

Bildreproduktion (1.1)

TNM097

Carl Melin
carme003

Edvin Nordin
edvno177

Jimmy Cedervall Lamin
jimla401

7 March 2022

Abstrakt

Att skapa bilder uppbyggda från mindre bilder används idag för konst och marknadsföring. Ett mindre område i originalbilden, en cell, bytts ut mot en bild från en databas. Detta ger effekten av att originalbilden är synlig från ett avstånd men när iakttagaren kommer närmare försvinner originalbilden och bilderna från databasen träder fram. Genom att filtrera en databas på 7129 färgbilder med hjälp av k-means clustering och de objektiva kvalitétsmåtten: Delta E och SSIM kan bilden som är mest lik cellen hittas. Den resulterade återskapade bilden jämförs med flera objektiva kvalitétsmått för att bedöma hur lik den är originalbilden. Flera optimeringar kan göras i koden t.ex. K-means clustering kan optimeras för att ge en jämnare spridning och istället för att hitta medelvärdet på varje cell kan den skalas ner till 1 pixel och användas för att jämföras med. Att använda en stor databas, stora bilder, liten cellstorlek och att vikta in strukturen på bilderna mycket resulterar i en bättre återskapad bild men beräkningstiden tar exponentiellt längre tid.

Innehåll

| | | |
|----------|---|-----------|
| 1 | Introduktion | 4 |
| 2 | Metod | 5 |
| 2.1 | Databas filtrering | 5 |
| 2.1.1 | K-means clustering | 5 |
| 2.2 | Bildreproduktion | 6 |
| 2.3 | Kvalitétsmått | 6 |
| 2.3.1 | Delta E / CIELAB / S-CIELAB | 6 |
| 2.3.2 | Signal-brusförhållande (SNR) | 7 |
| 2.3.3 | Strukturell likhetsindexmått (SSIM) | 7 |
| 3 | Resultat | 8 |
| 3.1 | SSIM | 9 |
| 3.2 | Kluster och klusterstorlek | 11 |
| 4 | Diskussion | 14 |
| 4.1 | Databas filtrering | 14 |
| 4.2 | Bildreprodukions algoritmen | 14 |
| 4.3 | Resultat av bilder | 14 |
| 4.3.1 | SSIM | 15 |
| 4.3.2 | Kluster och klusterstorlek | 15 |
| 5 | Slutsats | 16 |

1 Introduktion

Den här rapporten är baserat på ett projekt där bilder återskapas från mindre bilder. Denna metod används idag oftast inom marknadsföring och konst för att få fram fler budskap från en bild. Originalbilden syns från ett avstånd men när iakttagaren närmar sig försvinner den och de mindre bilderna, som bygger upp originalbilden, träder fram mer och mer. Projektet använder en databas på 7129 naturbilder som hittades på hemsidan Kaggle.com¹. Originalbilden bryts ner i celler som var och en byts ut av den bilden i databasen som matchar bäst beroende på dess färg och struktur. Databasen filteras till en mindre storlek genom att ta bort bilder som är lika varandra. Flera parametrar kan användare själv välja för att få fram det resultat som önskas. Användaren kan ändra bildernas storlek, antalet bilder i databasen, originalbild, cellstorlek och hur mycket av cellens struktur ska övervägas vid val av bilder från databasen. Projektet utvecklades i helhet med programmet *MATLAB*².

¹<https://www.kaggle.com/theblackmamba31/landscape-image-colorization>

²<https://se.mathworks.com/products/matlab.html>

2 Metod

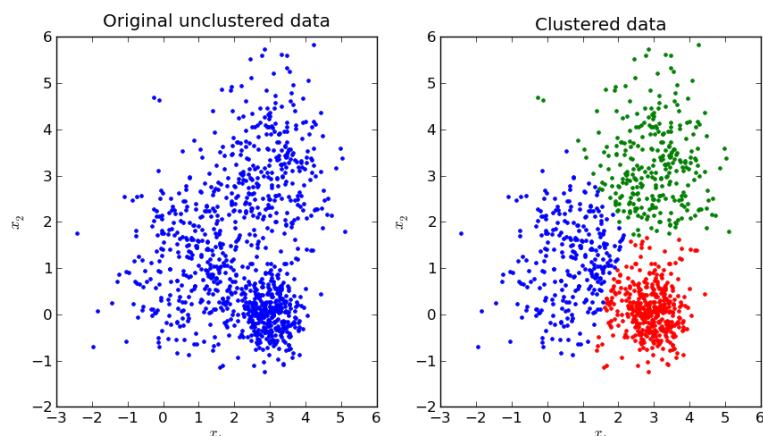
Programmet var uppdelad i två funktioner: databas filtrering och bildreproduktion för att kunna generera flera bilder med samma databas.

