

En introduksjon til bildesegmentering ved bruk av RGB-verdier

Elin Finstad

Institutt for informatikk

Universitetet i Oslo

Norge

November 28, 2020

I. INTRODUKSJON

Bildebehandling er en metode for å utføre oppgaver på bilder, blant annet for å forbedre bildekvaliteten eller for å gjenkjenne fysiske objekter i bildet. Det brukes gjerne for å gjøre videre analyser mindre tidkrevende. Dersom målet er å gjenkjenne fysiske objekter i bildet kan en bildegmentering begrense arbeidskapasiteten ved senere analyser. En slik segmentering består i å dele inn bilder i meningsfulle og/eller perceptuelle uniforme regioner, slik at det er mindre området i bildet som må gjennomsøkes i de videre analysene.

Bildegmentering kan gjennomføres på flere ulike måter, og jeg vil i dette prosjektet fokusere på å segmentere enkeltpiksler basert på farge og intensitet. Ved hjelp av en minimum feilrate-klassifikator skal jeg basert på RGB-verdien til hvert enkelt piksel bestemme hvilken klasse pikselet tilhører. Dette blir så visualisert i et segmentert bilde. Analysen blir gjennomført på både originale RGB-verdier og normaliserte RGB-verdier.

II. METODE

Metoden er tredelt, hvor den første delen består i å velge ut egnede regioner til treningsdata for hver av klassene. Deretter trenes det opp en klassifikator, som så brukes til å tilordne alle pikslene i et testbilde til en av de gitte klassene. Sistnevnte kan visualiseres ved hjelp av segmentering.

Treningssamplene velges ved å hente ut egnede rektangulære regioner av passe størrelse i bildet for hver av de ønskede klassene. Pikslene i disse områdene vil da være treningssamler for den gitt klassen. Når treningssamplene er hentet ut brukes disse til å trenere opp en klassifikator. I denne oppgaven er en minimum feilrate-klassifikator med normalfordelingsantakelse benyttet.

Når klassifikatoren har tilordnet alle piksler i både treningsbilde og et liknende testbilde, kan resultatet konverteres til et fargekodet bilderesultat. Hver klasse blir tilordnet en farge, og de tilhørende pikslene blir da

gitt denne fargen. På denne måten kan man enkelt visualisere hvor godt klassifikatoren har gjenkjent objektene i bildet. Denne metoden kalles segmentering.

Segmenteringen kan påvirkes dersom det er stor variasjon i intensitet i bildet. Dette kan være f.eks. lys/skygge eller refleksjoner. For å motvirke slike variasjoner kan en løsning være å normalisere RGB-verdiene. I denne oppgaven har følgende transformasjon blitt brukt

$$t_1 = \frac{R}{R + G + B},$$

$$t_2 = \frac{G}{R + G + B},$$

$$t_3 = \frac{B}{R + G + B}.$$

Ved en slik transformasjon blir de tre egenskapene lineært avhengige, ettersom en egenskap kan skrives som en lineærkombinasjon av de to andre. Dette fører til at minimum feilrate-klassifikatoren feiler, ettersom determinanten til kovariansmatrisene blir null, og de da ikke har en invers. Ved normalisering er det derfor tilstrekkelig å bruke kun to egenskaper til trening.

I dette prosjektet er det benyttet tre eksempelbilder Figures 1 to 3. Figure 1 brukes som både treningsbilde og testbilde, mens Figure 2 brukes som treningsbilde med Figure 3 som testbilde.

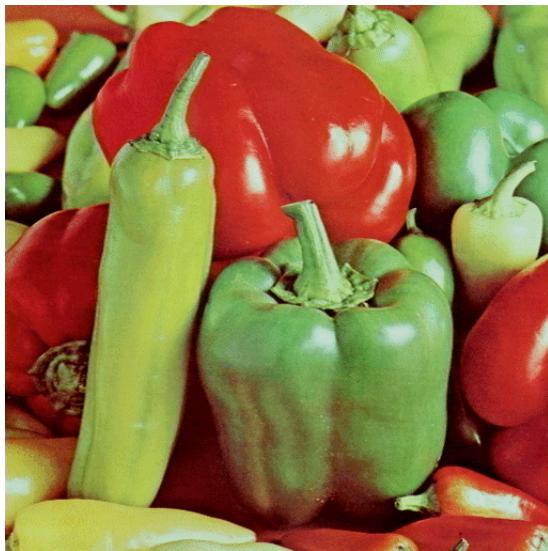


Figure 1. Eksempelbilde brukt til både trening og test for å skille rød paprika, grønn paprika og grønn chili fra hverandre.



