



**FFI** Forsvarets  
forskningsinstitutt

23/02227

FFI-RAPPORT

# Maskinlæring for tematisk inndeling og analyse av propagandabildene til IS (Den islamske staten)

Mathias Bynke  
Vidar B. Skretting  
Bernt Ivar Utstøl Nødland



# **Maskinl ring for tematisk inndeling og analyse av propagandabildene til IS (Den islamske staten)**

Mathias Bynke  
Vidar B. Skretting  
Bernt Ivar Utst l N dland

---

---

## **Emneord**

Kunstig intelligens  
Maskinl ring  
Bildeanalyse  
Terrorisme  
Jihad

## **FFI-rapport**

23/02227

## **Prosjektnummer**

582801

## **Engelsk tittel**

Machine learning for thematic categorisation and analysis of Islamic State propaganda pictures

## **Elektronisk ISBN**

978-82-464-3500-8

## **Godkjenner**

Stig Rune Sellev g, *forskningsleder*  
Janet Martha Blatny, *forskningsdirekt r*

*Dokumentet er elektronisk godkjent og har derfor ikke h ndskreven signatur.*

## **Opphavsrett**

  Forsvarets forskningsinstitutt (FFI). Publikasjonen kan siteres fritt med kildehenvisning.



---

---

## Sammendrag

Analytikere og forskere som studerer terroristgrupper og andre militante, ikke-statlige aktører, ser ofte til propagandamaterialer for å danne seg et bilde av hvem disse aktørene er, hva de ønsker, og hvordan de opererer. På grunn av den teknologiske utviklingen har det blitt mye lettere for slike aktører å produsere propaganda i store mengder og spre den på nettet. Derfor er det nærmest umulig for analytikere og forskere å få oversikt over denne propagandaen manuelt. Vi er med andre ord nødt til å utvikle nye metoder for å analysere disse store datamengdene. Denne studien utforsker hvordan maskinlæring kan støtte analysen av én av de mest utbredte formene for propaganda, nemlig bilder.

I denne rapporten tar vi for oss et typisk brukstilfelle – et korpus på over 30 000 propagandabilder produsert av terrororganisasjonen IS (Den islamske staten). Vi bruker maskinlæringsmetoder til å danne en oversikt over innholdet i dette korpuset og anvender denne oversikten som utgangspunkt for en analyse av propagandamaterialet. Målet med studien er todelt – både å bidra til metodeutvikling innen maskinlæring og å bidra til ny forståelse av IS' bildepropaganda.

For å analysere innholdet i hvert enkelt bilde bruker vi maskinlæringsmodellen Contrastive Language-Image Pre-training (CLIP). CLIP oversetter hvert bilde til en vektor som representerer innholdet i bildet. Deretter bruker vi klyngealgoritmen  $k$ -snitt til å dele inn disse vektorene i et antall klynger med så liten som mulig variasjon innad i hver klynge. Dette gir oss klynger av bilder med lignende tematisk innhold.

Vi bruker denne tematiske inndelingen som utgangspunkt for en analyse av IS' propagandabilder i perioden 2014–2022. Siden hvert av bildene inneholder metadata om produksjonstidspunkt og -sted, kan vi også analysere hvordan tematikken i IS' bildepropaganda har utviklet seg over tid og mellom IS' såkalte provinser.

Hovedfunnene i analysen er at IS har gått fra å fremstille seg både som en militær organisasjon og som et sivilt statsapparat tidlig i perioden (2014–2018) til nesten utelukkende å vise seg som militær bevegelse og opprørsgruppe fra 2019. Videre har den geografiske hovedtyngden i materialet beveget seg fra Irak og Syria til Vest- og Sentral-Afrika. Denne utviklingen i propagandamaterialet gjenspeiler de store endringene IS som organisasjon har vært gjennom i samme tidsrom, noe som blir særlig tydelig i materialet etter at IS mistet de siste landområdene de kontrollerte i Irak og Syria, i 2018–2019. Samtidig tyder funnene våre på at IS bevisst velger å nedtone de sivil-administrative delene av sin egen aktivitet i deler av Afrika der gruppen nå er i ferd med å bygge seg opp igjen. De afrikanske IS-provinsene viser også et gjennomgående mer voldelig innhold i bildene sine enn de andre provinsene, og de vier også vesentlig mindre oppmerksomhet til egne martyrer enn det som har vært vanlig i annen jihadistpropaganda. Disse utviklingstrekkene kan si noe om hvor IS er på vei i fremtiden.

Kombinasjonen av CLIP og klyngealgoritmer fremstår som et godt verktøy for å få rask oversikt over store mengder bilder, og det fungerer utmerket som analysestøtte. Metoden er enkel å bruke og lar seg lett overføre til andre brukstilfeller.

---

---

## Summary

For analysts and researchers studying terrorist groups and other non-state actors, propaganda materials provide important sources to understand who these actors are, what they want, and how they operate. Recent technological developments have made it significantly easier for such actors to produce propaganda and publish it online, in ever-increasing quantities. This has made it almost impossible for analysts and researchers to go through these materials manually. It is therefore imperative that we develop new methods to analyse such large amounts of data. This report explores how machine learning can support the analysis of one of the most common forms of propaganda: pictures.

In this report, we have chosen a typical use case – a corpus of over 30,000 propaganda pictures produced by the terrorist organisation IS (The Islamic State) – and explore how machine learning methods can be used to gain an overview of its contents. This is used as a starting point for an analysis of the propaganda pictures. The aim of the study is twofold: to contribute to method development in machine learning and to provide new insights into IS' visual propaganda.

We use the machine learning model Contrastive Language-Image Pre-training (CLIP) to analyse the content of each individual image. CLIP translates each image into a vector that represents its content. We then use the clustering algorithm  $k$ -means to divide these vectors into a number of clusters, minimising the internal variation between the vectors in each cluster. This results in clusters of images with similar thematic content.

We use this thematic division as a starting point for an analysis of IS' propaganda pictures in the period 2014–2022. Since each of the images contains metadata about the time and place of production, we can also analyse how the themes in IS' image propaganda have developed over time and between IS' so-called provinces.

Our main findings are that while IS portrayed itself both as a military organisation and as a civilian state apparatus early in the period (2014–2018), since 2019, it has almost exclusively presented itself as a military organisation and rebel group. Furthermore, the preponderance of the propaganda material has moved geographically from Iraq and Syria to West and Central Africa. The development in the propaganda materials reflects the major changes IS has undergone as an organisation in the same period. This is especially apparent after IS lost its last pockets of territory in Iraq and Syria in 2018–2019. However, our findings also suggest that IS deliberately chooses to downplay the civil-administrative parts of its own activity in parts of Africa where the group is on the rise. Moreover, the pictures produced by the African IS provinces feature significantly more violent content than the other provinces, while there is less content devoted to martyrs than what has been the norm in jihadist propaganda. These developments may give an indication of where IS is headed in the near future.

We find that combining CLIP and clustering algorithms is a quick way to get an overview of the contents of a large image corpus, providing a useful tool for analysts. The method is easy to employ and can easily be adapted to other use cases.

---

---

# Innhold

<b>Sammendrag</b>	3
<b>Summary</b>	4
<b>Forord</b>	7
<b>1 Introduksjon</b>	9
<b>2 Tilfellestudie: IS' bildepropaganda</b>	10
2.1 Den islamske staten (IS): En kort introduksjon	10
2.2 IS-propaganda	12
2.3 IS-propaganda i forskningslitteraturen	12
2.4 Om korpuset	14
<b>3 Metode</b>	15
3.1 Tolke bildenes innhold: CLIP	15
3.1.1 Arkitektur	16
3.1.2 Klassifikator med manuelle klasser	17
3.2 Finne lignende bilder: klynging	17
3.2.1 Klyngealgoritmer	18
3.2.2 Tolke klyngene	24
3.2.3 Bunntekster	26
<b>4 Strukturering av datamaterialet</b>	27
4.1 Finne metadata	27
4.1.1 Finne informasjon i bildeteksten: optisk tegngjenkjenning	27
4.1.2 Koding av provinser	29
4.1.3 Finne og kode datoer	29
4.2 Tematiske klynger	31
4.2.1 Valg av antall klynger og tematisk konsistens	31
4.2.2 Strukturering av tematiske klynger	32
4.2.3 Oversikt over tematiske klynger	34
4.2.4 Eksempelbilder fra klyngene	36
4.2.5 Test av tematisk konsistens i klyngene	40
<b>5 Analyse</b>	43
5.1 Bildeproduksjon per år og provins	43
5.2 Utvikling i tematisk innhold over tid	46
5.3 Tematisk innhold fordelt på provins	48
5.4 Oppsummering av analytiske funn	50

---

<b>6 Konklusjon</b>	52
<b>Referanser</b>	53
<b>Vedlegg</b>	
<b>A Genererte bilder</b>	56

---

---

## Forord

Forskningen som blir presentert i denne rapporten, er en del av “Maskinlæring og trusselaktører”, et samarbeidsprosjekt mellom FFIs forskningsmiljøer på terrorisme og kunstig intelligens/maskinlæring. Prosjektet er blitt finansiert av Forsvarsdepartementet gjennom tilskuddsordningen til akademiske institusjoner og universitetsmiljø innen sikkerhets- og forsvarspolitisk forskning.

Kjeller, 30. november 2023.

Vidar B. Skretting  
Mathias Bynke  
Bernt Ivar Utstøl Nødland



---

---

# 1 Introduksjon

Internett har gjort at man kan dele informasjon mer effektivt enn noen gang tidligere. I et sikkerhetsperspektiv betyr dette at fiendtlige aktører har fått langt større evne til å nå ut med propaganda. Dette gjelder naturligvis for statlige aktører, men endringen har vært særlig betydningsfull for militante ikke-statlige aktører, som i tiden før internett ble allemannseie, ikke kunne kommunisere direkte med omverdenen og var avhengige av andre medier for å nå ut med budskapet sitt. I dag har den teknologiske utviklingen gjort at selv aktører med begrensede ressurser kan lage propaganda av høy kvalitet og nå ut med denne til et globalt publikum.

For forskere og analytikere som studerer disse aktørene, utgjør propagandamaterialene – herunder tekst, video, lyd og bilder – viktige kilder til å forstå hvem aktørene er, hva de ønsker og hvordan de opererer. Men propagandaen produseres og spres nå i et så stort volum at det er vanskelig å holde oversikt. Derfor er det et klart behov for nye metoder til å analysere disse store datamengdene.

Denne rapporten undersøker hvordan maskinlæring kan brukes for å understøtte analysen av én av de mest utbredte formene for propaganda, nemlig bilder. Vi gjennomfører et tilfellestudie der vi bruker maskinlæringsmetoder til å analysere og kategorisere propagandabilder publisert av terrororganisasjonen IS (“Den islamske staten”). Datagrunnlaget er et unikt korpus på over 30 000 bilder publisert i perioden 2014–2022. Vi bruker maskinlæringsmodellen CLIP til å tolke innholdet i hvert bilde og deretter en “klyngealgoritme” ( $k$ -snitt) til å kategorisere bildene tematisk. Deretter bruker vi denne tematiske kategoriseringen til å studere hvordan IS’ propaganda har utviklet seg over tid og mellom ulike geografiske områder. Målet er å bidra både til metodeutvikling innen maskinlæring og til ny kunnskap om IS.

Rapporten er bygget opp som følger: Kapittel 2 gir en introduksjon til forskningslitteraturen om IS og presenterer bildekorpuset. I kapittel 3 presenterer vi maskinlæringsmetoden CLIP og viser hvordan man kan kombinere denne med “klyngealgoritmer” for å programmatisk kategorisere bildekorpuset etter tema. Kapittel 4 beskriver hvordan vi gikk frem for å ekstrahere metadata fra bildene og hvilke kvalitative valg vi tok for å operasjonalisere dataene fra kapittel 3 i analysen. I analysekapittelet (5) sammenstiller vi data om det tematiske innholdet i bildene – som vi fikk ved å kombinere CLIP og klyngealgoritmen  $k$ -snitt – med metadata om bildenes publiseringsdato og -sted. Slik får vi et unikt innblikk i hvordan IS’ propaganda har utviklet seg over tid og mellom forskjellige geografiske områder.

Vi konkluderer med at det å kombinere CLIP og klyngealgoritmer gir oss et svært nyttig verktøy for å analysere store mengder bilder. Tematisk kategorisering på denne måten gjør at man lett og effektivt kan få oversikt over bildekorpuset og se utviklingstrender i propagandabildene over tid. De viktigste trendene vi finner er at IS har gått fra å fremstille seg både som et sivilt statsapparat og militærmakt fra 2014 til 2018, til nesten utelukkende å vise militær aktivitet i årene som fulgte. På samme tid har den geografiske hovedvekten i materialet beveget seg fra Syria og Irak til Afrika. Metodene vi har brukt, lar seg lett overføre på tilsvarende brukstilfeller med store bildematerialer.

---

---

## 2 Tilfellestudie: IS' bildepropaganda

I dette kapitlet presenterer vi tilfellestudiet som vi skal ta for oss i resten av rapporten. Forskningsspørsmålet er det følgende: *Hvordan har IS' bildepropaganda utviklet seg fra 2014–2022, og hva forteller denne utviklingen oss om IS som gruppe?* Vi ønsker med andre ord å fokusere ikke bare på propagandaen i seg selv, men også på hva den kan si om IS' *historiske utvikling*.

Derfor begynner vi dette kapitlet med nettopp en introduksjon til IS' historie (2.1) og hovedtrekkene i gruppens propagandaproduksjon (2.2). I seksjon 2.3 gjennomgår vi forskningslitteraturen på IS-propaganda generelt og bildepropaganda spesielt. Vi avslutter med seksjon 2.4, som presenterer bildekorpuset vi baserer vårt arbeid på.

### 2.1 Den islamske staten (IS): En kort introduksjon

Den islamske staten (IS)<sup>1</sup> er en jihadistisk<sup>2</sup> terrororganisasjon.<sup>3</sup> IS' tilblivelse kan spores tilbake til gruppen *al-Tawhid wal-jihad*, en jihadistgruppe som ble etablert av jordaneren Abu Musab al-Zarqawi i 1999. Rett før invasjonen av Irak i 2003 etablerte al-Zarqawi og hans menn seg i landet for å slåss mot amerikanerne. Året etterpå sluttet denne gruppen seg til Osama bin Ladens al-Qaida-nettverk og tok navnet al-Qaida i Irak (AQI) (Byman 2016, s. 131–132).

AQI klarte å tiltrekke seg mange nye rekrutter, både fra utlandet og blant tidligere soldater i Saddam Husseins hær. Gruppen ble etterhvert en betydelig maktfaktor i Irak og forsøkte å fremstille seg som landets sunnimuslimske minoritets beskytter, blant annet ved å utføre en rekke terrorangrep mot den sjiamuslimske majoritetsbefolkningen. AQI gjennomførte også en rekke angrep mot koalisjonsstyrkene, ikke-jihadistiske militser og irakiske styresmakter.

Etter al-Zarqawis død i 2006 slo AQI seg sammen med flere andre sunni-muslimske militsgrupper og dannet Den islamske staten i Irak (ISI) (Tønnessen 2015, s. 48–49). Men ISIs ekstremisme og brutalitet gjorde at gruppen mistet mye av sin tidligere støtte blant sunnimuslimske sivile, og fra 2006 ble ISI påført betydelige militære nederlag av USA-støttede irakiske stammemilitser (Cigar 2014, s. 7–15). I 2008 lå ISI så langt nede at både amerikanske og irakiske styresmakter spådde at gruppens endelikt var nært forestående (Tønnessen 2019).

---

<sup>1</sup>Det er stor uenighet om hvilket navn man skal bruke om denne organisasjonen, som siden 2014 har kalt seg enten *al-dawla al-islāmiyya*, arabisk for “den islamske staten”, eller *al-khilāfa*, arabisk for “kalifatet”. Mange foretrekker å bruke en versjon av det navnet organisasjonen brukte i perioden 2013–2014: ISIL (Den islamske staten i Irak og Levanten); ISIS (Den islamske staten i Irak og *al-Sham*, et arabisk ord for Levanten); eller *Dā'ish* (et arabisk akronym som betyr det samme som ISIL). Vi anser det som mest hensiktsmessig å bruke det samme navnet organisasjonen bruker om seg selv, ikke minst fordi IS har vært langt mer aktive utenfor Irak og Levanten de siste årene (se kapittel 5 i denne rapporten).

<sup>2</sup>Begrepe “jihadist” og “jihadisme” refererer til militante sunni-muslimske bevegelser som kjemper mot det de anser som ikke-islamske regimer i den muslimske verden og mot vestlige land som støtter dem. Jihadistenes endelige mål er å opprette et nytt kalifat og styre etter islamsk lov. For mer utfyllende diskusjon rundt denne terminologien, se for eksempel Hegghammer (2014).

<sup>3</sup>Vi bruker begrepet terrororganisasjon, fordi blant andre FNs sikkerhetsråds sanksjonskomité anser IS og alle dens forgjengere for å være terrororganisasjoner. Dette betyr ikke at IS kun driver med terrorisme. Terrorangrep er kun ett av verktøyene i IS militære repertoar. IS driver også opprørskrig og regulær krig, og gruppen har tidvis også drevet en sivil administrasjon.



---

---

Mot alle odds skulle likevel 2010-tallet som fulgte, bli et høydepunkt for ISI. Etter at USA trakk ut sine styrker av Irak i 2011, ble de sekteriske skillelinjene på toppen av irakisk politikk tydeligere. Under den sjia-muslimske statsministeren Nouri al-Maliki bredte det seg en (ikke helt urimelig) oppfatning blant sunnimuslimer om at staten favoriserte den sjia-muslimske majoriteten, som igjen førte til at jihadistene igjen fikk økt oppslutning (Byman 2016, s. 133–134). Samtidig åpnet borgerkrigen i Syria opp nye muligheter for ISI, som grep sjansen til å opprette en ny avdeling i nabolandet.

Etter en konflikt med al-Qaida om strategien i Syria brøt ISI med moderorganisasjonen i 2013 og byttet navn til ISIL (Den islamske staten i Irak og Levanten) for å gjenspeile at gruppen nå offisielt opererte på begge sider av grensen. ISIL hadde enorm fremgang i året som fulgte, og erobret store deler av Syria og Irak, inkludert Mosul, Iraks nest største by (Byman 2016, s. 135–136). I juni 2014 erklærte ISIL at de hadde opprettet et nytt kalifat som alle muslimer pliktet å underkaste seg. Siden “kalifatet” i teorien skulle omfatte hele den muslimske verden, fjernet gruppen stedsbetegnelsene i navnet og har siden kalt seg “Den islamske staten” (Stern og Berger 2016, s. 95–96).

Fra 2014 til 2018 gjorde IS seg internasjonalt bemerket med sin brutale fremferd, som innebar innføringen av en reaksjonær form for islamsk lovgivning, henrettelser og amputasjoner av fanger, folkemord på minoriteter, offentlig kjøp og salg av slaver, samt en rekke terrorangrep mot sivile mål, både i Midtøsten, Asia, Afrika og Europa. Samtidig ønsket IS å vise at de hadde etablert en sann islamsk “stat” og ikke bare en terrororganisasjon. Gruppen la derfor vekt på å lage et sivilt statsapparat og opprettet institusjoner for lov og rett, skatteinnkreving, utdanning, helsestell og så videre. Denne “sivile” siden av IS ble, i tillegg til volden og brutaliteten, et gjennomgangstema i IS’ propagandaproduksjon. IS’ påstand om å ha etablert en islamsk stat skulle vise seg enormt tiltrekkende for folk rundt om i verden som delte gruppens ideologiske utgangspunkt. Mer enn 40 000 sympatisører reiste fra andre land for å slutte seg til IS i Syria og Irak, deriblant rundt 100 nordmenn (Cook og Vale 2018, s. 17). På denne tiden opprettet IS også de første såkalte provinsene utenfor kjerneområdene i Syria og Irak, med IS-filialer blant annet i Libya, på Sinai-halvøya og i Nigeria.

IS nådde sitt territorielle høydepunkt i 2015, da gruppen kontrollerte rundt en tredjedel av både Irak og Syria. Deretter ble IS gradvis slått tilbake av en bred internasjonal koalisjon, inntil “kalifatet” var blitt redusert til en liten enklave i øst-Syria i 2018. Dette siste området ble inntatt av kurdiske styrker i begynnelsen av 2019. Slik gikk IS i Syria og Irak fra å være en “stat” tilbake til en tilværelse som undergrunnsorganisasjon.

Selv om IS har blitt stadig mer marginalisert i sine gamle kjerneområder siden den gang, har IS’ filialer i andre deler av verden hatt betydelig fremgang. Dette gjelder særlig IS-provinsene i Afrika – Sahel-provinsen, Vest-Afrika-provinsen, og den Sentral-Afrikanske provinsen – og den såkalte Khorasan-provinsen som er aktiv i Afghanistan. Slik kan IS stadig gi inntrykk av å være en global organisasjon, med en “kalif” som leder og med et tilsynelatende aktivt samarbeid mellom forskjellige fraksjoner, til tross for at store deler av toppledelsen, inkludert flere “kalifer”, har blitt drept de siste årene. Ikke minst opprettholder IS et enhetlig og fortsatt betydelig propagandaapparat, som vi skal se nærmere på i neste seksjon.

