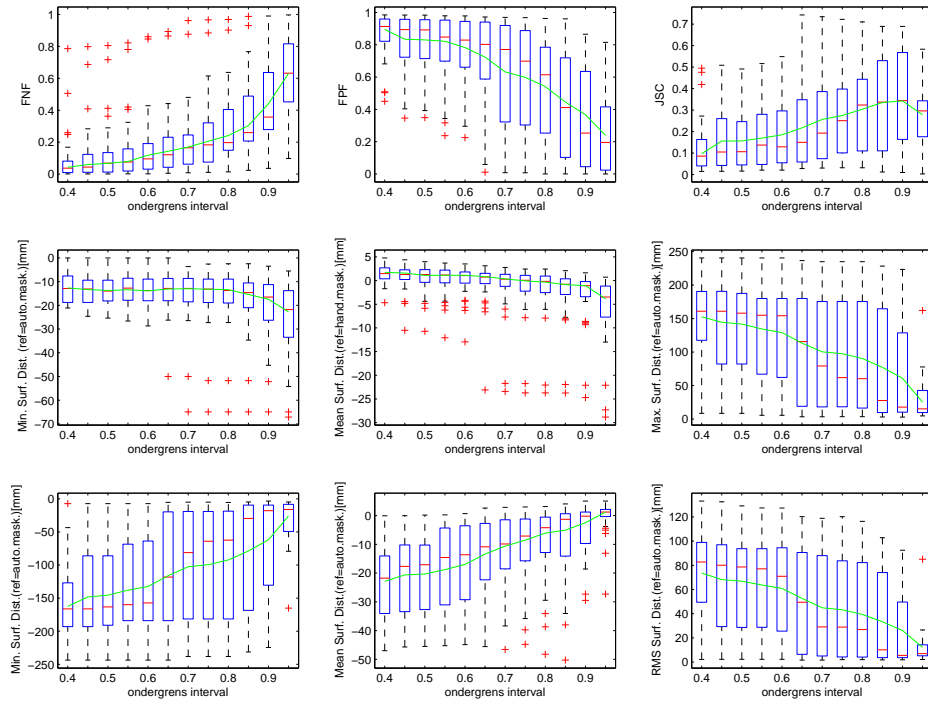
Figuur E.8: Overige validatiematen voor Region Growing met een cirkel als restrictie

Region Growing op basis van RECIST en de loodrechte hierop (Rechthoek)



Figuur E.9: Overige validatiematen voor Region Growing met een rechthoek als restrictie

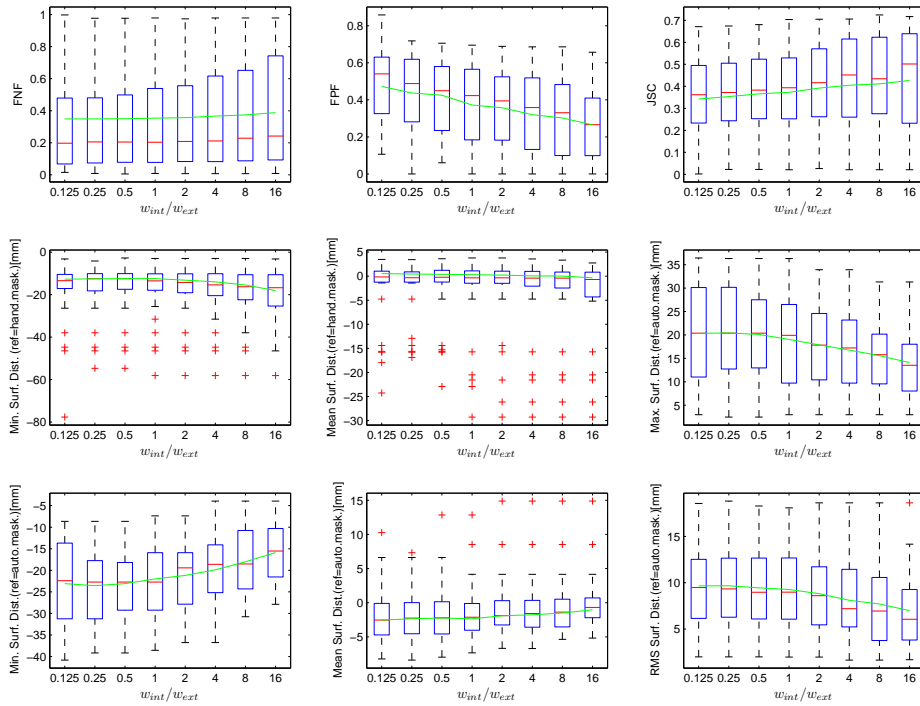
Region Growing gegeven een fuzzy classificatiebeeld



Figuur E.10: Overige validatiematen voor Region Growing van een fuzzy classificatiebeeld

