

Bruk av MR-bilder for estimering av skjelettalder

Notatnr

SAMBA/09/16

Forfattere

Line Eikvil

Dato

03.03.2016

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
office address
Gaustadalleen 23a
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Bankkonto · bank account
8200.01.48888
Org.nr. · enterprise no.
NO 952125001 VAT

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Tittel **Bruk av MR-bilder for estimering av skjelettalder**

Forfattere **Line Eikvil**
Dato 03.03.2016
År 2016
Publikasjonsnummer SAMBA/09/16

Sammendrag

Dette er et teknisk notat som beskriver metoder og mulig framgangsmåte for automatisk estimering av skjelettalder fra MR-bilder.

Beregning av skjelettalder fra hånd skjer i dag basert på røntgenbilder. Det er imidlertid ønskelig å finne fram til alternativer til dette som ikke benytter røntgenstråling. Dette vil blant annet gjøre det enklere å samle inn data for oppbygging av nye og større bildeatlas av skjelettutvikling. Et alternativ kan da være MR-bilder. Per i dag fins det ingen etablerte metoder for beregning av skjelettalder fra MR, men Universitetet i Roma arbeider med utvikling av en slik metode. Manuell visuell analyse og gradering av MR-bilder for estimering av skjelettalder krever imidlertid mye opplæring og det kan bli variasjoner avhengig av hvem som utfører analysen. Det er derfor ønskelig å kunne automatisere deler av analysen for å gjøre prosessen mer objektiv og reproducerbar.

Målet med denne studien har vært å se på om og hvordan det kan utvikles en metode for automatisk estimering av skjelettalder fra MR-bilder med utgangspunkt i den manuelt baserte metoden fra Universitetet i Roma. Gjennom arbeidet har vi satt oss inn i hvordan den manuelle metoden virker og laget en beskrivelse av denne. Videre har vi sett på studier som er gjort tidligere på automatisert beregning av skjelettalder fra røntgenbilder, samt helt ferske studier som tar for seg dette for MR-bilder. På bakgrunn av dette har vi foreslått en mulig framgangsmåte, utviklet metoder og implementert og testet ut dette for beregning av skjelettalder fra radialbeinet på et sett med bilder fra MR-opptak gjort ved Universitetet i Roma.

Resultatene viser at den foreslåtte automatiske metoden kan ha potensiale til å brukes i estimering av skjelettalder, men datasettet vi har hatt tilgjengelig har vært begrenset. Både den automatiske tilnærmingen og den manuelle MR-metoden trenger å valideres på et større og uavhengig datasett.

Emneord Estimering av skjelettalder, MR-bilder, automatisk bildeanalyse
Målgruppe Teknisk prosjektgruppe
Tilgjengelighet
Prosjektnummer 220 642
Satsningsfelt Bildeanalyse og mønstergjenkjenning
Antall sider 36
© Copyright Norsk Regnesentral

Innhold

1	Innledning.....	7
2	Beskrivelse av MR-metode fra Universitetet i Roma	7
2.1	Overordnet beskrivelse.....	7
2.2	Detaljert beskrivelse	8
2.2.1	Bildeopptak	8
2.2.2	Områdene i hånden som analyseres	8
2.2.3	Utvalg av bilder og bildesnitt.....	8
2.2.4	Egenskaper for utviklingstrinn.....	9
2.2.5	Sammenheng mellom skjelettalder og definerte stadier	10
2.2.6	Estimering av skjelettalder fra utviklingstrinn.....	11
2.3	Datasett.....	11
2.4	Validering	12
2.5	Diskusjon.....	12
2.6	Oppsummering.....	14
3	Andre MR-baserte studier	14
4	Automatisert beregning av skjelettalder	15
4.1	Røntgenbilder.....	15
4.1.1	BoneXpert.....	15
4.1.2	Andre tilnærminger	17
4.2	MR-bilder	18
4.3	Oppsummering og diskusjon.....	19
5	Automatisering av metode fra Universitetet i Roma	20
5.1	Vurdering av den manuelle prosessen med tanke på automatisering.....	21
5.1.1	Utvelgelse av bildesett og bildesnitt.....	21
5.1.2	Identifisering av område for videre analyse.....	21
5.1.3	Utvelgelse av egenskaper.....	22
5.1.4	Bestemmelse av utviklingstrinn.....	22
5.1.5	Estimering av skjelettalder fra utviklingstrinn.....	23
5.2	Forslag til automatisert metode for MR-bilder	23

5.2.1	Overordnet strategi	23
5.2.2	Datasett.....	24
5.3	Beskrivelse av automatisert analyse.....	25
5.3.1	Segmentering av radialbeinet	25
5.3.2	Utvalg av snitt.....	25
5.3.3	Bestemmelse av ROI	26
5.3.4	Uttrekking av egenskaper	26
5.3.5	Klassifikasjon	27
5.3.6	Aldersestimering.....	27
5.4	Eksperimenter og foreløpige resultater.....	27
5.4.1	Framgangsmåte	27
5.4.2	Resultater	30
5.4.3	Diskusjon	32
6	Oppsummering og konklusjon	33
7	Referanser	35

1 Innledning

Målet med studien beskrevet her har vært å se på om og hvordan det kan utvikles en metode for automatisk estimering av skjelettalder fra MR-bilder med utgangspunkt i en manuelt basert metode utviklet ved Universitetet i Roma [26][27].

I det følgende beskriver vi først i kapittel 2 den manuelle MR-metoden som er utviklet i Roma. Deretter har vi i kapittel 3 tatt med en kort oversikt over MR-arbeid fra andre miljøer. I kapittel 4 har vi sett på det som er gjort av tidligere arbeid på automatisert beregning av skjelettalder. Her omhandler de fleste studiene automatisert analyse av røntgenbilder, mens det bare er noen ganske få studier fra de siste 1-2 årene som har sett på dette for MR-bilder. Til slutt foreslår vi i kapittel 5 metoder for automatisert analyse av MR-bilder med utgangspunkt i den manuelle metoden fra Roma. Metoden er også implementert og testet ut, og i dette kapittelet presenterer vi også resultater fra denne uttestingen.

2 Beskrivelse av MR-metode fra Universitetet i Roma

Beskrivelsen i det følgende er i hovedsak basert på muntlig og skriftlig kommunikasjon med La Sapienza Universitet i Roma, samt noe underlagsmateriale i form av publikasjoner og presentasjoner fra dette miljøet. Der ikke annet er angitt er tabeller og figurer laget ved NR basert på informasjon vi har fått fra Universitetet i Roma.

Metoden har vært under stadig utvikling i Roma, og det eksisterte derfor ingen fullstendig og oppdatert beskrivelse av denne. En delvis og ikke helt oppdatert beskrivelse fantes i artikler [26][27], men det var ikke tilstrekkelig for vårt formål. Ettersom det viste seg å være vanskelig for italienerne å lage en beskrivelse på det nivået vi trengte, kom derfor i stedet Milvia Martino, som har arbeidet med denne delen ved Universitetet i Roma, til Oslo slik at vi kunne få en oppdatert og mer detaljert gjennomgang av deres metode.

2.1 Overordnet beskrivelse

Framgangsmåten fra Universitet i Roma baserer seg kun på manuell visuell analyse av MR-bildene. Den er inspirert av de manuelle og veietablerte metodene definert av Greulich og Pyle [7] og Tanner og Whitehouse [15] som benyttes for røntgenbilder. Metoden baserer seg på observasjoner av endringer for 9 ulike bein i hånden. Disse observeres fra to sett av MR-bilder tatt av øvre og nedre del av hånden, der hvert bildesett består av 26 snitt. Bildesett og snitt som analyseres velges for hvert bein ut fra hvor godt de interessante egenskapene ved beinet er avbildet.

Egenskapene som vurderes ved hvert enkelt bein i sammenvoksingsprosessen baserer seg på ting som avstand mellom bein, endringer i beinets form, tilstedeværelse av bruske rundt beinet og små endringer i tekstur i brusken rundt og i beinet. For hvert bein er det definert 12 utviklingstrinn hvor hvert trinn er beskrevet tekstlig basert på hvilke endringer som er observert i MR-bildene for de enkelte trinnene. Basert på analyser av et materiale på noen hundre personer, har de også utarbeidet en tabell for sammenheng mellom utviklingstrinn og alder. Ved aldersestimering bestemmes så utviklingstrinnene for hvert av de 9 beina i hånden, og basert på tabellen over sammenheng mellom trinn og alder, settes et aldersestimat. En mer detaljert beskrivelse av prosessen er gitt i de neste avsnittene.

2.2 Detaljert beskrivelse

2.2.1 Bildeopptak

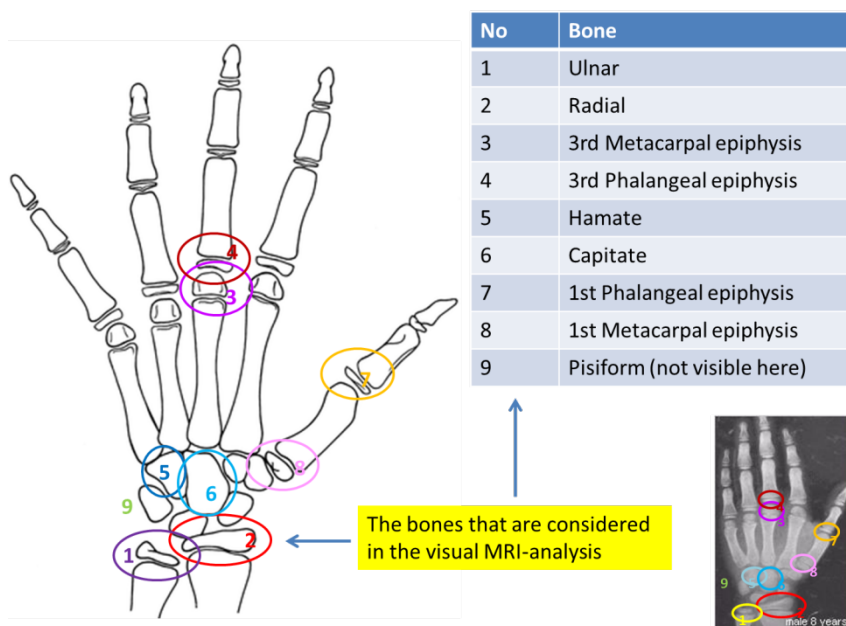
Bildene er tatt opp på en åpen 0,2 Tesla MR-maskin; G-Scan, Esaote, Genova, Italia. Parametre for opptaket var som følger: A single T1 spin echo sequence, repetition time [TR] 38/echo time [TE] 16/flip angle 65 degrees, matrix size 192_110_24; 1.3mm slice thickness with no interslice gap, field of view (FOV) 140_120_40 mm.

Maskinen er utstyrt spesielt for opptak av bilder av hånd og håndledd. Hånden blir holdt på plass av mindre pads for å begrense bevegelse mest mulig. For de minste barna er det tilstrekkelig med ett opptak som avbilder hele hånden med håndleddet. For de fleste er det imidlertid nødvendig med to opptak for å få med hele hånden, hvor hånden flyttes litt mellom de to opptakene. Totalt tar hele prosessen med posisjonering og bildeopptak cirka 10-12 minutter for de som trenger to opptak, og bare 5-6 minutter for de minste som klarer seg med ett opptak.

Bildene er lagret som DICOM i PACS og etterprosessert med dedikert software fra INFINITT, Healthcare Co., Ltd, Korea. Her kjenner vi ikke til hva denne etterprosesseringen består i.

2.2.2 Områdene i hånden som analyseres

Estimering av skjelettalder basert på hånd og håndledd baserer seg på å finne ut hvor langt utviklingen og sammenvoksingen av beina i hånden har kommet. I tilnærmingen de har kommet fram til i Roma, benyttes 9 forskjellige bein. Figuren nedenfor viser en oversikt over disse 9 beina og deres plassering i hånden.



Figur 1: Bein som brukes i den visuelle MR-analysen for aldersestimering ved Universitetet i Roma.

2.2.3 Utvalg av bilder og bildesnitt

For hver person vil bildesettet bestå av 26 bildesnitt fra hvert opptak, der det enten er gjort ett eller to opptak. Den videre visuelle analysen baserer seg på en del spesifikke detaljer som observeres i bildet (se neste avsnitt), det er derfor viktig å finne fram til de bildesnittene der de aktuelle detaljene kommer best fram. For hvert bein som skal analyseres velges derfor det

opptaket og det bildesnippet der de interessante egenskapene kommer best fram. Det betyr at kriteriet for hvilket snitt som skal velges i prinsippet kan variere både med hvilket av de 9 beina som skal analyseres og med hvilke detaljer som er viktige for utviklingstrinnene som er aktuelle for dette individet.

2.2.4 Egenskaper for utviklingstrinn

For hvert utviklingstrinn og for hvert av beina som analyseres er det laget tekstlige beskrivelser av endringer og egenskaper som forventes observert i MR-bildene. Definisjonen av disse utviklingstrinnene med de tilhørende beskrivelsene er kjernen i arbeidet som er gjort ved Universitet i Roma, hvor dette er et nybrottsarbeid i den forstand at ingen andre har beskrevet dette for MR-bilder av hånden. Ideen deres har vært å utvikle et atlas for skjelettalder basert på MR-bilde, bestående av definerte utviklingstrinn med tilhørende tekstlige beskrivelser og bilder.

