

FUND OG FORSKNING



BIND 64

2025



DET KGL. BIBLIOTEK

FUND OG FORSKNING

Bind 64

2025

FUND OG FORSKNING

I DET KONGELIGE BIBLIOTEKS

SAMLINGER

Bind 64

2025

With summaries

KØBENHAVN 2025

UDGIVET AF DET KGL. BIBLIOTEK

Det kronede monogram er tegnet af
Erik Ellegaard Frederiksen efter et bind fra Frederik 3.s bibliotek.

© Forfatterne og Det Kgl. Bibliotek

Redaktion:

Thomas Hvid Kromann (ansvarshavende), Caroline Nyvang og Stig Roar Svenningsen

Artikler i *Fund og Forskning* gennemgår
'double blinded peer-review' for at kunne antages.

Billedtekst til forsidebilledet:

Marie Topsøe (1847-1894) og Augusta Fenger (1844-1931) var lærere ved Københavns kommunale skolevæsen. De boede sammen fra ca. 1886 til Topsøes død i 1894. Sammen udgav de *Sundhedslære til Skolebrug* (1886), og her er de blevet fotograferet i A. Lønborgs atelier på Østergade 15 i København. Fotografiet indgår i artiklen "Vintage vision. Brug af object detection-modeller i historiske fotografier med fokus på kvindepar" af Henrik Kragh Sørensen, Mette Kia Krabbe Meyer og Laura Søvsø Thomasen.

Grafisk tilrettelæggelse: Narayana Press

ISSN 0069-9896

VINTAGE VISION

Brug af object detection-modeller i historiske
fotografier med fokus på kvindepar

AF

HENRIK KRAGH SØRENSEN, METTE KIA KRABBE MEYER
& LAURA SØVSØ THOMASEN

Kunstig intelligens (AI) giver helt nye muligheder for at arbejde med kulturarvssamlinger, og de seneste år er det blevet udbredt at anvende automatiseret analyse til at gruppere og 'genkende' skrift. Billedmateriale er imidlertid stadig en udfordring, for det er langt sværere at 'genkende' enkelte elementer på billeder end i tekst. Det gør sig særligt gældende, når det drejer sig om historiske billeder, hvor alt ser anderledes ud end i dag. At sige noget om *betydningen* af bestemte genstande, om relationer eller følelser er endnu mere kompliceret. Forventningerne er høje til teknologien, mulighederne er mange, men der er også begrænsninger, og i denne artikel afdækker vi nogle af dem, når vi undersøger en konkret case med brug af forskellige metoder.

I artiklen undersøger vi, hvordan vi kan træne en model – en 'vintagedetektor' – på historiske fotografier i Det Kgl. Biblioteks digitale samlinger (fra perioden 1880-1920). Vi forholder os kort til udfordringerne i 'simpel' objektgenkendelse i historiske samlinger og går derefter i dybden med en mere kompleks case, når vi efterprøver en påstand fremsat af historikeren Nina Søndergaard.¹

¹ For Mette Kia Krabbe Meyers vedkommende er artiklen skrevet som del af forskningsprojektet "Queer kvinder 1880-2020" finansieret af Danmarks Frie Forskningsfond. Bevillingsnummer 1024-00021B.

Corpus og metode

I vores analyse bruger vi både manuelle nærlæsninger (som vi kalder kvalitative undersøgelser) og automatiske fjernlæsninger baseret på maskinlæring (som vi kalder kvantitative metoder). Begge tilgange baserer sig på kuraterede corpora fra udvalgte digitale samlinger. I den manuelle analyse har vi brugt fotografier fra arkiv.dk udvalgt på baggrund af beskrivende metadata samt fotografier fra Visitkortsamlingen fra Det Kgl. Biblioteks Digitale Samlinger udvalgt manuelt. I den kvantitative analyse af de historiske fotografier har vi brugt objekt-detektorer på en delmængde af Det Kgl. Biblioteks digitaliserede samling af fotografier taget af kgl. hoffotograf Peter Elfelt (1866-1931). Denne samling omfatter i alt 140.774 fotografier, hvoraf 97.192 er portrætter (hvilket her betyder atelier- og kabinetkortfotografier). Da kun en begrænset del af disse fotografier er dateret med specifikke årstal, har vi afgrænset 'vintageperioden' til fotografierne, der er årstalsdaterede i perioden 1890-1920. Dette corpus omfatter i alt 26.071 fotografier.²

AI er vejen frem – skateboards i 1910'erne

Inden vi går i gang med Søndergaards tese, gør vi os nogle generelle overvejelser om objektgenkendelse. Det er nærliggende at tro, at de mange historiske fotografier, der ligger tilgængelige i forskellige digitale arkiver, og som kun har lidt eller måske ikke så brugbar metadata, nemt kan køres igennem en AI-model til objektgenkendelse (eng.: *object detection model*) for at opnå en masse præcis metadata.³ Vi er vant til at søge på billeder via Google og hurtigt få gode resultater. Der opstår imidlertid problemer, hvis man anvender gængse AI-modeller på fotografier fra før 1950. I de tilfælde 'genkender' modellen fx flade firkantede genstande som fjernbetjeninger (se ill. 1a) eller sabler som skateboards (ill. 1b). Den er trænet på nutidige billeder og gør det,

² Der er i samlingen desuden intervaldaterede portrætter fra perioderne 1890-1931 og 1931-1965, som altså ikke indgår i vores analyser. For tekniske specifikationer og dokumentation henvises til kodebasen (<https://github.com/hksorensen/vintage-vision>) samt datasættet (<https://doi.org/10.5281/zenodo.20269685>).

³ De modeller, som vi bruger, er alle af typen 'dybe neurale netværk', hvilket vil sige, at de er baseret på, at netværket er trænet på kendte data for at kunne lave forudsigelser om ukendte tilfælde.



a) Her har en model genkendt en fjernbetjening på et skrivebord i et vintagebillede.



b) Her har modellen identificeret en soldat med en sabel, som den tror er et skateboard.

Ill. 1: Når object detection-modeller bliver udsat for objekter uden for deres referenceramme, begynder de at 'hallucinere' og 'gætte' på betydningen af de objekter, som de identificerer.

der kaldes ‘at hallucinere’, og udviser et ‘lack of historicity’.⁴ For at forhindre dette kan man indføre en såkaldt ‘confidence score’, men det ændrer ikke ved, at modellen stadig kan være meget ‘sikker’ i sin tilskrivning af en kategori. I sådanne tilfælde af fejlagtig tilskrivning taler man om ‘falske positive’, og det er en generel udfordring at begrænse deres antal.