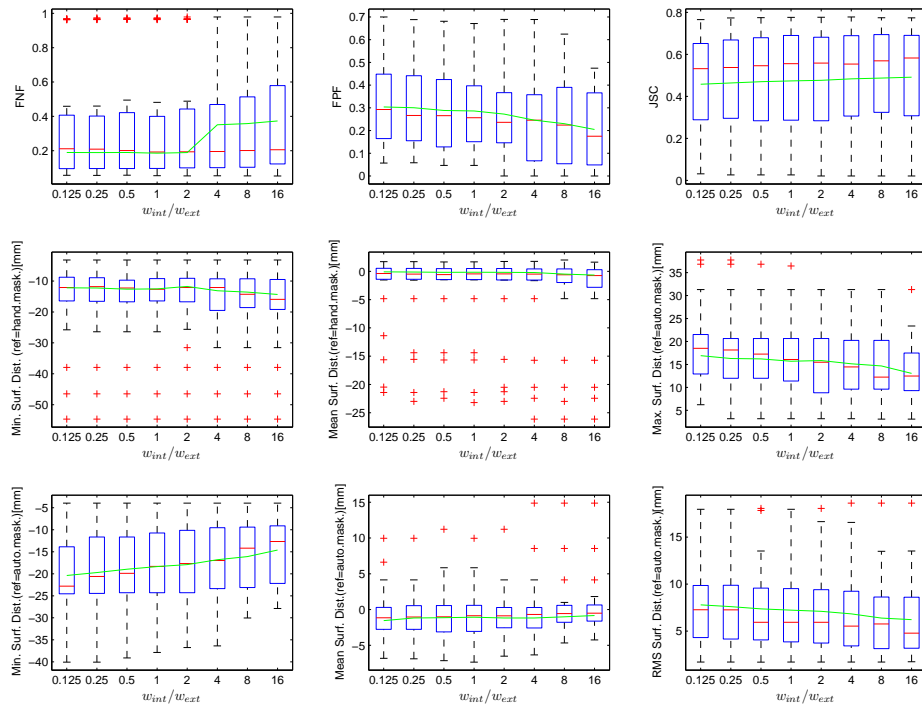
Spiral Scanning Technique

Volledig Beeld



Figuur E.11: Overige validatiematen voor Spiral Scanning

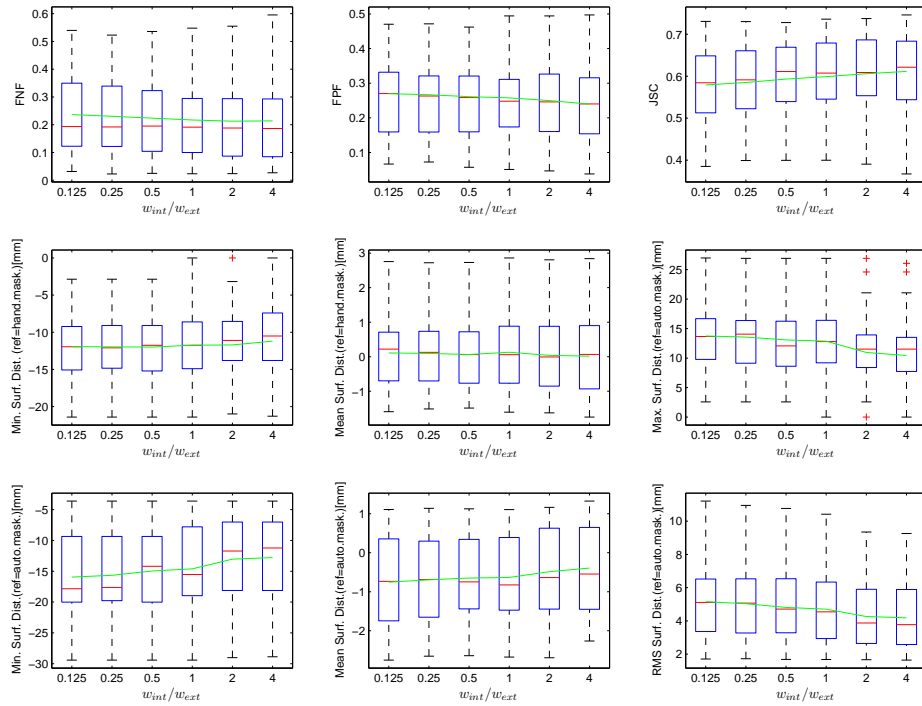
Afgelijnde Lever



Figuur E.12: Overige validatiematen voor Spiral Scanning met een afgelijnde lever

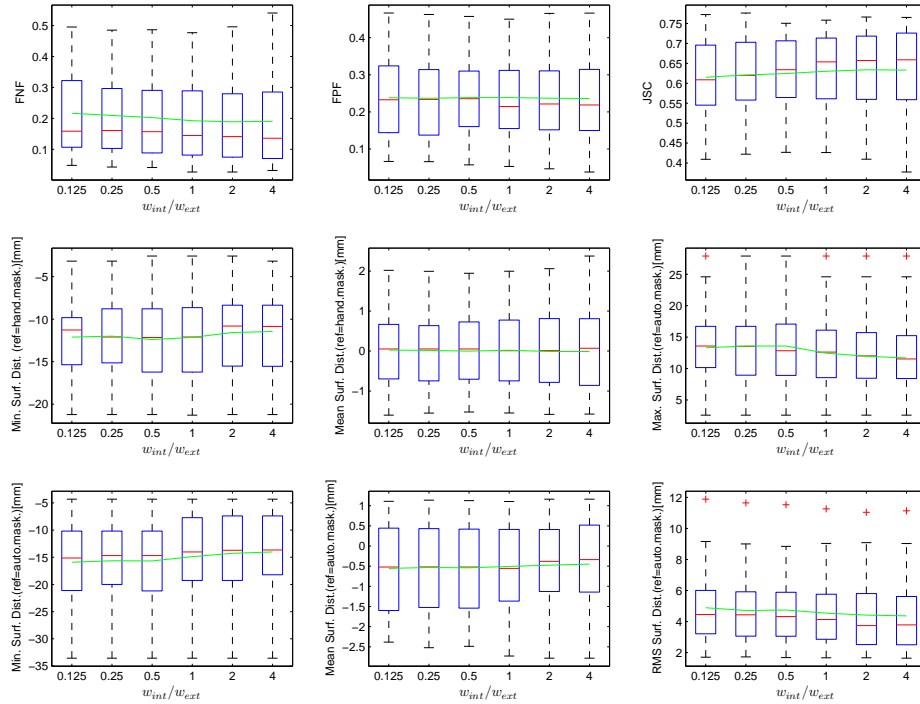
Circle Scanning Technique

Volledig Beeld



Figuur E.13: Overige validatiematen voor Circle Scanning

Afgelijnde Lever



Figuur E.14: Overige validatiematen voor Circle Scanning met een afgelijnde lever

Level Sets met Circle Scan

Volledig beeld

Afstellen van de parameters

Het afstellen van de advectieterm, de krommingsterm en de propagatieterm.