---

---

## 2.2 IS-propaganda

IS har revolusjonert måten ikke-statlige aktører driver propaganda på, og dette regnes som en av nøklene til gruppens suksess. Selv om jihadistgrupper som al-Qaida hadde produsert og distribuert propaganda på nett i en årrekke før IS slo seg opp i 2013–2014, fikk den sistnevnte gruppens propaganda langt større sprengkraft. Dette var det flere grunner til. Det at IS hadde kontrollen over et stort område, gjorde at de hadde mer penger og flere medlemmer som kunne jobbe som journalister, og de stadige kampene de var involvert i, gav dem mye råmateriale som kunne brukes i propagandaøyemed. Ikke minst var propaganda et tydelig prioritert område for IS som ble tildelt betydelige ressurser, som også bidro til at gruppen kunne opprettholde en betydelig større produksjon enn andre (Winter 2022, s. 13–14). Men første og fremst skilte IS seg ut ved at deres propaganda var *bedre* enn de andres. IS klarte å skape sin egen “merkevare” preget av høy produksjonskvalitet og et språklig og visuelt uttrykk som appellerte til de unge, kombinert med ekstremt brutalt innhold, der nærbilder av drap nærmest var en selvfølge. Samtidig var IS enda dyktigere enn andre grupper til å spre innholdet sitt i sosiale medier som Twitter, Facebook, YouTube, og etter hvert også på appen Telegram.<sup>4</sup> Slik fikk IS-propaganda enorm spredning i denne perioden, noe som er med på å forklare hvorfor så mange fremmedkrigere fra hele verden sluttet seg til gruppen.

I dag er det betydelig vanskeligere for IS å nå bredt ut med propaganda fordi de store sosiale mediene har begynt å ta ned slikt innhold omtrent med en gang det blir lagt ut. Likevel får IS spredt propaganda effektivt gjennom andre kanaler, og det er forholdsvis lett for sympatisører å finne frem til dem. IS’ nåværende distribusjonskanaler kan deles inn i tre hovedtyper: “private” sosiale medier som IS drifter på egne servere,<sup>5</sup> botter på det sosiale mediet Telegram, og vanlige, indekserte nettsider. Mens de private sosiale mediene har vært oppe i en årrekke, blir de to sistnevnte tatt ned fra nettet av administratorer og myndigheter med jevne mellomrom. Likevel blir de som regel gjenåpnet av IS under andre navn kort tid etterpå.

Oppsummert utgir IS fortsatt en jevn strøm av propagandatekster, videoer og bilder. Produksjonsvolumet er noe mindre enn det var i IS’ storhetstid, men det er fortsatt en betydelig produksjon av propaganda. Dette materialet er forholdsvis lett tilgjengelig for dem som ønsker å finne det, og det kjennetegnes stadig av høy produksjonskvalitet og ekstremt brutalt innhold.

## 2.3 IS-propaganda i forskningslitteraturen

Det er blitt gjort mye forskning på IS, og mye av denne litteraturen fokuserer spesifikt på gruppens propaganda. Et stort antall studier har tatt for seg IS’ skriftlige produksjoner – blant disse er det særlig mange som analyserer gruppens engelskspråklige magasiner *Dabiq* og *Rumiya*.<sup>6</sup> Flere studier har også gjort analyser av IS’ videoproduksjon, med fokus på hvordan gruppen iscenesetter sin egen brutalitet (Akil 2016; Euben 2017; Friis 2018; Mello 2018). Til tross for at bilder utgjør en stor del av IS’ samlede propagandaproduksjon, utgjør studier av disse bildene en liten del av den samlede

---

<sup>4</sup>Se for eksempel Prucha (2016) og Byman (2016, s. 147).

<sup>5</sup>Her har den viktigste kanalen i flere år vært en IS-styrt instans av *RocketChat*, en plattform for sosiale medier med åpen kildekode som man lett kan sette opp på sin egen server.

<sup>6</sup>Se for eksempel i oversikten i Tinnes (2023).

---

---

forskningslitteraturen på gruppens propaganda. Enkelte gode studier av IS' bildepropaganda finnes likevel, blant andre Milton (2018) og Baele, Boyd og Coan (2020).

Det viktigste enkeltbidraget til dette forskningsfeltet og det mest relevante for denne rapporten er Charlie Winters *The Terrorist Image* (Winter 2022). Boka presenterer en dybdestudie av IS' propagandabilder utgitt i perioden desember 2015 – september 2017, med spesielt fokus på narrativene og verdssystemet IS ønsker å formidle gjennom bildene. I den forbindelse gir også boka – i likhet med denne rapporten – en grundig gjennomgang og analyse av “temaene” som går igjen i materialet. Winter deler inn bildematerialet i hele 644 grunntemaer (*basic themes*), ordnet i 20 kategoriseringstemaer (*organising themes*), som igjen grovinndeles i to “globale temaer” (*global themes*) – *jihād* og *khilāfa* (kalifat). Hovedskillelinjen går mellom bilder som viser IS' militære aktiviteter i alle dets fasetter (*jihād*), og bilder som viser det sivile livet under IS' styre og bygger opp under gruppens påstand om å være en sann, islamsk stat (*khilāfa*). Dette hovedskillet – samt hovedtrekkene i den øvrige inndelingen – blir reproduisert av maskinlæringsmodellene vi tester, som vi diskuterer i delene 4.2.1 og 4.2.2. Siden *The Terrorist Image* jobber med en del av det samme materialet som denne rapporten og også foretar en tematisk analyse, utgjør den en svært nyttig grunnlinje å teste våre funn mot.

Samtidig skiller denne rapporten seg fra *The Terrorist Image* og den øvrige litteraturen på flere viktige punkter. Metodisk er dette den første studien som bruker maskinlæring til å analysere IS' propagandabilder. Her har Winter – som også analyserer det tematiske innholdet i bildene – kodet innholdet for hånd. Denne kodingen skiller seg også fra vår ved at den er basert primært på det *bildetekstene*<sup>7</sup> sier om det tematiske innholdet, og ikke på det faktiske innholdet i bildet, altså det man kan observere visuelt.<sup>8</sup> Vår koding baserer seg derimot kun på det visuelle innholdet.

Denne rapporten er den første som tar for seg utviklingen av IS' bildepropaganda etter “kalifatets” fall i 2018-2019. Vi fokuserer med andre ord ikke bare på perioden da IS var på høyden, slik tidligere studier har gjort, men også på det som har skjedd etterpå. Dette er viktig, fordi etter at “kalifatet” ble erklært i 2014, har IS faktisk operert lenger som gruppe *uten* å kontrollere territorium i Syria og Irak enn de gjorde *med*.<sup>9</sup> Til tross for at IS utgjør en vedvarende og økende trussel flere steder i verden, har vi ingen empiriske studier av hvordan bildepropagandaen har utviklet seg de siste årene. Vi behandler likevel ikke denne perioden isolert, men fokuserer på hele tidsspennet 2014-2022 for å forstå IS' bildepropaganda i sin helhet. Dette gir oss mulighet til å finne historiske kontinuiteter og brudd, og ikke minst forskjeller mellom IS' ulike provinser. Ingen tidligere analyser har basert seg på et like stort bildekorpus, eller et som dekker en lengre periode, og rapporten fyller dermed et viktig hull i forskningslitteraturen om IS.

---

<sup>7</sup>Hvert bilde har en bildetekst som “forklarer” hva man kan se på bildet, altså hvordan IS ønsker at mottakeren skal forstå innholdet i bildet.

<sup>8</sup>Winter baserer seg i førstehånd på *tittelen* til hver bilderapport når han kategoriserer (Winter 2022, s. 27). Denne tittelen er som oftest lik, eller forholdsvis lik, bildeteksten på hvert bilde i rapporten.

<sup>9</sup>Ikke medregnet at IS' forløpere hadde sporadisk territoriell kontroll over enkelte steder i Irak før “kalifatet” ble opprettet.

---

---

## 2.4 Om korpuset

Korpuset denne studien baserer seg på, kommer fra en av IS' arkiv-botter på det sosiale mediet Telegram.<sup>10</sup> Fra denne lastet vi ned et omfattende arkiv av IS' såkalte *taqārīr muṣawwara*, arabisk for "bilderapporter". Disse såkalte "bilderapportene" utgjør det overveldende flertallet av IS' samlede bildeproduksjon.<sup>11</sup>

Som vi går nærmere inn på i kapittel 5, viser bilderapportene forskjellige sider ved IS' aktiviteter – fra krigsscener og henrettelser av krigsfanger til vanlige gatebilder og barn på skolen. Disse er laget som *øyeblikksbilder* som skal formidle hva gruppen holder på med til enhver tid. Bildene gir et unikt innblikk i IS som organisasjon, i deres militære og ikke-militære aktiviteter, og i hvordan de over tid har ønsket å fremstille seg for omverdenen.

Hver "rapport" består typisk av mellom ett og tjue bilder (79 % av rapportene har ti eller færre bilder), med en median på fem. Det samlede korpuset består av over 30 000 bilder. Som regel viser alle bildene i hver rapport samme *scene*, f.eks. en skole, et torg, eller en slagmark. Ofte er ett av bildene i en rapport et slags "promo-bilde" som viser hva rapporten omhandler, og hvor bildene er tatt.<sup>12</sup>

Bilderapportene blir laget av lokale medieteam i hver av IS' provinser etter oppdrag fra IS' hovedkontor for medieproduksjon. Denne sentrale institusjonen har ansvar for å godkjenne hver rapport, og det er også denne som har eneansvaret for å publisere bildene gjennom IS' offisielle propagandakanaler.<sup>13</sup>

Siden bildene har blitt produsert jevnlig over mange år og hvert bilde (som regel) gir informasjon om publiseringsdato og -sted, gir dette korpuset en sjelden mulighet til å se utviklingen i IS' propaganda over tid. Det gir oss også muligheten til å sammenligne bildeproduksjon og tema i forskjellige IS-provinser. Samtidig er mengden bilder så stor at det er vanskelig å gjøre slike analyser rent kvalitativt. Det neste kapittelet beskriver hvordan vi har brukt maskinlæring for å gjøre en tematisk sortering av bildene som understøtter denne typen analysearbeid.

---

<sup>10</sup>IS har lastet opp arkiver med propaganda på Telegram. Disse arkivene er organisert etter sjanger, type, utgivelse, provins, og så videre. Hver bot inneholder det som fremstår som et komplett arkiv av én type propaganda.

<sup>11</sup>Selv om dette er den viktigste og desidert mest tallrike formen for bildepropaganda som IS produserer, er det ikke den eneste. IS publiserer også bilder i andre sammenhenger, for eksempel illustrasjonsbilder i magasinene *al-Naba* og *Dabiq*.

<sup>12</sup>Se figur 4.4, 4.5 og 4.6 for eksempelbilder.

<sup>13</sup>For en grundig innføring i hvordan IS' propagandaapparat fungerte i perioden 2014-2017, med ekstra fokus på bilderapportene, se Winter (2022, s. 13–26).

---

---

## 3 Metode

Dette kapittelet tar for seg maskinlæringsmetodene vi har brukt for å tolke og kategorisere bildene i korpuset.

Det metodiske hovedbidraget vårt er å kombinere to kjente metoder – en maskinlæringsbildemodell som tolker bildenes semantiske innhold, og en klyngealgoritme – for å kategorisere bildene i korpuset beskrevet i 2.4. Vi bruker bildemodellen CLIP, som representerer det semantiske innholdet i hvert bilde som en tallvektor. Vi bruker deretter klyngealgoritmer til å finne de vektorene som passer best sammen. Slik ender vi opp med en enkel måte å dele opp hele korpuset i tematiske grupper på, som igjen gjør det betydelig enklere å analysere innholdet.

### 3.1 Tolke bildenes innhold: CLIP

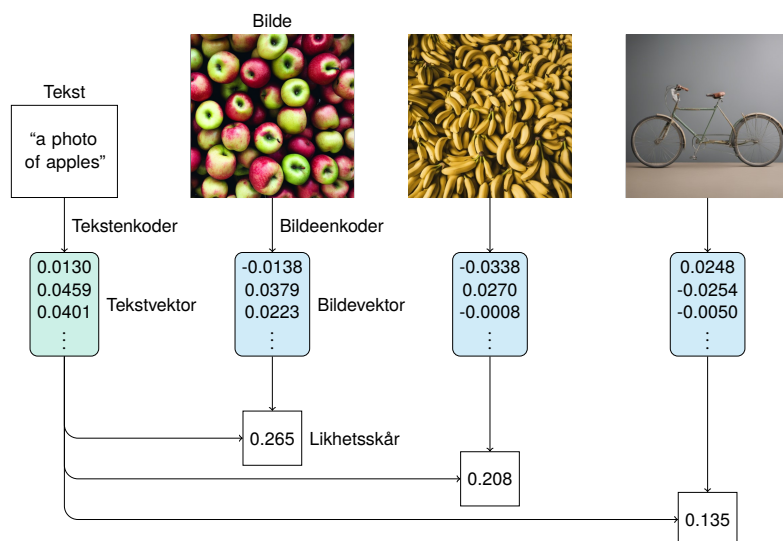
CLIP (Radford et al. 2021) – “Contrastive language-image pre-training” – er en fritt tilgjengelig maskinlæringsmodell utgitt av OpenAI som sammenligner bilder og tekst. I sin opprinnelige form tar modellen imot et bilde og en kort tekst og regner ut en likhetsskår for dem, som vist i figur 3.1. Den er trent slik at jo bedre egnet teksten er til å beskrive bildet, desto høyere skal likhetsskåren være. Treningsdataene<sup>14</sup> er hentet i en automatisk prosess fra Internett. Nærmere bestemt har man søkt gjennom offentlige nettsider og funnet bilder med tekstbeskrivelse. Til sammen er det samlet 400 millioner bilder, hvert med en tilhørende tekst.

Oppgaven som modellen er trent til, er å gi en høy skår dersom den blir presentert for et bilde–tekst-par som hører sammen i treningsdatasettet, og en lav skår dersom paret er tilfeldig sammensatt (figur 3.1). For å få til denne oppgaven var modellen nødt til å tilegne seg en viss forståelse for bilder og tekst.<sup>15</sup> Nærmere bestemt måtte modellen lære å kjenne igjen og fokusere på de samme egenskapene ved et bilde som et menneske ville gjort for å beskrive det i en bildetekst på Internett, og ignorere alt annet. Når vi i avsnitt 3.2 bruker CLIPs bildeforståelse til å dele inn vårt bildekorpus i kategorier, kan vi da håpe at disse egenskapene sammenfaller med det vi er interessert i, slik at bildene blir delt inn etter tema på en måte som samsvarer med en analytikers intuitive forståelse. Vi vil med andre ord bruke en modell som opprinnelig er trent til én oppgave – å sammenligne bilder med tekst, til en annen oppgave – å gruppere tematisk like bilder.

---

<sup>14</sup>I maskinlæring er *treningsdata* en samling med eksempler på inntutt som maskinlæringsmodellen kan få, samt ønsket utputt (fasit). Å trene modellen innebærer å optimalisere den for et konkret treningsdatasett slik at den – gitt eksempelinntutt – gir utputt som ligner på fasiten.

<sup>15</sup>Vi bruker begreper som vanligvis brukes om mennesker eller andre bevisste vesener i omtalen av slike modeller. For eksempel sier vi at modellen har “forståelse” eller “kunnskap”, eller at den “lærer” eller “prøver” å få til en oppgave. Vi mener ikke med dette å antyde at en slik matematisk modell har noen form for bevissthet, men bruker disse begrepene fordi de er nyttige metaforer for å beskrive det som foregår.



Figur 3.1 CLIP er trent til å sammenligne et bilde med en tekst. Jo bedre de passer sammen, desto høyere skal likhetsskåren være, som vi ser i dette eksempelet. Legg merke til at bananbildet gir en høyere likhetsskår i sammenligning med teksten enn sykkelpbildet, mens eplebildet gir høyest skår. Dette passer intuitivt med at bananer «ligner» mer på epler semantisk enn en sykkel gjør. Tallene i figuren er reelle, men vi viser her bare de første 3 tallene av de totalt 768 i hver bilde- og tekstvektor.

### 3.1.1 Arkitektur

CLIP består av to moduler: en bildeenkoder og en tekstenkoder. Bildeenkoderen er et nevralt nett som tar et bilde som innputt, og som utputt gir den en mer kompakt representasjon, eller oppsummering, av bildet. Rent faktisk er en slik representasjon en lang liste med tall, også kalt en vektor. I denne rapporten kaller vi denne representasjonsvektoren for en bildevektor. Nevralt nett har arkitekturen ViT (vision transformer, eller synstransformer) (Dosovitskiy et al. 2020). Tekstenkoderen er også et nevralt nett som lager en representasjonsvektor, men i motsetning til bildeenkoderen tar den tekst som innputt. I denne rapporten kaller vi representasjonsvektoren for en tekstvektor. Tekstenkoderen har også en transformer-arkitektur (Vaswani et al. 2017).

Når CLIP skal sammenligne et bilde med en tekst, er det i virkeligheten bildevektoren og tekstvektoren som sammenlignes med hverandre. Det fins flere mål på hvor like to vektorer er. CLIP bruker *cosinuslikhet*, som er definert i (3.1).

$$\text{cosinuslikhet}(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} = \frac{\vec{x}}{\|\vec{x}\|} \cdot \frac{\vec{y}}{\|\vec{y}\|} \quad (3.1)$$

Operatoren “ $\cdot$ ” viser til skalarproduktet av to vektorer, det vil si at man multipliserer elementene i vektorene parvis og summerer opp disse produktene. Navnet “cosinuslikhet” viser til at den kan tolkes geometrisk som cosinus til vinkelen mellom de to vektorene  $\vec{x}$  og  $\vec{y}$ . Dersom de peker i samme retning, er likheten 1, og dersom de peker i motsatt retning, er likheten -1. Dersom de står vinkelrett på hverandre, er likheten 0.

For at det å sammenligne bilde- og tekstvektorene skal være meningsfullt, er bildeenkoderen og tekstenkoderen trent sammen. Under trening ble kombinasjonen av de to enkoderne behandlet

---

---

som ett stort nevralt nett. Cosinuslikheten, eller likhetsskåren, ble behandlet som nettverkets utputt, som ble sammenlignet med fasit. Nettopp dette er den store fordelingen ved å trene modellen på denne måten: Man slipper å fortelle modellen hvordan bilde- og tekst-vektorene skal se ut; dette ble automatisk bestemt gjennom treningsprosessen. Gjennom at enkoderne i samarbeid lærte å produsere riktige likhetsskåre, har hver av dem også lært å produsere nyttige bilde- og tekstvektorer. Dette betyr at vi nå kan anvende bildeenkoderen *alene* til å prosessere bilder og få bildevektorer som forhåpentlig fanger den mest interessante informasjonen i bildene. Disse har vi brukt til å sammenligne bilder med hverandre og gruppere dem tematisk.

CLIP-modellen er utgitt i to ulike varianter, som hver fins i ulike størrelser. Vi bruker “ViT-L/14@336px”, som er den største blant dem som bruker arkitekturen ViT i bildeenkoderen, og som er den kraftigste av modellene som OpenAI har utgitt.<sup>16</sup> Denne modellen opererer med representasjonsvektorer med lengde 768. Å kjøre denne prosessen på de ca. 30 000 bildene tok en halv til én time ved bruk av et grafikkort. Dette trenger vi bare å gjøres én gang. Når bildevektorene først er regnet ut, kan de lagres og brukes mange ganger.

### 3.1.2 Klassifikator med manuelle klasser

CLIP kan også brukes på andre måter. Inspirert av forsøkene til Radford et al. (2021) forsøkte vi å bruke *både tekst- og bildeenkoderen* til å lage en bildeklassifikator som deler bildene inn i kategorier som vi hadde definert manuelt. Vi skrev korte tekstbeskrivelser av hver kategori, for eksempel: “a photo of war spoils”, “a photo showing a display of fighters or weapons” og “a photo of a celebration or a social event”. Vi skrev tekstene på engelsk fordi det meste av tekstmaterialet i CLIPs treningsdata er på engelsk. Så brukte vi CLIP til å sammenligne bildene i korpuset vårt med hver av disse tekstene og plassere dem i kategorier ut ifra hvilken tekst som passet best.

Vi fant imidlertid at en slik klassifikator var lite egnet til å analysere vårt bildekorpus. Dels var dette fordi resultatene ikke var veldig gode, men først og fremst fordi det var vanskelig å definere egnede kategorier uten å kjenne bildematerialet godt på forhånd. Vi valgte derfor å ikke bruke tekstenkoderen, men isteden bare bildevektorer kombinert med klynging, som beskrevet i 3.2.