2.1 Databas filtrering

Bilderna från databasen laddades in i en cell array och kunde därefter filtreras så de bästa bilderna återstår. Detta gjordes genom att använda *k-means clustering*, se 2.1.1. Bilderna i den nya databasen blev sedan förminskade från 150x150 pixlar för att minska storleken på den återskapade bilden. Antalet bilder och storlek på bilderna väljer användaren innan databasen är filtrerad. Medelvärdet på bildernas färgkanaler är också beräknade för att senare kunna hitta den som passar bäst. En del av bilderna innehåller inte tre kanaler vilket gör att de inte går att användas. Dessa tas bort och gör att den mindre databasen kan innehålla lite färre bilder än vad användaren specificerat.

2.1.1 K-means clustering

K-means clustering är en metod för att dela upp data i mindre grupper. Detta görs genom att dela upp databasen i kluster och iterativt uppdatera klusterna för att få en så jämn spridning som möjligt. K-means väljer en slumpad ursprunglig punkt som den gör den första klustern runt om. Detta gör att metoden ger olika resultat varje gång den körs. Ett exempel på kmeans clustering syns i figur 1 som visar hur data delas upp i tre olika kluster. För bilder så används färgkanalerna för att hitta vilka bilder som är mest lika. Databasen tar ett antal bilder från varje kluster som sedan används som den nya filtrerade databasen. Användaren väljer antalet bilder som databasen ska innehålla genom att specificerar hur många kluster och antalet bilder från varje kluster.



Figur 1: Exempel av k-means clustering.

2.2 Bildreproduktion

När databasbilderna blivit reducerade till ett mer hanterbart antal påbörjades implementationen av reproduktionen. I början av reproduktionen så delas originalbilden upp i mindre bilder. Dessa bilder sparas som celler för att lättare kunna hanteras i kod. Antalet små bilder som tas ut från originalbilden kan väljas av användaren utifrån en parameter. Parametern avgör hur stort MxM område som värden ska sampelas från. Om originalbilden inte är delbar på M i bredd eller höjd kommer bilden ändras i storlek så bredd samt höjd blir delbart på M. Detta resulterar i att lite information från originalbilden kommer att förloras, användaren blir notifierad av detta i ett meddelande. När hela originalbilden har delats upp i mindre bilder går samtliga bilder igenom för att bestämma utifrån den reducerade-databasen vilka bilder enligt Delta E 2 som är mest lik, dessa kvalitétsmått beskrivs mer detaljerat nedan i sektion 2.3. Utifrån de delta E värden som beräknas tas N antal bilder fram och sparas. Användaren får även här specificera i parametern N hur många bilder som önskas. När N bilder tagits fram går samtliga bilder igenom ytterligare en gång för att med hjälp av SSIM 3.1 avgöra vilken av dessa N bilder som passar bäst.

Efter en bild från den reducerade-databasen har valts ut läggs denna till i en stor duk som kommer innehålla samtliga 'bra/optimala' bilder och tillslut när alla bilder gätts igenom återskapas originalbilden med mindre bilder och därmed slutföra bildreproduktionen av originalbilden. Storleken på denna kanva (den reproducerade bilden) kommer bero på parametrar valda av användaren. Dessa parametrar inkluderar storleken på en bild från databasen, det vill säga hur stora dessa små bilder som ska återskapa originalbilden ska vara samt storleken på M. Den reproducerade bilden kommer alltså ha en skala X gånger större än originalbilden som kan beräknas enligt ekvation 1.

$$skala = \frac{bildstorlek}{cellstorlek} \quad (1)$$

2.3 Kvalitétsmått

Kvalitétsmått används för att säkerställa att de resulterade bilderna liknar originalbilden. Fyra kvalitétsmått har använts i detta projekt som nedan beskrivs i respektive sektion.