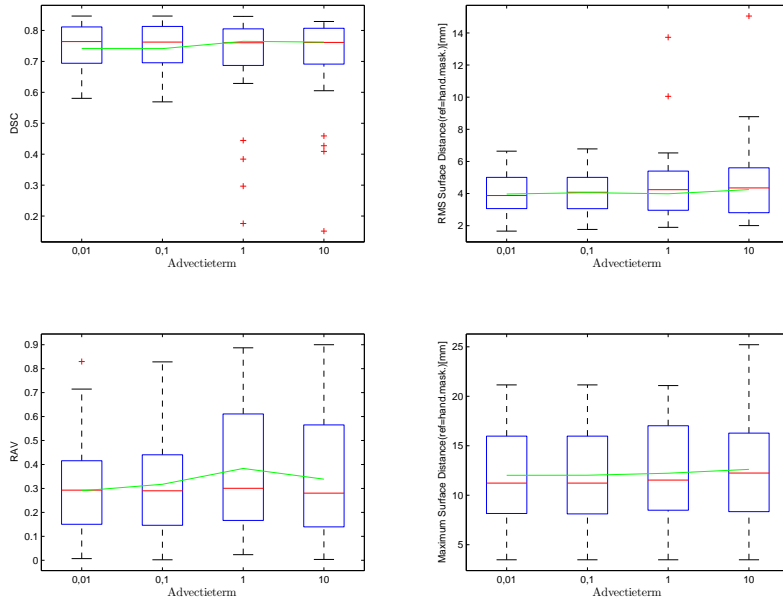
Overzicht van enkele validatiematen

E.2 De inter-observatorvariabiliteit: Vergelijking van twee studies

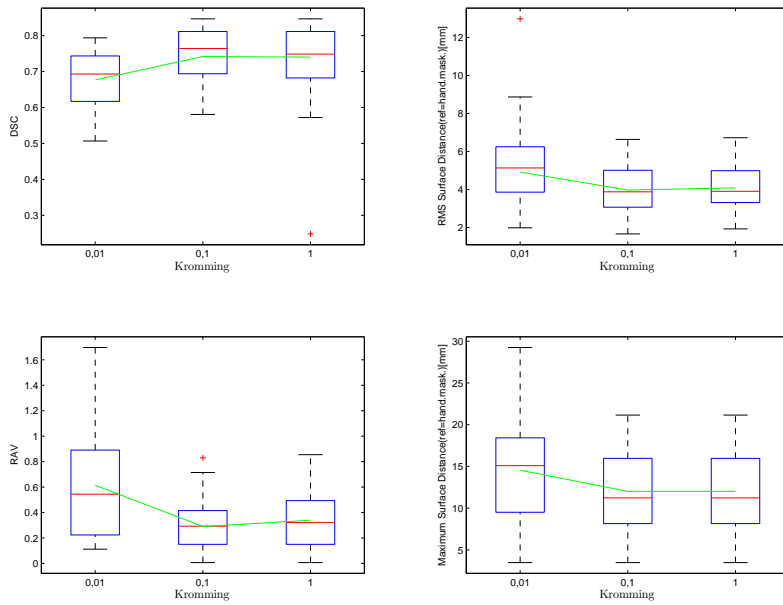
In deze bijlage is het resultaat van een studie rond inter-observatorvariabiliteit die we zelf hebben uitgevoerd, vergeleken met een studie uit de literatuur[12]. In de eigen studie is geprobeerd dezelfde opdeling te maken als in de literatuur. We delen de tumoren op volgens grootte en per groep wordt

	FNF	FPF'	JSC
Circle Scan	27,5	16,5	62,3
Level Sets	20,8	19,6	64,3
Level Sets (EM)	20,0	20,3	64,5

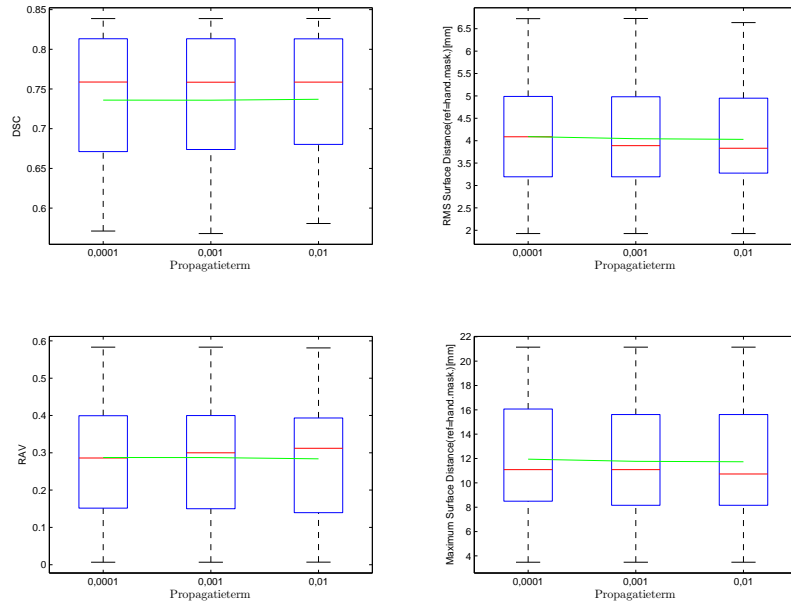
Tabel E.1: Andere validatiematen voor een afgelijnde lever



