

JPEG XL vs. AVIF

En empirisk jämförelse av moderna bildkomprimeringsalgoritmer med fokus på bildkvalitet

Johan Berg Engstrand

Gusten Berghäll

Kandidatexamensarbete i
Datavetenskap
Höstterminen 2024
Handledare: Tobias Falk
Institutionen för Data-
och Systemvetenskap



Sammanfattning

Introduktion Denna studie undersöker skillnader i bildkvalitet mellan de moderna komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering. Genom att analysera hur dessa algoritmer presterar i tre scenerier – mörka scener, skarpa kanter och detaljrika bilder – syftar studien till att bidra till en bättre förståelse av deras relativt styrkor och svagheter.

Forskningsfråga Den primära forskningsfrågan är: Vilka skillnader finns det gällande bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering? Studien undersöker dessutom tre underfrågor som fokuserar på specifika scenarier.

Metod Studien mäter bildkvalitet hos algoritmerna med hjälp av PSNR och SSIM. För att möjliggöra jämförelser mellan algoritmerna utvärderades deras prestanda i ett antal testfall som representerar de tre nämnda scenarierna. Statistiska analyser genomfördes med Mann-Whitney U-test och deskriptiv statistik.

Resultat Resultaten visar att JPEG XL överträffar AVIF i bildkvalitet för mörka scener medan AVIF presterar bättre vid skarpa kanter. För detaljrika scener uppvisade båda algoritmerna jämförbara resultat men med viss fördel för AVIF. Studien indikerar att algoritmernas prestanda är kontextberoende och att val av komprimeringsteknik bör anpassas efter användningsområde.

Diskussion Studien belyser behovet av att balansera bildkvalitet och filstorlek beroende på användarens krav. Begränsningar inkluderar antalet testfall och den specifika hårdvara som användes. Vidare forskning bör undersöka fler scenarier och kombinera objektiva mätvärden med subjektiva användarupplevelser för att ge en mer heltäckande bild av algoritmernas prestanda.

Abstract

This study investigates differences in image quality between the modern compression algorithms JPEG XL and AVIF under lossy compression. By analyzing their performance in three scenarios – dark scenes, sharp edges, and detailed images – the study aims to contribute to a better understanding of their relative strengths and weaknesses. The primary research question is: What differences in image quality exist between the compression algorithms JPEG XL and AVIF under lossy compression? Additionally, the study examines three sub-questions focusing on specific scenarios. The study evaluates the image quality of the algorithms using PSNR and SSIM. To facilitate comparisons, their performance was assessed in a number of test cases representing the three scenarios. Statistical analyses were conducted using the Mann-Whitney U test and descriptive statistics. The results indicate that JPEG XL outperforms AVIF in image quality for dark scenes, while AVIF performs better for sharp edges. For detailed images, both algorithms produced comparable results, with a slight advantage for AVIF. The study suggests that the algorithms' performance is context-dependent and that the choice of compression technique should be adapted to the specific use case. The study highlights the need to balance image quality and file size based on user requirements. Limitations include the number of test cases and the specific hardware used. Future research should investigate additional scenarios and combine objective measurements with subjective user experiences to provide a more comprehensive understanding of the algorithms' performance.

,

Förkortningar

Förkortning	Beskrivning
JPEG	Joint Photographic Experts Group
JPEG XL / JXL	En modern version av JPEG med stöd för avancerade funktioner
AVIF	AV1 Image Format
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity Index Measure
MSE	Mean Square Error
AI	Artificiell Intelligens
PNG	Portable Network Graphics
HDR	High Dynamic Range
ISO	International Organization for Standardization

Figurer

2.1	En grovt komprimerad JPEG. [4]	6
2.2	En jämförelse av en inzoomad och grovt komprimerad JPEG XL och den ursprungliga JPEG-bilden. Vänster: JPEG, Höger: JPEG XL. [5]	7
2.3	En jämförelse av en inzoomad och grovt komprimerad AVIF och den ursprungliga JPEG-bilden. Vänster: AVIF, Höger: JPEG. [5]	8
3.1	En bild som används i testfallet <i>Skarpa kanter och hög kontrast</i> [25].	16
3.2	Color banding. [22]	17
3.3	En detaljrik bild. [5]	18
4.1	Histogram över PSNR-värden vid mörka scener.	27
4.2	Histogram över SSIM-värden vid mörka scener.	28
4.3	Mann-Whitney U-testresultat för PSNR vid mörka scener. . .	28
4.4	Mann-Whitney U-testresultat för SSIM vid mörka scener. . .	29
4.5	Histogram över PSNR-värden vid skarpa kanter.	30
4.6	Histogram över SSIM-värden vid skarpa kanter.	31
4.7	Mann-Whitney U-testresultat för PSNR vid skarpa kanter. . .	32
4.8	Mann-Whitney U-testresultat för SSIM vid skarpa kanter. . .	32
4.9	Histogram över PSNR-värden vid detaljrika scener.	33
4.10	Histogram över SSIM-värden vid detaljrika scener.	34
4.11	Mann-Whitney U-testresultat för PSNR vid detaljrika scener.	35
4.12	Mann-Whitney U-testresultat för SSIM vid detaljrika scener.	35

Tabeller

Innehåll

Sammanfattning	iii
Abstract	v
Förkortningar	vii
Figurer	ix
Tabeller	xi
Innehåll	xiii
1 Introduktion	1
1.1 Problem	2
1.2 Frågeställning	2
1.2.1 Underfrågor med tillhörande hypoteser	3
2 Utvidgad bakgrund	5
2.1 Grundläggande komprimeringsteori	5
2.2 Vad är JPEG?	6
2.3 Vad är JPEG XL?	6
2.4 Vad är AVIF?	7
2.5 Bildkvalitet	9
2.6 Tidigare forskning	9
3 Metod	13
3.1 Forskningsstrategi	13
3.1.1 Alternativ forskningsstrategi	14
3.2 Datainsamling	15
3.2.1 Testfall	15
3.2.2 Utfallsbaserad parametermatchning	19
3.2.3 Tillämpning	20

3.2.4	Alternativa datainsamlingsmetoder	21
3.3	Analysmetod	21
3.3.1	Tillämpning	22
3.3.2	Alternativa analysmetoder	23
3.4	Forskningsetiska aspekter	23
4	Resultat	25
4.1	Villkor för Mann-Whitney U-test	25
4.2	Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med mörka scener	26
4.2.1	Deskriptiv statistik	26
4.2.2	Inferentiell statistik	28
4.3	Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med skarpa kanter	29
4.3.1	Deskriptiv statistik	30
4.3.2	Inferentiell statistik	31
4.4	Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med detaljrika scener	32
4.4.1	Deskriptiv statistik	33
4.4.2	Inferentiell statistik	34
5	Diskussion	37
5.1	Slutsatser	37
5.2	Begränsningar	39
5.2.1	Metodval	39
5.2.2	Testfall	40
5.2.3	Begränsande datorutrustning	40
5.2.4	Urval av mätverktyg	40
5.3	Resultatetiska aspekter	41
5.4	Fortsatt forskning	41
Litteratur	45	

1

Introduktion

Utvecklingen för digitala kameror, mobiltelefoner och digitala medier går i sprudlande takt framåt i den digitala tidsåldern. I samma takt som lagringsutrymme blir större blir kameraupplösningen bättre. För att lättare kunna lagra, transportera och ladda bilder måste avvägningar göras i bildkomprimeringen som används, så att balansen mellan filstorleken och bildkvalitet blir optimal. De standardiserade bildkomprimeringsteknikerna som används idag har dock inte hunnit med utvecklingen i samma takt.

Ett relevant exempel är JPEG som är ett av de mest etablerade bildformatet som används idag [27]. JPEG är över 30 år gammalt och har betydande brister i sin förmåga att balansera bildkvalitet med filstorlek. Till exempel kan JPEG XL transkoda JPEG-bilder till JPEG XL-formatet utan någon förlust av bildkvalitet (förlustfri komprimering) och samtidigt garantera en ~20% mindre filstorlek.

JPEG XL och AVIF är två framstående kandidater inom nästa generations bildkomprimeringstekniker. Tack vare förbättringar i prestanda, som bättre komprimeringsgrad och snabbare kodning, framstår de som intressanta alternativ för både industriella och individuella användare. Dessutom stödjer de avancerade funktioner som förlustfri och destruktiv komprimering vilket ger dem ett ännu bredare spektrum av potentiella tillämpningar

Trots alla fördelar är forskningen kring JPEG XL och AVIF fortfarande begränsad. Tidigare studier [11][24] har huvudsakligen fokuserat på aspekter som komprimeringshastighet och filstorlek men har ofta saknat en djupgående analys av bildkvalitet – en aspekt som är avgörande för många användare. Dessutom har dessa studier visat motstridiga resultat vilket gör

det svårt att dra definitiva slutsatser om vilken algoritm som är mest lämpad för specifika användningsområden.

Syftet med denna studie är att bidra till en bättre förståelse av hur JPEG XL och AVIF presterar med avseende på bildkvalitet vid destruktiv komprimering. Genom att fokusera på tre typiska utmaningar för bildkomprimering – mörka scener, skarpa kanter och detaljrika scener – syftar studien till att fylla en kritisk kunskapslucka och komplettera tidigare forskning. På så sätt skapas en samlad förståelse för dessa moderna tekniker vilket gynnar både forskarsamhället och praktiska användare.

1.1 Problem

I takt med att efterfrågan på högkvalitativ bildkomprimering ökar blir det allt viktigare att förstå hur moderna algoritmer presterar med avseende på centrala aspekter såsom bildkvalitet. Tidigare studier från aktörer som Google [24] och Cloudinary [11] har bidragit med insikter kring prestanda, exempelvis avkodningshastighet och komprimeringsgrad, även om resultaten kan uppfattas som motstridiga. Dock har bildkvalitet – en av de viktigaste aspekterna för många användare – ofta hamnat i skymundan. Detta försvarar för både forskare och praktiska användare att välja en lämplig komprimeringsteknik.

Problemet som denna studie adresserar är således avsaknaden av en objektiv jämförelse mellan JPEG XL och AVIF med fokus på bildkvalitet vid destruktiv komprimering. Genom att undersöka dessa algoritmers prestanda i tre olika testfall – mörka scener, skarpa kanter och detaljrika scener – strävar studien efter att fylla en kritisk kunskapslucka och bidra till en bättre förståelse för moderna bildkomprimeringstekniker.

1.2 Frågeställning

För att möta behovet av en djupare förståelse för hur moderna bildkomprimeringsalgoritmer presterar med avseende på bildkvalitet fokuserar denna studie på en empirisk jämförelse mellan JPEG XL och AVIF under kontrollerade förhållanden. Studien bidrar med nya insikter genom att analysera algoritmernas förmåga att bevara bildkvalitet i relation till destruktiv komprimering, en aspekt som tidigare forskning i stor utsträckning har förbisett. Den övergripande frågeställningen formuleras som:

- Vilka skillnader finns det gällande bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering?

För att ge ett nyanserat svar på denna frågeställning undersöks bildkvalitet i tre olika kontexter som representerar typiska utmaningar för bildkomprimering: mörka scener, skarpa kanter och detaljrika scener. Genom att fokusera på dessa områden kan studien analysera algoritmernas styrkor och svagheter på ett systematiskt och empiriskt sätt. Underfrågorna med tillhörande hypoteser specificeras nedan.

1.2.1 Underfrågor med tillhörande hypoteser

Underfråga 1: Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener?

H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener.

H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener.

Underfråga 2: Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter?

H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter.

H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter.