Til forskjell fra røntgenbilder som hovedsakelig avbilder beinstrukturer, er det med MR mulig å se mye mer detaljer og også strukturer i ulike typer av vev. Beskrivelsene av hvert utviklingstrinn omfatter derfor mange flere detaljer enn det som har vært beskrevet for røntgen, og det er derfor også definert mange flere utviklingstrinn. Totalt er det for hvert bein som analyseres, definert 12 utviklingstrinn. Egenskapene som er definert for hvert trinn er typisk knyttet til tilstedeværelse og størrelse av brusk, form på beinet, avstand til bein omkring og grad av sammenvoksing basert på observasjon av linjer og strukturer mellom bein.

Figur 2 viser en tabell som kort beskriver utviklingstrinnene 6-12 for radialbeinet. Beskrivelsen er basert på hvordan utviklingen framstår visuelt i MR-bildene, fra trinn 6 hvor beinet har en sommerfugl-fasong og hvordan denne gradvis endres og grensene mot de andre beina gradvis blir mindre og viskes ut. Figur 3 gir en illustrasjon av hvordan dette kan opptre i bildene.

	2 Radial (Similar to ulnar)	Age (Male)
I		
II		
III		
IV		
V		
VI	Epiphysis longer, covers whole plate.	8,5-11,5
VII	Styloid process: growth upper corner, from rounded to more sharp.	11,5-12,5
VIII	Hooks appear. Appearance of dots.	12,5-14
IX	Line going from black and white to solid black. Capping: hook grows, covers plate. Appearance of dots.	14-15,5
X	First bone bridge appears	15,5-16,5
XI	More bone bridges	16,5-17,5
XII	Closing, complete fusion. Lines within may still show.	17,5-19

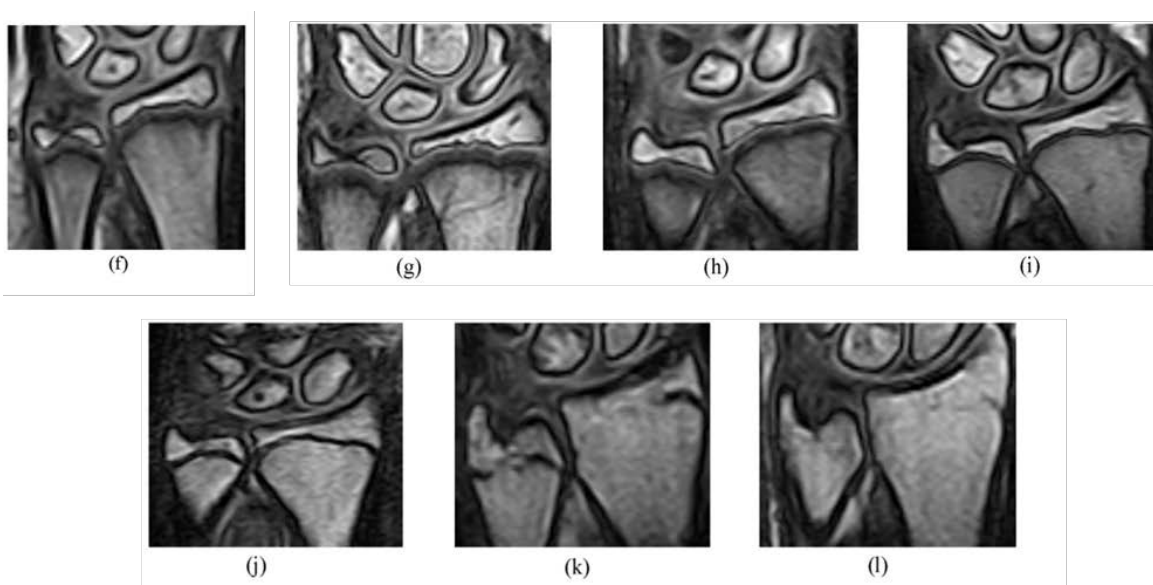
Figur 2: Beskrivelser av utviklingstrinn (fra VI og oppover) som observert i MR-bilder for radialbeinet og beskrevet av Universitetet i Roma.

En full oversikt med beskrivelser for alle 9 bein og alle utviklingstrinnene er gitt i kapittel 4 i boka «Text-atlas of skeletal age determination», av Tomei med flere, fra 2014 [27]. Beskrivelsene har imidlertid blitt noe oppdatert siden denne utgivelsen. I boka har også beskrivelsene noe mer fokus på den fysiske utviklingen, og forklarer ikke alltid hvordan dette opptrer i bildet.

2.2.5 Sammenheng mellom skjelettalder og definerte stadier

Basert på bildesett av friske barn med kjent alder og de utviklingstrinnene som beskrevet tidligere, har forskerne i Roma satt opp en sammenheng mellom utviklingstrinn og skjelettalder for hvert av de 9 beina i hånden. Dette er definert med én tabell for jenter og én for gutter, ettersom utviklingen normalt går forttere for jenter enn gutter.

Tabellen i Figur 4 viser sammenhengen som er definert mellom utviklingstrinn og alder for gutter. Her betyr spørsmålstegn at det har vært for lite data i denne kategorien til å gi noe godt anslag. En strek i tabellen betyr at det aktuelle beinet er fullt utvikst i dette stadiet. Fra denne tabellen ser vi det samme som er beskrevet i mange andre studier, at utviklingen av de fleste av beina er ferdig ved 16-års alder, mens det etter dette bare er ved ulnar og radial at det fortsatt skjer en utvikling.



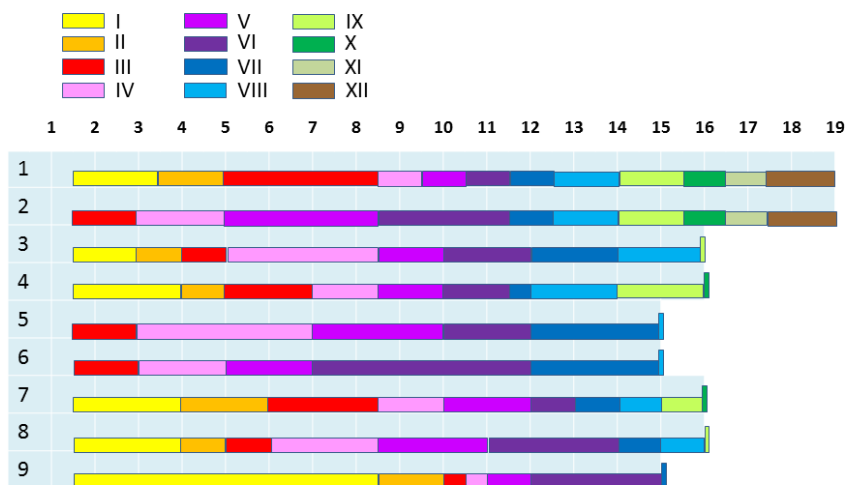
Figur 3: Utviklingen av radialbeinet som observert i MR-bilder for utviklingstrinn VI (f) til XII (l). Bilder fra Universitetet i Roma [27].

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1 - Ulnar	1.5-3.5	3.5-5	5-8.5	8.5-9.5	9.5-10.5	10.5-11.5	11.5-12.5	12.5-14	14-15.5	15.5-16.5	16.5-17.5	17.5-19
2 - Radial	?	?	1.5-3	3-5	5-8.5	8.5-11.5	11.5-12.5	12.5-14	14-15.5	15.5-16.5	16.5-17.5	17.5-19
3 - 3 rd metacarp	1.5-3	3-4	4-5	5-8.5	8.5-10	10-12	12-14	14-16	16	—	—	—
4 - 3 rd phalang.	1.5-4	4-5	5-7	7-8.5	8.5-10	10-11.5	11.5-12	12-14	14-16	16	—	—
5 - Hamate	?	?	1.5-3	3-7	7-10	10-12	12-15	15	—	—	—	—
6 - Capitate	?	?	1.5-3	3-5	5-7	7-12	12-15	15	—	—	—	—
7 - 1 st Phalang.	1.5-4	4-6	6-8.5	8.5-10	10-12	12-13	13-14	14-15	15-16	16	—	—
8 - 1 st Metacarp	1.5-4	4-5	5-6	6-8.5	8.5-11	11-14	14-15	15-16	16	—	—	—
9 - Pisiform	1.5-8.5	8.5-10	10-10.5	10.5-11	11-12	12-15	15	—	—	—	—	—

Figur 4: Tabellen viser sammenheng mellom alder og utviklingstrinn for gutter. Spørsmålstegn betyr at det har vært for lite data til å gi noe anslag, mens en strek betyr at beinet er fullt utvikst i dette stadiet. Aldersintervallene i tabellen er utarbeidet ved Universitetet i Roma.

Tabellen i Figur 5 viser det samme som tabellen i Figur 4, men visualisert på en litt annen måte. Her viser fargekoden hvilket utviklingstrinn hvert av de 9 beina forventes å være i for hver alder (1-19).

Age and stage correspondence (male)



Figur 5: Sammenheng mellom alder og utviklingstrinn for gutter. Figuren visualiserer hvilket utviklingstrinn (fargekodet, I-XII) hvert av de 9 beina (1-9) forventes å være i for ulike aldre (1-19).

2.2.6 Estimering av skjelettalder fra utviklingstrinn

For å estimere alder, er den videre framgangsmåten som definert ved Universitetet i Roma, at en finner en sammenheng mellom utviklingstrinnene og alder. Dette har de valgt å gjøre på følgende måte: For et sett med bilder av personer med kjent alder, bestemmes utviklingstrinnene for hvert av de 9 beina. Deretter summeres tallene som angir hvert av de 9 utviklingstrinnene til en samlet utviklingscore. Basert på regresjon lager de deretter en oppslagstabell for sammenhengen mellom summert utviklingscore og alder, der det er én tabell for jenter og én for gutter. Tanken er da senere at radiologen som skal estimere alder for nye individer, studerer MR-bildene og bestemmer utviklingstrinn for hvert bein, summerer trinnene til en utviklingscore og bruker denne scoren til å slå opp i tabellen og finne aldersestimatet. Bruken av denne typen utviklingscore er inspirert av tidligere arbeider av Tanner basert på røntgenbilder (1994 [16]), men med mindre differensiert vektning enn det Tanner bruker.

2.3 Datasett

Datasettet som er brukt i analysene fra Universitet I Roma, består av MR-opptak fra drøyt 100 personer av hvert kjønn. Opptakene er av friske barn rekruttert fra skoler i Roma, der barna er i alderen 4 til 18 år. Mer spesifikt er det totalt 126 opptak av gutter og 106 av jenter. Tabellen nedenfor viser antall personer i hver aldersgruppe for hvert kjønn. (Tallene her er hentet fra artikkelen "Value of MRI of the hand and the wrist in evaluation of bone age: preliminary results", fra 2014. [26])

Alder	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
# gutter	1	4	0	2	4	5	3	6	17	9	18	28	9	14	6
# jenter	2	1	0	2	4	6	13	10	16	10	7	10	16	9	0

Tabell 1: Oversikt over antall opptak for ulike aldersgrupper brukt i studien fra Roma [26].

Slik vi har forstått det er det dette datasettet som er benyttet for all analyse. Det vil si at det er dette datasettet som er benyttet både for å etablere den initielle beskrivelsen av utviklingstrinn, for å etablere sammenheng mellom alder og utviklingstrinn, for å gjøre regresjonen og etablere oppslagstabellen og for senere validering av resultatene. Dette går ikke helt klart fram av artiklene, men det er det eneste datasettet av friske individer som er beskrevet. Det er også dette inntrykket vi har fått gjennom direkte kommunikasjon med Roma.

2.4 Validering

Det følgende beskriver valideringen som er gjort ved Universitetet i Roma. Beskrivelsene er basert på artikkelen *“Value of MRI of the hand and the wrist in evaluation of bone age: preliminary results”* fra 2014 [26], og direkte skriftlig og muntlig kommunikasjon med Universitetet i Roma.

To radiologer som kjente personenes alder, etablerte prototypen for scoringssystemet basert på visuell analyse av MR-bilder fra settet med bilder som beskrevet i avsnittet over. De samme to radiologene bestemte i fellesskap (ved konsensus) skjelettalder fra de samme bildene. Etter tre måneder, analyserte de bildene på nytt, men nå hver for seg og estimerte skjelettalderen igjen – uten å se på det som var gjort tidligere, men datasettet var det samme. I etterkant ble følgende korrelasjons- og samsvarsanalyser gjort basert på disse resultatene:

Korrelasjon mellom skjelettalder og kronologisk alder: Pearson’s korrelasjonskoeffisient (R^2) ble brukt til å se på korrelasjonen mellom estimert skjelettalder og kronologisk alder bestemt av radiologer; AB, A og B, der AB=radiolog 1 og radiolog 2 i konsensus, A=radiolog 1 alene etter 3 mnd, B=radiolog 2 alene etter 3 mnd. Analysen viste god korrelasjon for de tre tilfellene; med $R^2=0.976$ for AB, $R^2=0.978$ for A og $R^2=0.977$ for B.

Samsvar mellom analyser: Cohen’s kappa koeffisient ble brukt til å se på samsvaret mellom radiologenes analyser. Det ble sett på repeterbarhet ved å sammenligne AB og A, samt AB og B, og reproducerbarhet ved å sammenligne A og B. Dette ga henholdsvis verdiene $k=0.686$, $k=0.684$ og $k=0.703$. Kappaverdier over 0.6 regnes vanligvis som godt samsvar, så for de gitte analysene er samsvaret bra.

Korrelasjon mellom sekvens av utviklingstrinn og alder: Spearman’s korrelasjonskoeffisient (definert som Pearsons korrelasjonskoeffisient mellom ordnede variable) ble brukt til å se på sammenhengen mellom utviklingstrinn og alder for hvert av de 9 beina som analyseres. Sterkest korrelasjon var det for ulnar og radial med henholdsvis $\rho=0.96$ og $\rho=0.963$, dårligst for hamate med $\rho=0.749$. Tilsvarende analyse er gjort for den totale utviklingsscoren, dvs. summen av utviklingstrinn anslått for hvert av de 9 beina, med $\rho=0.96$.