2.3.1 Delta E / CIELAB / S-CIELAB

Delta E beräknar färgskillnaden hos två bilder genom att beräkna det euklidiska avståndet. Med en RGB bild kan avståndet beräknas enligt ekvation 2.

$$\Delta E = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2} \quad (2)$$

Eftersom beräkningarna ovan inte är perceptuellt enhetsenhetliga som avsett så kan

beräkningarna istället göras i CIELAB-utrymmet enligt ekvation 3

$$\Delta E_{ab} = \sqrt{(L_2 - L_1)^2 + (a_2 - a_1)^2 + (b_2 - b_1)^2} \quad (3)$$

S-CIELAB är en förlängning till ovannämnda CIELAB / Delta E beräkningarna. Att inkludera ett rumsligt bearbetningssteg, innan CIELAB / Delta E beräknas, leder till bättre resultat och motsvarar den färgskillnadsupplevelsen av det mänskliga ögat. För att beräkna S-CIELAB krävs fler parametrar som avstånd till skärmen, skärmens *PPI* (*pixels per inch*) samt en vitpunktsljuskälla.

2.3.2 Signal-brusförhållande (SNR)

Signal-brusförhållandet är en term som definierar det förhållande som finns mellan en signal- och bruseffekt och beräknas enligt ekvation 4.

$$SNR = 10\log_{10} \frac{P_{signal}}{P_{brus}} \quad (4)$$

2.3.3 Strukturell likhetsindexmått (SSIM)

Detta mått används för att mäta likheten mellan två bilder utifrån den upplevda kvaliteten på exempelvis digitala bilder, videos eller digital-tv. Kortfattat beräknar SSIM den strukturella skillnaden hos bilder utifrån tanken av att pixlarna har starka ömsesidiga beroenden, särskilt när de är rumsligt nära, vars information förklrar strukturen hos objektet i den visuella scenen.³

SSIM skiljer sig från andra mätningar som t.ex. SNR vars tillvägagångssätt uppskattar det absoluta felet. Det vill säga att approximationen som beräknas inte ger någon information om felets tecken (om det är positivt eller negativt). Eftersom absolutfelet i en approximation är absolutbeloppet av differensen mellan det exakta värdet och det approximerade.

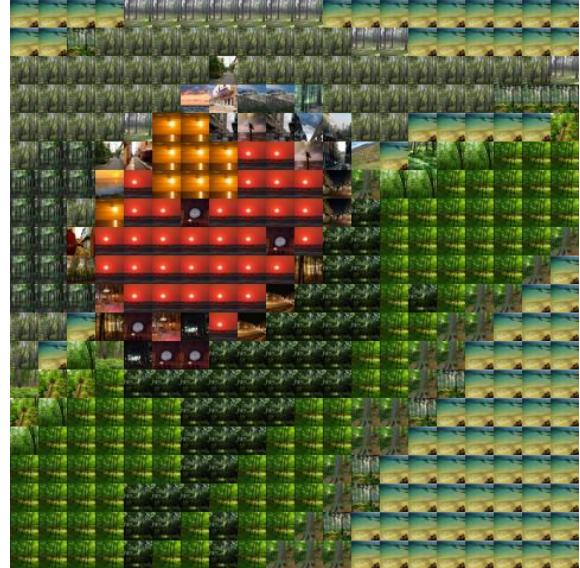
³<https://wikipedia.net/sv/SSIM>

3 Resultat

Resultaten från projektet visas i figurerna nedan. Det första exemplet är nyckelpigan som visas i figur 3 och kan jämföras med originalbilden i figur 2.



Figur 2: Originalbilden av en nyckelpiga.



Figur 3: Återskapad bild av databasbilder.

Figuren demonstrerar hur bilderna från databasen bygger upp den återskapade bilden. Vid närmare inspektion kan de mindre bilderna från databasen tydas och med ett steg bak går det att tyda originalbilden.

Då nyckelpigan kan vara svår att tyda genererades en ny bild med mindre celler vilket gör att den återskapade bilden byggs upp av flera databasbilder. Detta resultat visas i figur 5 som kan jämföras med originalbilden i figur 4.



Figur 4: Originalbilden av ett rådjur.



Figur 5: Återskapad bild med hög kvalitet.