Figuur E.15: De validatiematen i.f.v. de wegingsfactor voor de advectieterm

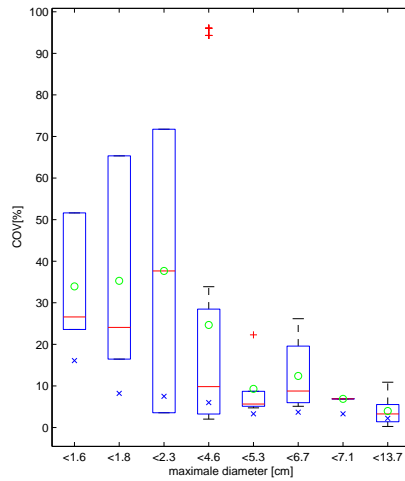


Figuur E.16: De validatiematen i.f.v. de wegingsfactor voor de krommingsterm



Figuur E.17: De validatiematen i.f.v. de wegingsfactor voor de propagatieterm

een boxplot gemaakt met daarop de gemiddelde COV van het volume volgens de eigen studie(\circ) en het gemiddelde volgens de literatuur(\times). Dit is weergegeven in Figuur E.18.



Figuur E.18: De COV van het volume gegroepeerd per grootte van de tumor. \circ =gemiddelde van de eigen studie, \times =gemiddelde uit de literatuur

De figuur toont duidelijk dat de aflijningen uit de eigen studie minder goed overeenstemmen met elkaar dan verwacht wordt volgens de studie. Er is wel dezelfde trend aanwezig als in de studie. Naarmate de tumoren een grotere diameter hebben, wordt de COV kleiner.

Bijlage F

Artikel MICCAI: Liver Tumor Segmentation Challenge

Deze bijlage bevat het artikel, opgestuurd naar “3D Segmentation in the Clinic: A Grand Challenge II” [46]. Het besproken algoritme haalde ter plaatse de eerste prijs in de categorie van de algoritmes met minimale gebruikersinteractie.

Segmentation of Liver Metastases Using a Level Set Method with Spiral-Scanning Technique and Supervised Fuzzy Pixel Classification

Abstract

In this paper a specific method is presented to facilitate the semi-automatic segmentation of liver metastases in CT images. Accurate and reliable segmentation of tumors is e.g. essential for the follow-up of cancer treatment. The core of the algorithm is a level set function. The initialization is provided by a spiral-scanning technique based on dynamic programming. The level set evolves according to a speed image that is the result of a statistical pixel classification algorithm with supervised learning. This method is tested on CT images of the abdomen and compared with manual delineations of liver tumors. The results show that the accuracy of the method does not reach that of the manual segmentation. The average overlap error is 34.6% and the average volume difference is 17.8%. The average, the RMS and the maximum surface are respectively 2.0, 2.7 and 10.1 mm.

F.1 Introduction

Segmentation is an image processing operation to distinguish an anatomical structure from the surrounding tissue. Tumor segmentation is an important issue for cancer follow-up, where the oncologist is interested to evaluate the change in size of the tumors. Early response prediction allows the oncologist to adapt the therapy, which can lead to a higher survival rate [25]. Measuring the response of a treatment can be done by uni-, bi- or tridimensional criteria. Clinical research [25, 26, 27] indicates that volume measurements (3D) give the best reflection of the tumor response. Volume measurements require the segmentation of the tumors, which is very time consuming when done manually. Moreover, the volume of manual delineations is subject to intra- and interobserver

variability, which is estimated at about 8% for liver tumors [15]. Therefore, automatic or semi-automatic tumor delineation algorithms are required.

This article focuses on the segmentation of liver tumors in contrast-enhanced CT images. Because tumors generally have different shapes and intensities, the segmentation is not straightforward. The gray values of a tumor depend on the delay between the contrast injection and the image acquisition, the contrast dose and the patient physiology. In general, liver tumors have a more or less round shape. However, the shape can be influenced by the vicinity of blood vessels, the edges of anatomical structures, the type of cancer,...

Several semi-automatic methods are already available for the segmentation of CT liver tumor segmentation. These include methods based on watershed like the paintbrush algorithm [14, 13], deformable models like the active contour algorithm in [21] and a region growing technique using several constraints [23]. These methods have in common that they all require an extensive amount of user interaction.

The segmentation algorithm described in this paper combines several techniques including a spiral scanning technique, pixel classification and level sets. The spiral scanning technique is based on [24], where it is implemented for lung nodules. Here the implementation is performed on liver tumor segmentation using a different parametrization and an additional gradient to enhance the tumor edges. The pixel classification and level set method are based on the standard implementation for both algorithms. The algorithm is validated on a test data set provided by “3D Segmentation in the Clinic: A Grand Challenge II” [46] using 5 evaluation measures: volumetric overlap, relative absolute volume difference, average symmetric absolute surface distance, symmetric RMS surface distance and maximum symmetric absolute surface distance.

After this introduction, we first give a detailed description of the method. Next, the validation results are given, followed by a discussion and conclusions.

F.2 Method

To start, the algorithm requires the user to place one point approximately in the middle of the tumor and to specify a maximal radius by placing a point in the surrounding liver tissue. This radius is preferably about two times the largest radius of the tumor. Next, this input is used to generate an initialization for the level set using a spiral-scanning technique [24]. The speed image guiding the propagation of the level set is calculated independently using fuzzy classification. The heart of the algorithm is the standard level set function [60]. After level set segmentation, post-processing is used to transform the distance map into a binary image. An overview of the method is given in Fig. F.1.