2.5 Diskusjon

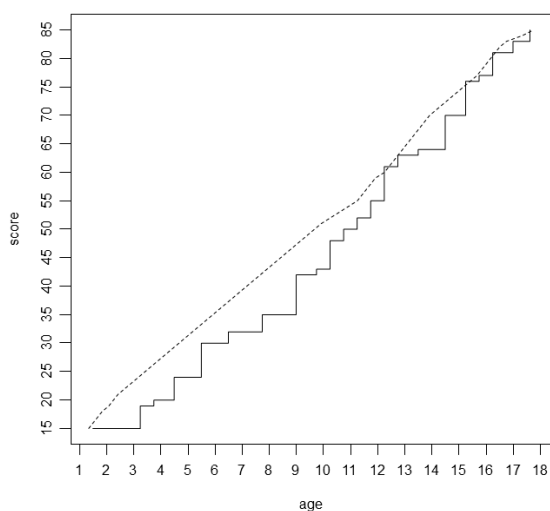
Slik vi har forstått det er alle trinnene i prosessen utført på samme datasett, dvs. både initiell definering av utviklingsstadier med beskrivelser, regresjonsanalyse basert på sammenheng mellom score og alder, samt validering. Datasettet som er brukt er også ganske lite. Med rundt 100 individer fordelt på 15 alderstrinn, blir det ganske få observasjoner for hvert alderstrinn. I tillegg ser det ut til at alle analysene er foretatt av de samme to ekspertene.

Slik vi har fått det forklart har to eksperter først studert bildene sammen med blant annet informasjonen om kronologisk alder for å etablere settet av beskrivelser av utviklingstrinnene med tilhørende aldersintervall for alle beina som skal analyseres. Deretter har de på det samme datasettet i fellesskap bestemt utviklingstrinn fra bildene. Denne informasjonen er

brukt til å gjøre en regresjon for å finne sammenhengen mellom summen av utviklingstrinn og kronologisk alder. Tre måneder senere har de samme to ekspertene analysert alle bildene på nytt, men denne gangen hver for seg uten at de har hatt kronologisk alder tilgjengelig på det tidspunktet.

Valideringen som er gjort i Roma viser god korrelasjon og godt samsvar på flere punkter. Slik vi har forstått det, er imidlertid både definering av utviklingstrinn og scoringsmodell, regresjon og validering gjort på ett og samme begrensede datasett. Analysene er også gjort av de samme radiologene som etter hvert har hatt god kjennskap til datamaterialet. Valideringen blir dermed ikke helt uavhengig og korrelasjonene som observeres kan da også skyldes at ekspertene er samkjørte og kjenner materialet veldig godt. Siden alt er gjort på samme datasett sier det også lite om hvor godt overførbart dette er til et datasett med nye individer.

På NR har vi gjort noen nærmere analyser av sammenhengen mellom utviklingstrinn som definert i oppslagstabellen (fra regresjonen) i forhold til det som er definert i de underliggende tabellene (Figur 4 og Figur 5). Vi har her observert noen inkonsistenser, hvor det for noen nivåer av score er stor forskjell mellom alder gitt i oppslagstabellen og alder som forventet ut fra de underliggende tabellene. Dette er illustrert i Figur 6. Vi vet lite om hvordan regresjonsanalysen er gjort, men en mulig årsak kan være at de har få datapunkter for gutter under 11 år.



Figur 6: Sammenheng mellom score og alder for oppslagstabell fra regresjon (prikkete linje) og for underliggende tabeller over utviklingstrinn (heltrukket linje).

Den visuelle analysen baserer seg på observasjon av veldig mange små detaljer. Det vil derfor kreves høy kompetanse på dette for å kunne gjøre en visuell analyse, og det kan være utfordrende å lære opp radiologer og å gjøre repeterbare og reproduerbare analyser. For en mest mulig objektiv analyse ville da en delvis automatisering av denne prosessen basert på bildeanalyse kunne være veldig nyttig. Samtidig gjør imidlertid alle detaljene at det å utvikle en mer automatisert prosess basert på de samme egenskapene også kan være ganske komplekst og omfattende.

2.6 Oppsummering

Metoden for estimering av skjelettalder fra MR-bilder som er utviklet ved Universitetet i Roma, baserer seg som mange røntgenbaserte metoder på å definere et atlas med utviklingstrinn bestående av tekstlige beskrivelser og bilder. Det er definert opp til 12 utviklingstrinn for 9 forskjellige bein i hånden. Videre er det basert på regresjon, laget en oppslagstabell som definerer sammenhengen mellom en score definert av summen av de 9 utviklingstrinnene og alder. Estimering av skjelettalder for en ny person er så tenkt gjort ved først å bestemme de 9 utviklingstrinnene fra bilder og beskrivelser, for deretter og summere disse til en score og finne aldersestimaten fra oppslagstabellen.

Datasettet som er benyttet for etablering av metoden består av rundt 100 individer av hvert kjønn fordelt over alderstrinnene 4-18 år. Scoringsmodellen og regresjonen bygger dermed på et ganske begrenset datamateriale. Foreløpig er metoden heller ikke validert for uavhengige datasett, og vi vet derfor lite om hvor godt overførbart dette er til et datasett med nye individer. Framgangsmåten er heller ikke testet for bruk av uavhengige radiologer. Siden den visuelle analysen for manuell tilordning til utviklingstrinn bygger på mange små detaljer, kan det bety at det vil være utfordrende å lære opp radiologer og å gjøre repeterbare og reproducerbare analyser.

3 Andre MR-baserte studier

Arbeidet gjort i Roma har fokusert på å etablere et eget atlas med detaljerte beskrivelser av utviklingstrinn. Det finnes imidlertid også noen få andre studier som beskriver framgangsmåter for å estimere skjelettalder fra MR-bilder. Her har formålet i noe mindre grad vært å utvikle en framgangsmåte, men mer å se på potensialet for estimering av skjelettalder ved bruk av MR.

Dvorak et al, 2007 [5], ser på bruk av MR av håndledd for estimering av alder på yngre mannlige fotballspillere. De har etablert et graderingssystem i 6 utviklingstrinn basert på sammenvoksingen av radialbeinet. Gradering inn i stadier basert på visuell analyse av MR-opptak ble så gjort uavhengig av tre personer for totalt 496 mannlige fotballspillere i aldersgruppen 14-19 år. Studien demonstrerte en god korrelasjon mellom utviklingstrinn og alder.

Terada et al, gjorde en MR-studie i 2013 [17] på 83 japanske barn. Graderingen her ble basert på TW2 RUS, dvs. Tanner and Whitehouse metode som ser på radial, ulnar og 11 korte bein i hånden. TW2 definerer 8 eller 9 utviklingstrinn, (A),B,...I, for hvert bein beskrevet ved figurer og bilder. For hvert bein og hvert utviklingstrinn er det tilordnet et tall som er utarbeidet i forhold til den ulike utviklingen av de forskjellige beina. Summering av tallene tilordnet hvert av utviklingstrinnene en kommer fram til gir en total utviklingsscore, hvor det er definert en sammenheng med skjelettalder [15]. I Teradas studie overføres TW2 trinnene basert på røntgen direkte til MR, der de konkluderer med at disse også kunne bedømmes presist fra disse bildene.

Begge disse studiene indikerer dermed at MR kan være godt egnet for å estimere skjelettalder.

4 Automatisert beregning av skjelettalder

Hovedproblemet med manuell visuell analyse og gradering av bilder for estimering av skjelettalder, er at det kan bli variasjoner avhengig av hvem som utfører analysen. Hensikten med å automatisere deler av analysen, er derfor først og fremst å gjøre prosessen mer objektiv og mindre avhengig av hvem som utfører analysen. Her vil vi først se litt på hva som er gjort på automatisert analyse av røntgenbilder siden historien her er lenger. For automatisert estimering av skjelettalder fra MR-bilder er det bare noen ganske få studier, alle fra 2014-2015 og alle fra samme miljø. Disse er oppsummert i avsnitt 4.2.

4.1 Røntgenbilder

Røntgenbilder har lenge vært brukt som basis for estimering av skjelettalder. Allerede i 1959 definerte Greulich og Pyle [7] et bildeatlas beregnet på visuell sammenligning som beskriver utviklingen knyttet til alderstrinn. Manuell estimering av skjelettalder ved denne metoden skjer typisk ved at røntgenbildet for hånden som skal analyseres sammenlignes med bildeatlasen for å finne referansen som ligner mest. Senere definerte Tanner og Whitehouse med flere [15] en litt annen tilnærming som baserer seg på en tilordning av score for ulike bein i hånden. Her bestemmes først hvert bein til ett av 8(9) utviklingstrinn (A), B,C,...I. For hvert av disse utviklingstrinnene defineres en score, og et samlet aldersestimat beregnes ved først å summere score for utviklingstrinnene som er bestemt for hvert bein og så benytte sammenhengen definert mellom summert utviklingsscore og alder. Tilordningen til utviklingstrinn var tenkt gjort basert på tekstlige beskrivelser, men bilder og tegninger av stadiene fulgte også med. Denne metoden har vært lansert i flere versjoner med ulike bein og score-systemer. I manuelle analyser av røntgenbilder er det i hovedsak disse to metodene som benyttes. Tanner og Whitehouse har et potensiale for å være mer nøyaktig enn Greulich og Pyle, men fordi det er mange bein som skal bestemmes enkeltvis blir den lett mer subjektiv samtidig som den er mer tidkrevende og kompleks å gjennomføre. I praksis benyttes derfor Greulich og Pyle mest.

Tanner så i utgangspunktet for seg at deres metode kunne brukes til å utvikle en automatisert regelbasert løsning, men innså etterhvert at dette ikke kunne løses fullstendig regelbasert, men også måtte være erfaringsbasert. Sammen med Gibbons foreslo Tanner senere en tilnærming basert på templatmatching av regioner definert rundt epifysen (knokkelendene) [16]. Dette ble sett på som en god løsning, men problemet var at områdene måtte finnes og avgrenses manuelt. Automatiseringen av dette var den store utfordringen.

Senere har ulike tilnærminger til automatisering av problemet vært beskrevet. En av dem som virkelig har lyktes i å lage en fullstendig automatisert løsning for dette er Hans Henrik Thodberg med BoneXpert systemet. Nedenfor er derfor en litt mer detaljert beskrivelse av dette systemet, mens i avsnittet etter følger en kort oppsummering av en del alternative teknikker som har vært foreslått.

4.1.1 BoneXpert

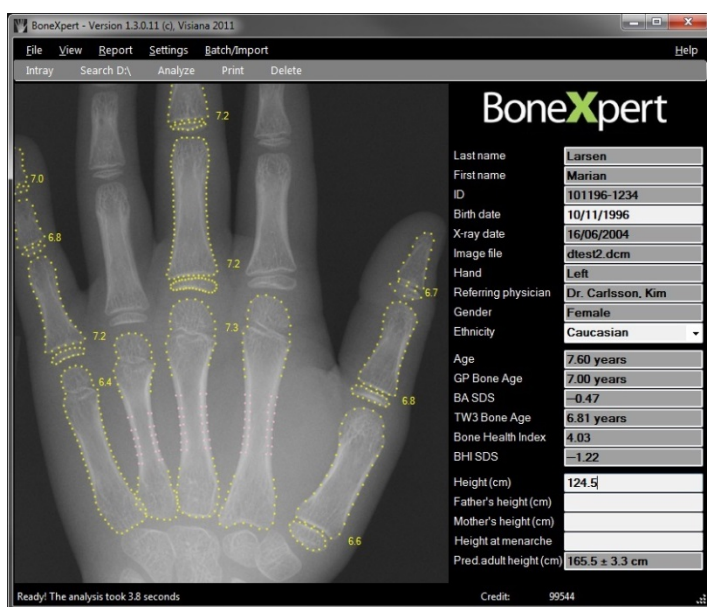
BoneXpert er et system for automatisert estimering av skjelettalder fra røntgenbilder. Systemet baserer seg på bildeanalyse av røntgenbilder og er utviklet i Danmark og tar utgangspunkt i metoder utviklet av Hans Henrik Thodberg med flere. Den automatiserte BoneXpert metoden er beskrevet blant annet i en artikkel fra 2009 [21]. Den er siden testet ut i flere studier [22][23][24].

Prosesen består av følgende tre hovedtrinn: i) I første trinn finnes automatisk omrisset av 15 bein i hånden og konturer med feil forkastes. ii) I andre trinn trekkes det ut egenskaper og bestemmes en initiell skjelettalder for hvert bein og potensielle feil forkastes, iii) I siste trinn transformeres resultatet fra forrige trinn til en skala som er kjent for radiologene.

I det første trinnet finner systemet automatisk omrisset av 15 ulike bein i hånden ved hjelp av opptrente deformérbare modeller [20], basert på aktive konturmodeller (ASM) som først foreslått av Cootes med flere på 90-tallet [2]. Dette er statistiske modeller for konturen av et objekt som iterativt deformeres til å passe til samme type objekter i nye bilder, der deformasjonen begrenses av variasjonene som er observert i treningssettet. Modellen som benyttes i BoneXpert er utviklet med basis i 1678 røntgenbilder av danske barn.