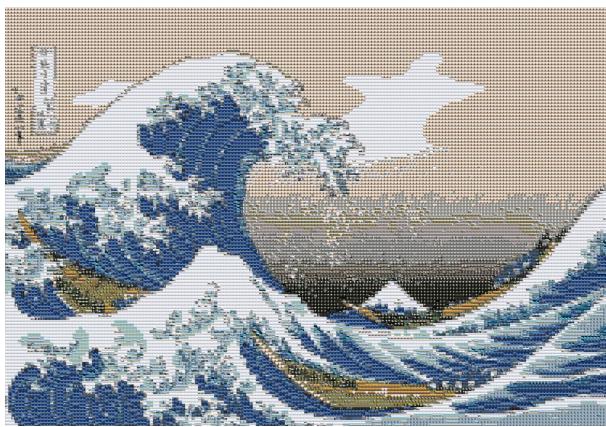
Då det är tydligare att se originalbilden i den återskapade bilden går det inte riktigt att se bilderna från databasen längre utan zoom. Utöver detta tar det markant längre tid att generera fram återskapningen jämfört med den återskapade bilden i figur 3.

3.1 SSIM

Nedan presenteras resultaten från olika N-värden på kvalitetsmåttet SSIM. Som tidigare nämnts är N-värdet antalet optimala bilder enligt deltaE som SSIM väljer mellan. Till exempel N = 3 betyder att SSIM har 3 bilder att välja mellan medan N=1 betyder att SSIM bara har en enstaka bild att välja mellan, med andra ord används inte SSIM då. Resultatet på SSIM kan ses i figurerna 7, 8, 9 samt 10 tillsammans med korresponderande kvalitetsvärden. Dessa figurer kan jämföras med originalbilden i figur 6.

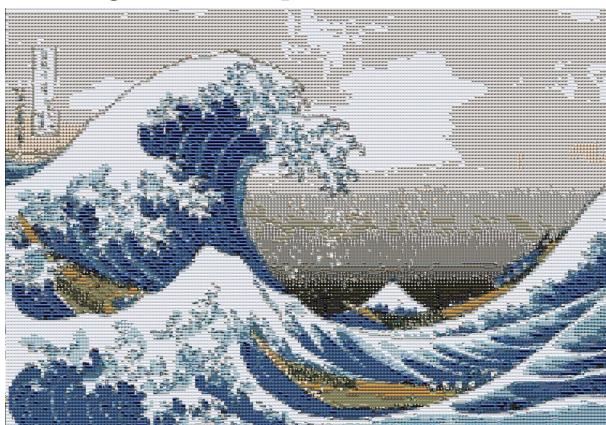


Figur 6: Originalbild av Hokusai vågen



SSIM:0.1571
SNR:11.1347
DELTA E:20.6032
DELTA E (scielab):1.6061
Elapsed time:2-min

Figur 7: SSIM påverkan med N:1.



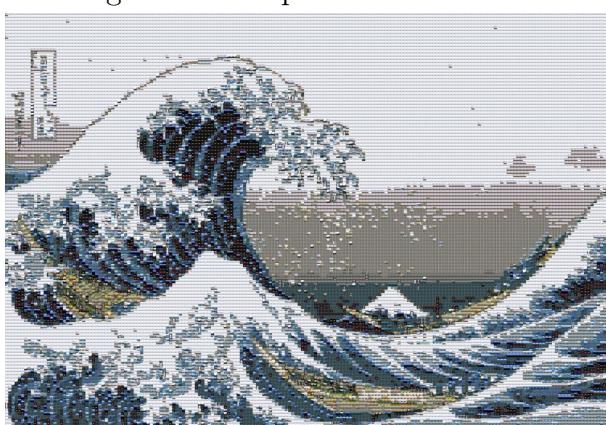
SSIM:0.2122
SNR:11.6876
DELTA E:20.4382
DELTA E (scielab):1.8127
Elapsed time:9-min

Figur 8: SSIM påverkan med N:3.



SSIM:0.2529
SNR:12.8406
DELTA E:20.7563
DELTA E (scielab):1.8485
Elapsed time:11-min

Figur 9: SSIM påverkan med N:7.