Objektgenkendelse over for billedbeskrivelse

Et andet problem opstår, fordi der er forskel på modeller for objektgenkendelse og det, man kalder modeller til billedbeskrivelse (på engelsk ‘image captioning’): Modeller beregnet til objektgenkendelse ‘leder’ efter bestemte klasser af objekter, fx bil, stol, hund, person osv. Det vil sige, at der er et endeligt og ret begrænset antal klasser, som modellen kan genkende og klassificere. Som en særlig anvendelse af objektgenkendelse findes en række modeller, som kombinerer objektgenkendelse med sprogmodeller til at lave ‘image captioning’. Det vil sige, at modellen giver en overordnet, kort beskrivelse af billedet ved hjælp fra en af de store sprogmodeller i stedet for blot at klassificere objekterne i billedet. Inddragelsen af sprogmodellerne rejser imidlertid adskillige nye spørgsmål og skaber nye udfordringer i forhold til metadatering af indhold. Sprogmodellerne vil nemlig typisk ikke respondere *konsekvent* i beskrivelserne, idet de oftest er trænet til en form for ‘samtalemed brugeren. Et simpelt eksempel fra vores analyser er et billede af en soldat klassificeret som ‘mand’ af en af modellerne, mens en anden model beskrev vedkommende som “en soldat”. Dvs. at denne form for ‘image captioning’ ofte er inkonsekvent, for så vidt som den giver to forskellige resultater, nemlig overordnet metadata og metadata, som er definerende inden for kategorien. Både ‘mand’ og ‘soldat’ er semantisk rigtige, men forsøger man at anvende disse ‘captions’ i en leksikal søgning på metadata, forsvinder billedet i det, man kalder ‘det usete hav’. I en fritekstsøgning ledes der efter de konkrete ord eller en streng af ord, som er tilgængelige i form af metadata (hvilket typisk er meget begrænset), mens semantisk søgning søger på billedets indhold. Derfor anvendes begge metoder. Hvis man tager det naturlige sprog

⁴ Francis Hunger: ‘Why so Many Windows?’ – How the ImageNet Image Database Influences Automated Image Recognition of Historical Images. *International Journal for Digital Art History* 6 (3), 2023, s. 71-85.

med i betragtning, er det endnu mere komplekst, fordi det indeholder et væld af ord, som de omtalte metoder ikke får med, fx 'korporal', 'mand i uniform' eller 'ung mand i uniform'. For at inkludere dem, skal de oversættes gennem fx en sprogmodel til en forsimplet søgestreng. Denne tilgang virker ikke i praksis, fordi skellet mellem beskrivelse og indhold er lagt forkert, og indholdssøgninger derfor kommer til at afhænge for meget af modellen og for lidt af den mening, der søges efter. Endnu mere kompliceret bliver situationen af, at det kan være – eller nærmest altid er – svært at vide, hvordan modellerne kommer frem til de resultater, de giver. Tilsammen fører alle disse forhold til, at vi i øjeblikket står i en situation, hvor de tilgængelige modeller er for begrænsede til at foretage effektive søgninger og få præcise resultater.

Ikke ulig bryllupsbilleder

Bevæger vi os videre fra objektgenkendelse til betydningen af relationerne mellem dem på basis af deres placering, bliver det endnu mere kompliceret. Den konkrete case, vi har arbejdet med, er en påstand fremsat af historikeren Nina Søndergaard, der skriver om kærlighed mellem kvinder omkring år 1900.⁵ I sit arbejde anvender hun som andre historikere skriftlige kilder som erindringer, folketællinger, breve m.m. for at belyse de forhold mellem kvinder, som der har været en tendens til at karakterisere som venindeforhold, og ser dem i et LGBTQ+-perspektiv. Samtidig inddrager hun motivanalysen af fotografier som metode, når hun om et fotografi af skolelederen Natalie Zahle og lærerinden Poulina Govl fra begyndelsen af 1860'erne samt ét af lærerinderne Marie Topsøe og Augusta Fenger fra 1880'erne skriver, at de henholdsvis "mimer datidens bryllupsfotografier" og er "ikke ulig datidens bryllupsfotografier" (se ill. 2).⁶ Ved at konstatere, at parrene er fotograferet som brudepar, tillægger hun billederne en betydning på grundlag af genren, dvs. hun lader fotograferingen selv – bryllupsfotograferingen – være et argument for kvindernes nære relation.⁷

⁵ Nina Søndergaard: Linnésgades 'lesbiske' lærerinder – kristen kvindekærlighed anno 1877, *Historiske meddelelser om København*, 2024, s. 69-111.

⁶ S. 98 og s. 79.

⁷ Søndergaards artikel benytter i stort omfang digitale og kvantitative metoder til literatursøgning og demografiske oplysninger, men lige disse to påstande om 'repræsentativiteten' af disse portrætter gives der ikke yderligere belæg for.



Ill. 2: Portræt af Marie Topsøe og Augusta Fenger (th) taget af Christian Adolph Barfod Lönborg i 1880'erne samt portræt af Pouline Govl og Nathalie Zahle (th) taget af Otto Hermann Emil Lange i 1860'erne.

En sjældent anvendt metode

I forskningen i romantiske forhold mellem kvinder er Søndergaards greb nyt og spændende. Historisk set har man i LGBTQ+-historieskrivningen anvendt portrætter eller fotografier af bestemte steder, men ikke som beviser på, at en person har tænkt på en bestemt måde eller haft en særlig relation til en anden portrætteret på et dobbeltportræt. Der er dog en interessant udgivelse i den danske LGBTQ+-forsknings historie, nemlig den første danske bog om kvinders forhold til kvinder, historikeren Karin Lützens *Hvad hjertet begærer*, der udkom første gang i 1986.⁸ I bogen optræder der forskellige såkaldte 'found photographs', dvs. billeder uden oplysninger om de portrætterede eller fotografen,

⁸ Karin Lützen: *Hvad hjertet begærer. Kvinders kærlighed til kvinder*, 1986.

som forestiller kvinder, der sidder tæt eller ser på hinanden med et 'særligt' blik. Lützen har selv udtalt, at fotografierne ikke optræder som historiske kilder, der skal dokumentere, at kvinderne havde et forhold, men som *freie* illustrationer, der i bogens tilfælde understøtter en pointe om det uafklarede eller uvisse i mange forhold – i virkeligheden og i eftertiden, hvor der kun er bevaret få kilder, og hvor erotik sjældent bliver dokumenteret. Man kan sige, at den tvetydighed, der altid er i et fotografi, hvis kontekst er ukendt, her optræder sammen med det, Lützen definerer som en generel tvetydighed i forholdene mellem kvinder dengang. En lignende indstilling finder man hos forfatteren David Deitcher, der i bogen *Dear Friends: American Photographs of Men Together* tilføjer et nyt aspekt, nemlig at det bringer beskueren i en nærmest flirtende relation til billedet, fordi engagementet i den uforløste situation på billedet svarer til det evigt nysgerrige og uforløste i at betragte et 'found photograph'.⁹

Søndergaards tilgang er væsensforskellig, for her er en tese om, at fotografierne *vidner* om en nær relation mellem Zahle og Govl samt Topsøe og Fenger, bygget på en aflæsning af deres position og en sammenligning med genren bryllupsfotografi, og det gør den som nævnt mulig at efterprøve empirisk.