I neste trinn beregnes det en initiell skjelettalder basert på 13 av de identifiserte beina (radial, ulnar og 11 korte bein, tilsvarende TW2 RUS). Skjelettalderen beregnes for hvert av de 13 beina basert på 30 egenskaper. Disse egenskapene består av 10 konturegenskaper fra PCA¹-representasjonen av konturen, 10 egenskaper fra PCA-representasjonen av bildeintensiteten (tettheten) for det samme området, og 10 teksturegenskaper basert på Gabor filtre som beskriver utvalgte områder langs konturen. De siste er spesielt beregnet på å fange opp lukningsprosessen der beina gror sammen. Egenskapene er dermed basert på hele konturen og modellen som er funnet for et bein og er ikke laget for å direkte reflektere egenskapene fra de tekstlige beskrivelsene fra Tanner & Whitehouse. En lineær regresjonsmodell er laget basert på egenskaper trukket ut fra bilder av 1966 danske barn med kjent alder og definerer sammenhengen mellom egenskapene og skjelettalder.

I det siste trinnet transformeres skjelettalderen som bestemt gjennom trinn to til en skala som er kjent for radiologene. Her er det skjelettalderen som definert av Greulich & Pyle som er mest brukt. Transformasjon til denne skjer gjennom å beregne gjennomsnittet for de 13 beina for deretter å benytte en ikke-lineær monoton funksjon for å transformere til denne skalaen.



Figur 7: Bildet av et screen shot fra BoneXpert, hentet fra <http://www.bonexpert.com>.

¹ Prinsippal komponent analyse (PCA) teknikk som brukes til å trekke ut den viktigste informasjonen fra store datasett.

4.1.2 Andre tilnæringer

Ulike tilnæringer har gjennom tiden vært brukt i forsøk på å utvikle løsninger for automatisert estimering av skjelettalder fra røntgenbilder. For å oppnå dette er det to hovedproblemstillinger som må løses: (i) Segmentering av området som skal analyseres (bein eller annet definert interesseområde), (ii) Finne en representasjon av karakteristika ved dette området som henger sammen med aldersutvikling og er egnet for bruk til automatisert estimering av alder.

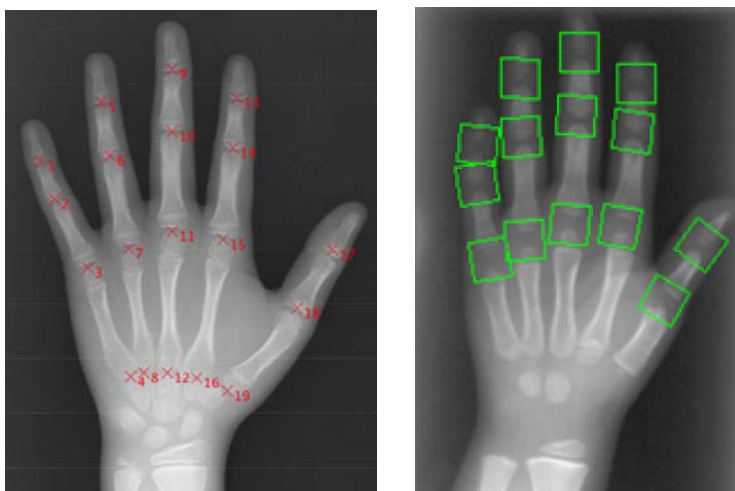
Det har blitt foreslått ulike metoder basert både på standard bildeanalyseteknikker, maskinlæring og modellbaserte metoder, mens Thodberg var den som virkelig lyktes i å utvikle dette helt fram til en ferdig løsning (2009). Tabellen nedenfor oppsummerer kort ulike framgangsmåter foreslått for automatisert estimering av skjelettalder fram til 2009 da BoneXpert ble introdusert. Oppsummeringen er i hovedsak basert på en litteraturoversikt fra en avhandling av Thangam som også er oppsummert i en artikkel fra 2012 [18]. Tabellen nedenfor fokuserer på de studiene som beskriver både segmentering, egenskapsuttrekking og klassifikasjon.

Årstall	Segmentering	Egenskaper	Forfatter
1989	Manuelt initiert adaptiv konturfølger	bredde-lengde mål	Michael, Nelson
1991	Fjerning av bakgrunn, kantdetektor	lengdemål	Pietka et al
1992	Finner ROI manuelt	Templatmatching (TW2)	Tanner, Gibbons
1993	Dynamisk terskling, morfologi	Størrelse, form, posisjon, gråtone	Pietka et al.
1994	Region groing og sammenslåing	Templatmatching (som TW2)	Manos et al
1994	Adaptiv terskling	tekstur	Cheng et al
1994	Fourier analyse	Templatmatching	Drayer, Cox
1997	Aktiv konturmodell	Formegenskaper	Mahmoodi et al
2001	Finner ROI manuelt	Diametre/ratio for epifyse, metafyse, gap	Pietka et al
2003	Aktiv konturmodell, samregistrering	Krysskorrelasjon (direkte eller med NN)	Niemeijer et al
2007	Adaptiv terskling, Canny kant detektor	Morfologiske egenskaper	Zhang et al
2007	DoG filtrering, adaptiv terskling	Ratio mellom bredder, avstander, areal	Giordano
2008	Semi-automatisk vha adaptiv clustering	89 egenskaper, NN-klassifikasjon	Tristan, Arribas
2009	Aktive konturmodeller	30 egenskaper langs konturen	Thodberg, BoneXpert

Det er også foreslått flere metoder etter dette. I det følgende beskrives noen nyere tilnæringer som skiller seg ut fra de som er nevnt i tabellen. Fischer et al fra 2010 [6], bruker en tilnærming basert på "content-based retrieval". Her angir brukeren først senterpunkter for knokkelendene (epifysen) som skal analyseres (se Figur 8, venstre). Interesseområdene rundt (epifyse ROI, ofte referert til som eROI) defineres så ved at det automatisk trekkes ut orienterte bokser rundt senterpunktet (Figur 8, høyre). Deretter brukes krysskorrelasjon til å sammenligne disse boksene med tilsvarende regioner i en database av 1102 røntgenbilder av personer med kjent alder. Skjelettalder estimeres så ut fra alderen for de eROI i databasen som ligner mest ut fra en k-nærmeste nabo (KNN) klassifikasjon.

To andre varianter av denne tilnærmingen basert på eROI er beskrevet av Harmsen som benytter Support Vector Machine (SVM) i stedet for KNN til klassifikasjon og Haak et al (2013) [8] som bruker support vektor regresjon. I en senere studie av Kashif et al, publisert i 2015 [9], brukes det i stedet for kryss-korrelasjon for hele boksen 6 SIFT (Scale Invariant Feature Transform) deskriptorer fra hver boks. Disse beregnes fra punkter som bestemmes etter en multi-nivå terskling og de 6 deskriptorene kombineres til en lang egenskapsvektor som representerer hver eROI. Deretter brukes SVM til å gjøre en klassifikasjon inn i én av 30 aldersklasser fordelt over alderstrinnene 1-18.

For flere av disse siste studiene er det gjort en sammenligning med BoneXpert ved å se på ytelse målt ved RMS-feil (Root Mean Square) på samme datasett, men ingen når helt opp til samme ytelse som BoneXpert. Metodene er imidlertid ikke trent opp på like mye data og er heller ikke på samme måte forbedret og tilpasset over lang tid. For de tidligere studiene er det i liten grad gjort sammenlignende studier.



Figur 8: Figur hentet fra Fischer et al. [6]. Bildet til venstre viser punkter definert av bruker som utgangspunkt for å finne eROI. Bildet til høyre viser de orienterte boksene som trekkes ut og analyseres videre.

4.2 MR-bilder

I det helt siste har det også blitt publisert noen få studier som adresserer automatisering av analyse av MR-bilder i sammenheng med estimering av skjelettalder. Det meste av dette arbeidet er gjort i Østerrike, hovedsakelig ved det tekniske Universitetet i Graz. Det følgende oppsummerer kort tilnærmingene de har brukt.

Fra «Determination of legal majority age from 3D magnetic resonance images of the radius bone» av Stern et al. 2014 [14]: Målet er å utvikle en automatisk metode for aldersbestemmelse basert på skjelettalder beregnet fra alle bein i hånden som sett i MR-bilder. Denne studien presenterer et første steg i denne retningen, og foreslår en metode basert på 3D segmentering av gapet mellom epifysen (knokkelenden) og metafysen (overgangen mellom leddenden (epifyse) og skaftet (diafyse)) for radialbeinet. De ser her spesielt på potensialet for å skille mellom personer som er over og under 18 år.

I studien antar de at en 3D segmentering av radialbeinet allerede er gjort. Ut fra en slik segmentering ser de så på hvordan en kan finne alle vokslene (pikslene i 3D) som utgjør gapet mellom epifysen og metafysen. Her bruker de random forest til å klassifisere vokslene som gap eller ikke-gap. Etter klassifikasjonen beregnes et normalisert mål, V , på volumet av gapet

basert på antallet vokslar og størrelsen på radialbeinet. Ettersom gapet gradvis lukkes gjennom utviklingsprosessen, modellerer de en sammenheng mellom stigende alder og avtagende volum og bruker lineær regresjon til å finne denne sammenhengen og senere estimere alder.

Fra “Towards automatic bone age estimation from MRI: Localization of 3D anatomical landmarks” av Ebner et al, 2014 [3]: Hovedbidraget i denne artikkelen er en metode for 3D lokalisering av anatomisk definerte interessepunkter gjennom en totrinns multiskala tilnærming basert på random regression forest. Først gjøres det en prediksjon av ganske grove posisjoner for punktene gjennom å analysere hele hånden. Resultatet fra dette trinnet brukes så i det vidare i en ny random regression forest som er begrenset av området fra det første trinnet og som fokuserer på lokale egenskaper. I denne studien gjøres det ingen aldersestimering.

Fra «Fully automatic bone age estimation from left hand MR images» av Stern et al, 2014 [13] og [28]: Her foreslås en fullstendig automatisk metode for aldersbestemmelse basert på skjelettalder beregnet fra alle bein i hånden basert på MR-bilder. Først brukes totrinns metoden basert på random regression forest (som beskrevet over) for å finne posisjonene for gapene mellom epifyse og metafyse for beina som skal analyseres. Deretter bestemmes automatisk en boks som omskriver gapet og det lages normaliserte «gap-bilder». Igjen brukes det så en regression random forest (RRF), denne gangen til å gjøre en egenskapsutvelgelse fra gap-bildet basert på noen enkle egenskaper fra bildeintensitet og differanser i tilfeldig genererte punkter og punktpar. Hensikten er å la Random Forest velge egenskapene som best mapper stadiet for sammenvoksing som observert i bildet til kronologisk alder. Det estimeres en skjelettalder for hvert gap, og den samlede skjelettalderen beregnes som en vektet sum av hver av disse.

4.3 Oppsummering og diskusjon

For automatisert estimering av skjelettalder er det to hovedproblemer som må løses:

(i) Automatisk segmentering av området som skal analyseres og (ii) uttrekking av karakteristika ved dette området som henger sammen med aldersutvikling. De ulike teknikkene adresserer dette på ulik måte.

I hovedsak er det to typer av områder i bildene som segmenteres (lokaliseres og avgrenses) for vidare analyse; enten konturen av beina i hånden slik som Thodberg gjør (Figur 7) eller en region rundt epifysen og/eller metafysen slik som for eksempel Fischer et al gjør (Figur 8). I noen tilnærmingar segmenteres beina først, for å kunne lokalisere ROI'ene etterpå. Der konturen av beina segmenteres, ser det ut til å være bruk av aktive konturmodeller som har hatt størst suksess.

De automatiserte metodene bruker stort sett ikke egenskaper som er basert direkte på karakteristikk som ville vært valgt ut dersom hensikten var å gjøre en visuell manuell analyse. I mange studier har de likevel en viss sammenheng med disse, gjennom at de beskriver for eksempel avstander, kanter, tekstur og lignende. I de fleste nyere studier brukes det i mye mindre grad spesifikke egenskaper. Her brukes enten bildeinformasjonen direkte, eller veldig generelle egenskaper som lokale deskriptorer. I studiene som er gjort på MR-bilder går de enda et skritt vidare og lar systemet lære seg egenskapene automatisk. Dette siste er også den generelle trenden innen bildeanalyse og gjenkjenning av mer komplekse objekter og strukturer. Her har teknikker basert på Deep Learning, som selv lærer egenskaper, vist en ytelse i form av gjenkjenningsrater som ikke har vært oppnådd tidligere.

Fram til 2009 baserer de fleste av de automatiserte metodene seg på samme prinsipp som de manuelle gjennom at de først klassifiserer hånden eller hvert enkelt bein inn i utviklingstrinn. Deretter bruker de en modell for sammenhengen mellom utviklingstrinn og alder (ofte regresjonsbasert) til å gi et aldersestimert. Fordelen ved denne tilnærmingen er at en enkelt kan bruke forskjellige regresjonsmodeller for ulike grupper inndelt for eksempel etter kjønn, etnisitet o.l., og så samme system for resten. Ved den andre tilnærmingen blir disse to delene noe tettere integrert og det vil kreve at en har tilgang på ulikt opptrente klassifikatorer. Fordelen ved den siste tilnærmingen er at en unngår et ekstra trinn med kvantisering der informasjon kan gå tapt.

5 Automatisering av metode fra Universitetet i Roma

Målet med arbeidet her har i utgangspunktet vært å se på hvordan en kan automatisere MR-metoden for estimering av skjelettalder som er utviklet ved Universitetet i Roma.

