

UNIVERSITETET I OSLO
Institutt for informatikk

Automatisk rangering
av bildetekster på
nettsider for
skjermlesere

Masteroppgave
(60 studiepoeng)

Nils Einar Eide

3. mai 2010



Sammendrag

Mange nettsider benytter bilder og grafikk i store mengder. Dette kan skape problemer for enkelte brukergrupper, for eksempel blinde og svaksynte. Det er et mål å utbedre denne situasjonen gjennom nasjonalt lovverk og internasjonalt standardiseringsarbeid som Web Accessibility Initiativ. Skjermleseren er her et viktig hjelpemiddel. Den formidler teksten på en nettside om til enten tale eller til en leseliste i Braillealfabet. Dette inkluderer også bildetekster der disse er tilgjengelige. Utfordringen med skjermlesere er å identifisere hvilken informasjon den skal formidle om bildene på siden når bildetekster ikke er tilstede. I denne oppgaven antas det at bilder som mangler bildetekst vil utgjøre støy på en nettside formidlet med skjermleser og det søkes å finne metoder for å filtrere bort denne støyen. For å oppnå dette vil både bildet og bildetekst analyseres. Et bilde vil således falle inn i én av tre kategorier: Informasjonbærer med bildetekst, informasjonbærer uten bildetekst og støy. I tekstmodalitet analyseres bildeteksten for å undersøke om den eksisterer, om ordene i den er korrekt stavet, om den inneholder bestemte nøkkelord, om den er et tall eller en filsti eller om samme tekst er gjentatt mange ganger på samme nettside. På tilsvarende måte vil en analyse av bildemodalitet rangere bildene etter hvor mye de utgjøre et blikkfang på nettsiden. Egenskaper som størrelse, kompleksitet i form og farger og kontrast mot bakgrunnen vil ha betydning for hvor godt blikkfang bildet utgjør på nettsiden. Inkludert i rangering i bildemodalitet er en algoritme for å lokalisere et bilde på en nettside uavhengig av kjennskap til den spesifikke nettlesers tolkning av HTML og CSS. Delrangeringene er alle i intervallet $[0.0, 1.0]$ slik at bildene tilslutt kan gis en samlet rangering ved å ta gjennomsnittet. Bilder som så faller under en spesifisert grense ignoreres av skjermleseren. De antas i denne oppgaven at denne grensen må settes separat for forskjellige kategorier av nettsider. Resultatene tyder på at det er mulig å rangere bilder etter relevans som informasjonbærer, men det krever større bevissthet blant utviklere av nettsider rundt bruk av ALT-attributten som alternativ bildetekst for å oppnå optimal effekt.

Kapittel 1

Forord

Denne oppgaven er skrevet i forbindelse med mastergraden i informatikk ved Institutt for informatikk, Universitetet i Oslo.

Da oppgaven har vært skrevet på deltid har det til tider vært krevende å balansere arbeid med fulltidsjobb. Selv om det så mørkt ut til tider er den nå endelig klar.

Det har faglig vært en krevende tid, da det finnes lite av tidligere arbeid i dette feltet til å bygge oppgaven på. Det har derfor vært mye prøving og feiling, men også dette har i seg selv vært lærerikt.

Jeg vil takke mine veiledere Frode Sandnes ved HIO og Dag Langmyr ved UIO for all støtte og veiledning jeg har fått i denne perioden.

Jeg vil også takke mine foreldre, familie og venner for den støtten jeg har fått under arbeidet med oppgaven.

Oslo, 3. mai 2010

Nils Einar Eide

Innhold

1	Forord	ii
I	Problemstilling	1
2	Innledning	2
2.1	Definisjon av problemområdet	3
2.2	Avgrensning av problemområdet	4
3	Problemstilling	5
3.1	Internett	5
3.1.1	Definisjon av Internett	5
3.1.2	Internetts fødsel	6
3.1.3	Vekst av Internettet	6
3.1.4	Dagens Internett	7
3.2	WWW - Verdensveven	7
3.2.1	WWW sine komponenter	7
3.2.2	WWW sitt bruksområde	8
3.3	WWW sitt teknologiske grunnlag	8
3.3.1	HTTP - HyperText Transfer Protocoll	9
3.3.2	URL - Uniform Resource Locator	9
3.3.3	HTML - HyperText Markup Language	10
3.4	Betydningen av WWW	12
3.4.1	Nettapplikasjoner	12
3.4.2	Nettportaler	12
3.4.3	Søkemotorer	13
3.4.4	Nettaviser	13
3.4.5	E-Handel	14
3.4.6	Nettbanker	14
3.4.7	Leksika: Wiki	14
3.4.8	Sosiale nettverk	15
3.5	Standardisering av Internett	15
3.5.1	Internet Engineering Task Force	16
3.5.2	World Wide Web Consortium	16

3.6	Brukergrupper med funksjonshemmelser	16
3.6.1	GAP-modellen	16
3.6.2	Universell utforming	17
3.6.3	Synshemming	18
3.7	Tilgjengelighet gjennom universell utforming	18
3.7.1	Nettinnhold og brukeragenter	19
3.7.2	Web Accessibility Initiative	19
3.7.3	Web Content Accessibility Guidelines	19
3.7.3.1	WCAG 1.0	19
3.7.3.2	WCAG 2.0	20
3.7.4	User Agent Accessibility Guidelines	21
3.8	Universell utforming for blinde og synshemmede	21
3.8.1	Braillealfabetet	21
3.8.1.1	Brailletastatur og leselist	22
3.8.2	Skjermlesere	22
3.8.2.1	Talesyntese	23
3.8.2.2	Skjermlesere og grafikk	23
3.9	Informative bilder	24
3.10	Bilder i nettsider: IMG-markøren	24
3.11	Målsetning	26
II	Tekstmodalitet	27
4	Tekst og språk	28
4.1	Tekst	28
4.2	Språk	29
4.2.1	Fonemer, bokstaver og alfabet	29
4.2.2	Morfemer, ord og ordklasser	29
4.2.3	Grammatikk	30
4.3	Tolkning av tekster	30
4.4	Datamaskiner og naturlige språk	31
4.4.1	Datalingvistikk	31
5	Tekstanalyse	32
5.1	Tekstklasser	33
5.1.1	Tom eller manglende ALT-attributt	33
5.1.2	Ikke i ordlisten	34
5.1.3	Bestemte nøkkelord	34
5.1.4	Filstier	35
5.1.5	Repeterte ALT-attributter på samme side	36
5.1.6	<i>Bare</i> et tall	36
5.2	Tester av tekstanalyser	37

III	Bildemodalitet	38
6	Farger og bilder	39
6.1	Bilder på nettsider	39
6.2	Digital representasjon av bilder	40
6.2.1	Rastergrafikk	40
6.2.2	Vektorgrafikk	41
6.2.3	Filformater	41
6.3	Fargemodeller	41
6.3.1	Primær- og sekundærfarger	42
6.3.2	Komplementærfarger, fargesirkelen og kontrast	43
6.3.3	Additiv og subtraktiv fargeblanding	43
6.3.4	RGB fargemodellen	44
6.3.4.1	Representasjon i en datamaskin	44
6.3.4.2	Alfakanal	44
6.3.5	HSV fargemodellen	45
6.4	Histogram og bilder	45
6.4.1	Formell definisjon	47
6.4.2	Normalisering av histogrammer	47
6.4.3	Like histogrammer, like datasett	47
6.4.3.1	Tilnærmet like datasett	48
7	Posisjon for delbilde	49
7.1	Rå kraft	49
7.2	Pikselplukk-algoritmen	51
7.2.1	Antakelser	52
7.2.2	Overordnet skisse	54
7.2.3	Et eksempel	54
7.2.3.1	Beregne innledende histogrammer	55
7.2.3.2	Velge nettsidens bakgrunnsfarge	56
7.2.3.3	Tilfeldig utplukk av piksler	57
7.2.3.4	Finn kandidatbilder	58
7.2.3.5	Sammenlikne histogrammer	58
7.2.4	Spesialtilfeller	59
7.2.4.1	Bildet deles ikke i to deler	60
7.2.4.2	Bildet deles i to deler	60
7.3	Testing av algoritmen	61
7.3.1	Utplukksratio	62
7.3.2	Antallet bilder på nettsiden	62
7.3.3	Skaleringer	63
7.3.3.1	Oppskalering - endimensjonal	63
7.3.3.2	Nedskalering - endimensjonal	63
7.3.3.3	Oppskalering - todimensjonal	63

7.3.3.4	Nedskalering - todimensjonal	63
7.3.4	Areal som dekker delbilde	64
8	Bildeanalyse	67
8.1	Bildets innhold og fargesammensetning	68
8.1.1	Komprimering	69
8.1.1.1	Singulærverdidekomposisjon (SVD)	69
8.1.2	Mønstergjenkjennelse	70
8.1.2.1	Bilder med kjent mening	70
8.1.3	Stor indre kontrast	70
8.2	Bildets størrelse og form	71
8.2.1	Etter størrelse	71
8.2.2	Størrelsen går mot 1x1 piksler	71
8.2.3	Misforhold mellom dimensjoner	72
8.3	Bildets plassering på nettsiden	72
8.3.1	Naturlig lesesti	72
8.3.2	Det gyldne snitt	72
8.4	Bildet og dets omgivelser på nettsiden	73
8.4.1	Kontrast mot bakgrunnen	74
8.4.1.1	Et bildes lokale bakgrunn	74
8.5	Test av utvalgte bildeanalyser	75
IV	Rangering av bilder	76
9	Samlet rangering	77
9.1	Grunnlaget for rangeringen	77
9.2	Rangeringsalgoritme	77
9.3	Velge bildetekster etter rangering	78
9.4	Test av rangeringsalgoritme	78
10	Vurdering av rangeringen	79
10.1	Kendall-Tau rangkorrelasjonskoeffisient	79
10.1.1	Beregne rangkorrelasjonskoeffisient	80
10.2	Tester på nettsteder	80
10.2.1	Test på <i>bra</i> og <i>dårlige</i> nettsteder	81
10.2.2	Test på nettaviser	82
V	Resultater og diskusjon	83
11	Resultater og diskusjon	84
11.1	Tekstanalyser	84
11.1.1	Resultater	84

11.1.2	Diskusjon	85
11.1.2.1	Manglende ALT-attributt	85
11.1.2.2	Tom ALT-attributt	87
11.1.2.3	ALT med rettskrevet tekst	87
11.1.2.4	ALT med ord i en liste nøkkelord	88
11.1.2.5	ALT som inneholder en filsti	89
11.1.2.6	ALT som er repetert på samme side	89
11.1.2.7	ALT som kun inneholder et tall	90
11.2	Pikselplukk algoritmen	90
11.2.1	Utplukksratio for piksler	90
11.2.1.1	Resultat	91
11.2.1.2	Diskusjon	91
11.2.2	Antall bilder på nettsiden	92
11.2.2.1	Resultat	92
11.2.2.2	Diskusjon	93
11.2.3	Oppskalering - 1 dimensjon	93
11.2.3.1	Resultat	93
11.2.3.2	Diskusjon	93
11.2.4	Nedskalering - 1 dimensjon	94
11.2.4.1	Resultat	94
11.2.4.2	Diskusjon	94
11.2.5	Oppskalering - 2 dimensjoner	95
11.2.5.1	Resultat	95
11.2.5.2	Diskusjon	95
11.2.6	Nedskalering - 2 dimensjoner	95
11.2.6.1	Resultat	97
11.2.6.2	Diskusjon	97
11.2.7	Delvis tildekkede bilder	97
11.2.7.1	Resultat	97
11.2.7.2	Diskusjon	97
11.2.8	Diskusjon av pikselplukk	98
11.3	Bildeanalyser	98
11.3.1	Resultat	98
11.3.1.1	Etter størrelse	99
11.3.1.2	Komprimeringsgrad	99
11.3.1.3	Kontrast mot bakgrunn	100
11.3.2	Diskusjon	101
11.4	Tekst -og bildeanalyser i sammenheng	101
11.4.1	Gode og dårlige nettsider	102
11.4.1.1	Resultater	102
11.4.1.2	Diskusjon	102
11.4.2	Ti nettaviser	103
11.4.2.1	Resultater	103

11.4.2.2	Diskusjon	104
VI	Konklusjon	106
12	Konklusjon	107
12.1	Rangere bildetekster innefor en nettside	107
12.2	Finne et bildes posisjon innenfor en nettside	108
12.3	Rangere bilder innenfor en nettside	108
12.4	Samlet rangering av bilde og bildetekst	108
12.5	Konklusjon	108
12.6	Implikasjoner av konklusjonen	109
	Bibliografi	109

Figurer

2.1	Eksempel på HTML kode	4
3.1	Strukturen på en Uniform Resource Locator	10
3.2	Eksempel på HTML kode	11
3.3	GAP modellen (Kilde: St. Meld. 40 (2002-2003))	17
3.4	Braille tegn for bokstaven a (Minuskel) (Kilde: wikipedia.org)	22
3.5	Brilletegn for bokstaven A (Majuskel) (Kilde: wikipedia.org)	22
3.6	Braille tegn for tallet 1 (Kilde: wikipedia.org)	23
3.7	Brilletegn for start anførselstegn eller spørsmålstegn (Kilde: wikipedia.org)	23
3.8	HTML sin IMG-markør for visning av bilder, med ALT- og SRC- attributtene	25
5.1	IMG-markøren	33
6.1	Primær -og sekundærfarger i RGB-modellen (Kilde: wikipedia.org)	42
6.2	Fargehjulet HSV-modellen (Kilde: wikipedia.org)	43
6.3	Farger og figurer	46
6.4	Pikselhistogram av farger fra figur 6.3	46
7.1	Meget enkel nettside, bare en rød boks	50
7.2	HTML-kode for figur 7.1, rendret av Internet Explorer 7.0	51
7.3	HTML-kode for figur 7.1, rendret av Firefox 3.5	52
7.4	Naiv metode for tekstsøk	53
7.5	Forenklet skisse av nettside med bilde	55
7.6	Bildet det skal søkes etter	56
7.7	Histogram av farger på nettsiden i Figur 7.5	56
7.8	Histogram farger i bildet i Figur 7.6	57
7.9	Utplukk av tilfeldige punkter, jevnt fordelt	58
7.10	Søk ut fra punktene, finn kanten	59
7.11	Spesialtilfelle 1	60
7.12	Spesialtilfelle 2	61
7.13	Spesialtilfelle 3	61
7.14	DIV-markør uten innhold gitt form av et kvadrat og plassert over et bilde ved hjelp av CSS	65
7.15	Bilde med overlappende firkant	66

8.1	Det gyldne snitt (Kilde: wikipedia.org)	73
10.1	Konseptskisse av en nettside	81
11.1	Endring i antall korrekt plasserte bilder ved økning av antallet utplukkede piksler	91
11.2	Endring i antall korrekt plasserte bilder ved økning av antallet bilder på siden	92
11.3	Endring i antall korrekt plasserte bilder ved oppskalering i bredde . . .	94
11.4	Endring i antall korrekt plasserte bilder ved nedskalering i bredde . . .	95
11.5	Endring i antall korrekt plasserte bilder ved oppskalering i begge retninger	96
11.6	Endring i antall korrekt plasserte bilder ved nedskalering i begge retninger	96
11.7	Endring i antall korrekt plasserte bilder ved økning av arealet av bildet som gjemmes/beskjæres	98
11.8	Nettside med 5 bilder i stingende størrelse	99
11.9	Nettside med 5 bilder i økende kompleksitet	100
11.10	Fem farger i økende fargekontrast fra rødt.	100
11.11	Nettside med 5 bilder med økende fargekontrast mot bakgrunnen . . .	101

Tabeller

11.1 IMG - markører og ALT - attributter	86
11.2 Kendall Tau for gode nettsier	102
11.3 Kendall Tau for tre dårlige nettsier	102
11.4 Ti nettaviser med mer enn 100 IMG-markører på hovedsiden	103
11.5 Kendall Tau for nettaviser i figur 11.4	103
11.6 ALT-tekster på www.nettavisen.no	104

Del I

Problemstilling

Kapittel 2

Innledning

Informasjon har gjennom tidene vært en kilde til makt og posisjon i samfunnet. De som hadde, og de som har, kunnskaper og ferdigheter vil kunne bruke disse til å oppnå fordeler.

Nåtiden kalles ofte *informasjonsalderen*. Dette skriver seg fra tradisjonen med å navngi førhistoriske tidsperioder etter periodens mest sentrale teknologiske nyvinning. Steinialder, bronsealder og jernalder er således eksempler på tidsaldre oppkalt etter det *nyeste* materialet for tilvirking av verktøy.

Det siste århundrets teknologiske fremgang har i løpet av kun få tiår gitt tre nye teknologiske tidsaldre. Atomalderen med utgangspunkt i utviklingen av atomvåpen og -reaktorer mot slutten av andre verdenskrig. Romalderen etter romkappløpet mellom USA og USSR på 1950 og 1960 med Sputnik og månelandinger. Til slutt informasjonsalderen fra personlige computere, PCer, ble allemansgjeld i løpet av 1980-tallet og fremveksten av Internett som ga vanlige mennesker tilgang til uante mengder informasjon.

Tilgangen til, og mulighet til å bruke, informasjon har alltid vært viktig. I middelalderen var tilgangen til kunnskap om religionen begrenset ved at kirken brukte latin som arbeidsspråk i tekster og i gudstjenester. Dette var et tiltak for å få et standardspråk for kommunikasjon innenfor kirken, men en konsekvens av dette ble at alle som ikke kunne latin ble utelukket fra det som i de dager var meget sentrale ting i det daglige liv. Fysiske og praktiske forhold bidro også til å holde kunnskap innen privilegerte grupper. Da alle tekster lenge måtte kopieres for hånd var bøker noe som kun de rikeste hadde mulighet til å eie. Dette bidro til å opprettholde mindretallets makt over flertallet.

Under opplysningstiden og reformasjonen endret dette seg og fordelene ved en mer informert befolkning ble mer forstått og akseptert. Trykkekunsten skapte muligheten til å lage store opplag av en tekst på en relativt rask og billig måte, sett i forhold til tidligere tiders kopiering for hånd. Men selv om trykkekunsten førte til at bøker ble mer vanlige, var de lange tider dyre gjenstander som var forbeholdt de velstående. Etersom Bøker gjennom århundrene ble billigere og mer tilgjengelige, var de fremdeles begrenset i tilgjengelighet av sin vekt og sitt volum. Selv med offentlige bibliotek er

2.1. DEFINISJON AV PROBLEMOMRÅDET

det begrenset hvor mye informasjonen som lagres der som er tilgjengelig, på grunn av det krevende arbeidet med å søke frem og systematisere relevant informasjon.

Gjennom Internett er løsningen på dette problemet endelig i sikte, ved hjelp av elektronisk lagrede medier, som er tilgjengelige og søkbare for de fleste mennesker. Men Internett og World Wide Web, og de nettsider de gjør tilgjengelige, er likevel ikke tilgjengelige for alle. Selv om det å bringe informasjon, via Internett, ut til alle kanskje er en utopisk drøm er det allikevel det som man, i et moderne demokrati, som hevder at alle er likeverdige, må tilstrebe. Det er derfor tatt mange initiativer for å forbedre dette, både nasjonalt [1] [2] [3] [4] og internasjonalt [5] [6].

Mange mennesker har en funksjonsnedsettelse som gjør det vanskelig for dem å tilegne seg informasjonen. Internett er et medium med vekt på tekst, bilder og andre grafiske elementer som forutsetter at man kan se for å gjøre seg nytte av disse. Nettsider er derfor ofte ikke spesielt tilgjengelige for blinde og svaksynte. Dette gjør at det fremdeles er et stykke igjen for å bringe informasjon ut til alle. Det er med denne oppgaven ønsket å bidra til å tette dette gapet.

2.1 Definisjon av problemområdet

Tidlige nettsider var først og fremst tekstbaserte formidlere av akademisk informasjon. Den fremste grunnen til dette var den gjennomsnittlige nettverksforbindelse sin lave båndbredde. Mange tidlige nettlesere, for eksempel Lynx ¹, var derfor også kun tekstbaserte. Etter som båndbredden økte, økte også bruken av grafikk, først som grafer og andre illustrasjoner til vitenskapelige tekster, men etter hvert som bruken av Internett bevegde seg utenfor det akademiske miljø økte også antallet grafiske elementer på en nettside betraktelig.

Bilder og grafikk på en nettside kan grovt deles inn i to kategorier, de som gir informasjon og de som gir struktur og layout. Denne siste gruppen kan inngi tillit ved å få en nettside til å fremstå som profesjonell og seriøs, men dette kan også være villedende, og kan være av liten verdi så lenge mottakeren ikke kan se siden.

Det finnes hjelpemidler for blinde og svaksynte, kalt *skjermlesere*, som kan lese opp innholdet på en nettside ved hjelp av leselist eller talesyntese. Disse har imidlertid begrensninger, og en av disse er vedrørende bilder som skjermleserne bare kan konstatere at er der, eventuelt lese opp filstien til bildet. I standarden for å plassere bilder på en nettside er det derfor inkludert måter man kan assosiere bildet med en alternativ bildetekst, som kan beskrive innholdet i bildet. Dessverre viser erfaring at dette ikke alltid er godt gjennomført på alle nettsider. Mulige grunner til dette kan være at utvikler og eier av nettstedet ikke er bevisst behovet eller at de ikke ønsker å ta ekstrautgiften med å gjennomføre dette. De siste tiders arbeid med bevisstgjøring av dette, både i industristandarder og i nasjonalt og internasjonalt lovverk, er i ferd med å bedre situasjonen noe, men det er fremdeles lang vei å gå.

¹<http://lynx.isc.org>

Hensikten med denne oppgaven er å undersøke muligheter for å filtrere bort irrelevante bilder fra opplesningen. Med irrelevante menes i denne sammenheng bilder som man bare kan konstatere at er tilstede. Det vil si at de er uten alternativ tekst som skjermleseren kan brukes ut å bare slå fast at her er det et bilde av noe ukjent.

Videre skal det også gjøres forsøk med å filtrere bort bilder som bare har estetikk som formål, som for eksempel rammer på tabeller. Disse er ikke nødvendigvis informasjonsbærende men bare til pynt, og derfor (kanskje?) ikke spesielt interessante for blinde og svaksynte, selv om de skulle ha en alternativ tekst. Videre kan det tenkes at bildene ikke bare er uinteressante, men kanskje også direkte forstyrrende nå det leses opp noe som ikke tilfører noe i forhold til teksten som leses opp.

2.2 Avgrensning av problemområdet

Det er flere måter i HTML [7] for å plassere bilder på en nettside. De kan være brukt som bakgrunn for siden, eller deler av den, som for eksempel tabeller, eller de kan være plassert i forgrunnen med IMG-markører. Et eksempel på dette er gitt i figur 2.1, som gir en nettside som består av kun et bilde. I tillegg er det gitt en bildetekst ved hjelp av en ALT-attributt (ALternativ).

```
<html>
  <body>
    <img src=' 'http://www.uio.no/logo.gif' '
        alt=' 'Universitetet i Oslo sin logo' ' >
  </body>
</html>
```

Figur 2.1: Eksempel på HTML kode

For å avgrense problemområdet er det blitt bestemt å fokusere på bilder som er plassert på en nettside med IMG-markøren, da dette antas å være den primære måten å inkludere informasjonsbærende bilder i en nettside. Det er andre måter å inkludere bilder i en nettside, primært som bakgrunnsbilder. De er derfor, i kontekst av denne oppgaven, ansett som primært brukt til layout og estetiske formål, selv om det også kan brukes som informasjonsbærere. IMG-markøren er også viktig i denne sammenheng fordi den har en metode å direkte tilordne bildet en alternativ tekst via dens ALT-attributt.

Kapittel 3

Problemstilling

Denne oppgaven skal belyse de utfordringer blinde og svaksynte møter i form av ned-satt tilgjengelighet av nettsider, som følge av disses utstrakte bruk av bilder og annen grafikk til å formidle informasjon.

En gjennomgang av Internettets utvikling og eksempler på hva Internett benyttes til vil illustrere dets betydning på flere områder i samfunnet. Deretter beskrives ulike gruppers utfordringer med bruk av Internett og hvilket arbeid som er gjort for å gjøre Internett mer tilgjengelig for disse brukerne. Oppgaven vil også kort beskrive de individuelle teknologiene som danner grunnlaget for Internett og WWW, da disse har betydning for hvilke muligheter som finnes vedrørende forbedringer av nettsiders tilgjengelighet.

3.1 Internett

I dagligtale benyttes ofte begrepet Internett når det omtalte objektet egentlig er *World Wide Web* (WWW). Dette er imidlertid to separate konsepter, selv om de er tett knyttet sammen. WWW [8] [9] [10], som omtales i detalj senere i dette kapittelet, er bare en av mange teknologier som benytter Internett som kommunikasjonsmedium. Andre sentrale teknologier er epost (SMTP [11]) og overføring av filer mellom datamaskiner (FTP [12]). En mulig forklaring på denne misforståelsen kan være at det er nettopp disse teknologiene som de fleste bruker i sin hverdag, mens Internettets indre liv stort sett er skjult.

3.1.1 Definisjon av Internett

Et internettverk, eller et internett, er en samling av datanettverk hvor datamaskiner som er koblet til ett eller flere av de tilknyttede datanettverkene kan kommunisere med de andre datamaskinene som også er koblet til et av de deltagende nettverkene [13]. Når Internett skrives som egennavn med stor forbokstav, refereres det vanligvis til det verdensdekkende internettverket som vi assosierer begrepet med i dagligtale.

En datamaskin som er koblet til to eller flere nettverk, og som dermed kan opptre som et mellomledd mellom datamaskiner som er koblet til forskjellige datanettverk kalles en ruter [13] [14]. Ved å gå veien gjennom en ruter kan en hvilken som helst datamaskin på et av de sammenknyttede nettverkene kommunisere med en hvilken som helst annen tilsvarende tilknyttet datamaskin.

Større organisasjoner kan også opprette internettverk mellom sine datanettverk for å kunne forbedre kommunikasjonen internt i organisasjonen. Slike interne internettverk omtales ofte som et *intranett*, selv om det ikke nødvendigvis er basert på annen teknologi.

3.1.2 Internetts fødsel

Utviklingen av *ARPANet* [15] på 1960-tallet omtales ofte som Internett sin fødsel. ARPANet ble utviklet av *Advanced Research Projects Agency* (ARPA), en forskningsinstitusjon tilknyttet USA sitt forsvar.

Dette er ikke helt korrekt da det var flere andre forsøk med nettverkskommunikasjon, både i USA og andre steder, som ga gode resultater, men ARPANet var helt klart det største [16]. Dette ble det bygget videre på i begynnelsen av 1970-tallet da man ønsket å binde sammen flere eksisterende datanettverk til et stort datanettverk.

Forsvaret i USA ønsket et kommunikasjonssystem som blant annet kunne fungere selv om deler av det ble slått ut under en atomkrig. Det var derfor et krav at data skulle kunne finne alternative ruter gjennom nettverket, dersom en rute ble ødelagt. På initiativ fra DARPA, nå omdøpt til *Defence Advanced Research Projects Agency*, ledet dette til utviklingen av TCP/IP, som er protokollene som styrer kommunikasjonstrafikken på Internett.

Ordet Internett ble først nevnt av Vint Cerf i RFC 675 - «Specification of Internet Transmission Control Protocol» (TCP) [17], publisert i 1974. Den andre viktige protokollen i dette arbeidet er RFC 791 - «Internet Protocol» (IP) [18]. Det er disse to protokollene som i samarbeid er ansvarlige for at data som applikasjonsprogrammer som WWW, Epost og FTP ønsker å formidle mellom to eller flere datamaskiner kommer frem til riktig datamaskin på nettet.

Det er TCP protokollens oppgave å sørge for at hele meldingen er kommet frem til adressaten og eventuelt ved behov sørge for at manglende deler sendes på nytt. Dette er ikke noen ulempe i mer fredelige situasjoner heller, da atomangrep ikke er den minst ekstreme situasjon hvor datasystemer bryter sammen.

Det er IP protokollens oppgave at meldinger finner frem til rett(e) maskin(er) hvor mottakerprogrammet kjøres. TCP/IP ble videre utviklet over det neste ti-året og ble fastsatt som eneste godkjente protokoll på ARPANet fra 1. januar 1983.

3.1.3 Vekst av Internettet

Selv om Internett først ble tatt i bruk til kommunikasjon internt i akademiske og militære miljøer ble det i løpet av 1980-tallet mer og mer åpnet for private og kommer-

3.2. WWW - VERDENSVEVEN

sielle interesser. I første omgang bestod tilbudet primært av epost og nyhetsgrupper, men etterhvert kom World Wide Web, som ble lansert offentlig av Tim Berners-Lee 6. august 1991. Den første nettsjeneren ble deretter tilgjengelig seks dager senere. Fra da vokste omfanget fra å være et verktøy som ble brukt av store organisasjoner til å bli et massemedia som mange har tilgang til i sitt eget hjem [16].

3.1.4 Dagens Internett

Dagen internett har vokst fra å være en arena for informasjonsutveksling i akademiske og statlige, spesielt militære, miljøer, til å være et sentralt element i stadig flere menneskers hverdag. Innkjøp, banktjenester, nettaviser, og i den senere tid ikke minst kommunikasjon med venner og kolleger gjennom nettsamfunn, er blant de applikasjoner som Internett har gjort mulig. Men Internett er som nevnt bare grunnlaget; et system for kommunikasjon mellom datamaskiner. Det som måtte på plass for å kunne gjennomføre de siste 15-20 årenes digitale revolusjon er verdensveven WWW.

3.2 WWW - Verdensveven

WWW, *World Wide Web* eller verdensveven, er et sett med standarder for å gjøre det mulig å kryssreferere dokumenter på tvers av en eller flere datamaskiner over et nettverk eller et intranett. Det er ofte dette som menes når det i dagligtale refereres til begrepet Internett.

WWW ble utviklet på 1980-tallet av Tim Berners-Lee. Berners-Lee jobbet ved *Conseil Européen pour la Recherche Nucléaire*¹ (CERN), den europeiske forskningsorganisasjonen for kjerne- og partikkelfysikk. Hensikten var å opprette en felles kilde til kunnskap slik at mennesker over hele verden, som jobbet med samme eller liknende problemer, skulle kunne dele informasjon på en enkel måte [19].

Tim Berners-Lee sin visjon med å utvikle WWW, bygger på kravet om at vitenskapelige tekster skal bygge på tidligere resultater. De skal ha kildereferanser som henviser til andres arbeid som tekstene, og forskningen som ligger til grunn for dem, bygger på. For å forenkle dette arbeidet så han for seg at alle dokumenter kunne være elektronisk lagret, og inneholde lenker til de refererte dokumentene, som så kunne være raskt tilgjengelige. Sammen med tre teknologier utgjør denne visjonen WWW sine komponenter.

3.2.1 WWW sine komponenter

WWW er bygget på følgende komponenter [19]:

- Idéen om en ubegrenset verden av informasjon hvor alle dokumenter kan refereres fra et annet dokument.

¹www.cern.ch

- En nettverksprotokoll som muliggjør at en nettleser kan spørre en nettsjener om en nettside og få den overført via nettverket (HTTP).
- Et adressesystem som muliggjør disse referansene (URL).
- Et språk som definerer hvordan informasjon skal presenteres (HTML).

I tillegg til visjonen for hvordan de tre kunne brukes sammen til dette formålet ble både HTTP, URL og HTML først utviklet ved CERN av Tim Berners-Lee, men de er senere videreutviklet av mange. Av de tre er spesielt HTML viktig i følge med design av tilgjengelige nettsteder og vil bli mye omtalt senere i oppgaven. HTTP og URL vil derimot bli noe mer overfladisk beskrevet.

3.2.2 WWW sitt bruksområde

Bruksområder for denne tilnærmelsen til struktureringen av data er imidlertid ikke begrenset til det vitenskapelige området. Andre aktører oppdaget hurtig at WWW kunne brukes til andre former for informasjonsutveksling. I løpet av et par år ble WWW derfor i stadig stigende grad benyttet på et stort antall forskjellige områder, fra akademia til e-handel og underholdning. Til tross for at den ennå ikke er fylt 20 år er WWW i dag en essensiell del av tilværelsen for mange.

Til tross for WWW sin unge alder har de fleste aktører fått forståelse for dens muligheter og begrensninger. Av spesiell interesse for denne oppgaven er nettsteder som det er viktig og nyttig å ha tilgang til for å fungere som en selvstendig borger i samfunnet. I denne kategorien faller nettbanker for betaling av regninger, sider for levering av selvangivelse og andre meldinger til og fra myndighetene, nettaviser og andre kilder for generell informasjon og nettbutikker.

Samfunnet har endret seg hurtig med tanke på hvordan vi utfører slike oppgaver. Imidlertid har enkelte grupper av forskjellige årsaker blitt stående utenfor, eller i hvertfall hatt begrenset mulighet til å nyttiggjøre seg nyvinningene. Blant disse skal det her fokuseres på personer som er blinde eller synshemmede, og på standarder og hjelpemidler som eksisterer for å gjøre nettsider tilgjengelige for disse brukergruppene. WWW sitt teknologiske grunnlag gir muligheter for å oppnå dette.

3.3 WWW sitt teknologiske grunnlag

WWW bygger på tre tekniske konsepter som virker sammen for å oppnå visjonen om sømløs kryssreferering av elektroniske dokumenter. Disse er HTTP, URL og HTML. Programmene som knytter disse sammen kalles *nettsjener* og *nettleser*. En nettsjener er en tjeneste som nettleseren kan komme i kontakt med via nettverk/internettverk. URL er adressen til nettsjeneren, mens HTTP beskriver hvordan nettsiden skal lastes ned og detaljer om hvordan den skal leses. HTML er kodespråket som bestemmer hvordan nettleseren skal presentere informasjonen den henter fra nettsjeneren.

3.3.1 HTTP - HyperText Transfer Protocol

HTTP [9] er en protokoll som støtter distribuerte, samarbeidende hypermedia informasjonssystemer. At det er distribuert og samarbeidende vil si at man fra et dokument på en nettsjener kan lenke til et dokument på enhver annen nettsjener som også ligger på samme nettverk/internettverk [10].

Foruten selve innholdet i filen inneholder meldingen som overføres et hode med opplysninger om filen. Dette kan være hva slags file det er (for eksempel bilde, html eller MS office dokument), hvor stor den er og, som kan ha betydning for tilgjengelighet, hvilket språk innholdet er skrevet i. Det er imidlertid viktig å være oppmerksom på at det ikke er noen innebygget kontroll i HTTP som sikrer at opplysningene oppgitt i hodet stemmer med de faktiske forhold, blant annet fordi en fil som er markert som et bilde kan være andre ting, for eksempel en vanlig HTML fil eller et virus.

Det er denne protokollen som danner grunnlaget for WWW ved at den beskriver hvordan en nettleser skal oppføre seg for å få tilsendt en nettside fra en nettsjener og hvordan nettsjeneren skal oppføre seg både på riktige og feilaktige forespørsler fra en nettleser. Teknologien som benyttes av HTTP for å finne den riktige nettsjeneren, og den riktige nettsiden på denne, er Uniform Resource Locator URL.

3.3.2 URL - Uniform Resource Locator

Hvert eneste dokument som skal være tilgjengelig for å refereres til under WWW må ha sin egen, unike adresse. Hvor mange dokumenter dette er, og vil bli, er ikke mulig å forutse. For å kunne referere til et større, på forhånd ukjent, antall dokumenter må det være en standard for å kunne gi hvert enkelt dokument en unik adresse.

I gjeldende terminologi omtales denne adressen som URL eller en filsti. En URL er en generell måte å adressere ressurser i et nettverk, det være seg ikke bare nettsider, bilder eller andre filer man ønsker å gjøre tilgjengelig for andre brukere gjennom nettet, men også tjenester som SMTP og FTP.