Initial Segmentation

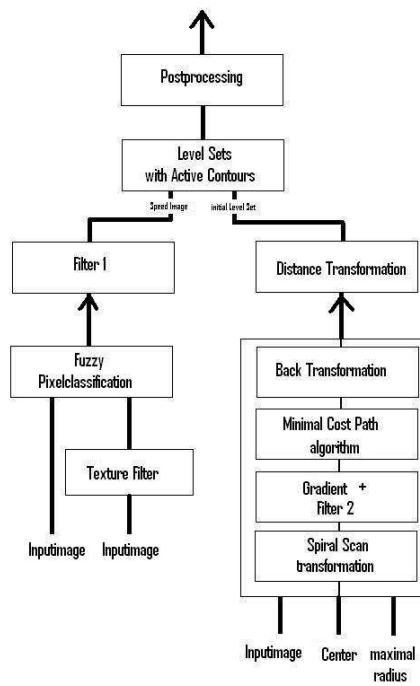
The initial segmentation is provided by a Spiral-Scanning Technique [24], represented by the first part of the right branch in Fig F.1. This technique transforms the 3D image to 2D space using a scanning technique, allowing to find a contour using dynamic programming for the segmentation. The 3D image is sampled along scan lines. The starting points of all scan lines coincide at the center of the tumor, as indicated by the user. The endpoints form a spiral on the surface of a sphere starting at the north pole and ending at the south pole. Our parametrization of this spiral is given by

$$x = x_C + \rho \sin(t) \cos(2Nt) \tag{F.1}$$

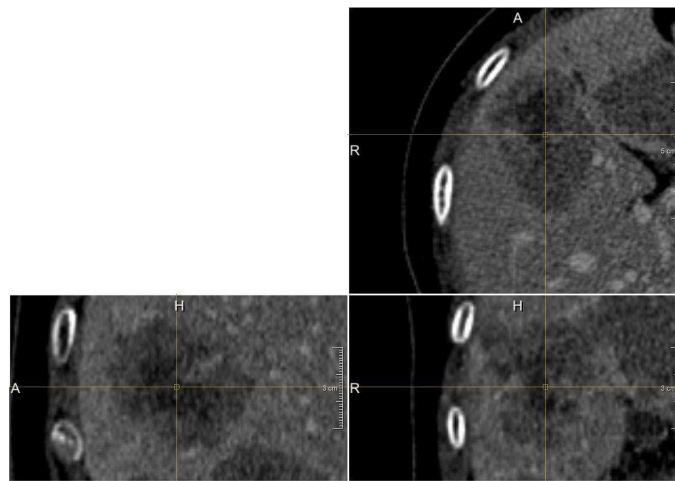
$$y = y_C + \rho \sin(t) \sin(2Nt) \tag{F.2}$$

$$z = z_C + \rho \cos(t) \tag{F.3}$$

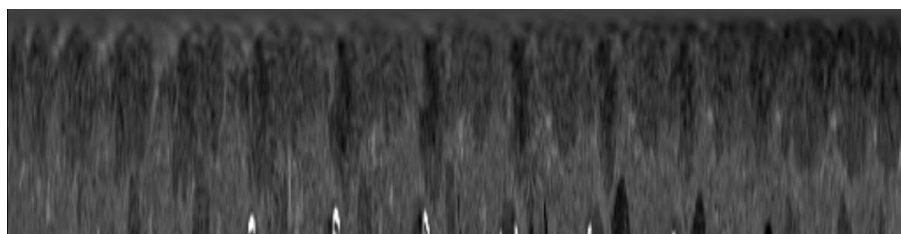
with $t = \frac{\pi}{2} + \arcsin(\frac{2k_1}{2N^2} - 1)$ en $\rho = \frac{\rho_{max} k_2}{M}$. N is the number of rotations, $2N^2$ the number of scan lines and M is the number of samples on each scan line, x' and y' are the coordinates of the 2D space. The coordinates of the center of the tumor are (x_C, y_C, z_C) . The result of the spiral-scan transformation is given in Fig. F.3.



Figuur F.1: Block Diagram of Segmentation Algorithm

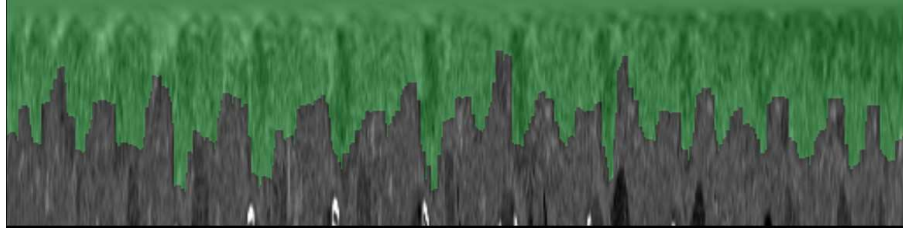


Figuur F.2: Original image with centrally a tumor



Figuur F.3: Scan converted image

In the transformed image, a dynamic programming algorithm performs the actual segmentation. This is done by calculation of the minimal cost path. The sum of internal cost, $E_C^{int}(i) = wt. |y'_i - y'_{i-1}|$, and the external cost, $E_C^{ex}(i) = |\nabla I|_{max} - |\nabla I(x'_i, y'_i)|$, is minimized to achieve a curve which separates the image in a tumor and a non-tumor part. Instead of just using the image gradient for the calculation of the external cost, the gradient of a binary image is added to the image gradient in order to enhance the tumor edges. The binary image is simply calculated by applying an interval thresholding operation. The used upper-threshold is chosen as the average of the estimates of the mean gray value of the tumor and the liver in the image. Those estimates are based on the intensity distribution in a sphere with a radius of 10 voxels around the user click respectively in the center of the tumor and outside the tumor. The result of the dynamic programming algorithm is a 2D binary segmentation, which is overlaid on the scan converted image in Fig. F.4.