Figure 2. Eksempelbilde brukt til trening for å skille de to ringpermene og gulvet fra hverandre.



Figure 3. Eksempelbilde brukt til testing av klassifikatoren trent på Figure 2.

III. RESULTATER

Starter med å se på Figure 1, som er brukt som både trenings- og testbilde. Henter først ut egnede treningsampler for hver av de ønskede klassene, som i dette tilfeller er rød paprika, grønn paprika og grønn chili..

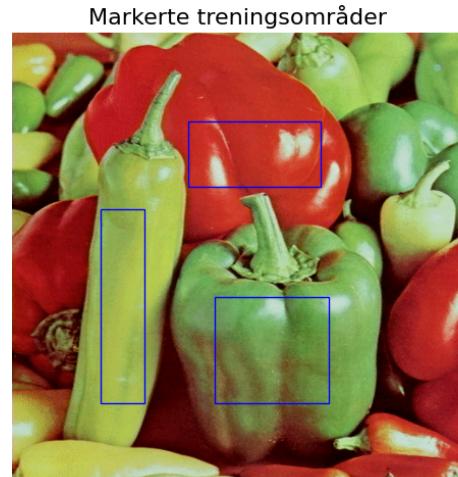


Figure 4. Områdene bruk til trening av de tre klassene markert med blå rektangler.

De samme treningsområdene er brukt på det normaliserte bildet også, hvor det normaliserte bildet ser ut som i Figure 5.

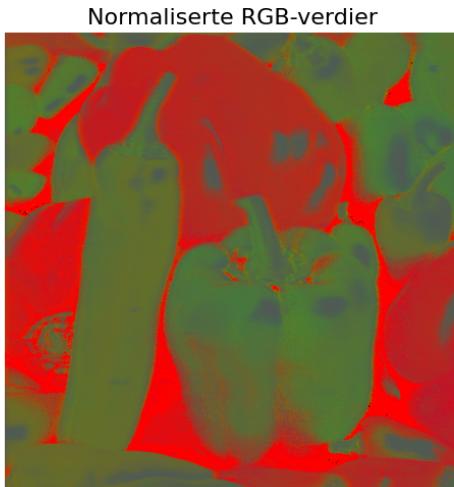


Figure 5. Figure 1 med normaliserte RGB-verdier.

Bruker så minimum feilrate-klassifikatoren til å klassifisere RGB-verdiene i bildet til en av de tre klassene og visualiserer dette ved segmentering i Figure 6. Rød paprika, grønn paprika og grønn chili er vist i henholdsvis rosa, grønn og oransje.



Figure 6. Bilde1 segmentert basert på originale RGB-verdier.

Deretter trenes klassifikatoren på normalisert RGB-verdier. De tre klassene er representert med de samme fargene som tidligere, og resultatet er vist i Figure 7.



Figure 7. Bilde1 segmentert basert på normaliserte RGB-verdier.

Videre går jeg over til å se på Figure 2 og Figure 3. Henter ut egnede treningssamler til hver klasse fra Figure 2 til å tren opp minimum feilrate-klassifikatoren, som deretter brukes til å segmentere Figure 3. De ulike klassene er her blå ringperm, rød ringperm og golv. Her har jeg valgt å ikke ha med metalldelen av ringpermene ettersom målet er å skille de to ringpermene basert på farge, mens metallet er felles for dem begge.



Figure 8. Områdene bruk til trening av de tre klassene markert med svarte rektangler.

Gjennomfører først en segmentering av Figure 2 med originale RGB-verdier og deretter med normaliserte RGB-verdier. Resultatet er vist i henholdsvis Figure 9 og Figure 10.

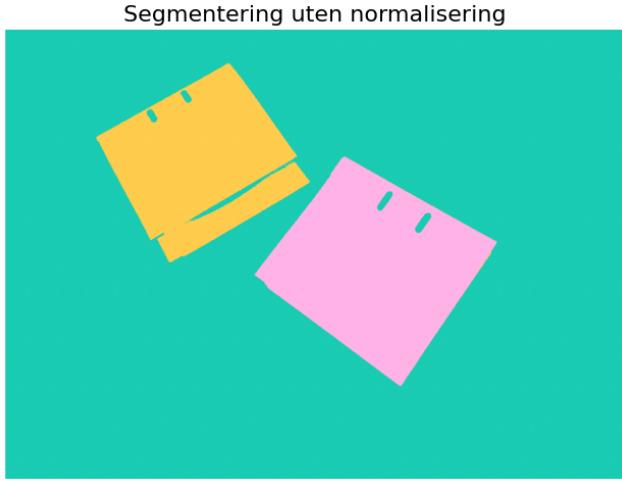


Figure 9. Figure 3 segmentert på originale RGB-verdier.