En URL [8] er kjent for mange fra lenken i adressefeltet til nettleseren, for eksempel <http://www.uio.no>. URL er standarden for hvordan lenker til nettsider, og andre ressurser på et nettverk, er bygget opp. I figur 3.1 illustreres de forskjellige delene av en URL som nærmere beskrevet er som følger:

- <schema>
Hva slags type applikasjon vi forventer skal svare på enden av lenken. I dette tilfelle en nettsjener, altså http.
- <host>
Adressen til tjeneren som ressursen er lokalisert på, for eksempel www.uio.no
- <port>
Hvilken nettverksport tjeneren kjører på. I http er port 80 reservert som standard. Brukes bare dersom porten er en annen en standardport for den aktuelle tjenesten.

- `<path>`
En filsti som sier hvor på tjeneren ressursen ligger.
- `<searchpart>`
en & separert liste av nøkkel/verdi par, `nøkkel1=verdi1&nøkkel2=verdi2`. Dette forteller ressursen mer om hva den skal utføre og brukes bare ved behov.
- `<anchor>`
Kan spesifisere at man skal gå til et bestemt punkt, et anker, i den resulterende HTML-siden, ikke bare starte fra toppen. Kan bare brukes hvis slike ankere er spesifisert i HTML-koden med ANCHOR-markøren [7].

Bemerk at det er brukt begrepet *ressurs*, og ikke HTML-fil, i beskrivelsen av URL. Dette skyldes at HTML ikke bare finnes som statiske tekstfiler, men kan også genereres fra gang til gang av forskjellige typer programmer og skript slik at HTML som returneres fra samme ressurs kan være forskjellige basert på parameteren i `<searchpath>` og innholdet i en database. URL også kan benyttes til å referere til, og laste ned, andre filer, som for eksempel bildefiler.

URL brukes ikke bare av HTTP, men også av protokoller som SMTP og FTP til å beskrive adresser til tjenester for å sende og motta epost eller til filtjeneren for å laste opp og ned filer. I denne oppgaven er det imidlertid HTML-filer, og tilhørende ressurser som bilder, som er fokusområdet. Den siste av teknologiene i WWW er derfor innholdet i de nedlastede HTML filer, altså en nettside formatert ved hjelp av HTML-markører.

```
<schema>://<host>:<port>/<path>?<searchpart>#<anchor>
```

Figur 3.1: Strukturen på en Uniform Resource Locator

3.3.3 HTML - HyperText Markup Language

Av de tre teknologiene er det HTML som er mest interessant i konteksten av denne oppgaven og vil derfor bli noe grundigere omtalt. Mens HTTP og URL er viktige i følge med at det er de som sørger for at riktig nettside finnes og lastes ned, er det HTML-koden som inneholder de faktiske opplysningene brukeren er ute etter og som er grunnlaget for hvordan disse presenteres i en nettleser. HTML inneholder også muligheter for å assosiere ytterligere opplysninger til HTML elementer som bilder og lenker, som blant annet kan brukes til å gi alternative opplysninger, hvis for eksempel bildet som skal vises av ulike årsaker er utilgjengelig.

Dokumentene som overføres via HTTP [10], og som utgjør innholdet i WWW, måtte ha en felles standard for hvordan de skulle se ut på skjermen i en nettleser. Ikke

3.3. WWW SITT TEKNOLOGISKE GRUNNLAG

bare måtte det være en mulighet for å legge inn hvor referanselenkene til andre dokumenter skulle være, men det måtte også være mulig å formattere layout for ulike dokumenter med tabeller, lister, overskrifter og bilder for å kunne presentere det aktuelle materialet på en så god måte som mulig.

Løsningen på dette var HTML som er et formateringspråk med en tre-struktur med markører inne i markører som bestemmer hvordan teksten, som er omgitt av markørene, skal presenteres for brukerne. Den første offisielle HTML standarden, som faktisk var andre versjon av HTML, ble publisert i november 1995 [9] og bygger på en ISO standard for generelle markørpråk (SGML) [20].

I figur 3.2 er det bygget videre på HTML-koden fra figur 2.1 for å illustrere flere grunnleggende elementer i en HTML-side. Selv om dette bare gir en side med et bilde og en lenke, så inneholder den allikevel mange av de mest sentrale begrepene i HTML. En HTML side er organisert som et tre med en <html> markør som rot, denne har igjen to barn. Disse er <head>, som inneholder ting som meta-informasjon om innholdet på siden og lenker til eksterne ressurser som for eksempel CSS filer, og <body> som er innholdet som faktisk vises i nettleseren. HTML-markører til tekst, bilder, lenker til andre nettsider og annet er plassert innenfor <body> markøren og tolkes av nettlesere som presenterer disse. Det er spesielt IMG-markøren og to av dens attributter, ALT og SRC, som skal studeres nærmere i oppgaven.

```
<html>
  <head>
    <title>Min Side</title>
  </head>
  <body>

    <img src=''http://www.uio.no/logo.gif''
          alt=''Universitetet i Oslo sin logo''>

    <a href=''http://www.uio.no''
        title=''Universitetet i Oslo sin hjemmeside''>

      Universitetet i Oslo

    </a>
  </body>
</html>
```

Figur 3.2: Eksempel på HTML kode

3.4 Betydningen av WWW

Siden WWW ble publisert 1. August 1990 har omfanget av nettsted på nettet økt både i antall og type. En rekke organisasjoner og bedrifter har etterhvert sett de muligheter som den nye teknologien gir og benytter den aktivt til å formidle sitt budskap, enten det er kommersielt, politisk, religiøst eller av et utall andre former og variasjoner.

Etterhvert som Internett ble mer og mer tilgjengelig i løpet av 1990 - tallet vokste det frem en rekke tjenester som banker, aviser, butikker, og i det siste, leksika og sosiale nettverk. Dette har i større og større grad gjort at mange organisasjoner har beveget seg bort fra tradisjonelle medier og metoder for å levere sine tjenester og over på elektronisk levering av de samme tjenestene via nettsider og nettapplikasjoner.

3.4.1 Nettapplikasjoner

WWW viste seg også snart å kunne være grunnlaget for en enkel og standardisert måte å lage klient delen av klient-tjener programmer. HTML ga en enkel og standardisert måte å designe grafiske brukergrensesnitt på, samtidig som HTTP og URL ga mulighet for å sende informasjon tilbake til tjeneren via spørrestrengen i URLen og de første nettapplikasjonene var født.

HTML inneholder en `<form>` markør som kan benyttes til å lage skjemaer til å fylle ut og sende inn informasjon som nettjeneren kan behandle. Det er denne markøren og tilhørende `<input>` markører som muliggjør nettskjemaer med grensesnittkomponenter for tekst og diverse andre former for knapper og lister til å velge inndata til applikasjonen.

Dette åpnet for en rekke muligheter der nettsider ikke bare kunne brukes til enveiskommunikasjon, men også til handel, banker, nettsamfunn og andre interaktive applikasjoner.

3.4.2 Nettportaler

Etterhvert som mengden av informasjon på nettet vokste, vokste også behovet for å søke og systematisere informasjonen. Informasjon og data er ikke samme sak, data er *bare* en samling av opplysninger av god eller dårlig kvalitet. Data blir først informasjon nå de er systematisert og katalogisert på en måte som gjør de tilgjengelige gjennom søk, slik at de er praktisk brukbare.

Nettportaler er nettsider som samler informasjon om forskjellige temaer på et sted og presenterer det på en konsistent måte. En nettportalen henter ofte også informasjon om et tema fra forskjellige kilder. Nettportaler kan være både private og statlige, og inneholde tjenester relatert til informasjonen de formidler [21].

Noen kjente nåværende og tidligere nettportaler er America online (AOL), Excite, Yahoo og Skandinavia Online (SOL). Ettersom Internett er blitt mer anerkjent som medium for seriøs kommunikasjon har også offentlige etater opprettet nettportaler. Disse kan ha som oppgave ikke bare å formidle informasjon til allmenheten,

3.4. BETYDNINGEN AV WWW

men også benyttes til innlevering av offentlige skjemaer som for eksempel selvangivelsen. Eksempler på dette er www.altinn.no og www.norge.no. Når det offentlige bruker nettprotaler til å samle inn opplysninger som alle borgere i landet er forpliktet til å levere, for eksempel selvangivelsen, er det spesielt viktig at løsningen som er valgt i nettportalen er tilgjengelig for alle [22].

Yahoo.com er en av de tidligste nettportalene, stiftet i 1994. Selskapet er et av de største som overlevde dot-com boblen rundt århundreskiftet leverer i dag foruten nettportalen en rekke andre relaterte Internettbaserte tjenester som søkemotorer og epost.

3.4.3 Søkemotorer

Nettportaler inneholder mye informasjon, men kan ikke inneholde lenker til alt. I hvertfall ikke på en nettportal alene uten å bli så uoversiktlig at den ikke er praktisk brukbar. En annen måte å gjøre informasjon tilgjengelig på nettet er søkemotorer.

En søkemotor består typisk av tre deler: en nettedderkopp som henter alle ord på et sett nettsider, en indeks denne lager med hvilke ord som opptrer på hvilke nettsider og et grensesnitt til å søke i de indekserte ordene. Indeksen vil også kunne inneholde en innbyrdes rangering av sidene det lenkes til, basert på gitte kriterier.

Google er rangert som det mektigste varemerket i verden ², foran langt eldre selskaper som IBM, General Electric, Microsoft og Coca-Cola. Det er ikke tilfeldig at denne tittelen tilfaller en søkemotor. Mengden av informasjon på Internett vokser eksponentielt og uten en effektiv og pålitelig metode for å finne frem rett informasjon til rett tid er det rett og slett overveldende. Google sin metode for å rangere sider er en viktig del av hemmeligheten bak deres suksess [23] [24].

3.4.4 Nettaviser

Et tidlig eksempel på nettsider som var antatt å revolusjonere en eksisterende bransje er nettaviser. Imidlertid viser det seg at dette har vært mer en evolusjon enn en revolusjon, da en overveldende mengde av nyhetene er kopiert fra de respektive nettavisers papirutgave [25]. Det ser derfor ikke ut til at papiraviser er i ferd med å bli utkonkurrert, men at nettaviser er blitt et naturlig supplement til disse.

De fleste norske aviser kommer i dag ut i elektronisk versjon. Eksempler er landsdekkende aviser som www.aftenposten.no, www.vg.no og www.dagbladet.no, og også lokalaviser som www.dagsavisen.no, www.tk.no og www.nordlys.no. Mens det finnes aviser som er rent elektroniske, som www.nettavisen.no, er det ikke funnet eksempler på trykte aviser som har gått over til bare elektronisk format.

Mange store internasjonale aviser med lange tradisjoner, som for eksempel The Times i London som er grunnlagt allerede i 1785, er i dag publisert også i digital

²<http://www.millwardbrown.com/Sites/mbOptimor/Ideas/BrandZTop100/BrandZTop100.aspx>

form³. I Norge er Nettavisen⁴ den første norske nettavisen som ble opprettet uten å ha en papirutgave. Nettavisen ble stiftet 1. november 1996 og har fremdeles som uttalt mål å dekke nyheter fra hele spekteret, uten å utkomme i papirversjon.

3.4.5 E-Handel

Handel og innkjøp ved å bestille varer på postordre via kataloger er en kjent ting og veien derfra til å bestille varer fra en elektronisk katalog er ikke lang. Nettbaserte butikker var derfor et viktig bruksområde for de tidligste nettapplikasjonene. Det er mange kjente eksempler på E-Handelsapplikasjoner, både norske og internasjonale. Noen kjente er nevnt her uten at det er gjort noen vurdering av disse nettstedenes kvalitet.

Amazon.com er et amerikansk nettsted som begynte som en elektronisk bokhandel. Forretningsidèen var å spare penger ved å unngå store utgifter til lager. Dette skulle oppnå ved å formidle bestillinger direkte fra kunden til trykkeri/grossist. Amazon.com har, etter å ha overlevd dot-com boblen, utvidet til å selge det meste, med fokus på bøker, musikk og filmer etter samme forretningside. Navnet spiller på at Amazonaselven er verdens største i vannvolum, amazon.com skal derfor ha verdens største salgsvolum.

3.4.6 Nettbanker

Banker så også tidlig en mulighet til å rasjonalisere sin drift og gjøre tjenester tilgjengelige for kundene til enhver tid gjennom å gi disse tilgang til å betale regninger og sjekke saldoer på nettet. Dette er også etterhvert utvidet til å gjelde en rekke banktjenester som innsending av lånesøknader og kjøp og salg av aksjer og fond.

Nettbanker er etterhvert blitt en naturlig del av tjenestene som leveres av de fleste banker. Igjen er det få helt rene nettbanker, med Skandiabanken.no som den første som etablerte seg i Norge. Som følge av at den ikke har utgifter til filialer har den kunnet tilby gebyrfrihet og generelt høyere renter på innskudd enn konkurrentene.

3.4.7 Leksika: Wiki

Leksikon og informasjonsutveksling var den opprinnelige motivasjonen bak utviklingen av WWW. Wiki er blitt synonymt med nettapplikasjoner i form av et leksikon som gir brukerne mulighet for selv og redigere artiklene som emner de har kjennskaper om. Ordet *Wiki* betyr *raskt* på språket til urfolket på Hawaii, og det er også et backronym⁵ for *What I Know Is*.

³ www.timesonline.co.uk

⁴ www.nettavisen.no

⁵ Akronym hvor ordene det forkorter ble innført etter akronymet

3.5. STANDARDISERING AV INTERNETT

En wiki kan være alt fra et konversjonsleksikon med artikler om alle mulige typer temaer, til spesialiserte wiki som konsentrerer seg om et spesifikt tema. Noen organisasjoner har også interne wiki til bruk for sine ansatte til å dokumentere opplysninger omkring organisasjonens virksomhet for hverandre.

Svakheten her har blitt hevdet å være kvalitetskontroll av at det som skrives faktisk er riktig og skrevet av personer som faktisk har kunnskaper om temaet de skriver om. Wiki sine tilhengere hevder imidlertid at peer review vil sikre kvaliteten over tid, samtidig som man må ha egne folk til å fjerne de groveste overtrampene. For eksempel ble for en tid tilbake artikkelen om Jens Stoltenberg endret til å indikere at han var dømt for overgrep mot barn, noe som ble fjernet samme dag.

www.wikipedia.org er et nettbasert konversasjonsleksikon med artikler som finnes i versjoner på flere språk. Wikipedia drives av Wikimedia Foundation ⁶, en veldedig stiftelse i USA som har som mål å samle inn og distribuere kunnskap gratis. Wikimedia Foundation driver også flere andre wikiprosjekter, blant annet www.wikiquote.org som er en samling sitater fra både virkelige og litterære personer av alle kategorier.

3.4.8 Sosiale nettverk

Et sosialt nettverk er en samling individer eller organisasjoner som er koplet sammen gjennom en eller flere relasjoner. Relasjonene kan være basert på for eksempel vennskap, slekt, forretningsinteresser, protest mot gitte tema, religiøs trosretning eller fagområde.

Formelle og uformelle sosiale nettverk har eksistert så lenge mennesker har eksistert, men med internett har muligheten til å utvide en persons sosiale nettverk til dimensjoner som tidligere ikke var praktiske, grunnet for eksempel avstand, blitt mulig [26].

I en senere tid har det derfor vært utviklet en rekke nettapplikasjoner som har som sin oppgave å skape og vedlikeholde sosiale nettverk. Navn som Facebook, MySpace og Twitter er alle eksempler på applikasjoner med forskjellige tilnærmelser til begrepet sosiale nettverk. Alle disse har til felles at man, i ulike formater, deler personlige opplysninger med andre brukere, eventuelt et utplukk av den totale brukermassen som man har definert til å være sine venner.

3.5 Standardisering av Internett

Etter hvert som Internett og WWW har vokst og blitt tatt i bruk på stadig flere områder har behovet for å etablere industristandarder for nettstedet og nettapplikasjoner vokst tilsvarende. Standarder inkluderer ikke bare selve teknologiene, som CSS og HTML, men også retningslinjer for hvordan de skal brukes for å oppnå mest mulig brukervennlige nettsteder. Det er derfor opprettet flere grupper som arbeider med dette.

⁶<http://wikimediafoundation.org>

3.5.1 Internet Engineering Task Force

IETF er en gruppe som jobber med å fremme tekniske standarder for Internett. Tatt fra deres egen nettside ⁷ er deres oppdrag;

«The mission of the IETF is to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet.»

IETFs viktigste middel for å løse dette oppdraget er publisering av anbefalte standarder via dokumenter som kalles *Requests For Comment* (RFC). Disse dokumentene beskriver anbefalte standarder relatert til forskjellige teknologier brukt i relasjon til Internett. Eksempler på dette er teknologier som HTTP [10], URL [8], HTML [9], FTP [12] og SMTP [11].

3.5.2 World Wide Web Consortium

World Wide Web Consortium (W3C) ⁸ er en internasjonal organisasjon ledet av Tim Berners-Lee. W3C har en rekke private og statlige virksomheter som medlemmer og har som oppgave å jobbe for standardisering av teknologier og metodikker brukt til utvikling av nettsider.

Et viktig satsingsområde for W3C er at WWW skal være tilgjengelig for alle, inkludert funksjonshemmede. I denne sammenheng har W3C, i samarbeid med blant annet USA sin regjering, dannet *Web Accessibility Initiative* (WAI) som jobber med standarder for hvordan både nettsider og nettlesere kan gjøres mest mulig tilgjengelige for flest mulige brukere.

3.6 Brukergrupper med funksjonshemmelser

Det er noen grupper som i større grad enn andre har hatt problemer med å ta den nye teknologien i bruk. Dette omfatter blant annet personer med forskjellige typer fysiske og psykiske funksjonshemmelser.

For å kunne beskrive tiltak som kan settes i verk for å gjøre nettstedet tilgjengelige for funksjonshemmede er det nødvendig å se nærmere på hva funksjonshemmelse er og hvilke utfordringer disse møter ved bruk av nettsider. Forskjellen mellom enkeltindividets kapasitet og samfunnets forventninger er beskrevet gjennom den såkalte GAP-modellen.

3.6.1 GAP-modellen

I Stortingsmelding 40 (2002-2003) [1] har Det Kongelige Norske Sosialdepartement definert nedsatt funksjonsevne som:

⁷ www.ietf.org

⁸ www.w3.org

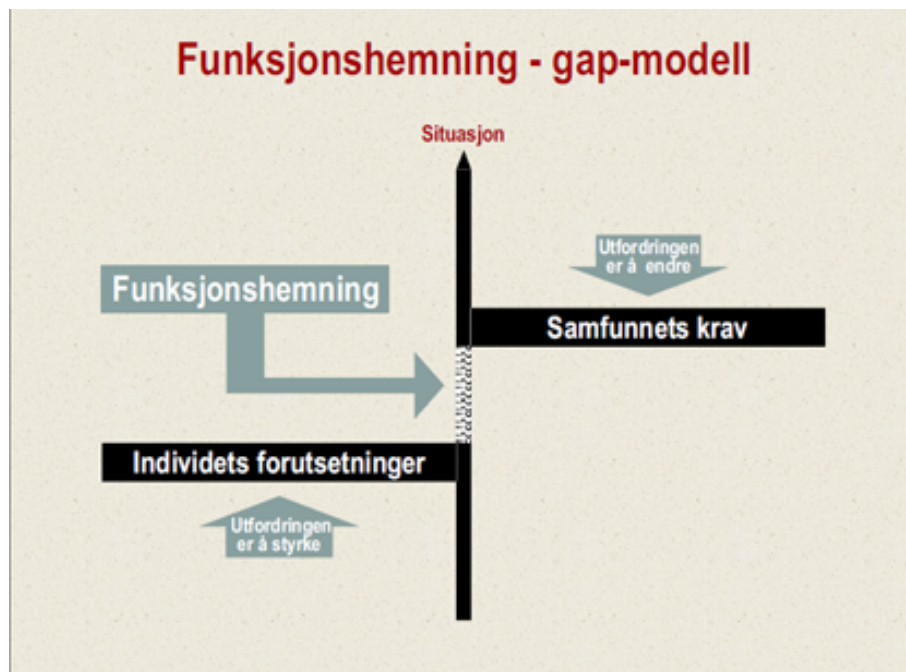
3.6. BRUKERGRUPPER MED FUNKSJONSHEMMELSER

«Med nedsatt funksjonsevne menes tap av eller skade på en kroppsdel eller i en av kroppens funksjoner. Dette kan for eksempel dreie seg om nedsatt bevegelses-, syns- eller hørselsfunksjon, nedsatt kognitiv funksjon, eller ulike funksjonsnedsettelse pga. allergi, hjerte- og lungesykdommer.»

Funksjonshemmelse er i samme dokument definert som

«Det er ingen selvfølge at personer med nedsatt funksjonsevne blir funksjonshemmet. En funksjonsnedsettelse behøver ikke resultere i begrensninger i samfunnsmessig deltakelse. Funksjonshemming oppstår når det foreligger et gap mellom individets forutsetninger og omgivelsenes utforming eller krav til funksjon.»

Gapet mellom et individs evne og omgivelsenes forventninger som kalles GAP-modellen. GAP-modellen er illustrert i figur 3.3 og det er dette gapet som ønskes gjort minst mulig, helst lukket helt, ved hjelp av universell utforming.



Figur 3.3: GAP modellen (Kilde: St. Meld. 40 (2002-2003))

3.6.2 Universell utforming

Universell utforming innebærer at bygninger, maskiner, nettstedet og annet som en person benytter seg av i arbeid og fritid skal kunne benyttes av er tilpasset til å kunne brukes uavhengig av hjelp. Ordet funksjonshemming brukes her i sin videste betydning og omfatter også grupper som ikke oppfattes som funksjonshemmet, for eksempel eldre mennesker som helt naturlig får redusert syn, hørsel og reaksjonsevne.

3.6.3 Synshemming

En gruppe som det er vanskelig å lage nettapplikasjoner for er blinde og synhemmede. Nettsteder er i sin grunnleggende natur visuelle medier, med de følger det får for utbytte av informasjonen som presenteres der dersom man ikke kan se sidene.

De som har vært innom en nettside kan være enige om at det først og fremst er en visuell overføring av informasjon. En slik side består nesten utelukkende av tekst og bilder presentert på en skjerm som som regel ikke kan utnytte blindeskriftsystemer som Braille. Ettersom det, i motsetning til fysiske banker og butikker, ikke er noe personale som kan tre støttende til kan dette fort bli en frustrerende opplevelse dersom en bruke ikke kan se hvor de forskjellige funksjonene er eller benytte seg av disse [27].

3.7 Tilgjengelighet gjennom universell utforming

Universell utforming er tiltak for å gi funksjonshemmede tilgang til områder de ellers ville vært utestengt fra. Bruk av universell utforming for å oppnå tilgjengelighet er blitt mer og mer vanlig i de senere årene, både på Internett og i den fysiske verden. I mange land er det også nedfelt i lovverket, i Norge gjennom Lov om forbud mot diskriminering på grunn av nedsatt funksjonsevne, diskriminerings- og tilgjengelighetsloven, med ikrafttredelse fra 01.01.2009⁹.

I §11 av denne lov pålegges virksomheter en plikt til å sørge for at IKT-løsninger relatert til virksomhetens hovedarbeid, eller som gjør informasjon og tjenester tilgjengelig for allmenheten, er tilrettelagt slik at de er tilgjengelig for funksjonshemmede.

Det er også opprettet et eget senter for universell utforming, Deltasenteret¹⁰, under Barne-, likestillings- og inkluderingsdepartementet, i samarbeid med Helsedirektoratet, for å jobbe med rådgivning og informasjon relatert til universell utforming. Senteret gir ut en rekke foldere og publikasjoner relatert til universell utforming for private og offentlige virksomheter [2] [3] [4].

Tilgjengelighet har lenge vært et viktig tema også internasjonalt. I 1990 vedtok kongressen i USA *Americans With Disabilities Act* (Public Law No: 101-336) som forbyr diskriminering av funksjonshemmede personer i forhold til ansettelse og levering av varer og tjenester. Det britiske parlamentet vedtok i 1995 en tilsvarende lov i *Disability Discrimination Act*.

Som følge av dette tok W3C i 1997 initiativet til å opprette *Web Accessibility Initiative*. Blant initiativtakerne var flere amerikanske og internasjonale universiteter samt USAs regjering¹¹. Flere har kommet til senere.

⁹www.lovdatab.no/all/tl-20080620-042-0.html

¹⁰www.helsedirektoratet.no/deltasenteret

¹¹www.w3.org/Press/WAI-Launch.html

3.7.1 Nettinnhold og brukeragenter

To sentral områder for å gjøre WWW tilgjengelig gjennom universell utforming er nettinnhold og brukeragenter. Nettinnhold er informasjon og tjenester som er gjort tilgjengelige på en nettside eller gjennom en nettapplikasjon. Begrepet omfatter, men er ikke begrenset til, tekst, bilde, skjemaer, lyd og film.

Brukeragenter er programvare som har som sin oppgave å lokalisere, laste ned og presentere data fra nettet. Det mest kjente av disse programmer, kalt *User Agents* på engelsk, er nettleseren [28]. For å oppnå en best mulig opplevelse for funksjonshemmede brukere bør utviklere innenfor både nettsider og nettlesere følge standarder som tar hensyn til deres behov og *Web Accessibility Initiative* omfatter derfor retningslinjer for begge.

3.7.2 Web Accessibility Initiative

Web Accessibility Initiative (WAI) ble etablert for å utvikle standarder for hvordan nettsider skal optimaliseres for dekke behovene til grupper som av forskjellige årsaker har spesielle behov som må tas hensyn til for at disse gruppene skal kunne få tilsvarende nytteverdi av Internett som resten av befolkningen [6]. WAI handler med andre ord om å realisere universell utforming på Internett.

Arbeidet i WAI og deres arbeidsgrupper har resultert i retningslinjer på flere områder for å gjøre Internett mer tilgjengelig for personer med diverse funksjonshemmelser. De viktigste i denne sammenhengen er *Web Content Accessibility Guidelines* (WCAG) og *User Agent Accessibility Guidelines* (UAAG).

WCAG beskriver hvordan innholdet på nettsider kan gjøres mer tilgjengelig for forskjellige brukergrupper. UAAG beskriver hvordan programmer, som brukes til å laste ned og tolke hvordan en nettside skal presenteres, skal lages for å gjøre disse brukervennlige og tilgjengelig for de samme brukergruppene.

3.7.3 Web Content Accessibility Guidelines

Web Content Accessibility Guidelines (WCAG) er retningslinjer for hvordan innholdet på nettsider kan tilpasses for brukere med ulike funksjonshemmelser. Det er primært rettet mot funksjonshemmede brukere som bruker nettlesere på en vanlig PC, men også for brukere som ønsker å se en nettside på enheter med begrensede ressurser, for eksempel mobiltelefoner [29]. I skrivende stund er WCAG kommet i versjon 2.0.

3.7.3.1 WCAG 1.0

WCAG 1.0 [30] ble publisert i mai i 1999 og består av 14 retningslinjer, hver med et antall sjekkpunkter. Sjekkpunktene er delt inn i tre kategorier, punkter som *MÅ* oppfylles, punkter som *BØR* oppfylles og punkter som *KAN* oppfylles. En nettside som

tilfredsstillende alle MÅ krav er sagt å være på konformitetsnivå A, dersom alle BØR krav også er med er nivået AA og hvis KAN også er med er det AAA.

Retningslinjene er høynivå abstraksjoner. For eksempel er den første retningslinjen «Inkluderer alternativ tekst til lyd og visuelt innhold». Denne formulerer et klart mål, men ikke hvordan dette skal kunne oppnås.

Hvordan målet med retningslinjene skal oppnås er konkretisert i sjekkpunktene. Det første sjekkpunktet i den nevnte retningslinjen er «Ha alltid et ekvivalent tekst alternativ til ikke-tekstlig innhold». Dette er et MÅ krav i WCAG 1.0, men det krever altså at ALLE IMG-markører har det, ikke bare de som formidler informasjon, men også de som gir layout og estetikk.

3.7.3.2 WCAG 2.0

Mens WCAG 1.0 bestod av retningslinjer og sjekkpunkter består WCAG 2.0 av fire prinsipper for tilgjengelighet [31][32]. Hvert prinsipp består av retningslinjer som igjen består av testbare suksesskriterier. Kriteriene er rangert A, AA, AAA. For at en nettside etter retningslinjen skal tilfredsstillende et nivå, må alle kriterier på det nivået være tilfredsstillende. De fire prinsippene for tilgjengelighet som definert i WCAG 2.0 er gjenkjennbart (Perceivable), anvendelig (Operable), forståelig (Understandable) og robust (Robust).

Gjenkjennbart At en nettside er gjenkjennbar vil si at informasjon og komponenter i brukergrensesnittet på nettsiden er presentert på en slik måte at brukeren kan nyttiggjøre seg av den ved å bruke minst en av sansene. Altså må det være mulig å benytte seg av nettsiden selv om brukeren har redusert funksjonsevne i en eller flere sanser.

Den retningslinjen som tilsvarer den som ble brukt som eksempel i omtalen av versjon 1.0 er lagt under dette prinsippet. Den er noe omformulert til «Ha alltid et ekvivalent tekst alternativ til ikke-tekstlig innhold slik at det kan endres til andre former som brukeren trenger, slik som større tekst, braille, tale, symboler eller enklere språk», noe som gir en mer detaljert beskrivelse av hvordan retningslinjen skal oppfylles. Også her er dette et MÅ krav, altså et suksesskriterier på A-nivå.

Anvendelig At en nettside er anvendelig vil si at alle operasjoner med komponenter i brukergrensesnittet, herunder navigasjon, må være enkle å utføre. Det må for eksempel være mulig å navigere frem og tilbake på et nettsted selv om man har redusert førighet i hender eller armer. Det bør for eksempel ikke være krav om at en serie med kommandoer skal gis innen bestemte tidsrom. Scenarioer av typen *For å aktivere modul A må man trykke knapp X, for så å trykke knapp y innen 0.5 sekunder* bør ikke forekomme.

Forståelig At en nettside er forståelig vil si at informasjonen som presenteres, og hvordan operasjoner med komponenter i brukergrensesnittet skal utføres, må være forståelig. Blant annet må det være mulig å identifisere, og ved behov bytte språket.

Dette gjelder både det som er sidens hovedspråk og dersom deler av siden er på andre språk. Et annet eksempel er at forkortelser ikke må forekomme uten at betydningen er tilgjengelig.

Robust At nettsiden er robust vil si at den tolkes på en konsistent måte over mange forskjellige nettlesere, inkludert støtteverktøy som skjermlesere.

3.7.4 User Agent Accessibility Guidelines

User Agent Accessibility Guidelines (UAAG) er et sett med retningslinjer for hvordan de programmer som benyttes til å hente ned og vise frem en nettside kan tilpasses funksjonshemmede [28]. Som for WCAG består de av retningslinjer med sjekkpunkter som en nettleser må overholde. Det vektlegges spesielt at nettleseren må støtte oppnåelsen av retningslinjene fra WCAG.

3.8 Universell utforming for blinde og synshemmede

Som følge av WWW sin grafiske natur er det en gruppe som er spesielt utestengt. Dette er de blinde og synshemmede brukerne. Det er derfor utviklet en rekke hjelpemidler for denne gruppen, hvor de mest sentrale er skjermlesere og brailletastaturer.

3.8.1 Braillealfabetet

Braillealfabetet det internasjonalt mest brukte system for blindeskrift og ble oppfunnet av franskmannen Louis Braille. Braille ble blindet i en ulykke da han var tre år gammel, noe som motiverte han til utviklet braillealfabetet. Dette var ferdig i 1824 da Braille var 15 år gammel [33].

Brailleskrift er basert på en teknikk for å la soldater dele informasjon om natten uten lys og lyd og fungerer ved at hver bokstav, tall og tegn er representert ved et sekspunktstegn, slik det er vist med bokstaven *a* i figur 3.4. Foruten tall og bokstaver finnes det også tegn for å indikere om en bokstav er majuskel ¹², se figur 3.5 og for skilletegn som komma, anførselstegn, punktum og parenteser.

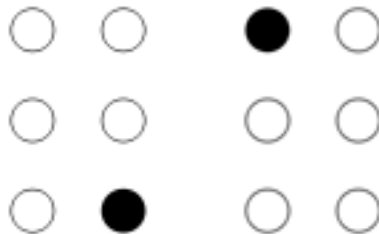
På grunn av det relativt lille antallet symboler, $63 (2^6 - 1 = 63)$, som kan representeres ved seks punkter er det nødvendig å gjenbruke tegn. Ved å sette et spesielt tegn foran brailletegnet for en bokstav endres tegnet sin betydning til å indikere at tegnet skal tolkes som det tallet som tilsvarer bokstavens plassering i alfabetet, eksempel på dette er gitt for tallet *1* i figur 3.6.

Av samme årsak er noen tegn gjenbrukt avhengig av kontekst. Tegnet i figur 3.7 betyr derfor start av sitat (anførselstegn) hvis plassert først i en setning og spørsmålstegn hvis plassert sist i en setning.

¹²Stor bokstav



Figur 3.4: Braille tegn for bokstaven a (Minuskel) (Kilde: wikipedia.org)



Figur 3.5: Brilletegn for bokstaven A (Majuskel) (Kilde: wikipedia.org)

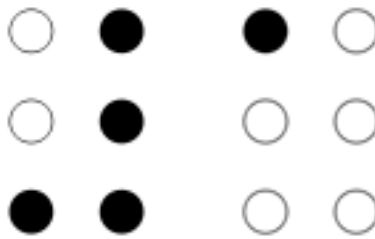
3.8.1.1 Brailletastatur og leselist

Et brailletastatur er et tastatur hvor bokstavene i for eksempel det latinske alfabetet er byttet ut med, eller supplementert av, de tilsvarende bokstaver, tall og tegn i braillealfabetet. Dette er gjort ganske enkelt ved at tastene har små prikker slik at brukeren kan føle seg frem til bokstavene med fingrene.

Tilsvarende er en leselist et apparat for å hente ut data som skriver teksten med brailletegn ved å heve og senke passende prikker på listen. Brukeren kan så lese teksten ved å kjenne hvilke tegn som presenteres på leselisten. Leselisten kan variere i antall tegn den kan presenter av gangen, vanligvis 40-80 tegn, og hvilke tegn som presenteres, og kan beveges frem og tilbake av brukeren for å tilpasse dennes lesehastighet [34].

3.8.2 Skjermlesere

En skjermleser er et program som brukes av blinde og svaksynte personer, eller andre grupper som av ulike årsaker ikke kan lese teksten selv [34]. Skjermleseren kontrolleres via hurtigtaster, da bruk av musepeker ikke er hensiktsmessig når brukeren ikke kan se pekeren på skjermen. Skjermlesere leser tekst på en dataskjerm enten ved å skrive teksten ut på en leselist eller ved å lese teksten høyt for brukeren ved hjelp av



Figur 3.6: Braille tegn for tallet 1 (Kilde: wikipedia.org)



Figur 3.7: Brailletegn for start anførselstegn eller spørsmålstegn (Kilde: wikipedia.org)

talesyntese. I Norge er skjermleseren *Job Access With Speech (JAWS)*¹³ en av de mest brukte skjermlesere [34].

3.8.2.1 Talesyntese

Talesyntese [34] er en teknikk som benyttes av skjermlesere, og andre programmer som genererer på forhånd ukjent tale, for så å omsette teksten på skjermen til lyd. Selv om talesyntese med tiden har blitt bedre er det fremdeles et stykke igjen til de produsere helt naturlig tale. Det kan derfor være slitsomt å høre på en syntetisk stemme over tid og mange blinde og svaksynte bruker derfor leselist som *avlastning* [34].

3.8.2.2 Skjermlesere og grafikk

Skjermlesere og leselister fungerer ved lineær gjennomlesning av en tekst [34]. De er derfor lite egnet til å lese data i to dimensjoner, for eksempel tabeller. Nettsider som nytter tabeller til å oppnå layout, for eksempel kolonner i en tekst, er derfor lite tilgjengelige for blind og svaksynte som benytter skjermleser.

En bedre metode for å oppnå samme resultat er å legge innholdet i de forskjellige sonene på nettsiden i bokser, definert ved hjelp av DIV-markører, som så kan plasseres

¹³<http://www.freedomsscientific.com>

på ønsket sted på nettside med CSS. Dermed kan nettsiden lineariseres ved å be nettleseren ignorere CSS definisjonene, slik at teksten kan leses lineært av skjermleseren.