Den kvalitative tilgang: Lille lighed med brudepar i samlinger

I undersøgelsen forsøger vi os først med en klassisk, manuel analyse uden anvendelse af automatisk objektgenkendelse. Først skaber vi en samling fotografier af brudepar ved at søge på fotografier, der er metadateret med 'bryllup' og 'brud' på arkiv.dk og i Elfelt-samlingen. I søgeresultatet er der en del billeder, som afviger fra atelieroptagelserne i Søndergaards artikel, fx. brudeoptog i gader etc., men også meget overraskende motiver som et fotografi af to brydere til en opvisning i forbindelse med prinsesse Ingeborgs bryllup i 1897.¹⁰ Efter en frasortering af disse billeder, tager vi også fotografier, hvor brudepar er fotograferet stående, fra, hvorefter vi har en samling fotografier med samme opstilling som billederne i Søndergaards artikel.

⁹ David Deitcher: *Dear Friends: American Photographs of Men Together*, 2001, s. 16.

¹⁰ <<https://digitalesamlinger.kb.dk/images/billed/2010/okt/billeder/object2112634/da/>>.

I Elfelt-samlingen er der i størrelsesordenen 150 fotografier med ordet 'bryllup' angivet i metadata. Ud af disse er der 22 parbilleder, deraf fem med en siddende person (4 kvinder og 1 mand). Tilsvarende er der 187 hits, hvis man søger på ordet 'brud'. Af disse er 13 parbilleder med en siddende person (12 kvinder og 1 mand). En søgning på arkiv.dk resulterer i mere end 300 fotografier markeret med 'bryllup' i metadata. Heraf 101 parbilleder, hvor 23 af billederne viser én siddende person (9 kvinder og 14 mænd).

På fotografierne af brudeparrene er der en tendens til, at den ene part er vendt mod den anden, en stol er drejet en smule, og en stående figur vender sig mod den siddende; de står med andre ord sjældent vendt frontalt mod kameraet, således som tilfældet er med Zahle og Govl. I Fenger og Topsøes tilfælde er der også forskel, for på bryllupsfotografierne er der ingen, der hviler armen på noget, således som Marie Topsøe gør. Resultatet af undersøgelsen peger altså på, at der kun er et overlap mellem fotografiet af Zahle og Govl og de fotografier, hvor parterne ikke er vendt mod hinanden. Det betyder, at resultatet af den kvalitative sammenligning af fotografier viser, at kvinderne ikke er optaget efter den bryllupsfotografiskabelon, som findes i de udvalg, vi har foretaget.

Andre fotografier af to kvinder

Hvis vi vender os fra bryllupsfotografierne, kan vi spørge, om der mon er fællestræk mellem den måde, som mænd og kvinder mere generelt er fotograferet sammen på, og den måde, som to kvinder er det. Dette medfører en ny opgave, og den løser vi først ved manuelt at sortere et udvalg af kvinde-kvinde-fotografier og kvinde-mand-fotografier i Visitkortfotografisamlingen, der består af fotografier fra hele landet optaget af mange forskellige fotografer.

I en delmængde af Visitkortfotografisamlingen på 4.000 fotografier kan man uddrage 36 fotografier af to kvinder, hvor den ene er stående og den anden siddende. Når man ser dem manuelt igennem, er det tydeligt, at kvinder er positioneret på mange forskellige måder, og at der findes flere eksempler, hvor en kvinde lægger armen om ryggen på en anden. Da fotografierne som angivet er uden metadata, er der en mulighed for, at de 36 fotografier forestiller kvinder i forhold, og at formodningen om, at de fx er mor-datter eller søster-søster kan være præget af en normativ tilgang.

Den kvantitative tilgang: Metode og motivation

Den manuelle udvælgelse af de 36 fotografier af to kvinder er et møjsommeligt arbejde og er her udført på et lille kildemateriale. Anvender man en automatisk analyse, vil man kunne analysere langt flere fotografier. Det kræver imidlertid, at man har et analyseværktøj, der kan identificere de portrætter, som indeholder præcis to voksne personer, og at man kan skelne mellem mænd og kvinder. Netop det værktøj forsøger vi at udvikle, og til det formål anvender vi visitkortfotografier alene fra Elfelt-samlingen. Opgaven skal løses ved at kombinere forskellige 'byggeklodser'. I vores corpus (der omfatter i alt 26.071 fotografier), er fordelingen som vist i figur 3.¹¹

Det virker oplagt, at opgaven med at identificere personer i et portræt og forudsige (eng.: *predict*) deres køn kan løses ved objektgenkendelse, men standardlisten af klasser, som modellerne bruger, omfatter kun klassen 'person' og ikke forskellige køn.¹² Da YOLO-detektoren, vi anvender, er trænet på moderne billeder, opstår der også et behov for at anvende en detektor, der er trænet på et mere varieret materiale.¹³

Den mest oplagte løsning er at bruge Det Kgl. Biblioteks samlinger til at specialtræne en persondetektor på vintagebilleder. Men før dette kan lade sig gøre, skal træningssættet annoteres, idet personer skal identificeres og kønsbestemmes. For at denne opgave, som vil være krævende, hvis den skal udføres af mennesker, kan automatiseres, bruger vi en anden model, nemlig Microsofts Florence-2-model,¹⁴ som bl.a. indeholder objektgenkendelse med mange flere klasser, herunder klasserne 'man', 'woman', 'girl', 'boy' og 'person'.¹⁵ Det viser sig, at

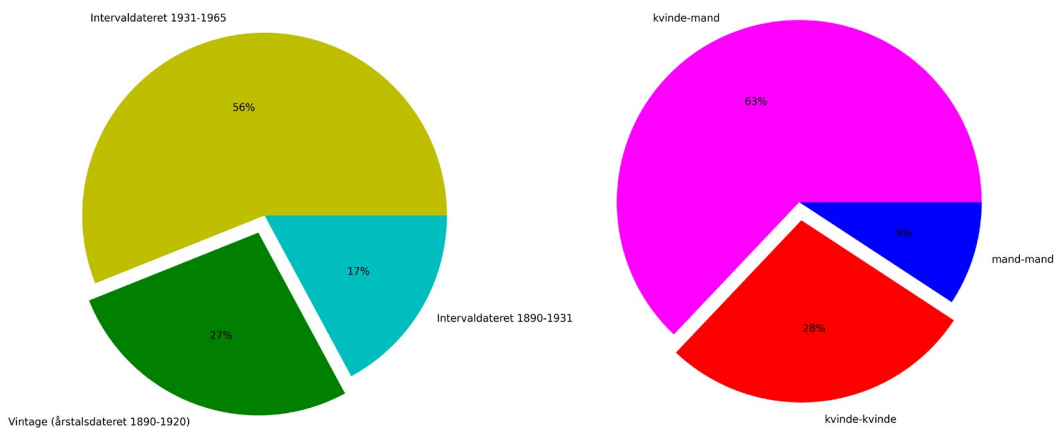
¹¹ Der er i samlingen desuden intervaldaterede portrætter fra perioderne 1890-1931 og 1931-1965, som altså ikke indgår i vores analyser.

¹² YOLO står for "you only look once" og er en familie af object detection-modeller, hvoraf den nyeste dags dato er YOLOv13. Joseph Redmon et al.: You Only Look Once. Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CV-PR)*, 2016, s. 779-788.