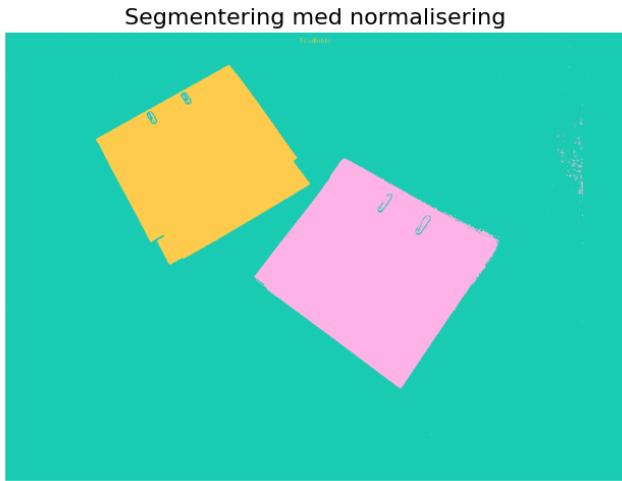


Figure 10. Figure 3 segmentert basert på normaliserte RGB-verdier.

IV. DISKUSJON

Segmenteringen av Figure 1 vist i Figure 6 og Figure 7 viser at normalisering av RGB-verdiene utgjør stor forskjell i resultatet. I Figure 6 blir skyggene i bildet klassifisert som grønn paprika, mens i Figure 7 blir skyggene i bildet klassifisert som rød paprika. Ved å se på originalbildet Figure 1 er det mulig å se at skyggene ligger

nærmere den grønne paprikaen i farge, mens skyggen har fått en betydelig mer rødtone i det normaliserte bilde Figure 5. Skyggene skaper størst utfordring i dette bildet ettersom de ikke hører til noen av klassene. Dette kunne alternativt vært løst ved å ha en egen klasse for skygger, men målet er i hovedsak å skille grønnsakene fra hverandre.

For den grønne chilien er noen området bedre klassifisert i bildet uten normalisering, mens andre områder er bedre klassifisert i bildet med normalisering. Det er antakelig vanskeligst å skille mellom den grønne paprikaen og chilien enn mellom den røde paprikaen og de to grønne grønnsakene, ettersom de to grønne ligger nærmere hverandre i RGB-verdier fra før av.

Ser fra Figure 9 og Figure 10 at normalisering av RGB-verdiene har mye mindre å si for bildene med ringpermer sammenliknet med bildet av grønnsaker. Dette skyldes antakelig at det er mindre refleksjoner i bildet, og det er mindre variasjoner i områder med mer lys/skygge. For den blå ringpermen blir resultatet bedre på normalisert data, ettersom det er tydelig fra Figure 3 at det er mer skygge i dette området. For den rød ringpermen utgjør normaliseringen minimalt med forskjell. Resultatet av segmenteringen blir også mye bedre for ringpermene enn grønnsakene ettersom det er større forskjell i RGB-verdiene til disse tre objektene sammenliknet med de tre grønnsakene.

V. KONKLUSJON

Normalisering av RGB-verdiene utgjør stor forskjell i segmenteringen dersom bildet inneholder store variasjoner i blant annet lysstyrke og refleksjoner. Dette gjør at resultatet for klassifiseringen av grønnsakene i Figure 1 varierer mye med normaliserte og ikke-normaliserte RGB-verdier, mens bildet med ringpermer Figure 3 ikke påvirkes i like stor grad.

Spredningen i RGB-verdiene påvirker også segmenteringen. I bildet med ringpermer er det større variasjon i RGB-verdiene til de forskjellige klassene enn det er for bildet med grønnsaker. Der har de to grønne objektene RGB-verdier som likner mer på hverandre, og som gjør dem vanskeligere å skille fra hverandre.

Appendix A: Kode for reproksjon

Koden brukt i dette prosjektet er tilgjengelig på følgende github-konto: <https://github.com/elinf1/TEK5020/tree/main/Prosjektoppgave2>.