Mange nettsider inneholder bilder og annen grafikk. WCAG standarden forutsetter at alle HTML IMG-markører skal inneholde en ALT-attributt som gir en alternativ tekst som beskriver bildet nettopp med tanke på slike skjermlesere. I mange tilfeller er imidlertid ALT-attributten full av intetsigende informasjon, for eksempel filnavnet, en tom streng (`alt=""`) eller rett og slett utelatt fullstendig. I disse tilfeller blir ALT-attributten støy, den gir ingen verdi for den som lytter til skjermleseren.

Om skjermleseren som bruker benytter seg av bare bruker talesyntese eller bare lese-list, eller en kombinasjon av dette, antas ikke å påvirke algoritmene senere i oppgaven. Algoritmene skal klassifisere og rangere bilder etter viktighet, men hvordan dette meddeles brukeren har ikke betydning.

3.9 Informative bilder

På en nettside kan det være bilder som det kan diskuteres hvorvidt det er av interesse å nevne dersom brukeren ikke kan se disse. Rammer på tabeller, skillelinjer til pynt under overskrifter og liknende har ofte en estetisk oppgave, og kan bidra til at en seende bruker oppfatter nettsiden som seriøs eller ikke [35], men for blinde og svaksynte kan det bli unødig støy når skjermleseren bare sier *bilde*.

Ved å identifisere de bildene som antas viktige som informasjonsbærere på nettsiden kan bilder uten informativ bildetekst utelates av skjermleseren. Et godt eksempel på dette er bilder som bare er pynt i følge med nettsidens layout, for eksempel rammer på tekster og tabeller. Disse antas å gi liten mening dersom brukeren ikke kan se dem.

Ved å se på sammenhengen med bildets struktur og plassering i teksten ønsker vi å finne ut om vi kan filtrere vekk irrelevant bilder, det vil si bilder som bare er med for å gi et penere visuelt inntrykk og ikke formidler noen nye opplysninger i seg selv.

3.10 Bilder i nettsider: IMG-markøren

Den primære metoden for å plassere et bilde som en del av innholdet i en nettside er IMG-markøren. Selv om det er mange andre metoder å vise et bilde i en nettside på, for eksempel ved hjelp av *background-image* attributten på en tabell, er det IMG-markøren som denne oppgaven vil konsentrere seg om. Grunnen til dette er at *background-image* og likenede skal brukes til design og layout og at det er ikke meningen å formidle informasjon med disse.

IMG-markøren inneholder en URL til bildet, SRC-attributten og en alternativ forklarende tekst, ALT-attributten slik det er vist i figur 3.8.

Det er i denne ALT-attributten som utvikleren av nettsiden plasserer teksten som skal forklare innholdet i bildet som vises av IMG-markøren. Det optimalt ønskelige er at det legges samme tekst i ALT-attributten som den som vises som synlig bilde-

3.10. BILDER I NETTSIDER: IMG-MARKØREN

```
<IMG src=''<URL til bilde>''  
      alt=''Tekst som forklarer bildet'' />
```

Figur 3.8: HTML sin IMG-markør for visning av bilder, med ALT- og SRC- attributtene

tekst på siden. Hvis det ikke er noen synlig bildetekst bør det allikevel legges inn en meningsfylt beskrivelse av bildet i ALT-attributten.

Erfaring, som resultatene i kapittel 11, viser imidlertid at det slurves mye med hva som legges inn i ALT-attributten. Ofte er den ikke inkludert i IMG-markøren i det hele tatt. Når den er inkludert inneholder den ofte tekst som ikke gir noen spesielle opplysninger som er relevante for en bruker som ikke ser selve bildet. Typiske eksempler på dette er:

- «Logo» eller «Foto» alene.
- Tall som «5» eller «343423», uten at noen meningen eller kontekst oppgis.
- «EmediateAd», navnet på et reklame byrå, men ikke HVA der er reklame for.
- «Mange drept» er tragisk, men er det nyheter eller filmanmeldelse?
- Kun navnet på bildefilen.

Det er rimelig å anta at en person som ikke kan se det aktuelle bildet selv vil oppfatte dette mer som støy enn som nyttige opplysninger. Det er derfor ønskelig at skjermleseren filtrerer bort disse og bare tar med tekster som den, basert på et gitt regelsett, oppfatter som å inneholde nyttig informasjon. Eksempel bildetekster som kan antas å være nyttige er:

- «Inngangen til Byfjordtunnelen i Rennfast (Foto: Wenche Lamo)»
- «Skya», «Lettsky», «Sol»
(Når man vet at man er på en værside og hvor og når varselet gjelder).
- «Illustrasjon: Røykvarsler»

En skjermleser er derfor avhengig av at utvikleren av nettsiden gir en meningsfylt beskrivelse av bildet. Dette er ikke alltid gjennomført og for å optimalisere skjermleseren er det derfor ønskelig å kunne filtrere bort bilder skjermleseren ikke kan gi noen meningsfylt beskrivelse av.

3.11 Målsetning

Med de temaer som er omtalt i dette kapittelet som bakgrunn er det ønskelig å finne regler som kan gi en indikasjon på om et gitt bilde på en nettside er informasjon eller støy for en blind eller svaksynt bruker som får nettsiden lest opp av en skjermleser.

Metoden som er tenkt brukt er å finne egenskaper ved bildet og dets bildetekst som kan brukes til å rangere bildet. Bilder som faller under en gitt terskel forkastes av skjermleseren. Egenskaper som er aktuelle vil bli gjennomgått i påfølgende kapitler og vil er basert på tekstens innhold og bildets antatte egenskaper som blikkfang på nettsiden.

Som første steg i bildeanalysen vil det være nødvendig å finne bildets eksakte posisjon på nettsiden slik den faktisk fremstår i den nettleseren brukeren benytter. Grunnet forskjellig implementasjon av nettlesere tolker disse ikke alltid standarder likt. Det må derfor søkes å finne en metode for å finne bildets posisjon på skjermen som er uavhengig av nettleserens implementasjon.

Del II

Tekstmodalitet

Kapittel 4

Tekst og språk

Muligheten til å forme og forstå tekst er en viktig forutsetning for at sivilisasjon og samfunn skal bestå og kjenne til sin historie, kultur og teknologi, og ikke minst utvikle disse videre over generasjoner [36].

Tekster skrevet i naturlige språk, og tolking av disse med datamaskiner, er et omfattende og komplisert tema [37]. Dette kommer først og fremst fordi datamaskiner er avhengige av en fast og uforanderlig struktur på innkommende data. De er flinke til å behandle store mengder data med stor presisjon og hurtighet, men ikke flinke til å håndtere uventede hendelser.

Strukturen på naturlige språk er ikke så rigid som i et programmeringsspråk, og naturlige språk trekker også på konteksten de uttrykkes i for å gi korrekt mening. For eksempel vil et menneske kunne forstå at en setning er uttalt i et sarkastisk tonefall, og derfor betyr det motsatte av det en ren bokstavtro tolkning av ordene skulle tilsi.

For å illustrere noen av utfordringene ved tolkning av naturlige språk med datamaskiner er det nødvendig med en kort innføring i noen grunnleggende begreper ved disse.

4.1 Tekst

En tekst er en samling ord og setninger som tilsammen gir en mening. Tekster faller inn i flere kategorier, for eksempel fag- eller skjønnlitterære tekster, med tilhørende underkategorier, kalt sjangere, som spenner fra teknisk håndbok til romantisk novelle. Alle disse har sin generelle form, men likevel med mulighet for individuelle tolkningert, og tolkes av leseren ut ifra dennes kulturelle ståsted. Tekster kan være av forskjellige lengder, fra korte beskjedder i en SMS til bøker over flere bind.

Felles for de fleste sjangere er at en tekst må være sammenhengende i betydning av at den beskriver en bestemt ting, hendelse eller mening på en konsistent måte. Ordet tekst kommer fra det latinske ordet *textus*, som betyr sammenføyning eller vev, i betydningen av å veve tøystoffet.

4.2 Språk

Et språk er et system for å formidle informasjon. Ting og begreper er representert ved hjelp av symboler, tegn og ord, som er kombinert ved hjelp av gitte regler for å gi bestemte meninger [36] [38]. Studiet av språks struktur og oppbygging kalles lingvistikk, og er meget mer omfattende enn den korte beskrivelse gitt i de påfølgende avsnitt. Det skulle imidlertid være omfattende nok til å illustrere de utfordringer som møtes ved prosessering av naturlige språk med datamaskiner.

4.2.1 Fonemer, bokstaver og alfabet

En bokstav er et symbol brukt i et skriftspråk til å symbolisere en uttalt lyd i talespråket. I en tekst plasseres bokstavene etter hverandre på en slik måte at de tilsammen utgjør ord, på samme måte som lydene uttalt sekvensielt etter hverandre ville utgjort det samme ordet i talespråket. Et språks bokstaver er samlet i språkets alfabet, som beskriver bokstavenes relasjoner til hverandre [38].

Navnet *alfabet* kommer fra de to første bokstavene i det greske alfabetet, alfa og beta, og er en ordnet mengde av de bokstaver som representerer de grunnleggende lydene i et språk. Disse lydene kalles også fonemer. De er den minste lyd som menneskets stemme kan produsere som det er mulig å tilordne noen meningsfull betydning, det vil si skilles fra andre lyder.

Siden bokstavene i et alfabet utgjør en ordnet mengde er det mulig å ordne ord og uttrykk skrevet med alfabetet sine bokstaver i en bestemt rekkefølge, som kalles alfabetisk rekkefølge. Alfabetet er imidlertid bare begynnelsen på hvordan en tekst bygges opp, det første som deretter må gjøres er å forme morfemer og ord fra bokstavene.

4.2.2 Morfemer, ord og ordklasser

Et *morfem* er den minste enheten i språket med betydning alene. Fonemer har mening i den forstand at de representerer en lyd, men denne lyden er et morfem hvis og bare hvis den formidler en mening. Et ord kan bestå av ett morfem som mat, fat, skole, barn, eller flere morfemer som i mat-fat-et og skole-barn-a.

Et ord er en meningsbærende enhet som kan uttales med et eget selvstendig trykk. De enkelte delene av et ord kan ikke bytte plass eller skilles fra hverandre ved at et annet ord blir skutt inn. Et ord er bygget opp av en stamme, som består av en rot/flere røtter og eventuelt avledningsmorfemer, samt eventuelle bøyningmorfem [38].

Ord grupperes i ordklasser basert på at de har like, eller liknende, morfologiske, syntaktiske og semantiske egenskaper. Hvilke ordklasser et ord tilhører vil påvirke hvilken rolle ordet kan ha i en setning. De viktigste roller i en setning er objekt, subjekt og verbal. Objekt er den eller det som gjøre noe, det som gjøres er verbalet og den eller det som det gjøres noe med er subjektet [36].

4.2.3 Grammatikk

For å kombinere ord sammen til større enheter med mening er det behov for bestemte regler. Disse reglene kalles språkets *grammatikk* og varierer i større og mindre grad fra språk til språk, selv om det også er likheter. Grammatikk er de regler et språk har for hvordan man kombinerer ord til fraser og setninger. Moderne grammatikk bestemmer både hva som er språkets form, morfologi og syntaks, og hva som er språkets mening, semantikk og pragmatikk [38].

Morfologi er den delen av grammatikken som beskriver regler for ordenes struktur. Den omfatter hvordan ord er bygget opp fra morfemer og hvordan ord kan settes sammen med andre ord og morfemer for å danne nye ord. Morfologi omhandler også hvordan ordene inndeles i ordklasser og reglene for hvordan ordene i hver ordklasse skal bøyes.

Syntaks omhandler hvordan ord settes sammen til fraser og setninger. Syntaksen bygger således videre på morfologien og ser ordene i sammenheng. Syntaks er det regelsett man bruker i setningsanalyse, hvor man ser på hvilken rolle de enkelte ord spiller i setningen, som for eksempel subjekt, det som gjør noe, objektet, der som det gjøres noe med og verbalet, det som subjektet gjør med objektet [38].

Semantikk beskriver hva ord, fraser og setninger betyr, hvilket innhold de har og hvilke ting, handlinger og konsepter de representerer. Semantikken henger derfor nøye sammen med syntaksen, fordi ordenes plassering i forhold til hverandre i en setning kan ha betydning for ordenes betydning. En syntaktisk korrekt setning kan være vrøvl, fordi den for eksempel omtaler ting som ikke eksisterer eller handlinger som ikke kan utføres. *Den grønne kua hoppet over månen* er en syntaktisk korrekt setning, men selv om kuer nok kan males grønne er det sjelden de er i stand til å hoppe over månen.

Syntaktisk like setninger kan ha forskjellige betydninger, når man ser på hvilken kontekst de er uttalt i. Pragmatikken ser derfor på hva den/de som har laget teksten ønsker å kommunisere. Her vil de kognitive og sosiale forutsetninger for kommunikasjon være sentrale temaer. Å oppfatte subtile konsepter, som ironi ved at man skriver det motsatte av det man mener, er tolkning av tekstens pragmatikk [38].

4.3 Tolkning av tekster

Hensikten med det foregående avsnittet er å gi et inntrykk av de utfordringene som oppstår ved tolkning av tekster med datamaskiner. De språk som mennesker nytter til å kommunisere seg imellom er ikke rigide og gir rom for subtile begreper som ironi og eufemismer. Det kan derfor være nødvendig å definere en passende delmengde av tekster som er aktuelle i problemområdet som behandles.

Den komplekse strukturen i naturlige språk, og de begrensede mulighetene datamaskiner har til å tolke disse, gjør det viktig å definere et så eksakt som mulig problemområde for hva applikasjonene eller algoritmen som skal utvikles skal løse.

Oppgavens mål er å filtrere ut bildetekster som ikke bidrar til å tydeliggjøre nettsi-

dens budskap. Disse teksten er som oftest ikke omfattende prosaiske verk da intensjonen er å gi en kort oppsummering av bildets innhold.

Oppgaven vil forutsette at tekstene som skal analyseres vil være forholdsvis korte. Bare unntaksvis vil de være et helt avsnitt eller mer og normen vil være en kort beskrivende setning med tre til fem ord. Eksempler på denne typen tekster kan være «Utsikt fra Prekestolen, Rogaland», «Graf over aldersgrupper i Norge» og «Illustrasjonsfoto av åstedet».

4.4 Datamaskiner og naturlige språk

Hensikten med denne korte repetisjonen av grunnleggende språklige begreper er å illustrere at tekster skrevet i naturlige språk er kompliserte og sterkt varierende strukturer. Samme setning kan ha forskjellig betydning, og forskjellige setninger kan ha samme betydning, avhengig av kontekst. Dette bunner ut i menneskets psykologi, hvor uttrykksformen kan velges basert på avsenderens agenda og mottakergruppens forutsetninger.

Det er imidlertid ikke slik at maskinell tolkning av tekster skrevet i naturlige språk i alle sammenhenger er en praktisk umulighet. Emnet innenfor datavitenskapen som omhandler dette kalles prosessering av naturlige språk (NLP) ¹ [39]. Betegnelsen naturlige språk er gitt for å skille språk som benyttes av mennesker fra mer strukturerte og formelle språk som programmeringsspråk eller matematisk og logisk notasjon. Selv om det i denne teksten fokuseres på analyse av tekster i naturlige språk omfatter NLP også generering av tekster, for eksempel i programmer som ved hjelp av kunstig intelligens [40] skal produsere svar på forespørsler, som for eksempel Eliza [41].

4.4.1 Datalingvistik

Forholdet mellom datamaskiner og språk er tema for *datalingvistikken* [42]. Hvilke av språkets egenskaper kan kodes i datamaskiner, og hva er det maskinene kan *forstå*? Vi ser nærmere på blant annet automatisk syntaktisk og semantisk analyse, noe som er sentralt for automatisk oversettelse mellom språk.

Morfologi og syntaks er de delene av grammatikken for naturlige språk som er lettest å bruke i databehandling, da disse er basert på et stort, men endelig regelsett. Semantikk og pragmatikk derimot kan være mer subtilt og en eksakt tolkning av semantisk og pragmatisk betydning av en tekst kan være avhengig av detaljkunnskaper om forfatteren og vedkommendes agenda.

¹Natural language processing

Kapittel 5

Tekstanalyse

Det første steget i å analysere relevansen av et bilde og tilhørende bildetekst er å finne ut om det er sannsynlig at bildeteksten i det hele tatt meddeler noen informasjon. Dersom bildeteksten skulle vise seg å mangle helt, eller at den er uten mening i den kontekst den leses i, er det antakelig ikke ønskelig at skjermleseren inkluderer den. Å gå videre med analyser av selve bildet kan unngås, da dette sparer ressurser i form av minne og prosessor.

Kapittel 4 omhandlet utfordringer med å tolke og klassifisere tekster i et naturlig språk ved hjelp av datamaskiner. Dersom det antas at bildetekster er en delmengde av alle mulige tekster i et naturlig språk er det mulig å definere et antall klasser av bildetekster. Disse klassene vil være basert på enkle regler i språkets grammatikk. Videre kreves det ikke at en tekst kun kan falle inn i en tekstklasse.

Det må også være mulig å rangere en tekst etter hvor interessant den er innenfor hver klasse. I denne oppgaven er det lukkede intervallet $[0.0, 1.0]$ valgt som skala for rangering hvor 0.0 angir at bildet antas å være støy og 1.0 angir at bildet antas å være interessant. Dette vil gjelde både for bildetekster i dette kapitlet og for rangering av bilder i kapittel 8.

For å kunne definere konkrete tekstklasser etter disse reglene er det viktig å begrense klassene til eksakt det delområdet av det aktuelle språket som behandles. I tilfellet bildetekster på nettsider vil dette være relativt korte tekster som gir en beskrivelse av et bilde for personer som ikke kan se bildet, men som allikevel kan tenkes å ha nytte av informasjonen som bildet er ment å formidle. Tekstene kan også være av typen enkeltord, uttrykk eller tall.

Dette kapitlet beskriver tekstbaserte metoder for å klassifisere bildetekster og rangere disse etter antatt verdi som informasjonsbærer. De resterende bildetekstene vil bli antatt å utgjøre støy for en blind eller synshemmet bruker som er avhengig av en skjermleser. Det antas at metodene for rangering er uavhengige av hvilket språk nettsiden er skrevet i så lenge de nødvendige ordlister er tilgjengelige.

5.1 Tekstklasser

På bakgrunn av innsamlet materiale fra forskjellige kategorier norske nettsider er det blitt valgt ut et antall forskjellige tekstklasser som bildetekstene i ALT-attributter kan falle inn i. Listen, som ikke kan regnes som fullstendig, inneholder følgende tekstklasser:

- Teksten mangler helt.
- Et flertall av ordene i teksten er oppgitt i ordlisten for det aktuelle språket.
- Teksten inneholder/inneholder ikke bestemte nøkkelord.
- Teksten er en filsti.
- Teksten består av et tall uten spesifisert mening som kilogram, meter eller antall.
- Samme tekst er repetert i to eller flere ALT-attributter på samme nettside.

Siden det er mulig for en bildetekst å tilhøre mer enn en klasse vil det være summen av rangeringene fra hver enkelt analyse som bestemmer teksten sin totale rangering. Hvordan den samlede rangeringen av tekst- og bilderangeringer fremkommer er omtalt i kapittel 9.

Hver analysefunksjon er gitt en formell definisjon. Konvensjonen for funksjonsnavn er $RT_{navn}(tekst)$, hvor RT er *rang av tekst*, $navn$ er navn på den enkelte analyse og *tekst*-parameteren er teksten som analyseres.

5.1.1 Tom eller manglende ALT-attributt

IMG-markøren har to attributter som er av betydning i denne oppgaven. Disse er, som vist i figur 5.1, SRC- og ALT-attributtene som henholdsvis forteller nettleseren hvor bildefilen finnes og hva som er alternativ tekst dersom denne bildefilen ikke kan vises. SRC-attributten som forteller hvor bildefilen kan hentes vil bli mer omtalt i kapittel 8.

```
<IMG SRC='filsti til bildet'  
      ALT='Alternativ bildetekst' />
```

Figur 5.1: IMG-markøren

ALT-attributten regnes som manglende når den er utelatt og som tom når den ikke inneholder noen tegn ($ALT=""$). At en IMG-markør mangler ALT-attributt er et brudd på det første sjekkpunktet i den første retningslinjen fra WCAG 1.0 [30], som krever at det skal være en alternativ tekst for alle ikke-tekstlige deler av en nettside. Det

er imidlertid ikke noe krav i HTML [7] at en IMG-markør *skal* inneholde en ALT-attributt, slik at en nettleser allikevel vil vise bildet.

At mange bilder inneholder en tom ALT-attributt er ikke uventet. Metoder som brukes til å verifisere at en nettsides HTML-kode er korrekt [43] [44], og at den oppfyller regler fra WCAG, stiller bare krav til at ALT-attributten er tilstede, ikke at den skal inneholde noen tekst. Forfatteren av en nettside kan legge disse til i den hensikt at siden skal kunne valideres når den sjekkes opp mot definisjonene for HTML [7] og CSS [45] eller WAI [6] og WCAG [29].

Den tomme eller manglende bildeteksten er i denne sammenheng definert til å være støy. Dette fordi den gir få eller ingen interessante opplysninger for en synshemmet bruker, som bare blir påminnet om at her er det noe som vedkommende ikke får vite hva er.

Reglene for tekstklasser som definert her gir bare to mulige rangeringer i denne klassen. Rangeringen må være 0.0 dersom teksten mangler og 1.0 hvis ikke gjør det. Den formelle rangeringsfunksjonen er gitt i uttrykk (5.1).

$$RT_{\emptyset}(tekst) = \begin{cases} 1.0 & : tekst \neq \emptyset \\ 0.0 & : tekst = \emptyset \end{cases} \quad (5.1)$$

5.1.2 Ikke i ordlisten

For at en tekst skal gi mening og ha troverdighet er det viktig at den er skrevet korrekt rettskrivningsmessig og grammatikalsk. Det å håndheve alle grammatikalske regler er en krevende oppgave. Det er imidlertid mulig å gjøre mye ved å kontrollere en av de mest grunnleggende, nemlig korrekt staving av ord [46].

En måte å gjøre dette på er å definere korrekt stavet ord til å være alle ord som forekommer i en gitt ordliste. Alle andre ord er så per definisjon ukorrekt stavet. Det vil således være nødvendig med en ordliste for hvert språk som er aktuelt for den enkelte bruker. Det kan også være nødvendig å spesifisere om en tekst skal kunne inneholde ord på mer enn et språk, for eksempel fordi de inneholder utenlandsk teknisk terminologi. Her kan også egne fagordlister for de fagområder bruker er interessert i benyttes. I tillegg må også tall som forekommer i teksten telles som korrekt stavete ord.

Formell definisjon av rangeringsfunksjonen er gitt ved uttrykket i (5.2) slik at en tekst hvor alle ord er korrekt stavet er rangert som 1.0 og et hvor ingen er korrekt stavet er rangert som 0.0.

$$RT_{Staving}(tekst) = \frac{\text{Antall korrekt stavede ord i teksten}}{\text{Totalt antall ord i teksten}} \quad (5.2)$$

5.1.3 Bestemte nøkkelord

En tekstklasse som ikke direkte fremkommer fra det innsamlede materialet, men som allikevel antas å kunne ha betydning, er rangering av tekster etter innhold av bestemte

5.1. TEKSTKLASSER

nøkkelord. Dette vil gi en bruker muligheten til selv å påvirke hvilke bildetekster som skal inkluderes av skjermleseren. Det er i denne rangeringsfunksjonen tatt høyde for at det både kan finnes ord som ønskes i teksten og ord som ikke ønskes i teksten.

Algoritmen for rangering etter innhold av nøkkelord er bygget opp rundt to lister av nøkkelord. Ord i den ene gir minuspoeng fordi det *ikke er* ønskelig med tekster som inneholder disse og en som gir plusspoeng for ord som *er* ønskelig at teksten skal inneholde. Det forutsettes også at samme ord ikke kan inngå i begge lister samtidig. Disse forutsetninger er gitt i uttrykk (5.3) sammen med selve rangeringsfunksjonen $RT_{Noekkel}(tekst)$. Systemer for automatisk klassifisering av tekster er komplisert og vil antakelig kreve at nøkkelordlistene oppdateres over tid [47].

$$\begin{aligned} T &= \{\text{Alle ord i teksten}\} \\ I &= \{\text{Alle nøkkelord som gir plusspoeng}\} \\ E &= \{\text{Alle nøkkelord som gir minuspoeng}\} \\ T_I &= \{\text{Alle ord som inngår i T og I, inkludert duplikater fra T}\} \\ T_E &= \{\text{Alle ord som inngår i T og E, inkludert duplikater fra T}\} \\ I \cap E &= \emptyset \\ RT_{Noekkel}(tekst) &= \frac{1.0 - \frac{|T_I| + |T_E|}{|T|}}{2} \end{aligned} \tag{5.3}$$

$RT_{Noekkel}(tekst)$ gir gjennomsnitt av pluss- og minuspoengene. For plusspoengene er det tatt utgangspunkt i 0.0 poeng, med en økning som tilsvarer forholdet mellom antallet ord i teksten som er i I og det totale antallet ord i T . Tilsvarende er gjort for minuspoengene, men her er de trukket fra 1.0. Den samlede vurdering er gjennomsnittet av disse to verdier. $RT_{Noekkel}(tekst)$ er således utledet som vist i uttrykkene (5.4), (5.5) og (5.6)

$$\frac{\left(1.0 - \frac{|T_I|}{|T|}\right) + \left(0.0 + \frac{|T_I|}{|T|}\right)}{2} \Rightarrow \tag{5.4}$$

$$\frac{\left(1.0 - \frac{|T_I|}{|T|}\right) + \frac{|T_I|}{|T|}}{2} \Rightarrow \tag{5.5}$$

$$\frac{1.0 - \frac{|T_I| + |T_E|}{|T|}}{2} \tag{5.6}$$

5.1.4 Filstier

Noen av tekstene fra de innsamlede ALT-attributtene er filstier [8]. Hvis dette er samme filsti som i SRC-attributten kan dette antas å være noe nettsidens forfatter har lagt inn bare for å ha noe i ALT-attributten og kan derfor antas å være støy.

Hvis det er en annen filsti er det allikevel antakelig at den ikke sier så mye for den jevne bruker, men muligheten for at den gjør det øker. Hvis ALT-attributten inneholder en annen filsti enn den i SRC-attributten kan det være aktuelt å se om bildet er et ikon som brukes til å identifisere filtypen filstien i ALT peker mot. På bakgrunn av dette er det formelle uttrykket for rangeringen av en tekst mot denne tekstklassen gitt i uttrykk (5.7).

$$RT_{Filsti}(tekst) = \begin{cases} 1.0 & : \text{hvis tekst ikke er en filsti} \\ 0.5 & : \text{hvis tekst er en filsti og tekst er ulik SRC} \\ 0.0 & : \text{hvis tekst er en filsti og tekst er lik SRC} \end{cases} \quad (5.7)$$

Et eksempel på filtypeikoner som ofte inngår på nettsider er RSS-feeds for publisering av artikkellister [48]. Hvis ALT-attributten inneholder en filsti til en RSS-fil, bildet er et RSS-ikon og IMG-markøren ligger inne i en ANCHOR-markør med lenke til samme RSS-fil som i ALT-attributten kan det være interessant for skjermleseren å inkludere denne ALT-attributten.

5.1.5 Repeterte ALT-attributter på samme side

Det hender at samme tekst forekommer i to eller flere ALT-attributter på en nettside. Det kan være forskjellige grunner til dette. En mulighet er at det forteller om egenskaper ved bildet, for eksempel at det er en transparent gif. Dette kan tyde på at bildet har med grafisk utforming [49] og ikke informasjonsformidling. En annen mulighet er at bildet leveres fra en ekstern tjeneste, som for eksempel en reklameleverandør. SRC-attributten kan da inneholde en lenke til en nettressurs som leverer forskjellige reklamebilder over tid. I begge tilfeller kan opplesning av teksten i ALT-attributten for bildet virke som støy.

Som rangering av disse tekstene er det derfor benyttet forholdet mellom antall ganger teksten forekommer og det totale antallet IMG-markører på siden, som vist i uttrykk (5.8).

$$RT_{Repetert}(tekst) = \frac{\text{Antall ganger tekst opptrer i en ALT-attributt på nettsiden}}{\text{Totalt antall IMG-markører på nettsiden}} \quad (5.8)$$

Selv om dette teoretisk gir en mulighet for en rangering i området 0.0 til 1.0 antas det at forholdet i uttrykk (5.8) mer ofte enn ikke vil ligge ned mot 0.0. Nettsider hvor alle ALT-attributter er den samme antas å være sjelden, untatt i det tilfellet hvor de fleste eller alle ALT-attributter på nettsiden gir $RT_{\emptyset}(tekst) = 0.0$.

5.1.6 Bare et tall

Det forekommer at bildeteksten er et tall, uten benevning som kilo, meter eller antall. Et slikt enkeltstående tall kan antas å virke som støy, fordi det ikke sier noen om hva

det betyr. Den formelle definisjonen av denne rangeringsfunksjonen er dermed gitt ved uttrykk (5.9).

$$RT_{Tall}(tekst) = \begin{cases} 1.0 & : \text{ hvis tekst ikke er et tall} \\ 0.0 & : \text{ hvis tekst er et tall} \end{cases} \quad (5.9)$$

Dersom det senere gjennom bildeanalysen skulle vise seg at tallet har en fremtredende plass i bildet kan det allikevel være interessant å inkludere det i opplesningen. Et eksempel på dette er hvis det som er avbildet er siden på en terning med samme valør som tallet. Da kan det også være mulig å generere en bildetekst slik at det som leses opp blir *Terningkast 5*, selv om ALT-attributten bare inneholdt 5.

5.2 Tester av tekstanalyser

De algoritmer som er skissert her for å rangere bildetekster innenfor de gitte tekstklasser er grunnleggende, men viktige. Å utvikle en algoritme som bestemmer om en tekst er et tall eller en filsti er ikke det sentrale, da dette ofte kan gjøres med standardfunksjoner i biblioteker til mange programmeringsspråk, slik som i `Java.lang.String`-objektet i Java ¹.

Hovedmålet med testene her er derfor å undersøke hvorvidt de foreslåtte tekstklassene faktisk er relevante for forskjellige kategorier av nettsider. Dette må sjekkes ved innsamling av empiriske data, og det er derfor plukket ut seks kategorier av nettsted. Disse er *Nettaviser*, *Bank og finans*, *Nettbutikker*, *Høyere utdanning*, *Statlige institusjoner* og *Andre nettsteder*.

Disse nettstedene er valgt fordi de antas å være mye i bruk og det er også tatt med forskjellige kategorier innenfor både offentlige og private nettsteder. At det er valgt ut kun norske nettsteder skyldes utelukkende tilgang til ordlister. Gitt tilgang på de rette ordlister antas det ikke å være noen forskjell mellom forskjellige språk.

Resultatene og diskusjon av de data som er samlet inn ifølge med denne oppgaven finnes i kapittel 11. En samlet oversikt over kategorier, hvilke nettsteder som er med i hvilken kategori, og hvor mange ALT-attributter som faller inn i hver tekstklasse for det enkelte nettsted er listet i tabell 11.1 i kapittel 11.

¹<http://java.sun.com/j2se/1.5.0/docs/api/java/lang/String.html>

Del III

Bildemodalitet

Kapittel 6

Farger og bilder

Å analysere innholdet av et bilde ved hjelp av en datamaskin er et komplisert tema. En datamaskin virker ved å utføre eksakt spesifiserte programmer på data som ikke forandrer struktur. Datamaskinens evne til å reagere på uventede hendelser er derfor begrenset til det som er forutsett av mennesket som skrev programmet. Å kjenne igjen mønstre i data som er *nesten* likt det mønsteret det søkes etter er derfor noe en datamaskin normalt ikke håndterer godt. Mennesker har derimot denne evnen, basert på blant annet erfaring, intuisjon og fantasi. Men denne evnen er ikke nøyaktig forstått og er derfor ikke lett å kopiere. Å bygge en applikasjon som kan imitere denne egenskapen er derfor en krevende oppgave [50] [40].

Det er imidlertid mulig å lage regler som klassifiserer bilder basert på målbare egenskaper i gitte situasjoner. Å rangere bilder etter betydning under den antakelse at bildets viktighet øker proporsjonalt med hvor godt bildet fanger menneskets blikk er et slikt scenario. Egenskaper som kan benyttes til å vurdere et bildes betydning som blikkfang er bildets størrelse, fargenes kontrast mot hverandre, hvor kompliserte former det inneholder og hvor det er plassert på siden.

6.1 Bilder på nettsider

Grafiske elementer, som fotografier eller tegninger, er en meget viktig del av informasjonsformidlingen på nettsider. Men de har også en annen viktig oppgave, som er å gi et godt inntrykk av nettsiden gjennom estetisk pen layout. Det antas i denne oppgaven at layout ikke har betydning for brukerens oppfattelse av en nettside dersom brukeren ikke kan se siden, men får den formidlet via metoder som talesyntese eller leseliste. Disse bildene kan derfor regnes som støy på siden og kan ignoreres.

HTML og CSS gir muligheter for å endre et bildes størrelse i forhold til den oppløsning som er gitt i grafikkfilen [51] [49]. For å utlede et regelsett for klassifisering av bilder på en nettside er det derfor nødvendig å ta utgangspunkt i HTML og hvordan HTML behandler bilder på en nettside. Blant de muligheter som HTML gir for å inkludere grafikk på en nettside er i denne oppgaven IMG-markøren valgt ut for

nærmere undersøkelse.

IMG-markøren definerer et område i nettsiden hvor bildet skal plasseres. Nettleseren laster så ned bildefilen separat og viser bildet i det avholdte området. Med CSS kan dette området flyttes til andre steder på nettsiden enn det IMG-markørens plassering i HTML-koden alene skulle tilsi.

Med CSS er det også mulig å endre bildets størrelse. For et bilde som er lagret med en oppløsning på 300x500 piksler er det mulig å spesifisere et område på 150x1000 piksler det skal fylle. Nettleseren beregner så hvilke piksler som må fjernes og/eller legges til for å fylle dette området. Dette får betydning for hvordan bildet blir seende ut på nettsiden og klassifiseringen må ta hensyn til dette.

6.2 Digital representasjon av bilder

I et digitalt bilde av et motiv kan det ikke lagres en representasjon av alle farger, til det er det alt for mange mulige nyanser. Spekteret av synlig lys er en kontinuerlig serie med uendelig mange farger. Å representere alle disse fargene i en datamaskin ville derfor være altfor krevende i både tid og rom (prosessorkraft og minne).

Dette er heller ikke nødvendig da menneskets øyne ikke kan se forskjell på alle disse nyansene. Det er derfor mulig å diskretisere fargespekteret og se på intervaller med lysfrekvenser som ligger nært inntil hverandre som samme farge. Disse fargene kan så representeres på forskjellige måter i datamaskinen, basert på hvilke egenskaper ved fargen som er hensiktsmessig for bruksområdet, for eksempel på trykk eller på skjerm. Denne representasjonen kalles en fargemodell og er omtalt i detalj senere i kapittelet.

Foruten fargemodellen er måten bildets struktur lagres på viktig. De mest brukte formatene å lagre et bilde er som raster- eller vektorgrafikk.

6.2.1 Rastergrafikk

Et bilde kan representeres i form av et todimensjonal rutenett av fargepunkter som kalles *piksler*. Denne måten å representere et bilde på kalles *rastergrafikk* [52] [53]. For hvert piksel i rutenettet lagres en farge og antallet piksler som lagres kalles bildets *oppløsning*.

Detaljnivået som kan lagres er avhengig av oppløsningen. Et digitalt kamera vil ikke kunne lagre alle farger (lysfrekvenser) som reflekteres fra motivet. Det kan bare lagre en gjennomsnittsverdi for hvert piksel i sin oppløsning. Små gjenstander som er plassert nære hverandre kan derfor smelte sammen til et fargefelt dersom oppløsningen blir mindre.