¹³ Se også Melvin Wevers og Thomas Smits: The Visual Digital Turn: Using Neural Networks to Study Historical Images. *Digital Scholarship in the Humanities*, 35(1), 2020, s. 194-207.

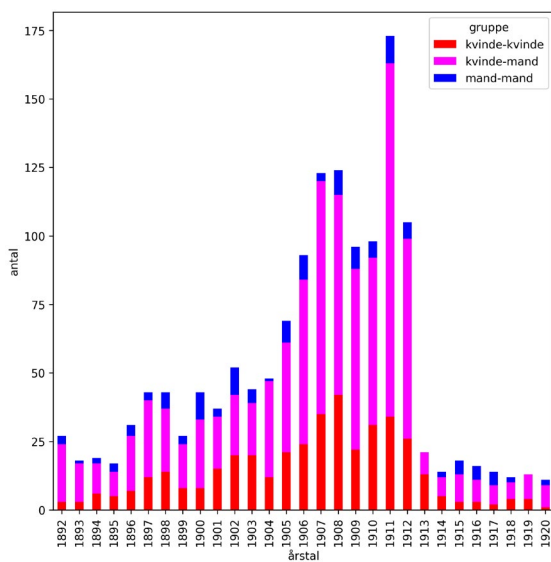
¹⁴ Bin Xiao et al.: Florence-2. Advancing a Unified Representation for a Variety of Vision Tasks. *arXiv*. 2311.06242, 2023.

¹⁵ Det er ikke beskrevet, hvordan Florence-2-modellen definerer køn, og vi er klar over, at det i en LGBTQ+-kontekst er vigtigt at arbejde med diversitet og undgå udbredte stereotyper og bias. Det vil være oplagt at udforske i en undersøgelse af kompleksite-



a) Fordeling af portrætter i Elfelt-samlingen på tidlige subcorpora (N=97.192).

b) Fordelingen af topersportsportrætter på grupper (N=1449).



c) Fordeling over tid af topersportsportrætter i Elfelt-samlingen (N=1449).

III. 3: Visualisering af informationer om datasættet.



Ill. 4: Skematisk oversigt over den pipeline, der tilvejebringer det endelige datasæt.

denne tilgang til *det konkrete problem* faktisk gør det overflødigt at træne en specialiseret persondetektor.

I stedet implementerede vi en arbejdsgang, som benytter Florence-2-modellen til at identificere ansigter og knytte dem til personer, hvis køn Florence-2 også kan forudsige. Disse ansigter sender vi så videre til en model baseret på RetinaNet,¹⁶ som kan detektere øjne i ansigter. Som det sidste skridt i forberedelsen af datasættet benytter vi forskellige visualiseringer til at bekræfte datasættets generelle kvalitet og til at fjerne ‘outliers’, der ville forstyrre analyserne.¹⁷ Denne proces for dataforberedelse er vist i skematisk form i ill. 4.

Pålidelig detektion i vintagebilleder med minimal human-in-the-loop

For at processen senere kan bruges til at træne en detektor, er det som sagt vigtigt, at dens træningsgrundlag er så præcist som muligt. Vi vælger derfor at sortere en del fotografier fra, hvor vi statistisk er i tvivl om identifikation eller klassifikation. Med andre ord forsøger vi at maksimere andelen af korrekte identifikationer (eng.: *true positives*, forkortet TP) ved at begrænse datasættet.

Klassen ‘person’ omfatter personer, hvis køn ikke kan bestemmes ud fra modellens kategorier. Modellen er dog overraskende præcis i bestemmelsen af køn. Når man ser på de mennesker, der ikke kønsbestemmes, er det formentlig baseret på, at fotografiet kun viser et udsnit af dem. Klasserne ‘boy’ og ‘girl’ har ret vage grænser til de tilsvarende ‘man’ og ‘woman’, så vi ser bort fra børneklasserne.

ten i 1800-tallets klædedragt for kvinder, men i nærværende analyse opererer vi med Florence-2’s binære kategorisering.

¹⁶ Tsung-Yi Lin, Priya Goyal et al.: Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, s. 2980-2988, 2017.

¹⁷ En ‘outlier’ er en eller få observationer, der adskiller sig abnormt meget fra hovedparten af en gruppe. Til statistiske undersøgelser så simple som gennemsnitsberegninger vil outliers forstyrre resultaterne for den langt største hovedpart af gruppen. Derfor kan man – efter grundig overvejelse – vælge at se bort fra outliers.

Modellen har yderligere en klasse, 'human face', som er vigtig for os af to grunde. For det første tillader den yderligere en betingelse, som er relevant for at maksimere andelen af TP: Det er nødvendigt, at antallet af ansigter er det samme som antallet af voksne personer ('man' og 'woman'). For det andet gør klassen 'human face' det muligt at udsondre den del af fotografiet, hvori en persons øjne befinder sig.

De instanser, som opfylder ovenstående filtreringskriterier, kalder vi for vores 'maskinvaliderede corpus' (MVC), og når vi begrænser dette til de fotografier, der netop indeholder to personer, ender vi med et corpus bestående af 1.315 fotografier, hvori der er detekteret to personer, hvoraf mindst en er kvinde, og vi ved, hvilke personer der er mænd, og hvilke der er kvinder. Vi kalder denne samling for 'topersonsprofilerne' og underinddeler den i portrætgrupperne (i det følgende blot grupperne) kvinde-kvinde (403 stk.) og kvinde-mand (912 stk.) (se ill. 3b).¹⁸ Man observerer her, at filtreringsprocessen gør det muligt og nok sandsynligt, at den fuldt automatiserede metode har 'overset' enkelte instanser (FN), eller at den har medtaget enkelte falske positive (FP). Men til det aktuelle formål vurderer vi ud fra stikprøver, at vores corpus besidder en tilstrækkelig kvalitet – ikke mindst fordi vi har fundet meget få falske positive. Det er i denne inspektion og vurdering, at den menneskelige intervention (human-in-the-loop) optræder i vores datakonstruktion. En visualisering af gruppernes størrelse over tid (se ill. 3c) antyder en udvikling og især en drastisk reduktion i antallet af kvinde-mand-portrætter fra 1911 til 1913. Men signifikansen og fortolkningen af sådanne hypoteser, som er dannet på grundlag af ret oplagte og let tilgængelige visualiseringer, efterlades som åbne spørgsmål til et andet studie.