Underfråga 3: Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener?

H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener.

H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener.

2

Utvidgad bakgrund

2.1 Grundläggande komprimeringsteori

Det finns två typer av tekniker vid komprimering av bilder: lossless (förlustfri komprimering) samt lossy (destruktiv komprimering).

Många faktorer kommer i spel när det gäller destruktiv komprimering av bilder. Bland annat nyttjar bildkomprimeringstekniker i allmänhet människans visuella perception för att bortse information. Det märks framförallt i form av människans känslighet till vissa ljusfrekvenser samt dess ignorans vid förändring av små detaljer [17, p. 866]. Vid komprimering av bilder tas därför psykoviella modeller i beaktande för att kunna komprimera mer där det inte tros märkas av.

Förlustfri komprimering av bilder skiljer sig från destruktiv komprimering då all originell data måste kvarstå för att räknas som förlustfritt [17, p. 867]. Det gör att komprimeringen inte sker på bekostnad av detalj och bildtrohet utan att exklusivt matematiska modeller, informationsteori och kodnings scheman används för att paketera samma data på en mindre filstorlek [15, p. 532]. Exempelvis kan en bild som normalt sett tar upp cirka 6,2 MB (antag 1920 x 1080 pixlar x 8 bitar färgdjup per kanal) komprimeras till cirka 5 MB (-20%) utan att bildkvaliteten försämras. Avvägningen som görs är den beräkningsmässiga komplexiteten som behövs för att avkoda bilden vilken kan innebära prestationsmässiga utmaningar för det underliggande systemet.

2.2 Vad är JPEG?

Den vanligaste metoden av bildkomprimering i dagsläget är den originella utgåvan av JPEG och utvecklades av Joint Photographic Experts Group som skapades 1986 och gav ut JPEG 1992 [6]. Det uppskattas att flera miljarder bilder komprimeras med JPEG dagligen [3].

Tillhörande anatomin av JPEG-komprimering är dess blockiga sätt att bilda artefakter samt längden den går för att minska mängden färger som används. Det gör att JPEG har visuellt ett enkelt igenkännbart sätt att komprimera bilder vid i synnerhet grov komprimering. Detta illustreras i Figur 2.1 där de typiska artefakterna som uppstår vid grov JPEG-komprimering visas.



Figur 2.1: En grovt komprimerad JPEG. [4]

2.3 Vad är JPEG XL?

JPEG XL har utvecklats av hobbyister, Google såväl som Cloudinary under 2010-talet [12] och erkändes som en ISO-standard mellan 2021 och 2022. Så sent som juli 2024 släpptes den andra JPEG XL ISO-utgåvan.

Googles webbläsare Google Chrome stödde tillfälligt avläsning av JPEG XL-bilder mellan april 2021 och december 2022. Google anförde bland annat att det saknades ett intresse bland intressenter, att JPEG XL inte erbjuder betydande fördelar jämfört med andra bildformat samt underhållskostnader som anledningar till att de valde att sluta stödja JPEG XL [11]. Idag är Apples Safari den enda etablerade webbläsaren som stödjer avläsning av JPEG XL [8].

JPEG XL har likt JPEG ett blockigt sätt att komprimera (även kallat Ringing Artifacts). Vidare har den också en tydlig söm av brus längst de flesta kanter vid grov komprimering. I Figur 2.2 visas ett exempel på detta. JPEG XL har dock en bättre förmåga att behålla färgtrogenheten än vad JPEG har.



Figur 2.2: En jämförelse av en inzoomad och grovt komprimerad JPEG XL och den ursprungliga JPEG-bilden. Vänster: JPEG, Höger: JPEG XL. [5]

2.4 Vad är AVIF?

AVIF har sitt ursprung i AV1-video codec:en som börjades utvecklas 2015 av techjättar som Google, Mozilla, Cisco och Netflix. Tillsammans med andra

aktörer skapades The Alliance for Open Media, AOMedia, med syftet att skapa en ny video-codec med öppen källkod som kunde komprimera videor effektivare och med bättre videokvalitet än kontemporära codec:s som H.264 och VP9. AV1 blev en succé inom sin respektive domän vilket skapade ett intresse för en eventuell nästa generation bildkomprimeringsteknik. 2018 släpptes den första versionen av AVIF och 2020 blev det en erkänd ISO-standard [10].

Till skillnad från JPEG XL så stödjer alla huvudsakliga webbläsare avläsning av AVIF-bilder [7]. Konsekvent kan publicister använda AVIF-bilder utan att kompromissa tillgänglighet.

AVIF till skillnad från JPEG och JPEG XL har vad som kan bäst beskrivas som ett smetigt eller mjukt sätt att komprimera. Det gör att AVIF komprimering inte erhåller något märkbart brus men att det vid grov komprimering kan ge resultat som påminner om en oljemålning där detaljerna går förlorade i en mer sammanflätad textur. Ett exempel på detta visas i Figur 2.3.



Figur 2.3: En jämförelse av en inzoomad och grovt komprimerad AVIF och den ursprungliga JPEG-bilden. Vänster: AVIF, Höger: JPEG. [5]

2.5 Bildkvalitet

Bildkvalitet är en central aspekt inom bildkomprimering och används ofta som ett mått för att bedöma hur väl en komprimeringsalgoritm kan bevara detaljer och färger i en bild. I denna studie utvärderas bildkvaliteten med hjälp av två vedertagna metoder: PSNR (Peak Signal-to-Noise Ratio) och SSIM (Structural Similarity Index Measure).

PSNR är ett mått på skillnaden mellan den ursprungliga och den komprimerade bilden. Det beräknas baserat på den genomsnittliga kvadratiska felet (MSE) mellan pixlar i de två bilderna. Ju högre PSNR desto mindre skillnad finns mellan de två bilderna, vilket indikerar högre bildkvalitet. PSNR uttrycks i decibel (dB) och används ofta för att jämföra komprimeringsalgoritmer på ett kvantitativt sätt.

SSIM är ett perceptuellt mått som jämför likheter i ljusstyrka, kontrast och struktur mellan två bilder. Till skillnad från PSNR som endast mäter pixelbaserade skillnader är SSIM utformat för att bättre efterlikna mänsklig visuell perception. SSIM-värden ligger mellan -1 och 1, där 1 indikerar att bilderna är identiska.

Genom att använda både PSNR och SSIM kan studien få en mer omfattande bild av hur de två komprimeringsalgoritmerna presterar i olika scenarier.

2.6 Tidigare forskning

Forskningen kring komprimeringstekniker är omfattande och sträcker sig över många områden, från text och data till bilder och video. Generellt handlar komprimering om att minska mängden data som behövs för att representera information vilket är avgörande för att minska lagringskostnader och öka effektiviteten i dataöverföring.

Modernta komprimeringstekniker bygger på matematiska modeller och informationsteori. Gupta et al. [16] ger en översikt av förlustfria komprimeringstekniker och visar hur dessa används inom olika domäner, medan Ketshabetswe et al. [19] diskuterar komprimering för trådlösa sensornätverk där optimering av lagring och överföring är avgörande för systemens prestanda. Inom textkomprimering har Kodituwakku och Amarasinghe [20] jämfört förlustfria tekniker och betonat vikten av algoritmernas balans mellan komprimeringshastighet och precision. Dessa studier visar att valet av komprimeringsmetod påverkas starkt av det specifika användningsområdet och de krav som ställs på datahantering.

Inom bildkomprimering har fokus legat på att balansera bibehållen bildkvalitet och filstorlek. Subramanya [23] och Dhawan [14] erbjuder insikter i tekniker för bildkomprimering och lyfter fram viktiga överväganden för att balansera bibehållen detaljåtergivning och komprimeringsgrad. Subramanya ger en grundläggande översikt över bildkomprimeringstekniker medan Dhawan jämför algoritmer för destruktiv och förlustfri komprimering och betonar vikten av att välja teknik baserat på krav på detaljåtergivning och filstorlek. Tillsammans belyser deras arbete hur valet av komprimeringsteknik bör anpassas efter specifika användningsområden. Vijayvargiya et al. [29] erbjuder en omfattande sammanställning av tekniker för bildkomprimering och deras tillämpningar. Studien framhåller särskilt hur olika algoritmers prestanda varierar beroende på bildinnehållets specifika egenskaper vilket är relevant för att förstå skillnader i effektivitet för olika komprimeringstekniker.

Mer nyligen har AI-baserade tekniker för bildkomprimering fått stort forskningsintresse. Ballé et al. [1] utvecklade en end-to-end optimerad metod för bildkomprimering som använder neurala nätverk för att förbättra komprimeringens effektivitet. Detta visar att bildkomprimering är ett område i snabb utveckling där nya metoder ständigt utmanar traditionella tekniker.

Specifikt för JPEG XL och AVIF har forskning visat på deras styrkor inom olika tillämpningar. Cho et al. [9] föreslog en metod för att förbättra den destruktiva komprimeringen i JPEG XL genom adaptiv blockindelning med särskild framgång för grader och homogena bildsegment. Mandel et al. [21] har jämfört JPEG XL med andra format som WebP och BPG och funnit att JPEG XL är särskilt effektivt för bilder med högre bitdjup, en funktion som blir allt mer relevant i moderna tillämpningar.

För AVIF har Barman och Martini [2] undersökt dess prestanda i jämförelse med äldre format som JPEG samt andra moderna algoritmer. Resultaten visar att AVIF erbjuder betydande fördelar i både 4:2:0 och 4:4:4 färgsubsampling, vilket indikerar dess potential för högkvalitativa bildapplikationer. Dessutom har Tulabandu et al. [26] undersökt optimeringar av AVIF:s encoders och visat att förbättrad kodningseffektivitet kan uppnås utan att kompromissa med bildkvaliteten vilket gör AVIF mer praktiskt användbar för webbaserade och mobila applikationer.

Trots den vetenskapliga forskningen som finns på JPEG XL och AVIF saknas fortfarande oberoende jämförelser av algoritmerna under kontrollerade och likvärdiga förhållanden. Googles rapport [24] framhäver AVIF:s

fördelar vad gäller komprimeringshastighet och filstorlek medan Cloudinary [11] argumenterar för att JPEG XL erbjuder bättre prestanda vid högre kodningshastigheter och högre bildkvalitet. Dessa motstridiga resultat illustrerar behovet av en självständig studie som kan bidra till att ge klarhet i vilken algoritm som erbjuder bättre balans mellan bildkvalitet och komprimeringsgrad under olika scenarier.

3

Metod

Detta kapitel redogör för de metodval som har gjorts för att besvara studiens forskningsfråga och beskriver hur dessa har tillämpats i praktiken. Studien använder experiment som forskningsstrategi med empiriska observationer och mätning som datainsamlingsmetoder samt statistisk analys som metod för dataanalys. Valet av dessa metoder baseras på studiens syfte att objektivt undersöka skillnader i bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprivering.

Kapitlet inkluderar en motivering av valda metoder och hur dessa har utformats för att säkerställa validitet och reliabilitet. Vidare diskuteras alternativa tillvägagångssätt och deras potentiella tillämpbarhet. Avslutningsvis presenteras forskningsetiska överväganden.

3.1 Forskningsstrategi

En forskningsstrategi kan ses som en generell plan för att utföra forskning genom att den guidar forskaren i planeringen, utförandet och utvärdering av studien. Vid valet av forskningsstrategi är det viktigt att överväga hur väl strategin lämpar sig för studiens syfte och frågeställning, hur praktiskt genomförbar den är samt vilka etiska aspekter som påverkar studien [18, pp. 39–40].