Det er lagres normalt ikke nok data til å beregne hvordan bildet ville sett ut dersom det hadde hatt høyere oppløsning. Det som er mulig er å gjøre bildet større ved å la hvert piksel fra bildet få et større areal ved at det dekker flere piksler av skjermens oppløsning. Dette fører etter en hvis størrelse til at det blir synlig at bildet er bygget

6.3. FARGEMODELLER

opp av ensfargede enkeltfirkanter, som igjen gir et mer utydelig bilde. Denne effekten kalles *pikselisering*.

Fordelen med rastergrafikk er at den kan lagre bilder hvor det er mange farger og former, for eksempel fotografier. Ulempen er at dette krever mer lagringsplass enn for bilder med enklere sammensetning, som logoer. Tegnede bilder lagres derfor ofte som vektorgrafikk.

6.2.2 Vektorgrafikk

I stedet for å lagre alle fargepunkter i et rutenett bruker vektorgrafikk matematiske likninger som representasjoner av geometriske figurer som punkter, linjer, kurver og polygoner [54]. Dette gjør at en firkant kan representeres ved hjelp av koordinater for plassering, firkantens lengde og bredde, farge og bredde på rammen og fargen på området innenfor. Dette gir to fordeler ovenfor rastergrafikk.

For det første er det mulig å representere en firkant som er så stor som man ønsker ved samme forbruk av minne/disk. For eksempel kan en blå firkant på 300x300 piksler lagres i vektorgrafikk som tre heltall for lengde, bredde og fargekoden for blå. Sett i forhold til rastergrafikk, hvor det hadde blitt en tabell med 300x300 like celler hvor alle inneholder fargekoden for blå, er dette en vesentlig innsparing.

For det andre er det mulig å skalere opp størrelsen på samme firkant ved å multiplisere opp lengde og/eller bredde med så mye som det er behov for og fremdeles få en firkant med samme farge over hele området, uten pikselisering.

Dette gjør at vektorgrafikk er godt egnet til tegninger, men også at den ikke er godt egnet til å representere fotografier. Dette fordi et fotografi kan inneholde så små og mange figurer at det ikke gir noen innsparing. Vektor- og rastergrafikk utfyller derfor hverandre [55].

6.2.3 Filformater

Det er mange forskjellige filformater for å lagre bilder og grafikk, men GIF, JPG og PNG er de mest brukte i nettsider i dag. Disse har forskjellige egenskaper blant annet når det gjelder antallet farger og hvor stor kompresjonsgrad de har [51] [49].

Algoritmene som er beskrevet i kapittel 7 og 8 forutsetter at bildene kan omdannes til rastergrafikk og representeres i fargemodellene RGB og HSV. Da en dataskjerm er et rutenett tilsvarende rastergrafikk må dette være mulig for filer i både vektor- og rastergrafikk. Filformater antas derfor å ikke ha noen betydning for algoritmene i påfølgende kapitler.

6.3 Fargemodeller

En *fargemodell* er konstruert for å beskrive en farges egenskaper i et gitt bruksområde, for eksempel på en skjerm eller på trykk. De fargene som en fargemodell beskriver

vil være en delmengde av alle de mulige fargene i det synlige spekteret og kalles modellens *fargerom* [56] [57].

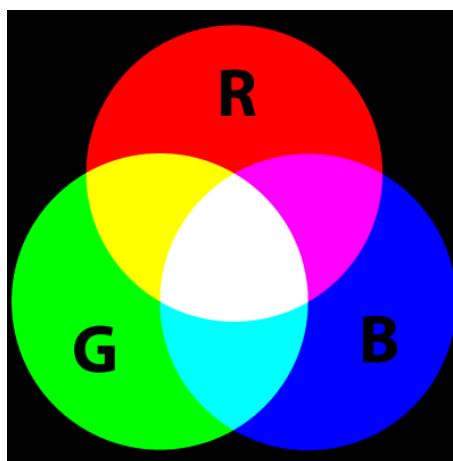
Stråling i frekvensområde 416.4 – 749.5 megahertz i det elektromagnetiske spektrum kalles synlig lys. Hvis det antas at hver frekvens i dette spektrumet utgjør en distinkt farge er det uendelig mange farger. Menneskets øye er imidlertid bare i stand til å se forskjell på cirka 10 000 000 forskjellige fargenyanser. Antallet forskjellige farger det er hensiktsmessig å lagre er enda mindre, da skjermer, film eller fargetykk sjelden er i stand til å representere så mange forskjellige fargenyanser [58].

I denne oppgaven er det benyttet to forskjellige fargemodeller, RGB og HSV. Disse har forskjellige egenskaper som er utnyttet på forskjellig vis i de påfølgende kapitler. Det er derfor valgt å beskrive noen aspekter ved begge modeller, etter først å ha omtalt noen felles begreper.

6.3.1 Primær- og sekundærfarger

Fargene rødt, gult og blått kalles ofte *primærfargene*. Dette er en betegnelse som benyttes i farvelære fordi disse fargene ikke fremkommer ved å blande andre farger. Derimot kan andre farger fremkomme ved å finne riktig forhold mellom to eller tre primærfarger.

Resultatfargen av å blande to primærfarger i forholdet 1 : 1 kalles en *sekundærfarge*. Fargene som fremkommer ved igjen å blande to av av de tre primære og tre sekundærfargene kalles en *tertiærfarge*. Disse primær- og sekundærfargene er illustrert i figur 6.1 som også viser at blanding av alle tre primærfarger gir hvit, mens fravær av all farge gir sort.



Figur 6.1: Primær -og sekundærfarger i RGB-modellen (Kilde: wikipedia.org)

6.3.2 Komplementærfarger, fargesirkelen og kontrast

Fargene som fremkommer ved blanding kan organiseres i et fargehjul som vist i figur 6.2. To fargetoner som er plassert 180 grader fra hverandre på hjulet kalles komplementærfarger. Ved blandingsforhold 1 : 1 vil de parvise komplementærfargene gi en gråbrun farge, de nøytraliserer hverandre.



Figur 6.2: Fargehjul HSV-modellen (Kilde: wikipedia.org)

To farger som ligger rett ved siden av hverandre på fargesirkelen kan være vanskelig å skille fra hverandre. Jo flere farger det er i fargesirkelen, jo mer tilnærmet like er de. To farger som er komplementære er derimot lett å skille fra hverandre, de er ulike. Fra figur 6.2 kan det utledes at rød og grønn er komplementærfarger. Dette er grunnen til at de brukes til henholdsvis stopp og kjør i trafikklys, det er lett å se forskjell på disse. Dette kalles kontrasten mellom to farger. To farger som ligger langt fra hverandre på fargehjulet sies å ha høy kontrast. Kontrast er et viktig virkemiddel som blikkfang og for å se forskjell på elementer i nettsiden som tekst og lenker [59].

6.3.3 Additiv og subtraktiv fargeblanding

En fargemodell kan kombinere primærfargene på to forskjellige måter, additivt eller subtraktivt. I begge tilfeller kalles det totale antall fargetoner som kan oppnås for modellens *fargerom*.

Ved additiv fargeblanding er det summen av primærfargene i forskjellig styrke som danner den ønskede fargetonen. Dersom alle tre primærfarger har full styrke blir resultatet hvit mens fravær av alle gir sort. Additiv fargeblanding brukes i prosjektører og skjermer, hvor lys i de tre primærfargene blandes for å oppnå de forskjellige fargene i en fargeskjerm.

Subtraktiv fargeblanding fungerer ved å fjerne bølgelengder av reflektert lys fra en flate. Dette er derfor mye benyttet i trykte medier. Ved riktig blandingsforhold av pigmenter i sekundærfargene kan det bestemmes hvor mye av primærfargen som skal

absorberes eller reflekteres. Dette kan gjøres fordi hver primærfarge har en sekundærfarge som komplementærfarge, som vist i figur 6.1.

En fargemodell som er mye brukt til trykk kalles *CMYK* etter sekundærfargene cyan, magenta og gul, samt *blacK*. Et sort pigment er inkludert fordi det er vanskelig i praksis å skaffe CMY-pigmenter som ved blanding gir en helt sort farge. Den kalles *blacK* fordi *B* er opptatt i form av *Blått* i primærfargene.

Additiv og subtraktiv fargeblanding er ikke egenskaper ved lyset, men ved hvordan menneskets øye oppfatter farger. Øyet ser ikke forskjell på en lyskilde som sender ut gult lys og en lyskilde som sender ut like kraftig rødt og grønt lys, stråling på henholdsvis én og to bølgelengder. I begge tilfellet vil dette bli oppfattet som gult lys.

Siden et sentralt poeng i denne oppgaven er hvordan bilder oppfattes når de projiseres fra en dataskjerm vil kun additive fargemodeller være omtalt videre. De fargemodellene som er benyttet i oppgaven kalles RGB og HSV. RGB og HSV har forskjellige egenskaper som komplementerer hverandre i algoritmene for bildeanalyse.

6.3.4 RGB fargemodellen

RGB modellen er en additiv fargemodell basert på å blande primærfargene rødt, grønt, og blått i forskjellige styrker. De tre komponentene av hver farge lagres som regel separat for hvert piksel i et bilde. Dette gjør at et bilde lagret i RGB-modellen lett kan skrives til en dataskjerm.

Siden modellen bygger på blandingen av farger i spektrumet er den godt egnet til å representere farger direkte fra innsamlede data over hvilke lysfrekvenser en gjenstand utstråler, et digitalt fotografi. Det er imidlertid ikke så lett for et menneske å se for seg hvordan en farge som er oppgitt i et styrkeforhold mellom primærfargene ser ut.

6.3.4.1 Representasjon i en datamaskin

En utfordring med RGB-modellen er at det ikke er presist definert hvilken frekvens i det synlige spektrum som utgjør hver av de tre primærfargene. Dette kan derfor variere noe fra produsent til produsent, modellen er hardware-sentrisk. Dette vil si at fargene som skapes må sees relativt til den enkelte produsents tolkning. Dette antas ikke å ha betydning for algoritmene som er utviklet i denne oppgaven.

Den lagrede styrken, et tall mellom 0 og 255, av hver primærfarge kalles en kanal. Med tre fargekanaler gir dette $256 \cdot 256 \cdot 256 = 16\,777\,216$ forskjellige fargenyanser. Sammenliknet med de cirka 10 000 000 fargenyanser menneskets øye kan se forskjell på gir dette en fargemodell med et tilstrekkelig fargerom for de fleste formål.

6.3.4.2 Alfakanal

I mange tilfeller er det ønskelig at bakgrunnen i forskjellig grad skal kunne syntes gjennom et bilde. RGB modellen kan derfor utvides med en alfakanal og kalles da

6.4. HISTOGRAM OG BILDER

ARGB [60]. Alfakanalen forandrer ikke fargen, men bestemmer hvor mye av bildets bakgrunn som er synlig.

Dette kan være aktuelt for bilder som er plassert oppå hverandre med CSS [45]. Dersom fargene i det øverste bilde har en lav styrke på alfakanalen er det mulig å se bildet som ligger bak. Dette kan sammenliknes med å se på noe gjennom farget glass. Som fargekanalene oppgis også alfakanalen i en verdi mellom 0 og 256 hvor 0 er gjennomsiktig og 256 er ugjennomsiktig.

6.3.5 HSV fargemodellen

HSV representerer farger på en annen måte enn som en kombinasjon av primærfarger. HSV står for *Hue*, *Saturation* og *Value* og modellen organiserer farger med disse verdiene som henholdsvis vinkel, radius og høyde i sylinderkoordinater. Dette representerer fargene på en måte som er en mer presis tilnærming til måten mennesker oppfatter farger på. Den gjør det også enklere å beregne forhold mellom to farger, som hvor stor kontrast det er mellom dem.

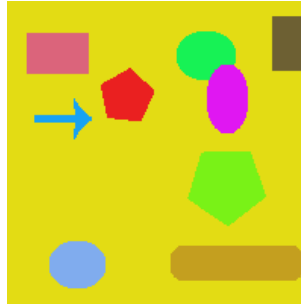
Hue er selve fargetonen, den egenskap ved en farge som beskrives med ord som rød eller blå. Fargetonen i HSV er beskrevet som et punkt langs fargesirkelen. Dette gjør at det er enkelt å beregne hvor stor kontrast det er mellom to farger ved å måle hvor stor vinkelen mellom disse er. Komplementærfarger ligger således 180 grader fra hverandre. Flere algoritmer i den påfølgende diskusjon er basert på å beregne kontrast.

Saturation betyr metning. Dette er en betegnelse på intensiteten i fargetonen. En farge med stor metning har en klart definert farge, mens ved minkende metning får den en mer grålig tone. Metning kan sees på som hvor mye hvitt pigment som er blandet inn i fargen. Metning avbildes i modellen som verdien langs sylinderens radius. Fargen er helt hvit i sentrum og går over mot å blir mer skarpt definert som den aktuelle fargetone jo mer den fjerner seg fra sentrum.

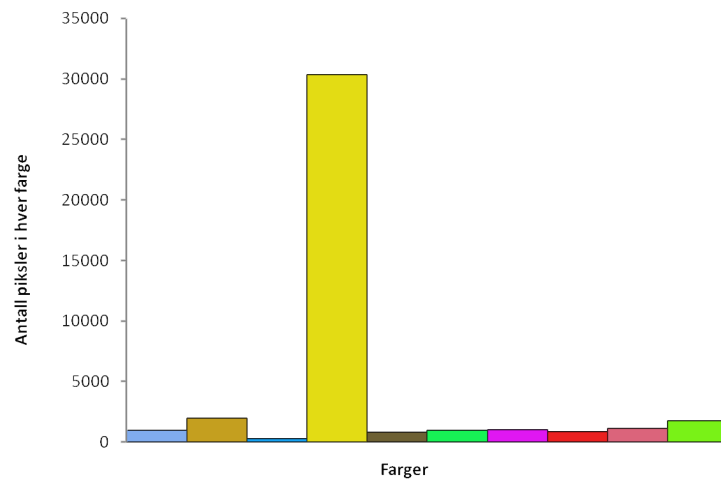
Value er fargens lyshetsgrad, hvor fargen ligger på akse mellom hvitt og svart. Ved høy lyshetsgrad er fargen den rene *Hue*, men den beveger seg mer og mer mot svart med minkende *Value*. Siden alle farger går over i sort i bunnen av sylindren vises denne ofte også som en omvendt kjegle.

6.4 Histogram og bilder

Et histogram er en gruppering av et datasetts elementer i diskrete kategorier. I figur 6.4 vises et eksempel på dette i form av et histogram som viser hvor mange piksler det er for hver farge i bildet i figur 6.3. Ved å studere fordelingen av farger over pikslene i et bilde er mulig å trekke slutninger om bildet. Legg merke til den store søylen med gule piksler i figur 6.4. Siden den gule søylen er så markant mye større enn de andre kan dette tolkes til å være bildets bakgrunnsfarge, noe som bekreftes ved å studere selve bildet i figur 6.3.



Figur 6.3: Farger og figurer



Figur 6.4: Pikkshistogram av farger fra figur 6.3

6.4.1 Formell definisjon

Gitt et datasett M med n elementer er et histogram et sett funksjoner m_i som bestemmer hvilke elementer fra M som tilhører gruppe i . Alle elementer i datasettet må tilhøre en gruppe. Gitt at det er k grupper kreves det at likheten i uttrykk (6.1) oppfylles.

$$\sum_{i=0}^k m_i = n \quad (6.1)$$

Intervallbredden på hver gruppe er avhengig av anvendelsesområdet. For eksempel kan hver lysfrekvens være en farge, selv om mennesket ikke kan se forskjell på den og farger i et frekvensintervall rundt. Dette frekvensintervallet kan da utgjøre en gruppe. Når intervallbredden er bestemt er det å finne antallet grupper redusert til likningen i uttrykk (6.2) hvor k er antallet grupper, h er bredden på intervallet og M er datasettet som grupperes. M_{max} og M_{min} er da henholdsvis den største og den minste verdien i datasettet M .

$$k = \left\lceil \frac{M_{max} - M_{min}}{h} \right\rceil \quad (6.2)$$

6.4.2 Normalisering av histogrammer

Det kan i noen tilfeller være interessant å sammenlikne to datasett for å se om forholdet mellom antallet elementer i hver gruppe og det totale antallet elementer i datasettene er det samme. Dette gjøres ved å dele antallet i hver gruppe med det totale antallet i datasettet slik at summen av gruppene er lik 1, som vist i uttrykk (6.3). Dette kalles *normalisering* og i stedet for å vise antallet elementer i hver gruppe viser det normaliserte histogrammet sannsynligheten for at et element forekommer i datasettet.

$$\frac{1}{n} \sum_{i=0}^k m_i = 1.0 \quad (6.3)$$

Et scenario hvor dette er aktuelt er når bilder er gitt en annen størrelse på en nettside enn det bildefilens oppløsning tilsier. Når et bilde er skalert opp eller ned vil allikevel forholdet mellom antallet piksler i hver farge og det totale antallet piksler være tilnærmet det samme. Ved å sammenlikne normaliserte histogrammer vil det derfor være mulig å finne like bilder selv om de er i forskjellig størrelse.

6.4.3 Like histogrammer, like datasett

I algoritmen beskrevet i kapittel 7 er det nødvendig å på enkelt vis sammenlikne to bilder for finne ut om de er identiske. Dette vil basere seg på hypotesen om at datasett med samme histogram er samme datasett.

Identisk like er her definert til å være at histogrammene har samme antall grupper, at gruppene har samme bredde og samme gruppe har samme antall elementer i begge histogrammer. Hvis disse betingelser er oppfylt vil histogrammene A og B være like hvis summen i uttrykk (6.4) er 0.

$$\sum_{i=0}^k (a_i - b_i) = 0.0 \quad (6.4)$$

I tilfellet hvor bildene som skal sammenliknes ikke har samme skala er det imidlertid ikke alltid at histogrammene vil være helt identiske, selv om de er normalisert.

6.4.3.1 Tilnærmet like datasett

Algoritmer som benyttes til å skalere opp bilder innebærer ikke bare at det legges til flere piksler i samme farge. I grenseområdet mellom regioner av forskjellige farger legges det også til piksler med farger som er beregnet basert på pikslene rundt. Dette innebærer at også histogrammene blir noe forskjellige for forskjellige skaleringer av samme bilde.

Ved å anta at selv bilder med meget like motiver er tilstrekkelig forskjellige, og at det derfor er tilstrekkelig at histogrammene er tilnærmet like, kan to histogrammer ansees som tilnærmet like hvis ϵ verdien i uttrykk (6.5) er tilstrekkelig liten.

$$\sum_{i=0}^k |a_i - b_i| \leq \epsilon \quad (6.5)$$

I uttrykk (6.5) blir k antallet grupper i unionen av de to histogrammene. Dersom en gruppe mangler i et av histogramene er antallet piksler i denne gruppen 0 for det histogrammet. Verdien ϵ avhenger av hvor like bildene på nettsiden det letes i er. Dersom det er veldig mange bilder av tilsvarende motiver, for eksempel personer mot samme bakgrunn, vil det kreves en mindre verdi. En mindre verdi gir imidlertid økt sjanse for å forkaste bilder som skulle vært inkludert.

Kapittel 7

Posisjon for delbilde

Et bildes betydning som blikkfang på en nettside er ikke bare avhengig av dets egne egenskaper. Bildet må blant annet også sees i forhold til hvor på siden det ligger og fargespillet mellom bildet og bakgrunnen på siden i området rundt bildet. Som første steg i bildeanalysene diskutert i kapittel 8 er det derfor nødvendig å kjenne til et gitt bildes eksakte posisjon på nettsiden. Dette kapittelet skissere en metode for dette.

HTML [7] og CSS [45] er detaljert definerte industristandarder, men de er ikke implementert likt i alle nettlesere. Figur 7.1 inneholder HTML/CSS-kode for en nettside med teksten «Sentret, rød boks?» inne i en boks med rød og stiplet ramme. Boksen har en bredde på 120 piksler. Figur 7.2 viser hvordan koden fra figur 7.1 vises i Internet Explorer 7.0 og figur 7.3 hvordan den vises i Firefox 3.5.

Sentralt i eksemplet er de to CSS-reglene «margin-left: auto;» og «margin-right: auto;». Etter standarden skal de sørge for at det tilgjengelige området inne i markøren som omslutter DIV-markøren fordeles likt på begge sider av DIV-markøren, i form av like stor marg på begge sider. I dette eksemplet er det BODY-markøren, og boksen skal derfor plasseres midt på toppen av siden.

Dette eksemplet viser at selv meget enkle nettsider kan tolkes tilstrekkelig forskjellig fra nettleser til nettleser til at elementer ikke ligger der man forventer at de skal ligge dersom man kun foretar en tekstbasert analyse av strukturen i et HTML-dokument, med tilhørende CSS.

For å unngå å inkludere spesifikk kode for bestemte nettlesere i implementasjonen er det derfor antatt i denne oppgaven at det er bedre å la den aktuelle nettleseren rendere nettsiden slik dens implementasjon tilsier, for så å ta et skjermbilde av denne rendringen. De enkelte bilder sine eksakte posisjoner i den rendrede nettsiden kan så lokaliseres ut i fra dette skjermbildet.

7.1 Rå kraft

Som navnet sier er dette en algoritme basert på å finne delbildet ved bruk av rå prosessorkraft. I denne algoritmen ble det tatt utgangspunkt i metoder hentet fra [61] for

```
<html>
  <head>
    <title>Rød boks</title>
    <style type="text/css">
      div
      {
        width: 120;
        margin-left: auto;
        margin-right: auto;
        border: 1px dotted red;
      }
    </style>
  </head>
  <body>
    <div>
      Sentrert, rød boks?
    </div>
  </body>
</html>
```

Figur 7.1: Meget enkel nettside, bare en rød boks

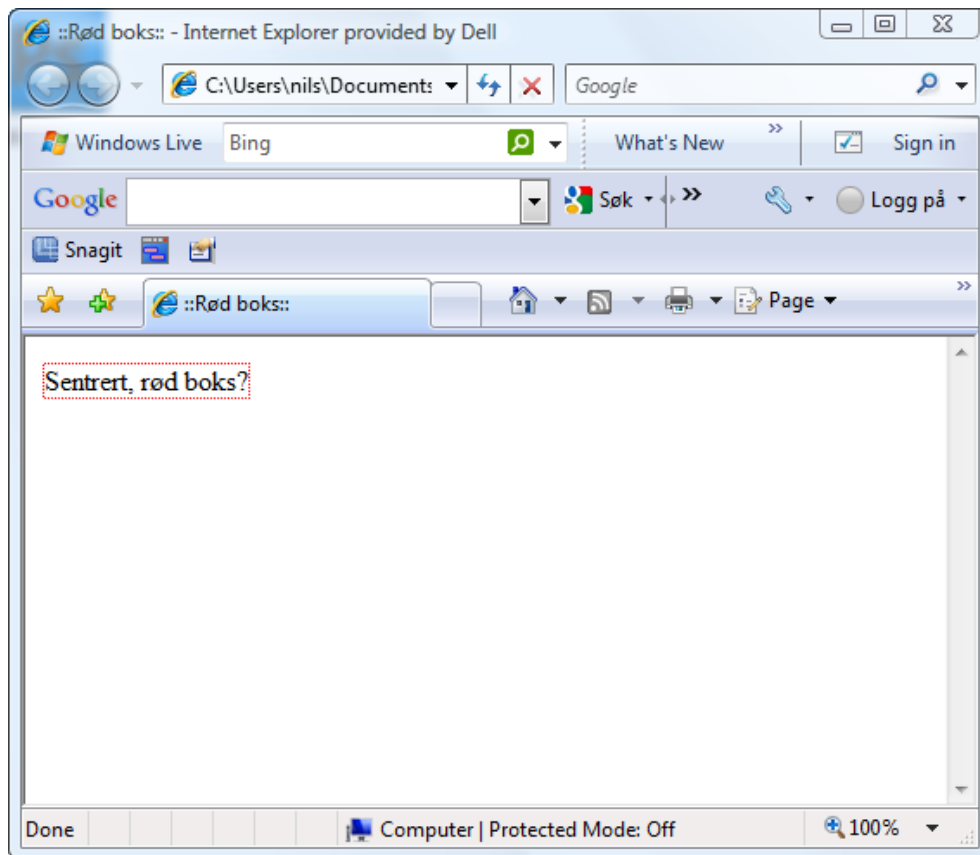
søk etter mønster i tekster. Algoritmen som ble testet er basert på den naive algoritmen som er skissert i [61] for å finne en deltekst med lengde $m \leq n$ i en tekst med lengde n . Dette gjøres ved å se på teksten, og mønsteret man søker, som lister av symboler/bokstaver.

Ved å sammenlikne mønsterets første bokstav med tekstens første bokstav, mønsterets andre med tekstens andre og så videre til man enten kommer til mønsterets siste bokstav, og har funnet en forekomst av mønsteret i teksten, eller til første bokstav hvor mønsteret og teksten ikke samsvarer. Dersom det finnes en bokstav hvor det ikke er samsvar mellom mønsteret og teksten, flyttes mønsteret ett hakk lenger fram i teksten og starter sammenlikningen på nytt. Dersom mønsteret flyttes så langt fram i teksten at det ikke lenger er nok bokstaver til å gjøre en ny sammenlikning med mønstere, er det konstatert at mønsteret ikke finnes i teksten. Et eksempel på denne algoritmen kan sees i figur 7.4.

Hypotesen var at fargen på hver piksel kunne overta rollen til bokstaver, skjerm-bildet blir teksten det søkes i og bildet det søkes etter er mønsteret. Søket ble da i to dimensjoner, slik at når enden på en linje er nådd hopper man tilbake til begynnelsen av neste linje.

Algoritmen antas derfor, selv om den er meget nøyaktig, å være for lite effektiv til å kunne brukes i en sanntidsanalyse som det her er snakk om. Mer effektive algoritmer for søk etter mønster i tekster er Morris-Pratt algoritmen, som også er beskrevet i [61].

7.2. PIKSELPLUKK-ALGORITMEN



Figur 7.2: HTML-kode for figur 7.1, rendret av Internet Explorer 7.0

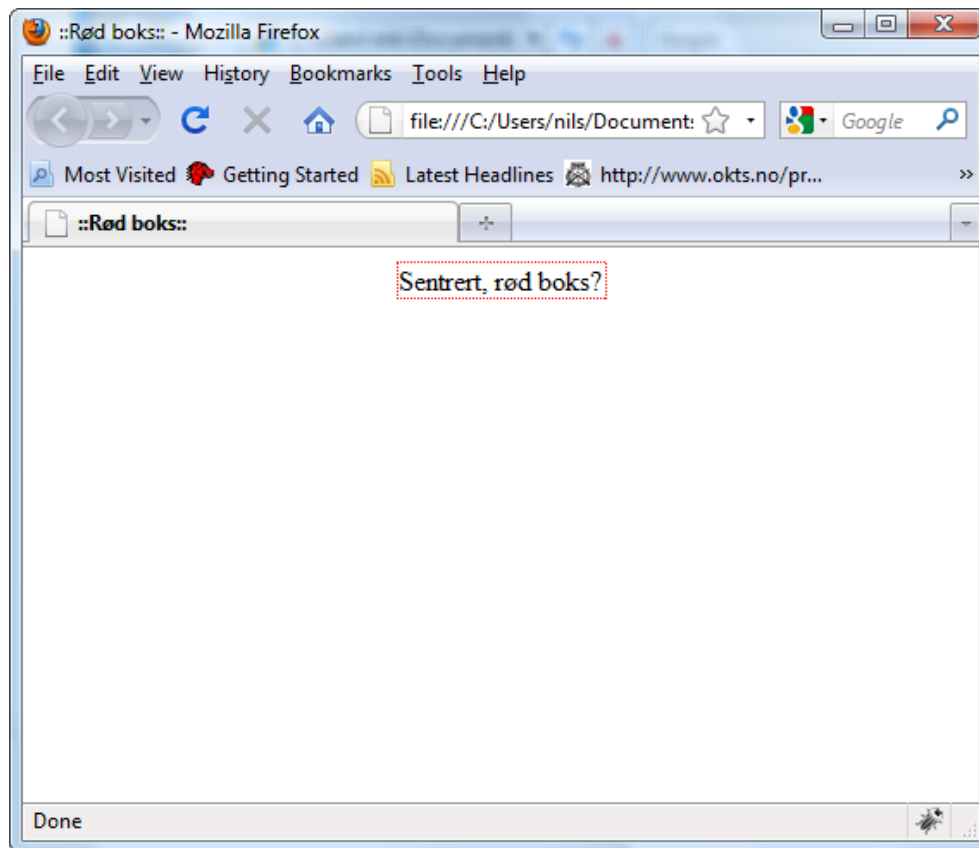
Denne har imidlertid, selv om den er meget effektiv, og kunne vært generalisert til å matche et to dimensjonalt mønster, et problem i forhold til denne oppgavens formål.

Akkurat som den naive algoritmen beskrevet ovenfor er den basert på å finne deltekster som er eksakt like med det mønsteret de søker etter. I denne oppgaven er algoritmene imidlertid også avhengige av å kunne finne et bilde på en nettside selv om bildet er skalert opp eller ned ved hjelp av HTML og CSS.

I et skalert bilde antas det at forholde mellom antallet piksler i hver fargenyans og det totale antall piksler i bildet er tilnærmet konstant. Som beskrevet i kapittel 6 burde det derfor være mulig å undersøke om to bilder egentlig er samme bilde i forskjellig oppløsning.

7.2 Pikselplukk-algoritmen

Utifra kravet om tilstrekkelig ytelse til å kunne brukes i sanntidssystemer ble det klart at en algoritme basert på sekvensiell gjennom søkning av samtlige elementer i store todimensjonale tabeller ikke var akseptabelt. Uten å komme med noen uttømmende argumenter på hva som er minste og største størrelse på bilder på en nettside vises det



Figur 7.3: HTML-kode for figur 7.1, rendret av Firefox 3.5

til at et bilde på 200x300 piksler, et ikke spesielt stort bilde, allerede gir en tabell på 60 000 piksler.

Et alternativ til en sekvensiell gjennomgang av hele tabellen er å gjøre et tilfeldig utplukk av punkter, fordelt over hele nettsiden, for deretter å undersøke om et av punktene ligger i det bildet man søker etter.

7.2.1 Antakelser

Da ingen tilsvarende eller liknende algoritmer er funnet i litteraturen, har det vært nødvendig å gjøre noen antakelser for hvordan en slik algoritme bør være konstruert og hva den bør ta hensyn til. Dette er basert på erfaring med, og på eksisterende standarder og retningslinjer [49] [6] [5] for utvikling av nettsider. Disse antakelser er som følger:

- At hovedbildet, et skjermbilde av en nettside, har en distinkt bakgrunnsfarge.
- At denne bakgrunnsfargen er den fargen som forekommer oftest på nettsiden, inkludert forekomster i bilder og annen grafikk.

7.2. PIKSELPLUKK-ALGORITMEN

Tekst: "Dette er en test."

Mønster: "test"

Søk 1: Ikke match

```
[D][e][t][t][e][ ][e][r][ ][e][n][ ][t][e][s][t][.]  
[t][e][s][t]
```

Søk 2: Ikke match

```
[D][e][t][t][e][ ][e][r][ ][e][n][ ][t][e][s][t][.]  
[t][e][s][t]
```

Søk 3: Ikke match

```
[D][e][t][t][e][ ][e][r][ ][e][n][ ][t][e][s][t][.]  
[t][e][s][t]
```

.
. .
.

Søk 13: Match

```
[D][e][t][t][e][ ][e][r][ ][e][n][ ][t][e][s][t][.]  
[t][e][s][t]
```

Figur 7.4: Naiv metode for tekstsøk

- At den fargen som forekommer nest oftest er fargen som hoveddelen av teksten på nettsiden har.
- At delbildet det søkes etter er rektangulært, i betydningen at fargene i bildets ytterkant danner en definert ramme som har andre fargetoner enn nettsidens bakgrunnsfarge.
- Dersom det plukkes ut et tilstrekkelig antall tilfeldige piksler fra bildet er det rimelig å anta at minst en av disse ligger i det delbildet det søkes etter. Piksler med samme farge som bakgrunnsfargen forkastes.
- At det fra det utplukkede punktet i bildet kan søkes frem til rammen som dannes mot bakgrunnsfargen og dermed finne posisjonen til bildet på nettsiden.
- At bilder som har samme histogram er det samme bildet.
- At bilder som er skalert i forhold til bildefilens oppløsning kan sammenliknes ved at histogrammene normaliseres.

Dette vil i korthet si at de delbilder det søkes etter har bakgrunnsfargen i hovedbildet som en markant ramme, og at det er denne rammen det søkes etter. Dette vil gi et sett med delbilder, kandidatbildene. Det bildet det søkes etter kan deretter finnes ved å sammenlike bildets histogram med histogrammene til hvert kandidatbilde.

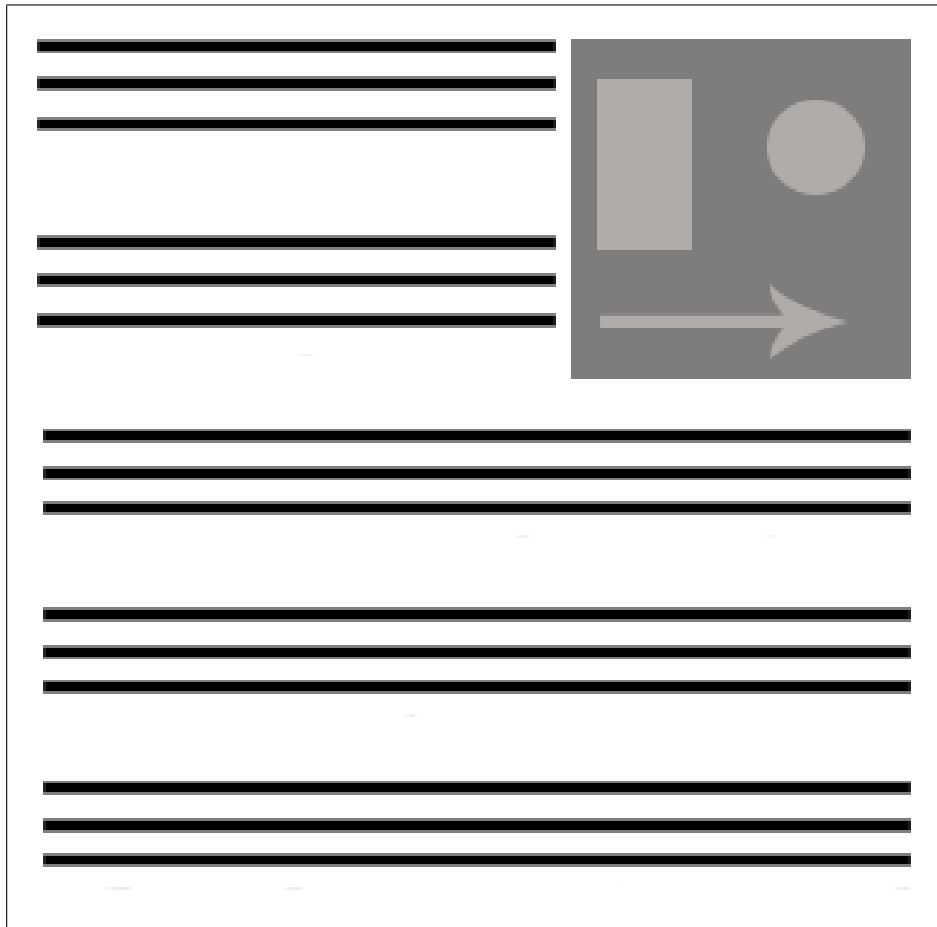
7.2.2 Overordnet skisse

Grunnlaget for algoritmen er ønsket om å finne et rektangulært delbilde inne i et hovedbilde. Dette hovedbilde kan for eksempel være et skjermbilde av en nettside. Bildet ansees som funnet dersom histogrammet av bildet som det søkes etter er tilnærmet likt histogrammet for en region (delbilde) av skjermbildet (hovedbilde). Algoritmen går i korte trekk ut på:

1. Generer pikselhistogram for hovedbilde og bildet det søkes etter.
2. Velg den fargen som forekommer oftest i histogrammet for hovedbilde som antatt bakgrunnsfarge for hovedbilde. Tilsvarende velges den fargen som forekommer nest mest som antatt farge på testen på nettsiden.
3. Plukk ut et antall tilfeldige punkter i hovedbildet. Punktene må være fordelt utover hele hovedbildet.
4. Fra de punkter som ikke har samme farge som bakgrunns- eller tekstfargen søkes det utover i alle retninger til en piksel med bakgrunns- eller tekstfargen finnes. Disse fire punkter antas å ligge på delbildets ramme.
5. Dersom delbilder fremkommet fra et av de andre pikslene ligger innenfor det nye delbildets rammer forkastes de eksisterende delbildene.
6. Dersom det nye delbildet ligger innenfor et delbilde fremkommet fra en av de andre pikslene forkastes det nye delbildet.
7. Generer histogram for det nye delbildet og sammenlikning med histogrammet for det bildet det søkes etter.
8. Det delbildet som har et histogram nærmest det bildet det søkes etter ansees å være bildet det søkes etter.