## 3.2 Finne lignende bilder: klynging

For å skape orden i den store mengden bilder deler vi dem inn i grupper etter tematikk. Dette kan vi gjøre automatisk ved hjelp av eksisterende algoritmer som deler vektorer inn i grupper kalt *klynger* (engelsk: *clusters*). Vektorene får vi fra CLIPs bildeenkoder, som konverterer bilder til vektorer, som beskrevet i 3.1.

En klynge er i denne sammenhengen et utvalg vektorer som er relativt like hverandre. Litt mer presist kan vi si at oppgaven til en klyngealgoritme er å ta en mengde vektorer og gruppere dem i klynger slik at vektorene innad i hver klynge ligner mer på hverandre enn på vektorer i andre klynger. Denne definisjonen er fortsatt ikke helt presis. Blant annet avhenger den av hvilket mål

---

<sup>16</sup>Andre har reproduisert OpenAIs arbeid og utgitt enda større modeller av samme type, se spesielt Cherti et al. (2023).

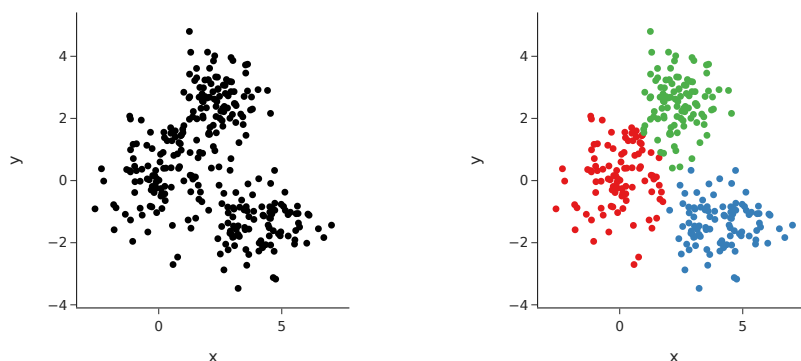
---

---

man bruker på «likhet» mellom to vektorer. Ulike måter å tolke denne oppgaven på kan resultere i klyngealgoritmer som gir svært ulikt resultat, som vi skal se.

Vi kunne kjørt klyngealgoritmene på bildenes pikselverdier direkte og dermed ikke trengt å lage en vektor med CLIP. Dette ville imidlertid ført til at vi sammenlignet bildene med hverandre piksel for piksel og grupperte bilder som har lignende farger på de samme stedene i bildet. Dette kan være nyttig dersom man ser etter identiske bilder, men er lite hensiktsmessig når vi er interessert i bildenes semantiske innhold. For eksempel vil pikselverdiene i et bilde av en person være svært avhengig av lysforholdene, fargen på klærne til personen eller den nøyaktige plasseringen av elementene i bildet. Isteden bruker vi altså vektorene produsert av CLIPs bildeenkoder. Som beskrevet i 3.1 koder disse vektorene for viktig semantisk informasjon i bildene istedenfor «irrelevant» informasjon som nøyaktig plassering av elementer i bildet. De vil derfor sannsynligvis være bedre egnet til å dele inn bildene i klynger som gjenspeiler tematikken deres.

I forbindelse med klynging er det vanlig å tenke på vektorer som punkter i et flerdimensjonalt rom (én dimensjon for hvert tall i vektorene). Vektorene danner da en sky av punkter. Å dele inn vektorene i klynger kan tolkes som å dele inn dette rommet i områder. Dersom antall dimensjoner ikke er mer enn tre, kan vi visualisere dette i en figur. Et eksempel på dette i to dimensjoner er vist i figur 3.2. Siden vi jobber med vektorer av lengde 768, bruker vi algoritmen t-SNE for å redusere dette til to dimensjoner, slik at vi kan inspisere resultatet i en todimensjonal figur. t-SNE er beskrevet i 3.2.1.2.



(a) Uten inndeling i klynger

(b) Med inndeling i klynger, vist som ulike farger

*Figur 3.2 Eksempel på todimensjonale vektorer visualisert som punkter i todimensjonalt rom med x- og y-koordinater. Klyngeinndelingen vist her vil ikke nødvendigvis være den “beste” for et gitt formål.*

### 3.2.1 Klyngealgoritmer

Her beskriver vi hva de ulike algoritmene vi har brukt, gjør, og hvordan vi har brukt dem. Vi bruker Python-pakken “scikit-learn” sin implementasjon av  $k$ -snitt, DBSCAN og t-SNE.



---

---

### 3.2.1.1 *k*-snitt

*k*-gjennomsnitt-klynging, eller *k*-snitt-klynging (engelsk: *k*-means clustering) (Lloyd 1982) er en mye brukt klyngealgoritme. Gitt et forhåndsdefinert antall klynger *k* prøver den å minimere variansen innad i hver klynge.<sup>17</sup> De *k* klyngene er definert ved hvert sitt *klynge*sentrum slik at hver vektor tilhører den klyngen som den ligger nærmest sentrumet til. Algoritmen er iterativ; den begynner med en tilfeldig inndeling og forbedrer denne litt og litt for hver iterasjon.

Første steg i algoritmen (initialiseringssteget) er å velge *k* vilkårlige vektorer som klyngecentrum for hver sin klynge. Alle de øvrige vektorene tilhører nå klyngen til det sentrumet de ligger nærmest, men enn så lenge er denne inndelingen sannsynligvis svært lite optimal. Neste steg (iterasjonssteget) er å flytte sentrumet for hver klynge til gjennomsnittet av alle vektorene i klyngen.<sup>18</sup> Dermed ligger nå “sentrumet” rent faktisk i sentrum av vektorene i sin klynge. Når vi nå har flyttet klyngecentrene, vil gjerne noen vektorer ha flyttet seg fra én klynge til en annen fordi et annet sentrum kom nærmere eller sentrumet til deres opprinnelige klynge flyttet seg lenger unna. Deretter gjentas iterasjonssteget igjen og igjen. Det kan bevises at klyngene alltid etter hvert vil stabilisere seg slik at det til slutt ikke lenger er noen vektorer som flytter seg fra en klynge til en annen. Når dette skjer, er algoritmen ferdig. Nå har hvert klyngecentrum to roller: Det er lik gjennomsnittet av alle vektorene i klyngen sin samtidig som det sammen med de andre sentrene definerer klyngenes grenser.

Selv om klyngene til slutt stabiliserer seg, er det ikke garantert at løsningen er optimal. Dette avhenger av hvilke vektorer som ble valgt ut i initialiseringssteget til å være de første klyngecentrene. For å bøte på dette er det vanlig å utføre algoritmen flere ganger der man trekker nye initiale klyngecentra hver gang, og beholde det beste resultatet.

Vi bruker minibatch-*k*-snitt-klynging (eng: minibatch *k*-means clustering) (Sculley 2010), en variant av *k*-snitt-klynging som i hver iterasjon av algoritmen ser på et utvalg av vektorene istedenfor alle hver gang. Dette gjør algoritmen raskere når man jobber med et stort antall vektorer, men til gjengjeld finner den gjerne løsninger som er noe mindre optimale.

Man må selv velge antallet klynger som vektorene skal deles inn i, altså *k*. Et høyt antall klynger vil gi en finkornet inndeling, og tilsvarende vil et lavt antall klynger gi en grovere inndeling. Vi ønsker å finne en passelig mellomting slik at inndelingen blir mest mulig nyttig i vår analyse, som beskrevet i 4.2.1. Å velge riktig *k* blir i vårt tilfelle vanskeliggjort av at vi ikke kjenner dataene inngående fra før – å bli kjent med dataene er nettopp grunnen til at vi ønsker å lage klynger.

En kjent måte å velge *k* på er å prøve mange klynginger med hver sin verdi av *k* og visualisere resultatet ved å plote summen av klyngevariansene (som algoritmen prøver å minimere) som funksjon av *k*. Figur 3.3 viser en slik visualisering. Da kan det vise seg en eller flere verdier av *k* som utpeker seg som spesielt gunstige. Generelt ønsker man en lav varians, men ikke på bekostning av et overdrevent stort antall klynger. Jo høyere man setter *k*, desto lavere varians kan man få, ettersom det er mulig å dele inn vektorene mer og mer finkornet. Dette kulminerer når antall klynger er lik antall vektorer, slik at hver klynge har nøyaktig én vektor, og variansen blir null. En slik inndeling er perfekt i den forstand at hver klynge har null varians, men verdiløs fordi én klynge

---

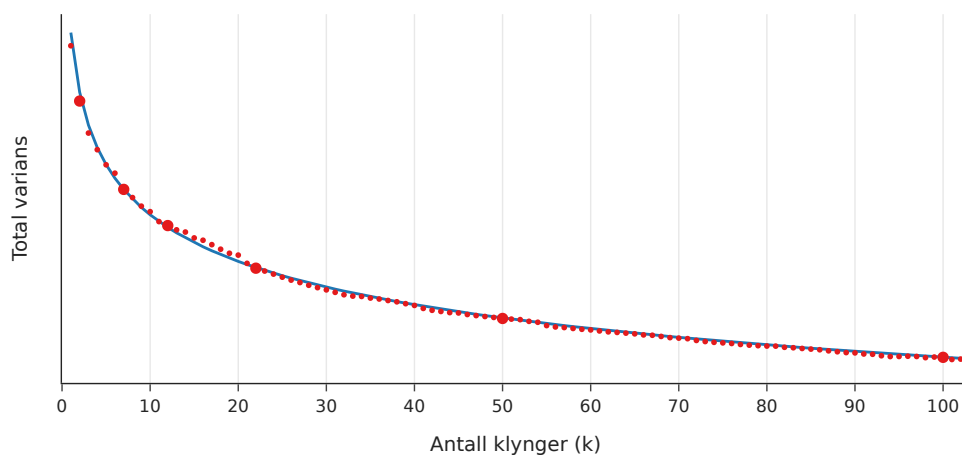
<sup>17</sup>Dette er ekvivalent med å minimere den euklidske avstanden fra hver vektor til dens klynge sentrum, nærmere bestemt summen av kvadratet av hver av disse avstandene. Intuitivt kan man forstå det som at algoritmen prøver å gi hver vektor kortest mulig vei til sentrum av en klynge.

<sup>18</sup>Å finne gjennomsnittet av en samling vektorer innebærer å finne gjennomsnittet av hvert enkelt tall i vektoren.

---

---

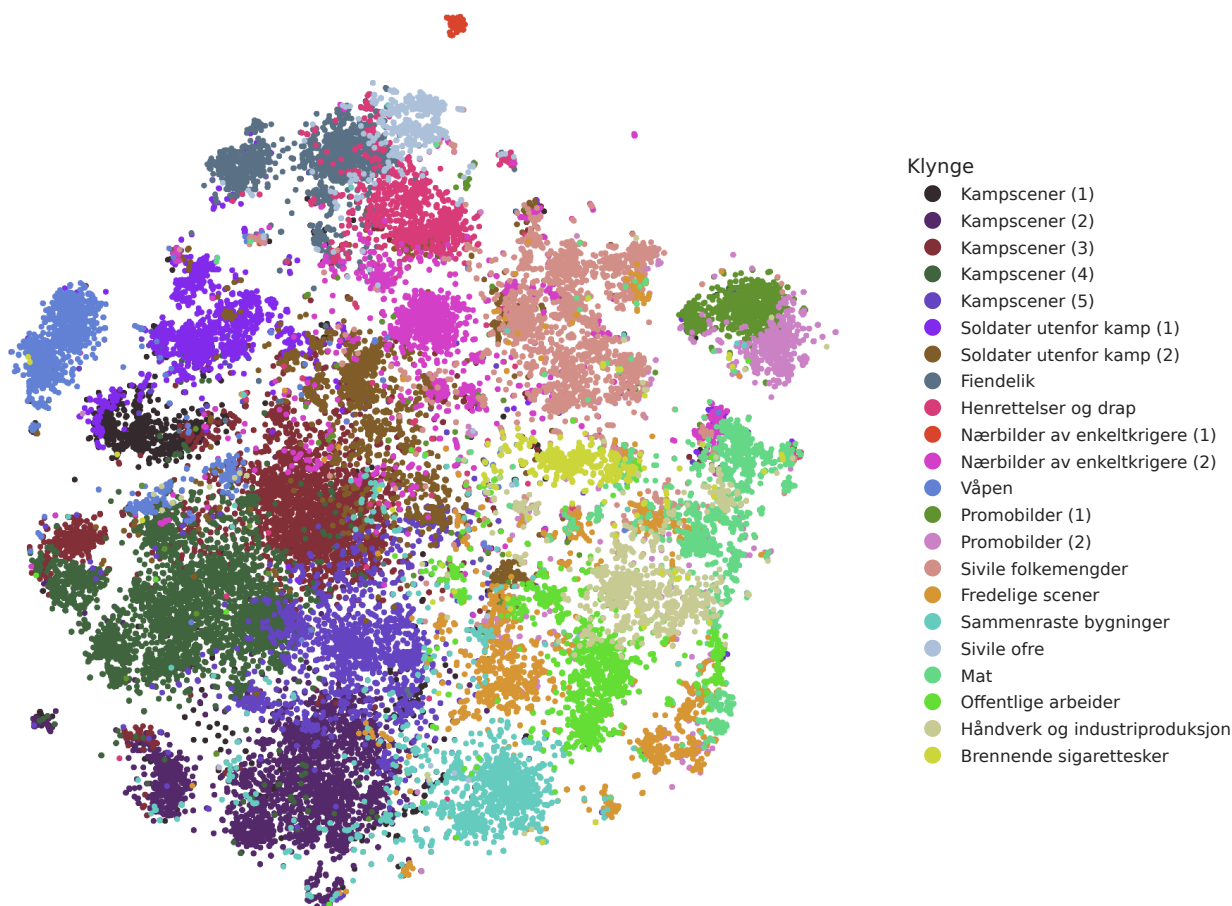
per vektor er en triviell løsning som ikke gir noen innsikt. Når vi ser på grafen, ser vi altså etter en relativt lav verdi av  $k$  som gir relativt lav varians, slik at å øke  $k$  mer gir en relativt liten nedgang i variansen. Visuelt kunne vi sett at grafen synker raskt fram til denne verdien, mens den synker sakte etterpå. Dersom det eksisterer en slik god verdi, kan det bety at vektorene er klumpet sammen i naturlige grupper, og at vi har funnet riktig antall klynger slik at klyngene vi finner, stemmer med disse gruppene. Dette er nettopp det vi ønsker oss.



*Figur 3.3 De røde punktene viser den totale variansen (som  $k$ -snitt prøver å minimere) for hver verdi av  $k$ . De verdiene av  $k$  vi vurderte manuelt, er markert med større punkter. Den blå kurven viser til sammenligning den overordnede trenden, funnet ved regresjon ( $f(k) = 12134k^{-0,107}$ ). Vi ser at punktene følger trenden tett, og dermed at ingen utmerker seg som spesielt gode.*

Vi konstruerte en slik graf i et forsøk på å finne en god verdi av  $k$  (figur 3.3). Dessverre framgår det ikke av grafen at det fins noen verdi av  $k$  som markerer seg som spesielt gunstig. Tvert imot går variansen jevnt og forutsigbart nedover med økt  $k$ , kanskje med unntak av området rundt 13–20, der variansen er noe høyere enn den øvrige trenden skulle tilsi. Vi tar dette som et tegn på at datasettet ikke er tydelig delt inn i naturlige grupper som vi kan identifisere med  $k$ -snitt-klynging.

Vi valgte isteden å vurdere et utvalg verdier av  $k$  manuelt ved å studere bilder i hver klynge og vurdere hvor nyttig inndelingen så ut til å være. Vi valgte verdier av  $k$  logaritmisk omtrentlig jevnt fordelt, bare justert noen steder for å treffe punkter der grafen i 3.3 synker raskere enn i området rundt. Vi valgte disse verdiene av  $k$ : 7, 12, 22, 50 og 100, samt 2. Vi konkluderte med at  $k = 22$  ga den mest nyttige inndelingen for vår analyse, noe vi kommer tilbake til i 4.2.1. Figur 3.4 viser  $k$ -snitts inndeling av bildevektorene i 22 klynger.



Figur 3.4 Inndelingen av bildevektorene i 22 klynger utført med  $k$ -snitt. Her vises klyngenavnene som vi etablerer i 4.2.2. De 768-dimensjonale vektorene er redusert til 2 dimensjoner ved hjelp av metoden  $t$ -SNE for å kunne gjøre denne visualiseringen. Vi kan se at det er en viss sammenheng mellom plassering og bildenes tema. Blant annet har militære temaer grovt sett havnet til venstre og sivile temaer til høyre, hvilket svarer til inndelingen i to klynger beskrevet i 4.2.1. “Fiendelik”, “Henrettelser og drap” og “Sivile ofre” ligger også nær hverandre. Reduksjonen til to dimensjoner gjør at det ser ut som at klyngene overlapper hverandre. I det opprinnelige 768-dimensjonale rommet er disse klart avgrenset, slik som i figur 3.2b. Det kan likevel ligge mange vektorer nær disse grensene, og disse kan vi se på som tvilstilfeller.

---

---

### 3.2.1.2 *t-SNE + DBSCAN*

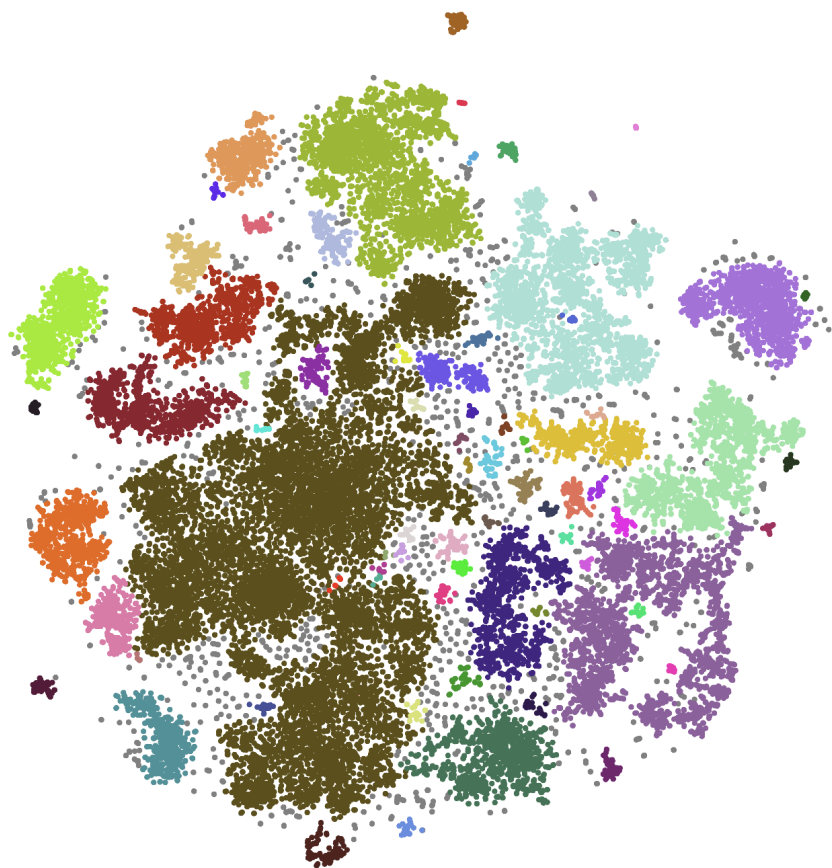
DBSCAN (Ester et al. 1996) er en annen klyngealgoritme. Intuitivt kan man si at den finner sammenhengende områder med relativt høy tetthet av vektorer og plasserer vektorene i et slikt område i samme klynge. I motsetning til  $k$ -snitt sine klynger, som alltid er konvekse i formen, kan DBSCAN sine klynger ha mye mer varierte former, og en klynge kan for eksempel være viklet inn i eller omringe en annen. Mer presist identifiserer DBSCAN områder med høy tetthet ved å finne alle vektorer som har minimum et visst antall andre vektorer innenfor en viss avstand  $\epsilon$  fra seg. Dette antallet og denne avstanden må defineres av brukeren som parametere. Vektorene som blir funnet, kalles for *kjernevektorer*, og dersom to kjernevektorer ligger innenfor en avstand på  $\epsilon$  fra hverandre, plasseres de i samme klynge. Dette betyr at dersom avstanden mellom vektorene  $\vec{a}$  og  $\vec{b}$  og avstanden mellom  $\vec{b}$  og  $\vec{c}$  er mindre enn  $\epsilon$ , vil også  $\vec{a}$  og  $\vec{c}$  være indirekte koblet gjennom  $\vec{b}$  og dermed høre til samme klynge. Dette vil fortsette så lenge det fins en slik «kjede» av kjernevektorer som ligger nær hverandre. Vektorer som ikke er kjernevektorer, blir plassert i klyngen til nærmeste kjernevektor dersom avstanden til denne er mindre enn  $\epsilon$ . Slike vektorer ligger i utkanten av en klynge. Dersom det derimot ikke er noen kjernevektor innenfor denne avstanden, ligger vektoren utenfor klyngene, og den blir ansett som støy og ikke tildelt noen klynge. Til sammenligning har ikke  $k$ -snitt-klynging noe konsept om støy og kan la seg sterkt påvirke av vektorer som ligger langt borte.