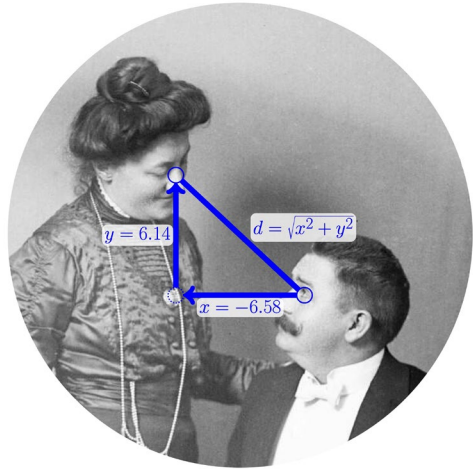
Beregning af øjereationer

For at behandle opstillingen i fotografierne har vi valgt en proxy, dvs. en kvantificerbar størrelse, som på en eller anden måde måler de egenskaber (opstilling), som vi egentlig er interesserede i. Den proxy, som vi har valgt, fokuserer på afstanden, såvel vandret som lodret, mellem de to personer på billedet. Af pragmatiske grunde har vi valgt at måle afstand som en normeret afstand mellem midtpunktet mellem hver persons

¹⁸ Der findes desuden 134 stk. mand-mand-billeder i samlingen, som ikke er medtaget i vores corpus, da de ikke er relevante for vores forsknings spørgsmål.



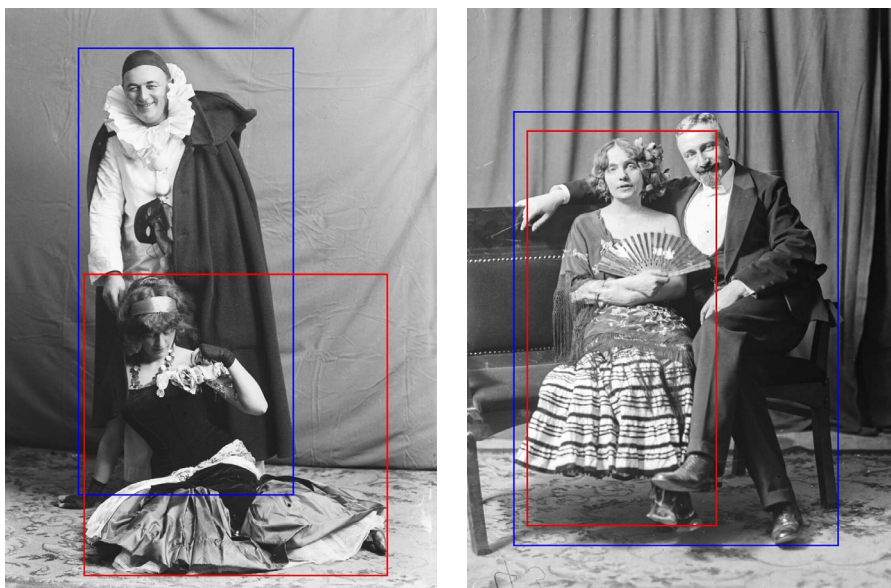
a) Orienteringen i billederne er sådan, at *opad* svarer til positive y -værdier, og *til højre* svarer til positive x -værdier.



b) Alle beregninger tager udgangspunkt i enten mandens øjne (i kvinde-mand-billeder) eller (i kvinde-kvinde-billeder) i øjnene på den kvinde, hvis øjne er placeret højest i billedet.

Ill. 5: Forklaring af fortegn i koordinatangivelser. En positiv y -værdi i kvinde-mand-billeder svarer til, at kvindens øjne er højere placeret i billedet end mandens (positive y -værdier giver derimod slet ikke mening i kvinde-kvinde-billeder). Positive x -værdier svarer til, at kvinden (hhv. den nederste kvinde) befinder sig til højre for manden (hhv. den øverste kvinde). Alle afstande er udtrykt i den enhed, som er beregnet for det enkelte billede.

øjne. Denne proxy står ikke i en en-til-en-relation til opstillingen, men den tillader os at operationalisere forskningsspørgsmålet på en måde, som er teknisk mulig, samtidig med at den virker plausibel i forhold til formålet. Når vi har valgt at normalisere afstanden mellem øjenparrene, er det for at tage hensyn til et andet vigtigt aspekt af portrætanalyser, nemlig beskæringen: Afstand i billedet, som er målt i pixels, vil i virkeligheden være bestemt af, hvor stor en del af motivet, der er synligt. Vi indfører derfor en enhed som den største afstand mellem en persons øjne. Denne enhed har den vigtige fordel, at den gør afstande dimensionsløse, hvilket tillader sammenligninger mellem billeder.



a) Elfelt 1911. Mandens ansigt falder kun i én boks.
 b) Elfelt 1911. Begge ansigter overlapper med begge bokse.

III. 6: Udredning af match mellem øjne og personer.

Identifikationen af øjne er foretaget ved hjælp af modellen Batch-Face, som er baseret på RetinaFace.¹⁹ Denne model muliggør effektiv identifikation af højre og venstre øje i et billede, og for at sikre forudsigelsernes kvalitet har vi anvendt modellen direkte på de ansigter, som vores brug af Florence-2 havde udpeget.

En udfordring opstår, fordi personerne i nogle af billederne dækker over hinanden, og modellen derfor ikke kan bestemme ansigter i billederne. Da vi gerne vil knytte ansigterne til de kønsbestemte personer, har vi foretaget en simpel matematisk reduktion (se ill. 6).

¹⁹ Elliott Zheng et al.: Batch Face for Modern Research, <<https://github.com/elliottzheng/batch-face>>, 2025, og Deng Jiankang et al.: RetinaFace. Single-stage Dense Face Localisation in the Wild, *arXiv*. 1905.00641, 2019.

Behandling af dobbeltfotografier

Når man anvender machine learning i humanistiske projekter, er det vigtigt at inkludere en menneskelig indsats på udvalgte steder i processen. Indledningsvis har vi anvendt en automatiseret metode til at identificere objekter i billederne og til at danne vores centrale MVC. Derefter har vi set på resultaterne og identificeret 'outliers' og andre fænomener for at få et homogent datasæt tilpasset undersøgelsen.

Under inspektion af resultaterne blev vi opmærksomme på, at der var nogle fotografier, hvor modellen påviste usædvanligt store afstande mellem øjenparrene. Da vi så nærmere på dem, fandt vi ud af, at det skyldes, at nogle af fotografierne bestod af to ens portrætter på samme negativ, og at de derfor måtte fjernes fra søgesættet.

Da vores fokus er på bryllups- og andre parbilleder, antager vi, at de to mennesker ikke står langt fra hinanden, og da der endvidere er få 'dobbeltfotografier', er det ikke væsentligt for analysen at trække dem ud af corpus. Til at udskille dobbeltportrætterne har vi benyttet 'template matching'. Det er en statistisk, såkaldt 'unsupervised' metode til at finde forekomster af gentagelser af et mønster i et billede. I vores tilfælde vælger vi at søge efter forekomster af den ene kvindes ansigt i billeder, som indeholder to kvinder. Metoden er statistisk, og vi har derfor sat en grænseværdi for, hvor sikre matches vi ønsker at fange. Denne grænseværdi har vi sat retrospektivt, dvs. efter at den statistiske analyse var gennemført, idet der ikke på forhånd var nogen fornuftig grund til at vælge den ene eller den anden grænseværdi. Ved manuelt at inspicere de billeder, som faldt under forskellige grænseværdier, kunne vi se, at med en grænseværdi på 0,8 ville vi ramme et godt forhold mellem falske og sande positive.