7.2.3 Et eksempel

Betrakt den forenklede nettsiden med bilde i figur 7.5. I og for seg nokså intetsigende, men den skulle være tilstrekkelig for å illustrere prinsippene bak algoritmen; en nettside med bakgrunnsfarge, tekstfarge og et rektangulært bilde med tilstrekkelig kontrast mot denne bakgrunnsfargen, slik at det dannes en ramme rundt delbildet.



Figur 7.5: Forenklet skisse av nettside med bilde

7.2.3.1 Beregne innledende histogrammer

Første steg i algoritmen er å beregne histogrammer for skjermbildet av hele nettsiden, hovedbildet, og av det bildet det søkes etter. Dette gjøres som beskrevet i avsnitt 6.4. Histogrammene for hovedbilde og bildet det søkes etter i dette eksemplet finnes i henholdsvis figur 7.7 og 7.8. Legg merke til at histogrammene i figurene ikke er normalisert. Dette er for lettere å kunne vise søylene fra det lille bildet i det store. For å ta hensyn til eventuelle opp- eller nedskalerte bilder i nettsiden normaliseres alle histogrammene.

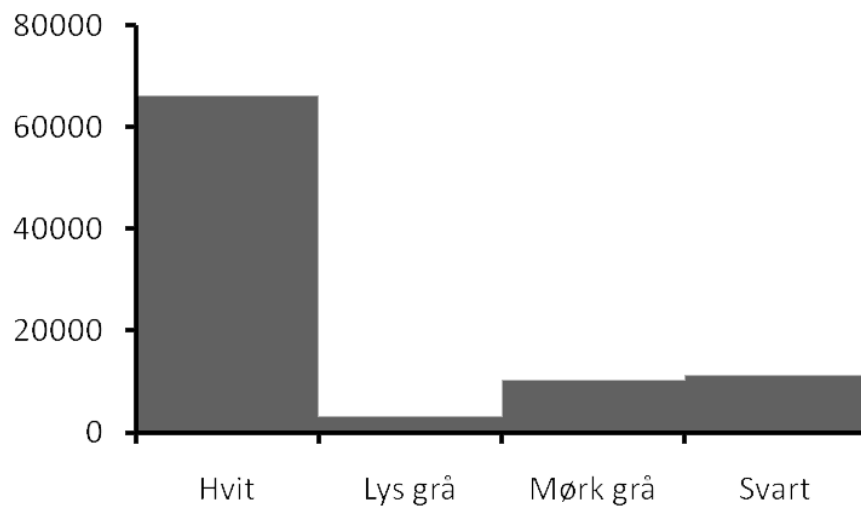
Algoritmen vet at bildet finnes på nettsiden fordi det er funnet en IMG-markør som henviser til bildet. Dermed er det neste steget å identifisere det rektangulære delbildet av skjermbildet av nettsiden som utgjør bildet det søkes etter.

Histogrammet for bildet er nødvendig for å kunne sammenlikne det med de regionene av skjermbildet som søkes frem som kandidatbilder. Kandidatbildene vil være alle rektangulære regioner i hovedbildet hvis ramme utgjøres av hovedbildets bakgrunnsfarge. Histogrammet for skjermbildet brukes til å finne de antatte bakgrunns-



Figur 7.6: Bildet det skal søkes etter

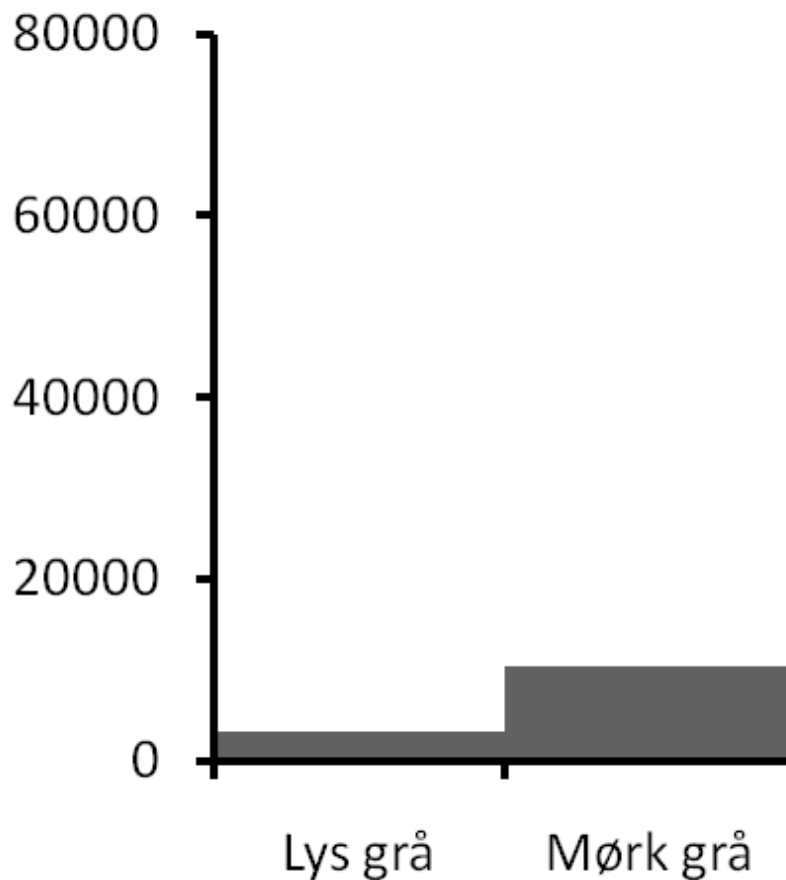
og tekstfarger for nettsiden.



Figur 7.7: Histogram av farger på nettsiden i Figur 7.5

7.2.3.2 Velge nettsidens bakgrunnsfarge

Den fargen som forekommer oftest i histogrammet for hovedbildet antas å være bakgrunnsfargen. Den fargen som forekommer nest oftest antas å være fargen på teksten. Bakgrunnsfargen brukes til å definere rammen rundt de delbilder som utgjør kandidatene til å være det bildet det søkes etter. Dette vil i praksis ikke bare si alle bilder på nettsiden, men alle rektangulære regioner på nettsiden som har en ramme bestående av piksler i samme farge som bakgrunnsfargen.

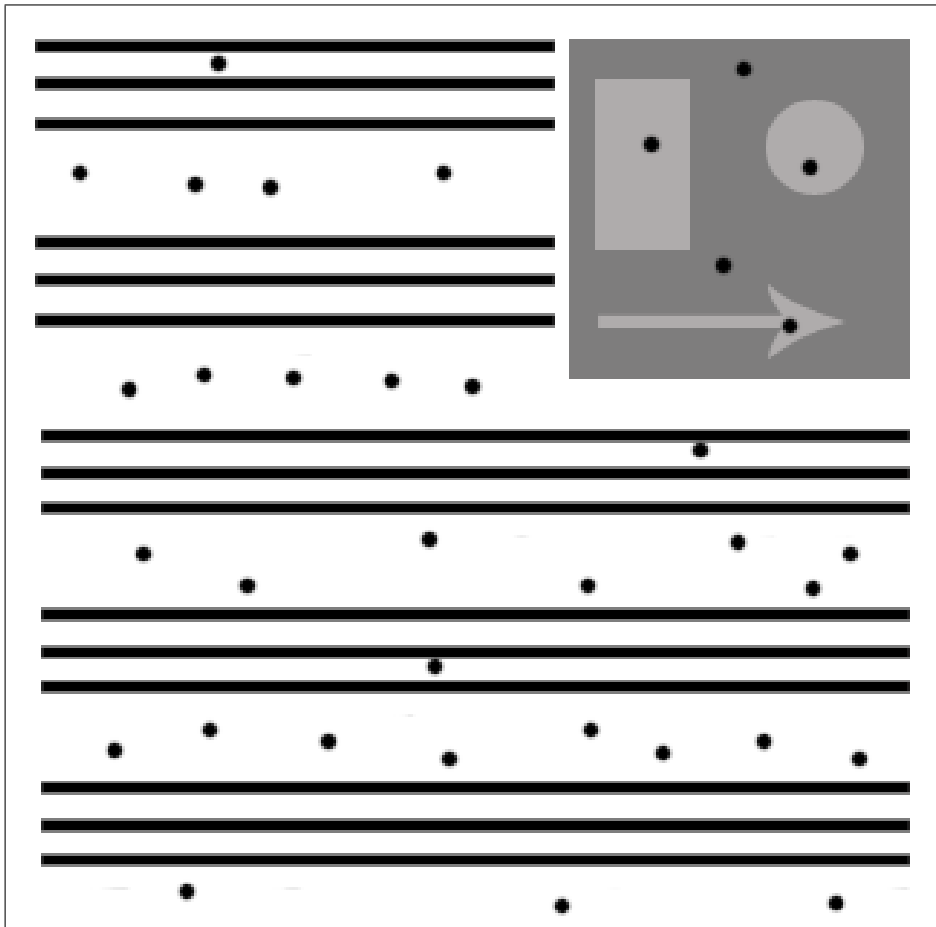


Figur 7.8: Histogram farger i bildet i Figur 7.6

Dersom boksen fra figur 7.1 gis en egen bakgrunnsfarge, som skiller seg fra nettsidens bakgrunnsfarge, vil den kunne være blant kandidatbildene. Det vil da være opp til sammenlikningen av dennes histogram med histogrammet til det bildet som søkes å avgjøre at denne boksen ikke er bildet det søkes etter.

7.2.3.3 Tilfeldig utplukk av piksler

Utgangspunktet for å finne kandidatbilder er en sekvens av tilfeldig valgte punkter jevnt distribuert utover hele hovedbildet. Uten for mye å foregripe gjennomgangen av testresultatene senere i oppgaven, viser disse at det er tilstrekkelig å plukke ut 0.05% av pikslene i hovedbildet for å finne ønsket delbilde i over 80% av tilfellene. Figur 7.9 anskueliggjør disse punktene, noen i bakgrunnen og noen i delbildet. Punkter som har samme farge som antatt bakgrunnsfarge forkastes, selv om de ligger i et kandidatbilde. Dette ansees ikke som noe problem, da det må antas at det plukkes ut mer enn et piksel i hvert bilde på nettsiden.



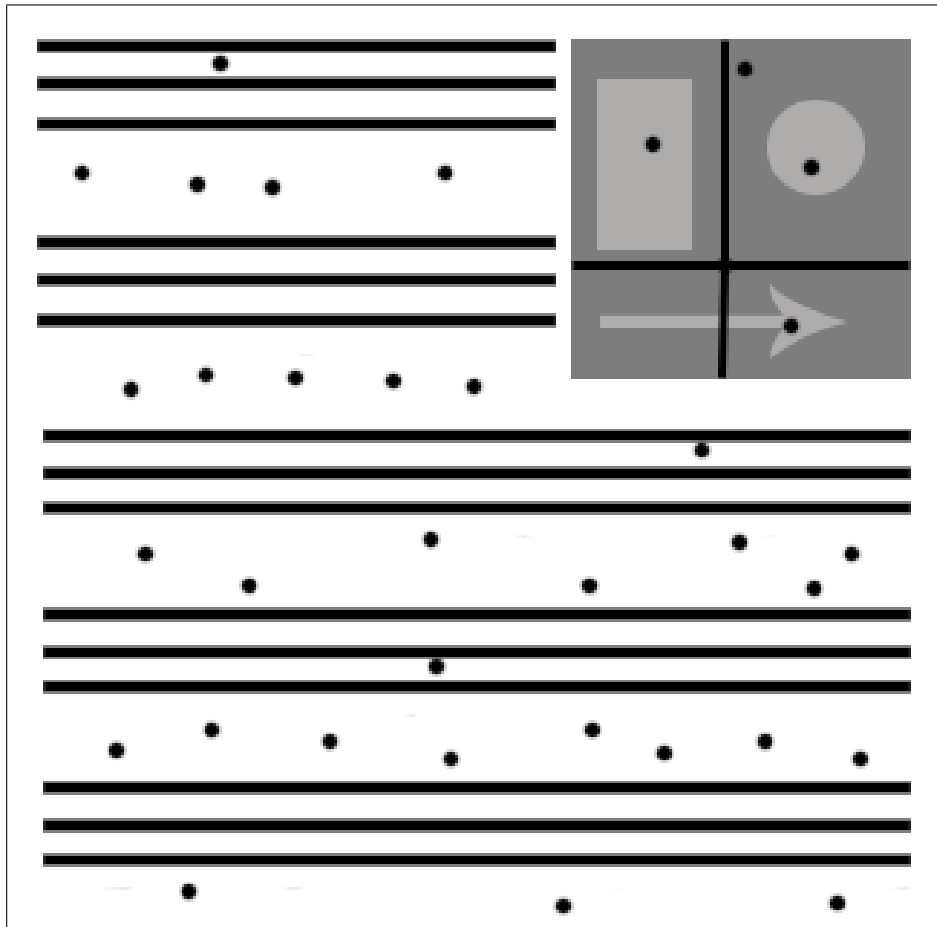
Figur 7.9: Utplukk av tilfeldige punkter, jevnt fordelt

7.2.3.4 Finn kandidatbilder

For de utvalgte punkter søkes det utover i alle fire retninger til første piksel med nettsidens bakgrunnsfarge i hver retning finnes. Dette steget i algoritmen er skissert ved krysset i bildet i figur 7.10. Disse punktene ansees så for å ligge på rammen bakgrunnen danner rundt bildet. Selv om dette også er en gjennomgang av delbildet piksel for piksel, er det allikevel meget færre piksler som undersøkes per kandidatbilde, enn ved den råkraft-algoritme som ble skissert tidligere i kapitlet.

7.2.3.5 Sammenlikne histogrammer

Avsnitt 6.4 beskrev en metode for å avgjøre om to bilder var like basert på sammenlikning av histogrammer for bildenes farger. Rektangelet som utspennes av krysset i figur 7.10 er et kandidatbilde. For å avgjøre om dette kandidatbildet er bildet det søkes etter beregnes histogrammet for dette delbildet. Dette histogrammet sammenlikens så med histogrammet for det bildet det søkes etter. Histogrammene normaliseres for å



Figur 7.10: Søk ut fra punktene, finn kanten

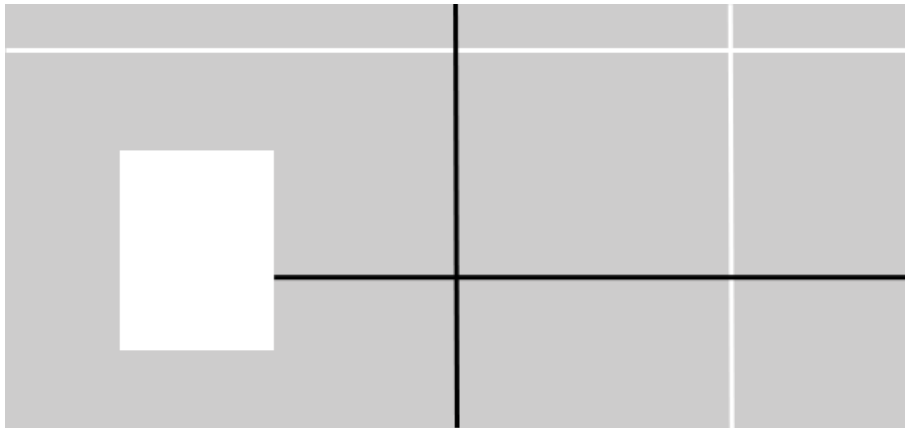
kompensere for bilder som er skalert opp eller ned.

Dersom det er flere kandidatbilder med samme histogram må det sjekkes i HTML-koden om det samme bildet er brukt flere steder på samme nettside. Dette kan være problematisk, da samme bilde kan skjule seg bak flere forskjellige filstier. Det antas derfor i denne oppgaven at hvis en nettutvikler bruker det samme bildet flere steder på en nettside benytter vedkommende også samme fil på tjeneren og dermed samme filsti.

Figur 7.7 og 7.8 viser histogrammer av henholdsvis nettsiden i figur 7.5 og bildet i figur 7.5. I histogrammet av nettsiden kan histogrammet fra bildet kjønes igjen på de to minste søylene, mørk grå og lyst grå. Den største søylen, hvit, er bakgrunnsfargen og den nest største, sort, er tekstfargen.

7.2.4 Spesialtilfeller

Algoritmen er sårbar for enkelte spesialtilfeller. Blant disse er de tilfeller hvor bildene inneholder regioner, av forskjellige form, med samme farge som nettsidens bakgrunnsfarge. Algoritmens evne til å finne bildets eksakte posisjon avhenger da av regionens



Figur 7.11: Spesialtilfelle 1

form. De minst alvorlige formene medfører bare at algoritmen bruker mer tid, fordi flere piksler må sjekkes, mens andre medfører at bildet ikke kan lokaliseres. Disse spesialtilfellene kan deles i to grupper, de som deler bildet i to eller flere regioner og de som ikke gjøre det.

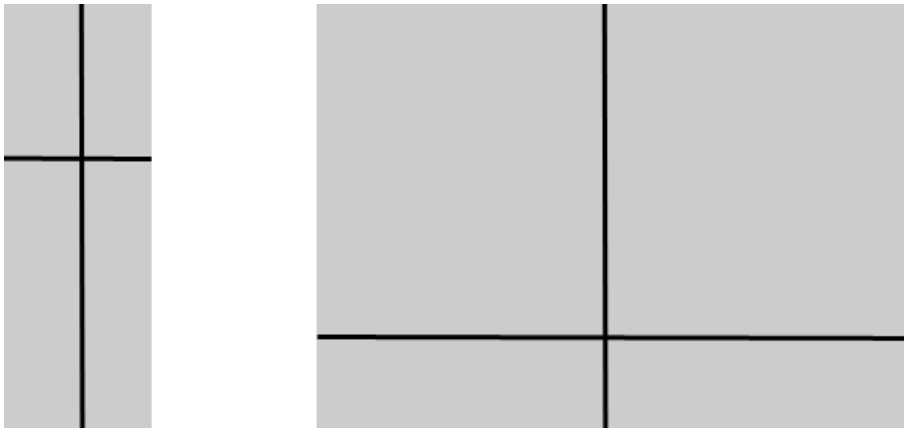
For å unngå at bilder forkastes ubegrunnet fordi de inneholder piksler med samme farge som bakgrunnsfargen velges alltid den største regionen dersom to punkter faller innenfor overlappende regioner. Figur 7.11 gir et eksempel på dette hvor det finnes to kandidatbilder, de som spennes ut med henholdsvis sort og hvit strek. Den minste boksen, utspent av de sorte linjene er i sin helhet inneholdt i boksen utspent av de hvite linjene og den forkastes derfor som kandidatbilde. Boksen som spennes ut av de hvite linjene har ingen omsluttende boks og er derfor ansett som et kandidatbilde.

7.2.4.1 Bildet deles ikke i to deler

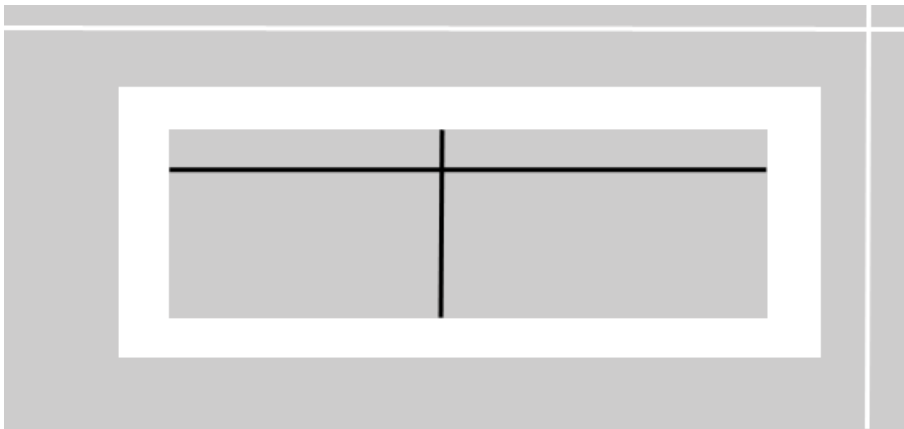
Figur 7.11 viser et tilfelle hvor et bilde inneholder en region hvor fargen er samme farge som nettsidens bakgrunnsfarge. Det søkes ut ifra det første utplukkede pikselet finner ikke det korrekte delbildet fordi de treffer denne regionen og antar at dette er bildets venstre ramme. Dette er illustrert ved de sorte søkelinjene. Neste utplukte piksler har imidlertid fri linje til rett ramme i alle retninger, som vist med de hvite søkelinjene. Kandidatbildet som er utspent av de sorte søkelinjene ligger da i sin helhet innenfor kandidatbildet utspent av de hvite. Det sorte kandidatbildet kan derfor forkastes. Hvis mange bilder forkastes på et slikt grunnlag kan dette løses ved å øke antallet utplukkede piksler.

7.2.4.2 Bildet deles i to deler

Dersom en region i et bilde har samme farge som nettsidens bakgrunn deler hele bildet i to deler, slik som i figur 7.12, vil algoritmen ikke fungere. En mulighet for å løse dette er å søke etter den eller de delene som er funnet ved å benytte algoritmen med



Figur 7.12: Spesialtilfelle 2



Figur 7.13: Spesialtilfelle 3

det bilde det opprinnelig søkes etter som hovedbilde og søke etter kandidatbildene i dette.

Bildet det søkes etter tar da plassen til skjermbildet av nettsiden og kandidatbildene overtar plassen som bildet det søkes etter. Dersom kandidatbildet kan identifiseres til å være en del av det bildet det opprinnelig søkes etter har man funnet dette bildet.

Figur 7.13 gir et eksempel på et tilfelle hvor bildet er delt i to regioner av en region med samme farge som nettsidens bakgrunn, men som likevel gir en mulighet for å lokalisere bildet. Dette siden den indre regionen er helt omsluttet av den ytre region. Den ytre regionen ble funnet av de hvite søkelinjene.

7.3 Testing av algoritmen

En algoritme av denne typen kan ikke lett bevises ved et matematisk bevis. I stedet er det valgt å teste algoritmen i en serie testscenarier. I alle testene er ratio av korrekt

lokaliserte bilder den avhengige variable. De uavhengige variabler har vært følgende:

- Utplukksratio for piksler.
- Antallet bilder på samme nettside.
- Skalering samme bilde opp i en dimensjon.
- Skalering samme bilde opp i to dimensjoner.
- Skalering samme bilde ned i en dimensjon.
- Skalering samme bilde ned i to dimensjoner.
- Areal som dekker delbilde.

Dette kapitlet beskriver hva hvert testscenario skal teste. Resultatene er samlet i kapittel 11, hvor også detaljer som antall bilder i hver test er omtalt. Det påstås ikke at dette testsettet dekker alle eventualiteter, men de antas å dekke tilfeller som er viktige i forhold til hvordan bilder brukes på nettsider.

Testscenarioene er autogeneratede HTML-filer, med tilhørende skjermbilder. Hver HTML-fil inneholder et eller flere bilder, avhengig av typen test. Filstien til hvert av bildene er lagret i en oversiktsfil, det samme er hvor de er plassert på nettsiden. Denne oversiktsfilen brukes til å verifisere om bildene er korrekt lokaliserte eller ikke. Oversiktsfilen sendes så til et testprogram som kjører algoritmen på alle testene og produserer en resultatfil med x og y koordinater algoritmen har funnet for hvert bilde. Grafer hvor disse er plottet finnes i kapittel 11.

7.3.1 Utplukksratio

Dette testscenarioet er for å undersøke om algoritmen fungerer under optimale forhold. Testsettet er en serie med skjermbilder av nettsider med ett tilfeldig plassert bilde på hver side. Bildene er uten spesielle fargesammensetninger, skaleringer eller overlappende bilder. Kan ikke algoritmen finne et enslig bilde på en hvit bakgrunn er det lite sannsynlig at den kan brukes til noe nyttig.

Videre skal denne testen brukes til å finne hvor mange piksler som er nødvendig får å kunne lokalisere bildene nøyaktig. Testsettet er derfor gjentatt med økende antall utplukkede piksler for å finne punktet hvor antallet bilder som finnes stabiliseres. Dette tallet er deretter nyttet i påfølgende testscenarioer.

7.3.2 Antallet bilder på nettsiden

Dette testscenarioet undersøker algoritmens oppførsel ved økende antall bilder på en nettside. For hver gang antallet bilder økes med én er det testet med flere nettsider med dette antallet bilder på hver nettside. Hensikten med denne testen er å gå et skritt videre fra et bilde på hver nettside og undersøke om algoritmen er i stand til å skille et økende antall bilder fra hverandre.

7.3.3 Skaleringer

Skalering av bilder benyttes ofte på nettsider. Dette gjelder særlig for elementer som for eksempel logoer i vektorgrafikk. I disse tilfellene kan man da benytte samme bildefil og så skalere denne opp og ned med høyde og bredde attributtene til IMG-markøren [7] eller med CSS [45] etter behov. Fire forskjellige skaleringer vil derfor bli testet, disse er en- og todimensjonal opp- og nedskalering.

7.3.3.1 Oppskalering - endimensjonal

Denne testserien skal se på om algoritmen håndterer bilder som er skalert opp i en dimensjon, for eksempel i lengderetningen. Det antas at hvilken dimensjon som skaleres ikke utgjør noen forskjell på algoritmens evne til å lokalisere et bilde. I alle skaleringstestene er det samme bildet som skaleres gradvis opp og ned.

Endimensjonal oppskalering kan være aktuelt i tilfeller der et bilde benyttes til å endre layout ved å *dytte* andre elementer på nettsiden i en bestemt retning. Disse består da gjerne av ensfargede 1x1 piksel bilder, som skaleres opp ved behov.

7.3.3.2 Nedskalering - endimensjonal

Bilder i nettsider kan skaleres ned på samme måte som de skaleres opp. Det er bedre å skalere ned et stort bilde enn å skalere opp et lite. Grunnen til dette er at det er enklere å beskjære piksler fra et stort bilde enn å anta hvilke piksler som må legges til et lite. Dette utnyttes ofte for å lage thumbnail versjoner av bilder i et fotoalbum på nett. Det antas derfor at nedskalering er mer vanlig enn oppskalering, spesielt av fotografier.

Det er derfor valgt å utføre tester på både opp- og nedskalering. Når et bilde skaleres ned er det i HTML/CSS mulig å skalere det helt ned til 0 piksels bredde. Dette er det samme som å skjule bildet helt og likheten for histogrammet nå og histogrammet for originalbildet er minimalt. Det er derfor ønskelig å finne frem til hvor lite bildet kan bli før det har mistet så mange piksler at det ikke kan sammenliknes med originalen.

7.3.3.3 Oppskalering - todimensjonal

Dette testsenarioet utføres som for endimensjonal skalering, men skaleringen foregår proporsjonalt i begge retninger. Oppskalering i to dimensjoner antas være mer aktuelt enn i en dimensjon. Et eksempel på dette er oppskalering av logoer i vektorgrafikk.

7.3.3.4 Nedskalering - todimensjonal

Den siste testen i kategorien skaleringer er todimensjonal nedskalering. På grunn av det store tapet av piksler ved reduksjon av oppløsning i bildet forventes dette å være den testen hvor funnratio faller mest og hurtigst.

Det er også dette testscenarioet som antas å være det viktigste av skaleringsscenarioene. Dette fordi det normaliserte histogrammet til bildet forventes å miste rett forhold mellom antallet piksler i hver farge og det totale antall piksler etterhvert som bildet blir mindre. Se eksemplet med thumbnail-bilder under endimensjonal nedskalering for et typisk tilfelle hvor dette er aktuelt.

7.3.4 Areal som dekker delbilde

I dette testscenarioet blir det evaluert hvor mye av bildet som kan klippes bort før histogrammene blir for ulike til at det kan være mulig å finne rett match ved direkte sammenlikning. Bildet er fremdeles innrammet i en rektangulær boks bestående av nettsidens bakgrunnsfarge.

I figurene 7.14 og 7.15 er det vist et eksempel på HTML-kode og tilhørende skjerm bilde med den rendrede nettsiden som viser tilfellet hvor et bilde er delvis *gjemt* under en DIV-markør ved hjelp av CSS.

Målet er å se om det fremdeles er mulig å lokalisere et bilde redusert på denne måte. Spørsmålet blir om histogrammene fremdeles blir tilstrekkelig like til at det ansees som et treff mot det bildet det søkes. Det antas at forholdene mellom hver farge og det totale antallet piksler ikke er like, slik det var antatt for skaleringer.

7.3. TESTING AV ALGORITMEN

```
<html>
  <head>
    <title>Overlappende bilde</title>
    <style type="text/css">
      #div1
      {
        border: 1px dotted red;
        width: 100px;
        height: 100px;
        position: absolute;
        top: 50px;
        left: 50px;
        background-color: magenta;
        z-index: 1;
      }

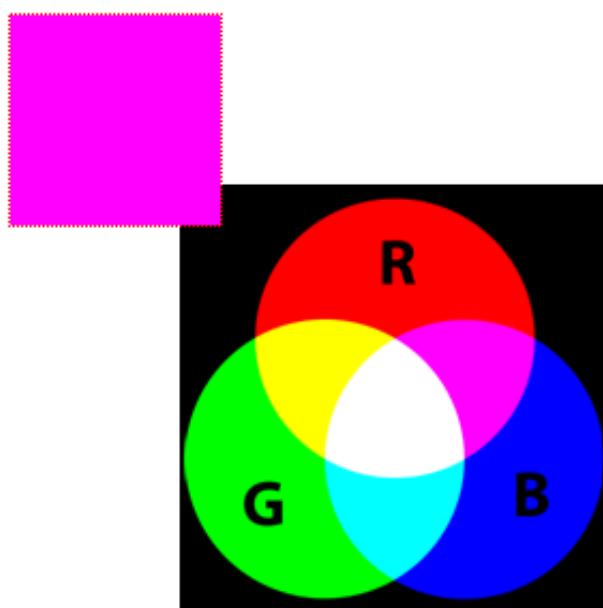
      #img1
      {
        position: absolute;
        top: 130px;
        left: 130px;
        width: 200px;
        height: 200px;
        z-index: 0;
      }
    </style>
  </head>
  <body>

    <div id="div1"></div>

  </body>
</html>
```

Figur 7.14: DIV-markør uten innhold gitt form av et kvadrat og plassert over et bilde ved hjelp av CSS



Figur 7.15: Bilde med overlappende firkant

Kapittel 8

Bildeanalyse

e Dersom en bruker av en nettside er blind eller svaksynt må det tas i bruk andre metoder enn de visuelle for å oppnå en likeverdig brukeropplevelse i forhold til en seende. Gjennom riktig bruk av HTML er det mulig å utvikle nettsteder som oppnår dette målet. Det er utviklet standarder som WCAG [6] [29] [31] for hvordan HTML kan brukes til å gjøre et nettsted mer tilgjengelig for funksjonshemmede brukere, inkludert blinde og svaksynte.

I denne sammenheng er ALT-attributten på IMG-markøren et viktig hjelpemiddel. Som illustrert i figur 3.2 er denne tenkt brukt til å inneholde en alternativ bildetekst som kan vises der bildet ikke kan vises. I figur 3.2 er også TITLE-attributten på ANCHOR-markøren tatt med som ytterligere et eksempel på hvilke muligheter HTML har for å hjelpe funksjonshemmede. I dette tilfelle i form av en beskrivelse av nettsiden lenken i ANCHOR-markøren fører til.

Teksten i ALT-attributten var opprinnelig tenkt brukt når bildet ikke var tilgjengelig. Dette kan være i rent tekstbaserte nettlesere, som for eksempel Lynx ¹, eller i situasjoner med liten båndbredden hvor nettleseren kan vise teksten i ALT-attributten i stedet for å laste ned bildet. Ettersom båndbredde har økt er dette blitt en mindre aktuell problemstilling. Det er imidlertid funnet et nytt bruksområde for ALT-attributten som bildetekst til hjelp for blinde og svaksynte. Hensikten med denne oppgaven er å se på metoder for å identifisere hvilke ALT-attributter som antas å inneholde nyttig informasjon og hvilke som kan sees på som forstyrrende og som derfor bør ignoreres av skjermleseren.

På samme måte som kapittel 5 omtalte analyse av bildetekster skal dette kapitlet beskrive forskjellige bildeklasser for å bestemme en rangering av et sett med bilder. Videre skal det vurderes om disse rangeringene gir grunnlag for å bestemme bildenes nytteverdi som informasjonsbærer på en nettside. Det antas at bildets viktighet er koblet til dets rolle som blikkfang på nettsiden. Det sees derfor på om egenskaper ved bildets innhold, dets posisjon på nettsiden og dets relasjon til omgivelsene kan ha betydning for graden av blikkfang et bilde har på nettsiden.

¹<http://lynx.isc.org>

Følgende bildeklasser er omhandlet i oppgaven, hvorav tre er implementert og testet:

- Bildets innhold og fargesammensetning.
- Bildets størrelse og form.
- Bildets farger i forhold til omgivelsene.
- Bildets posisjon på nettsiden.
- Bildet ansees å ha en bestemt betydning som er alminnelig kjent, for eksempel filikoner.

Listen er på ingen måte fullstendig, men er valgt med tanke på hva som er hovedmomentene i utvalg og plassering av bilder på en nettside [49] [51] [7] [45]. Med bilde menes her enhver grafisk representasjon, uavhengig av om det er fotografier eller tegnet grafikk. I motsetning til for tekstklassene kan hvert bilde rangeres innenfor hver klasse. Animert grafikk, som gif-animasjoner eller flash, er utelatt fra undersøkelsen. Dette er gjort selv om disse på grunn av bevegelse kan være nyttige som blikkfang [62].

Hver analysefunksjon er gitt en formell definisjon. Konvensjonen for funksjonsnavn er $RB_{navn}(Src)$, hvor RB er *rang av bilde*, *navn* er navn på den enkelte analyse og *Src*-parameteren representerer bildet som analyseres. *Src*-parameteren er gitt navn etter SRC-attributten på IMG-markøren, som inneholder filsti til bildet som representeres av IMG-markøren.

8.1 Bildets innhold og fargesammensetning

I prosessen med å vurdere et bilde og dets relevans som informasjonsbærer er det først og fremst viktig å se på egenskaper ved bildet og dets innhold. Selv om en datamaskin vanskelig kan si eksakt hva et gitt bilde forestiller, kan den allikevel trekke ut en del opplysninger som kan brukes til å gjøre antakelser om det er sannsynlig at det bildet vil være et blikkfang for mennesker.

Eksempler på dette vil være antallet forskjellige farger i bildet og deres kontrast mot hverandre, samt antall og kompleksitet av forskjellige former og figurer. Det er imidlertid ikke gitt at det mest komplekse bildet er det mest interessante i konteksten. For eksempel vil et fargefotografi være mer komplisert og detaljert enn en graf i sort/hvit, men det er ikke dermed sikkert at fotografiet er mer betydningsfullt enn grafen i relasjon til teksten på nettsiden. Det er totalinntrykket av tekst- og bildeanalyser som avgjør om bildet skal inkluderes i skjermleserens opplesning av nettsiden.

8.1.1 Komprimering

Å måle hvor *komplisert* et bilde er, i form av antall farger og former, kan være et vanskelig spørsmål. En mulig måte å måle visuell kompleksitet på er ved å studere hvor mye et bilde kan komprimeres. Dette er fordi bildekompresjon i praksis vil si at overflødig informasjon fjernes. Den informasjon som er tatt bort kan (tilnærmet) beregnes utifra den resterende delen ved dekomprimering, slik at det opprinnelige bildet kan gjenstapes. Jo mer variasjon det er i farge og form innenfor et bilde, jo mindre kan det komprimeres uten å minste så mye data at det gjenskapte bildet ikke likner på originalen. Dette innebærer at et bilde der få farger dekker store flater vil kunne komprimeres mer enn et bilde med mange små fargeflater.