Som vi har sett av de ulike studiene av automatiske tilnærminger til problemet for røntgen, er det sjelden slik at metodene beregnet på manuell visuell analyse, direkte lar seg overføre til en automatisk prosess. En av årsakene til dette er, som Tanner observerte, at selv om den manuelt definerte metoden i utgangspunktet er regelbasert, vil spesialistene som utfører analysen etter hvert også benytte mye erfaringsbasert informasjon som ikke inngår i selve metodebeskrivelsen. En annen årsak er at det ikke alltid er samsvar mellom hva som lett fanges opp av øyet og hva som enkelt kan trekkes ut med automatiske metoder. Samtidig har automatiske metoder andre fordeler, for eksempel gjennom å huske bedre, bedre generalisere over mange eksempler, se på flere bilder i sammenheng, raskere kunne gjøre ulike beregninger basert på bildet osv.

Det er derfor usikkert om det er nyttig å ta utgangspunkt i en framgangsmåte utarbeidet til bruk for manuell estimering av skjelettalder når en skal utvikle automatiske metoder, eller om det er bedre å starte med blanke ark. Det er derfor nødvendig med en vurdering av de ulike delene av den manuelle prosessen med tanke på hva som kan ha noe for seg også i en automatisk prosess.

Da prosjektet opprinnelig ble definert (høsten 2013), var det foreløpig ingen som hadde sett på metoder for automatisert bestemmelse av skjelettalder basert på MR-bilder. Men etter dette har det blitt gjort noen interessante arbeider på dette i Graz, Østerrike. Det kan derfor være aktuelt å vurdere om vi bør søke å se på løsninger som er inspirert av det som er gjort her eller prøve ut tilnærminger som er veldig forskjellig fra dette.

Som vist i beskrivelsene i forrige avsnitt, vil et system for estimering av skjelettalder måtte løse flere forskjellige problemer. I de studiene som har kommet fram til gode løsninger, har arbeidet da også pågått over mange år med mange involverte. Innenfor dette prosjektet derimot er ressursene svært begrenset og fokuset må derfor også begrenses. Det betyr at vi bare vil ha mulighet til å se på enkelte utvalgte deler av problemet.

Basert på disse aspektene vil vi i det følgende gjøre en vurdering av de ulike delene av den manuelle metoden med tanke på om og hvordan disse kan automatiseres innenfor denne studien.

5.1 Vurdering av den manuelle prosessen med tanke på automatisering

For å rekapitulere består den manuelle prosessen som definert av Universitetet i Roma av følgende steg. Først, for hvert av 9 bein:

- Utvelgelse av bilde-sett og snitt der beinet og utvikling er best synlig
- Identifisere hvert bein/område i bildet
- Analysere området med hensyn på de definerte egenskapene.
- Bestemme utviklingstrinn basert på observerte egenskaper

Deretter, samlet for alle analyserte bein:

- Skjelettalder estimeres ut fra utviklingstrinnene bestemt for hvert bein og ved hjelp av en oppslagstabell som definerer sammenhengen mellom summen av utviklingstrinn og alder.

5.1.1 Utvelgelse av bildesett og bildesnitt

I motsetning til røntgen gir MR mange bildesnitt og mulighet for å gjøre analyser i 3D. Den manuelle prosessen som definert over, velger ut enkeltsnitt og er dermed i hovedsak 2D-basert. Dette er naturlig for en manuell prosess ettersom det å gjøre en 3D-analyse manuelt er vanskelig. En automatisk analyse derimot gjør det mulig å utnytte informasjon fra alle snittene, og i tilnærmingene beskrevet fra Graz, gjøres det en 3D-analyse. Generelt gir også mer informasjon bedre resultater, i hvert fall hvis det fins nok treningsdata. Det fins foreløpig ingen andre kjente studier på automatisering av denne prosessen for MR-bilder, og dermed ingen som har sett på bruk av 2D med bare enkeltsnitt.

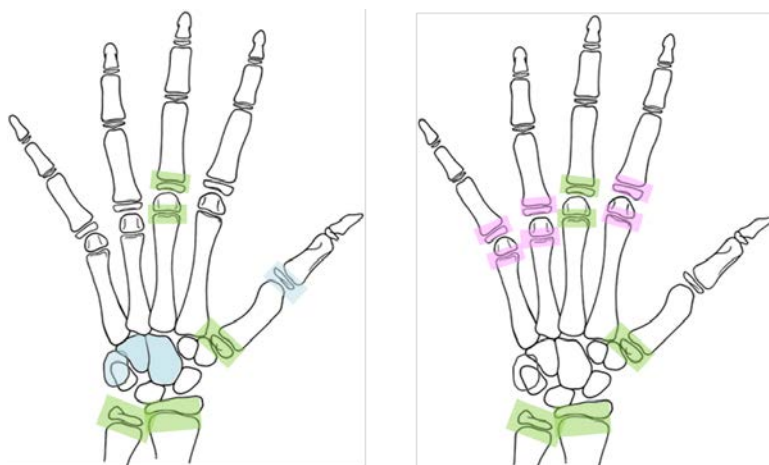
I studien fra Roma vil et opptak vanligvis bestå av to bildesett, ett av nedre og ett av øvre del av hånden. I andre studier vi har sett, har det vært nok med ett sett med bilder. Årsaken til at det i de andre studiene ikke er nødvendig med to sett for å dekke hele hånden, er trolig at det der er brukt en større MR-maskin. Problemstillingen rundt valg av bildesett ser dermed ut til å være spesifikk for denne typen MR-opptak.

Hvert bildesett vil bestå av 26 snitt. I en eventuell fullstendig automatisk prosess må alle disse bildene inngå i analysen. Enten gjennom at det automatisk velges ut hvilke bildesett og bildesnitt som skal analyseres for hvert bein, eller gjennom at alle bildene brukes til å lage og analysere en 3D-representasjon av hånden. I den manuelle prosessen gjøres det et utvalg av enkeltsnitt for hvert bein.

5.1.2 Identifisering av område for videre analyse

Enten analysen foregår i 2D eller i 3D, må området som skal analyseres identifiseres. Den manuelle analysen brukt i Roma og den automatiske analysen som beskrevet i Graz ser på litt forskjellige områder og bein i hånden, som illustrert i Figur 9.

Hvilke områder i hånden en fokuserer på, vil også kunne henge sammen med hvilke aldersgrupper en typisk vil komme til å se på, ettersom de ulike beina i hånden forventes å være ferdig utviklet ved ulike alderstrinn. Figur 5 som er basert på studiene fra Roma og viser sammenheng mellom alderstrinn og utvikling av de ulike beina i hånden, illustrerer dette. Den viser at håndrotsbeina (bein 5, 6 og 9 i Figur 1) for gutter er ferdig utviklet rundt 15-års alder, radial og ulnar er ferdig utviklet rundt 19-års alder, mens de resterende er ferdig utviklet rundt 16-års alder.



Figur 9: Figuren indikerer hvilke områder som benyttes i den manuelle analysen som definert av Universitetet i Roma (venstre) og i den automatiske metoden som utviklet i Graz (høyre). Områdene merket med grønt er de som begge metodene ser på, mens de merket i blått bare benyttes i Roma-metoden og de rosa bare benyttes i Graz-metoden.

For beina i håndroten er form på beinet og avstand til andre bein de viktigste egenskapene. For de andre beina er detaljer ved formen også en del av den visuelle analysen, men hovedfokus er på egenskaper ved sammenvokningsflatene. Hvilke aspekter som er viktige for analysen, kan ha noe å si for hvordan interesseområdene defineres. I hovedsak er det to måter å definere interesseområder på som typisk brukes, der områdene enten defineres gjennom å finne konturen av området eller gjennom å posisjonere en rektangulær ROI (Region Of Interest). For beina inni håndroten, kan et område som er definert ut fra konturen kanskje være mest hensiktsmessig, mens for sammenvokningsflatene, ser det for MR ut til at ROier kan være best.

5.1.3 Utvelgelse av egenskaper

I den manuelle MR-metoden fra Roma er det definert egenskaper i form av tekstlige beskrivelser av utviklingen slik den framstår visuelt for hvert utviklingstrinn og for hvert bein som observeres.

For håndrotsbeina kapitat og hamat er utviklingen definert i forhold til tre lag i og rundt beinet, omtalt som indre, mellomliggende og ytre lag. For pisiform ses bare indre og ytre lag. Beskrivelsene av utviklingen er knyttet til endringer i størrelse, form, bredde på konturen og avstand til andre bein. For første metakarpal (innerste tommelbein) er utviklingen definert for epifysen og vekstsonen, mens for de resterende områdene er egenskaper beskrevet for utviklingen av epifysen, vekstsonen og metafysen. Her beskrives utviklingen gjennom endringer i størrelse, form, bredde, tekstur i vekstsonen og lukking av denne.

Det er vanskelig å lage automatisk beregnede egenskaper som direkte etterligner de manuelt definerte, men det er mulig å definere spesifikke bildebaserte egenskaper som forsøker å fange opp noe av det samme. Trenden i nyere studier ser ellers heller ut til å være å bruke mer generelle egenskaper som beskriver områdene.

5.1.4 Bestemmelse av utviklingstrinn

Den manuelle metoden går via utviklingstrinn for å komme fram til alder. For en manuell analyse kan dette være en naturlig framgangsmåte. Radiologen behøver da bare å kjenne til

definisjonen av utviklingstrinn, uten å kjenne til hvordan disse henger sammen med alder for ulike grupper (kjønn, etnisiteter e.l.). Ulempen er at det introduseres et ekstra trinn som innebærer kvantisering.

For en automatisk analyse kan en like gjerne modellere sammenheng mellom alder og bilde direkte der informasjon ikke går tapt i en kvantisering underveis. Men dersom analysen skal brukes på nye grupper, der denne sammenhengen avviker fra det systemet er trent opp for, er en avhengig av å ha tilgjengelig nye bildesett som er tilstrekkelig store til å gjøre en slik ny opptrening. Da er det kanskje ikke behov for like mye treningsdata for å modellere sammenhengen mellom utviklingstrinn og alder for en ny gruppe, slik at dette likevel kan være en fordel. Dette vil da også være et prinsipp radiologene kjenner seg igjen i.

5.1.5 Estimering av skjelettalder fra utviklingstrinn

Italienerne definerer sammenhengen mellom utviklingstrinn og alder gjennom å summere numrene definert for utviklingstrinnene identifisert for hvert av de 9 beina og å lage en regresjonsmodell for sammenhengen mellom summen og alder. Her kunne det vært valgt andre strategier, som for eksempel multivariat regresjon basert direkte på de 9 variablene som utgjøres av utviklingstrinnene bestemt for hvert bein. Det kan imidlertid se ut som det er litt for lite data for en slik type analyse. I vår studie vil vi fokusere på bildeanalysen og vil derfor ikke gå inn på denne delen av metoden.

5.2 Forslag til automatisert metode for MR-bilder

Basert på en vurdering av punktene diskutert i forrige avsnitt, har vi valgt en strategi som utnytter mye av kompetansen på denne typen MR-bilder som er opparbeidet i Roma samtidig som den nyttegjør seg potensialet som finnes i en mer automatisert framgangsmåte. Her er det da ikke nødvendigvis fornuftig å basere en automatisk analyse fullstendig på en manuell prosess. Dette fordi automatiske tilnærminger har andre styrker og svakheter enn manuelle.

I det følgende beskrives den overordnede strategien vi har valgt, samt datasettet vi vil bruke i analysen. I avsnitt 5.3 beskrives framgangsmåten i mer detalj, mens avsnitt 5.4 presenterer de foreløpige resultatene.

5.2.1 Overordnet strategi

Som en overordnet strategi for bildeanalysen vil vi ta utgangspunkt i utviklingstrinn og bein som definert i Roma, og se på om og hvordan en kan bestemme utviklingstrinn for definerte bein i hånden basert på automatisk bildeanalyse. Vi velger å gå via utviklingstrinn både fordi det er slik den manuelle framgangsmåten er definert og fordi vi tror det er en fornuftig strategi når det er behov for både å kunne ha ulike modeller for sammenheng mellom utviklingstrinn og alder for ulike grupper og hvor det underveis kan være nødvendig å utvide modellene til å takle nye grupper med annen sammenheng mellom alder og utvikling.

For å kunne gå via utviklingstrinn i vår studie er det nødvendig med en fasit i form av hva som er riktig utviklingstrinn for hvert bein for hvert individ. Til dette vil vi benytte de manuelt tilordnede utviklingstrinnene oppnådd gjennom analysen i Italia.

Vi vil begrense oss til å se på bare ett av beina som er benyttet i den manuelle analysen, hvor vi har valgt å fokusere på radialbeinet. Dette fordi vi er nødt til å begrense omfanget, samtidig er dette (sammen med ulnar) det beinet i hånden som er i utvikling helt fram til 18/19-årsalderen. Om vi klarer å demonstrere en metode for radial, vil det indikere at det kan være tilsvarende potensiale for de andre.

For utvalg av bildesett og bildesnitt for analysen, vil vi i denne omgang bruke veldig enkle strategier. For automatisk valg av bildesett velger vi å ikke utvikle noe spesielt, men bruker bare bildesettet av nederste del av hånden (mot håndleddet) siden vi forventer at radialbeinet vil være omfattet av bildene fra denne delen. For utvalg av sentrale snitt for radialbeinet vil vi bruke en enkel algoritme som tar utgangspunkt i snittet der radialbeinet har størst utstrekning. Vi vil i analysen ikke benytte full 3D, men ser det som aktuelt å benytte mer enn ett snitt. Ved bruk av flere snitt, vil vi vurdere litt ulike strategier der vi kan trekke ut egenskaper fra hvert snitt for seg.