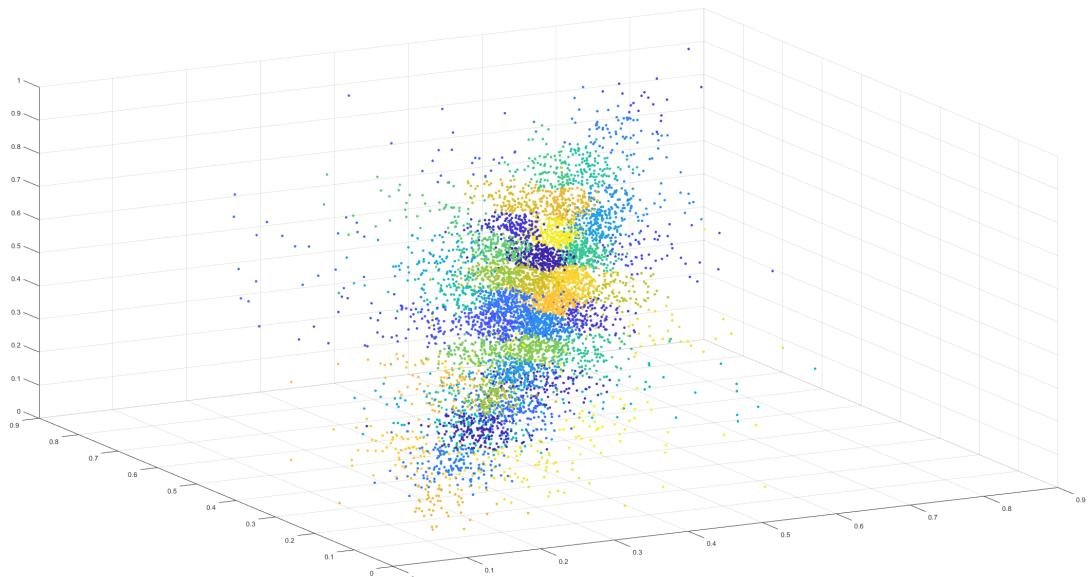
SSIM:0.2953
SNR:13.2773
DELTA E:21.4327
DELTA E (scielab):1.9802
Elapsed time:100-min

Figur 10: SSIM påverkan med N:20.

I och med att SSIMs påverkan på den återskapade bilden ökar, ökar även liknelsen mellan återskapad och original enligt kvalitetsmåtten vid respektive bild. Däremot tar genereringen mer och mer tid desto fler bilder SSIM behöver välja mellan.

3.2 Kluster och klusterstorlek

Då projektet tillåter användaren att välja antal kluster och klusterstorlek att dela upp bilderna från databasen, kan antalet unika bilder som bygger upp den återskapade bilden variera. I figur 11 visualiseras sådana kluster där varje punkt är en bild från databasen. Skillnaden mellan återskapade bilder med olika antal unika databasbilder visas nedan i figur 13, 14 samt 15 och kan jämföras med originalet i figur 12.



Figur 11: Alla bilder indelade i 50 kluster.



Figur 12: Originalbild.



Figur 13: 100st unika bilder.



Figur 14: 200st unika bilder.



Figur 15: 400st unika bilder.

I dessa figurer visas att antalet unika bilder ändrar resultatet. Den största skillnaden är mellan 100 unika bilder jämfört med 200 unika bilder medan skillnaden mellan 200 och 400 unika bilder relativt liten.

I figur 17 och 18 visas skillnaden mellan en återskapad bild med 500 kluster och 5 i klusterstorlek jämfört med en återskapad bild med 5 kluster och 500 i klusterstorlek. Originalbilden som visas i figur 16.



Figur 16: Originalbild av tempel.



SSIM:0.1337
SNR:9.6906
DELTA E:25.9560
DELTA E (scielab):2.0169
Elapsed time:55-sec

Figur 17: 500 kluster med 5 klusterstorlek.



SSIM:0.1576
SNR:10.1804
DELTA E:23.2274
DELTA E (scielab):1.7391
Elapsed time:3-min

Figur 18: 5 kluster med 500 klusterstorlek.

Kvalitetsmåtten visar att skillnaden mellan dessa inte är stor, dock är den med färre antal kluster och större antal i klusterstorlek bättre. Utöver detta tar bilden i figur 17 längre tid att generera.

4 Diskussion

Nedan tas problem och förbättringspunkter upp i samtliga områden som projektet behandlat.

4.1 Databas filtrering

Filtreringen av databasen kan effektiviseras och göras bättre. Ett sätt att förbättra den är ta bort de bilder som inte har r-, g- och b-kanaler medan de importeras in i *MATLAB*. Detta skulle betyda att färre bilder behövs gå igenom och databasen blir alltid så många som användaren specificerat. Detta gör dock endast en liten skillnad då det handlar om runt en procent av databasen som försvisser. Eftersom k-means ger olika output varje gång den körs kan en förbättring vara att den körs fler gånger och endast den bästa outputen används. Koden skulle då vara längsammare men om användaren behöver en väldigt optimerad databas kan det vara värt. Eftersom det beror på vad användaren behöver kan det finnas en variabel som kan ändras för hur många gånger k-means ska itereras.