Det blir dermed mulig å rangere bilder etter effektiviteten av komprimeringen, definert som forholdet mellom størrelsen på originalbildet og på den komprimerte versjonen. Et bilde som er meget enkelt, i betydningen at det har store områder med samme farge, gir en høy komprimeringsgrad. Mer komplekse strukturer og fargesammensetninger gir lav komprimeringsgrad. Det er dermed mulig å rangere bildet etter antatt kompleksitet i intervallet $[0.0, 1.0]$ hvor et bilde som komprimeres til 0.0 er så enkelt at det ikke antas å gi noen informasjon, mens et bilde 1.0 er så fargesprakende at det må antas at det legges merke til. I praksis vil verdiene i de fleste tilfeller ligge mellom disse, da 0.0 vil innebære at det opprinnelige bildet ikke inneholdt noen piksler og 1.0 vil bety at hvert piksel i bildet har forskjellig farge fra de andre.

8.1.1.1 Singulærverdidekomposisjon (SVD)

Som komprimeringsalgoritme er det her valgt en algoritme for faktorisering av matriser som kalles singulærverdidekomposisjon (SVD). Dette er en metode innen lineær algebra [63] med mange nyttige bruksområder, da inkludert bildekompimering.

Fordi det med SVD er lett å beregne komprimeringsgrad er denne valgt over andre mer konvensjonelle komprimeringsteknikker for bilder. Imidlertid vil enhver algoritme hvor graden av komprimering kan bestemmes kunne brukes. En mer detaljert beskrivelse av den underliggende matematikken er beskrevet i [64] og [65].

Det som er tilstrekkelig i denne sammenheng er at av de tre matrisene som bildet faktoreriseres til vil en av disse være en diagonal matrise i samme dimensjon som bildet. Verdiene langs diagonalen vil være synkende og i et antall av de siste kolonnene vil de være 0.0. Ratio mellom antall kolonner i bildet og antallet kolonner i singulærverdimatrisen hvor verdien er ulik 0.0 vil da indikere hvor mye bildet er komprimert. Uttrykk (8.1) gir da en rangering av bildet i intervallet 0.0 og 1.0.

$$RB_{Komprimering}(Src) = \frac{\text{Antall singulærverdier}}{\text{Antall kolonner i bildet}} \quad (8.1)$$

8.1.2 Mønstergjennkjennelse

Mønstergjennkjennelse går ut på å analysere rådata, i dette tilfelle en matrise med fargekodene fra et bilde, og gjenkjenne gitte mønster i disse. Et eksempel på dette kan være å finne områder i bildet som inneholder et ansikt, en bil eller et tall.

I kapittel 5 ble det nevnt et tilfellet hvor en ALT-attributt kun inneholdt et tall uten benevnelse. I dette tilfellet ble det skissert et scenario hvor tallet allikevel kunne være interessant for en skjermleser å inkludere hvis tallet i en eller annen form inngår i bildet. Som eksempel på dette kan det antas at det er mer interessant for en blind eller bruker svaksynt å få lest opp *Terningkast 5* enn bare 5 i det tilfelle hvor ALT-attributten kun inneholder tallet 5, men hvor mønstergjennkjennelse kan si at bildet forestiller 5-siden på en terning.

Et annet eksempel på bruksområde for mønstergjennkjennelse i denne sammenheng er for å identifisere bilder med betydning som ansees kjent for store brukergrupper som for eksempel emotikons og filtypeikoner [66].

8.1.2.1 Bilder med kjent mening

Framveksten av operativsystemer som Windows og Macintosh med grafiske brukergrensensitt har ført til en stor mengde ikoner som representerer blant annet programmer og filtyper assosiert med disse [66].

Internett, for eksempel gjennom chat-programmer for direkte kommunikasjon mellom brukere [67], har medført enda flere ikoner. I slike chat-programmer er det ikke lett å lese samtalepartnerens kroppsspråk og stemme. Det har som følge av dette utviklet seg en kultur for å kommunisere ved hjelp av ikoner som formidler blant annet følelser, såkalte emotikons.

Denne studiens interesse for disse er, uansett betydning, deres enkle utforming og store utbredelse. Ved hjelp av mønstergjennkjennende algoritmer kan det være mulig å avgjøre om et bilde forestiller et ikon og eventuelt hvilket ikon det er snakk om. Foruten å avgjøre om bildet skal leses opp eller ikke, kan det derfor også informeres om hvilket ikon det er snakk om, også dersom bildet ikke har oppgitt noen ALT-attributt.

8.1.3 Stor indre kontrast

Bilder som inneholder farger med stor kontrast til hverandre kan antas å være et større blikkfang enn bilder med en fargesammensetning som gir mindre kontrast. En algoritme for å undersøke dette kan bestå av å beregne gjennomsnittet av den parvise kontrasten mellom de 10% av fargene i bildet som forekommer oftest. En algoritme for å beregne kontrast av to farger, gjennom å måle avstand mellom fargene på en farge-sirkel, er skissert senere i dette kapitlet.

Bildene som er størst blikkfang er ikke nødvendigvis de som formidler mest informasjon ut i fra mottakers synspunkt. Det er imidlertid rimelig å anta at de er av betydning for den som har publisert nettsiden. For eksempel kan det antas at det er

viktig for en nettavis at reklamebannerene på deres side er gode blikkfang, for å sikre reklameinntektene. For leseren som forsøker å få med seg de siste nyhetene er det derimot ikke sikkert dette alltid er like nyttig [68].

8.2 Bildets størrelse og form

Kan størrelsen på et bilde være en god indikasjon på hvor relevant det er? Et stort bilde kan antas å være et større blikkfang enn et mindre. Foregående avsnitt tok for seg analyse av bildets innhold i form av farger og struktur. I dette avsnittet sees det på hvor viktig bildets størrelse og form kan antas å være for bildets rolle som blikkfang.

Det kan for eksempel tenkes at et bilde som er meget lite kan antas å være en del av layout, heller enn en illustrasjon som gir bedre forståelse av informasjonen som nettsiden ønsker å formidle. Meget lite betyr i dette tilfelle at bildets størrelse nærmer seg en bredde på én piksel i én eller begge dimensjoner.

I de nettsider som ble studert under arbeidet med denne oppgaven inneholdt ALT-attributten i noen tilfeller tekster på formen «line.gif» eller «dot.gif», altså en filsti eller et filnavn. Et bilde som har liten utstrekning i én eller begge dimensjoner, og en bildetekst som er en filsti, kan derfor være et bilde det er ønskelig at en skjermleser utelater. Det det same kan også antas dersom ALT-attributten bare inneholder ord som «line», «pil» eller «punkt».

8.2.1 Etter størrelse

Et større bilde kan antas å tiltrekke mer oppmerksomhet enn et mindre. Det er derfor grunn til å tro at en rangering etter størrelse kan gi en indikasjon på bildets betydning som blikkfang. Algoritmen som er foreslått rangerer det største, i areal, bildet på nettsiden som 1.0 og deretter de resterende bildene i forhold til dette etter uttrykket i (8.2).

$$RB_{Stoerrelse}(Src) = \frac{\text{Areal av bildet som rangeres}}{\text{Areal av det største bilde på siden}} \quad (8.2)$$

8.2.2 Størrelsen går mot 1x1 piksler

Dersom et bilde er veldig lite, og kanskje bare består av én eller to farger, kan det være en indikasjon på at bildets funksjon på nettsiden har med estetisk design å gjøre og ikke informasjonsformidling.

Et eksempel på dette kan være et bilde som gir et avrundet hjørne i rammen på en tabell. Tabeller i HTML kan gis rammer med CSS, men de er da kun firkantet. En teknikk som brukes for å gi disse pene, avrundede rammer er å bruke bilder som bare inneholder rette linjer på sidene og så bilde med bare den runde kanten i hjørnene.

8.2.3 Misforhold mellom dimensjoner

Et tilfelle hvor bilder benytte på nettsider er som rammer på tabeller. I dette tilfellet vil bilder som er lange og tynne utgjøre sidene på tabellen. Disse bildene vil være bare noen få piksler brede, med en lengde som tilsvarer lengden på tabellen noe som kan være flere hundre piksler. Bilder som har en slik form kan være en god kandidat for skjermleseren å utelate.

8.3 Bildets plassering på nettsiden

I denne gruppen av rangeringsalgoritmer ser ikke på egenskaper ved selve bildet, men på betydningen av bildets plassering på nettsiden. Et bildets plassering på en nettside kan si ting om dets antatte betydning for å forsterke forståelsen av nettsidens innhold.

8.3.1 Naturlig lesesti

Bilder som kommer tidlig i en tekst kan være plassert der nettopp i den hensikt å vekke interesse og motivere til å lese teksten ferdig. Rekkefølgen bildene er plassert i den retning brukeren leser dokumentet kan derfor antas å ha betydning for hvor viktige de er på nettsiden.

Hva som er naturlig lesesti for den enkelte bruker er en kulturelt betinget sak. Mens brukere fra den vestlige verden leser dokumenter linje for linje ovenfra og fra venstre, leses tekster på arabisk eller hebraisk fra høyre. Dette antas imidlertid ikke å spille noen rolle så lenge det er en fast rekkefølge på hvordan man leser en tekst. Det kan derfor være interessant å se på om det har betydning for bildets betydning om blikkfang hvor det er plassert i denne lesestien.

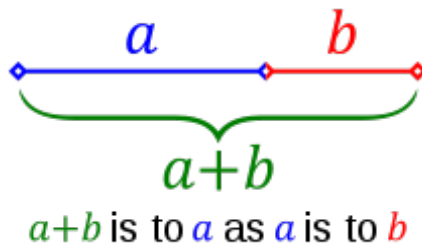
8.3.2 Det gyldne snitt

Det punktet på en linje som deler linjen i to deler slik at forholdet mellom linjen og den lengste delen er den samme som forholdet mellom de to delene kalles det gyldne snitt [69]. Figur 8.1 gir en grafisk, og uttrykk (8.3) en matematisk, definisjon av det gyldne snitt.

$$\frac{a+b}{a} = \frac{a}{b} = \phi = \frac{1+\sqrt{5}}{2} \approx 1.6180339887 \quad (8.3)$$

Det gyldne snitt opptrer i mange tilfeller i geometrien og i naturen, og da særlig i menneskekroppen. Eksempler på det siste er forholdet mellom lengden fra skulder til fingertuppen og fra albuen til fingertuppene. Kunstneren og matematikeren Leonardo Da Vinci var av den mening at på en perfekt kropp måtte en rekke forhold på denne tilsvare det gyldne snitt, noe hans statuer og malerier bærer preg av ².

²<http://www.maths.surrey.ac.uk/hosted-sites/R.Knott/Fibonacci/fibInArt.html>



Figur 8.1: Det gyldne snitt (Kilde: wikipedia.org)

En teknikk som er benyttet av malere i mange århundrer er å plassere viktige ting i et maleri i punktet for det gyldne snitt langs maleriets diagonal, for eksempel hodet til den mest sentrale personen i bildet. Dette har vært ansett av å være det punktet i maleriet som blikket trekkes mot først. Denne teknikken er så i moderne tid overført til andre former for grafikk. En mulig metrikk for å rangere bilder kan derfor være å rangere bildene etter deres avstand til punktet for det gyldne snitt på nettsiden sin diagonal.

8.4 Bildet og dets omgivelser på nettsiden

Mange nettsider bruker mye grafikk, for layout og for å se mest mulig innbydende og presentable ut, uten at dette egentlig bidrar til større forståelse av den informasjon nettsiden er utformet for å formidle. Med dette menes det konkrete innholdet og ikke den delen av formidlingen som ligger i å bygge troverdighet gjennom at nettsiden fremstår som estetisk pen og strukturert [35].

Bruken av pynt på nettsider er forståelig, mennesker liker orden og at det de holder på med er pent å se på. Men for en person som helt eller delvis ikke kan se gir dette antakeligvis ikke noe ekstra informasjon eller troverdighet. Den innsamlede informasjon fra studien tyder også på at slike bilder heller ikke er prioritert når det gjelder tekster i ALT-attributter, det er ikke funnet eksempler av typen *Dette bildet er rammen av en tabell*.

Bildene det er snakk om brukes ofte til å oppnå en helhet på siden og det er da naturlig å forvente at bildets relevans på siden kan sees opp mot dets omgivelser. Dersom et bilde for eksempel består av en ensfarget flate og omgivelsene har samme farge, eller i hvert fall nyanser som ligger nær hverandre, kan det antas at bildet har pynt eller struktur som oppgave.

For å kunne evaluere bildet opp i mot sin bakgrunn kan det være nødvendig å ta et skjermbilde av hele nettsiden og så å forsøke å finne igjen bildet som en del av dette skjermbildet, for eksempel ved hjelp av pikselplukk algoritmen fra kapittel 7. Dersom bildet ikke kan gjenfinnes, men fargene i det består primært eller utelukkende av samme farge som bakgrunnsfargen på nettsiden, kan det være rimelig å anta at bildet har estetikk eller struktur som oppgave og ikke informasjonsformidling.

8.4.1 Kontrast mot bakgrunnen

Et bilde som har høy kontrast mot bakgrunnen antas å være et blikkfang. Når bildets plassering har blitt funnet er sammenlikning av kontrasten mellom delbildet og den lokale bakgrunnen enkel. Kontrast er en egenskap ved forholdet mellom to farger som sier noe om hvor tydelig forskjell det er mellom fargene, ofte beskrevet som hvor mye to farger *skriker* mot hverandre. Hvordan kontrast måles er en egenskap definert i de fargemodellene som er beskrevet i kapittel 6. I dette tilfellet er det fokusert på kontrast mellom fargetoner. Kontrast kan imidlertid også oppfattes som kontrast mellom lyse og mørke farger i samme fargetone [70].

Fargemodellen som enklest gir mulighet for å beregne kontrasten mellom to fargetoner som en konkret størrelse er HSV-modellen hvor fargetonene er definert som grader på en sirkel. Den fargen som står rett ovenfor en gitt farge, med 180° avstand, er den som gir mest kontrast. Det kan derfor tenkes at kontrasten mellom den fargen som forekommer mest i bildet og den tilsvarende fargen i den lokale bakgrunnen kan benyttes som et mål på hvor mye et bilde virker som blikkfang på siden.

I uttrykk (8.4) er denne rangeringen definert som forholdet mellom avstanden i grader mellom den fargen i bildet som har flest piksler og den tilsvarende fargen i den lokale bakgrunnen og 180° . En avstand på 180 grader gir da rang 1.0 for størst mulig kontrast og en avstand på 0 grader gir 0.0 for minst mulig kontrast. En rang av 0.0 vil da si at det er snakk om samme farge i både bildet og bakgrunnen, med det resultatet at bildet forsvinner eller blir meget utydelig opp i mot bakgrunnen.

$$RB_{Kontrast}(Src) = \frac{\text{Grader kontrast mellom bilde og lokal bakgrunn}}{180^\circ} \quad (8.4)$$

Det kan tenkes at det kan være ønskelig å måle kontrasten av de N mest brukte fargene i bildet mot de $\leq N$ mest brukte fargene i den lokale bakgrunnen. Bildets totale rangeringen vil i så fall være et gjennomsnitt av de parvise kontraster.

Begrepet *den lokale bakgrunnen* har så langt vært brukt uten nærmere definisjon på det området av nettsiden som er i umiddelbart nærhet av bildet. Det er imidlertid nødvendig å definere denne presist for å kunne beregne rangering med uttrykk (8.4).

8.4.1.1 Et bildes lokale bakgrunn

Den lokale bakgrunnen er her definert som området rundt bildet som strekker seg ut fra kantene og utover til en avstand tilsvarende $\frac{1}{2}$ av lengden av bildet i den retningen. Selve beregningen av kontrasten er basert på fargehjulet som beskrevet i kapittel 6 og den samlede algoritmen for å rangere bildet etter kontrast mot lokal bakgrunn blir dermed:

1. Generer histogram av bildet.
2. Generer histogram av den lokale bakgrunnen, altså uten bildet i midten.

8.5. TEST AV UTVALGTE BILDEANALYSER

3. Finn fargen(e) som forekommer oftest i begge.
4. Beregn bildets rangering med $RB_{Kontrast}(Src)$.
5. Er det samme farge er rangen 0.0
6. Er det komplementærfarger er rangen 1.0

8.5 Test av utvalgte bildeanalyser

Blant de bildeklasser som er diskutert i dette kapitlet er det plukket ut tre for videre implementering og testing. Disse er rangering etter komprimeringsgrad, størrelse og kontrast mot bakgrunnen. Beskrivelse av testen og resultater av dette finnes i kapittel 11.

Del IV

Rangering av bilder

Kapittel 9

Samlet rangering

I de foregående kapitlene er det beskrevet algoritmer for å rangere bilder og deres bildetekster etter antatt viktighet som informasjonsbærer på en nettside. Både selve bildet og tilhørende bildetekst er i den hensikt blitt analysert via et sett med utvalgte metrikker som hver for seg sier noe om enkeltaspekter ved bildene eller bildetekstene.

Dette kapitlet beskriver hvordan en samlet rangering av bildet, basert på rangeringene fra både tekst- og bildemodalitet, skal settes sammen. Denne rangeringen skal så benyttes til å avgjøre hvilke bilder som bør tas med og hvilke som bør ignoreres av skjermleseren.

9.1 Grunnlaget for rangeringen

Rangeringsalgoritmen bygger på at det foreligger to sett med rangeringer, et med fra 0 til m rangeringer av bildeteksten og ett med fra 0 til n rangeringer av bildet. I begge tilfeller ligger hver rangering i det lukkede intervallet $[0.0, 1.0]$. Den samlede rangeringen utgjøres av gjennomsnittet av disse $n + m$ rangeringer. Det er dette gjennomsnittstallet som skal avgjøre hvorvidt en tekst skal inkluderes av skjermleseren eller ikke.

Det antas at alle rangeringer, i både tekst- og bildemodalitet, har samme betydning for den samlede rangering. For å kunne vurdere hvorvidt dette er en korrekt antakelse kreves mer erfaring med bruk av rangeringsalgoritmene over et mer omfattende antall og spekter av nettsider. Hvis det viser seg at vekting av rangeringen er nødvendig, kan det også tenkes at de må vektas forskjellig i ulike kategorier av nettstedet.

9.2 Rangeringsalgoritme

Rangeringen er gjennomsnittet av enkeltrangeringer fra tekst- og bildemodalitet. Uttrykk (9.1) gir det enkelte bildets totale rangering, i kontekst av den konkrete nettsiden som er under evaluering. Samme bilde-bildetekst paret kan ha forskjellig rangering på

forskjellige nettsider. Dette skyldes for eksempel at bildet kan være størst på en nettside og minst på en annen eller at teksten er repetert flere ganger på en nettside men ikke på en annen.

$$R = \frac{1}{n_T + n_B} \left(\sum_{i=1}^{n_T} T_i + \sum_{j=1}^{n_B} B_j \right) \quad (9.1)$$

I uttrykket i (9.1) er R er bildets totale rangering, T er settet av resultater fra tekstanalysene og B settet av resultater fra bildeanalysene. Videre er n_T antallet tekstanalyser i T og n_B antall bildeanalyser i B . Siden R er et gjennomsnitt av tall i intervallet $[0.0, 1.0]$ er R selv et tall i samme intervall.

9.3 Velge bildetekster etter rangering

Når rangeringen av alle bilde-bildetekst par på en nettside er beregnet er det nødvendig med en grenseverdi for å kunne vurdere om en skjermleser skal inkludere bildeteksten. Denne grenseverdien antas være avhengig av faktorer som kategori på nettstedet og personlige preferanser.

Når denne grenseverdien er satt er imidlertid avgjørelsen om hvilke bilder som skal inkluderes av skjermleseren redusert å avgjøre om de har rank R som oppfyller ulikheten i uttrykk (9.2). Verdien ϵ_c er da grenseverdien for nettsidekategorien c .

$$R \geq \epsilon_c \quad (9.2)$$

I uttrykk (9.2) må ϵ_c velges på grunnlag av hvilke rangeringer som er utført, hvordan disse er satt opp og den enkelte brukeres personlige preferanser. Eksempler på variasjon i oppsett av rangeringsalgoritmer kan være hvilke nøkkelord som er med i listene til $RT_{noekkel}(tekst)$.

9.4 Test av rangeringsalgoritme

Den samlede rangering som er beregnet av algoritmen i dette kapitlet er antatt å være et mål for hvor viktig et bilde er i relasjon til de andre bildene på nettsiden. I neste kapittel beskrives en metode for å teste algoritmen gjennom å sammenlikne den rekkefølgen som er frambrakt av rangeringen gitt av algoritmene med rekkefølgen på IMG-markørene i nettsidens HTML-kode.

Kapittel 10

Vurdering av rangeringen

De foregående kapitlene har beskrevet algoritmer for rangering av bildetekster og bilder, samt utregning av samlet rangering basert på disse enkeltrangeringer. Både de tekstbaserte, de bildebaserte og totalresultatet rangerer bildene på en nettside etter *viktigheten*, basert på bildeteksternes innhold og bildenes egenskaper som blikkfang på nettsiden. Viktighet er her definert som hvor stort blikkfang bildet utgjør på nettsiden og om bildetekstens innhold gir noen informasjon.

Det antas at bilder helt øverst og nederst på siden utgjør overskrift og faste elementer og at de informasjonsbærende bildene kommer noe lenger ned på nettsiden. Eksempel på dette er nettaviser med et stort banner med avisen navn og reklame øverst og illustrasjonsbilder for artiklene lenger ned.

Det kunne derfor være interessant å se om rangeringen gitt av algoritmene i denne oppgaven avviker, og i så fall graden av avvik, fra bildenes *naturlige* rekkefølge. Dette gjøres gjennom å sammenlikne rangeringen algoritmene gir med rekkefølgen på bildene slik IMG-markørene opptrer ved lineær gjennomlesning av HTML-koden, som i dette tilfelle defineres som bildenes *naturlige* rekkefølge.

En metode for å bestemme avstanden mellom to forskjellige sorteringer av en liste kalles *rangkorrelasjonskoeffisient*. Denne metoden vil først bli generelt beskrevet i detalj, for så å bli benyttet til å beregne avstanden mellom de forskjellige rangeringer som er funnet i oppgaven med bildenes *naturlige* rangering.

10.1 Kendall-Tau rangkorrelasjonskoeffisient

En rangkorrelasjonskoeffisient er en metode innenfor statistikk som benyttes til å måle hvor godt to forskjellige rangeringer av en liste, rangert etter to forskjellige sett kriterier, korresponderer og hvor nære denne korrespondansen er. Hvis de to rangeringen er helt like er koeffisienten 1.0 og hvis de to rangeringene er motsatte av hverandre er koeffisienten -1.0 . For alle mulige rangeringer mellom dette er koeffisienten i det åpne intervallet $< -1.0, 1.0 >$.

To metoder for å beregne rangkorrelasjonskoeffisienter er funnet, Kendalls Tau

[71] og Spearmans Rho [72]. Av disse er Kendalls Tau valgt fordi den er lettere å implementere og tolke enn Spearmans Rho. Det er også antatt at resultatene av de to algoritmene ikke varierer betydelig for de fleste applikasjoner [73].

Algoritmen ble utviklet på slutten av 1930-tallet av den britiske statistikeren Maurice Kendall. Algoritmen måler avstanden mellom to forskjellige sorteringer av en liste [71]. Hensikten med dette er å kunne oppdage sammenhenger, for eksempel å se om det er en sammenheng mellom høyde og vekt i en gruppe mennesker. Hvis de to sorteringene er tilnærmet like er det mulig å konkludere med at det kan være en sammenheng mellom høyde og vekt i denne gruppen.

10.1.1 Beregne rangkorrelasjonskoeffisient

Kendall Tau algoritmen teller antallet parvise uoverensstemmelser mellom listerne som skal sammenliknes. Dette gjøres ved at elementene i listene først nummereres fra 1 til n . Deretter sammenliknes differansen i rangeringen av alle par i listen X med alle par i listen Y .

La da nc være antallet par i listene hvor $x_i < x_j$ og $y_i < y_j$, nd er antallet par hvor $x_i > x_j$ og $y_i > y_j$. Videre er $extraX$ er antallet par i X hvor $x_i = x_j$, tilsvarende for $extraY$ i liste Y .

Kendall Tau rangkorrelasjonskoeffisient mellom listene X og Y er da gitt ved uttrykk (10.1) [73].

$$\tau = \frac{nc - nd}{\sqrt{nc + nd + extraX} + \sqrt{nc + nd + extraY}} \quad (10.1)$$

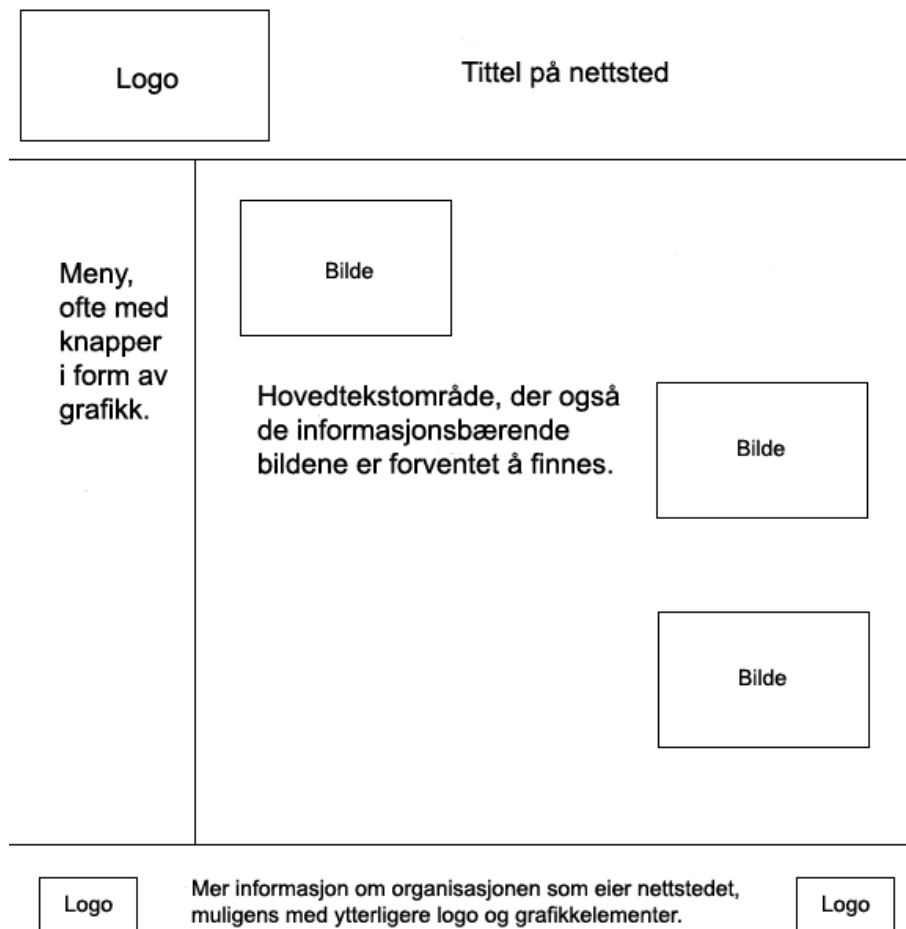
Dette kan sammenliknes med å telle hvor mange steg en bubblesorteringsalgoritme vil bruke på å endre sorteringen av X til Y , eller omvendt [73].

10.2 Tester på nettsteder

Rangeringen av bildetekster og bilder kan sees på som en alternativ rangering. For å gjøre betraktninger om disse rangeringene gir noe relevant informasjon er det nødvendig å kunne sammenlikne resultatene med en referansesortering. I dette tilfelle er det valgt å bruke rekkefølgen IMG-markørene kommer i HTML-koden til nettsiden. Dette er ikke et tilfeldig valg, men basert på antakelsen om at bilder som kommer først og sist i en nettside er mindre viktige enn de som kommer midt i.

Ved å lese en nettside med start i øvre høyre hjørne og slutt i nedre venstre hjørne, finnes som oftest hovedinnholdet midt på siden. I tillegg kommer ofte et hode med tittel på siden, organisasjonens logo og eventuelt lenker deler av siden som ønskes fremhevet. Videre forekommer det at nettsiden har en fot med kontaktopplysninger til for eksempel ansvarlig redaktør, gjerne også gjentakelse av logo i mindre format. Tilslutt er det ofte også en meny med lenker til andre sider i samme nettsted.

10.2. TESTER PÅ NETTSTEDER



Figur 10.1: Konseptskisse av en nettside

Figur 10.1 viser en skisse med en mulig layout hvor disse områdene er med. Selv om alle områdene ikke nødvendigvis er med, og i samme grad som her, er dette i mange tilfeller hoveddelene i en nettside.

10.2.1 Test på *bra* og *dårlige* nettsteder

Det er valgt å bruke finalistene i Direktoratet for forvaltning og IKT sin ¹ pris for tilgjengelighet ² som eksempler på *bra* nettsteder. Finalistene for 2009 var

- www.steinkjer.kommune.no
- www.sorum.kommune.no

¹<http://www.difi.no>

²<http://kvalitet.difi.no/arets-nettsteder>

- www.tromso.kommune.no.

Som *dårlige* nettsider er det valgt ut de tre som kom dårligst ut i tilgjengelighetstesten ³ som MediaLT ⁴ har utført på oppdrag fra Helsedirektoratet ved Deltasenteret. Disse var

- www.nav.no
- www.yr.no
- www.oslo.kommune.no.

Hensikten med disse testene er å se om, og i så fall hvor mye, rangeringen er forskjellig i disse to gruppene. Resultater og diskusjon av disse testene finnes i kapittel 11.

10.2.2 Test på nettaviser

Nettsidene definert som *bra* og *dårlig* i forrige avsnitt er imidlertid sider med relativt får IMG-markører, sett i forhold til sidene i kategorien *Nettaviser* i tabell 11.1. Dette gjør at nettsidene i de to nevnte tilfellene ikke er av en type som algoritmen er rettet mot. Derfor er denne testen gjennomført på nettaviser, som i følge tabell 11.1 er en kategori nettsider som bruker flere bilder, i form av IMG-markører, enn de andre målte kategoriene.

Til testen er det tilfeldig plukket ut ti norske nettaviser blant de som hadde mer enn 100 IMG-markører på sin side. Resultatene er beskrevet og diskutert i kapittel 11.

³<http://www.medialt.no/news/test-av-tilgjengelighet-paa-poppulaere-nettsteder/579.aspx>

⁴<http://www.medialt.no>

Del V

Resultater og diskusjon

Kapittel 11

Resultater og diskusjon

De foregående kapitler (5, 7, 8 og 10) foreslår algoritmer for å vurdere/rangere bildetekster (i form av ALT-attributter) på nettsider. I de samme kapitler er det foreslått tester for å kunne verifisere gyldigheten til disse. Algoritmene faller inn i fire kategorier.

De første analysene er av bildene for rangering etter antatt viktighet basert på bildetekstene. Dernest kommer lokalisering av rektangulære delbilder i et større bilde, med påfølgende analyse av de lokaliserte bildene for rangering etter antatt viktighet basert på egenskaper ved bildene.

Kapittelet avsluttes med samlet analyse av billedlig og tekstuell rangering med tanke på å skille ut hvilke bilder som er informative og hvilke som har andre oppgaver på en nettside, for eksempel estetisk pynt. I dette kapittelet blir resultatet av disse testene presentert og diskutert hver for seg.

11.1 Tekstanalyser

I kapittel 5 ble det nevnt flere mulige klasser av bildetekster som var antatt å ofte forekomme som innhold i en ALT-attributt. For å undersøke om dette er relevante klasser er det gjennomført opptelling av IMG-markører og tilhørende ALT-attributter og deres innhold på flere norske offentlige og private nettsider for å bringe på det rene hvilke klasser som forekommer oftest, hvis i det hele tatt.

11.1.1 Resultater

Resultatene av denne opptellingen er sammenfattet i tabell 11.1. Tabellen inneholder 48 nettsider fordelt på kategoriene *Nettavisser*, *Bank og finans*, *Nettbutikker*, *Høyere utdanning*, *Statlige institusjoner* og *Andre nettsteder*. For hvert nettstedets hovedside er det målt hvor mange IMG-markører det er på hver side, hvor mange av disse som har ALT-attributt og hvor mange av disse igjen som faller inn i de oppgitte tekstklasser.

11.1. TEKSTANALYSER

Merk at bildeteksten i én ALT-attributt kan falle inn i mer enn én tekstklasse, den kan for eksempel ha alle ord korrekt stavet og inneholde definerte nøkkelord.

For hver klasse av bildetekster er det for hver kategori oppgitt hvor mange det er totalt for alle sidene i kategorien og hvor hvor mange dette utgjør i prosent av det totale antallet IMG-markører.

11.1.2 Diskusjon

Hver tekstklasse i tabell 11.1 gjennomgås i detalj. Hvordan hver kategori sine nettsider faller inn i hver tekstklasse beskrives og en antatt grunn for hvorfor eller hvorfor ikke presenteres. Der enkeltnettsider i en kategori utmerker seg spesielt innenfor en eller flere tekstklasser er dette også forsøkt forklart.

11.1.2.1 Manglende ALT-attributt

Av det innsamlede tallmaterialet går det frem at de sidene som oftest mangler ALT-attributter er de sidene som det må antas oppdateres oftest, dette er *Nettbutikker* og *Nettavisser*. Innenfor disse to kategoriene mangler ALT-attributten på henholdsvis 13.4% og 23.3% av IMG-markørene. At nettbutikker og nettavisser er sider som oppdateres relativt ofte antas å være en del av grunnen til dette.

Nettavisser, som er den kategorien som utmerker seg med høyest prosent manglende ALT-attributter, oppdateres vanligvis daglig, ofte fra minutt til minutt når viktige saker er under utvikling. Det enkelte bildet er i mange tilfeller bare tilstede en kort tid, mens nyheten er fersk. En annen mulighet er at artikkelen må raskt ut, men oppdateres underveis, og at detaljer som bildetekster derfor legges på etterhvert. For å undersøke dette nærmere må imidlertid måten en nettavisredaksjon opererer på studeres nøyer, noe som ligger utenfor denne studien.

At *Nettbutikker* mangler mange ALT-attributter er imidlertid mer overraskende da man skulle anta at disse, utifra et reklamemessig synspunkt, ønsker å fremheve mest mulig informasjon om produktene. Det kan derfor være rimelig å anta at de manglende ALT-attributter i denne kategorien er for IMG-markører som ikke viser konkrete varer, noe som igjen impliserer at det kan være designelementer av mindre informasjonsbærende betydning.

I kategorien *Andre nettsteder* er prosent manglende ALT-attributter enda høyere, 31.2%, enn for nettbutikker og nettavisser. Det er imidlertid vanskelig å konkludere noe på grunnlag av denne tekstklassen i denne kategorien da den høye verdien skyldes et enkelt nettsted¹ av de åtte som er undersøkt.

I kategoriene *Bank og finans*, *Høyere utdanning* og *Statlige institusjoner* er antallet manglende ALT-attributter så lite at det antas å ikke kunne gi noe generelt mønster å følge ved klassifisering av bildetekster for disse kategoriene.