Denna studie syftar till att undersöka skillnader i bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprivering. Eftersom forskningsfrågan kräver en objektiv jämförelse av algoritmernas tekniska prestanda har experiment som forskningsstrategi valts. Ett

experiment möjliggör direkta mätningar av algoritmernas prestanda under kontrollerade förhållanden vilket säkerställer att resultaten är reproducera och objektiva [18, p. 40].

Genom att kontrollera och mäta specifika variabler, såsom PSNR, SSIM, komprimeringstid och komprimeringsgrad, kan studien isolera och analysera skillnader mellan algoritmerna. Experiment möjliggör även att externa faktorer, exempelvis skillnader i testmiljö eller bildmaterial, kan minimeras för att säkerställa en rättvis och konsekvent jämförelse.

Experiment bedöms därmed vara den mest lämpliga strategin för att besvara forskningsfrågan, eftersom den ger tillgång till kvantitativ och reproducera data som är relevanta för studiens syfte.

3.1.1 Alternativ forskningsstrategi

Ett möjligt alternativ till experiment hade varit en kartläggning, exempelvis i form av en kvalitativ studie med intervjuer eller enkäter. Kartläggning används ofta för att samla in data om individers uppfattningar och erfarenheter och det hade kunnat bidra med insikter om hur algoritmerna upplevs vid praktisk användning [13, p. 7].

Dock innebär kartläggning ett antal begränsningar i detta sammanhang. Algoritmernas egenskaper är objektiva och mätbara vilket innebär att subjektiva åsikter inte är tillräckliga för att ge ett heltäckande svar på forskningsfrågan. Även om subjektiva åsikter kan ge insyn i algoritmernas praktiska tillämpbarhet blir det svårare att argumentera för resultatets validitet när externa faktorer kan påverka användarens upplevelse och bedömning av algoritmerna. Exempel på externa faktorer kan vara hur algoritmerna har implementerats ur ett användbarhetsperspektiv eller att implementationen påverkar algoritmernas prestanda.

Att förklara ett fenomen, snarare än att genomföra explorativ forskning, kräver enligt kartläggning som forskningsstrategi att ett representativt urval av individer identifieras. Ett sådant urval är nödvändigt för att säkerställa att forskningsresultatet blir tillförlitligt och generaliserbart [13, p. 32]. I denna studie bedöms det dock vara svårt att identifiera ett tillräckligt stort och relevant urval av experter eller användare som kan representera en bred population av algoritmanvändare. Detta begränsar möjligheterna att dra generaliseringar och riskerar att minska validiteten i insamlad data.

Ett ytterligare problem med en kvalitativ ansats och kartläggning som forskningsstrategi för att besvara studiens forskningsfråga är att kvalitativ

forskning ofta betonar forskarens aktiva roll i konstruktionen och tolkningen av data [13, p. 244]. Detta kan påverka både resultatets validitet och tillförlitlighet då forskarens subjektiva tolkningar riskerar introducera bias i hur data definieras och analyseras. Därmed bedöms kartläggning inte vara en optimal strategi för denna studie.

3.2 Datainsamling

Valet av datainsamlingsmetod är en central del av studiens utformning eftersom det påverkar möjligheten att besvara forskningsfrågan och uppnå studiens syfte. I denna studie är målet att undersöka skillnader mellan komprimeringsalgoritmerna JPEG XL och AVIF med avseende på prestanda, såsom komprimeringsgrad, hastighet och bildkvalitet. För att samla in data om dessa egenskaper har empiriska observationer och mätning använts som en del av en experimentbaserad ansats [13, p. 66].

Experiment innebär att forskaren genomför kontrollerade tester där specifika variabler observeras och mäts systematiskt. Mätning i detta sammanhang avser användning av instrument och verktyg för att samla in kvantitativa data, exempelvis komprimeringstid och bildkvalitet, på ett objektivt och reproducerbart sätt. Genom denna metod kan störande faktorer kontrolleras vilket möjliggör en rättvis jämförelse mellan algoritmerna.

I praktiken har empiriska observationer och mätning tillämpats genom att köra komprimeringsalgoritmerna på en uppsättning testbilder. Data samlades in om komprimeringstid, komprimeringsgrad och bildkvalitet, där den senare mättes med hjälp av de etablerade metoderna PSNR och SSIM. Dessa mätvärden registrerades under kontrollerade förhållanden för att minimera externa faktorer som kan påverka resultaten.

Experiment, kombinerad med systematisk mätning, stärker studiens validitet och trovärdighet genom att säkerställa att resultaten är objektiva och reproducerbara. Detta är särskilt viktigt för tekniska och kvantitativa studier där precision och jämförbarhet är centrala.

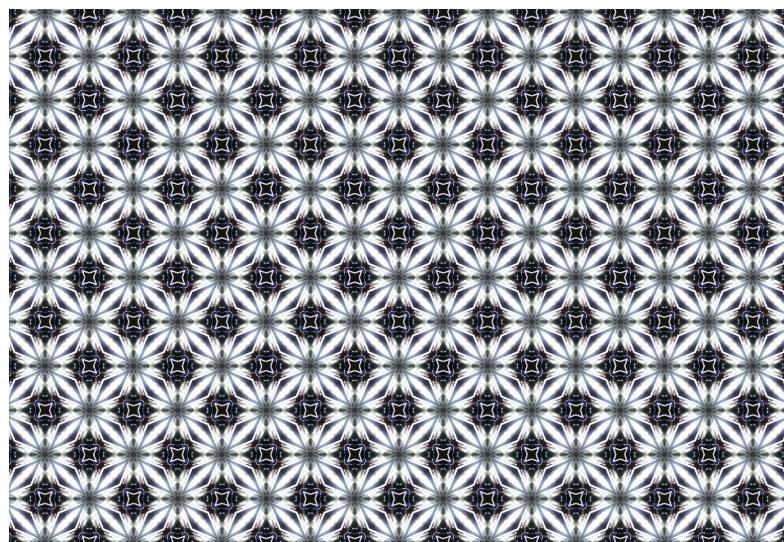
3.2.1 Testfall

För att besvara forskningsfrågan definierades tre specifika testfall, var och en med unika utmaningar för destruktiv bildkomprimering. Till varje testfall hör ett urval av bilder som har valts utifrån specifika kriterier för att både spegla testfallets unika egenskaper och säkerställa relevans för studiens syfte.

Alla testfall baserades på handplockade bilder som licensierades under Creative Commons. Bilderna var i PNG-format och hämtades från källor som Wikimedia Commons och NASA. För varje testfall användes specifika kriterier för att välja bilder, exempelvis genomsnittlig ljusstyrka för mörka scener eller förekomst av repetitiva mönster för skarpa kanter. Alla bilder är baserade på rastergrafik.

Skarpa kanter och hög kontrast

Enligt Jon Sneyers, en expert inom fältet, är ett i synnerhet intressant testfall för destruktiv komprimering tunna linjer med hög kontrast (Personlig kommunikation, 20 november 2024). Se Figur 3.1 för ett exempel. Detta då det finns många skarpa kanter som skapar många snabbt växlande pixelvärden. Eftersom destruktiv komprimering till stort förlitar sig på att (1) gruppera pixlar med snarlika pixelvärden och (2) bortse överflödig information för att minska mängden data skapar det besvärs när (1) kontrasterande kanter inte har snarlika pixelvärden och (2) kanterna är detaljrika. Det gör att destruktiva algoritmer har svårt att komprimera kanter utan att orsaka bi-effekter som artefakter.



Figur 3.1: En bild som används i testfallet *Skarpa kanter och hög kontrast* [25].

Testfallet för skarpa kanter har bilder som blandar repetitiva mönster med varierande pixelvärde. Bilderna varierar mellan enkla mönster i svart och vitt (rutnät och diagonala linjer) och mer komplicerade mönster som kalejdoskop-inspirerade färgglada bilder. Totalt hittades 22 testfall med en

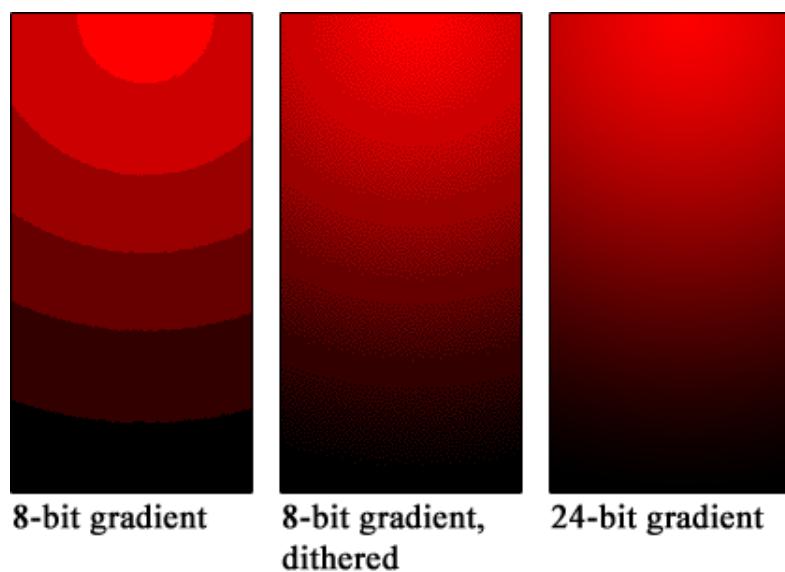
genomsnittlig okomprimerad filstorlek på 1.3 MB. 12 av bilderna är samma 3 testfall i 4 olika storlekar, främst för att se om storleken kan ha en inverkan på komprimeringsförmågan.

Mörka scener

Scener med låga nivåer av ljus kan göra bilder särskilt dunkla och bidra till att subtila detaljer behövs behandlas med omsorg. Eftersom komprimeringstekniker i grunden helst vill bortse från så mycket information som möjligt behöver en hårfin avvägning göras för vilka detaljer som anses viktiga (till exempel nyansen i skuggor) och vilka som är mindre betydelsefulla (till exempel bakgrund).

Mörka bilder kan även ses som mindre visuellt komplexa på grund av en genomgående färgpalett vilket kan leda till att mindre resurser än vad som kan anses som nödvändigt används. I synnerhet mörka scener präglas, vid komprimering, av vad som kallas Color banding vilket är när skiftning av snarlika färger inte är jämn utan att det istället synliggörs tydliga band av färger. Exempel på Color banding illustreras i Figur 3.2.

Avslutningsvis är även mörka scener inom fotografier särskilt benägna att lida av brus vilket ger brus-aspektet ett mer praktiskt förankrat användningsfall än vad datorgenererade superlativa verk har. Hur vardera av teknikerna förhåller sig till att avväga att behålla viktiga detaljer och samtidigt effektivt komprimera mörka scener är även därför ett intressant område att utforska.