Basert på nyere arbeid på røntgenbilder, samt det som har vært gjort på automatiserte metoder for MR, velger vi å benytte en tilnærming som benytter egenskaper definert fra en rektangulær «Region of Interest», ROI. Segmenteringen av ROIen vil vi måtte automatisere for å kunne kjøre analysen på et større antall bilder, men i denne studien gjør vi dette enklest mulig. I en reell situasjon antar vi at dette finnes, ettersom det allerede er mange som har utviklet metoder for dette basert både på røntgen og MR.

Fra den definerte ROIen vil vi trekke ut bildebaserte egenskaper som kan egne seg for å klassifisere ROI til utviklingstrinn. Vi vil ikke basere disse egenskapene direkte på de tekstlige beskrivelsene som er laget for den manuelle bestemmelsen av utviklingstrinnene, men vil i stedet vurdere mer generelle egenskaper som kan fange opp noe av det samme. Basert på egenskapene vil vi gjøre en klassifikasjon inn i utviklingstrinn.

5.2.2 Datasett

Vi tar utgangspunkt i bildene fra Roma tatt opp med liten MR, 0.2 T. Dette er en litt annen type maskin og opptak enn det vi ser andre har brukt. Det er imidlertid for denne maskinen vi har tilgang til et større bildemateriale med tilhørende informasjon, og det er også basert på disse opptakene den manuelle metoden fra Roma er utviklet og basert. Selv om metoden som sådan trolig er overførbart til opptak fra andre typer av MR-maskiner.

Vi benytter bildene med geometri og oppløsning som mottatt fra Roma. Dette ser også ut til å stemme overens med det de har brukt i sine studier og med bildene som er gjengitt i deres publikasjoner.

Siden det er nødvendig å begrense omfanget av arbeidet, har vi valgt å se på en begrenset aldersgruppe i analysen. Vi har da valgt å se på opptak fra gutter i alderen 14-18/19. Fra Roma har vi fått følgende bildeopptak av italienske gutter i denne aldersgruppen:

- 14 healthy males included in the range age of 14y1m-14y11m (CA)
- 19 healthy males included in the range age of 15y1m-15y11m (CA)
- 7 healthy males included in the range age of 16y.o.-16y11m (CA)
- 12 healthy males included in the range age of 17y1m-17y11m (CA)
- 8 healthy males included in the range age of 18y2m-18y11m (CA)
- 3 healthy males of 19 y.o. (2) and 19y1m (1) (CA)

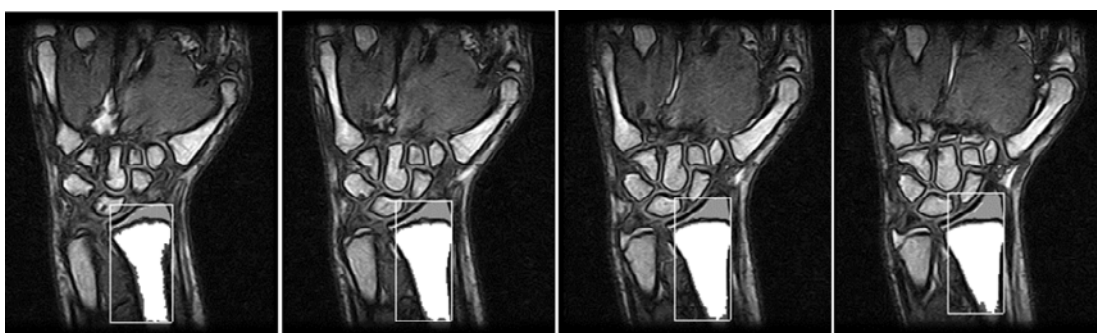
Det betyr at vi har totalt 63 opptak. Det er en del i underkant av det som ville vært ideelt, men det er bare dette som er tilgjengelig. For noen av aldersklassene er det også veldig få opptak. Bildesettene er anonymisert, men for hvert opptak følger det med kronologisk alder. I tillegg har vi for hvert sett fått informasjon om hvilket utviklingstrinn som er bestemt for radialbeinet i den manuelle analysen. Vi har også fått informasjon om hvilke(t) bildesnitt analysen av radialbeinet i hovedsak er basert på.

5.3 Beskrivelse av automatisert analyse

I det følgende beskrives trinnene i den automatiserte analysen som består av (i) Segmentering av radialbeinet, (ii) automatisk valg av snitt, (iii) bestemmelse av ROI, (iv) beregning av egenskaper og (v) klassifikasjon i utviklingstrinn.

5.3.1 Segmentering av radialbeinet

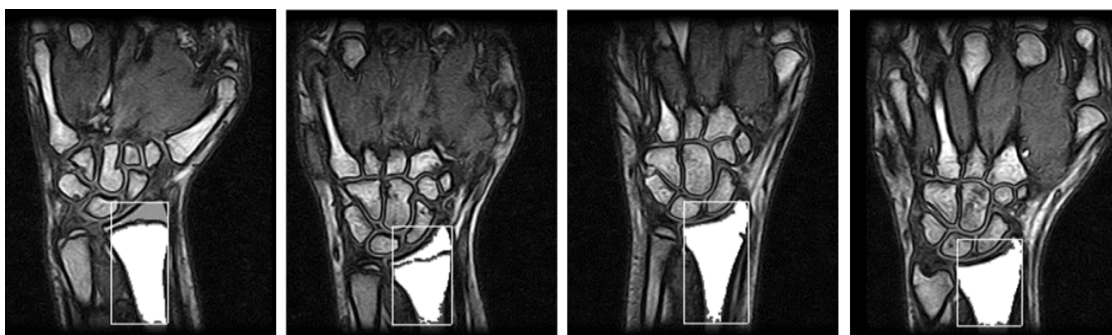
Her er det flere som har utviklet og demonstrert med gode resultater hvordan dette kan gjøres både for røntgen og MR og for 2D og 3D. Vi antar derfor at en slik type segmentering vil kunne være tilgjengelig og har ikke lagt vekt på utvikling av en robust metode for dette. For å kunne gjøre den videre analysen er vi likevel avhengige av å ha gjort en slik segmentering. Ettersom fullstendig manuell segmentering vil være veldig tidkrevende når det er snakk om en del bilder, har vi derfor basert segmenteringen på en veldig enkel algoritme. Algoritmen baserer seg på å finne store sammenhengende lyse komponenter i nedre høyre kvadrant av bildet. Figuren nedenfor viser resultatet av segmenteringen for 4 etterfølgende snitt for ett opptak fra en gutt på 15 år.



Figur 10: Segmentering av radialbeinet. Figuren viser resultatet av segmenteringen for 4 etterfølgende snitt for ett MR-opptak fra en gutt på 15 år. Segmentert område er markert i hvitt og grått og den omskrevne boksen er tegnet inn.

5.3.2 Utvalg av snitt

Ved hjelp av algoritmen beskrevet over analyseres hvert snitt i MR-opptaket. Det er bare de snittene der beinet synes godt vi vil være interessert i for den videre analysen. Vi bruker her også en enkel metode som velger det snittet der radialbeinet er størst som utgangspunkt fordi vi ønsker å bestemme størrelse og posisjon for regionen som skal analyseres ut fra dette. Utvalg av snitt kan eventuelt også baseres på dette, eller velges som et fast snitt i sekvensen. Vi kan deretter i tillegg velge ut et antall snitt på hver side av dette. Figur 11 viser snittet metoden har valgt ut som utgangspunkt for fire ulike MR-opptak av fire gutter på hhv. 15, 16, 17 og 18 år.

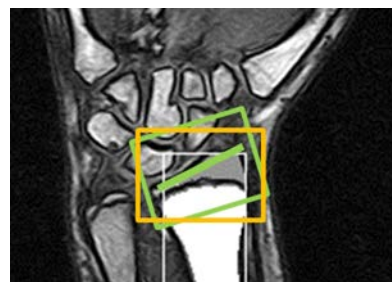


Figur 11: Automatisk utvalg av MR-snitt. Figuren viser snittene som er valgt ut som utgangspunkt for fire ulike MR-opptak av fire gutter på hhv. 15, 16, 17 og 18 år.

5.3.3 Bestemmelse av ROI

For å trekke ut egenskaper ønsker vi å bestemme en region som omskriver området som skal analyseres, en «Region Of Interest» (ROI). Basert på analysen gjort over, ser vi noen mulige alternativer for hvordan en slik ROI kan defineres:

1. ROI defineres som et rektangel orientert i henhold til sammenvokningsflaten. Orienteringen vil da være relativt lik bildets orientering og for å unngå resampling (som kan degradere bildet) kan dette enkelt gjøres ved bare å plukke ut øvre del av identifisert boks fra segmenteringen (oransje rektangel i Figur 12).
2. ROI defineres som et rektangel orientert i henhold til øvre kant av radialbeinet ved å finne øverste venstre og høyre punkt for radialbeinet og deretter lage et skrått orientert rektangel basert på dette (grønt rektangel i Figur 12).
3. ROI defineres som en mellomting av de to foregående ved å plukke ut øvre del av boksen som i punkt 1, men så fjerne informasjonen i overkant av radialbeinet for ikke å ta med informasjon fra beina inne i hånden (se Figur 13, høyre).



Figur 12: Alternativer for utplukking av ROI rundt radialbeinet.



Figur 13: ROI plukket i henhold til alternativ 1 (venstre) og alternativ 3 (høyre).

I analysene som er gjort i denne studien har vi valgt alternativ 1 for utplukking av ROI, slik at vi får utsnitt tilsvarende det som er vist til venstre i Figur 13.

5.3.4 Uttrekking av egenskaper

Vi ønsker et sett med egenskaper som er i stand til å fange opp mange av de ulike karakteristikkene ved utviklingstrinnene definert for radialbeinet. De manuelle beskrivelsene fra Roma omfatter både form, tekstur og avstander relatert både til bein og brusk. Det er vanskelig å omsette dette direkte til veldefinerte egenskaper som lar seg trekke ut automatisk. I stedet ønsker vi derfor å basere oss på et sett av nye metoder som har hatt stor suksess i det siste på andre områder innenfor bildegjenkjenning og som selv er i stand til å lære egenskaper fra dataene. Disse metodene, omtalt som Deep Learning, lærer gode datarepresentasjoner direkte fra bildedataene og har ført til store gjennombrudd innen bildebasert gjenkjenning (Krizhevsky, 2012) [10]. Dette er mulig gjort gjennom en hierarkisk representasjon ved hjelp av dype nevrale nettverk som er trent opp på et stort antall bilder merket med klasser.

I dette tilfellet vil vi ikke ha tilgang på et slikt stort treningsmateriale. Helt nye studier har imidlertid vist at det er mulig å utnytte såkalt «transfer learning» til å lære generiske datarepresentasjoner fra et område med mye treningsdata for så å benytte den lærte representasjonen på et problem fra et annet område, som i Azizpour et al, 2014 [1]. Vi vil derfor ta utgangspunkt i dette og benytte egenskaper beregnet med såkalte convolutional

neural networks (CNN), der disse nettverkene har vært trent opp på andre områder. Her finnes det blant annet flere slike ferdig trente nettverk med litt ulike arkitekturer, som har vært trent på et sett av 1.2 millioner bilder merket med 1000 klasser (ImageNet, [4]), hvor flere av disse nettverkene er åpent tilgjengelig.

5.3.5 Klassifikasjon

Vi ønsker å gjøre en klassifikasjon inn i utviklingstrinn for radialbeinet som definert gjennom arbeidet i Roma. Settene vi har fått fra Roma er merket med utviklingstrinn for radialbeinet per individ som bestemt gjennom deres manuelle analyse. Med bare 63 personer, har vi alt for lite data til å gjøre selve klassifikasjonen med et nettverk. Her vil vi derfor i stedet bruke en lineær SVM (Support Vector Machine). For ikke å trene og teste på samme individer, vil vi bruke leave-one-out på individ-nivå. Det vil si at vi for hver av de 63, gjør klassifikasjon på den ene, mens vi trener opp klassifikatoren på de 62 andre.

Vi gjør treningen og klassifikasjonen per snitt. Per person blir det da et sett med klassifikasjonsresultater som kan være litt ulike. Vi kan da for eksempel velge den klassen (utviklingstrinnet) som opptrer oftest som endelig resultat.

5.3.6 Aldersestimering

For enklere sammenligning med andre studier, ønsker vi også å kunne beregne et aldersestimat. Siden vi i denne analysen bare ser på radialbeinet, kan vi ikke bruke metoden som er definert i Roma der det tas hensyn til beregnet utviklingstrinn for 9 ulike bein i hånden. Vi har derfor valgt en enkel tilnærming der vi beregner aldersestimatet fra midtpunktet av aldersintervallet for utviklingstrinnet som er bestemt. Når vi skal beregne aldersestimatet, bruker vi så midtpunktet for klassifikasjonsresultatet fra hvert snitt som er brukt i analysen:

$$alder = \left(\sum_{i=1}^N m(k_i) \right) / N$$

N = antall snitt brukt i analysen

k_i = klassifisert utviklingstrinn for snitt i , $k=\{8:12\}$.

$m(k)$ = midtpunkt for aldersintervallet for utviklingstrinn k .

Utviklingstrinn	8	9	10	11	12
Alder intervall	12.5 - 14	14 - 15.5	15.5 - 16.5	16.5 - 17.5	17.5 - 19
Alder midtpunkt	13.25	14.75	16	17	18.25

Tabell 2: Oversikt over sammenheng mellom utviklingstrinn og aldersintervall (som bestemt av metoden utviklet i Roma) hvor vi også har angitt midtpunktet for hvert intervall.