¹www.statoil.no

KAPITTEL 11. RESULTATER OG DISKUSJON

URL til nettside	IMG markører	Mangler ALT	ALT er tom	Rettskriving	Har nøkkelord	Er URL	Er repetert	Er tall
<i>Nettavisser:</i>								
www.vg.no	116	4	79	7	12	0	20	0
www.dagbladet.no	143	25	0	20	0	0	39	0
www.nettavisen.no	106	13	0	51	0	0	25	0
www.aftenposten.no	104	3	44	28	1	0	12	0
www.budstikka.no	115	19	8	27	0	0	5	0
www.nordlys.no	97	67	0	0	0	0	5	0
www.bt.no	133	22	0	8	1	0	6	0
www.tk.no	115	63	0	1	0	0	5	0
Total	929	216	131	142	14	0	117	0
Prosent		23.3	14.1	15.3	1.5	0.0	12.6	0.0
<i>Bank og finans:</i>								
www.skandiabanken.no	19	1	0	10	0	0	3	0
www.lillestrombanken.no	8	0	0	6	0	0	0	0
www.sparebank1.no	8	0	0	2	1	0	0	0
www.bnbank.no	14	1	0	5	0	0	5	0
www.gjensidige.no	1	0	0	0	0	0	0	0
www.spv.no	40	0	0	8	1	0	2	0
www.nordea.no	12	0	1	1	1	0	0	0
www.sparebanken-hedmark.no	7	0	0	2	1	0	0	0
Total	109	2	1	34	4	0	10	0
Prosent		1.8	0.9	31.2	3.7	0.0	9.2	0.0
<i>Nettbutikker:</i>								
www.clasohlson.no	22	5	0	2	0	0	0	0
www.jernia.no	7	0	0	5	1	0	0	0
www.midwaynorve.com	43	7	0	19	0	0	15	0
www.norli.no	64	1	0	48	20	0	37	9
www.xxl.no	11	0	0	4	1	0	2	0
www.akademika.no	65	10	0	42	0	0	8	0
www.komplett.no	36	0	0	24	0	0	0	18
www.gsport.no	51	17	0	1	0	0	17	0
Total	299	40	0	145	22	0	79	27
Prosent		13.4	0.0	48.5	7.4	0.0	26.4	9.0
<i>Høyere utdanning:</i>								
www.uio.no	44	0	0	9	1	0	5	0
www.hio.no	9	0	0	4	1	0	0	0
www.hist.no	31	0	0	7	0	0	15	0
www.ntnu.no	7	0	0	4	0	0	0	0
www.uib.no	8	0	0	2	0	0	0	0
www.uit.no	7	5	0	0	0	0	2	0
www.hil.no	23	0	0	8	2	0	2	0
www.hiak.no	35	0	15	10	2	0	0	0
Total	164	5	15	44	6	0	24	0
Prosent		3.1	9.2	26.8	3.7	0.0	14.6	0.0
<i>Statlige institusjoner:</i>								
www.regjeringen.no	14	0	0	2	0	0	0	0
www.stortinget.no	15	0	0	1	0	0	0	0
www.difi.no	9	0	0	6	4	0	0	0
www.helsedirektoratet.no/deltasenteret	100	3	0	13	0	0	86	0
www.oslo.kommune.no	8	0	0	3	1	0	0	0
www.politi.no	23	0	0	13	0	0	12	0
www.vegvesen.no	14	0	0	8	0	0	4	0
www.mil.no	92	18	0	1	0	0	0	0
Total	275	21	0	47	5	0	102	0
Prosent		7.6	0.0	17.1	1.8	0.0	37.1	0.0
<i>Andre nettsteder:</i>								
www.startsiden.no	35	2	0	5	1	0	15	0
www.qxl.no	36	0	0	15	0	0	6	0
www.finn.no	6	1	0	0	0	0	0	0
www.get.no	42	6	0	14	0	0	0	0
www.telenor.no	4	0	0	1	0	0	0	0
www.statoil.no	59	50	0	2	0	0	6	0
www.hydro.no	6	0	0	0	0	0	0	0
www.aker.no	1	0	0	0	0	0	0	0
Total	189	59	0	37	1	0	27	0
Prosent		31.2	0.0	19.6	0.5	0.0	14.3	0.0

Tabell 11.1: IMG - markører og ALT - attributter

11.1.2.2 Tom ALT-attributt

Med tanke på brukbarheten av de algoritmer for tekstanalyse som er diskutert i denne oppgaven er det meget oppløftende at antallet manglende ALT-attributter er så lavt som det er antydnet i forrige avsnitt. Av de gjenværende IMG-markører ser det imidlertid ut til at antallet tomme ALT-attributter også utgjør et betydelig antall av det totale antallet i noen kategorier, i dette tilfelle er kategoriene *Nettavisser* og *Høyere utdanning* høyest representert med henholdsvis 14.1% og 9.2% av IMG-markørene registrert med tomme ALT-attributter.

I kategorien *Nettavisser* var dette ikke uventet av samme grunner som for manglende ALT-attributter i denne kategorien i forrige avsnitt. At en nettside inneholder mange IMG-markører med manglende eller tomme ALT-attributter kan derfor tyde på at det er en side med hyppige oppdateringer. Dette kan igjen være en indikasjon på at siden brukes til hyppig publisering av nye artikler.

At institusjoner for høyere utdanning skulle komme så relativt dårlig ut i denne tekstklassen i forhold til de andre kategoriene var en overraskelse, da det var antatt at disse hadde regler for nettsider som krevde at standarder for tilgjengelighet var på plass. Imidlertid ser det ut til at også i dette tilfellet skyldes resultatene verdier fra en enkelt nettside ², og at det er derfor vanskelig å gi noen generell tolkning for denne kategoriene.

I de fire siste kategoriene *Bank og finans*, *Nettbutikker*, *Statlige institusjoner* og *Andre nettsteder* er få eller ingen IMG-markører i denne tekstklassen. Dette er bra siden det innebærer at det er relativt mange ALT-attributter som faktisk inneholder tekst i disse kategoriene og muligheten for å gjøre ytterligere tekstanalyser øker.

I sum for tekstklassene manglene eller tomme ALT-attributt er det derfor kun i kategorien *Nettavisser* at en betydelig mengde IMG-markører, 37.4%, ikke kan rangeres på grunnlag av tekstmodalitet, men må henvises til bildemodalitet alene. For alle kategorier gir dette allikevel et stort antall tilstedeværende tekster for analyse og rangering.

11.1.2.3 ALT med rettskrevet tekst

I hvor stor grad en tekst følger det aktuelle språkets regler for rettskrivning og grammatikk er av betydning for hvordan mottaker oppfatter innholdets troverdighet [46]. En ellers innholdsmessig korrekt tekst kan miste troverdighet dersom disse reglene ikke følges. Og kanskje ingen av et språks grammatikalske regler er så viktige i dette henseende som at ord er korrekt stavet. I denne tekstklassen er det derfor målt ratio av korrekt stavede ord i forhold til det totale antallet av ord i teksten. ALT-attributter som er talt med i denne kolonnen inneholder minst 50% korrekt stavede norske ord.

Ordlstens kvantitative og kvalitative innhold spiller her en stor rolle som feilkilde, spesielt ved forskjellige ordbøyninger, derfor er grensen satt til et forhold på 50/50. Allikevel ser det ut til at et relativt lite antall ALT-attributter oppfyller dette kravet.

²www.hiak.no

Det er også mulig at en kontroll med andre språk, med tanke på hvor populære særlig engelske ord er blitt.

En ytterligere grunn til det forholdsvis lave innholdet av rettskrevne ord som kreves i teksten er at ord som normalt ikke finnes i en generell ordliste kan ha forekommet. Dette kan være ord eller uttrykk som er konstruert og innarbeidet som merkevarenavn. Resultatene i denne tekstklassen bør derfor sees i sammenheng med nøkkelordene i neste avsnitt.

På tross av at kvaliteten på rettskriving antas å ha stor betydning for mottakers oppfatning av den ser det ut til at relativt få tekster oppfyller dette kravet. Den klassen med det laveste forholdet ser ut til å være *Nettavisser* der kun 15.29% av teksten har halvparten eller mer av ordene korrekt stavet ut ifra den benyttede ordlisten. Dette kan sies å være som forventet, da nettavisser har tekster som de ønsker å få publisert raskt, og helst før konkurrentene. Artikkelenes relativt korte levetid på forsiden antas også å være en grunn til at korrekturlesing ikke er en prioritet.

I den andre enden av skalaen er kategorien *Nettbutikker*. Vareutvalget her antas å være mer stabilt og varig enn en enkelt nyhetsartikkel, og dermed er muligheten og behovet for korrekte tekster større. Videre er troverdighet antatt å være viktigere for en nettbutikk da den skal motivere brukeren til å kjøpe fra denne heller enn konkurrentene.

Det er heller ingen av de resterende kategoriene i tabell 11.1 utmerker seg med et høyt ratio av rettskrevne ord i sine ALT-attributter. Dette kan tyde på at selv om ALT-attributter er blitt mer vanlig som en del av organisasjoners formelle krav til publisering på sine hjemmesider er det fremdeles noe vei igjen før teksten de inneholder ansees som en viktig og seriøs formidler av informasjon.

11.1.2.4 ALT med ord i en liste nøkkelord

I denne testen er det plukket ut noen få ord som opptrer ofte på de undersøkte nettsidene, for å teste algoritmen. Listene for ord som skal inkluderes og ekskluderes har i dette tilfellet vært som vist i uttrykk (11.1).

$$\begin{aligned}
 \text{Inkluder} &= \{ \text{"logo"}, \text{"handlekurv"}, \text{"vgtv"} \} \\
 \text{Ekskluder} &= \{ \text{"transparent"}, \text{"emediatead"}, \text{"foto"}, \text{"prikk"}, \text{"pil"} \} \\
 \text{Inkluder} \cap \text{Ekskluder} &= \emptyset
 \end{aligned}
 \tag{11.1}$$

Ordene er valgt ut blant flere kandidater som eksempler på nøkkelord som går igjen over flere sider. I inkluder listen er ordene «logo», «handlekurv» og «vgtv» lagt inn. De to første fordi de forekommer ofte i det innsamlede materialet og antas å være koplet til viktige bilder på nettside. Den første er generell og kan forventes funnet på alle typer nettsider, mens den andre er mer knyttet til nettbutikker og er valgt som eksempel på kategorisert oppsett av algoritmen. Dette gjelder også de tredje elementet i inkluderlisten, «vgtv», som er et eksempel på oppsett mot en spesifikk nettside.

I ekskluderlisten er ordet «Transparent» lagt til fordi det går igjen ofte. «Transparent» ser ut til å beskrive en teknisk egenskap ved bildefilen og ikke hva bildet

11.1. TEKSTANALYSER

inneholder. «EmediateAd» er lagt til i ekskluderlisten fordi det betegner bilder som er reklameannonser og ikke nødvendigvis relatert direkte til hovedinnholdet på den aktuelle nettsiden. Dette er oppsett basert på individuelle preferanser hos brukeren. De tre siste eksemplene i ekskluderlisten, «foto», «prikk» og «pil», er eksempler på nøkkelord som det antas å ikke gi betydning alene. Tekster som inneholder disse må sees i sammenheng med andre analyser, for eksempel at bildeteksten «pil» er koplet mot et bildet hvor en pil utgjør en vesentlig del av bildet. Det er satt som krav at en tekst rangeres over 0.5 for å bli telt med i resultatet, altså at det er overvekt av ord fra inkluderlisten.

For denne tekstklassen er det generelt få treff, spesielt i kategoriene *Andre nettsider* og *Statlige institusjoner*. En betraktning som kan gjøres er at *Nettavisene* har meget få treff i forhold til at det totale antall bilder i denne kategorien utgjør mer enn halvparten av de undersøkte IMG-markørene i hele testsettet. Dette kan skyldes at det av nøkkelord som opptrer mest, og som dermed bidrar mest til resultatet, er *Foto*. Nettavisene benytter seg av bilder fra mange kilder, og i de fleste tilfeller krever dette at kilden oppgis med en tekst på formen *Foto: navn på fotograf/kilde*. Dette gjør at mange bilder hos nettavisene rangeres lavt i denne tekstklassen. Det antas at mange av nettsidene i de andre kategoriene i større grad benytter seg av internt produserte bilder, og at de derfor heller ikke oppgir en eksplisitt kildehenvisning.

Det generelt lave antallet treff kan tolkes dit at begge nøkkelordlistene er meget korte og derfor utelukker mange tekster som ikke inneholder tekster fra noen av listene. Disse vil da få rangering 0.5, som er tatt som terskelverdi for i dette tilfellet. Dette er den eneste av tekstanalysene som gir brukeren mulighet til å påvirke rangering av tekster etter innhold. Dette kan tyde på at de testede inkluder- og ekskluderlistene er noe for snevre i forhold til antallet bilder på nettsidene. Dette ser ut til å gjelde for alle kategorier.

11.1.2.5 ALT som inneholder en filsti

Resultatene støtter ikke hypotesen at ALT-attributten noen ganger inneholder samme filsti som SRC-attributten. Det er ikke funnet ALT-attributter som inneholder filstier i noen av kategoriene. Dette inkluderer også ALT-attributter som kun inneholder et filnavn, for eksempel «dot.gif».

11.1.2.6 ALT som er repetert på samme side

I noen tilfeller opptrer samme tekst i flere ALT-attributter på samme side. Dette er ofte tekster som enten beskriver innholdsleverandørene eller egenskaper ved bildene. Et eksempel på det første er «EmediateAd», som ser ut til å være Dagbladets annonser fra annonser.dagbladet.no, eller «vgtv», som er video-artikler fra www.vg.no.

I gruppen egenskaper ved bildene er det særlig en nettside som utmerker seg. Dette er www.helsedirektoratet.no/deltasenteret som benytter et 1x1 gjennomsiktig bilde, med teksten «Transperant gif» i sin ALT-attributt, som en viktig element i utseende

på sin nettside. Det er imidlertid ikke urimelig å anta at å få opplest «Transperant gif» 86 ganger er støy i de fleste tilfeller og at denne bildeteksten burde filtreres bort ved kombinasjon av denne regelen og bildeanalyse av bildet med tanke på størrelse.

At tekster i ALT-attributter gjentas flere ganger på samme nettside ser ut til være noe som forekommer ofte. I de fleste kategorier ligger antallet forekomster i intervallet 9% – 15%. Ekstremverdier her finnes i *Nettbutikker* og *Statlige institusjoner* med henholdsvis 26.4% og 37.1%. I *Statlige institusjoner* er imidlertid dette grunnet de mange «Transperant gif» som hos <http://www.helsedirektoratet.no/deltasenteret>. Det skulle derfor ikke være grunnlag for å si at dette er en viktig test på nettsider i denne kategorien. I *Nettbutikker* er det mer jevnt fordelt internt i kategorien og dermed antakeligvis mer beskrivende for denne kategorien.

11.1.2.7 ALT som kun inneholder et tall

Bildtekster som inneholder et tall forekommer i testdataene relativt sjelden og da utelukkende i kategorien nettbutikker. Det er også innenfor denne kategorien begrenset til få, to av åtte, nettsider www.norli.no og www.komplett.no.

I Nordli sitt tilfelle ser dette ut til å være verdien på terningkastet denne boken har fått. Her burde teksten i ALT-attributten være påført en bedre tekst, «Terningkast 4» inneholder betydningsskillende mer informasjon enn bare «4». Her er det et godt eksempel på behov for samvirke mellom forskjellige typer analyser. Tekstanalysen alene kan ikke si noe konkret om hva dette tallet betyr. Men dersom en bildeanalyse kan fastslå at dette er bildet av en terning med 4 øyne kan de to analysene bekrefte hverandre og på egenhånd generere bildeteksten «Terningkast 4».

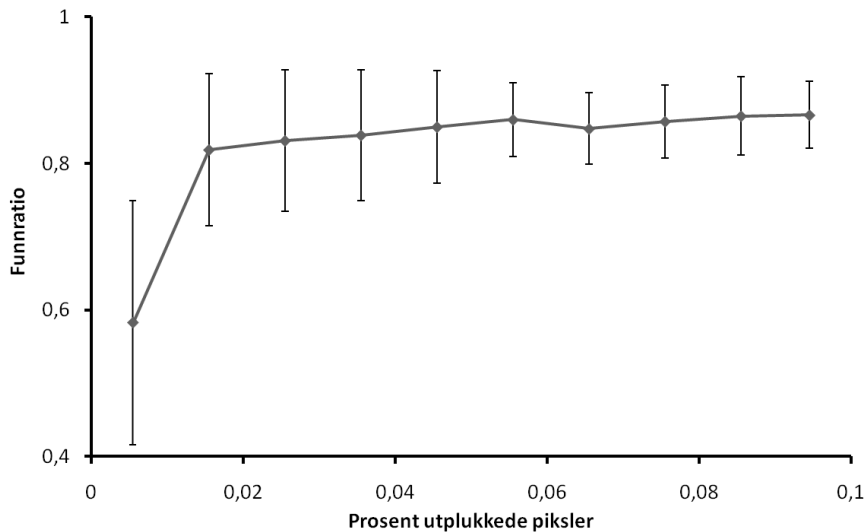
11.2 Pikselpukk algoritmen

I kapittel 7 ble det satt opp et antall tester for å verifisere pikselpukk algoritmen for lokalisering av bilder sin faktiske posisjon på en nettside. Bildene som er brukt i testene er hentet fra norske nettavisers forsider over flere dager. Punktene i grafene er fremkommet ved at det er beregnet gjennomsnittet av verdien i intervaller på ti måleresultater. I tillegg er det plottet standardavvik for verdiene i disse intervallene som er vist som feillinjer i horisontal retning ved punktet. Bildene i utvalget er maks på 300x300 piksler og nettsidene der er plassert på har en fast størrelse på 2000x2000 piksler.

11.2.1 Utplukksratio for piksler

Den første testen er tenkt å bekrefte at algoritmen fungerer under optimale forhold, samt finne et passende minste antall piksler som må undersøkes for å kunne finne bilder med denne metoden.

11.2. PIKSELPLUKK ALGORITMEN



Figur 11.1: Endring i antall korrekt plasserte bilder ved økning av antallet utplukkede piksler

Testen ble gjennomført ved å plukke ut fra 0.01% til 0.1% av pikslene i skjermbildene av nettsidene. Testsettet besto av 80 nettsider, hver bestående av et tilfeldig valgt bilde på hvit bakgrunn. For hver stegvis økning i antall utplukkede piksler er algoritmen kjørt på samtlige skjermbilder og antallet korrekt plasserte bilder er registrert.

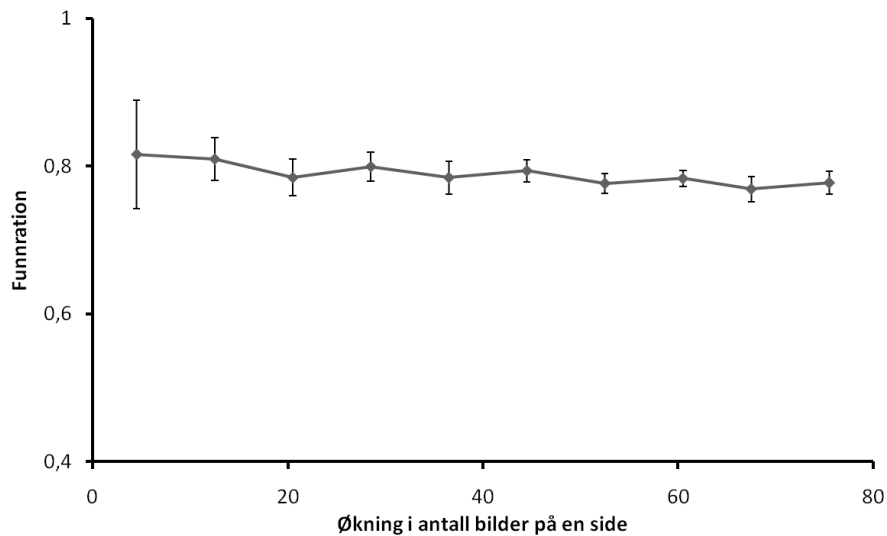
11.2.1.1 Resultat

En graf med resultatet er vist i figur 11.1. Funratio øker markant fra 59% til 82% når antallet utplukkede piksler øker 0.01% til 0.02%, etter at antallet passerer over 0.02% stabiliseres funratio på rundt 85%. Samtidig med at antall utplukkede piksler øker, minker også standardavviket. Dette innebærer ikke bare at antallet korrekt lokaliserte bilder øker, men også at spredningen mellom verdien i de forskjellige målinger minker etterhvert som antallet utplukkede piksler øker. Grafen viser totalt sett en lineært økende tendens, selv om den fra 0.02% til 0.1% bare øker fra 82% til 85%.

11.2.1.2 Diskusjon

Resultatet av denne testen tyder på en lovende algoritme. Allerede ved utplukk av kun 0.015% av pikslene i nettsiden er funratioen steget til over 80%. Ved 0.065% utplukkede piksler er det en liten knekk i grafen. Dette er imidlertid minimalt, verdien i dette punktet er bare et prosentpoeng mindre enn nabopunktene slik at funratioen også her er over 84%.

Avviket kan forklares med at algoritmen er basert på et *tilfeldig* utvalg av piksler. Det er derfor teoretisk mulig for algoritmen å bomme helt på et bilde, uansett hvor



Figur 11.2: Endring i antall korrekt plasserte bilder ved økning av antallet bilder på siden

godt fordelt funksjonen som plukker ut bildet er. Sannsynligheten for dette øker når størrelsen på bildet det søkes etter blir mindre i forhold til størrelsen på den totale nettsiden.

Siden antallet piksler er så lite i forhold til det totale antall piksler i skjermbildet er det i de påfølgende tester valgt å bruke en utplukksratio på 0.05%. Dette er gjort for å få en sikkerhetsmargin, men forventes ikke å gi et betydningsfullt utslag på algoritmens effektivitet i tid siden det allikevel er et meget lite antall av pikslene i skjermbildet som er involvert.

11.2.2 Antall bilder på nettsiden

Denne testen skal undersøke algoritmens evne til å korrekt lokalisere bilder på en nettside med mer enn et bilde. Antallet bilder på samme nettside økes med 1 av gangen fra 1 til 80. Det er testet 10 forskjellige nettsider for hver økning; 10 med 1 bilde, 10 med 2 bilder også videre. Funnratio er dermed det totale antall lokaliserte bilder, av alle bildene på de 10 nettstedene.

Testen ble gjennomført ved å bruke et utplukk på 0.05% av pikslene i skjermbildet, dette ble valgt på grunnlag av resultatet fra grafen i figur 11.1.

11.2.2.1 Resultat

Resultatet i figur 11.2 viser en lineært avtagende funnratio etterhvert som antallet bilder øker. Funnratio er imidlertid fremdeles på over 75% når det er over 75 bilder på samme nettside. Standardavviket minker også etterhvert som antallet bilder øker.

11.2.2.2 Diskusjon

At antallet korrekt lokaliserte bilder skulle minke etterhvert som antallet bilder på nettsiden økte var som forventet. Selv om sjansen for at et av de utplukkede pikslene skal havne innenfor et av bildene øker, øker også sjansen for at to forskjellige bilder skal være tilstrekkelig like hverandre eller at de skal inneholde noen av feilkildene som er beskrevet i kapittel 7. Det ser imidlertid ikke ut til at dette har gitt så store utslag som det var antatt før testene ble gjennomført. Grafen har noen knekker, men disse utslagene er også relativt små og kan antas å skyldes samme grunn som knekken i grafen i figur 11.1.

En nettside med over 75 bilder er en ganske omfattende nettside. I tallmaterialet i tabell 11.1 i avsnitt 11.1 er det summert antallet IMG-markører på forskjellige norske nettsider i flere kategorier. I nettsiden med maksimalt antall her er www.dagbladet.no med 145, mens de fleste har et antall langt under dette. Hvis endringen i funnratio fortsetter som antydnet i 11.2 vil den allikevel ligge på over 70% på en side med 160 IMG-markører. Fallet i funnratio med økende antall bilder kan derfor antas å ligge innenfor et akseptabelt nivå for de nettsider som er representert i det innsamlede tallmateriale.

11.2.3 Oppskalering - 1 dimensjon

Testen er gjennomført ved å velge et bilde som skal skaleres og 12 andre bilder som *bakgrunnsstøy*. Skaleringen er gjort med stegvis økning på et prosentpoeng fra 1% til 80% i bildets bredde og det er 10 testsider for hver økning. Hver nettside har et bilde som endrer størrelse. Dette bildet er forskjellig fra det tilsvarende bildet på de ni andre nettsidene. Også i denne testen, og i de tre påfølgende skaleringstester, er antallet piksler som plukkes ut satt til 0.05%.

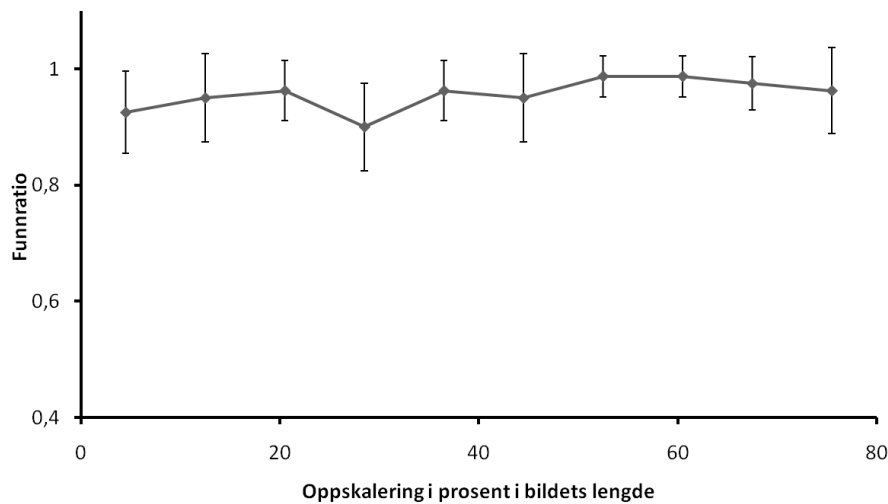
11.2.3.1 Resultat

Resultatet i figur 11.3 viser at endringen i funnratio endrer seg lite eller intet når et bilde skaleres opp til nærmest dobbel størrelse. Standardavviket er noe større der hvor funnratio er minst.

11.2.3.2 Diskusjon

Fordi skalering er en mye brukt teknikk i følge med bilder på nettsider er det viktig i analysen av en nettside å kunne finne igjen et skalert bilde. Fra resultatene av denne testen ser dette ut til å være oppfylt innenfor det intervallet som er testet, 1% til 80% oppskalering i lengderetning.

Resultatene tyder på at å sammenligne normaliserte histogrammer er en lovende metode for å sammenligne to bilder og at det relative innholdet i bildet av hver farge i forhold til det totale antallet piksler ikke endres vesentlig når bildet skaleres opp i en



Figur 11.3: Endring i antall korrekt plasserte bilder ved oppskalering i bredde

side. Det antas at det spiller liten rolle om skaleringen er gjort i bredde eller høyde, selv om testene i dette tilfelle er gjort i bredde skulle det gi samme resultat i høyde.

11.2.4 Nedskalering - 1 dimensjon

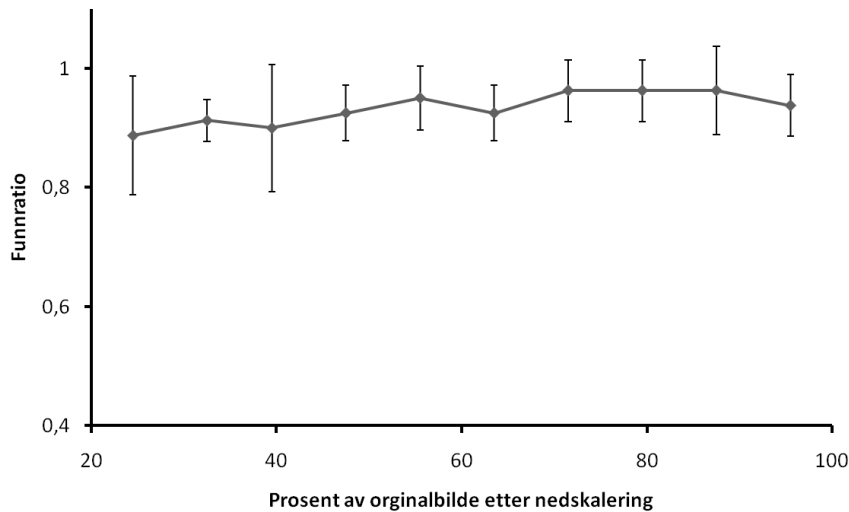
Denne testen er utført på helt samme måte som 1 dimensjonal oppskalering, bare at her er bildet skalert ned. Grafens horisontale akse viser imidlertid hvor mange prosent av bildet som er igjen og ikke hvor mange prosent bilde er redusert med.

11.2.4.1 Resultat

Resultatene i figur 11.4 viser at antallet korrekt lokaliserte bilder minker ettersom bildet blir mindre. Standardavviket øker også etterhvert som bildet blir mindre og mindre.

11.2.4.2 Diskusjon

På samme måte som det var forventet at resultatet ved 1 dimensjonal oppskalering skulle gi et forholdsvis jevnt resultat, var det også antatt at nedskalering skulle gi et fall i funnratio. Dette på grunn av forskjellen på algoritmer for å fjerne og å legge til piksler ved henholdsvis ned- og oppskalering. Dette er beskrevet i kapittel 6. Fallet er imidlertid ikke meget stort, etter at bildet er skalert ned med 90% er fremdeles over 80% av bildene korrekt lokalisert.



Figur 11.4: Endring i antall korrekt plasserte bilder ved nedskalering i bredde

11.2.5 Oppskalering - 2 dimensjoner

Denne testen er gjennomført på samme måte som ved oppskalering i en dimensjon, men med inkrementell økning i samme prosent i begge retninger.

11.2.5.1 Resultat

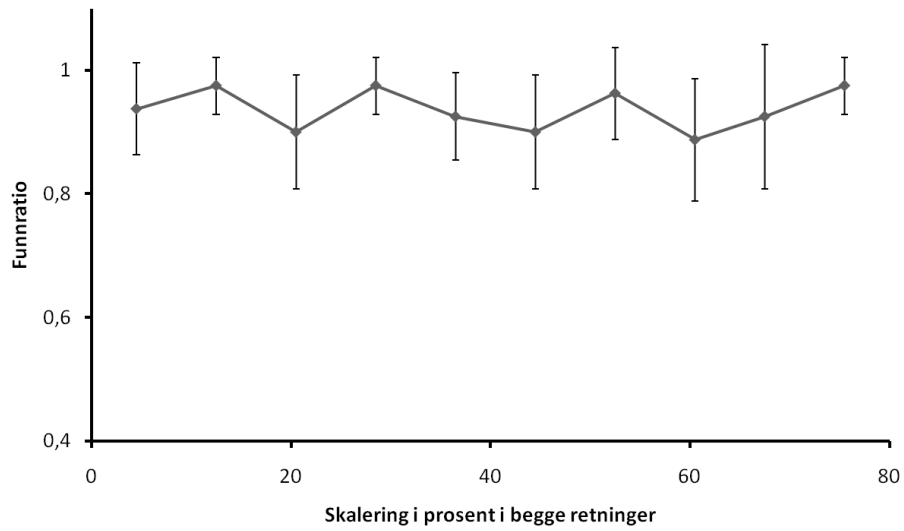
Resultatet i figur 11.5 viser at variasjonen i funnratio endrer seg lite etterhvert som bildet skaleres oppover. Variasjonene i standardavviket er også små, med størst avvik for de minste funnratio.

11.2.5.2 Diskusjon

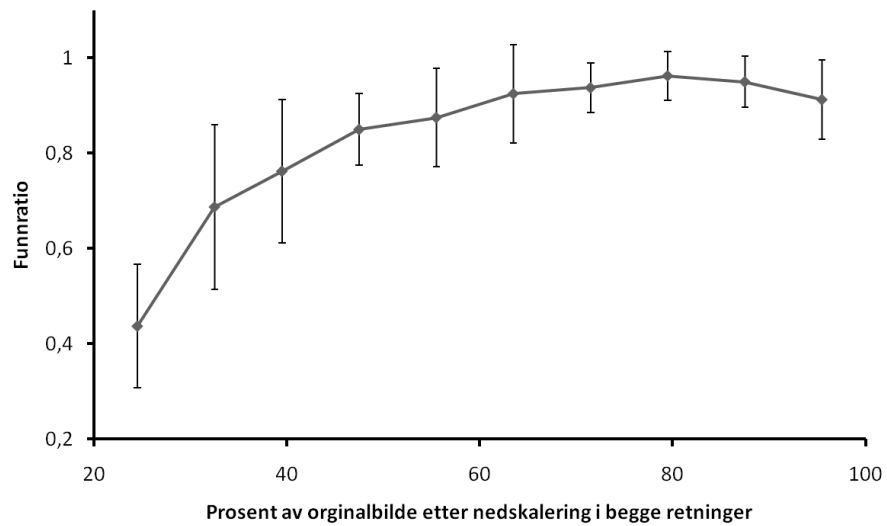
At skalering i 2 dimensjoner skulle gi lite utslag på funnratio var som forventet, da forholdet mellom antallet piksler i samme farge og det totale antallet piksler også her er tilnærmet konstant. Variasjonen i funnratio antas å skyldes den tilfeldige utplukkingen av piksler som er diskutert tidligere for de andre testene.

11.2.6 Nedskalering - 2 dimensjoner

Denne testen er gjennomført på samme måte som ved oppskalering i 2 dimensjon, men med inkrementell reduksjon i samme prosent i begge retninger. Som for 1 dimensjonal nedskalering viser grafens horisontale akse hvor mange prosent av bildet som er igjen og ikke prosent reduksjon.



Figur 11.5: Endring i antall korrekt plasserte bilder ved oppskalering i begge retninger



Figur 11.6: Endring i antall korrekt plasserte bilder ved nedskalering i begge retninger

11.2. PIKSELPLUKK ALGORITMEN

11.2.6.1 Resultat

Resultatet er vist i figur 11.6 viser at funnratio øker etterhvert som det er mer og mer igjen av bildet. Dette er det samme som for 1 dimensjonal nedskalering, men fallet i antall korrekt lokaliserte bilder er mer markant etterhvert som bildet blir mindre enn hva som er tilfelle med 1 dimensjonal nedskalering.

11.2.6.2 Diskusjon

At reduksjon i funnratio her er tydelig mer markant enn tilsvarende reduksjon for 1 dimensjonal må antas å skyldes at antallet piksler som skaleringsalgoritmen slår sammen er mye større og at histogrammet derfor blir tilsvarende redusert i forhold til histogrammet for originalen. Det blir dermed lettere for algoritmen å gjøre feil, fordi antallet piksler som er igjen i bildet er mindre enn ved en tilsvarende reduksjon i prosent i bare én dimensjon.

11.2.7 Delvis tildekkede bilder

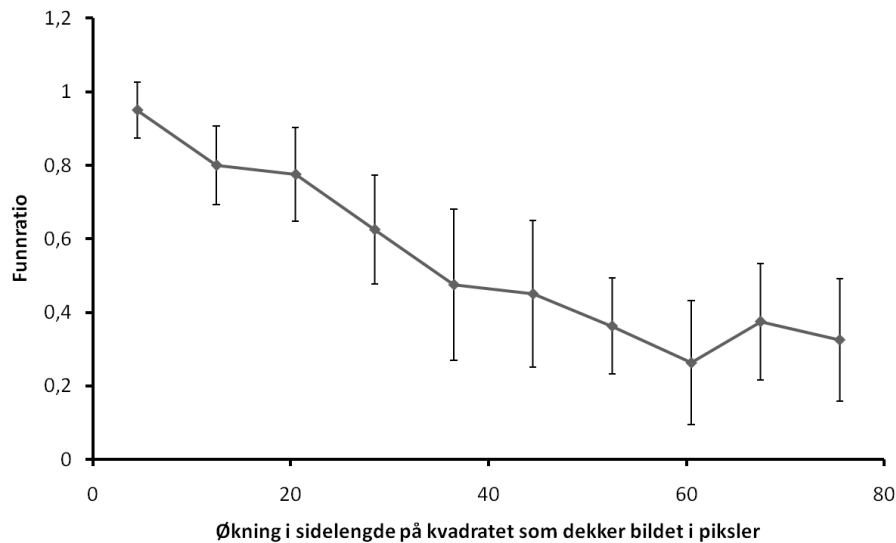
Denne testen skal demonstrere hva som skjer dersom et bilde er delvis tildekket av andre elementer i en nettside. I dette tilfellet vil et ensfarget rektangel flyttes langs med bildets diagonal, med en økning på 1 piksel av gangen, slik at det delvis dekker til det utplukkede bildet. I hver økning er det igjen undersøkt 10 nettsteder og funnratio er hvor mange bilder som er korrekt lokalisert etter hver økning. Nettsidene inneholder også andre bilder enn det som tildekkes som støy. Testen ble gjennomført ved å bruke same utplukk på 0.05% av pikslene i skjermbildet som er i bruk i alle tester etter den første.