'Outliers' og endelig trimning af datasættet

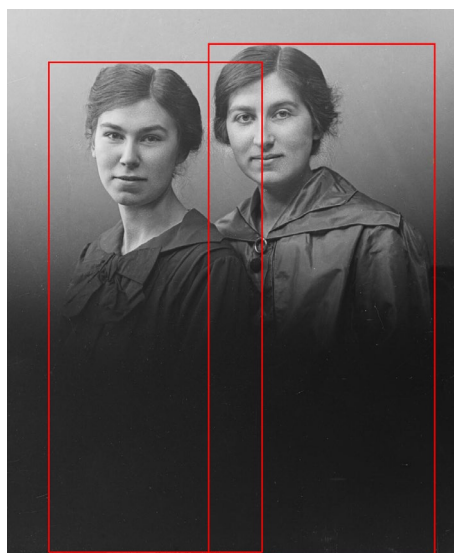
Det sidste skridt i forberedelsen af datasættet er at filtrere 'outliers' bort. Vi har valgt at filtrere ud fra afstanden mellem øjnene, og vi har valgt at fastsætte en grænseværdi på 12 enheder. Dette er en grænseværdi, som vi har valgt ud fra vores egen erfaring med at inspicere datasættet, og den er dermed retrospektivt valgt ligesom grænseværdien for dobbeltfotografier.



a) Elfelt, 1902.

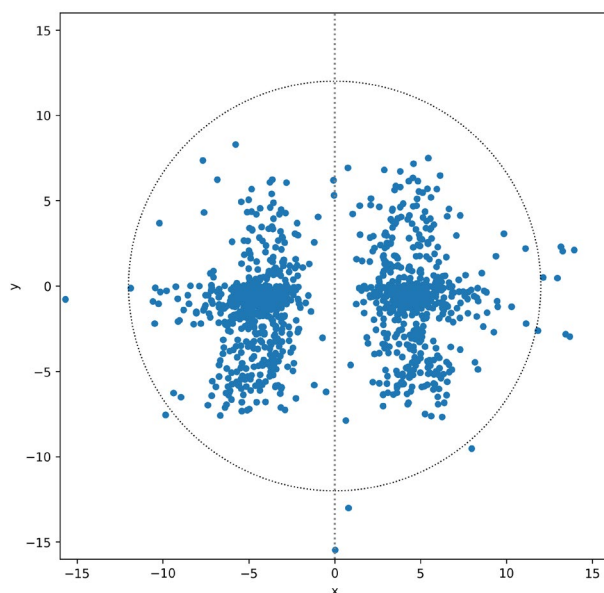


b) Elfelt, 1914.



c) Elfelt, 1918.

Ill. 7: I Elfelt-samlingen findes eksempler på 'dobbeltfotografier', som vil 'forurene' datasættet. Eksemplet 7a viser en detektion med meget høj sikkerhed, mens sikkerheden i 7b er lavere, men stadig over den grænseværdi, vi retrospektivt har fastsat.



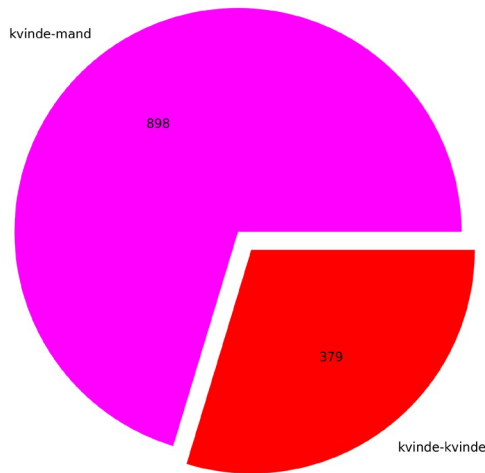
Ill. 8. Ved for kvinde-mand-gruppen at fastsætte en maksimal radius på 12 enheder kasseres 1,47 % af datasættet; det resterende datasæt bliver derved mere homogent.

Efter at have fjernet dobbeltfotografier og ‘outliers’ som vist i ill. 8 og 9, nåede vi til det endelige datasæt, som består af 898 portrætter i kvinde-mand-gruppen og 379 i kvinde-kvinde-gruppen (se ill. 9).

Den ovenfor beskrevne kvantitative metode til at uddrage, behandle og filtrere datasættet er dikteret af forskningsspørgsmålet og rammesat af de tekniske muligheder. Kun få skridt i processen og valg af grænseværdier var beskrevet på forhånd, og selve processen foregik som en vekselvirkning mellem automatisk analyse, vurdering af resultater og valg og udvikling af modeller.

Resultater af de kvantitative big data-analyser

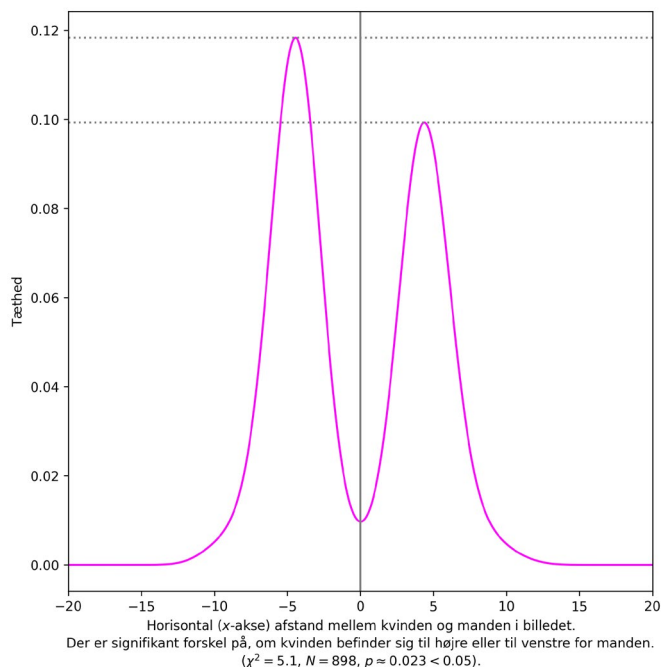
I det følgende beskriver vi de resultater, vi er nået frem til. Vi har valgt at inddrage både visuelle og statistiske analyser, og vi har valgt at fokusere på de resultater, som er relevante for den konkrete case.



Ill. 9: Det endelige datasæt omfatter de to grupper kvinde-kvinde (379) og kvinde-mand (898), i alt 1.277 portrætter.

1. Vi foretog først en statistisk analyse af, hvordan mand og kvinde er placeret i parbillederne. Analysen er baseret på en visualisering af den horisontale afstand mellem kvinden og manden i billedet, som ledte til en hypotese, der kunne testes statistisk.
2. Dernæst beskrev vi et kombineret statistisk og visuelt studie af klyngestrukturen i dette kvinde-mand-datasæt. Vi var nærmere bestemt interesseret i at undersøge, hvorvidt der er generelle mønstre i opstillingen i disse billeder. Vi kunne derefter overføre denne tilgang til undersøgelsen af klyngestrukturen (eng.: *cluster*) i kvinde-kvinde-datasættet.
3. Endelig brugte vi disse dele til at afklare spørgsmålet om, hvorvidt et konkret kvinde-kvinde-billede er typisk i corpus eller ej.