4.2 Bildreproduktions algoritmen

Den nuvarande algoritmen som används för att återskapa en bild av små bilder fungerar enligt oss bra med avseende på resultat och tiden programmet tar att köra. Det finns många utvecklingsområden som kan göras för att både resultat och effektivitet av kod ska bli betydligt bättre. Till en början skulle vi kunna ändra storleken på bilden vi skickar in så att bredd och höjd matchar antalet ”pixlar / celler” som önskas. Då kan vi istället iterera igenom varje pixel i bilden och hitta respektive bild från databasen. Detta skulle ta bort behovet av att behöva göra beräkningar för olika ’delar’ av originalbilden samt dela upp dessa i små ’celler’. Det finns även många delar i koden som skulle kunna effektiviseras ytterligare genom mindre kopior eller funktioner.

4.3 Resultat av bilder

De resultat som togs fram är stora återskapningar med hög kvalitet där det tydligt går att se likheten till originalbilden. Däremot går det inte att se bilderna från databasen som bygger upp återskapningen utan att zooma-in. Detta beror på att gruppen valde att fokusera mer på likheten med originalbilden, vilket också var en anledning till att använda SSIM tillsammans med Delta E för att generera fram de återskapade bilderna. På grund av de resulterade bildernas stora storlek är det möjligt att artefakter uppstår. Som exempelvis i figur 5 går det att se ett vitt nät över den mörkare delen. Dessa uppstår på grund av att den mörkaste bilden i databasen fortfarande har ljus i sig. Tillsammans med att bilderna är stora kommer detta ljus tydligt synas utmed bilden som ett brus.

4.3.1 SSIM

Som visas i sektion 3.1 har SSIM en hög påverkan i resultaten. Enligt samtliga kvalitetsmått blir likheten mellan original och återskapade bilder högre desto mer vikt SSIM har förutom Delta E måttet, detta beror på att Delta E endast kontrollerar färgvärdet medan SSIM kontrollerar struktur. Ju mer vikt SSIM har i en bild, desto bättre likhet i struktur får resultaten medan färgvalet blir sämre. Detta syns i figurerna 7, 8, 9 och 10 där den övre delen i bilden blir mer och mer olik originalet medan vågen får tydligare struktur ju mer SSIMs påverkan ökas. På grund av att kvalitetsmåtten mäter på olika sätt går det inte att fullt lita på den ena eller den andra. Istället behövs en kompromiss tas alternativt att blanda in det mänskliga tyckandet i ekvationen för att ta fram det bästa resultatet.

4.3.2 Kluster och klusterstorlek

Av det som visas i sektion 3.2 ser vi hur det totala antalet unika bilder påverkar resultatet. Som tidigare nämnts är det större skillnaden mellan 100 och 200 unika bilder jämfört mellan 200 och 400 unika bilder. Utav detta kunde slutsatsen dras att fler unika bilder behöver inte ge ett bättre resultat utan kräver endast mer köringstid och drar ut på genereringstiden. Skillnaden på antalet kluster och klusterstorlek hade inte någon större påverkan på resultatet och därför kunde inte många slutsatser dras. Däremot var den bild med färre kluster och större klusterstorlek lite bättre enligt kvalitetsmåtten vilket kan indikera att vid ökning av klusterstorlek förbättras resultatet något, dock ökar även tiden det tar att generera fram resultatet.

5 Slutsats

Vi har sett att det går att återskapa en bild med flera små bilder samtidigt som motivet i originalbilden tydligt kan synas. Upplösning och storlek på 'cell' spelar stor betydelse på hur väl resultatet speglar originalbilden, där högre upplösning och mindre cell storlek alltid ger bättre resultat. Om en cellstorlek är större än vissa motiv i bilden blir det väldigt svårt för programmet att få med dessa detaljer i resultatet. Om många bilder väljes att sparas beroende av delta E, det vill säga att många bilder kommer jämföras med SSIM, blir resultatet en bild med överlag bättre kontraster. Körningstiden för programmet blir dock exponentiellt längre vilket medför att alltför högt N värde, även om resultatet ibland blir bättre, är inte alltid att önskas.

Vi är nöjda med resultatet av detta projekt och de lärdomar som följt. Skulle vi göra om projektet för framtida syfte är det till stor fördel att ta del av alla punkter som tas upp i diskussion då största problemet med nuvarande program är körningstiden för väldigt stora bilder med högt N värde och en stor databas.