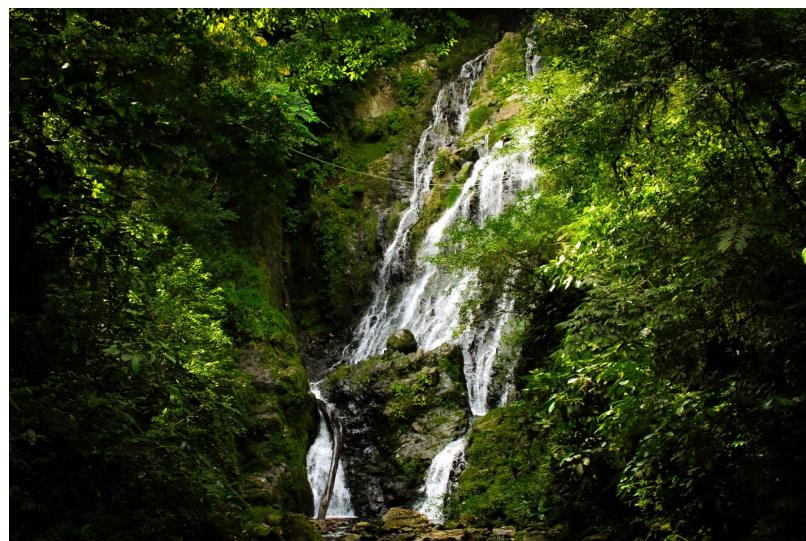


Figur 3.2: Color banding. [22]

För att objektivt bedöma vad som räknas som en mörk scen tillämpas ett program som beräknar om det genomsnittliga ljusstyrkan för varje pixel är under 50, där 0 är inget ljus och 255 är maximalt ljus. Om en bild hade en genomsnittlig ljusstyrka under 50 ansågs det vara en mörk bild. För att mäta ljusstyrkan hos alla bilder skrevs ett Python-program som använde Pillow version v11.0.0 samt NumPy version v2.1.3. Totalt hittades 17 bilder som passade testfallet med en genomsnittlig okomprimerad filstorlek på 37.4 MB.

Detaljrika scener

Det sista avskilda testfallet är detaljerade bilder. I och med att bilder rika på detalj (natur, stadsbilder) har en hög entropi (varierande färgpalett, varierande pixelvärden) utmanas teknikerna i att hålla en god bild-trohet i ett kaosartat sammanhang. Konsekvent krävs det att vardera av teknikerna prioriterar områden med hög detalj effektivt. Syftet är att mäta hur vardera av komprimeringsteknikerna förhåller sig till att behålla detalj gentemot dess komprimeringsgrad i en utmanande miljö. Detaljrika bilder är även ett vanligt användningsfall inom fotografering vilket gör detta testfall speciellt intressant i en pragmatisk kontext. En eventuell nackdel med det här testfallet är att komprimeringstekniker använder sig av psykovisuella modeller för att ta bort information där det är mest troligt att människan inte kommer märka det (se: Utvidgad Bakgrund) vilket PSNR och SSIM inte tar hänsyn till. Exempel på detaljrik bild kan ses i Figur 3.3.



Figur 3.3: En detaljrik bild. [5]

I studien är detaljrika bilder främst fokuserade på högkvalitéfotografier med ett brett utbud av färger, ofta med hjälp av HDR-tekniken. Framför allt har bilderna komplexa texturer (fåglar, skogar) och nyanserade färgskiftnings (himmel, nebulosor). Att kvantifiera vad som innebär med en detaljrik bild är svårt men en gemensam faktor är den exceptionellt stora filstorleken som komplexa scener bidrar med. Totalt hittades 12 lämpliga bilder med en genomsnittlig okomprimerad filstorlek på 68 MB.

3.2.2 Utfallsbaserad parametermatchning

För att jämföra teknikerna behövdes parametrar som hade snarlik komprimeringstid och komprimeringsgrad hittas. I och med att algoritmerna har olika underliggande strukturer för att hantera bildkomprimering innebär det att det inte finns ett sätt att översätta en mängd parametrar från en algoritm till en annan. Det innebär att det behövdes en empirisk utforskning av varsin teknik.

För att göra detta valdes ett urval av varsitt testfall ut som basfall som sedan testades med ~400 olika parametrar. Alla kombinationer mättes och dokumenterades sedan i komprimeringstid samt komprimeringsgrad. Exempel på hur visas i Bilaga A 1. Avslutningsvis valdes tre par kombinationer av parametrar med jämförbara nyckelindikationer för varje testfall. Val av parametrar för varje testfall kan ses i Tabell 3.1.

Det första paret är menat att vara en snabb men icke-precis komprimering och det tredje en långsam men precis komprimering. Det mellanliggande paret är en kompromiss mellan hastighet och precision. Ingen av parametrarna skiljer sig i något testfall mer än 5% i komprimeringstid eller komprimeringsgrad och sannolikt är de mycket närmre varandra än så. För att möjliggöra bäst lämpade par av parametrar finjusterades de slutgiltiga parametrarna manuellt i den slutgiltiga komprimeringen.

KATEGORI	KOMB.	JXL PARAMETRAR	AVIF PARAMETRAR
Mörka	1	--effort 4 --distance 2	-s 6 -q 5
	2	--effort 5 --distance 4	-s 6 -q 50
	3	--effort 7 --distance 4	-s 5 -q 20
Skarpa	1	--effort 1 --distance 11	-s 9 -q 35
	2	--effort 7 --distance 23	-s 3 -q 17
	3	--effort 9 --distance 3	-s 3 -q 40
Detaljrika	1	--effort 5 --distance 10	-s 7 -q 35
	2	--effort 8 --distance 4	-s 2 -q 44
	3	--effort 9 --distance 1	-s 2 -q 74

Tabell 3.1: Parametrar för JPEG XL och AVIF under olika scenarier.

3.2.3 Tillämpning

För att utföra testerna användes en Windows 11 version 22H2 hemdator med en modell AMD Ryzen 3700x processor samt 16 GB DDR4 arbetsminne. Vid komprimeringen användes libjxl version v0.11.0 för JPEG XL och libavif version v1.1.1 för AVIF. Under exekveringen användes inget övrigt krävande program på datorn för att försäkra att alla datorresurser kunde nyttjas av varsin komprimeringsalgoritm. Python version 3.12.7 användes för att bygga tre speciellt anpassade program. Det första programmet sekventiellt kör alla testfall genom kommandotolken. Efter körningen kompileras respektive tekniks körningstid och komprimeringsgrad i ett CSV-dokument, så väl som hur de presterade gentemot varandra i tid (millisekunder) och i komprimeringsgrad (%).

Efter körningen av testfallen kodades alla bilder om förlustfritt från sitt respektive filformat till PNG-filformatet genom ett ytterligare skräddarsytt Python-program. Programmet använde respektive filformats inbyggda avkodare och användes endast för att effektivisera massavkodning av bilder. Detta steg är endast till för att möjliggöra en mätning av de komprimerade bilderna genom det sista Python-programmet. Genom att omvandla alla komprimerade bilder till PNG-bildformatet undviks potentiella kompatibilitetsproblem som kan uppstå vid programmatisk mätning av AVIF och JPEG XL bilder.

Avslutningsvis användes ett sista skräddarsytt Python-program för att mäta de avkodade bilderna inom respektive filformat. Vid mätning av SSIM användes structural_similarity-funktionen ifrån Python-biblioteket scikit-image version v0.24.0. För att mäta PSNR användes NumPy version v2.1.3. Resultaten sammanställdes i två CSV-filer; en för alla bilders individuella PSNR och SSIM resultat och ett baserat på det tidigare CSV-dokumentet för alla testfalls genomsnittliga resultat inom PSNR, SSIM samt skillnad i MSE (en faktor i PSNR). Skillnaden i MSE mellan testfallen beskrevs i procentuell förändring där en negativ procent antyder att JPEG XL har ett bättre MSE-värde och vice versa för AVIF. Om resultatet för PSNR blev oändligt (en perfekt återskapelse, alltså inget brus på grund av komprimering) så beskrevs det genom positivt oändligt (inf) för AVIF och negativt oändligt (-inf) för JPEG XL i kolumnen för MSE skillnaden. Alla resultat i dokumentet avrundades till 3 decimaler genom Pythons round-funktion.

3.2.4 Alternativa datainsamlingsmetoder

Flera alternativa metoder övervägdes för att samla in data om algoritmernas prestanda.

Observation är en metod som innebär att studera och dokumentera fenomen i deras naturliga miljö utan att direkt ingripa i skeendet [13, p. 208]. För denna studie hade metoden kunnat användas för att observera hur individer eller system interagerar med komprimeringsalgoritmerna i praktiska tillämpningar, som webbapplikationer eller bildbehandling. Även om observation kan ge kontextuell förståelse är den inte väl lämpad för att mäta algoritmernas tekniska prestanda vilket är studiens huvudsakliga fokus.

Enkäter och intervjuer har också övervägts som potentiella metoder. Enkäter kan användas för att samla in data från en större population av användare [13, p. 166], såsom utvecklare eller grafiska designers, medan intervjuer kan ge mer djupgående insikter om användarnas upplevelser och åsikter om algoritmernas prestanda [13, p. 184]. Dessa metoder är dock främst inriktade på att fånga subjektiva upplevelser och åsikter vilket gör dem mindre relevanta för denna studie som fokuserar på objektiva och mätbara skillnader.

Sammantaget kan dessa metoder bidra med användbar kontext men bedöms inte vara optimala för att besvara studiens forskningsfråga. Istället krävs en metod som möjliggör insamling av objektiv, reproducerbar data. Detta motiverar valet av empiriska observationer och mätning som datainsamlingsmetod inom ramen för experiment.

3.3 Analysmetod

Dataanalys handlar om att bearbeta och tolka insamlad data för att beskriva eller förklara ett fenomen. Data i sig saknar ofta kontext och behöver därför förberedas, analyseras och presenteras för att slutsatser ska kunna dras. Det finns två huvudtyper av dataanalys: kvantitativ och kvalitativ analys. Kvantitativ analys fokuserar på numeriska data medan kvalitativ analys behandlar kvalitativ data, såsom text, bilder och video [18, p. 61]. Eftersom denna studie fokuserar på att jämföra tekniska egenskaper hos två komprimeringsalgoritmer, krävs en kvantitativ analys som möjliggör en objektiv och systematisk jämförelse baserad på numeriska data.

Kvantitativ analys inkluderar både deskriptiv och inferentiell statistik. Deskriptiv statistik används för att sammanfatta och presentera de insamlade datamängderna medan inferentiell statistik undersöker om skillnader mellan grupper är statistiskt signifikanta. I denna studie valdes Mann-Whitney U-test som inferentiell metod eftersom det är en icke-parametrisk metod som inte ställer krav på att den beroende variabeln är normalfördelad.

För att säkerställa validiteten i Mann-Whitney U-testet genomfördes även Kolmogorov-Smirnov-test för att jämföra fördelningarnas form mellan grupperna. Detta ger en grund för att tolka testresultaten på ett korrekt sätt.

De valda metoderna gör det möjligt att dra slutsatser om skillnader i algoritmernas prestanda på ett transparent och reproducerbart sätt vilket stärker resultatens trovärdighet.

3.3.1 Tillämpning

Underlaget från datainsamlingen importerades till SPSS där både deskriptiv och inferentiell statistik användes för att analysera bildkvalitet för de två algoritmerna, JPEG XL och AVIF. Eftersom studiens syfte var att undersöka skillnader mellan algoritmerna valdes Mann-Whitney U-test som statistisk metod.

Ett vanligt alternativ för att jämföra två grupper är t-test men detta kräver att den beroende variabeln är normalfördelad. Vid en initial granskning av datan observerades att värdena för bildkvalitet, mätta i PSNR och SSIM, inte uppfyllde kravet på normalfördelning. Därför valdes Mann-Whitney U-test som är en icke-parametrisk metod och inte ställer samma krav på datafördelningen.

Mann-Whitney U-test jämför fördelningarna mellan två grupper genom att rangordna alla observationer och beräkna medelrankingens för varje grupp. Resultatet visar om det finns en statistiskt signifikant skillnad mellan grupperna men eftersom testets validitet påverkas av fördelningarnas form genomfördes först en deskriptiv analys. Här analyserades frekvensfördelningarna visuellt med hjälp av histogram och för att statistiskt säkerställa att fördelningarna hade liknande eller olika form användes Kolmogorov-Smirnov-test.