11.2.7.1 Resultat

Resultatet i figur 11.7 viser et fall i funnratio som følger det antallet piksler som dekkes av den fargede fikanten. Standardavviket vokser på samme måte også i takt med reduksjonen i funnratio.

11.2.7.2 Diskusjon

Dette er som forventet. Etterhvert som antallet piksler i kandidatbildet reduseres vil histogrammet for dette i økende grad fjerne seg fra histogrammet for det bildet man leter etter idet piksler fra den overlappende regionen kommer mer og mer med i histogrammet for kandidatbildet. Kandidatbildet vil fremdeles kunne finnes så lenge det er mulig å søke ut til rammen som omslutter det, men histogrammet som beregnes vil etterhvert være for forskjellig fra histogrammet til bildet det søkes til å kunne fastslå om det dreier seg om samme bilde.



Figur 11.7: Endring i antall korrekt plasserte bilder ved økning av arealet av bildet som gjemmes/beskjæres

11.2.8 Diskusjon av pikselplukk

Testene viser at algoritmen fungerer med mellom 80% og 90% funnratio i de fleste testtilfeller unntatt 2 dimensjonal nedskalering og når bildet er delvis tildekket. Dette er som forventet da det er i disse to tilfeller at antallet piksler i hver farge minkermest markant. I tilfellet med nedskalering vil også mange piksler kombineres til en gjennomsnittsverdi ettersom bildets oppløsning reduseres. Histogrammene som skal sammenliknes vil dermed bli for ulike, selv om de er normalisert. I sum tyder derfor testene på at pikselplukk-algoritmen er et godt grunnlag for de bildeanalysene som er testet og diskutert i påfølgende avsnitt.

11.3 Bildeanalyser

Dette avsnittet beskriver resultater av tester gjennomført på algoritmer fra kapittel 8. Hensikten her er å måle om funksjonene for å rangere bilder etter størrelse, kompleksitet i mønster og farger og som følge av kontrast mot bakgrunnen fungerer som forutsatt.

11.3.1 Resultat

Å teste algoritmene for bildeanalyse viste seg å være noe mer omfattende enn for tekstanalysene. I stedet for å benytte de samme nettsider som for tekstanalyser ble bildeanalysene derfor testet ved å kjøre analysene på konstruerte nettsider.

11.3. BILDEANALYSER

I hvert av disse er det fem bilder med IMG-markører i det som antas være motsatt rekkefølge i forhold til den rekkefølgen de respektive rangeringsfunksjoner vil gi. I alle testene er avstanden mellom rangeringen og IMG-markørens rekkefølge målt ved hjelp av Kendalls Tau som definert i kapittel 10.

11.3.1.1 Etter størrelse

Rangering etter størrelse er gjort ved å plassere fem ensfargede bilder på samme nettside. Bildene er plassert etter økende størrelse på nettsiden slik at rangeringsfunksjonen for størrelse skal gi den helt motsatte rangeringen. Utskrift av rangeringen er gitt i figur 11.8.

IMG/ALT statistikk:

IMG-markører på siden:	5
IMG med ALT-attributt:	5 (100%)
IMG uten ALT-attributt:	0 (0%)
ALT-attributt med innhold:	5 (100%)

Rekkefølge på nettsiden (Rank, ALT-tekst):

```
1.000000, 1
2.000000, 2
3.000000, 3
4.000000, 4
5.000000, 5
```

Rekkefølge etter størrelse (Rank, ALT-tekst):

```
1.000000, 5
0.871111, 4
0.751111, 3
0.640000, 2
0.537778, 1
```

Kendall-Tau opprinnelig/etter rangering:

```
-1.000000
```

Figur 11.8: Nettside med 5 bilder i stingende størrelse

11.3.1.2 Komprimeringsgrad

Komprimeringsgrad er testet på en nettside med fem like store kvadratiske bilder. Langs bildenes diagonal er det plassert fra en til fem kvadrater, hvor bildet med fem kvadrater anses å være mer komplekst enn det med et kvadrat. Bildene er plassert på

nettsiden i økende kompleksitet, slik at rangeringsfunksjonen for komprimeringsgrad skal gi den helt motsatt rangering med den mest komplekse først. Utskrift av rangeringen er gitt i figur 11.9.

```

IMG/ALT statistikk:
IMG-markører på siden:          5
IMG med ALT-attributt:         5 (100%)
IMG uten ALT-attributt:        0 (0%)
ALT-attributt med innhold:     5 (100%)
    
```

Rekkefølge på nettsiden (Rank, ALT-tekst):

```

1.000000, 1
2.000000, 2
3.000000, 3
4.000000, 4
5.000000, 5
    
```

Rekkefølge etter komprimeringsgrad (Rank, ALT-tekst):

```

0.600000, 5
0.480000, 4
0.375000, 3
0.295000, 2
0.225000, 1
    
```

Kendall-Tau opprinnelig/etter rangering:
-1.000000

Figur 11.9: Nettside med 5 bilder i økende kompleksitet

11.3.1.3 Kontrast mot bakgrunn

Kontrast mot bakgrunnen er testet på en nettside med fem bilder med diverse fargerike figurer på rød bakgrunn. Bildene er omkranset av en ensfarget, bred ramme og hvert bildets ramme sin farge er 0 deg, 30 deg, 60 deg, 90 deg og 120 deg fra rødt på fargehjulet. Disse fargene er illustrert i figur 11.10.



Figur 11.10: Fem farger i økende fargekontrast fra rødt.

11.4. TEKST -OG BILDEANALYSER I SAMMENHENG

IMG/ALT statistikk:

IMG-markører på siden:	5
IMG med ALT-attributt:	5 (100%)
IMG uten ALT-attributt:	0 (0%)
ALT-attributt med innhold:	5 (100%)

Rekkefølge på nettsiden (Rank, ALT-tekst):

```
1.000000, 1
2.000000, 2
3.000000, 3
4.000000, 4
5.000000, 5
```

Rekkefølge etter kontrast (Rank, ALT-tekst):

```
0.666667, 5
0.500000, 2
0.494444, 4
0.333333, 3
0.000000, 1
```

Kendall-Tau opprinnelig/etter rangering:

```
-1.000000
```

Figur 11.11: Nettside med 5 bilder med økende fargekontrast mot bakgrunnen

11.3.2 Diskusjon

I tilfellene størrelse og kompresjon er rekkefølgen etter rangering i disse bildeklassene som forventet. Rekkefølgen etter størrelse er endret til at størst er på førsteplass i rangeringen, den var siste plass etter IMG-markørenes rekkefølge. Det samme gjelder kompleksitet som er endret til at det antatt mest komplekse bildet er først i rangeringen.

Når det gjelder kontrast mot lokal bakgrunn har denne også rangert bildene etter det forventede resultat med et unntak. Bilde nummer to, med 30 deg avstand i fargekontrast, har ikke forandret plass i rangeringen. Dette skyldes antakelig at denne rangeringsfunksjonen er avhengig av at pikselplukk-algoritmen lokaliserer bildet korrekt. Det antas at dette ikke er gjort i dette tilfellet og bildet er derfor analysert mot *feil* bakgrunn.

11.4 Tekst -og bildeanalyser i sammenheng

Testene så langt har testet enkeltfunksjoner for å klassifisere bilder og bildetekster. De påfølgende testene skal avgjøre om disse egenskapene tilsammen er tilstrekkelig for å

avgjøre om en bildetekst er *viktig* nok til å bli inkludert i en skjermleserens tolkning av nettsiden. Testene er gjennomført etter beskrivelse av algoritmer i kapittel 10. Siden tekstklassene *filsti* og *tall* opptrer så sjelden er det besluttet å ta disse ut av den samlede test. Det samme med testfunksjonen om teksten ikke har innhold da dette er implisert i de tre resterende tekster ved at et en tekst uten ord får 0.0 poeng ved kontroll mot ordlisten.

11.4.1 Gode og dårlige nettsider

I avsnitt 10.2.1 ble det definert tre antatt gode og tre antatt dårlige nettsider, i forhold til deres tilgjengelighet for funksjonshemmede. For disse sidene er både tekst -, bilde - og samlet rangering sammenliket med IMG-markørens rekkefølge på i HTML-koden ved hjelp av Kendalls Tau.

11.4.1.1 Resultater

Resultatene for de gode nettsidene er gitt i tabell 11.2 og for de dårlige i tabell 11.2. NaN i tabellen indikerer at det ikke har vært noen bilder/tekster i den kategorier og at det derfor ikke har vært noen alternativ ordning å måle Kendalls Tau avstand til.

Nettside	Samlet	Nøkkelord	Staving	Repetert	Komprimering	Størrelse	Kontrast
www.tromso.kommune.no	-0.971825	-0.374448	-0.889385	-0.590937	-0.877406	-0.966366	-0.519462
www.sorum.kommune.no	-0.953105	NaN	-0.855998	-0.867325	-0.930788	-0.901284	-0.819250
www.tromso.kommune.no	-0.969223	-0.408248	-0.807165	NaN	-0.937437	-0.945484	-0.674200

Tabell 11.2: Kendall Tau for gode nettsier

Nettside	Samlet	Nøkkelord	Staving	Repetert	Komprimering	Størrelse	Kontrast
www.yr.no	-0.987214	-0.464727	-0.886137	-0.868554	-0.961761	-0.874108	-0.869598
www.nav.no	-0.967153	-0.701646	-0.630018	-0.384308	-0.947872	-0.952729	-0.786423
www.oslo.kommune.no	-1.000000	-0.731925	-0.681385	NaN	-1.000000	-1.000000	-0.981981

Tabell 11.3: Kendall Tau for tre dårlige nettsier

11.4.1.2 Diskusjon

Resultatene viser at rekkefølgen etter rangering i begge kategorier er nesten totalt reversert i forhold til rekkefølgen på IMG-markørene i HTML-koden. Dette er som forventet. I skissen i figur 10.1 er de fleste informasjonsbærende bilder plassert midt på siden med logoer og annen layout-grafikk plassert primært på toppen av siden. Det samme er tilfelle for sidene som undersøkes i dette avsnittet. Imidlertid er det så få IMG-markører på disse sidene at det var ønskelig å få dette bekreftet ved å undersøke en kategori nettsider med antatt mer IMG-markører. Denne kategorien er nettaviser.

11.4. TEKST -OG BILDEANALYSER I SAMMENHENG

11.4.2 Ti nettaviser

I denne testen er det plukket ut ti tilfeldige norske nettaviser som i testøyeblikket hadde mer enn 100 IMG-markører på sin hovedside. En oversikt over hvilke tekstklasser disse markørenes ALT-attributter falt inn i er gitt i tabell 11.4.

URL til nettside	IMG markører	Mangler ALT	ALT er tom	Rettskriving	Nøkkelord	Er repetert
Nettavisere:						
www.budstikka.no	110	18	4	25	0	9
www.dagbladet.no	131	25	0	19	0	32
www.f-b.no	154	33	3	36	0	10
www.harstad-tidende.no	149	1	0	24	0	2
www.laagendalsposten.no	118	16	0	32	0	14
www.moss-avis.no	152	39	5	35	0	24
www.nettavisen.no	113	13	0	51	0	26
www.stavanger-aftenblad.no	157	32	0	29	0	28
www.varden.no	112	37	2	30	0	5
www.ostlendingen.no	124	22	2	34	0	11
Total	1320	236	16	315	0	161
Prosent		17.9	1.2	23.9	0.0	12.2

Tabell 11.4: Ti nettaviser med mer enn 100 IMG-markører på hovedsiden

Bildene, og tilhørende bildetekster, på disse ti nettavisene sin forside er så rangert på samme måte som de gode og dårlige nettstedene i avsnitt 11.4.1.

11.4.2.1 Resultater

Ved å rangere bildene på nettavisene i tabell 11.4 fremkommer Kendall-Tau distansene mellom tekst -, bilde -og samlet rangering sammenliknet med IMG-markørenes rekkefølge i HTML-koden ved hjelp av Kendalls Tau i tabell 11.5. Av disse resultatene er så alle tekster fra ALT-attributter fra www.nettavisen.no tatt med i tabell 11.6.

Nettside	Samlet	Nøkkelord	Staving	Repetert	Komprimering	Størrelse	Kontrast
www.varden.no	-0.993314	-0.718538	-0.771148	-0.380420	-0.951070	-0.973947	-0.817295
www.stavanger-aftenblad.no	-0.982386	-0.637885	-0.607649	-0.540610	-0.960177	-0.966430	-0.771068
www.ostlendingen.no	-0.994975	-0.661419	-0.725968	-0.410702	-0.942170	-0.974891	-0.857635
www.nettavisen.no	-0.963374	-0.647612	-0.893005	-0.624082	-0.807283	-0.852074	-0.557648
www.moss-avis.no	-0.986319	-0.718149	-0.779974	-0.553626	-0.947903	-0.968546	-0.793311
www.laagendalsposten.no	-0.977893	-0.691578	-0.708264	-0.427645	-0.918251	-0.958400	-0.788935
www.harstad-tidende.no	-0.993612	-0.625797	-0.662352	-0.160121	-0.956665	-0.971612	-0.947999
www.f-b.no	-0.979912	-0.715837	-0.763170	-0.358446	-0.950896	-0.974788	-0.715639
www.dagbladet.no	-0.987252	-0.670324	-0.473878	-0.629530	-0.894319	-0.968204	-0.800789
www.budstikka.no	-0.979530	-0.735360	-0.735677	-0.476591	-0.930933	-0.968692	-0.838449

Tabell 11.5: Kendall Tau for nettaviser i figur 11.4

KAPITTEL 11. RESULTATER OG DISKUSJON

Rank	Rekkefølge	Tekst
0.724522	68	Vaktsjef
0.705185	109	Bli fan av Nettavisen på Facebook
0.694354	63	Nett på sak
0.601980	6	Oslo Børs
0.593425	91	Skibilder.no
0.585715	69	Klikk.no
0.573406	1	1881.no
0.550002	23	Newcastle har hatt en kanonsesong og vinner igjen i dag.
0.546424	7	Dow Jones Industrial
0.544031	108	Facebook
0.535705	95	Tips
0.534695	93	
0.534316	50	Med vår anbudsjeneste kan du lett finne en dyktig eiendomsmegler som ikke flår deg helt.
0.527943	103	Side2.no
0.527906	5	USD/NOK
0.525864	12	RYKKET UT: Politiet fant mannen livløs da de rykket ut til en adresse i Narvik fredag kveld etter melding om en konflikt på en privat fest.
0.525643	34	UD oppfordrer nordmenn som allerede befinner seg i Bangkok mot å bevege seg i områder med store folkemengder. samt holde seg unna demonstrasjonene.
0.522729	11	Suksesstrener: Alex Ferguson gir seg ikke med det første.
0.520835	21	ASKEFAST: Hva er din historie fra askeperioden? Mange reisende opplevde venting og lang hjemreise
0.519843	46	JAGER GULL: Disse Barcelona-supporterne kommer nok til å fyre blussene sine igjen hvis det blir seriegull i Spania.
0.516669	59	Det er ikke påtrengende mange som kommer til den vakre øya i Bahamas.
0.515696	98	VIF-fotball.no
0.513891	106	Ciara holder det naturlig og enkelt i sin usjenerte video (Sony)
0.512927	Samme rank ^a	Si din mening!
0.510419	96	Spillmagasinets samarbeidspartner 007.com har plukket ut tre nettspill til helgens mindre alvorlige gjøremål.
0.507578	36	SKAL PÅ KINO: Ancelotti skal ikke se United-kampen.
0.507578	61	Trond Hansen viser frem regningen han har fått av Travelpartner.
0.500546	85	Match
0.500002	8	I HARDT VÆR: Paul-Christian Rieber er i hardt vær etter påstått tolljuks i eget selskap. Nå velger han å gå av som NHO-president.
0.500002	20	LESERNE STILTE OPP I WOMAN: Her bikinimodellene som redaksjonen var svært fornøyde med.
0.500002	52	RAMPELYSFESTIVALEN: Mette-Marit tok med seg barna til åpningen av barne- og ungdomsfestivalen.
0.496916	51	ORD I ØRET: Per Arne Olsen. som her hviker partileder Siv Jensen noen ord i øret. etterlyser regjeringens definisjon av et lokalsykehus.
0.496379	13	STOLER IKKE PÅ OBAMA: En demonstrant uttrykker tydelig mistillit til sittende president Barack Obama under et av Tea Party-bevegelsens skattedemonstrasjoner.
0.495789	80	Norsk Rikstoto
0.495461	87	
0.494050	30	KAN FÅ GULL: Tore Reginiussen har kun spilt 19 minutter for Schalke A-lag, men får likevel gull klubben tar Bayern i gullkampen.
0.493058	39	GÅR FOR GULL: Therese Johaug håper på sol, blå himmel, god stemning og full utklassing i Skarverennet.
0.487376	32	BARE NESTEN: Stefan Gislason får Sead Ramovic til å strekke seg under eliteseriekampen i fotball mellom Viking og Tromsø, på Viking stadion. Kampen endte 1-1.
0.485296	44	STÅR DISTANSEN: Lars Petter Nordhaug har imponert stort under debutsesongen for Team Sky. Det britiske laget bør kanskje sende en takk til Gino van Oudenhove som lurte ham av terrengsykkelen.
0.483335	81	Trineline er fredagens Biri-banker.
0.479169	4	DENZEL WASHINGTON i «The Book of Eli».
0.479169	78	West Ham har alt i egne hender når Wigan kommer på besøk og plassen i Premier League neste sesong skal sikres.
0.478682	107	RSS
0.475002	102	Henrik Østby er synes studiene har blitt morsommere, etter at han lærte seg å bruke husketeknikker.
0.473960	25	HOPP: Det befinner seg et stort antall trampoliner rundt omkring i norske hager, og det skjer stadig ulykker. Tar man klare forhåndsregler, kan mange skader unngås.
0.472224	48	PASSIV RØYKING: Barna kan ikke gå sin vei når de voksne røyker. Nå krever flere nye regler.
0.466962	92	fotovideo.no
0.465653	54	Twitterlisten
0.462123	18	Linn Skåber stiller lettkledd i det nyeste nummeret av Woman.
0.450189	74	Norsk Tipping
0.448522	79	Norsk Rikstoto
0.446899	83	
0.444941	82	Flere oppskrifter på Tine.no
0.420914	72	
0.412581	71	
0.398920	70	
0.374886	94	Premier
0.374298	15	Side2.no
0.358424	86	
0.356698	2	Nettavisen Til forsiden
0.351016	Samme rank ^b	
0.317677	84	Norges største datingtjeneste
0.294366	55	Annonsebilag
0.195341	73	
0.166669	112	

^a 14,17,19,22,24,26,27,29,31,33,35,37,38,40,42,45,47,49,53,60,62

^b 3,10,16,28,41,43,56,57,58,64,65,66,67,75,76,77,88,89,90,97,99,100,101,114,105,110,111

Tabell 11.6: ALT-tekster på www.nettavisen.no

11.4.2.2 Diskusjon

Kendalls Tau distansene i tabell 11.4 indikerer at det også i kategorien nettaviser fremkommer en nesten omvendt rekkefølge i forhold til hvordan IMG-markørene opp-

11.4. TEKST -OG BILDEANALYSER I SAMMENHENG

trer i HTML-koden. Dette er som forventet da nettsider ser ut til å være *topptunge* når det gjelder å plassere pynt på toppen av siden og informasjonsbærende bilder lenger ned. I tilfellet nettaviser har disse ofte reklamebannere og siden egen logo øverst på siden og bilder relatert til artikler.

Tekstene fra ALT-attributtene på www.nettavisen.no i tabell 11.6 tyder på at tekster med mer innhold prioriteres fremfor de med mindre tekst. I de tilfeller hvor mindre tekster får høyere totalrangering antas dette å skyldes de respektive bilders rangering.

Del VI

Konklusjon

Kapittel 12

Konklusjon

Oppgaven tar utgangspunkt i hypotesen om at bilder som har manglende, eller mangelfulle, bildetekster kan utgjøre et forstyrrende element for blinde og svaksynte som studerer innholdet på en nettside med ved hjelp av en skjermleser. Spørsmålet som ønskes bevart er om det er mulig å rangere bilder, og deres bildetekster, etter antatt betydning som informasjonsbærere på en nettside.

Det er også undersøkt om det er mulig å fastsette når et bilde etter denne rangeringen bør inkluderes av skjermlesere og når det bør antas å være forstyrrende at brukeren blir gjort oppmerksom på bildet. For å oppnå dette er bildetekster rangert etter hvor mye relevant informasjon de antas å inneholde og bilder etter deres egenskaper som blikkfang på en nettside.

12.1 Rangere bildetekster innefor en nettside

Oppgaven ser på et sett med seks forskjellige måter å rangere bildeteksters innhold, som diskutert i kapittel 4 og 5. Tekstrangeringene ser på om hvorvidt bildeteksten eksisterer, om den er et tall eller en filsti, om den er repetert flere ganger på samme nettside, hvor mange prosent riktig stavede ord den inneholder og hvor mange ord i teksten som finnes blant et sett med søkeord.

Resultatene i tabell 11.1 i kapittel 11 viser at tre av disse testene måler egenskaper det er vanlig at bildetekster på de utvalgte nettsidene har. Disse egenskapene er at teksten mangler, at samme tekst er repetert to eller flere ganger og at det er over 50% riktig stavede ord i teksten.

Av de tre resterende reangeringene ser det ut til at tall og filstier forekommer mer sjelden på de nettsidene som er undersøkt. Rangeringen etter innhold av søkeord er mer vanskelig å gi en konklusjon på, basert på de resultater som foreligger. Dette fordi den er avhengig av hvilke søkeord som er med. Det antas at denne testen må justeres etter brukers preferanser innenfor en gitt nettside, eller kategori av nettsider, for å kunne påvirke resultatet.

12.2 Finne et bildes posisjon innenfor en nettside

For å rangere et bilde etter dets funksjon som blikkfang kreves kjennskap til bildets posisjon på nettsiden slik den er rendret av den aktuelle nettleseren. Pikselpukk-algoritmen som håndterer dette er diskutert i kapittel 6 og 7.

Resultatene av tester av pikselpukk-algoritmen fra kapittel 7 finnes i kapittel 11. Resultatene indikerer at funnratio i de fleste tilfeller ligger på mellom 80% og 85%. Unntakene er ved 2 dimensjonal nedskalering og når bildet delvis dekkes av noe annet som går fram av grafene i henholdsvis figur 11.6 og 11.7.

Dette er imidlertid antatt å være ekstreme situasjoner som sjelden oppstår på en nettside. Det antas derfor at det er mulig å finne posisjonen til et bilde på en nettside ved å undersøke et skjermbilde av denne nettsiden ved hjelp av denne algoritmen.

12.3 Rangere bilder innenfor en nettside

Å rangere bilder etter egenskaper som gjør det til et blikkfang har vært mer komplekst enn å rangere bildetekster. *At et bilde skal virke som blikkfang på en nettside er et mindre presist definert begrep enn at alle ord i teksten bør være korrekt stavet.*

Noen metrikker for å måle et bilde sin grad av blikkfang har vært utprøvd. Disse er rangering etter størrelse, kompleksitet i innhold og kontrast mot bakgrunnen. Resultatene av disse testene finnes i avsnitt 11.3 og tyder på at dette er målbare størrelser. De foreslåtte metrikker kan derfor antas å gi en indikasjon på et bildes rolle som blikkfang på en nettside.

12.4 Samlet rangering av bilde og bildetekst

I Kapittel 10 ble det fremsatt en hypotese om at bilder på en nettside er plassert slik at de som er antatt minst viktige som informasjonsbærende øverst på siden og de antatt mer viktige lenger ned. Et eksempel på dette er en nettavis med logo, overskrift og reklame øverst og illustrasjonsbilder til artiklene lenger ned.

Resultatene i avsnitt 11.4 kan se ut til å støtte dette. Kendalls Tau rangkorrelasjonskoeffisient mellom de forskjellige rangeringer og IMG-markørene rekkefølge i HTML-koden er i de fleste tilfeller nesten -1.0 , altså nesten det motsatte av hverandre. Dette kan tyde på at hypotesen stemmer for de undersøkte nettsidene og at rangeringen for disse kan gi en pekepinn på hvilke bilder som gir størst blikkfang.

12.5 Konklusjon

Som samlet konklusjon ser det ut til at resultatene er lovende. Dette gjelder både for de individuelle rangeringsalgoritmer i tekst- og bildemodalitet, pikselpukk-algoritmen og

den samlede rangering. Disse algoritmene kan derfor anbefales som en grunnlag for videre arbeide.

Resultatene i tabell 11.1 og 11.6 tyder imidlertid på at innholdet i ALT-attributter fremdels er noe tynt på mange nettsider. Det forutsettes derfor at innholdet i ALT-attributtene på nettsider generelt forbedres i forhold til kravene i WCAG. Dette antas være den største utfordringen for nettopplevelsen til blinde og svaksynte brukere.

12.6 Implikasjoner av konklusjonen

Graden av motivasjon hos nettutviklere til å følge WCAG er fremhevet som den antatt viktigste forutsetningen for at metodikken som er omtalt i denne oppgaven skal ha en fremtid. En viktig forutsetning for å kunne jobbe videre er derfor fortsatt økende fokus på oppfølging av regler for WCAG. Dette er et holdings spørsmål hos både nettutviklere og de organisasjoner de jobber for.

I tekstmodalitet er det rangering basert på nøkkelord som antas ha det største potensiale for videre studier, i form av utarbeidelse av nøkkelordlister for forskjellige kategorier av nettsider. Videre bør det også sees på muligheten for å kategorisere tekster etter innhold og om de forskjellige rangringsalgoritmene bør ha forskjellig vektning i sluttresultatet.

I bildemodalitet bør bildets plassering på nettsiden gis videre oppmerksomhet. Også i bildemodalitet bør det vurderes om de forskjellige rangeringer skal vektet forskjellig.

Bibliografi

- [1] Det Kongelige Sosialdepartement. Nedbygging av funksjonshemmende barrierer. strategier, mål og tiltak i politikken for personer med nedsatt funksjonsevne. Stortingsmelding 40 (2002-2003), 2002.
- [2] Deltasenteret Helsedirektoratet. Tilgjengelige nettsteder, 1:3 oversikt og innhold-sproduksjon, 2006.
- [3] Deltasenteret Helsedirektoratet. Tilgjengelige nettsteder, 2:3 design og koding, 2006.
- [4] Deltasenteret Helsedirektoratet. Tilgjengelige nettsteder, 3:3 universell utform-ing, 2007.
- [5] ETSI. Human factors (hf); guidelines for ict products and services; 'design for all', 2007.
- [6] World Wide Web Consortium. Introduction to wai. <http://www.w3.org/WAI>.
- [7] Arnaud Le Hors Dave Raggett and Ian Jacobs. Html 4.0 specification. <http://www.w3.org/TR/1998/REC-html40-19980424>.
- [8] T. Berners-Lee, L. Masinter, and M. McCahill. Uniform Resource Locators (URL), 1994.
- [9] T. Berners-Lee. HTML 2.0, 1995.
- [10] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext transfer protocol - http/1.1, 1999.
- [11] Jonathan B. Postel. Simple mail transfer protocol - smtp, 1982.
- [12] J. Reynolds. File transfer protocol - ftp, 1985.
- [13] Andrew Tanenbaum. *Computer Networks*. Prentice Hall, 2002.
- [14] Candace Leiden og Marshall Wilensky. *TCP/IP for Dummies*. IDG Books Norge, 2001.

BIBLIOGRAFI

- [15] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, and Stephen Wolff. A brief history of the internet. *SIGCOMM Comput. Commun. Rev.*, 39(5):22–31, 2009.
- [16] Vinton Cerf. How the internet came to be. *The Online User's Encyclopedia*, 1993.
- [17] Vinton Cerf. Specification of internet transmission control program (TCP), 1974.
- [18] Jonathan B. Postel. Internet protocol (IP), 1981.
- [19] Tim Berners-Lee et. al. The world wide web. *Communications of the ACM*, 37(8):76–82, 1994.
- [20] International Organization for Standardization. Information processing, text and office systems, standard generalized markup language (sgml), 1986.
- [21] Pedro DeRose, Warren Shen, Fei Chen, AnHai Doan, and Raghu Ramakrishnan. Building structured web community portals: a top-down, compositional, and incremental approach. In *VLDB '07: Proceedings of the 33rd international conference on Very large data bases*, pages 399–410. VLDB Endowment, 2007.
- [22] Henrike Gappa and Gabriele Nordbrock. Applying web accessibility to internet portals. *Univers. Access Inf. Soc.*, 3(1):80–87, 2004.
- [23] K. Bryan and T. Leise. The 25,000,000,000 eigenvector: The linear algebra behind google. *SIAM Review*, (48), 2006.
- [24] KA. N. Langville and C. D. Meyer. *Google's pagerank and beyond: the science of search engine rankings*. Princeton University Press, 2006.
- [25] Pablo J. Boczkowski. Digitizing the news: Innovation in online newspapers. *Political Communication*, 24:213–214, 2007.
- [26] Lene Tolstrup Sørensen and Knud Erik Skouby. Requirements on next generation social networking—a user's perspective. *Wirel. Pers. Commun.*, 51(4):811–823, 2009.
- [27] Jonathan Lazar et. al. What frustrates screen reader users on the web: A study of 100 blind users. *International journal of human-computer interaction*, pages 247–269, 2007.
- [28] World Wide Web Consortium. User agent accessibility guidelines overview. <http://www.w3.org/WAI/intro/uaag.html>.
- [29] World Wide Web Consortium. Web content accessibility guidelines overview. <http://www.w3.org/WAI/intro/wcag.php>.

- [30] World Wide Web Consortium. Web content accessibility guidelines 1.0. <http://www.w3.org/TR/WCAG10/>.
- [31] World Wide Web Consortium. Web content accessibility guidelines 2.0. <http://www.w3.org/TR/WCAG20/>.
- [32] World Wide Web Consortium. How wcag 2.0 differs from wcag 1.0. <http://www.w3.org/WAI/WCAG20/from10/diff.php>.
- [33] J. Torres I. Alonso D. Harder K. Fischer J. Jiménez, J. Olea. Biography of louis braille and invention of the braille alphabet. *Survey of Ophthalmology*, 54:142–149, 2009.
- [34] Aud Marie Hauge og Frank Fardal. Universell utforming. Master's thesis, Universitetet i Oslo, Institutt for informatikk, 2005.
- [35] Farah Alsudani and Matthew Casey. The effect of aesthetics on web credibility. In *BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction*, pages 512–519. British Computer Society, 2009.
- [36] R.L. Trask and Bill Mayblin. *Introducing Linguistics*. Icon Books Ltd, 2002.
- [37] John R. Hammen. The linguistic reason why the computer will never think. *SIGLASH Newsl.*, 13(4):8–16, 1980.
- [38] George Yule. *The Study of Language*. Cambridge University Press, 2010.
- [39] ed. Leonard Bolc. *Representation and Processing of Natural Language*. Hanser, 1980.
- [40] George F Luger. *Artificial Intelligence Fourth Edition*. Pearson Addison Wesley, 2002.
- [41] J. Weizenbaum. Eliza - a computer program for the study of natural language communication between man and machine. *Comm. A.C.M.*, 9(1):36–45, 1966.
- [42] Chris Brew, Markus Dickinson, and W. Detmar Meurers. language and computers"creating an introduction for a general undergraduate audience. In *TeachNLP '05: Proceedings of the Second ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*, pages 15–22. Association for Computational Linguistics, 2005.
- [43] Tom Brinck and Erik Hofer. Automatically evaluating the usability of web sites. In *Proceedings of ACM CHI 2002 Conference on Human Factors in Computing Systems*, volume 2 of *Workshops*, pages 906–907, 2002.
- [44] Shan Chen, Dan Hong, and Vincent Y. Shen. An experimental study on validation problems with existing html webpages, 2005.

BIBLIOGRAFI

- [45] Håkon Wium Lie and Bert Bos. Cascading style sheets, level 1. <http://www.w3.org/TR/1999/REC-CSS1-19990111>.
- [46] C. Nadine Wathen and Jacquelyn Burkell. Believe it or not: Factors influencing credibility on the web. *JASIST*, 53(2):134–144, 2002.
- [47] Nils Einar Eide, Andreas N. Blaaftadt, Baard H. Rehn Johansen, and Frode Eikka Sandnes. DIGIMIMIR: A tool for rapid situation analysis of helpdesk and support email. In *LISA*, pages 21–32. USENIX, 2004.
- [48] Ben Hammersley. *Developing Feeds with Rss and Atom*. O’Reilly Media, 2005.
- [49] Jenifer Niederst. *Web Design in a Nutshell*. O’REILY, 2001.
- [50] Russel C. Eberhart og Roy W. Dobbins. *Nural Networks PC Tools*. Academic Press, 1990.
- [51] Deborah S. Ray and Eric j. Ray. *Mastering HTML 4 Second Edition*. Sybex, 1999.
- [52] *Raster Graphics Handbook*. Van Nostrand Reinhold Co., 1985. Preface By-Deusen, Edmund Van.
- [53] James Acquah, James Foley, John Sibert, and Patricia Wenner. A conceptual model of raster graphics systems. In *SIGGRAPH ’82: Proceedings of the 9th annual conference on Computer graphics and interactive techniques*, pages 321–328. ACM, 1982.
- [54] James R. Miller. Vector geometry for computer graphics. *IEEE Comput. Graph. Appl.*, 19(3):66–73, 1999.
- [55] Charles M. Eastman. Vector versus raster: A funtional comparison of drawing technologies. *IEEE Computer Graphics and Applications*, 10(5):68–80, 1990.
- [56] George H. Joblove and Donald Greenberg. Color spaces for computer graphics. *Computer Graphics*, 12(3):20–25, 1978.
- [57] William Hansen. *Den elementære farvelære*. Høst & Søns forlag, 1967.
- [58] Gerald M. Murch. Physiological principles for the effective use of color. *IEEE Computer Graphics and Applications*, 4(11):49–54, 1984.
- [59] Melody Y. Ivory and Rodrick Megraw. Evolution of web site design patterns. *ACM Trans. Inf. Syst.*, 23(4):463–497, 2005.
- [60] Thomas Porter and Tom Duff. Composing digital images. *Computer graphics 18*, pages 253–259, 1984.

-
- [61] Kenneth A. Berman and Jerome L. Paul. *Algorithms: Sequential, Paralell, and Distributed*. Thompson Course Technology, 2005.
- [62] A. P. Hillstrom and S. Yantis. Visual motion and attentional capture. *Perception and Psychophysics*, 55:399–411, 1994.
- [63] David C. Lay. *Linear Algebra and its applications, Third Edition Updated*. Addison Wesley, 2006.
- [64] Adam Abrahamsen and David Richards. Image compression using singular value decomposition. *online.redwoods.edu*, 2001.
- [65] H. S. Prasantha, H. L. Shashidhara, and K. N. Balasubramanya Murthy. Image compression using svd. In *ICCIMA '07: Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, pages 143–145. IEEE Computer Society, 2007.
- [66] William Horton. *The Icon Book: Visual Symbols for Computer Systems and Documentation*. John Wiley & Sons, 1994.
- [67] Kathryn Toyer. *Learn Internet Relay Chat*. Wordware, 1997.
- [68] Geore Buscher, Edward Cutrell, and Meredith Ringel Morris. What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, pages 21–30. ACM, 2009.
- [69] A.P. Stakhov. The generalized principle of the golden section and its applications in mathematics, science, and engineering. *Chaos, Solitons and Fractals*, 26(2):263 – 289, 2005.
- [70] A. Naiman. Colour spaces and colour contrast. In *Graphics Interface '85 Proceedings*, pages 313–320, 1985.
- [71] M. Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.
- [72] C. Spearman. The proof and measurement of association between two things. *Amer. J. Psychol.*, 15, 1904.
- [73] Brian P. Flannery Saul A. Teukolsky William H. Press, William H. Press and William T. Vetterling. *Numerical Recipes in Fortran 77, Second Edition*. Cambridge University Press, 1996.