5.4 Eksperimenter og foreløpige resultater

5.4.1 Framgangsmåte

Det følgende beskriver framgangsmåten mer konkret gjennom hvilke valg av parametre og andre valg som er gjort.

5.4.1.1 Segmentering

Segmenteringen av radialbeinet er gjort som beskrevet i avsnitt 5.3.1. Deretter bestemmes det snittet hvor radialbeinet har størst utstrekning. ROIen for dette opptaket bestemmes som den omskrevne boksen for radialbeinet i dette snittet (alternativ 1 i avsnitt 5.3.3). Dette

alternativet er valgt på grunn av enkelhet og fordi vi da unngår en resampling av bildene. Vi har i første omgang også valgt å ta med all informasjonen innenfor regionen.

Vi har som tidligere nevnt ikke lagt mye vekt på segmenteringsalgoritmen i denne studien. Her finnes det mange gode metoder, men disse er mer tidkrevende å implementere. For noen av MR-opptakene ser vi at den enkle metoden bommer noe. Den videre metoden er ikke veldig følsom for helt nøyaktig posisjonering, men hele radialbeinet bør være innenfor regionen. For disse har vi derfor manuelt justert koordinatene for boksen som skal trekkes ut. Når ROIen for snittet der radialbeinet med størst utstrekning er bestemt, brukes denne til å trekke ut regioner for alle snittene som skal analyseres fra dette opptaket.

Bredden på ROIen slik vi har definert det, bestemmes av bredden på radialbeinet. For høyden på boksen har vi sett på noen ulike alternativer hvor høyden enten settes lik bredden eller noe lavere for ikke å ta med så mye informasjon nedenfor sammenvokningsflaten.

5.4.1.2 Utvalg av snitt

I eksperimentene har vi sett på litt ulike måter å velge snitt på til analysen. Her er det noen ulike valg som kan gjøres:

- 1) *Valg av sentralsnitt.* Dette er snittet det tas utgangspunkt i for utvalg av resten. Her har vi sett på tre ulike alternativer:
 - Snitt med størst radialbein
 - Snitt som angitt av Roma (første når de angir to eller flere)
 - Et fast snitt for alle (valgt som gjennomsnitt av valg fra Roma)
- 2) *Speiling av snitt.* I utvalget av bilder vi har fått fra Roma, er 7 av de 63 bildene tatt av høyre hånd (høyre hånd er valgt for venstrehendte). I analysen er det viktig at alle snittene opptrer med samme orientering(er). Dette oppnår vi på en av følgende måter:
 - Speiling av alle snitt, der vi tar med begge varianter for alle. Totalt sett får vi da dobbelt antall snitt til bruk i analysen.
 - Speiling av de 7 høyrehåndssnittene, sånn at vi får et sett der radialbeinet er venstreorientert i alle snitt.
- 3) *Valg av antall snitt.* Ut fra sentralsnittet velges et antall snitt sentrert om dette. Her har vi testet med følgende ulike antall: 5, 7, 9, og 11 snitt fra de totalt 26 snittene. Dersom speiling av alle snitt velges, tilsvarer det et totalt antall snitt på hhv. 10, 14, 18 og 22 per opptak.
- 4) *Valg av y-dimensjon (høyde) på snitt.* Bredden på snittet er alltid bestemt av bredden observert på radialbeinet i snittet der dette har størst utstrekning. For høyden på snittet har vi sett på tre ulike alternativer:
 - Høyde lik bredde (ca. 64 piksler).
 - Høyde rett i underkant av sammenvokningsflaten (ca. 44 piksler).
 - Høyden som en mellomting av disse (ca. 54 piksler).

For alle utvalgene benyttes regionen (koordinatene) som er bestemt (og i blant manuelt justert) for snittet der radialbeinet er funnet å ha størst utstrekning. I henhold til valgene over, varierer antallet snitt som er brukt per person fra 5 (uten speiling) til 22 (med speiling).

Alle kombinasjonene beskrevet over er ikke testet ut. I stedet har vi forsøkt å avgrense dette gjennom noen innledende eksperimenter:

- 1) *Valg av sentralsnitt*: Her ga valg av et fast snitt best resultater. Det faste snittet var satt til snitt nummer 16, som er et gjennomsnitt av valgene fra Roma. Mer robust segmentering kan bety at estimatet for største radialbein blir bedre. Dermed kan denne måten å velge ut snitt på fungere bedre enn det vi har sett av eksperimentene. Romas måte å velge på er basert på at de bare studerer noen veldig få snitt (1-3). Det snittet de ser på kan derfor i enkelte tilfeller ligge langt ut mot endene, slik at det ikke egner seg som utgangspunkt når vi velger et antall snitt på hver side av dette. Deres valg av snitt varierer fra snitt 11 til snitt 21 for MR-opptakene vi har sett på i denne studien.
- 2) *Speiling av snitt*: Av de innledende testene ser det ut til å være litt bedre å ta med både høyre- og venstrespeilet bilde, i stedet for bare venstre. Bruk av begge varianter kan forenkle prosessen noe, da systemet ikke er avhengig av å vite om opptaket er gjort av venstre eller høyre hånd. I det videre har vi derfor brukt begge variantene i klassifikasjonen.
- 3) *Valg av antall snitt*: Her har vi sett på bruk av 2x5, 2x7, 2x9 og 2x11 snitt. Dårligst ut kom 2x11 og 2x9, mens 2x5 og 2x7 ga ganske like resultater.
- 4) *Valg av høyde på snitt*: Her ga alle varianter ganske like resultater, men mellomhøyden var hårfint bedre.

Basert på dette har vi i det følgende (hvis ikke annet er angitt) brukt fast valg av sentralsnitt, speilet alle valgte snitt, valgt 5 snitt fra hver speiling og valgt mellomhøyden for utsnittene.

5.4.1.3 Beregning av egenskaper

Beregningen av egenskaper er gjort med ferdig trent CNN (convolutional neural nets). Nettverkene er først pretrent på bilder fra ImageNet, og deretter er det samme nettet brukt til å trekke ut egenskaper fra alle ROIene fra alle snittene. (Totalt 315 til 1386 snitt).

Vi har sett på to ulike nett. Det ene, såkalt AlexNet [10] gir 4096 egenskaper og er trent opp på ImageNet [4] i full oppløsning. Det tar som input 227x227 bilder. For våre bilder som er i størrelsesorden 64x64, betyr det en del oppskalering. Det betyr igjen at det er mange egenskaper som blir null. Det kunne vært aktuelt å gjøre en egenskapsreduksjon, men vi har ikke prioritert det i denne studien. Det andre nettet vi har sett på tar en input på 64x64 bilder og gir 128 egenskaper. Dette er også trent opp på ImageNet, men siden bildene her er større enn 64x64 blir de først nedsamlet, noe som kan gi artefakter og en uheldig bildekvalitet for egenskapsberegning. Det kunne også vært interessant å se på noen flere, eventuelt trent opp noen nett selv, men det har det da heller ikke vært rom for.

Innledende tester viste at 64x64-nettet ga generelt noe dårligere resultater enn AlexNet, så resultatene vi presenterer i det videre er basert på egenskapsuttrekking med AlexNet.

5.4.1.4 Klassifikasjon

Klassifikasjonen er gjort med lineær SVM per snitt og leave-one-out på individnivå. Resultatet per person er beregnet ved «majority vote» for alle snitt fra den personen.

5.4.1.5 Aldersestimat

Aldersestimatene er beregnet som beskrevet i avsnitt 5.3.6 fra resultatene av klassifikasjonen av alle snitt for hver person.

5.4.2 Resultater

All analysen er gjort på datasettet beskrevet i avsnitt 5.2.2, bestående av MR-opptak fra 63 gutter i alderen 14-19 år. Resultatene er her oppsummert for klassifisering i alderstrinn og beregning av aldersestimat basert på metodene beskrevet over. Resultatene er ellers basert på valgene som beskrevet over for hvert av punktene. Datasettet vi har sett på i denne studien er imidlertid såpass lite, at dette for en endelig automatisert metode er noe som burde testes grundigere ut på et betydelig større sett.

5.4.2.1 Klassifikasjon i utviklingstrinn

Klassifikasjonen i utviklingstrinn er gjort som beskrevet i 5.3.5, hvor hvert snitt klassifiseres separat, og deretter telles det opp hvilken klasse som opptrer oftest. Som fasit for hva som er riktig trinn, bruker vi de manuelt bestemte trinnene for radialbeinet fra Roma. For de ulike testene som er kjørt ligger den korrekte klassifikasjonsraten da typisk på 70-75%. Det vil si at vi for 70-75% av de 63 personene får samsvar med det manuelt bestemte utviklingstrinnet. For de 25-30% hvor vi ikke får samsvar, ser vi at vi får maksimalt ett trinn i avvik. Tabell 3 viser resultatet fra klassifikasjonen i form av en forvirringsmatrise, der det også framgår hvor forvekslingene skjer. Her er klassifikasjonsraten 74.6%.

	Manuelt bestemt utviklingstrinn					
	Trinn	8	9	10	11	12
Automatisk bestemt utviklingstrinn	8	4				
	9	1	15	5		
	10		4	5	2	
	11			1	7	1
	12				2	16

Tabell 3: Forvirringsmatrise for klassifikasjon i utviklingstrinn.

Fordi utviklingstrinnene er kategoriske (ikke kontinuerlige), fanger ikke forskjellen mellom manuelt og automatisk bestemt utviklingstrinn nødvendigvis opp hvordan denne forskjellen kan slå ut i form av estimert alder. Som et eksempel på dette viser vi et enkeltresultat:

- Kronologisk alder: 15,5 år
- Manuelt bestemt utviklingstrinn for radialbein fra Roma:
 - 10 (tilsv. aldersgruppe 15,5-16,5)
- Klassifisering av utviklingstrinn for hvert snitt med automatisk metode:
 - Sekvens (2x5 snitt): 10, 10, 9, 9, 9, 10, 11, 9, 9, 9,
 - Majoritet: 9 (tilsv. aldersgruppe 14-15.5)
 - Gjennomsnitt: 9.5
- Estimert alder Roma: 15,67 år (avvik +0.17)

- Estimert alder NR: 15,35 år (avvik -0.15)

Her ser vi at manuelt og automatisk bestemt utviklingstrinn avviker med ett trinn, men om vi beregner et aldersestimat som foreslått i 5.3.6, er absoluttavviket fra kronologisk alder ganske likt for de to tilfellene. På bakgrunn av dette har vi i neste avsnitt også sett på og sammenlignet kronologisk alder med manuelt og automatisk beregnede aldersestimater.

5.4.2.2 Aldersestimering

I det følgende har vi sammenlignet avviket fra kronologisk alder for aldersestimering gjort med tre ulike metoder:

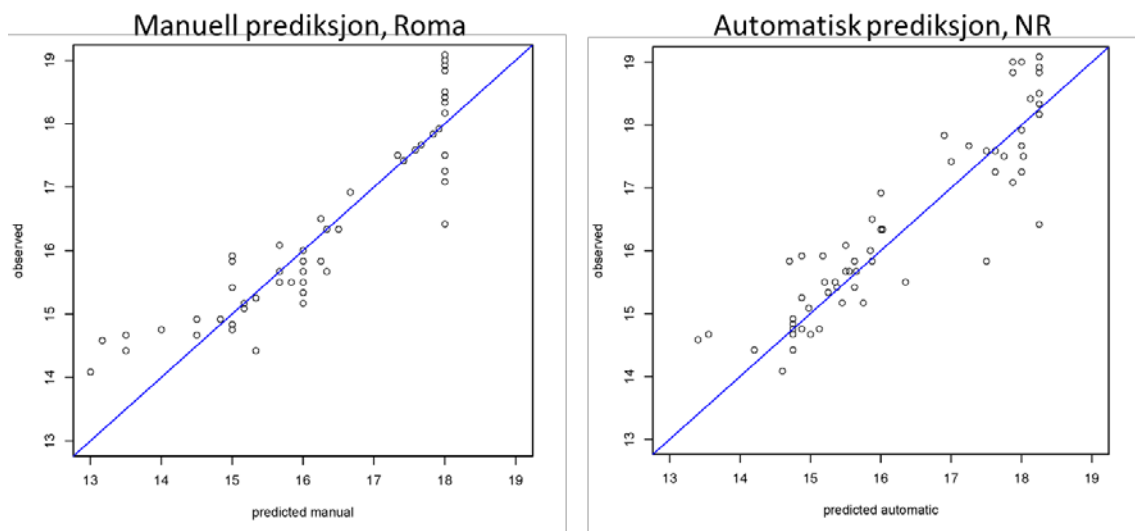
- *Manuell metode, Roma*: Metoden her er basert på at to personer manuelt har tilordnet utviklingstrinn for 9 bein som observert i MR-bildet. Deretter har de ved konsensus bestemt endelige utviklingstrinn, summert score og avledet aldersestimatet fra dette. For de over 16 år vil det her i hovedsak bare være radial og ulnar som fortsatt er i utvikling.
- *Automatisk metode, NR*: Som beskrevet i tidligere avsnitt, der utviklingstrinn og aldersestimat foreløpig bare beregnes fra radialbeinet.
- *Automatisk metode, Graz [13]*: Aldersestimering med denne metoden er ikke gjort på samme datamateriale som de andre. Her er estimeringen gjort på et eget sett tatt opp med en annen type MR-maskin fra 102 gutter i alderen 13 til 20 år. Sånn sett er ikke tallene nødvendigvis direkte sammenlignbare, men de gir en indikasjon. Metoden baserer seg på analyse av sammenvokningsflaten for 11 områder i hånden (se Figur 9).