Kolmogorov-Smirnov-testet jämför de kumulativa fördelningarna för två grupper och avgör om skillnaden mellan dem är statistiskt signifikant. Om testet visade att grupperna hade olika fördelningar tolkades resultaten från Mann-Whitney U-testet i termer av medelrank snarare än medianvärdet.

Medelrankningen ger en indikation på vilken grupp som i genomsnitt har högre värden, beroende av fördelningarnas form.

På grund av hur PSNR mäts innebär det att en exakt återskapelse av en bild teoretiskt skulle ge ett oändligt stort PSNR-värde. Under databehandlingen identifierades vissa oändliga värden i data för algoritmerna. För att undvika att dessa påverkar resultaten negativt ersattes de med det högsta observerade värdet i respektive dataset. Denna metod valdes för att behålla datans jämförbarhet samtidigt som extrema värden hanterades på ett rimligt och konsekvent sätt.

3.3.2 Alternativa analysmetoder

Hade kvalitativ data samlats in finns det ett antal vedertagna analysmetoder att tillämpa. Denscombe listar exempelvis innehållsanalys (eng: content analysis) som används för att identifiera dolda mönster eller medde- landen, grundad teori (eng: grounded theory) som syftar till att utveckla nya koncept eller teorier och kritisk diskursanalys (eng: discourse analysis) som belyser hur maktstrukturer uttrycks genom språk [13, p. 282]. Dessa metoder är emellertid inte tillämpliga i denna studie eftersom forskningsfrågan kräver en objektiv jämförelse av tekniska egenskaper snarare än en tolkande eller explorativ ansats.

Som tidigare nämnts är syftet med denna studie att objektivt jämföra komprimeringsalgoritmernas prestanda. För att uppnå detta krävs kvantitativa analysmetoder som möjliggör en systematisk och reproducerbar jämförelse baserad på mätbara data. Den kvantitativa analysen kommer att innehålla både deskriptiv statistik för att sammanfatta och presentera de insamlade datamängderna och inferentiell statistik för att undersöka om skillnader mellan algoritmerna är statistiskt signifikanta. Detta upplägg säkerställer en noggrann och objektiv analys av resultaten vilket gör kvantitativ analys till det naturliga valet för att besvara forskningsfrågan.

3.4 Forskningsetiska aspekter

Enligt God Forskningssed är tillförlitligheten en av de centrala principerna för god forskningssed. Forskare har ett ansvar att säkerställa forskningens kvalitet genom noggrant val av metod, korrekt analys och transparent rapportering av resultat. [28, p. 11]. I denna studie finns en låg men potentiell

risk att spridning av felaktiga resultat kan påverka intressenterna till algoritmerna JPEG XL och AVIF negativt. För att minimera denna risk eftersträvas hög validitet och tillförlitlighet genom en öppen redovisning av metodval och metodtillämpning. Detta möjliggör att eventuella fel kan härledas och förklaras vilket är i linje med principerna för ärlig och ansvarsfull informationsspridning [28, p. 77] och stärker forskningens transparens och trovärdighet [28, pp. 11–12].

Ett centralt skyddsintresse inom forskningsetik är att respektera individens integritet och rättigheter [28, p. 13]. För att undvika onödiga etiska komplikationer och skydda individers personuppgifter har studien valt att utesluta bilder som innehåller mäniskor. Detta beslut möjliggör att forskningsfrågan kan besvaras utan att införa extra åtgärder för att säkerställa integritetsskydd vilket samtidigt förenklar genomförandet. Detta är i linje med principen om att minimera risker och säkerställa en rimlig balans mellan forskningsintresset och skyddsintresset [28, p. 14].

Enligt God forskningssed är tillförlitlighet och ärlighet centrala principer för god forskning vilket inkluderar att all forskning ska genomföras med respekt för immateriella rättigheter och upphovsrätt [28, p. 12]. I denna studie används bilder som är fria från upphovsrätt eller licensierade för fritt användande. Detta säkerställer att inga upphovsrättsliga överträdelser sker och att forskningen bedrivs i enlighet med etiska riktlinjer. Användningen av fritt tillgängligt material är även i linje med principerna om transparens och öppenhet i forskning [28, p. 83]. Genom att välja licensfria bilder minimeras risken för rättsliga komplikationer och potentiella intressekonflikter samtidigt som det säkerställs att forskningen är etiskt försvarbar och reproducerbar.

4

Resultat

Detta avsnitt presenterar resultaten från de statistiska analyser som genomförts för att besvara studiens underfrågor samt huvudfrågan: *Vilka skillnader finns det gällande bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering?* Resultaten redovisas med hjälp av deskriptiv och inferentiell statistik för att belysa eventuella skillnader mellan algoritmerna.

För att använda Mann-Whitney U-test i de analyser som genomförts behöver ett antal villkor vara uppfyllda. Eftersom alla underfrågor analyseras med hjälp av samma typ av test kan majoriteten av villkoren redovisas gemensamt i nästa avsnitt. Det sista villkoret som handlar om huruvida distributionen av data för de båda algoritmerna har samma eller liknande form redovisas under respektive underfråga i samband med den deskriptiva statistiken.

4.1 Villkor för Mann-Whitney U-test

Mann-Whitney U-test är ett icke-parametriskt test som används för att jämföra två oberoende grupper. För att detta test ska vara tillämpligt krävs att följande villkor är uppfyllda:

1. Beroende variabeln (bildkvalitet) ska vara mätt i minst ordinalskala. I denna studie används PSNR och SSIM som mått på bildkvalitet vilka båda är kvotskalemått och därmed uppfyller detta krav.

2. Oberoende variabeln (algoritm) måste bestå av två oberoende grupper. I denna studie jämförs algoritmerna JPEG XL och AVIF vilka representerar två separata och oberoende grupper.
3. Oberoende observationer handlar om att observationerna inom varje grupp måste vara oberoende av varandra. I denna studie analyseras varje bild som en enskild enhet och påverkas inte av andra observationer.

4.2 Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med mörka scener

Den första underfrågan som undersöks är: *Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener?* För att besvara denna fråga testades följande hypoteser:

H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener.

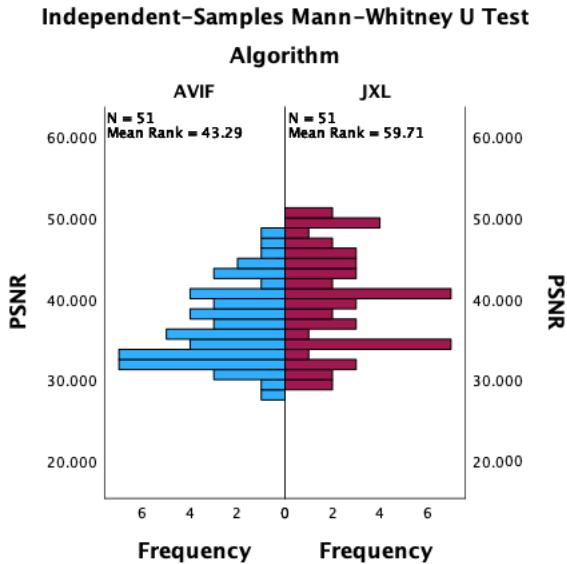
H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid mörka scener.

4.2.1 Deskriptiv statistik

För att analysera och jämföra bildkvalitet för JPEG XL och AVIF vid mörka scener utfördes en deskriptiv analys. Histogram genererades för att visuellt bedöma fördelningarna av PSNR och SSIM-värden vilka representerar bildkvalitet. Syftet var att undersöka om gruppernas distributioner har en liknande form vilket är det sista och fjärde villkoret för Mann-Whitney U-test.

PSNR

Histogrammen i Figur 4.1 jämför frekvensfördelningarna av PSNR mellan JPEG XL och AVIF. Vid en visuell inspektion framgår att distributionerna har olika form där JPEG XL visar en något mer utdragen fördelning mot högre PSNR-värden medan AVIF har en tydligare koncentration i det lägre intervallet. Detta indikerar att algoritmerna beter sig annorlunda vid mörka scener. För att bekräfta denna observation genomfördes ett Kolmogorov-Smirnov-test.



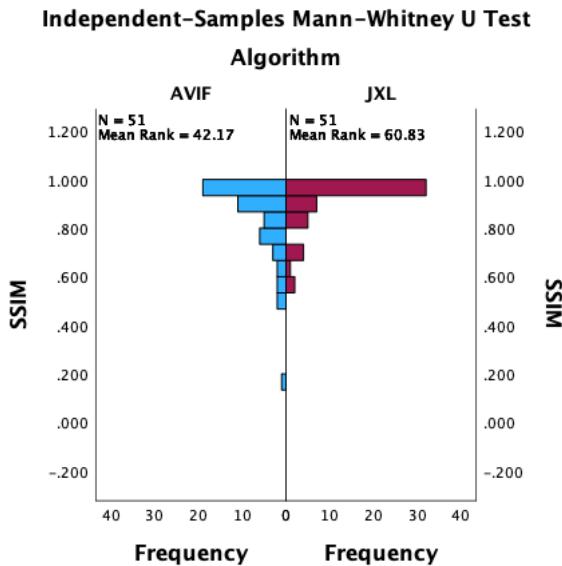
Figur 4.1: Histogram över PSNR-värden vid mörka scener.

Kolmogorov-Smirnov-testet visade en statistiskt signifikant skillnad mellan distributionerna ($D = .294$, $p = .024$). Detta innebär att distributionerna för PSNR mellan de två algoritmerna inte har samma form. Eftersom fördelningarna inte uppfyller kravet på liknande form är en direkt jämförelse av medianvärden inte lämplig. Istället fokuseras analysen på att tolka mean rank från Mann-Whitney U-testet för att identifiera vilken algoritm som presterar bättre i detta scenario.

SSIM

Histogrammen i Figur 4.2 jämför frekvensfördelningarna av SSIM mellan JPEG XL och AVIF. Vid en visuell tolkning framgår att distributionerna har olika form där SSIM-värdena för JPEG XL är mer koncentrerade runt högre värden (nära 1.0) medan AVIF har större spridning mot lägre värden. För att säkerställa denna observation genomfördes ett Kolmogorov-Smirnov-test.

Kolmogorov-Smirnov-testet visade en statistiskt signifikant skillnad mellan distributionerna ($D = .353$, $p = .003$) vilket bekräftar att SSIM-distributionerna för de båda algoritmerna inte har samma form. Eftersom fördelningarna är olika kan en direkt jämförelse av medianvärden inte göras och analysen fokuserar därför på att tolka mean rank från Mann-Whitney U-testet.



Figur 4.2: Histogram över SSIM-värden vid mörka scener.

4.2.2 Inferentiell statistik

För att undersöka skillnader i PSNR- och SSIM-värden mellan JPEG XL och AVIF vid mörka scener genomfördes Mann-Whitney U-test.

PSNR

Ett Mann-Whitney U-test genomfördes för att undersöka skillnader i PSNR mellan JPEG XL och AVIF vid mörka scener. Resultatet, se Figur 4.3, visade en statistiskt signifikant skillnad mellan grupperna ($U = 1719$, $p = .005$). Medelrankningen för JPEG XL var 59.71 medan medelrankningen för AVIF var 43.29 vilket indikerar att JPEG XL presterade bättre i detta scenario.

Independent-Samples Mann-Whitney U Test Summary	
Total N	102
Mann-Whitney U	1719.000
Wilcoxon W	3045.000
Test Statistic	1719.000
Standard Error	149.416
Standardized Test Statistic	2.801
Asymptotic Sig.(2-sided test)	.005

Figur 4.3: Mann-Whitney U-testresultat för PSNR vid mörka scener.