Det viste seg svært vanskelig å finne verdier av de to parameterne til DBSCAN som fungerte godt på datasettet med vektorer i 768 dimensjoner. Vi forsøkte derfor å først bruke *t-SNE* til å redusere antall dimensjoner til to<sup>19</sup> for deretter å kjøre DBSCAN på disse reduserte vektorene. Vi kaller denne framgangsmåten for *t-SNE + DBSCAN*. *t-SNE* er en av flere algoritmer for å overføre en mengde vektorer fra et vektorrom med mange dimensjoner til et med færre dimensjoner samtidig som man forsøker å bevare informasjon om vektorenes plassering i forhold til hverandre. *t-SNE* forsøker å bevare vektorenes nærhet til hverandre. Det vil si at vektorer som ligger nært hverandre i det opprinnelige rommet, også skal gjøre det i det nye, og motsatt for vektorer som ligger langt fra hverandre. Å redusere antall dimensjoner vil i de fleste tilfeller innebære tap av informasjon, og det er definitivt mye informasjon som går tapt når vi reduserer vår mengde av vektorer fra 768 til to dimensjoner. *t-SNE* har likevel en viss evne til å bevare naturlige “klynger” på bekostning av global posisjon i rommet. Av figurene 3.4 og 3.5, der vi har brukt *t-SNE*, kan man også se at *t-SNE* har bevart flere områder med høy tetthet av vektorer. Derfor var det en mulighet for at å bruke DBSCAN til å finne disse områdene ville gi klynger som kunne være nyttige i en analyse. Se figur 3.5 for en visualisering av resultatene av dette.

Under manuell inspeksjon av klyngene ble det klart at klyngene fra *t-SNE + DBSCAN* var mindre enhetlige og mindre tolkbare enn klyngene fra  $k$ -snitt. Vi valgte derfor å ikke bruke disse i analysen. Av figur 3.5 kan vi se en grunn til at disse klyngene ikke ble så nyttige: Flere av klyngene ser ut til å være satt sammen av flere undergrupper som noen steder ligger så tett inntil hverandre at de har “smeltet” sammen. Disse klyngene strekker seg over så store områder at det er usannsynlig at bildene i hver ende hører naturlig sammen, og dermed at vi ikke vil ønske å skille mellom dem i analysen.

---

<sup>19</sup>Selv med bare tre dimensjoner var det vanskelig å få DBSCAN til å gi et godt resultat.



*Figur 3.5 Eksempel på inndeling av bildevektorene i klynger utført med t-SNE + DBSCAN. Vi har ikke gitt navn til disse klyngene. De grå punktene er vektorer som ikke kom med i noen klynge. De 768-dimensjonale vektorene er redusert til 2 dimensjoner ved hjelp av t-SNE for å kunne gjøre denne visualiseringen, og den samme reduksjonen er brukt som innputt til DBSCAN. En konsekvens av dette er at klyngene ikke kan overlape i figuren, hvilket kan gi et misvisende inntrykk av at denne inndelingen er bedre enn k-snitt-inndelingen (vist i figur 3.4). Punktenes plassering er den samme her som i figur 3.4, så det er mulig å sammenligne.*

---

---

Vi forsøkte å unngå slike store klynger ved å justere parameterne til DBSCAN, men da ble det enten urimelig mange vektorer som ikke ble tildelt klynge, eller det ble upraktisk mange og veldig små klynger. Selv i figur 3.5 fins det på samme tid både veldig store klynger og mange små klynger som knapt er synlige i visualiseringen. Totalt er det 78 klynger i denne inndelingen, og av disse er det bare 16 som utgjør minst 1 % hver av de rundt 30 000 bildevektorene. Til sammenligning gjelder dette 21 av de 22 klyngene vi fikk med  $k$ -snitt.

DBSCAN sin evne til å finne klynger med irregulære former ble altså i vårt tilfelle mer til ulempe enn til hjelp, ettersom vi var ute etter klynger der alle bildene ligner så mye på hverandre som mulig.<sup>20</sup> Dersom vi hadde fått til å bruke DBSCAN direkte på de 768-dimensjonale vektorene, kan det hende at de ekstra dimensjonene ville gjort det mulig å dele opp de store klyngene på en god måte.  $t$ -SNE + DBSCAN med et passende valg av parametere viste seg imidlertid å være svært godt egnet til å identifisere “promo-bildene” fra bilderapportene for å filtrere dem bort. Vi kan se disse bildene tydelig isolert oppe til høyre i figur 3.5.

### 3.2.2 Tolke klyngene

En viktig del av analysen var å studere resultatet fra klyngealgoritmene og vurdere om klyngene var rimelige, og hva som kjennetegnet bildene i hver av dem. For å slippe å se på alle bildene i hver klynge hver gang, så vi etter metoder som kunne hjelpe oss å raskere få et overblikk over klyngene.

#### 3.2.2.1 Se på bildene

En naturlig metode er å trekke noen bilder tilfeldig fra hver klynge og se på bare dem. Antallet bilder må være stort nok til at de til sammen gir et representativt inntrykk av hele klyngen. Vi begynte med å trekke 5–10 bilder fra hver klynge, men det viste seg at dette ikke alltid var nok til å få en god representasjon av klyngen. Det var såpass stor variasjon i noen av klyngene at mer sjeldne tilfeller hadde for lav sannsynlighet for å bli representert.

En metode som viste seg svært nyttig, var å regne ut avstanden fra hvert bildes representasjonvektor til dens respektive klyngesentrum og vise alle bildene i klyngen sortert på denne avstanden. Vi bruker den euklidske avstanden, som i vårt tilfelle alltid ligger mellom 0 og 2,<sup>21</sup> og for det meste mellom 0,3 og 0,8 (figur 4.2). Ved å bla gjennom en slik liste kan vedkommende raskt danne seg et overblikk over (1) bildene som ligger nærmest sentrum (i den ene enden av listen), (2) dem som ligger lengst fra sentrum (i den andre enden), og (3) dem som ligger i “typisk” avstand fra sentrum (midten av listen, der medianavstanden vil ligge). Disse gruppene kan vi tolke som henholdsvis (1) de mest prototypiske bildene for klyngen, (2) bildene helt i utkanten av klyngen, som viser den maksimale variasjonen, og (3) «alminnelige» bilder som viser den normale variasjonen. En slik tolkning er spesielt relevant for  $k$ -snitt-klynging, ettersom klyngesentrumet ikke bare representerer

<sup>20</sup>Klynger der alle vektorene ligner så mye på hverandre som mulig, er jo også nettopp det  $k$ -snitt optimerer for.

<sup>21</sup>Dette er fordi alle vektorene er L2-normalisert slik at de har lengde (norm) 1. Geometrisk vil det si at de ligger på et 768-dimensjonalt kuleskall. Den største avstanden som det er mulig å oppnå mellom to vektorer vil dermed være om de er på motsatte sider av kulens sentrum (origo). Avstanden vil da være lik avstanden fra den første vektoren til origo pluss avstanden fra origo til den andre vektoren. Siden begge vektorene har lengde 1, er kulens radius også 1, og avstanden blir  $1 + 1 = 2$ .

---

---

en typisk vektor i klyngen, men faktisk definerer klyngen. Analytikeren kan selv vurdere hvor mange bilder det er nødvendig å se på i hvert område av listen for å danne seg et tilstrekkelig godt inntrykk. Man kan svært raskt få et overblikk over hva klyngen inneholder, men man kan også velge å se på flere bilder når klyngeinndelingen ser lovende ut, og man vil få et mer nøyaktig inntrykk.

Vi vurderte det slik at bildene langt fra sentrum var forstyrrende elementer, ettersom de etter vårt skjønn ikke hadde en meningsfull tilknytning til klyngen for øvrig. Mens en metode som DBSCAN selv identifiserer vektorer som ikke passer inn i noen klynge, vil  $k$ -snitt-klynging tvinge også disse inn i en klynge, og det er nettopp slike tilfeller vi kan finne i vår sorterte liste som bilder med stor avstand til klyngesentrumet. Vi brukte slike lister til å bestemme oss for en terskelavstand og ekskludere bildevektorer som hadde en større avstand til klyngesenteret enn denne terskelen, som nevnt i 4.2.1.

### 3.2.2.2 Generere nye bilder

Et alternativ til å vurdere bilder fra datasettet er å generere nye bilder som kan fungere som eksempler. Slik kan vi se hvordan et bilde potensielt kunne sett ut på steder i bildevektorrommet der det ikke er noen eksempler i datasettet. Vi kan for eksempel få et prototypisk eksempel på en klynge ved å generere et bilde ut ifra et klyngesentrum. Det vil typisk være mange bilder i klyngen som ligger nært dette sentrumet, men i praksis vil ingen ligge nøyaktig i sentrum. Også dette er spesielt meningsfullt når vi ser på  $k$ -snitt-klynging, ettersom det da er klyngesentrumene som definerer klyngene. Å studere slike genererte prototyper fra hver klynge kan dermed gi en viss innsikt i hvordan bildene er inndelt. Dersom avstander i bildevektorrommet svarer til semantisk ulikhet mellom bildene, slik vi håper, kan vi intuitivt forstå at et enkeltbilde plasseres i klyngen til den prototypen det ligner mest på. Vi kan også forstå at det vil finnes gråsoner mellom disse prototypene der det er uklart hvilken klynge et bilde skal tilhøre, samt at det kan finnes bilder som ikke ligner på noen av prototypene, men som likevel vil bli tvunget inn i en av klyngene. Dette kan igjen hjelpe oss å vurdere om inndelingen er god, og om vi har valgt riktige parametere for klyngingen, f.eks. antall klynger for  $k$ -snitt-klynging.

Vi har generert slike bilder ved hjelp av bildegenereringsmodellen Stable UnCLIP 2.1<sup>22</sup>, en variant av Stable diffusion 2.1 (Rombach et al. 2022) inspirert av UnCLIP fra Ramesh et al. (2022). I motsetning til den opprinnelige varianten av denne modellen, som genererer et bilde ut ifra en tekst, genererer denne bilder ut ifra en CLIP-bildevektor, slik som dem vi bruker.<sup>23</sup> Mens CLIPs bildeenkoder tar et bilde og regner ut en representasjonsvektor (eller bildevektor), kan vi forstå det som at denne bildegenereringsmodellen går motsatt vei og gjetter hvordan bildet kan ha sett ut, gitt representasjonsvektoren. Dette er bakgrunnen for navnet *UnCLIP*. Vi bruker ordet “gjetter” fordi det i prinsippet vil finnes mange ulike bilder som kodes til samme representasjonsvektor.<sup>24</sup>

<sup>22</sup><https://github.com/Stability-AI/stablediffusion>

<sup>23</sup>Vi bruker L2-normaliserte bildevektorer, mens UnCLIP bruker ikke-normaliserte vektorer. Derfor ganger vi våre vektorer med en typisk vektorlengde før vi gir dem til UnCLIP.

<sup>24</sup>Vi kan innse dette ved å legge merke til at representasjonsvektoren, som med 768 tall og 2 B per tall utgjør 1,5 kB, er drastisk mindre enn et typisk bilde fra datasettet, som har en oppløsning  $1280 \times 720$  og dermed inneholder ca. en million piksler, hver på tre tall og 1 B per tall, til sammen 2700 kB, eller rundt 2000 ganger så mye som representasjonsvektoren. Selv etter JPG-komprimeringen som er brukt på bildene, er de rundt 100 ganger så store som representasjonsvektorene. Det er rett og slett ikke på langt nær nok lagringsplass i representasjonsvektoren til å kunne lagre all informasjon ved et av disse bildene. Ekvivalent kan vi si at det ikke fins mange nok mulige representasjonsvektorer til at alle mulige bilder kan få tildelt en hver, og med andre ord må flere bilder dele på samme representasjonsvektor.

---

---

Modellen kan kjøres flere ganger på samme bildevektor og gi en ny gjetning hver gang. Så kan vi se hva som er felles for disse genererte bildene. Siden de kommer fra samme vektor, kan vi forvente at likheter mellom dem skriver seg fra informasjon i bildevektoren, og at eventuelle ulikheter er et resultat av gjettingen, som vi ønsker se bort ifra. Bildene vi har generert, kan ses i appendiks A.

Verken CLIPs bildeenkoder eller UnCLIP er perfekte. Begge kan ha svakheter som gjør at de fungerer dårlig i visse sammenhenger, for eksempel på grunn av treningsdata som er mangelfulle eller ikke representerer de bildene man egentlig er interessert i. Når vi går fra bilde til bildevektor (med CLIP) og tilbake til bilde (med UnCLIP), kan vi observere forandringer mellom start- og sluttbildet, som ideelt sett skulle vært like. Det er da vanskelig å vite om det var CLIP eller UnCLIP som introduserte disse forandringene, for eksempel ved at CLIP ikke lagret et visst aspekt ved bildet i bildevektoren, eller ved at UnCLIP ikke klarte å tyde den eller ikke visste hvordan den skulle rekonstruere bildet. Vi har sett tydelige eksempler på dette i tilfeller der UnCLIP skal gjengi groteske og voldelige bilder, som det er mye av i IS' propaganda. De genererte bildene i disse tilfellene blir nedtonet og ofte svært urealistiske og mangler det groteske innholdet. Denne effekten er tydelig i figur A.1f og A.2c.

Vi mistenker at dette skyldes at slike bilder var lite eller ikke representert i treningsdataene til UnCLIP, og at den derfor ikke har sett noen eksempler på dette i sin trening. Ettersom utgiverne av UnCLIP ikke har gått ut med informasjon om treningssettet, er dette spekulasjon. Det er imidlertid vanlig i treningen av slike bildegenereringsmodeller, blant annet andre Stable diffusion-modeller, at man forsøker å rense datasettet for ting man ikke ønsker at modellen skal kunne generere. Det er derfor ikke usannsynlig at det er gjort også i dette tilfellet. En annen mulighet er at slike groteske bilder som IS publiserer, er svært utypiske for bilder på Internett forøvrig, og at de derfor er dårlig representert i datasettet som UnCLIP er trent på. Det er også mulig at CLIPs bildeenkoder ikke klarer å fange opp disse elementene. Vi mener at det er usannsynlig fordi vi under klyngingen og øvrig analyse av bildevektorene ser at det skilles på disse tingene, også mellom grupper av bilder som ser like ut i UnCLIP sine rekonstruksjoner. Det er også mindre naturlig at CLIP sitt datasett er renset for uønskede bilder på samme måte, fordi det ikke er en genererende modell og dermed ikke står i fare for å generere "problematisk" bilder. At UnCLIP ikke klarer denne oppgaven, vil naturligvis være en fordel innen bruksområder der man vil unngå å lage groteske bilder, men er en ulempe for oss som ønsker å undersøke nettopp slike.

### **3.2.3 Bunntekster**

De fleste bildene har bunntekster – grafikk med tekst, provins og dato langs bunnen. I vår analyse er denne med når bildene kodes om til bildevektorer av CLIPs bildeenkoder. I og med at vi gjør en egen analyse av innholdet i bunntekstene (beskrevet i 4.1.1) kan det tenkes at det er en fordel å holde disse utenfor i bildeanalysen og se bare på fotografidelen av bildet. Dette for å unngå at bildeenkoderen plukker opp stilen på bunnteksten eller det står der, og at dette dermed påvirker klyngingen. Vi forsøkte å fjerne bunnen av bildene, lage nye bildevektorer og utføre klyngingen på nytt. Vi klarte imidlertid ikke å se noen meningsfull forandring i resultatet og brukte ikke disse nye klyngene i analysen. At bunntekstene ikke påvirker resultatet i stor grad, støttes også av at UnCLIP konsekvent ikke rekonstruerer dem i bildene den genererer. Det hender at man ser spor av dem, men da bare i form av noe som ligner på tekst på tilfeldige steder i bildet. Av og til ligner dette på arabisk tekst, men ofte ligner det på latinske bokstaver. Vi tar disse observasjonene som tegn på at CLIP ikke vektlegger dette i nevneverdig grad, eller iallfall ikke på en måte som forstyrrer våre resultater.



---

---

## 4 Strukturering av datamaterialet

Maskinlæringsmetodene som ble presentert i kapittel 3, er nyttige verktøy for å få oversikt over store datamengder, men selve analysen gjøres ikke programmatisk. For å analysere datamaterialet må man først bestemme hva man ønsker å undersøke. I vårt tilfelle ønsker vi å studere utviklingen til IS' bildepropaganda, og hva det kan fortelle oss om utviklingen til IS som gruppe, som beskrevet i introduksjonen til kapittel 2. Deretter må analytikeren vurdere hvilke faktorer som er relevante, og strukturere dataene på en hensiktsmessige måte for å finne ut av det man er ute etter.

Dette kapittelet har to hoveddeler. Første del viser hvordan vi gikk frem for å ekstrahere metadata om publiseringstidspunkt og -sted for hvert av bildene, og diskuterer deretter valgene vi måtte ta for å operasjonalisere disse dataene i analysen. Andre del viser hvordan vi har strukturert dataene som maskinlæringsmetodene CLIP og  $k$ -snitt gav oss. Seksjon 4.2.3 gir en oversikt over de tematiske klyngene vi bruker i analysen, med eksempler på hver klynge. Vi avslutter med en test på hvor godt CLIP og  $k$ -snitt faktisk klarer å klynge innhold på en måte som gir mening for en analytiker, og diskuterer hvilke forbehold som bør tas, i seksjon 4.2.5.

### 4.1 Finne metadata

For å kunne gjøre en analyse av hvordan det tematiske innholdet i bildene utvikler seg over tid, og for å utforske forskjeller mellom IS' forskjellige provinser, ønsket vi å finne informasjon om hvert bildes publiseringstidspunkt og -sted. Vi kombinerte ulike metoder for å finne denne informasjonen, som var lettere eller vanskeligere tilgjengelig avhengig av designen på hvert bilde og av hvorvidt metadataene i hver enkelt fil var blitt endret før publisering.

#### 4.1.1 Finne informasjon i bildeteksten: *optisk tegngjenkjenning*

For mange av bildene i korpuset var alle metadataene vi var ute etter – publiseringstidspunkt og -sted – oppgitt på selve bildet. For å hente ut denne informasjonen brukte vi *optisk tegngjenkjenning*<sup>25</sup> – en metode for å lese ut tekstdata av bilder. Vi brukte verktøyet Google Document AI,<sup>26</sup> som er regnet som det beste til å arbeide med arabisk skrift (Hegghammer 2022). Vi brukte kun *bunnteksten* i bildet i den optiske tegngjenkjenningen, fordi hvis det var annen tekst i selve bildet (for eksempel på et skilt eller en plakate), ble også denne gjenkjent, og det førte til unødvendig rot i det tekstmaterialet vi ønsket å ekstrahere.

Mengden informasjon vi kunne ekstrahere kun med optisk tegngjenkjenning, var avhengig av den grafiske utformingen på hvert bilde. Det var lett å finne all metainformasjon for bildene som var publisert etter 2018, fordi IS da innførte en enhetlig grafisk utforming for alle bilderapporter, med bildetekst, provinsnavn og dato i bunnteksten på hvert bilde (figur 4.1a og 4.1b). Dette var derimot betydelig vanskeligere for den delen av korpuset som er publisert i perioden 2014–2017. På denne

---

<sup>25</sup>Engelsk: *Optical Character Recognition (OCR)*

<sup>26</sup>Tilgjengelig på <https://cloud.google.com/document-ai>

tiden stod de lokale mediesenterne – altså de som produserte bilderapporter i hver provins – fritt til å velge egen grafiske utforming (se figur 4.1c og 4.1d for eksempler på forskjellige utforminger). Disse har likevel alltid en beskrivende billedtekst og en logo som viser hvilken provins bildet er fra.<sup>27</sup> Slik kunne vi bruke optisk tegngjenkjenning til å finne *publiseringssted* for disse bildene (se 4.1.2).

For bildene som ikke oppgav publiseringsdato, måtte vi skaffe denne på andre måter, som beskrevet i seksjon 4.1.3.



(a) Eksempel på den enhetlige utformingen som alle IS-provins har brukt siden 2018. Dette bildet er fra provinsen Irak, 2020. Bunnteksten viser (fra høyre til venstre), navnet på provinsen, deretter en kalligrafisk fremstilling av provinsnavnet, dato (måned og år), og en billedtekst som beskriver det som foregår på bildet.



(b) Eksempel på den enhetlige utformingen, denne gangen med provinsnavn på engelsk, fra provinsen Vest-Afrika, 2022.



(c) Eksempel på en design fra provinsen Raqqa fra 2014. Logo med provinsnavn i arabisk kalligrafi og engelsk tekst til høyre, billedtekst i midten, ingen dato



(d) Eksempel på en design fra provinsen Nineveh fra 2015. Logo med provinsnavn i kalligrafi til høyre, billedtekst i midten. Helt til venstre i bunnteksten er det en kalligrafisk fremstilling av “Den islamske staten”.