Figuur F.4: Binary segmentation overlaid on the scan converted image

The back transformation to 3D uses the nearest neighbor principle. For each voxel lying on a distance $\rho < \rho_{max}$ from the center of the spiral, the algorithm searches the nearest sample on a spiral with radius ρ .

The final step to achieve an initial level set is a distance transformation. For each voxel, the shortest distance to the surface is calculated and stored. Voxels lying within the segmented surface get a negative value. The result is a signed distance function [52].

Speed Function

The left branch in Fig. F.1 results in a speed image that will deform the initial segmentation, obtained in (F.2), to the final result. This speed image is based on supervised fuzzy pixel classification. Fuzzy pixel classification calculates the probability that a certain pixel belongs to a certain image class c_j [9]. Within the liver two classes can be distinguished: liver tissue and tumor tissue. The statistical model uses a normal distribution where the parameters μ and σ are estimated based on the gray values in spheres around the point inside and outside the tumor. To increase reliability, not only the image intensities could be taken into account, but also some texture features [36, 35]. Initial experiments show that correlation and variance could best distinguish liver tissue from tumor tissue, but the computational complexity to calculate the features is too high to allow the use in semi-automatic algorithms. After having calculated the fuzzy image, Filter 1 (see Fig. F.1) transforms the image in a speed image by applying

$$I_V(x) = f_3(p(c_{tumor}|I_k)) = |2 p(c_{tumor}|I_k) - 1|$$

The speed function has high values far from the tumor edges. To reduce noise without eliminating the edges an anisotropic diffusion filter is used to filter the speed image. A speed image is given in Fig.F.5

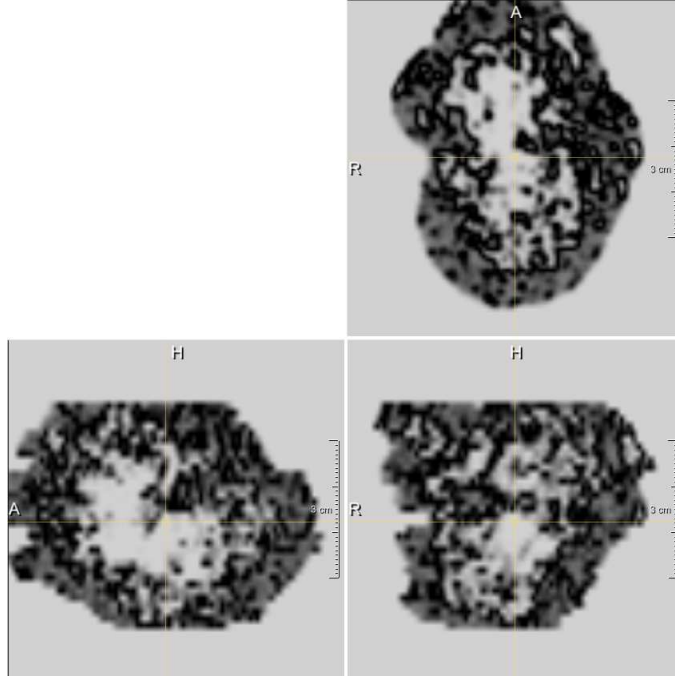


Figure F.5: Speed Image

The speed function is only calculated in a restricted area defined based on the initial segmentation. This can be seen in Fig. F.5.

Level Sets

The initial level set is represented by an implicit function $\phi(x, t)$ and deformed under influence of a speed image. For computational reasons, a local level set method is used, because it defines $\phi(x, t + \delta t)$ only near the zero level set [20]. The partial differential equation which deforms the initial level set, is given by

$$\frac{d}{dt}\phi = -\alpha\mathbf{A}(\mathbf{x}) \cdot \nabla\phi - \beta P(\mathbf{x})|\nabla\phi| + \gamma Z(\mathbf{x})\kappa|\nabla\phi| \quad (\text{F.4})$$

where \mathbf{A} is an advection term, P is a propagation term and Z is a spatial modifier term for the mean curvature κ [40]. The advection term $\mathbf{A}(\mathbf{x}) = -\nabla I_V(\mathbf{x})$ evolves the level set to the edges of the tumor. When the gradient of the speed function is negative, the level set expands. A positive gradient causes the level set to shrink. The propagation term $P(\mathbf{x}) = I_V(\mathbf{x})$ causes a speed dependent expansion of the level set and the curvature term $Z(\mathbf{x}) = I_V(\mathbf{x})$ reduces sharp edges in the level set. The maximum RMS change between different iterations is used to determine when the solution has converged. The value is chosen on 0.01. Further, the number of iterations is limited to avoid infinite loops.

Training

11 data sets of CT images of the abdomen from the university hospitals of Leuven are used for the training of the algorithm. For calculation of the minimal cost path, this training determines the relative weight of the external cost with respect to the internal cost. Besides that, parameter tuning of the weighting factors for the different terms (α, β, γ) for the level set equation (Eq. F.4) is necessary to obtain the optimal results. Afterwards the tuned parameters are checked and further adjusted with the training data sets of “3D Segmentation in the Clinic: A Grand Challenge II” [46].

Evaluation

To evaluate the data, five measures are used: the overlap error (1 - volume overlap), the relative absolute volume difference between manual and semi-automatic segmentation, the average surface distance, the RMS surface distance and the maximum surface distance between the surface of manual and the surface of the semi-automatic segmentation. Those measures are compared with other independent manual delineations by the organization of the challenge [46]. Based on this comparison a score is given to each evaluation measure and to each segmentation. A score of 90% means that the semi-automatic segmentation is as good as the manual delineation.