SSIM

Ett Mann-Whitney U-test genomfördes också för att analysera skillnader i SSIM mellan JPEG XL och AVIF vid mörka scener. Resultatet, se Figur 4.4, visade en statistiskt signifikant skillnad mellan grupperna ($U = 1776.5$, $p = .001$). Medelrankningen för JPEG XL var 60.83 medan medelrankningen för AVIF var 42.17 vilket återigen indikerar att JPEG XL presterade bättre i detta scenario.

Independent-Samples Mann-Whitney U Test Summary

Total N	102
Mann-Whitney U	1776.500
Wilcoxon W	3102.500
Test Statistic	1776.500
Standard Error	149.403
Standardized Test Statistic	3.186
Asymptotic Sig.(2-sided test)	.001

Figur 4.4: Mann-Whitney U-testresultat för SSIM vid mörka scener.

Detta innebär att JPEG XL presterar signifikant bättre än AVIF både när det gäller PSNR och SSIM för bilder med mörka scener.

4.3 Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med skarpa kanter

Den andra underfrågan som undersöks är: *Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter?* För att besvara denna fråga testades följande hypoteser:

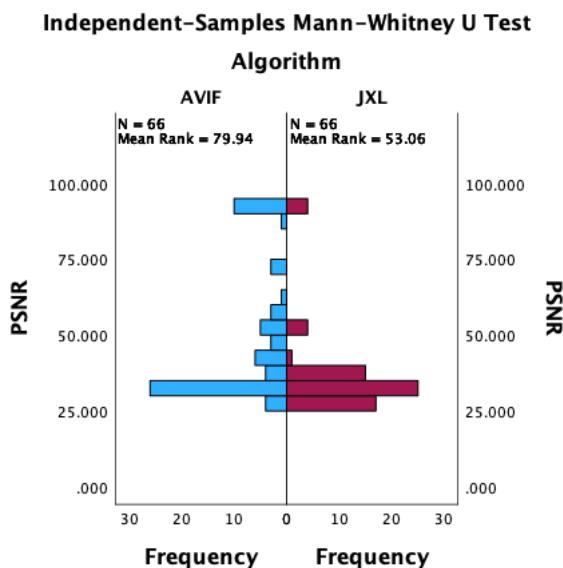
H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter.

H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid skarpa kanter.

4.3.1 Deskriptiv statistik

PSNR

Histogrammen i Figur 4.5 jämför frekvensfördelningarna av PSNR mellan JPEG XL och AVIF. Vid en visuell tolkning framgår att fördelningarna har olika form där AVIF uppvisar en större koncentration av högre PSNR-värden medan JPEG XL är något mer spridd mot lägre värden. För att säkerställa denna observation genomfördes ett Kolmogorov-Smirnov-test.

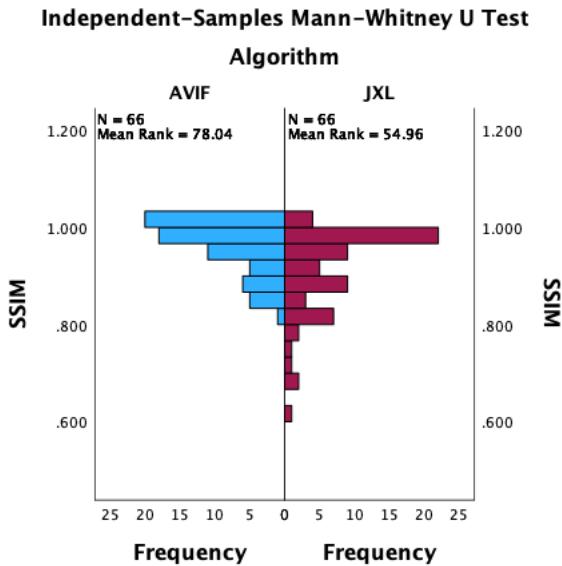


Figur 4.5: Histogram över PSNR-värden vid skarpa kanter.

Kolmogorov-Smirnov-testet visade en statistiskt signifikant skillnad mellan distributionerna ($D = .379$, $p < .001$) vilket innebär att fördelningarna för PSNR mellan algoritmerna inte har samma form. Eftersom distributionerna inte är likformiga fokuserar analysen på att tolka mean rank från Mann-Whitney U-testet.

SSIM

Histogrammen i Figur 4.6 jämför frekvensfördelningarna av SSIM mellan JPEG XL och AVIF. Vid en visuell tolkning framgår att AVIF uppvisar en koncentration av värden nära det maximala SSIM-värdet på 1.0 medan JPEG XL har en något större spridning mot lägre värden. För att verifiera skillnaden i distribution genomfördes ett Kolmogorov-Smirnov-test.



Figur 4.6: Histogram över SSIM-värden vid skarpa kanter.

Kolmogorov-Smirnov-testet visade en statistiskt signifikant skillnad mellan grupperna ($D = .273$, $p = .015$). Detta resultat tyder på att distributionerna för SSIM mellan algoritmerna inte är likformiga vilket innebär att analysen även här fokuserar på tolkningen av mean rank från Mann-Whitney U-testet.

4.3.2 Inferentiell statistik

PSNR

Ett Mann-Whitney U-test genomfördes för att undersöka skillnader i PSNR mellan JPEG XL och AVIF vid skarpa kanter. Resultatet, se Figur 4.7, visade en statistiskt signifikant skillnad mellan grupperna ($U = 1291.000$, $p < .001$). Medelrankningen för AVIF var 79.94 medan den för JPEG XL var 53.06. Detta indikerar att AVIF presterade bättre i detta scenario.

SSIM

Ett Mann-Whitney U-test genomfördes även för att analysera skillnader i SSIM mellan JPEG XL och AVIF vid skarpa kanter. Resultatet, se Figur 4.8, visade en statistiskt signifikant skillnad mellan grupperna ($U = 1416.500$, $p < .001$). Medelrankningen för AVIF var 78.04 medan den för JPEG XL var 54.96 vilket innebär att AVIF även här presterade bättre i detta scenario.

Independent-Samples Mann-Whitney U Test Summary	
Total N	132
Mann-Whitney U	1291.000
Wilcoxon W	3502.000
Test Statistic	1291.000
Standard Error	219.594
Standardized Test Statistic	-4.039
Asymptotic Sig.(2-sided test)	<.001

Figur 4.7: Mann-Whitney U-testresultat för PSNR vid skarpa kanter.

Independent-Samples Mann-Whitney U Test Summary	
Total N	132
Mann-Whitney U	1416.500
Wilcoxon W	3627.500
Test Statistic	1416.500
Standard Error	219.015
Standardized Test Statistic	-3.477
Asymptotic Sig.(2-sided test)	<.001

Figur 4.8: Mann-Whitney U-testresultat för SSIM vid skarpa kanter.

Sammanfattningsvis visar resultaten att AVIF presterar signifikant bättre än JPEG XL både när det gäller PSNR och SSIM för bilder med skarpa kanter.

4.4 Skillnad i bildkvalitet mellan JPEG XL och AVIF vid komprimering av bilder med detaljrika scener

Den tredje underfrågan som undersöks är: *Finns det någon skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener?* För att besvara denna fråga testades följande hypoteser:

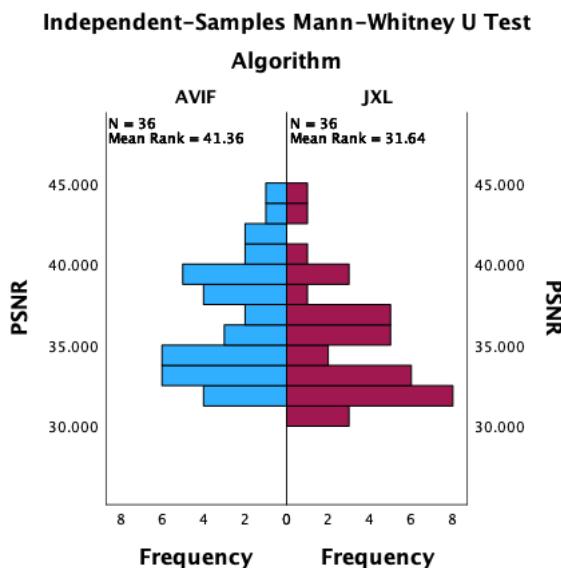
H_0 : Det finns ingen skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener.

H_A : Det finns en skillnad i bildkvalitet mellan JPEG XL och AVIF vid detaljrika scener.

4.4.1 Deskriptiv statistik

PSNR

Histogrammen i Figur 4.9 visar frekvensfördelningarna av PSNR-värden mellan JPEG XL och AVIF för detaljrika scener. Vid en visuell tolkning framgår att distributionerna för de två algoritmerna har en liknande form med en tydlig koncentration runt PSNR-värden mellan 30 och 40 för båda grupperna. Det finns dock viss skillnad i fördelningens spridning där JPEG XL verkar ha något högre spridning.



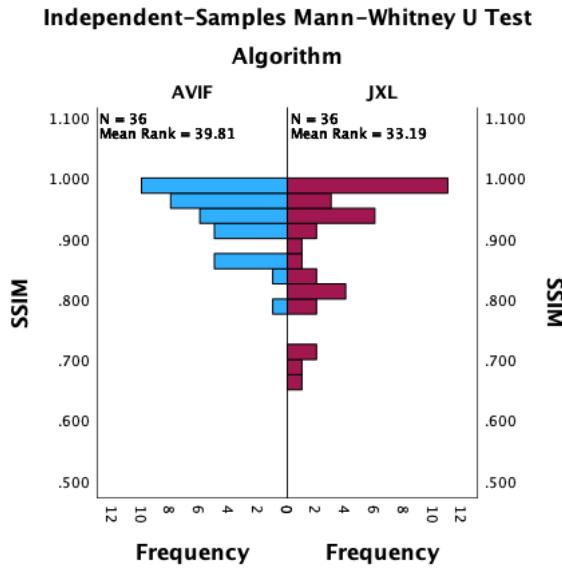
Figur 4.9: Histogram över PSNR-värden vid detaljrika scener.

För att ytterligare undersöka om distributionerna har liknande form genomfördes ett Kolmogorov-Smirnov-test. Resultatet visade ingen statistiskt signifikant skillnad mellan gruppernas fördelningar ($D = .250$, $p = .211$). Detta innebär att distributionerna kan anses ha liknande form vilket gör det möjligt att jämföra medianvärden med hjälp av Mann-Whitney U-test.

SSIM

Histogrammen i Figur 4.10 jämför frekvensfördelningarna av SSIM-värden mellan JPEG XL och AVIF. Vid en visuell tolkning är fördelningarna liknande med en tydlig koncentration av värden nära det maximala SSIM-värdet

1.0 för båda algoritmerna. AVIF har en något högre frekvens av värden närmare 1.0 jämfört med JPEG XL samtidigt som JPEG XL uppvisar en större spridning av värden mot lägre nivåer.



Figur 4.10: Histogram över SSIM-värden vid detaljrika scener.

Kolmogorov-Smirnov-testet bekräftade denna observation genom att inte påvisa någon statistiskt signifikant skillnad mellan distributionerna ($D = .306$, $p = .069$). Detta innebär att distributionerna har liknande form vilket gör det möjligt att jämföra medianvärden mellan grupperna med hjälp av Mann-Whitney U-test.