Figur 4.1

<sup>27</sup> Mediesenteret til IS’ undergruppe Khalid Ibn al-Walid brukte en design helt ulik de andre, der provinsnavnet ikke var i bunnteksten, men i en egen logo oppe til høyre i bildet. Siden logoen var på samme sted hver gang, kunne vi lett klassifisere bildene som tilhørende dette mediesenteret, ved å se om pikslene i denne delen av bildet hadde en viss farge.

---

---

### 4.1.2 Koding av provinser

Ved å bruke optisk tegngjenkjenning fant vi informasjon om provinstillhørighet for de fleste bildene. Det viste seg videre at det originale datamaterialet vi lastet ned fra Telegram, hadde plassert hver rapport i en egen mappe, og navnet på mappen indikerte som oftest hvor bildene kom fra. Med disse to datakildene fikk vi informasjon om provinstillhørighet for nesten alle bildene. Vi måtte likevel gjøre enkelte kvalitative grep for å kunne bruke denne informasjonen i analysen, fordi IS-provinsene har endret seg mye over tid.

I utgangspunktet tilsvarte IS-provinsene omtrent de eksisterende administrative provinsene (tilsvarende norske fylker) i Irak og Syria. Men ettersom IS ekspanderte utover Irak og Syria fra 2014, begynte betydningen av det å være en “provins” å endre seg. Enkelte steder – som i Jemen og Libya – fortsatte disse IS-provinsene å overlappes med eksisterende administrative enheter i landet. Andre IS-provinser kunne omfatte hele stater – som IS-provinsene Algerie og Tunisia – eller hele regioner, som IS’ vestafrikanske provins, som omfattet to distinkte grupper i henholdsvis Nigeria og Sahel-landene Mali, Burkina Faso og Niger. Etter kalifatets fall i 2018 sluttet man å operere med distinkte provinser i Syria og Irak og gikk over til å omtale hvert av landene som provinser. Fra 2022 har enkelte av de større provinsene blitt delt opp i mindre enheter – for eksempel har IS-Sahel blitt separert fra IS Vest-Afrika, mens IS Mosambik, som tidligere var regnet som del av IS Sentral-Afrika, nå også blir omtalt som en egen gruppe. Med andre ord, fra å være en rent administrativ inndeling i Syria og Irak har det å være en IS-provins blitt synonymt med at en lokal jihadistgruppe var anerkjent som en offisiell “filial” av IS i utlandet.

For å gjøre det mulig å sammenligne produksjonen fra forskjellige IS-provinser over tid har vi valgt å operere med så store provinser som mulig. Det vil si at vi regner alle IS-provinser i Syria som én syrisk provins for å kunne sammenligne produksjonen fra før og etter 2018. På samme måte regner vi IS-Sahel som en del av IS Vest-Afrika, selv om førstnevnte begynte å utgi bildepropaganda i eget navn i 2022. Tabell 4.1 viser en oversikt over provinsene vi ser under ett på denne måten. Figur 5.2 viser hvor mange bilder som hører til hvert område.

### 4.1.3 Finne og kode datoer

For de bildene som ikke oppgav dato i bunnteksten, måtte vi ty til andre metoder for å finne publiseringstidspunkt. Vi fant at mange av bildene kunne plasseres i tid ved å se på modifikasjonstidspunktet i bildefilenes metadata. Dette viste seg å ikke være helt pålitelig, for mange av filene hadde tilsynelatende blitt modifisert etter at de opprinnelig ble produsert – så mange som en tredel av bildene hadde metadata som tilsa at de var modifisert i mars 2020, og disse datoene måtte vi altså se bort fra. Det viste seg imidlertid at bildene var publisert i kronologisk rekkefølge i det originale datamaterialet vi hadde lastet ned fra Telegram. Dermed kunne vi se hvor bildene uten tidsangivelse lå i denne kronologiske rekkefølgen, og vi kunne sammenligne med bilder før og etter i rekkefølgen som *hadde* tidsangivelse enten fra OCR eller metadata. Slik kunne vi anslå publiseringstidspunkt også for disse bildene. Tidspunktet er som regel oppgitt i måned og år, men vi har valgt å kun bruke året i analysen som følger. Dette er fordi de fleste av datoene er oppgitt i den islamske *hijri*-kalenderen, der månedene skiller seg fra våre. Det var derfor ikke mulig å finne nøyaktig hvilken av “våre” måneder et gitt bilde er publisert i, hvis kun den islamske måneden er oppgitt.

Provinser		Område
Al-Badiya	Damascus	Syria
Al-Baraka	Hama	
Aleppo	Hawran	
Al-Jazeera	Homs	
Al-Khayr	Idlib	
Al-Raqqa	Jaysh Khalid Ibn al-Walid	
Al-Sham	Eufrat	
Al-Anbar	Kirkuk	Irak
Al-Junub	Nineveh	
Baghdad	Nord-Baghdad	
Diyala	Saladin	
Falluja	Tigris	
Irak		
Aden	Lahij	Jemen
Aden-Abyan	San'a	
Al-Bayda	Shabwa	
Hadramaut	Jemen	
Barqa	Libya	Libya
Fezzan	Tripoli	
Vest-Afrika	Sahel	Vest-Afrika
Sentral-Afrika	Mosambik	Sentral-Afrika
Khorasan	Pakistan	Khorasan

Tabell 4.1 Oversikt over provinsene vi har slått sammen til større områder i analysen. Bare områder med mer enn 50 bilder vises.

---

---

Enkelte bilder ble tildelt datoer som ikke kunne stemme – det vil si at de var registrert med et årstall fra før IS ble opprettet, eller fra etter datoen bildet ble lastet ned. Disse feilene skyldes sannsynligvis unøyaktigheter i den optiske tegngjenkjenningen. Vi valgte følgelig å se bort fra disse bildene når vi gjorde analyser av utvikling over tid i 5.1 og 5.2.

## 4.2 Tematiske klynger

### 4.2.1 Valg av antall klynger og tematisk konsistens

I kapittel 3 brukte vi CLIP til å lage en vektorrepresentasjon av hvert bilde og brukte deretter algoritmen  $k$ -snitt til å “klynge” disse bildevektorene. Slik fikk vi kategorisert bildene i tematisk like klynger.

Som beskrevet i 3.2.1.1 finnes det ingen fasit på hvor stor  $k$  man skal operere med – altså hvor mange tematiske klynger man skal ha. Med for få klynger risikerer man at det blir stor intern variasjon i hver klynge, og man kan dermed gå glipp av viktig informasjon i datamaterialet ved at bilder som skiller seg ut tematisk, *ikke* blir skilt ut i en egen klynge. Videre kan analysen bli upresis hvis man trekker konklusjoner på bakgrunn av store klynger med lite tematisk konsistens. Bruker man derimot for mange klynger, kan mange av dem være fordelt etter temaer som er helt irrelevante for analysen. Det er også viktig å merke seg at jo flere klynger man opererer med, jo vanskeligere blir det å få en rask oversikt over materialet, og det er tross alt dette som er poenget med å bruke denne metoden.

I 3.2.1.1 kom vi frem til at 22 klynger gav oss et godt kompromiss mellom analytisk presisjon og tematisk relevans. Det betyr ikke at andre  $k$  ikke kan gi oss nyttig informasjon. Ved inndeling i bare to klynger fikk vi for eksempel en grovinndeling mellom militær/voldelig og ikke-militær tematikk, tilsvarende Charlie Winters inndeling av IS-propaganda i hovedtemaene *jihād* og *khilāfa* beskrevet i 2.3. Slik kunne vi enkelt bruke  $k$ -snitt til å finne et av de viktigste særtrekkene i dette bildematerialet.

Selv om de 22 klyngene var forholdsvis tematisk konsistente, var det fortsatt en del variasjon, og ikke alle bildene hørte åpenbart under samme “tema” som de andre bildene i sin klynge. Som beskrevet i 3.2.2.1 forsøkte vi å måle hvor tematisk likt hvert bilde var sin klynge, ved å regne ut hvor nærme hver bildevektor var sentrum i sin klynge.

For å sikre høy integritet i analysen valgte vi å utelate bilder som ikke var sentrale nok i sin klynge, fra analysen av tematisk innhold.<sup>28</sup> Vi bestemte oss for en terskelavstand på 0,6, der kun bilder med kortere avstand til sentrum enn dette ble medregnet. Målet med terskelavstanden var at rundt 90 % av bildene skulle være plassert i en klynge som analytikeren kunne gjenkjenne som tematisk “riktig”. Her er det selvsagt en glidende overgang, og tallet 0,6 var basert på en kvalitativ vurdering av når mer enn hvert tiende bilde fremsto tematisk ulikt resten av klyngen.<sup>29</sup> Vi måler hvor robust denne inndelingen er, i 4.2.5.

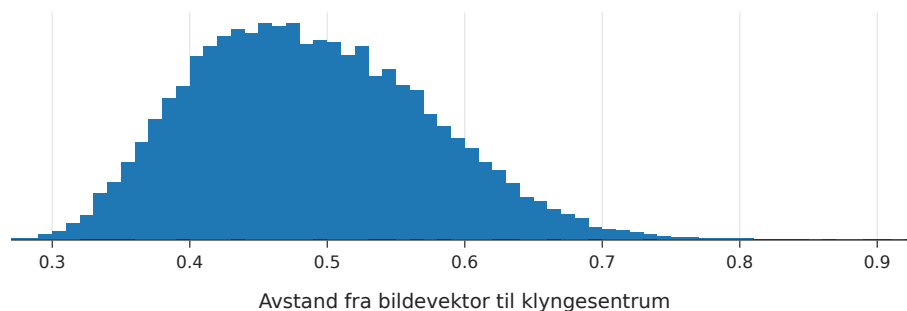
<sup>28</sup>Det vil si at vi beholder disse i analysen når vi kun ser på totalt antall bilder per år og per provins i seksjon 5.1.

<sup>29</sup>Denne terskelen kunne blitt satt separat for hver klynge for å kompensere for ulik variasjon innad i klyngene. Vi valgte imidlertid å sette en felles terskel for alle klynger for å unngå unødvendig kompleksitet og fordi det ville krevd en egen vurdering for hver klynge. Dermed ville metoden dårligere latt seg generalisere til et større antall klynger.

---

---

Bildene som ble utelatt, utgjorde rundt 10 % av det totale antallet. Fordelingen av bildenes avstand til klyngesentrumene kan ses i figur 4.2.



Figur 4.2 Histogram over avstanden fra hver bildevektor til dens klyngesentrum. Vi ekskluderte alle bilder med en avstand større enn 0,6 fra tematisk analyse.

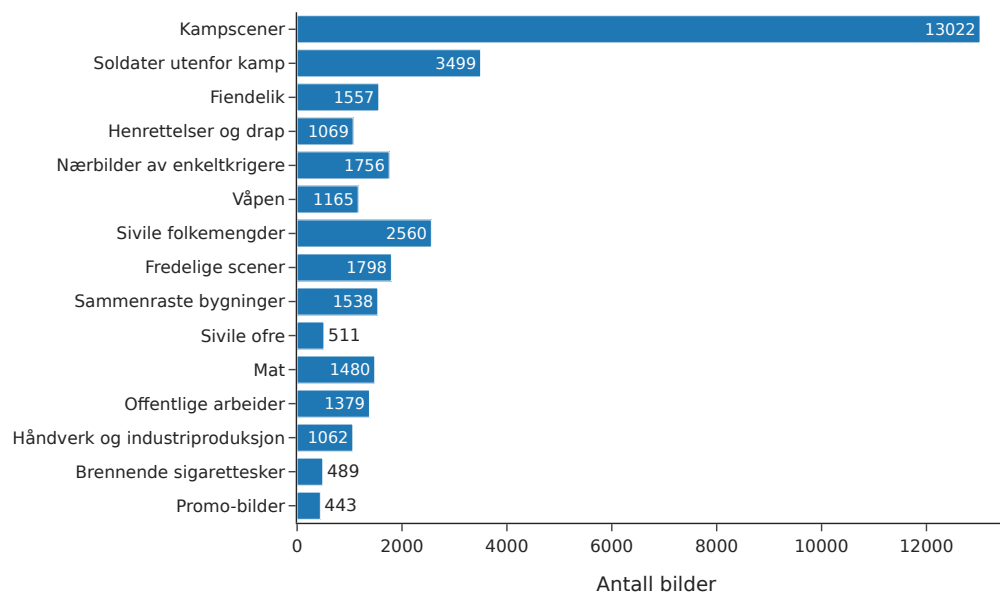
#### 4.2.2 Strukturering av tematiske klynger

Når klyngealgoritmen fordeler bildene inn i klynger etter tematisk likhet, er det ikke alltid åpenbart for et menneskelig øye hva det er som er “temaet” i en gitt, programmatisk definert klynge. Dette må til syvende og sist avgjøres simpelthen ved å tolke og observere innholdet, selv om også dette arbeidet kan understøttes programmatisk, for eksempel ved å rangere bilderevektorene etter avstand til klyngesentrum, eller ved å generere nye bilder basert på vektoren til klyngesenteret, som beskrevet i 3.2.2. Vi anså alle de 22 klyngene vi endte opp med å bruke, som analytisk relevante, det vil si at de viste ting som var interessant for oss å vite om datamaterialet. Likevel måtte vi gjøre enkelte grep for å kunne bruke disse klyngene videre i analysen.

Først gav vi hver klynge et beskrivende navn, selv om det ikke nødvendigvis er fullstendig dekkende for det tematiske innholdet. Dette gjorde vi fordi det er lettere for en analytiker å operere med slike navn enn nummererte klynger. Dette gav oss også muligheten til deretter å slå sammen enkelte av de originale klyngene under et felles navn. Dette var hensiktsmessig fordi noen av klyngene var tematisk så like hverandre at det gav mening å se dem under ett i den analysen vi skulle gjøre, selv om skillelinjene kunne ha vært nyttige i en annen type analyse. Vi endte opp med tre slike “metaklynger” som er satt sammen av flere av de originale klyngene. Dette gjelder “metaklyngen” *kampscener*, som er satt sammen av fem andre klynger, *soldater utenfor kamp*, som er satt sammen av to klynger, og *nærbilder av enkeltkrigere*, som også er satt sammen av to klynger. Disse vil bli beskrevet i mer detalj i det som følger. For enkelthets skyld kommer vi fra nå av til å omtale de opprinnelige klyngene og de manuelt definerte “metaklyngene” kollektivt bare som *klynger*.

Vi valgte også å fjerne to klynger fra analysen som bestod av såkalte “promo-bilder” for andre bildeserier (se figur 4.6g). Promo-bildene er tematisk ulike hverandre, men siden de har en design som er veldig ulik de andre bildene i korpuset, blir de likevel sortert sammen. Selv om det i og for seg er nyttig å få skilt ut promo-bildene på denne måten, er selve bildene lite relevante for de spørsmålene vi ønsker å svare på.

Dermed endte vi til slutt opp med 14 navngitte klynger til analysen (figur 4.3).



Figur 4.3 Antall bilder i hver metaklynge før ekskludering av bilder

Disse klyngene ligner på, men er ikke identiske med, de 20 kategoriseringstemaene Charlie Winter opererer med i den tidligere nevnte studien *The Terrorist Image* (se 2.3). I noen tilfeller har vi slått sammen klynger som Winter betrakter som separate kategorier, for eksempel kategoriene *offensive operations* og *projectile warfare*. Disse blir identifisert som forskjellige klynger av algoritmen, men vi har valgt å slå dem sammen i metaklyngen *kampscener*. I andre tilfeller er bildene så visuelt like at klyngealgoritmen ikke skiller mellom dem, mens Winter kan dele dem inn i forskjellige kategorier ut ifra bildeteksten. Dette gjelder blant annet de visuelt like kategoriene *offensive* og *defensive operations*, som vi har i klyngen *kampscener*, samt flere av Winters “sivile” kategorier; her blir blant annet bilder som Winter kategoriserer som *outreach*, *religious life* og *education* typisk plassert i klyngen *sivile folkemengder*. Enkelte av klyngene blir også behandlet separat av oss, mens de figurerer som del av samme kategori hos Winter. Dette gjelder blant andre *våpen* og *fiendelik*, som Winter plasserer sammen under kategorien *aftermath* (altså ting som skjer i etterkant av et angrep). Oppsummert er våre automatisk inndelte klynger forholdsvis like Winters, kategorier til tross for at de har blitt laget helt uavhengig av hverandre. Utover enkelte ulike tolkningsvalg, skyldes forskjellene først og fremst at vår modell baserer seg på det visuelle innholdet i bildene, mens Winters inndeling baserer seg på bildetekstene.

I tillegg til å gi våre klynger beskrivende navn, introduserte vi to manuelt definerte binære variabler for hver klynge. Dette var for å nærmere undersøke enkelte temaer som har fått mye oppmerksomhet i litteraturen om IS-propaganda. Dette er de binære variablene vi introduserer:

1. Militær eller sivil tematikk (1/0)
2. Ekstremvold (1/0)

Den første variabelen er lagt til for å undersøke i hvilken grad IS fremstiller seg som en sivil styresmakt. Som nevnt i kapittel 2 har IS lagt stor vekt på å fremstille seg som en “stat” og ikke bare en militær- eller terrororganisasjon. Vi forventer at forekomsten av sivil tematikk vil fortsette å gå ned i bildene som er publisert etter 2017.

Variabelen “ekstremvold” refererer til groteske bilder av drap, lik, henrettelser og så videre. Vi tar med denne variabelen fordi dette kanskje er det IS-propaganda er mest kjent for, og vi ønsker å undersøke hvordan dette temaet fordeler seg over tid og fra provins til provins.

### 4.2.3 Oversikt over tematiske klynger

Vi gir her en beskrivelse av hver av de tematiske klyngene. Tabell 4.2 viser en oversikt med navn, antall bilder i hver klynge og angitte binære variabler.

Navn	Antall	Militær	Ekstremvold
Kampscener	13022	✓	
Soldater utenfor kamp	3499	✓	
Fiendelik	1557	✓	✓
Henrettelser og drap	1069		✓
Nærbilder av enkeltkrigere	1756	✓	
Våpen	1165	✓	
Sivile folkemengder	2560		
Fredelige scener	1798		
Sammenraste bygninger	1538		
Sivile ofre	511		
Mat	1480		
Offentlige arbeider	1379		
Håndverk og industriproduksjon	1062		
Brennende sigarettesker	489		
Promo-bilder	443		

Tabell 4.2 Oversikt over de navngitte klyngene med antall bilder og binære variabler.

*Kampscener:* Kampscener er det desidert vanligste temaet i bildematerialet. Denne “metaklyngen” består av fem andre klynger som viser forskjellige typer kamper. Her er det interessant å merke seg at modellen kan plukke ut relativt små forskjeller mellom bildene. For eksempel er hovedskillet mellom to av de originale klyngene etter alt å dømme om det blir skutt med stort eller lite kaliber (se figur 4.4c og 4.4d). Å kategorisere bilder etter slike skillelinjer kan være nyttig i noen sammenhenger, men i dette analysearbeidet var det mest hensiktsmessig å slå dem sammen.

*Soldater utenfor kamp:* Denne “metaklyngen” viser soldater som gjør ting som ikke er direkte relatert til kamp. Dette kan være soldater som leser Koranen, feirer Id, poserer med våpen, sverger troskap til en ny leder, og så videre. “Metaklyngen” er satt sammen av to nær identiske klynger, der hovedskillet mellom dem er at soldatene i den ene klyngen som regel er fra Midtøsten (figur



---

---

4.4f), mens soldatene i den andre klyngen er fra en av IS' afrikanske provinser. Det er nærliggende å tenke at både etnisiteten til personene på bildet og forskjeller i landskap og farger har bidratt til at disse har blitt kategorisert i separate klynger. Dette reflekteres også i de genererte bildene fra disse klyngene (figur A.1d og A.1e).

*Fiendelik:* Denne klyngen viser blodige, ille tilredte lik av IS' fiender, ofte tatt etter et slag eller trefning. Bildene er ofte svært groteske, og klyngen kodes som ekstremvold.

*Henrettelser og drap:* Denne klyngen viser henrettelser, drap og fanger. De fleste bildene viser drap på tilfangetatte soldater og kodes derfor som militær – selv om enkelte av bildene også viser henrettelser av sivile i IS-territorium. Flere av bildene er svært groteske og viser blant annet halshugginger og folk som blir kastet ned fra bygninger. Her kan man også legge merke til at *k*-snitt også har gruppert enkelte bilder av sauer som blir slaktet i denne klyngen, men ikke mange nok til at det vil gjøre stort utslag i analysen.

*Nærbilder av enkeltkrigere:* Denne "metaklyngen" viser nærbilder av enkeltkrigere, som oftest (men ikke alltid) i konteksten at soldaten på bildet er blitt "martyr". Bildene er uansett tatt *før* disse er blitt martyrer – IS viser kun ille tilredte lik av sine fiender. Metaklyngen er blitt satt sammen av to ulike klynger med likt tematisk innhold, men der den ene klyngen har en spesiell grafisk utforming.