Da disse resultater er svar på forskellige spørgsmål, kan sikkerheden af dem ikke sammenlignes direkte. Nogle af resultaterne er baseret på en statistisk test, mens andre er baseret på en visuel inspektion og fortolkning.



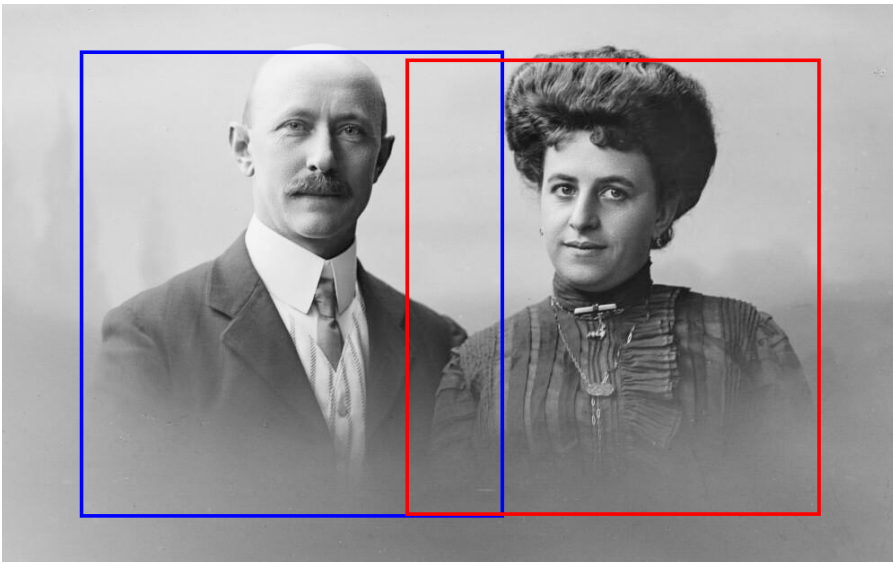
Ill. 10: Grafisk og numerisk illustration af fordelingen af kvinder til venstre og højre for mænd i kvinde-mand-gruppen.

Analysen af kvinde-mand-billeder

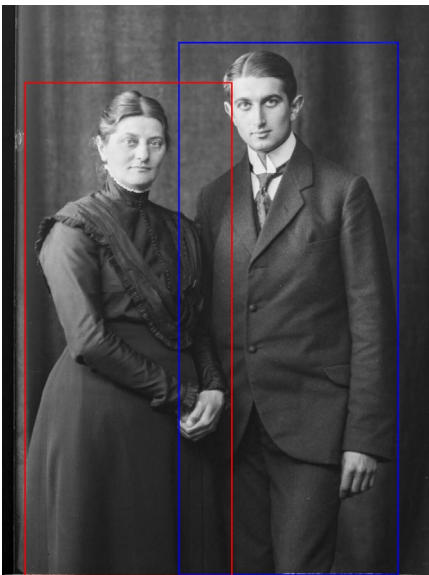
Det fremgår umiddelbart af ill. 8, at der er meget få billeder omkring akse $x = 0$, hvilket også er forventeligt, idet billeder med meget små værdier for den horisontale afstand, svarer til, at ansigterne befinder sig lodret over hinanden. Vi kunne have formodet på forhånd, at en sådan opstilling ville være sjælden, og nu er det så også visuelt bekræftet.

Som en del af vores indsatser for at vurdere og validere vores datasæt har vi plottet forskellige kvantitative data, som er lette at beregne.²⁰ Vi har for eksempel plottet fordelingen af den horisontale afstand mellem kvinden og manden i billedet i form af et 'histogram' med en 'kernel density estimation' (se ill. 10).

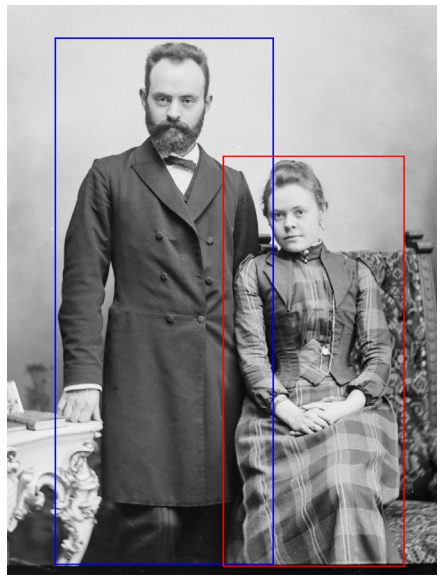
²⁰ Det er vigtigt at være påpasselig med, hvad man kan udlede af denne slags eksplorative plots, hvis man ikke forbinder dem med en teoretisk model eller en statistisk test.



a) Elfelt, 1909.



b) Elfelt, 1916.



c) Elfelt, 1892.

Ill. 11: Eksempler på typiske øjenplaceringer horisontalt i kvinde-mand-gruppen.

Af ill. 10 fremgår det, at kvinden befinder sig til venstre for manden i de fleste billeder. Dette er en hypotese, der kan testes og påvises statistisk, fx ved hjælp af en χ^2 -test, og den viser, at der er en statistisk signifikant forskel på, om kvinden befinder sig til højre eller til venstre for manden ($\chi^2 = 4.2$, $N = 912$, $p \approx 0.040 < 0.05$). Vi kan dermed fremsætte vores første kvantitativt underbyggede konklusion: I vintagebillederne i Elfelt-samlingen står kvinden typisk længere til venstre end manden. Oversat til fotografiets personers perspektiv betyder det, at kvinden står til mandens højre side.

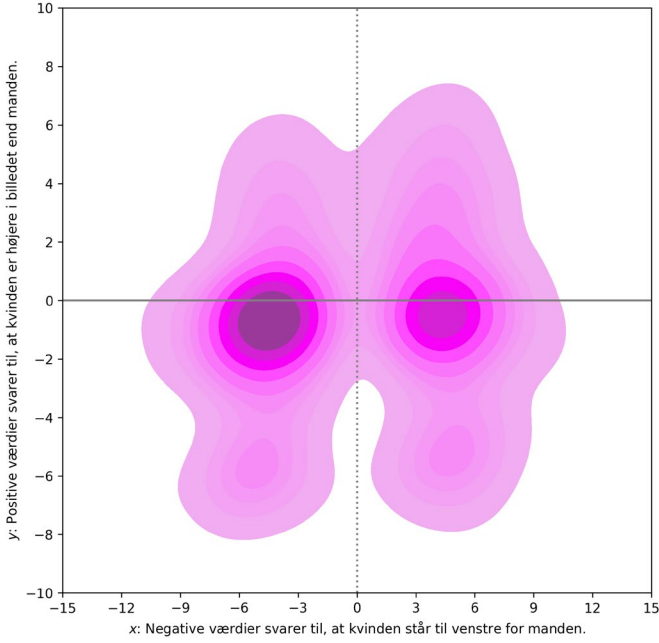
Når den kvantitative analyse er gennemført, kan man bruge den til at udtrække eksempler på typiske opstillinger i kvinde-mand-gruppen, som vi har plottet i ill. 10. Man kan ikke generalisere ud fra de konkrete billeder, men de antyder dog, at de to personer typisk er opstillet således, at deres skuldre overlapper. På den måde kan man forsøge at koble enheden, som er målt som afstand mellem øjnene, til yderligere udsagn om motivets opstilling i billedet.