4.4.2 Inferentiell statistik

PSNR

Ett Mann-Whitney U-test genomfördes för att undersöka skillnader i PSNR mellan JPEG XL och AVIF vid komprimering av bilder med detaljrika scener. Resultatet, se Figur 4.11, visade en statistiskt signifikant skillnad mellan grupperna ($U = 473.000$, $p = .049$). Medelrankningen för AVIF var 41.36 medan den för JPEG XL var 31.64 vilket indikerar att AVIF presterade bättre i detta scenario.

Independent-Samples Mann-Whitney U Test Summary	
Total N	72
Mann-Whitney U	473.000
Wilcoxon W	1139.000
Test Statistic	473.000
Standard Error	88.792
Standardized Test Statistic	-1.971
Asymptotic Sig.(2-sided test)	.049

Figur 4.11: Mann-Whitney U-testresultat för PSNR vid detaljrika scener.

SSIM

Ett Mann-Whitney U-test genomfördes även för att analysera skillnader i SSIM mellan JPEG XL och AVIF för detaljrika scener. Resultatet, se Figur 4.12 visade ingen statistiskt signifikant skillnad mellan grupperna ($U = 529.000$, $p = .180$). Detta innebär att nollhypotesen inte kunde förkastas vilket tyder på att det inte går att påvisa någon signifikant skillnad i bildkvalitet mellan algoritmerna när SSIM används som mått.

Independent-Samples Mann-Whitney U Test Summary	
Total N	72
Mann-Whitney U	529.000
Wilcoxon W	1195.000
Test Statistic	529.000
Standard Error	88.770
Standardized Test Statistic	-1.341
Asymptotic Sig.(2-sided test)	.180

Figur 4.12: Mann-Whitney U-testresultat för SSIM vid detaljrika scener.

Sammanfattningsvis presterade AVIF signifikant bättre än JPEG XL när PSNR användes som mått för detaljrika scener medan ingen signifikant skillnad kunde påvisas när bildkvaliteten mättes med SSIM.

5

Diskussion

I detta avsnitt diskuteras studiens resultat i relation till forskningsfrågan, tidigare forskning och studiens metodologiska val. Genom att analysera de observerade skillnaderna mellan JPEG XL och AVIF vid destruktiv komprimering belyses både styrkor och begränsningar hos de undersökta algoritmerna. Vidare reflekteras kring studiens begränsningar, resultatens praktiska betydelse samt förslag på fortsatt forskning för att ytterligare fördjupa kunskapen inom området.

5.1 Slutsatser

Studiens syfte var att undersöka om det finns skillnader i bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering. Studien kontrollerade för övriga variabler såsom komprimeringsgrad och komprimeringshastighet vilket innebär att jämförelsen isolerades till bildkvalitet mätt genom PSNR och SSIM. Den generella frågeställningen formulerades som: *Finns det skillnader i bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF vid destruktiv komprimering?*

För att besvara huvudfrågan analyserades bildkvalitet inom tre olika testfall: mörka scener, skarpa kanter och detaljrika scener.

För mörka scener visade Mann-Whitney U-testet att JPEG XL presterade signifikant bättre än AVIF både för PSNR och SSIM. Detta indikerar att JPEG XL i högre grad bevarar bildkvalitet under destruktiv komprimering när bildinnehållet domineras av mörkare nyanser. Resultatet är särskilt intressant med tanke på JPEG XL:s utformning för att hantera gradienter och

mörka bildsegment effektivt vilket framhävts i tidigare studier på dess tekniska struktur [9]. JPEG XL var även 31,6% snabbare än AVIF i det här testfallet, vilket konkretiseras att JPEG XL är mer effektiv än AVIF för mörka bilder. Å andra sidan hade JPEG XL en 227% större filstorlek än AVIF, se Bilaga B 2.

För skarpa kanter presterade AVIF signifikant bättre än JPEG XL för både PSNR och SSIM. Detta tyder på att AVIF hanterar skarpa kontraster mellan ljusa och mörka områden mer effektivt. Fyndet är i linje med Vijayvargiya et al. [29] som betonar att algoritmers prestanda är beroende av bildinnehållets specifika egenskaper. Här kan AVIF:s tekniska implementation ha en fördel i att komprimera komplexa kontrastmönster utan större kvalitetstapp. JPEG XL var dock 63,9% snabbare än AVIF i det här testfallet samt endast 1% större filstorlek, se Bilaga B 2. Det tyder på att JPEG XL är förhållandevis tidsbesparande med marginella skillnader i komprimeringsgraden.

För detaljrika scener var resultaten delvis signifikanta. AVIF presterade bättre än JPEG XL för PSNR medan ingen signifikant skillnad kunde påvisas för SSIM. Detta antyder att AVIF har en viss fördel i att bevara bildkvalitet för detaljrika bilder men att skillnaden inte är lika tydlig som för skarpa kanter. Resultatet kan relateras till Subramanya [23] och Dhawan [14] som båda påpekar vikten av att välja komprimeringsteknik utifrån krav på detaljåtergivning och balans mellan kvalitet och komprimeringsgrad. Samtidigt var AVIF 94,3% snabbare än JPEG XL med endast 11,2% större filstorlek, se Bilaga B 2.

Sammantaget indikerar studiens resultat att det finns skillnader i bildkvalitet mellan JPEG XL och AVIF vid destruktiv komprimering men att dessa skillnader varierar beroende på bildtyp:

- JPEG XL presterar bättre för mörka scener.
- AVIF presterar bättre för skarpa kanter och detaljrika scener.

Resultaten i denna studie utgör ett viktigt komplement till den begränsade forskningen kring JPEG XL och AVIF. Tidigare studier, såsom Googles [24] och Cloudinarys [11] analyser, har främst fokuserat på komprimeringshastighet och filstorlek. Genom att isolera och undersöka bildkvalitet ger denna studie ny empirisk kunskap om hur dessa algoritmer presterar för olika typer av bildinnehåll.

Resultaten är betydelsefulla eftersom de ger praktiska insikter för aktörer som hanterar bildkomprimering i olika sammanhang. För användningsområden där mörka scener är vanliga, exempelvis filmproduktion,

kan JPEG XL vara att föredra. För situationer som kräver bevarande av hög kontrast och detaljer, exempelvis medicinsk bildbehandling eller tekniska illustrationer, kan AVIF vara ett bättre alternativ.

5.2 Begränsningar

Det är viktigt att beakta studiens begränsningar för att sätta resultaten i ett korrekt sammanhang och tydliggöra dess generaliserbarhet och tillförlitlighet. Begränsningarna härrör främst från metodval, resurstillgång och de valda mätningsverktygen. För det första har studiens parametrar och testfall valts ut med hänsyn till de praktiska begränsningarna som tid och datorresurser medför vilket påverkat både omfattningen och urvalet av data. Vidare kan val av bilder, komprimeringsteckniska aspekter och specifik datorutrustning ha haft en inverkan på resultatens applicerbarhet i större skala. Slutligen medför valet av mätningsverktyg, PSNR och SSIM, en viss begränsning i studiens förmåga att spegla den faktiska perceptuella upplevelsen av bildkvalitet.

Följande avsnitt belyser dessa begränsningar i detalj med fokus på metodval, testfallens omfattning, datorresurser och mätningsverktygens begränsningar.

5.2.1 Metodval

I ett optimalt scenario skulle det vara möjligt att köra alla kombinationer av parametrar över alla testfall och välja rätt parametrar baserat på det faktiska testet. Detta är dock inte genomförbart på grund av begränsad tillgång till tid och datorresurser. För att identifiera parametrar med liknande komprimeringstid och komprimeringsgrad användes istället utfallsbaserad parametermatchning vilket innebär att en delmängd av varje testfall användes som bas för att hitta matchande parametrar mellan varje teknik. I detta fall fungerar 1–2 bilder som bas för varje testfall. Detta möjliggör testning av ett stort antal parametrar på en representativ delmängd av det faktiska testfallet. Det finns dock ingen garanti för att parametrar med snarlik resursanvändning i en komprimering av en viss mängd bilder kommer att ha samma resursanvändning i en annan köring av en annan mängd bilder.

5.2.2 Testfall

Totalt komprimerades endast 51 bilder över tre testfall. Begränsningen berodde främst på bristande resurser i form av datorkraft och tid. Den slutgiltiga komprimeringen tog cirka 25 minuter och tillsammans med avkodning och mätning strax över en timme. Detta ledde till att antalet bilder som kunde komprimeras begränsades vilket resulterade i ett restriktivt urval för analysen.

Bilderna som komprimerades var i ett okomprimerat format och hade generellt sett exceptionellt stora filstorlekar vilket gör analysen endimensionell i förhållande till hur bildens filstorlek kan påverka respektive tekniks prestanda. En alternativ strategi hade varit att inkludera redan komprimerade bilder för att ge en bredare förståelse av teknikernas flexibilitet och för att belysa eventuella begränsningar i deras förmåga att hantera den typen av bilddata.

5.2.3 Begränsande datorutrustning

Testerna utfördes på en begränsande hemdator vilket skiljer sig från hur företag som Netflix och Google skulle nyttja varsin komprimeringsteknik. Datorn som komprimerade hade exempelvis 16 GB arbetsminne vilket nästan under hela körningen var maximerat. Det kan orsaka att vardera av algoritmerna spenderade en oproportionerlig stor mängd tid på att ladda in och ur data från ett virtuellt arbetsminne istället för att komprimera.

I samma veva observerades det även under körningen att algoritmerna använde processorkraft i olika utsträckning. Exempelvis var AVIF benägen att använda all processorkraft på alla kärnor vilket gjorde datorn obrukbar till andra ändamål under tiden som AVIF komprimerade. JPEG XL var å andra sidan resurssnålare och orsakade inte problem som hämmade vardagligt datorbruk. Detta ska dock ses som en anekdot då det alltid är möjligt att parallellt köra flera instanser av JPEG XL och i viss utsträckning AVIF om det önskas ett mer omfattande bruk av resurser.

5.2.4 Urval av mätverktyg

För att mäta bilddegraderingen som komprimeringen medför används PSNR och SSIM i denna studie. Moderna bildkomprimeringstekniker är dock avancerade och använder psykovisuella modeller för att ta bort detaljer och komprimera där det är mest sannolikt att den mänskliga hjärnan inte märker av förändringen. Detta gör inte bara mängden tillämpad komprimering

intressant utan även hur komprimeringen tillämpas. Exempelvis är män-niskor ofta mer benägna att fokusera på mitten av en bild snarare än på hörn och kanter. Därför kan det vara mer relevant att mäta lokaliseringen av komprimeringen i förhållande till dess perceptuella betydelse, något som varken SSIM eller PSNR tar hänsyn till.

5.3 Resultatetiska aspekter

I denna studie har stor vikt lagts vid att säkerställa resultatens trovärdighet och transparens. Ett av de centrala etiska övervägandena har därför varit att minimera risken för att publicera felaktiga eller missvisande resultat. För att uppnå detta har ett antal åtgärder vidtagits.

Val av data och metod för datainsamling har redovisats öppet för att säkerställa transparens genom hela forskningsprocessen. Detta innefattar hur testfall och parametrar valts ut samt hur data har bearbetats för att möjliggöra en rättvis och jämförbar analys. Resultaten har rapporterats i sin helhet, inklusive fall där skillnader inte var signifikanta, för att undvika selektiv rapportering. Resultaten har tolkats med försiktighet och med hänsyn till studiens begränsningar. Inga slutsatser har dragits bortom vad datan kan stödja. Denna transparens underlättar för andra att granska och reproducera studien vilket ytterligare stärker dess trovärdighet.