*Våpen:* Denne klyngen viser våpen og annet utstyr som blir stilt ut – konteksten er gjerne at dette er våpen IS har tatt i krigsbytte.

*Sivile folkemengder:* Denne klyngen viser sivile folkemengder i forskjellige kontekster. Det kan være i moskeen, på skolen, på markedet, osv. Mange av bildene viser folkemengder som overvårer offentlige avstraffelser eller henrettelser – en viktig del av IS' propaganda.

*Fredelige scener:* Denne klyngen viser rolige gatebilder og naturbilder. IS brukte slike bilder for å vise at kalifatet var et fint sted å bo, og at det ikke bare var krig og elendighet, og klyngen kodes derfor som sivil.

*Sammenraste bygninger:* Denne klyngen viser sammenraste – som regel utbombede – bygninger. Vi koder denne klyngen som sivil, fordi den som regel viser bygninger i IS' egne territorier som har blitt bombet av koalisjonsstyrker. IS viser frem disse bildene for å vise hvor brutalt deres fiender behandler muslimske sivile.

*Sivile ofre:* Denne klyngen viser skadde eller døde sivile, ofte barn. Vi koder denne klyngen som sivil fordi bildene som regel viser sivile i IS' territorium som er blitt såret eller drept av koalisjonsstyrkenes bombeangrep. Selv om innholdet i disse bildene kan være brutalt og groteskt, kodes det ikke som ekstremvold, fordi disse ikke er voldshandlinger som IS selv har utført. Altså er ikke målet å skremme fienden med IS' brutalitet, men å skape sympati med IS' sak ved å vise hvor brutalt deres fiender behandler muslimske sivile.

*Mat:* Denne klyngen viser mat som blir produsert eller tilberedt, matvarer som blir vist frem på markeder, og mat som blir spist. Denne klyngen er kodet som sivil fordi de fleste bildene viser dagligliv i IS' territorier. Hensikten er blant annet å vise at IS klarer å skaffe nok mat til de som lever under deres styre. En del av bildene viser også mat som blir tilberedt i felten, og dette gjør det noe vanskeligere å klart kode klyngen som rent sivil.

---

---

*Offentlige arbeider:* Denne klyngen viser sivile som arbeider ute – veiarbeid, bygningsarbeid, innhøsting osv.

*Håndverk og industriproduksjon:* Denne klyngen viser også sivile som arbeider, som regel innendørs, med håndverk og industri.

*Brennende sigarettesker:* Denne klyngen viser at IS brenner sigarettesker. IS anser sigaretter, alkohol og narkotika som umoralsk, og kampen mot dette har vært en viktig symbolsak i IS' sivile administrasjon.

#### **4.2.4 Eksempelbilder fra klyngene**

Figurene 4.4, 4.5 og 4.6 viser et eksempelbilde fra hver klynge. Vi har valgt det bildet som er nærmest sentrum i sin klynge, for å få et representativt eksempel (jamfør 3.2.2.1 og 4.2.1).



(a) Eksempelbilde fra første klynge i metaklyngen kampscener. Denne klyngen viser angrep i mørke.



(b) Eksempelbilde fra andre klynge i metaklyngen kampscener. Denne klyngen viser eksplosjoner sett på lang avstand.



(c) Eksempelbilde fra tredje klynge i metaklyngen kampscener. Denne klyngen viser skyting med mindre våpen.



(d) Eksempelbilde fra fjerde klynge i metaklyngen kampscener. Denne klyngen viser skyting med tungt kaliber.



(e) Eksempelbilde fra femte klynge i metaklyngen kampscener. De fleste bildene i denne klyngen viser kampscener der minst én bil er involvert.



(f) Eksempelbilde fra andre klynge i metaklyngen soldater utenfor kamp. Disse bildene er som regel fra Midtøsten, mens den første klyngen i metaklyngen viser soldater fra IS' afrikanske provinser.

Figur 4.4



(a) Eksempelbilde fra klyngen fiendelik



(b) Eksempelbilde fra klyngen henrettelser og drap



(c) Eksempelbilde fra metaklyngen nærbilder av enkeltkrigere



(d) Eksempelbilde fra klyngen våpen



(e) Eksempelbilde fra klyngen sivile folkemengder



(f) Eksempelbilde fra klyngen fredelige scener

Figur 4.5





(a) Eksempelbilde fra klyngen sammenraste bygninger



(b) Eksempelbilde fra klyngen sivile ofre



(c) Eksempelbilde fra klyngen mat



(d) Eksempelbilde fra klyngen offentlige arbeider



(e) Eksempelbilde fra klyngen håndverk og industriproduksjon



(f) Eksempelbilde fra klyngen brennende sigarettesker



(g) Eksempel på et såkalt "promo-bilde" som er utelatt fra analysen.

Figur 4.6

#### 4.2.5 Test av tematisk konsistens i klyngene

Som nevnt i 4.2.1 satte vi en avstand fra hver bildevektor til klyngesentrum på 0,6 som en grenseverdi for hvilke bilder som skulle tas med i den tematiske analysen. Dette var basert på en kvalitativ vurdering av at minst 90 % av bildene under denne grenseverdien fremstod som tematisk riktig plassert ved gjennomsyn, mens det var noe mindre enn dette når verdiene overskred 0,6. Dette varierte også noe fra klynge til klynge, men 0,6 anså vi som et konservativt anslag på 90 % riktig.

For å teste hvor godt den programmatisk inndelingen ville samsvare med en analytikers, gjorde vi en blindtest med 200 bilder. Disse bildene hadde alle en skår under 0,6 og var tilfeldig utvalgt<sup>30</sup>. Analytikeren plasserte hvert av bildene i den klyngen han mente det burde høre hjemme i. Deretter sammenlignet vi denne manuelle inndelingen med den programmatisk.

Resultatet for hver klynge kan sees i forvekslingsmatrisen i figur 4.7.

Metaklynge ifølge k-snitt

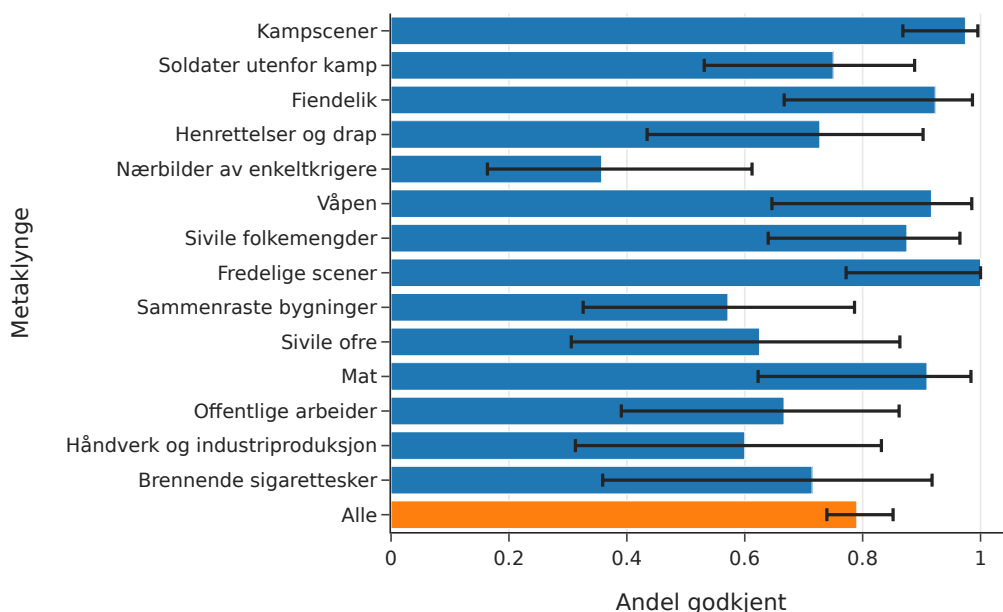
		Nærbilder av enkeltkrigere	Henrettelser og drap	Fiendelik	Soldater utenfor kamp	Kampscener	Sivile folkemengder	Våpen	Fredelige scener	Sammenraste bygninger	Sivile ofre	Offentlige arbeider	Håndverk og industriproduksjon	Brennende sigarettesker		
Manuell klassifisering	Kampscener	38	4	0	0	0	0	0	0	0	4	0	0	1	0	0
	Soldater utenfor kamp	0	15	0	0	3	0	0	0	0	0	0	0	0	0	0
	Fiendelik	0	0	12	3	0	0	0	0	0	0	3	0	0	0	0
	Henrettelser og drap	0	0	1	8	1	0	0	0	0	0	0	0	0	0	0
	Nærbilder av enkeltkrigere	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
	Våpen	0	1	0	0	0	11	0	0	0	0	0	0	0	0	0
	Sivile folkemengder	0	0	0	0	2	0	14	0	0	0	0	0	0	0	1
	Fredelige scener	0	0	0	0	0	0	0	13	1	0	0	0	0	0	0
	Sammenraste bygninger	0	0	0	0	0	0	0	0	8	0	0	1	0	0	0
	Sivile ofre	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0
	Mat	0	0	0	0	2	0	1	0	0	0	10	1	3	0	0
	Offentlige arbeider	0	0	0	0	0	0	0	0	0	1	0	0	8	1	0
	Håndverk og industriproduksjon	1	0	0	0	1	1	1	0	0	0	0	1	1	6	1
	Brennende sigarettesker	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5

Figur 4.7 Forvekslingsmatrise for test av tematisk konsistens i klyngene. Tallet i rad *R* og kolonne *K* er antallet bilder som ble plassert i klynge *K* av klyngealgoritmen og plassert i klynge *R* av analytikeren.

I alt plasserte analytikeren 158 av 200 bilder i samme klynge som *k*-snitt. Dette utgjør 79 %. Vektet

<sup>30</sup>For å få en jevn fordeling av de ulike klyngene, trakk vi et forhåndsbestemt antall bilder fra hver klynge. Dette antallet var proporsjonalt med kvadratrotten av antall bilder i klyngen. Dette betyr at små klynger er overrepresentert. Dette var for å få et noe bedre tallgrunnlag fra disse klyngene, som ellers bare ville blitt representert med noen ytterst få eksempler. Innenfor hver klynge ble bildene trukket med uniform sannsynlighet.

for størrelsen på klyngene stiger denne andelen til 84 %.<sup>31</sup> Figur 4.8 viser hvordan dette fordeler seg per klynge.



Figur 4.8 Andel av stikkprøvene for hver klynge der vi plasserer bildet i samme klynge som klyngealgoritmen. Figuren viser 95 % Wilson skår-konfidensintervaller. For noen klynges er datagrunnlaget svært lite, og dette gjenspeiles i store konfidensintervaller.

Det endelige resultatet (84 % vektet for størrelsen på klyngene) er litt lavere enn det vi hadde forventet (rundt 90 %). Her er det imidlertid viktig å merke seg at analytikerens inndeling ikke nødvendigvis utgjør “fasiten”. De fleste bildene som ble plassert i en annen klynge, har analytikerens karakterisert som tvilstilfeller, der det er uklart hvilket tema bildet skal plasseres under, og der analytikerens ikke regner  $k$ -snitt-inndelingen som *feil* men som en annen tolkningsmulighet. Hvis for eksempel et bilde viser en *kampscene* blant *sammenraste bygninger*, er det ikke åpenbart hvor denne skal plasseres (se figur 4.9a). Tilsvarende var det vanskelig å avgjøre – både for analytikerens og  $k$ -snitt – om forskjellige typer matproduksjon hørte best hjemme under *mat* eller *håndverk og industriproduksjon* (se figur 4.9b). Begge disse eksemplene kan også forstås ved å studere figur 3.4 og se at nettopp disse klyngene har bildevektorer som ligger svært nær hverandre. Siden disse tvilstilfellene uansett forekommer mellom klynges som tematisk ligner på hverandre, anser vi det ikke som noe stort problem for det analytiske potensialet.

<sup>31</sup> Dette er fordi denne andelen er størst for de store klyngene, som er underrepresentert blant de 200 stikkprøvene. Særlig gjelder dette *kampscener*, som utgjør en svært stor andel av korpuset, og der andelen er betydelig høyere enn de fleste øvrige klyngene.





(a) Bildet viser en kampscene blant sammenraste bygninger



(b) Honningproduksjon

Figur 4.9

Når det er sagt, er det enkelte bilder som åpenbart er plassert feil i forhold til det som er nyttig for analysen, og uten at de kan ansees som tvilstilfeller. Dette gjelder imidlertid bare 6,5 % av bildene i blindtesten – med andre ord lavere enn de 10 % som vi tok høyde for når vi satte grenseverdien til 0,6. Her er det også verdt å merke seg at de feilplasserte bildene kan gi en nyttig indikasjon på hvilke egenskaper ved bildene  $k$ -snitt faktisk har brukt til å lage tematiske klynger. Et eksempel på dette kan man se i figur 4.10. Dette bildet er plassert i klyngen *brennende sigarettesker*, selv om bildet faktisk viser en mengde folk som gjør seg klare til å steine en dødsdømt. Bildet skulle med andre ord helst vært plassert i klyngen *sivile folkemengder*. Her kan man imidlertid observere at *steinene* som folk er i ferd med å plukke opp, er nærmest firkantete – som sigarettesker – og ligger i en haug på bakken, slik sigaretteskene pleier å gjøre på bildene der de brennes. Dermed “lures”  $k$ -snitt til å sortere disse bildene sammen tematisk. Med andre ord viser det at de egenskapene ved et bilde som  $k$ -snitt bruker i den tematiske inndelingen, ikke alltid er de samme som vi ville brukt, selv om resultatet man ender opp med, jevnt over er nyttig for en analytiker. Vi kan også se dette i figur A.2g, der det har blitt generert et bilde med firkantede esker og en haug – uten sigaretter eller ild. Dette må ses i sammenheng med at CLIP-modellen med sine bildevektorer – som  $k$ -snitt har forholdt seg til – ikke i enhver sammenheng legger vekt på det samme som vi er interessert i, som nevnt i 3.1.



Figur 4.10 Haugen med steiner på bildet gjør at  $k$ -snitt grupperer det sammen med sigarettesker (se figur 4.6f).



---

---

## 5 Analyse

I de foregående kapitlene har vi vist hvordan vi samlet data om tre distinkte variabler for hvert bilde. Den første er det tematiske innholdet, operasjonalisert i 14 tematiske klynger som igjen er tildelt 2 binære variabler. De to andre variablene er publiseringstidspunktet for hvert bilde og provinstillhørigheten. Dette gir oss grunnlaget for å besvare forskningsspørsmålet vi har stilt i denne tilfellestudien: *Hvordan har IS' bildepropaganda utviklet seg fra 2014–2022, og hva forteller denne utviklingen oss om IS som gruppe?*

For å besvare dette spørsmålet undersøker vi de tre variablene og sammenligner dem to og to. Først undersøker vi hvordan hver provins' bildeproduksjon – altså det totale antallet bilder en provins publiserer – har utviklet seg fra år til år (5.1). Deretter ser vi på hvordan det tematisk innholdet i bildene har utviklet seg i perioden 2014–2022 (5.2). Til slutt tar vi for oss likheter og forskjeller i det tematiske innholdet i bildene etter hvilken provins de kommer fra (5.3).

### 5.1 Bildeproduksjon per år og provins

I denne delen begynner vi med å undersøke det totale antallet propagandabilder IS *som helhet* har produsert hvert år (figur 5.1) og det totale antallet *hver av IS' provinser* har publisert (figur 5.2). Deretter kombinerer vi data om bildenes publiseringstidspunkt og provinstillhørighet for å se hvordan bildeproduksjonen har utviklet seg i hver provins gjennom årenes løp. Dette viser oss hvilke provinser som har vært mest produktive (og uproduktive) i hvilke perioder, noe som igjen kan gi en viss indikasjon på hvor aktive provinsene har vært og hvor viktige de er og har vært for IS.<sup>32</sup>

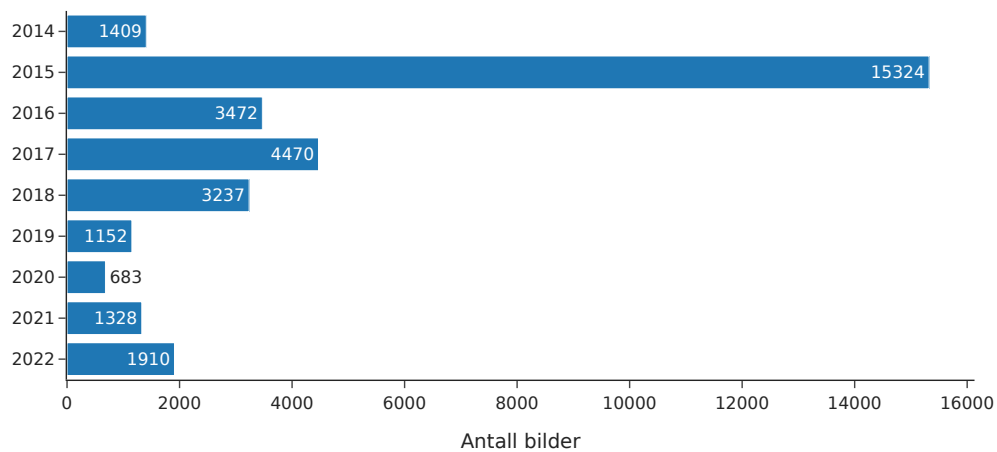
Figur 5.1 viser det totale antallet propagandabilder IS har produsert hvert år. De første bildene i vårt korpus ble publisert i 2014, som var året gruppen tok over store deler av Syria og Irak og erklærte gjenopprettelsen av kalifatet. Over halvparten av bildene i korpuset er fra året 2015, da gruppen nådde sitt høydepunkt militært og kontrollerte store deler av Syria og Irak. Nedgangen i 2016 kan delvis forklares med at det militære presset på gruppen øker, territorier går tapt og mediepersonell blir drept, men det er sannsynlig at en del bilder fra dette tidsrommet mangler i vårt korpus.<sup>33</sup> Etter at gruppen mistet sine siste territorier i Syria og Irak i 2018 går produksjonsnivået ytterligere ned, til et foreløpig lavmål i 2020. I årene etterpå har imidlertid bildeproduksjonen begynt å ta seg opp igjen. Slik ser vi at vi kun ved å se på antall bilder per år får et – riktignok svært grovkornet – inntrykk av hvordan IS har utviklet seg i perioden 2014–2022.

Vi ser også på hvordan antallet bilder fordeler seg på provinser (figur 5.2). Her kommer det ikke som noen overraskelse at Syria og Irak har produsert nesten to tredeler av det totale antallet, siden dette var territoriene IS kontrollerte da gruppen var på høyden. Provinsene Vest-Afrika og Libya – som vi kommer tilbake til senere – ligger på henholdsvis tredje og fjerdeplass som bildeprodusenter, men fortsatt betydelig lavere enn Irak på nummer to.

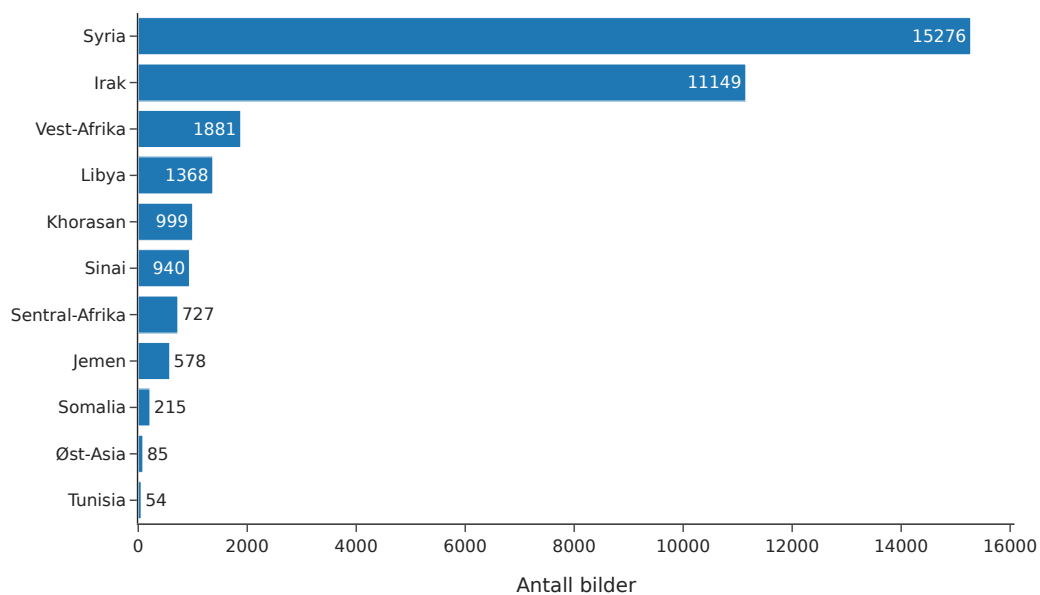
---

<sup>32</sup>Det gir en *indikasjon*, men ikke noe bevis på hvor aktiv en provins er. Som diskutert i 5.4 kan det tenkes at IS i noen tilfeller i kan eller ønsker å vise frem hvor mye aktivitet de har i enkelte provinser.