F.3 Results

The method was developed on a data set of the University Hospitals Leuven (Leuven data), acquired on two single slice and two 4-slice CT scanners with slice thickness of 5-7 mm and an in-plane resolution of 0.68 – 0.82 mm. Next, the parameters were further tuned on the training data provided for the challenge [46]. Finally, the algorithm was evaluated on actual challenge data. The liver tumor data of the challenge were acquired on one 64-slice and two 40-slice CT scanners using a standard four-phase contrast enhanced imaging protocol with slice thickness of 1 mm or 1.5 mm and an in-plane resolution of 0.6 – 0.9 mm.

Leuven data

In previous work, a 2D variant of the algorithm described here is, is developed, because the data sets had a poor axial resolution (slice thickness of 5-7 mm). The parameters for this algorithm are tuned on 11 data sets of the university hospitals of Leuven containing 31 metastases. This tuning reveals that the most relevant parameters are the weight of the external cost with respect to the internal cost in the calculation of the minimal cost path, the weight of the binary image gradient for edge enhancement with respect to the normal gradient in the spiral-scan transformation and the three weighting factors (α, β, γ) in Eq. F.4. Other parameters like the weighting factor of the active contour algorithm (within the level set method), the maximum number of iterations, the maximum RMS error,... have not an important influence on the results. Therefore, those parameters are not finetuned further. Furthermore, several speed functions are tested and an equilibrium is sought for the size of the restricted area in which the speed function is calculated. The larger this area, the better the algorithm is able to correct high surface distance errors, but the higher the risk that the level set algorithm will deform the initial segmentation to a wrong (strong) edge.

The average DSC [43], volume difference, RMS and maximum surface distance¹ for those training data sets are respectively 74.1%, 28.3%, 4.0 and 12.2 mm with a restricted area, calculated by applying a spherical dilation filter with radius 25 on the initial segmentation.

The 2D variant of the algorithm was also validated, which resulted in a DSC of 76.5%, a volume difference of 27.3, a RMS and maximum surface distance of 8.5 and 10.8 mm. Those results were compared with the interobserver variability, which was significantly higher than in [15]. This indicates that the used tumors were rather difficult to segment.

Training data

Afterwards, the results of the semi-automatic liver tumor segmentation of the 10 clinical training data sets of the challenge [46] coming from 4 patients were assessed first visually and then quantitatively. Based on the visual inspection, the parameters of the level set method and the dynamic programming algorithm were further adjusted. To find the optimal parameters mainly the volume overlap was used. An example of a segmentation after volume rendering is shown in Fig. F.6,

¹Here, not the symmetric surface distance is taken, only the distances from the manual segmentation to the semi-automatic segmentation

where the true positives, the false positives and the false negatives are represented respectively by a green, blue and red color.

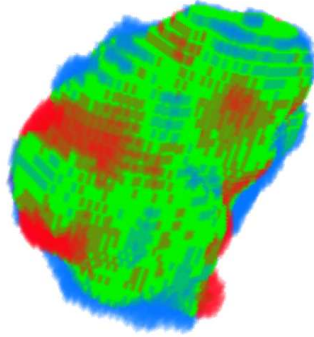


Figure F.6: Example of the segmentation

The quantitative results for the training data are given in Table F.2 and Table F.3 (Appendix F.5). The latter gives the surface distances in mm. There is made a distinction based on which surface is the reference surface. Calculating the total score using the same interobserver variability as the test data, gives a score of $79.1 \pm 9.2\%$. With a score of 92.3% for the average volume difference, this difference is smaller than the interobserver variability. This is not true for the scores of the other evaluation measures: 81.8% for the overlap error, $83.5\%^2$ 70.6% and 67.1% for respectively the average, the RMS and the maximum surface distance. With an average DSC of 86.1%, a volume difference of 7.4%, a RMS and maximum surface difference of 2.3 and 6.6 mm, the results are obviously better than the 2D variant on historical data.

Also the sensitivity of the result when changing the initial points is examined. The algorithm is tested several times with different starting conditions. Sometimes those initial points are deliberately chosen to be not near to the ideal places. This leads to a coefficient of variation of 28.95% for the segmentation volume.

Test data

The algorithm is tested by [46] on 10 tumors coming from 6 patients. The results of the test data are worse than the results of the training data, as can be seen in Table F.1.

The average total score \pm standard deviation is $68.9 \pm 14.1\%$. The median is 74%. All average measures have a value below 90%. Generally, the best results are obtained with small tumors with a strong gradient at the edges (tum. 2, pat. 6) and tumors surrounded by homogeneous liver tissue (tum. 2, pat. 7). The large tumors within different structures have a lower score (tum. 1, pat. 7 and tum. 1 pat. 8), and also the tumors with a weak gradient (tum. 1, pat. 9). Combination of a high gradient and a small amount of surrounding liver tissue gives an average score (tum. 2, pat. 5 and tum. 1 pat. 10). The time needed for segmentation varies from less than 20 s to almost 2 min. The largest tumors need the highest segmentation time.

F.4 Discussion

First, we will discuss the influence of the different steps of the algorithm. We see tumor segmentation as an optimization algorithm. The initialization is provided by the spiral scanning technique, whereas the level set algorithm will adjust the segmentation to the local optimum. Both the spiral scanning and level set parts of the algorithm have to balance between following the image information and providing a smooth tumor. The main finetuning of the algorithm goes down to this

²This value is differently calculated as the mean surface distance for the test data, because here distances within a surface are taken negative