För att undvika att oändliga värden påverkar resultaten negativt har dessa ersatts med det högsta observerade värdet inom respektive dataset. Denna metod valdes för att säkerställa en rättvis och konsistent analys utan att data snedvrids.

Endast bilder som är fria från upphovsrättsliga begränsningar har använts vilket stärker studiens reproducertbarhet och trovärdighet. Detta val är särskilt viktigt i jämförande studier där bilddata utgör grunden för analys och tolkning av resultat.

Sammanfattningsvis har studiens resultat rapporterats på ett sätt som är transparent, nyanserat och i linje med god forskningsetik. Genom att tydligt redovisa val av metoder, datahantering och tolkning av resultat har studien strävat efter att upprätthålla hög kvalitet och trovärdighet i studien.

5.4 Fortsatt forskning

Denna studie har bidragit med insikter kring skillnader i bildkvalitet mellan komprimeringsalgoritmerna JPEG XL och AVIF men flera frågor kvarstår

som kan utforskas i framtida forskning. En central begränsning i denna studie har varit de begränsade resurserna vad gäller datorkraft, tid och urvalsstorlek. Framtida studier bör inkludera ett större antal testfall och bilder för att skapa en mer robust och generaliseringbar analys. Med mer omfattande resurser skulle det vara möjligt att testa fler kombinationer av parametrar och få en djupare förståelse för algoritmernas prestanda.

Vidare bör framtida forskning kombinera objektiva mått som PSNR och SSIM med subjektiva bedömningar av bildkvalitet från användare. Moderna komprimeringstekniker är ofta optimerade för att anpassa sig till mänsklig perception vilket gör det viktigt att även inkludera mänskliga aspekter i bedömningen av bildkvalitet. Det finns även en stor potential i att utforska AI-baserade metoder för bildkomprimering. AI och maskininlärning kan användas för att ytterligare optimera algoritmerna, exempelvis genom att anpassa komprimeringen efter hur människor uppfattar olika delar av en bild.

Slutligen behövs fler oberoende och standardiserade studier för att lösa den nuvarande oenigheten mellan intressenter kring vilken algoritm som presterar bäst. Genom att analysera hur JPEG XL och AVIF presterar i olika användningsområden, såsom medicinsk bildbehandling, tekniska illustrationer eller redan komprimerade bilder, kan framtida forskning bidra till en mer heltäckande förståelse av respektive algoritms styrkor och svagheter.

Litteratur

- [1] Johannes Ballé, Valero Laparra och Eero P Simoncelli. "End-to-End Optimized Image Compression". I: *CoRR* abs/1611.01704 (2016). arXiv: 1611.01704.
- [2] Nabajeet Barman och Maria G. Martini. "An Evaluation of the Next-Generation Image Coding Standard AVIF". I: *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, s. 1–4. DOI: 10.1109/QoMEX48832.2020.9123131.
- [3] BBC News. *JPEG lockdown: Restriction options sought by committee*. <https://www.bbc.com/news/technology-34538705>. Accessed: 2025-01-12. 2015.
- [4] Tilda Berghäll. *Private image, used with permission*. Private collection. 2023.
- [5] Tilda Berghäll. *Private image, used with permission*. Private collection. 2024.
- [6] BT.com. *What is a JPEG? The internet's most popular image format explained*. <https://web.archive.org/web/20190220125937/http://home.bt.com/tech-gadgets/photography/what-is-a-jpeg-11364206889349>. Accessed: 2025-01-12. 2018.
- [7] Can I Use? *AVIF browser support*. <https://caniuse.com/?search=avif>. Accessed: 2024-11-20. 2024.
- [8] Can I Use? *JPEG XL browser support*. <https://caniuse.com/?search=jpeg%20xl>. Accessed: 2024-11-20. 2024.
- [9] Joonhyung Cho, Oh-Jin Kwon och Seungcheol Choi. "Improvement of JPEG XL lossy image coding using region adaptive dct block partitioning structure". I: *IEEE Access* 9 (2021), s. 113213–113225. DOI: 10.1109/ACCESS.2021.3102235.

- [10] Cloudinary. *AVIF Format: A Next-Gen Image Format to Rule them All?* <https://cloudinary.com/guides/image-formats/avif-format-a-next-gen-image-format-to-rule-them-all>. Accessed: 2025-01-12. 2024.
- [11] Cloudinary. *The case for JPEG XL*. <https://cloudinary.com/blog/the-case-for-jpeg-xl>. Accessed: 2025-01-12. 2022.
- [12] JPEG XL Community. *JPEG XL Official Website*. <https://jpegxl.info/>. Accessed: 2024-12-03. 2024.
- [13] Martyn Denscombe. *The Good Research Guide: For Small-Scale Societal Research Projects*. 5. utg. McGraw-Hill Education, 2014.
- [14] Sachin Dhawan. "A Review of Image Compression and Comparison of Its Algorithms". I: *International Journal of Electronics & Communication Technology* 2.1 (2011), s. 22–26.
- [15] Rafael C Gonzalez och Rirchard E Woods. *Digital Image Processing*. 3. utg. Upper Saddle River, NJ: Pearson Education International, 2008.
- [16] Apoorv Gupta, Aman Bansal och Vidhi Khanduja. "Modern Lossless Compression Techniques: Review, Comparison and Analysis". I: *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. IEEE, 2017, s. 1–8. DOI: 10.1109/ICECCT.2017.8117850.
- [17] Murooj Khalid Ibraheem Ibraheem, Alexander Viktorovich Dvorkovich och Israa M Abdalameer Al-khafaji. "A Comprehensive Literature Review on Image and Video Compression: Trends, Algorithms, and Techniques". I: *Ingénierie des Systèmes d'Information* 29.3 (2024). DOI: 10.18280/isi.290307.
- [18] Paul Johannesson och Erik Perjons. *An Introduction to Design Science*. Vol. 10. Cham: Springer, 2014.
- [19] Keleadile Lucia Ketshabetswe m. fl. "Data Compression Algorithms for Wireless Sensor Networks: A Review and Comparison". I: *IEEE Access* 9 (2021), s. 136872–136891. DOI: 10.1109/ACCESS.2021.3116311.
- [20] Saluka Ranasinghe Kodituwakku och Upali S Amarasinghe. "Comparison of Lossless Data Compression Algorithms for Text Data". I: *Indian Journal of Computer Science and Engineering* 1.4 (2010), s. 416–425.

- [21] Thulfiqar H Mandeel m. fl. "A Comparative Study on Lossless Compression Mode in WebP, Better Portable Graphics (BPG), and JPEG XL Image Compression Algorithms". I: *2021 8th International Conference on Computer and Communication Engineering (ICCCE)*. IEEE, 2021, s. 17–22. DOI: 10.1109/ICCCE50029.2021.9467224.
- [22] Phlake. *File:Colour banding example01.png*. https://commons.wikimedia.org/wiki/File:Colour_banding_example01.png. Accessed: 2024-12-03. 2004.
- [23] A Subramanya. "Image Compression Technique". I: *IEEE Potentials* 20.1 (2001), s. 19–23. DOI: 10.1109/45.913206.
- [24] The AVIF Team. *Image coding comparisons*. <https://storage.googleapis.com/avif-comparison/index.html>. Accessed: 2025-01-12. 2022.
- [25] Tris-T7. *File:Digital graphic pattern 80.png*. https://commons.wikimedia.org/wiki/File:Digital_graphic_pattern_80.png. Accessed: 2024-12-03. 2021.
- [26] Ranjit Kumar Tulabandu m. fl. "Evolution of AVIF Encoder: Speed and Memory Optimizations". I: *2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2022, s. 90–95. DOI: 10.1109/MIPR54900.2022.00023.
- [27] W3Techs. *Usage statistics of image file formats for websites*. https://w3techs.com/technologies/overview/image_format. Accessed: 2025-01-12. 2025.
- [28] Vetenskapsrådet. *God forskningssed*. <https://www.vr.se/download/18.4c9f221a191e4edf9053a474/1727853946433/God%20forsknings%20VR%202024.pdf>. Accessed: 2024-12-03. 2024.
- [29] Gaurav Vijayvargiya, Sanjay Silakari och Rajeev Pandey. "A Survey: Various Techniques of Image Compression". I: *CoRR* abs/1311.6877 (2013). arXiv: 1311.6877.

Bilaga A — Exempel Basfall

File	Format	Effort	Distance	Speed	Quality	ms	bytes	ID
testfall\high-contrast\Digital_Graphic_Pattern_35.png	JPEG XL	1.0	11.0	-	-	113	58024	1
testfall\high-contrast\Digital_Graphic_Pattern_35.png	AVIF	-	-	9.0	35.0	113	57112	1
testfall\high-contrast\Digital_Graphic_Pattern_35.png	JPEG XL	1.0	12.0	-	-	110	54671	2
testfall\high-contrast\Digital_Graphic_Pattern_35.png	AVIF	-	-	9.0	35.0	113	57112	2
testfall\high-contrast\Digital_Graphic_Pattern_35.png	JPEG XL	1.0	13.0	-	-	117	51736	3
testfall\high-contrast\Digital_Graphic_Pattern_35.png	AVIF	-	-	9.0	32.0	119	49172	3
testfall\high-contrast\Digital_Graphic_Pattern_35.png	JPEG XL	1.0	14.0	-	-	120	48315	4
testfall\high-contrast\Digital_Graphic_Pattern_35.png	AVIF	-	-	9.0	32.0	119	49172	4
testfall\high-contrast\Digital_Graphic_Pattern_35.png	JPEG XL	1.0	21.0	-	-	251	36259	5
testfall\high-contrast\Digital_Graphic_Pattern_35.png	AVIF	-	-	7.0	29.0	262	35021	5

Tabell 1: Exempeldata för JPEG XL och AVIF vid olika inställningar.

Bilaga B — Testfallöversikt

FOLDER	COMMAND	JPEG	AVIF	JPEG	AVIF	FASTER TE- CHNIQUE	RELATIVE REDUC- TION (%)	BEST COM- PRESSION DIF. (%)	COMPRESSION DIF. (%)	PSNR JXL	PSNR AVIF	SSIM JXL	SSIM AVIF	MSE IMPROVE- MENT (%)
dark	1	22727	52976	98.228	99.942	JPEG XL	57.099	AVIF	1.714	40.919	33.382	0.942	0.723	82.367
dark	2	33043	63323	99.371	99.382	JPEG XL	47.818	AVIF	0.011	39.496	39.889	0.884	0.937	-9.487
dark	3	97006	106983	99.364	99.874	JPEG XL	9.326	AVIF	0.509	39.587	36.666	0.893	0.838	48.963
high-contrast	1	3476	3928	89.283	81.732	JPEG XL	11.517	JPEG XL	7.551	inf	36.598	0.936	0.951	-inf
high-contrast	2	105281	202434	96.332	95.920	JPEG XL	47.993	JPEG XL	0.412	31.624	inf	0.821	0.939	inf
high-contrast	3	141216	203320	82.693	90.981	JPEG XL	30.545	AVIF	8.288	38.391	inf	0.978	0.985	inf
high-detail	1	29923	31201	98.403	97.552	JPEG XL	4.098	JPEG XL	0.851	32.904	33.650	0.816	0.892	-18.749
high-detail	2	700663	390225	97.063	96.358	AVIF	44.306	JPEG XL	0.706	34.270	35.495	0.903	0.937	-32.586
high-detail	3	1028508	483991	90.221	90.107	AVIF	52.942	JPEG XL	0.114	38.169	40.569	0.960	0.980	-73.775

Tabell 2: Resultat för komprimeringstid, reduktion, bildkvalitet och prestanda för JPEG XL och AVIF.