I Tabell 4 nedenfor har vi sett på avvikene mellom kronologisk og estimert alder i form av gjennomsnittlig avvik fra kronologisk alder, absoluttverdien av gjennomsnittlig avvik, standardavviket for dette, samt RMS feilen.

Metode/Mål	Avvik gj. snitt	Abs avvik gj. snitt	Standardavvik	RMS feil
Roma manuell konsensus	-0.17	0.50	0.41	1.61
NR automatisk, N=2x5	-0.14	0.47	0.42	1.58
NR automatisk, N=2x7	-0.18	0.48	0.41	1.56
Graz automatisk*	NA	0.85	0.58	NA

Tabell 4: Tabellen viser gjennomsnittlige avvik mellom kronologisk og estimert alder for de ulike metodene.*) Metoden fra Graz er ikke testet ut på samme datasett som de andre.

I Figur 14 har vi også plottet avvikene mellom predikert og kronologisk alder for den manuelle prediksjonen (Roma) og den automatiske prediksjonen (NR).



Figur 14: Plott av kronologisk alder mot estimert alder for den manuelle prediksjonen fra Universitetet i Roma (venstre) og den automatiske prediksjonen fra NR (høyre)

5.4.3 Diskusjon

Fra Tabell 4 ser vi at for dette datasettet er den manuelle metoden fra Universitetet i Roma og den automatiske NR-metoden veldig like. Den automatiske metoden er dermed i veldig stor grad i stand til å gjenskape resultatene som oppnås ved manuell analyse. Dette til tross for at NR-metoden bare benytter egenskaper ved radialbeinet i analysen.

Videre ser vi at nivået på avvikene er betydelig mindre enn det som rapporteres med den automatiske metoden utviklet i Graz. Metoden utviklet ved Graz tar bare hensyn til egenskaper ved selve sammenvokningsflaten, mens metoden fra Roma derimot, og dermed også NRs metode, benytter egenskaper fra større områder. Dermed inkorporeres også informasjon om form, beinstruktur og bruskdannelse. Dette kan være noe av årsaken til at disse metodene gjør det bedre. På den annen side er resultatene fra Graz oppnådd på et annet datasett og for en litt videre aldersgruppe (13-20 mot 14-19), slik at noe av årsaken også kan ligge her.

Det skal også bemerkes at, slik vi har forstått, er arbeidet gjort i Roma med å definere og beskrive utviklingsstadier basert på datasettet som vi i denne studien har brukt et utvalg fra. Både den manuelle metoden utviklet i Roma og den automatiske NR-metoden bør derfor valideres på et nytt og helt uavhengig datasett.

Fordi datasettet er såpass lite, kan også noen av metode- og parametervalgene knyttet til den automatiske tilnærmingen ha blitt for godt tilpasset det begrensede utvalget. Enkelte tilpasninger har vi dessuten bevisst latt være å se på, nettopp fordi datasettet er så lite. Dette er derfor ting som bør testes ut for et større og uavhengig datasett.

I denne studien har vi bare sett på analyse av områder rundt radialbeinet, men metoden bør enkelt kunne tilpasses og utvides til å omfatte de andre åtte områdene som er brukt i den manuelle prosedyren utviklet i Roma. I en utvidet studie på et større datasett, er dette noe det også vil være interessant å se på.

6 Oppsummering og konklusjon

Målet med denne studien har vært å se på om og hvordan det kan utvikles en metode for automatisk estimering av skjelettalder fra MR-bilder med utgangspunkt i en manuelt basert metode utviklet ved Universitetet i Roma. Gjennom arbeidet har vi satt oss i inn i hvordan den manuelle metoden virker og laget en beskrivelse av denne. Videre har vi sett på studier som er gjort tidligere på automatisert beregning av skjelettalder fra røntgenbilder, samt helt ferske studier som tar for seg dette for MR-bilder.

Den manuelle metoden som er utviklet ved Universitetet i Roma, baserer seg på å definere et atlas med utviklingstrinn bestående av tekstlige beskrivelser og bilder. Det er definert opp til 12 utviklingstrinn for 9 forskjellige bein i hånden. Videre er det, basert på regresjon, laget en oppslagstabell som definerer sammenhengen mellom en score definert av summen av de 9 utviklingstrinnene og alder. Aldersestimering gjøres ved først å bestemme de 9 utviklingstrinnene fra bilder og beskrivelser, for deretter å summere disse til en score og finne aldersestimatet fra oppslagstabellen. Metoden er foreløpig ikke validert for uavhengige datasett og i liten grad testet for bruk av uavhengige radiologer.

Basert på en vurdering av trinnene i denne manuelle prosessen samt hva som er state-of-the-art innenfor automatisering av slike prosesser, har vi valgt en strategi som utnytter mye av kompetansen på denne typen MR-bilder som er opparbeidet i Roma samtidig som den nyttegjør seg potensialet som finnes i en mer automatisert framgangsmåte. Her er det da ikke nødvendigvis fornuftig å basere en automatisk analyse fullstendig på en manuell prosess. Dette fordi automatiske tilnærminger har andre styrker og svakheter enn manuelle.

For den automatiske metoden vi foreslår har vi, basert på nyere arbeid på røntgenbilder samt det som har vært gjort på automatiserte metoder for MR, valgt å benytte en tilnærming som benytter egenskaper beregnet fra rektangulære regioner i interesseområdet. Lokalisasjonen og avgrensningen av disse områdene har vi i denne studien gjort svært enkelt. Vi har valgt å ikke fokusere på dette ettersom det allerede er mange som har utviklet slike metoder både på røntgen og etter hvert også MR. I denne studien har vi begrenset oss til å se bare på radialbeinet, men metoden kan enkelt utvides til alle områdene som benyttes i den manuelle metoden.

Vi trekker ut de definerte regionene fra flere etterfølgende snitt fra MR-opptakene. Fra hvert snitt trekkes det ut et sett med generiske egenskaper. Egenskapene er bestemt ved hjelp av metoder basert på såkalt Deep Learning, som er i stand til å lære gode datarepresentasjoner direkte fra bildedata. Fordi vi har begrenset med treningsdata har vi her utnyttet at disse metodene er i stand til å lære generiske datarepresentasjoner fra et område med mye treningsdata, for så å benytte den lærte representasjonen på et problem fra et annet område.

Klassifikasjonen inn i utviklingstrinn gjøres per snitt, slik at vi for hvert opptak får flere klassifikasjonsresultater. I denne studien har vi sett på både hvordan denne klassifikasjonen samsvarer med resultatene fra den manuelle analysen, og også på hvordan aldersestimater utledet fra dette samsvarer med kronologisk alder.

Tilnærmingen har vært testet ut på et datasett bestående av MR-opptak fra 63 gutter i alderen 14-19 år. For ca. 75% får vi samsvar med det manuelt bestemte utviklingstrinnet, mens der vi ikke får samsvar er det maksimalt ett trinn i avvik. Fordi utviklingstrinnene er kategoriske (ikke kontinuerlige), fanger ikke forskjellen mellom manuelt og automatisk bestemt utviklingstrinn

nødvendigvis opp hvordan denne forskjellen kan slå ut i form av estimert alder. Vi har derfor også estimert alder fra utviklingstrinn og sammenlignet dette med resultatene som oppnås med den manuelle metoden. Vi ser da at for dette datasettet gir den manuelle metoden fra Roma og den automatiske metoden veldig sammenlignbare resultater.

Konklusjonen så langt er dermed at den automatiske metoden, for det datasettet vi har sett på, er i stand til å gi estimater innenfor samme feilmargin som den manuelt baserte metoden. Dette tyder på at den foreslåtte automatiske metoden kan ha potensiale til å brukes i estimering av skjelettalder. Både den automatiske tilnærmingen og den manuelle MR-metoden trenger imidlertid å valideres på et større og uavhengig datasett.

7 Referanser

- [1] Azizpour, H., Razavian, A. S., Sullivan, J., Maki, A., & Carlsson, S. (2014). *From generic to specific deep representations for visual recognition*. arXiv preprint arXiv:1406.5774.
- [2] Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. *Active shape models-their training and application*. *Computer vision and image understanding*, 61(1), 38-59. 1995.
- [3] Ebner T., Stern D., Donner R., Bischof H., Urschler M.: *Towards Automatic Bone Age Estimation from MRI: Localization of 3D Anatomical Landmarks*. MICCAI (2) 2014: 421-428.
- [4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, *ImageNet: A Large-Scale Hierarchical Image Database*. IEEE Computer Vision and Pattern Recognition (CVPR), 2009.
- [5] Dvorak, J., George, J., Junge, A., & Hodler, J. *Age determination by magnetic resonance imaging of the wrist in adolescent male football players*. *British journal of sports medicine*, 41(1), 45-52. 2007.
- [6] Fischer B.; Brosig A.; Welter P.; Grouls, C. G., Rolf W.; Deserno T. M. *Content-based image retrieval applied to bone age assessment*. *Proceedings of the SPIE, Volume 7624*, id. 762412 (2010).
- [7] Greulich WW, Pyle SI: *Radiographic Atlas of Skeletal Development of the Hand and Wrist*, 2nd edition. Stanford, CA: Stanford University Press, 1959.
- [8] Haak, D., Yu, J., Simon, H., Schramm, H., Seidl, T., & Deserno, T. M. (2013, February). *Bone age assessment using support vector regression with smart class mapping*. In *SPIE Medical Imaging* (pp. 86700A-86700A). International Society for Optics and Photonics.
- [9] Kashif M., Jonas S., Haak D., Deserno T.M. *Bone age assessment meets SIFT*. *Proc SPIE* 2015; 9414: 391-7.
- [10] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks*. In *Advances in neural information processing systems* (pp. 1097-1105).
- [11] van Rijn R.R., Thodberg H.H. *Bone age assessment: automated techniques coming of age?* *Acta Radiol*. 2013 Nov;54(9):1024-9.
- [12] Serinelli S., Panebianco V., Martino M., Battisti S., Rodacki K., Marinelli E., Zaccagna F., Semelka R.C., Tomei E. *Accuracy of MRI skeletal age estimation for subjects 12-19. Potential use for subjects of unknown age*. *Int J Legal Med*. 2015 May;129(3):609-17.
- [13] Stern D., Ebner T., Bischof H., Grassegger S., Ehammer T., Urschler M. *Fully automatic bone age estimation from left hand MR images*. *Med Image Comput Comput Assist Interv*. 2014;17(Pt 2):220-7.
- [14] Stern D., Ebner T., Bischof H., Urschler M.: *Determination of legal majority age from 3D magnetic resonance images of the radius bone*. *ISBI 2014*: 1119-1122.

- [15] Tanner J.M., Whitehouse R.H., Marshall W.A., et al.: *Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method)*. New York: Academic Press, 1975.
- [16] Tanner, J. M., and R. D. Gibbons. "A computerized image analysis system for estimating Tanner-Whitehouse 2 bone age." *Hormone Research in Paediatrics* 42.6 (1994): 282-287.
- [17] Terada Y, Kono S, Tamada D, Uchiumi T, Kose K, Miyagi R, Yamabe E, Yoshioka H. *Skeletal age assessment in children using an open compact MRI system*. *Magnetic Resonance in Medicine*, 2013, 69(6):1697-1702.
- [18] Thangam, P., K. Thanushkodi, and T. V. Mahendiran. "Comparative Study of Skeletal Bone Age Assessment Approaches using Partitioning Technique." *International Journal of Computer Applications* 45.18 (2012): 15-20.
- [19] Hans Henrik Thodberg, A. Rosholm: *Application of the Active Shape Model in a commercial medical device for bone densitometry*. *BMVC* 2001: 1-10
- [20] Thodberg H.H., Rosholm A. *Application of the active shape model in a commercial medical device for bone densitometry*. *Image and Vision Computing*, 2003 Vol. 21 No. 13-14 pp. 1155-1161.
- [21] Thodberg H. H. , Kreiborg S., Juul A., Pedersen K.D. *The BoneXpert method for automated determination of skeletal maturity*. *IEEE Trans Med Imaging*. 2009 Jan;28(1):52-66.
- [22] Thodberg H.H. *Clinical review: An automated method for determination of bone age*. *J Clin Endocrinol Metab*. 2009 Jul;94(7):2239-44.
- [23] Thodberg H.H., Neuhof J., Ranke M.B., Jenni O.G., Martin D.D. *Validation of bone age methods by their ability to predict adult height*. *Horm Res Paediatr*. 2010;74(1):15-22.
- [24] Thodberg H.H., Säwendahl L. *Validation and reference values of automated bone age determination for four ethnicities*. *Acad Radiol*. 2010 Nov;17(11):1425-32.
- [25] Thodberg H.H., Jenni O.G., Ranke M.B., Martin D.D. *Standardization of the Tanner-Whitehouse bone age method in the context of automated image analysis*. *Ann Hum Biol*. 2012 Jan;39(1):68-75.
- [26] Tomei E., Sartori A., Nissman D., Al Ansari N., Battisti S., Rubini A., Stagnitti A., Martino M., Marini M., Barbato E., Semelka R.C. *Value of MRI of the hand and the wrist in evaluation of bone age: preliminary results*. *Journal of magnetic resonance imaging: JMIR* 39:5 2014 May pg 1198-205.
- [27] Tomei E., Semelka R.E., Nissman D. *Text-Atlas of Skeletal Age Determination: MRI of the Hand and Wrist in Children*. Wiley-Blackwell, ISBN: 978-1-118-69227-1, Februar 2014.
- [28] Urschler M., Grassegger S., Stern D. *What automated age estimation of hand and wrist MRI data tells us about skeletal maturation in male adolescents*. *Annals of Human Biology*, 42 (4), p. 358-367. 2015.