<sup>33</sup>Charlie Winter registrerer ikke den samme brå nedgangen i 2016 (Winter 2022, s. 27–29). Se forøvrig 2.4

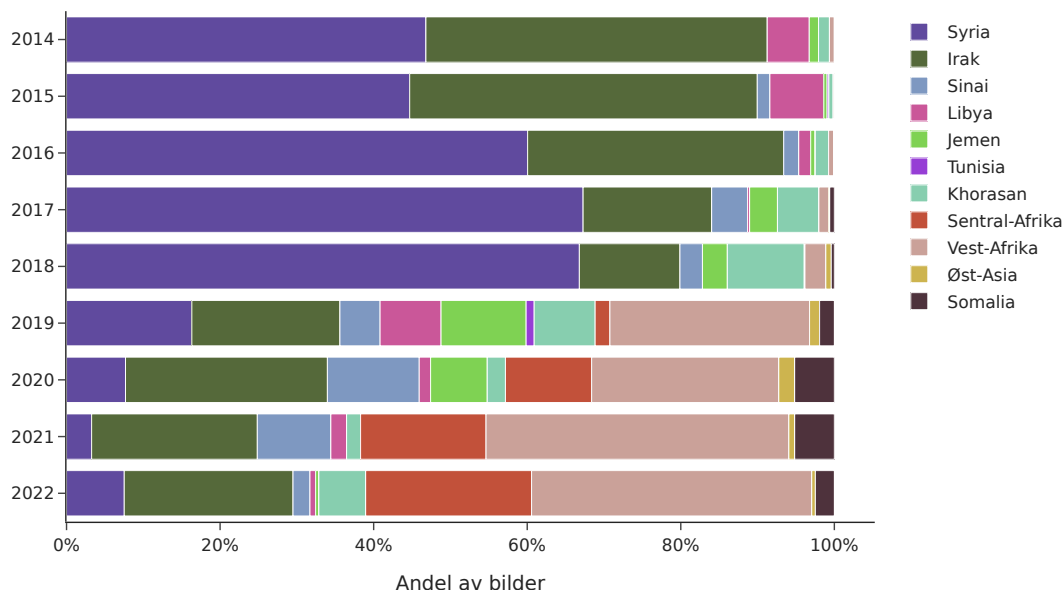


*Figur 5.1 Antall bilder fra hvert år*



*Figur 5.2 Antall bilder for hvert område. Bare områder med minst 50 bilder er vist.*

Hvis vi derimot ser hvordan bildeproduksjonen har utviklet seg per år per provins, blir utviklingstrekkene enda tydeligere. Figur 5.3 viser det relative antallet bilder hver provins har produsert i perioden 2014–2022. Som forventet produserte Irak og Syria det store flertallet av bilder i perioden 2014–2018, som forteller oss at nedgangen i det totale antall bilder fra 2015 og utover (figur 5.1) kan forklares nesten utelukkende med nedgangen i Syria og Irak.



Figur 5.3 Antall bilder fra hvert område fra hvert år. Bare områder med minst 50 bilder er vist.

Men det mest iøynefallende utviklingstrekket er at hovedtyngden i bildeproduksjonen har forflyttet seg fra Syria og Irak til Afrika.<sup>34</sup> Vest-Afrika overtar som den mest produktive enkeltprovinsen allerede i 2019 og er større enn Irak og Syria *til sammen* i 2021. Sentral-Afrika blir først anerkjent som IS-provins i 2020, men er allerede i 2022 den største produsenten av bildepropaganda etter Vest-Afrika og Irak. Til sammen står de afrikanske provinsene (medregnet Somalia) for 61 % av produksjonen i 2022. Dette gjenspeiler en tendens som har vært tydelig for observatører de siste årene, nemlig at IS er i raskt fremmarsj på det afrikanske kontinentet, og at gruppens militære hovedtyngde nå befinner seg i Afrika, ikke i Midtøsten.

Utover dette hovedtrekket i utviklingen er det også andre interessante observasjoner man kan gjøre i disse dataene. Som vi kan se i figur 5.3, var det i 2015 også et betydelig antall bilder fra Libya. IS utnyttet en pågående borgerkrig i landet til å opprette en svært slagkraftig fraksjon med hovedsete i byen Sirte. Men Libya forsvinner fra datasettet i 2017 og 2018 etter Sirte ble inntatt av libyske regjeringsstyrker i 2016, og dukker ikke opp igjen før de slipper enkelte angrepsbilder og *baya*-bilder i 2019.<sup>35</sup> Man kan også merke seg at IS-Khorasan, som av flere observatører har vært

<sup>34</sup>Dette skyldes på ingen måte bare at produksjonen har gått ned i Syria og Irak. I 2018 publiserte Vest- og Sentral-Afrika henholdsvis 88 og ingen bilder. I 2022 var tilsvarende tall 694 og 411.

<sup>35</sup>I perioden april-juni dette året gjennomførte IS Libya elleve angrep, før gruppen igjen ble slått ned av den libyske

---

---

regnet som den av IS-provinsene som utgjør den største terrortrusselen mot Vesten,<sup>36</sup> på ingen måte utmerker seg som en stor produsent av bildepropaganda.

Oppsummert kan man få øye på hovedtrekkene i IS' historie fra 2014–2022 ved kun å se på volumet av bildepropaganda som er blitt publisert over tid i forskjellige provinser. IS var på sitt største i 2015 da gruppen kontrollerte store områder i Syria og Irak. Nå har gruppen blitt mer eller mindre nedkjempet i de gamle kjerneområdene, men er i ferd med å bygge seg opp igjen i Afrika.

## 5.2 Utvikling i tematisk innhold over tid

Etter kun å ha sett på hvordan antallet bilder fordeler seg over tid og provinser, skal vi nå se nærmere på det tematiske innholdet i bildene. I denne delen undersøker vi hvordan innholdet i bildene – operasjonalisert som tematiske klynger – har forandret seg i perioden 2014–2022. Man kan lære mye om gruppen ved å se hvilke temaer de legger vekt på – og ikke legger vekt på – i forskjellige perioder – dette gjenspeiler både IS' operasjonelle utvikling og hvordan gruppen ønsker å fremstille seg utad.

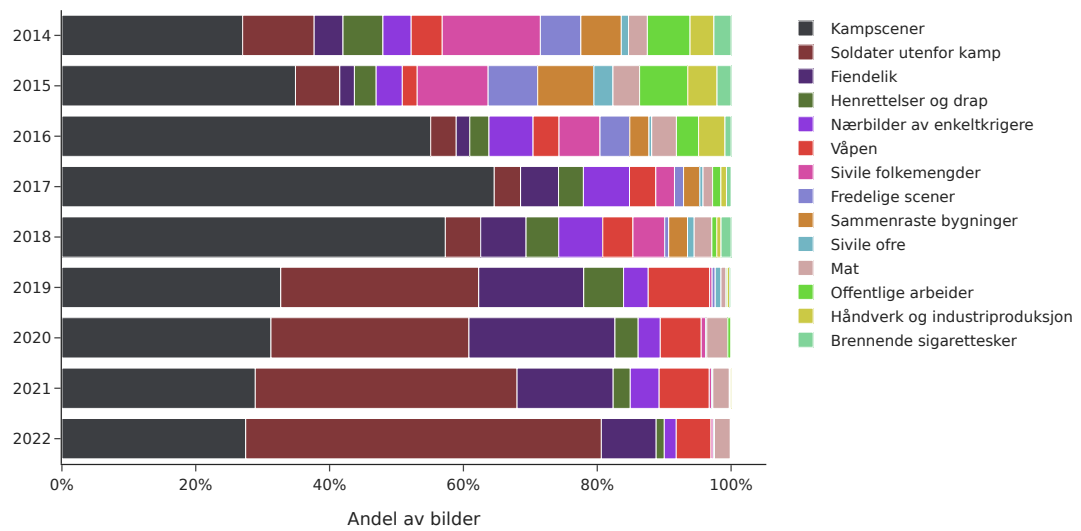
Figur 5.4 viser prosentvis fordeling av klyngetemaer per år. Det mest iøynefallende utviklingstrekket er at det tematiske innholdet blir mer ensartet ettersom tiden går – særlig etter 2018. Her kan vi se at flere av temaene som er viktige i perioden 2014–2018 forsvinner helt etterpå. For eksempel utgjør *sivile folkemengder* nesten 15 % av bildene i 2014, men blir nærmest utradert etter 2018. I samme tidsrom ser vi at temaene *produksjon*, *offentlige arbeidere*, *fredelige scener* og *brennende sigarettesker*, som til sammen utgjorde rundt 20 % på toppen i 2015, går mot null. Tilsvarende kan vi observere at temaene som omhandler brutaliteten til IS' fiender, *sammenraste bygninger* og *sivile ofre*, med henholdsvis 8,4 % og 2,8 % av materialet i 2015, også blir borte. Med andre ord er det de temaene som omhandler de *sivile* sidene av IS som går nedover etter 2015 og nærmest forsvinner etter 2018, noe som selvsagt har nær sammenheng med at IS mistet territoriene i Irak og Syria gjennom denne tiden. IS har operert primært som opprørs- og terrorgruppe – ikke som styresmakt – i tiden etterpå, og dette gjenspeiles i temaene som går igjen i bildepropagandaen.

Denne utviklingen mot stadig mer militært innhold blir enda tydeligere når vi bare fokuserer på den binære variabelen *militær/ikke-militær* (figur 5.5). Mens vi ser en tilnærmet 50–50 fordeling av militære og ikke-militære temaer i 2014 og 2015, har sivile temaer sunket til 4,2 % i 2022. Dette er en betydelig nedgang, men den reelle andelen bilder som viser sivile temaer er sannsynligvis enda mindre enn det som vises her. Dette er fordi vi valgte å kode klyngen *mat* som sivil – dette er korrekt for de fleste av bildene, men en del av disse viser også mat som blir spist eller tilberedt i felten, og burde ideelt sett blitt kodet som militær. Dette gjelder særlig *mat*-bildene som er publisert etter kalifatets fall – og dette er grunnen til at *mat*-klyngen også er den eneste sivile klyngen som holder seg relativt stabil også etter 2018.

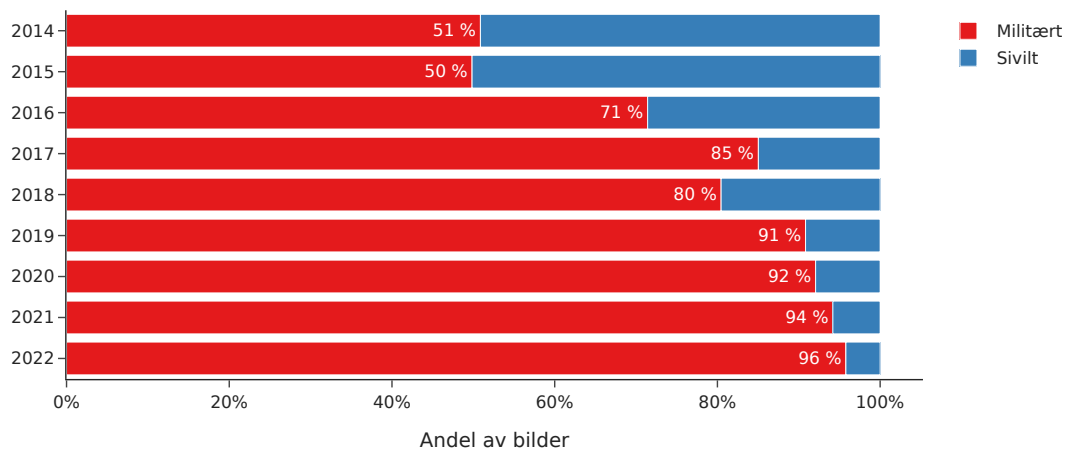
Ellers er det interessant å merke seg at temaet med størst prosentvis økning etter 2018 er *soldater utenfor kamp*. Dette har selvsagt sammenheng med at militære temaer i det hele tatt blir viktigere.

hæren (LNA). I årene som følger, kommer det også enkelte bilder fra Libya – primært *baya*-bilder – men den samlede produksjonen er betydelig lavere i perioden fra 2019 og utover enn den var da IS holdt territorium i Libya i perioden 2014–2016. Se Zelin (2019).

<sup>36</sup>Se for eksempel Sayed og Hamming (2023, s. 21)



Figur 5.4 Fordelingen av metaklynger for hvert år



Figur 5.5 Andel av bildene klassifisert under kategorien "militært" for hvert år.

---

---

Men en annen viktig grunn til denne økningen er de såkalte *baya*-bildene. *Baya* er en arabisk term for en troskapsed som hver av IS' provinser avlegger når en ny kalif (leder) blir utpekt, og disse troskapsedene blir grundig dokumentert i en bildeserie fra hver provins (se for eksempel figur 4.4f). Da amerikanske styrker drepte IS' første kalif Abu Bakr al-Baghdadi i oktober 2019, utløste det et skred av *baya*-bilder med forskjellige IS-provinser som sverget troskap til hans etterfølger Abu Ibrahim al-Hashimi al-Qurashi. Disse bildene ble publisert i slutten av 2019 og begynnelsen av 2020. Tilsvarende *baya*-kampanjer kom da sistnevnte ble drept i februar 2022, og da hans etterfølger igjen, Abu al-Hasan al-Hashimi al-Qurashi, ble drept i oktober 2022. Disse *baya*-bildene utgjør en stor del av klyngen *soldater utenfor kamp* i disse tidsperiodene, men det faktum at det er et like stort antall bilder i denne klyngen også i 2021 – et år uten en *baya*-kampanje – tyder på at dette temaet har blitt mer fremtredende i IS-propaganda i det hele tatt.

Samlet sett ser vi altså en gradvis *ensretting* av det tematiske innholdet, der alle klyngene som viser *sivile* temaer, nærmest forsvinner etter 2018 ettersom IS' "statsapparat" kollapser. Selv om militære temaer tar over, er det ikke *kampscener* som har den største prosentvise økningen – denne går faktisk nedover fra toppåret 2017. I stedet er det bilder av andre militære temaer – særlig av *soldater utenfor kamp* – som står for den største relative økningen.

### 5.3 Tematisk innhold fordelt på provins

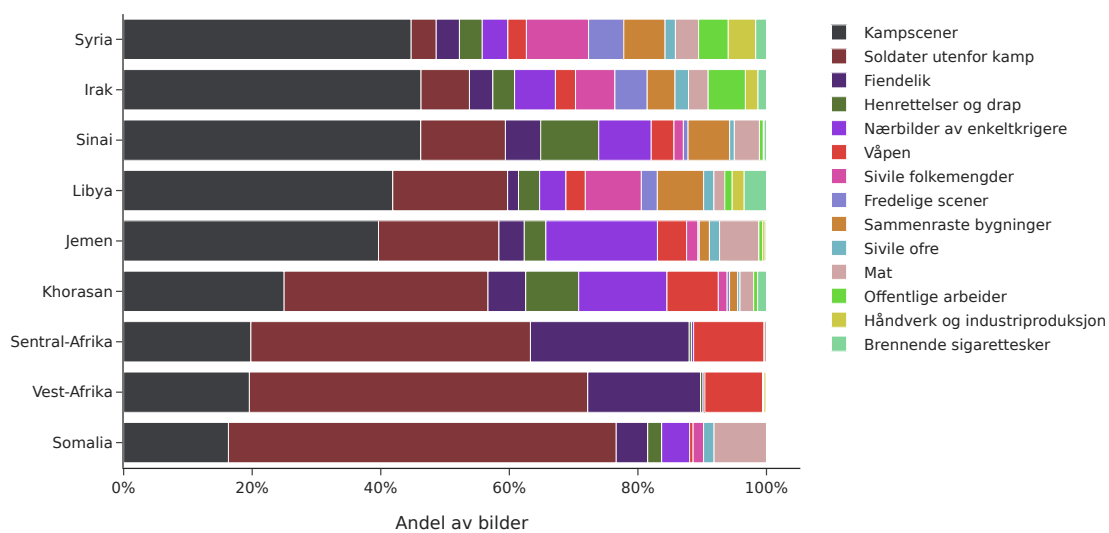
Til slutt sammenstiller vi informasjon om bildenes tematiske innhold med deres provinstillhørighet. Dette kan vise oss forskjeller i hvordan IS' forskjellige underavdelinger opererer og ønsker å fremstille seg selv. Vi begynner med den relative fordelingen av tematiske klynger per område (figur 5.6 og 5.7).

Som man kan forvente, er de "sivile" klyngetemaene *sivile folkemengder, produksjon, offentlige arbeider, fredelige scener, brennende sigarettesker, sammenraste bygninger* og *sivile ofre* best representert i Syria og Irak, der IS hadde kontroll over store territorier i perioden 2014–2018 og ønsket å fremstille seg som en "stat". Sivile klyngetemaer er også fremtredende i bilder fra Libya, til tross for at IS' kontroll over byen Sirte og andre steder i Libya var relativt kortvarig. Samtidig var IS-provinsen i Libya IS' første vellykkete ekspansjon utover Syria og Irak, og bildematerialet tyder på at provinsen skulle være et utstillingsvindu for hvordan statsstyret som var etablert i IS' kjerneområder, kunne innføres andre steder i verden. I de øvrige IS-provinsene i Midtøsten og Asia er sivile klyngetemaer sjeldne.

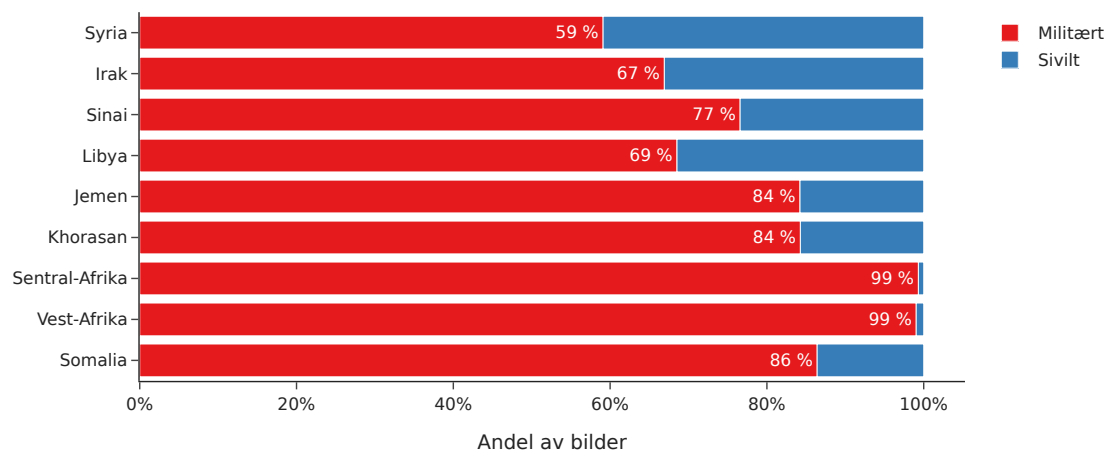
Provinsene Vest-Afrika og Sentral-Afrika skiller seg tematisk ut fra de andre. I begge disse provinsene er de "sivile" klyngetemaene fullstendig fraværende. Disse provinsene har ikke hatt mulighet til å kontrollere territorium over tid og etablere en sofistikert sivil administrasjon lik den IS hadde i Syria og Irak. Dette er likevel ikke *hele* forklaringen. Selv om disse afrikanske provinsene ikke har kontrollert et klart definert territorium som i Irak og Syria, har de tatt over områder over kortere tid og også utøvd "indirekte" kontroll over større territorier. IS Vest-Afrika – den utvilsomt sterkeste og best organiserte av de to – har også innført en rudimentær siviladministrasjon, som skattelegging og domstoler, i områder de kontrollerer.<sup>37</sup> Det faktum at denne sivile siden ved operasjonene deres ikke blir viet oppmerksomhet i bildepropagandaen, kan tyde på at de har gjort et bevisst valg om å fremstille seg selv utelukkende som opprørsgrupper.

---

<sup>37</sup>Se for eksempel *After Shekau: Confronting Jihadists in Nigeria's North East* (2022).



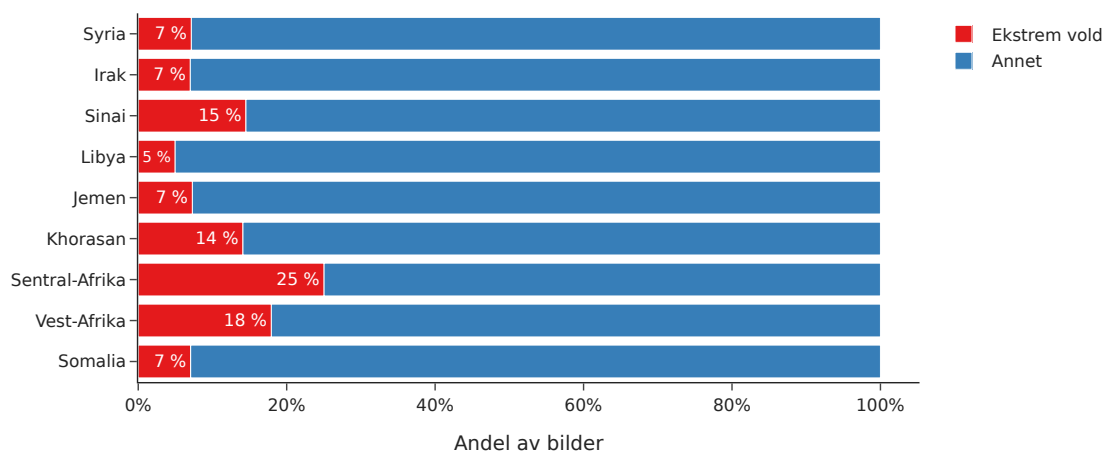
Figur 5.6 Fordelingen av metaklynger for hvert område. Bare områder med minst 100 bilder er vist.