Tabel F.1: Results

Patient	Tumor	Overlap Error		Volume Diff.		Ave. Surf. Dist.		RMS Surf. Dist.		Max. Surf. Dist.		Time (s)	Total Score
		(%)	Score	(%)	Score	(mm)	Score	(mm)	Score	(mm)	Score		
5	1	32.49	75	13.64	86	2.58	35	3.47	52	12.91	68	40.8	63
	2	38.30	70	29.64	69	1.35	66	1.72	76	5.00	87	25.0	74
	3	31.23	76	22.91	76	1.12	72	1.51	79	5.97	85	24.1	78
6	1	37.42	71	22.74	76	1.06	73	1.44	80	5.45	86	20.3	77
	2	33.84	74	18.27	81	0.72	82	1.10	85	4.46	89	19.2	82
7	1	38.89	70	28.39	71	5.00	0	6.51	9	26.45	34	118.8	37
	2	24.53	81	5.17	95	1.05	73	1.52	79	6.39	84	24.0	82
8	1	31.64	76	2.93	97	3.65	8	4.52	37	16.13	60	92.2	55
9	1	50.52	61	8.58	91	2.07	48	2.90	60	10.35	74	19.5	67
10	1	26.92	79	25.61	73	1.46	63	1.99	72	7.82	80	25.0	74
Average		34.58	73	17.79	82	2.01	52	2.67	63	10.09	75	40.9	69

balance. In the spiral scanning technique, this is obtained by the external cost, while the level set is penalized by the curvature term.

By inverting the gradient when the tumor has higher intensity than the surrounding liver, the algorithm is able to segment dark and light tumors with respect to liver tissue. Another strength of the algorithm is the use of gray level intensities, gradients and local shape. Taking more tumor information into account, increases the reliability. An improvement could be obtained by implementing an efficient texture filter (see Fig. F.1) which can further increase the reliability of the pixel classification algorithm. Among the Haralick features [55], we suggest using correlation and variance, but also other features could be investigated. Another possible improvement of the pixel classification algorithm is adding outlier detection [51].

A shortcoming of the algorithm is the assumption that the tumor is surrounded by a single tissue class (i.e. liver tissue). By adding a rest class, the pixel classification algorithm is capable to distinguish tissues with other gray levels from tumor tissue. If the tumor is adjacent to liver tissue and other structures in the abdomen, the chance on a poor segmentation is higher. We expect the algorithm to perform well for the segmentation of homogeneous round shaped structures, surrounded by homogeneous tissue. However, the worst results in Table F.1 are mainly due to high surface distance errors. The cause is that those surface distances are not relative measures. In consequence big tumors have higher surface distance errors. Most of this errors occur at the poles of the sphere, the starting point and the end point of the spiral (see south pole in Fig. F.6). The reason could be that at the poles the initial segmentation is influenced more by underlying voxels than voxels that lay in the axial plane through the pole. In that direction, the resolution is less than the resolution in the axial plane.

The discrepancy between the results of test and training data is remarkable. There are several possible explanations. Because we had no knowledge about the manual segmentation of the test data, our user intelligence was reduced when placing the two seeds. In comparison to us, a trained physician does have that knowledge. In consequence he is capable to provide a qualitatively better segmentation. Another explanation could be that the tumors of the test data are inherently more difficult to segment. Therefore, it is important to keep in mind that comparing results of different algorithms that use different data sets has to be done with sufficient caution. The last and probable possible reason is the high sensitivity to the initial conditions of the algorithm as shown in the reproducibility study (Sect. F.3). With respect to 7.9% from [12] for manual segmentations, the coefficient of variation of the segmentation volume of 28.95% is very high. It has to be remarked that the goal in [12] is to have a low coefficient of variation, while tests here wanted also test the robustness. Nevertheless, placing one point too far from the center of the tumor or placing the other point not at the right distance from the center in liver tissue could give poor segmentations.

F.5 Conclusions

The semi-automatic algorithm described here is able to segment liver tumors by only two user placed points. The tumor center and the maximum radius, both calculated from those two points,

are used to provide a primary segmentation using the spiral-scanning technique. A level set method adjusts this first segmentation using a speed function obtained from a pixel classification algorithm. The accuracy is only sufficient in a small number of cases. The algorithm performs better for tumors that have an obvious edge, a high difference between the average gray level intensity of the tumor and the liver combined with a rather small standard deviation for both intensity distributions. Besides, the user is required to have some knowledge about tumors.

Acknowledgements

We wish to thank W. Coudyser, prof. dr. S. Tejpar and dr. H. Prenen of the University Hospitals Leuven for providing the data sets with manual delineations.

Appendix: Details Results Training Data

Tabel F.2: Results of training set

patient	tumor	dsc [%]	jsc [%]	Vol. Diff. [%]
1	1	62.42	45.37	3.15
	2	87.82	78.29	15.60
2	1	86.96	76.93	0.93
	2	87.83	78.30	9.17
3	3	92.26	85.64	0.89
	1	82.88	70.77	10.38
4	1	87.13	77.20	9.10
	2	89.96	81.75	14.20
av	3	91.17	83.78	6.19
	4	92.62	86.26	4.13
std		86.11	76.43	7.37
		8.82	11.87	5.19

Tabel F.3: Results surface distances of training set

patient	tumor	manual			algorithm		
		SD_{mean}	SD_{max}	SD_{RMS}	SD_{mean}	SD_{max}	SD_{RMS}
1	1	-3.46	-13.29	5.23	-1.42	-16.59	6.46
	2	-1.57	-5.14	1.86	1.36	4.80	1.67
2	1	-1.31	-6.70	2.02	0.99	4.61	3.04
	2	-0.73	4.04	1.69	0.44	-6.76	1.82
3	3	-1.08	-3.40	1.55	0.92	3.97	1.52
	1	-1.88	-4.44	2.21	1.29	3.94	1.94
4	1	-0.88	-12.11	3.05	-0.02	8.29	2.98
	2	-0.72	-3.11	1.52	0.42	-3.53	1.59
av.	3	-1.70	-8.90	2.51	1.58	8.58	2.56
	4	-0.92	-4.64	1.71	0.82	5.32	1.73
std.		-1.42	6.57	2.33	0.64	6.64	2.40
		0.83	3.65	1.12	0.87	3.93	1.50