Figur 5.7 Fordeling av militære og sivile temaer per område. Bare områder med minst 100 bilder er vist.

Videre finnes det nesten ingen bilder fra klyngen *nærbilder av enkeltkrigere* – som for det meste viser bilder av IS’ såkalte “martyrer” – i materialet fra Vest- og Sentral-Afrika. Dette klyngetemaet er godt representert i alle andre provinser – fra 3,9 % i Syria til så mye som 17 % i Jemen. Martyrbiografier og -bilder har vært et yndet tema i jihadistpropaganda fra 1980-tallet til i dag, og det er vanskelig å gi noe klart svar på hvorfor disse afrikanske provinsene bryter med denne trenden.

Sist men ikke minst skiller Vest- og Sentral-Afrika seg ut ved at klyngetemaet *fiendelik* er betydelig større enn i andre provinser, med 17 % for Vest-Afrika og hele 25 % av materialet for Sentral-Afrika. Denne typen bilder er ofte svært groteske. Når det er sagt, er ikke groteske bilder utpreget vanlige for de afrikanske provinsene. Selv om de ikke-afrikanske provinsene har færre bilder i klyngetemaet *fiendelik*, har disse som regel flere bilder i klyngetemaet *henrettelser og drap*. Når vi koder disse klyngetemaene sammen i variabelen *ekstrem vold* (figur 5.8) ser vi en viss overrepresentasjon i Vest- og Sentral-Afrika, men ikke betydelig.



Figur 5.8 Andel ekstrem vold per område. Bare områder med minst 100 bilder er vist.

## 5.4 Oppsummering av analytiske funn

Som *propaganda* har bildene vi har analysert i denne rapporten, en spesiell funksjon – nemlig å videreformidle IS’ egne narrativ. Man kan derfor regne med at IS vil overdrive egne seire og underkommunisere nederlag. Men siden IS har produsert så *mange* av disse bildene over tid, og fordi bildene viser ting som IS’ fraksjoner faktisk holder på med, utgjør de uansett en viktig kilde til å forstå IS’ utvikling. Situasjonen IS til enhver tid befinner seg i setter nødvendigvis noen føringer for hva som fremstilles i bildepropagandaen. Hvis IS for eksempel ikke lenger har noe sivilt administrasjonsapparat, kan det heller ikke vises frem. Tilsvarende gir produksjonsvolumet oss også en indikasjon på hvordan det går med gruppen. Hvis IS for eksempel ikke *har* noen til å dokumentere aktiviteten i en provins, eller det ikke er noe aktivitet å dokumentere, må produksjonsvolumet nødvendigvis gå ned.



---

---

Samtidig er det viktig å være klar over at forholdet mellom det IS *gjør*, og det de viser frem i propagandaen, ikke nødvendigvis er én-til-én. Forskere har pekt på at IS – særlig i senere år – har underrapportert hvor mange angrep de har gjennomført i enkelte provinser. Dette har delvis sammenheng med at lokale grupper har manglet utstyr både til å dokumentere aktivitet og til å kommunisere med IS' sentrale medieorgan. Men mye tyder også på at IS bevisst har latt være å kringkaste sin egen (angreps-)aktivitet enkelte steder for å unngå å tiltrekke seg fiendens oppmerksomhet og gi inntrykk av at de er svakere enn de egentlig er.<sup>38</sup> Med andre ord gir ikke bildepropagandaen noen “fasit” på hvor IS står som organisasjon, i hvert fall ikke på detaljplan.

Når det er sagt, sammenfaller de store linjene i bildematerialet med IS' egen historie. IS' bildepropaganda – og IS selv – er noe helt annet i 2022 enn da gruppen var på sitt sterkeste i 2015. Fra toppen i 2015 kollapset bildeproduksjonen i takt med IS' territorielle kontroll til et foreløpig lavmål i 2020 før det tok seg noe opp igjen. Fra å vise omtrent like mye sivilt som militært innhold for å speile at IS var både en *stat* og en *militærorganisasjon*, viser bilderapportene nå kun militære temaer. Og mens hovedtyngden i bildematerialet pleide å være fra Midtøsten, kommer nå flertallet av bildene fra Afrika.

Ved å se nærmere på bildene fra nettopp de afrikanske provinsene der IS er på vei oppover, kan man få et innblikk i hvor IS er på vei i fremtiden. IS har ikke bare forflyttet seg, men kanskje også blitt noe annet i Afrika enn de var i Midtøsten. Selv-iscenesettelsen i propagandaen er annerledes – som vi så, er det langt mindre fokus på individuelle jihadister og martyrer i de afrikanske IS-provinsene enn i andre, samtidig som propagandabildene jevnt over er enda mer brutale. Videre har bildepropagandaen fra de afrikanske IS-provinsene i likhet med andre provinser blitt stadig mer dominert av militært innhold, som selvsagt gjenspeiler det faktum at de er mest aktive som militærorganisasjoner.

Men IS-provinsene i Vest-Afrika og Sahel *er* faktisk involvert i sivil administrasjon i områder som er helt eller delvis under deres kontroll, noe som kunne blitt lagt vekt på i propagandabildene. Dette kan tyde på at gruppen har mindre pretensjoner om “statlighet” og at dette er blitt en mindre viktig del av selvbildet og propagandaen til IS, ettersom “staten” ligger flere år bak i tid. IS' *identitet* synes med andre ord delvis å ha skiftet til å være en effektiv – og svært brutal – opprørsgruppe i Afrika. Paradoksalt nok kan dette være en styrke for IS – nemlig at legitimiteten til gruppen *ikke* er direkte koblet til hvorvidt de styrer en faktisk stat. Slik kan IS leve godt med status quo inntil det igjen blir aktuelt å erklære en ny “stat” som er like slagkraftig som det såkalte “kalifatet” var fra 2014–2018.

---

<sup>38</sup>For en mer utfyllende diskusjon om dette poenget, se Zelin og Margolin (2023).

---

---

## 6 Konklusjon

Denne studien har undersøkt hvordan man kan bruke maskinlæringsmetoder til å understøtte analysen av store mengder propagandabilder. Vi har benyttet maskinlæringsmetoder på et spesifikt brukstilfelle, nemlig IS' propagandabilder i perioden 2014–2022. Hensikten har vært todelt: Metodisk har vi ønsket å finne ut hvor godt disse metodene kan hjelpe oss å analysere et slikt bildemateriale. Den analytiske ambisjonen har vært å undersøke hvordan IS' propagandabilder har utviklet seg over tid og mellom forskjellige provinser.

Vi har funnet at CLIP, som er en maskinlæringsmodell for bildeforståelse trent ved hjelp av bildetekster, kombinert med klyngealgoritmen  $k$ -snitt gir oss et svært nyttig verktøy i analyse av store bildematerialer – i dette tilfellet over 30 000 bilder. Vi brukte CLIPs bildeenkoder til å lage representasjonsvektorer som fanger det semantiske innholdet i hvert bilde, og deretter brukte vi  $k$ -snitt til å gruppere disse i klynger av lignende vektorer. Dette gav oss en oversiktlig tematisk inndeling av bildene med minimalt med arbeid. Inndelingen i tematisk relevante klynger var ikke “perfekt” fra en analytikers perspektiv, siden en viss prosentandel fremstod som feilplassert. Like fullt bidrar dette til å gjøre analysearbeidet betydelig raskere.

Selv om vi har sett på propagandabilder i denne studien, er ikke metoden begrenset til dette brukstilfellet. Etersom modellen CLIP ikke er trent spesifikt på denne typen bilder, er det grunn til å tro at metodene kan fungere tilsvarende godt på bilder med et annet innhold enn jihadistpropaganda. Metoden er relativt enkel og skalerbar, og man kan lett bruke den til å dele inn andre samlinger med bilder i tematiske klynger, også samlinger som er betydelig større enn 30 000 bilder. Det er nettopp i møte med storesamlinger metoden først og fremst er nyttig, og særlig dersom man vet lite om innholdet.

Studien har vist at det tematiske innholdet i IS' propagandabilder har endret seg betydelig i løpet av den perioden vi har undersøkt, i takt med at gruppen selv har gått igjennom opp- og nedgangstider. Det mest sentrale utviklingstrekket er at IS har gått fra å fremstille seg både som en sivil administrasjon og som et militært apparat til nesten utelukkende å vise militær aktivitet. Samtidig har hovedvekten i materialet beveget seg fra Syria og Irak til IS' provinser i Afrika. Sammenlignet med provinsene i Midtøsten vier de afrikanske IS-provinsene oppsiktsvekkende lite oppmerksomhet til egne enkeltkrigere og martyrer, men de viser enda flere, groteske bilder av døde fiender, samtidig som de legger mindre vekt på sivilt innhold enn man kan forvente utifra deres posisjon på bakken. Disse forskjellene i det tematiske innholdet mellom provinsene i Afrika og Midtøsten gir en indikasjon på at IS' selvforståelse er i ferd med å endre seg ettersom det geografiske tyngdepunktet forflyttes.

Med en stadig mer uoversiktlig informasjonssfære på nett ser vi det som hensiktsmessig å videreutvikle maskinlæringsmetoder for å kartlegge propaganda fra statlige og ikke-statlige aktører. Her er det ikke bare programmatisk *bilde*analyse som kan være relevant, men også tilsvarende metoder for lett å få oversikt over store tekst-, lyd- og videomaterialer. For tematisk analyse av bilder fremstår CLIP – eller en tilsvarende modell – som et av de mest beste verktøyene som er åpent tilgjengelig for forskere og analytikere. Kombinasjonen av CLIP og klyngealgoritmer, som vi har brukt i denne rapporten, er spesielt godt egnet når man vil få rask oversikt over et omfattende ukjent bildemateriale.

---

---

## Referanser

- After Shekau: Confronting Jihadists in Nigeria's North East* (29. mar. 2022). Africa Briefing 180. International Crisis Group.
- Akil, Hatem N. (okt. 2016). «Cinematic Terrorism: Deleuze, ISIS and Delirium». I: *Journal for Cultural Research* 20.4, s. 366–379. ISSN: 1479-7585, 1740-1666. DOI: 10.1080/14797585.2016.1168973.
- Baele, Stephane J, Katharine A Boyd og Travis G Coan (7. okt. 2020). «Lethal Images: Analyzing Extremist Visual Propaganda from ISIS and Beyond». I: *Journal of Global Security Studies* 5.4, s. 634–657. ISSN: 2057-3170, 2057-3189. DOI: 10.1093/jogss/ogz058.
- Byman, Daniel (2016). «Understanding the Islamic state—a review essay». I: *International Security* 40.4, s. 127–165.
- Cherti, Mehdi, Romain Beaumont, Ross Wightman, Mitchell Wortsman, Gabriel Ilharco, Cade Gordon, Christoph Schuhmann, Ludwig Schmidt og Jenia Jitsev (2023). «Reproducible scaling laws for contrastive language-image learning». I: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, s. 2818–2829.
- Cigar, Norman (2014). *Tribal Militias: An Effective Tool to Counter Al-Qaida and Its Affiliates?* Tekn. rapp. Strategic Studies Institute (SSI), s. 116.
- Cook, Joana og Gina Vale (2018). *From Daesh to 'Diaspora': Tracing the Women and Minors of Islamic State*. Tekn. rapp. London: ICSR, King's College, s. 72.
- Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly et al. (2020). «An image is worth 16x16 words: Transformers for image recognition at scale». I: *arXiv preprint arXiv:2010.11929*.
- Ester, Martin, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu et al. (1996). «A density-based algorithm for discovering clusters in large spatial databases with noise». I: *kdd*. Bd. 96. 34, s. 226–231.
- Euben, Roxanne L. (des. 2017). «Spectacles of Sovereignty in Digital Time: ISIS Executions, Visual Rhetoric and Sovereign Power». I: *Perspectives on Politics* 15.4, s. 1007–1033. ISSN: 1537-5927, 1541-0986. DOI: 10.1017/S1537592717002134.
- Friis, Simone Molin (jun. 2018). «'Behead, Burn, Crucify, Crush': Theorizing the Islamic State's Public Displays of Violence». I: *European Journal of International Relations* 24.2, s. 243–267. ISSN: 1354-0661, 1460-3713. DOI: 10.1177/1354066117714416.
- Hegghammer, Thomas (16. jan. 2014). «Jihadi-Salafis Or Revolutionaries?» I: *Global Salafism*. Red. av Roel Meijer. Oxford University Press, s. 245–266. ISBN: 978-0-19-933343-1. DOI: 10.1093/acprof:oso/9780199333431.003.0011.
- Hegghammer, Thomas (mai 2022). «OCR with Tesseract, Amazon Textract, and Google Document AI: A Benchmarking Experiment». I: *Journal of Computational Social Science* 5.1, s. 861–882. ISSN: 2432-2717, 2432-2725. DOI: 10.1007/s42001-021-00149-1.

- 
- 
- Lloyd, Stuart (1982). «Least squares quantization in PCM». I: *IEEE transactions on information theory* 28.2, s. 129–137.
- Mello, Brian (1. jun. 2018). «The Islamic State: Violence and Ideology in a Post-colonial Revolutionary Regime». I: *International Political Sociology* 12.2, s. 139–155. ISSN: 1749-5679, 1749-5687. DOI: 10.1093/ips/oly003.
- Milton, Daniel James (2018). *Down, but not out: An updated examination of the Islamic State's visual propaganda*. Combating Terrorism Center at West Point, United States Military Academy.
- Prucha, Nico (2016). «IS and the Jihadist Information Highway – Projecting Influence and Religious Identity via Telegram». I: *Perspectives on Terrorism* 10.6.
- Radford, Alec, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger og Ilya Sutskever (jul. 2021). «Learning Transferable Visual Models From Natural Language Supervision». I: *Proceedings of the 38th International Conference on Machine Learning*. Red. av Marina Meila og Tong Zhang. Bd. 139. Proceedings of Machine Learning Research. PMLR, s. 8748–8763. URL: <https://proceedings.mlr.press/v139/radford21a.html>.
- Ramesh, Aditya, Prafulla Dhariwal, Alex Nichol, Casey Chu og Mark Chen (2022). «Hierarchical text-conditional image generation with clip latents». I: *arXiv preprint arXiv:2204.06125* 1.2, s. 3.
- Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser og Björn Ommer (2022). «High-resolution image synthesis with latent diffusion models». I: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, s. 10684–10695.
- Sayed, Abdul og Tore Refslund Hamming (jun. 2023). *The Growing Threat of the Islamic State in Afghanistan and South Asia*. 520. United States Institute of Peace.
- Sculley, David (2010). «Web-scale k-means clustering». I: *Proceedings of the 19th international conference on World wide web*, s. 1177–1178.
- Stern, Jessica og J. M. Berger (2016). *ISIS: The State of Terror*. New York: Ecco Press, an imprint of HarperCollins Publishers. 398 s. ISBN: 978-0-06-239555-9.
- Tinnes, Judith (mar. 2023). «Islamic State (IS, ISIS, ISIL, Daesh) Part 7». I: *Perspectives on Terrorism* 17.1.
- Tønnessen, Truls Hallberg (aug. 2015). «Heirs of Zarqawi or Saddam? The Relationship between al-Qaida in Iraq and the Islamic State». I: *Perspectives on Terrorism* 9.4, s. 48–60.
- Tønnessen, Truls Hallberg (feb. 2019). «The Islamic State after the Caliphate». I: *Perspectives on Terrorism* 13.1.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser og Illia Polosukhin (2017). «Attention is all you need». I: *Advances in neural information processing systems* 30.
- Winter, Charlie (2022). *The Terrorist Image: Decoding the Islamic State's Photo-Propaganda*. London: Hurst & Company. 259 s. ISBN: 978-1-78738-567-2.

---

Zelin, Aaron Y. (6. des. 2019). *The Islamic State in Libya Has Yet to Recover*. The Washington Institute. URL: <https://www.washingtoninstitute.org/policy-analysis/islamic-state-libya-has-yet-recover> (lest 30.10.2023).

Zelin, Aaron Y. og Devorah Margolin (20. sep. 2023). «The Islamic State's Shadow Governance in Eastern Syria Since the Fall of Baghuz». I: *Combating Terrorism Center at West Point* 16.9. URL: <https://ctc.westpoint.edu/the-islamic-states-shadow-governance-in-eastern-syria-since-the-fall-of-baghuz/> (lest 18.10.2023).

## A Genererte bilder



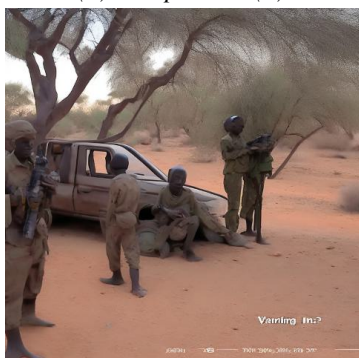
(a) Kampscener (1)



(b) Kampscener (2)



(c) Kampscener (3)



(d) Soldater utenfor kamp (1)



(e) Soldater utenfor kamp (2)



(f) Henrettelser og drap



(g) Nærbilder av enkeltkrigere (2)



(h) Våpen



(i) Sivile folkemengder

Figur A.1 Bilder automatisk generert fra sentrumsvektoren i et utvalg av klyngene, som beskrevet i 3.2.2.2.





Figur A.2 Bilder automatisk generert fra sentrumsvektoren i et utvalg av klyngene, som beskrevet i 3.2.2.2.

## Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan med særskilte fullmakter underlagt Forsvarsdepartementet.

## FFIs formål

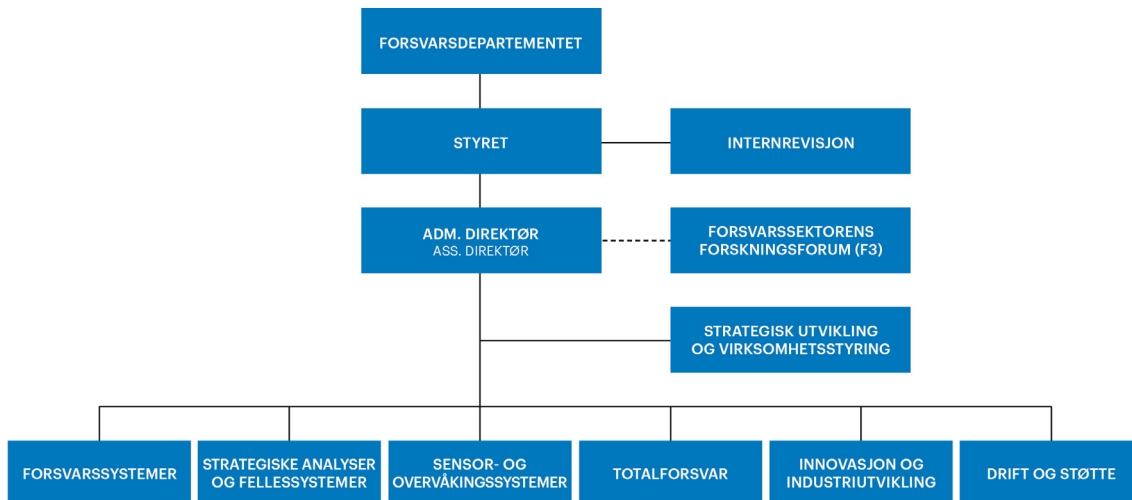
Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

## FFIs visjon

FFI gjør kunnskap og ideer til et effektivt forsvar.

## FFIs verdier

Skapende, drivende, vidsynt og ansvarlig.





Forsvarets forskningsinstitutt (FFI)  
Postboks 25  
2027 Kjeller

Besøksadresse:  
Kjeller: Instituttveien 20, Kjeller  
Horten: Nedre vei 16, Karljohansvern, Horten

Telefon: 91 50 30 03  
E-post: [post@ffi.no](mailto:post@ffi.no)  
[ffi.no](http://ffi.no)

Norwegian Defence Research Establishment (FFI)  
PO box 25  
NO-2027 Kjeller  
NORWAY

Visitor address:  
Kjeller: Instituttveien 20, Kjeller  
Horten: Nedre vei 16, Karljohansvern, Horten

Telephone: +47 91 50 30 03  
E-mail: [post@ffi.no](mailto:post@ffi.no)  
[ffi.no/en](http://ffi.no/en)