Typiske opstillinger

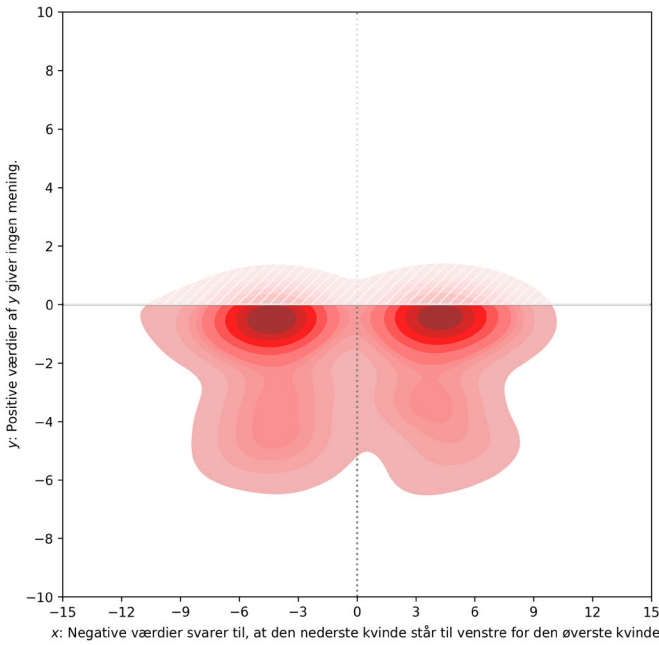
Hvor vi indtil nu primært har analyseret vandrette afstande mellem øjenpar, er det også interessant at se på deres afstand i lodret linje. Dette har vi gjort ved at skabe et 'heatmap' af afstandene mellem øjnene i såvel kvinde-mand-billederne som kvinde-kvinde-billederne (se ill. 12). Et heatmap kan læses som en slags topografisk kort, hvor højdekurverne viser, hvordan øjnene er placeret i forhold til hinanden.

I kvinde-mand-billederne tager vi udgangspunkt i, at mandens øjne er placeret i koordinatsystemets centrum (0,0), og derudfra måles så retningen og afstanden til kvindens øjne. I kvinde-kvinde-billederne har vi i stedet placeret det par øjne, der ligger højest i billedet, i koordinatsystemets centrum; det giver derfor ikke mening at have positive y -værdier, da det ville betyde, at den anden kvindes øjne var placeret højere i billedet end udgangspunktets.

De to 'heatmaps' i ill. 12 viser umiddelbart, at der findes strukturer i portrætternes opstilling. For at analysere fordelingen af typiske portrætter kan man fx benytte k -means-klynger. Det er en statistisk metode, der ud fra et på forhånd givet antal klynger (grupper) fordeler datapunkterne, således at data inden for en klynge hører mere sammen (ligger tættere på hinanden) end data på tværs af klynger. Antallet af klynger skal angives på forhånd, og da vi ikke har nogen a priori-antagelser om



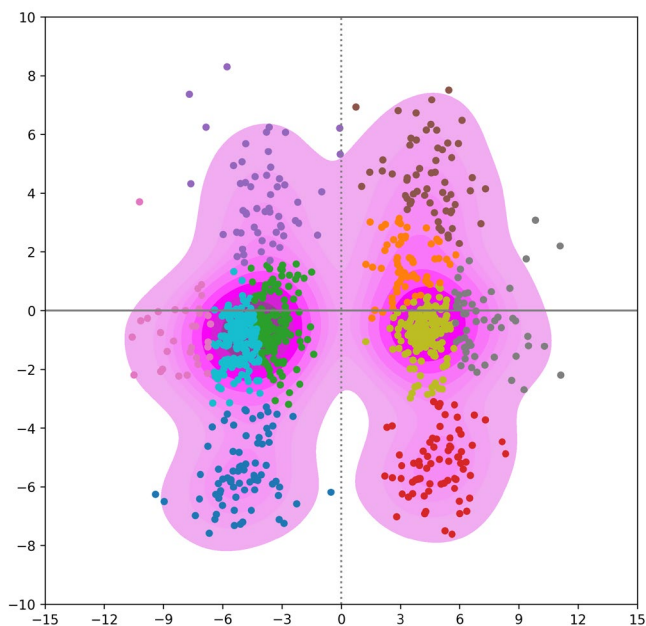
a) Kvinde-mand-billeder.



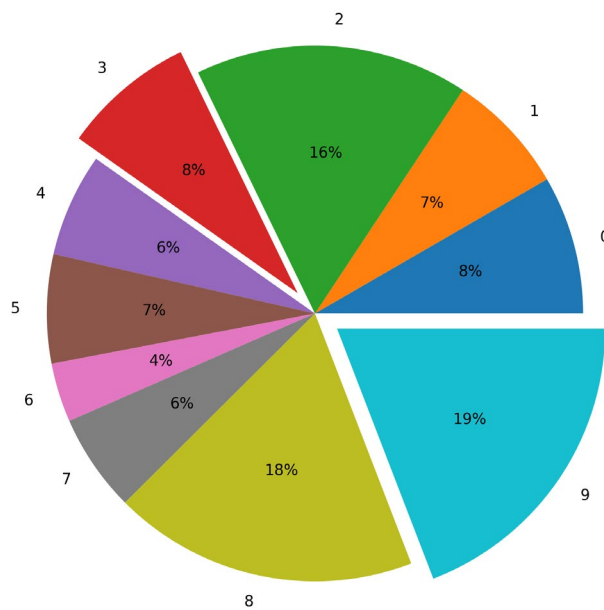
b) Kvinde-kvinde-billeder.

Ill. 12: 'Heatmap' af afstandene mellem øjnene i kvinde-mand-billederne og kvinde-kvinde-billederne. b) udgør en direkte udvidelse af ill. 10 til to dimensioner.

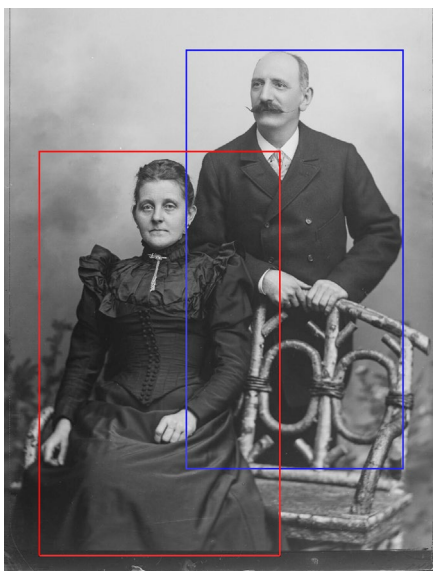
a) 10 klynger i kvinde-
mand-gruppen.



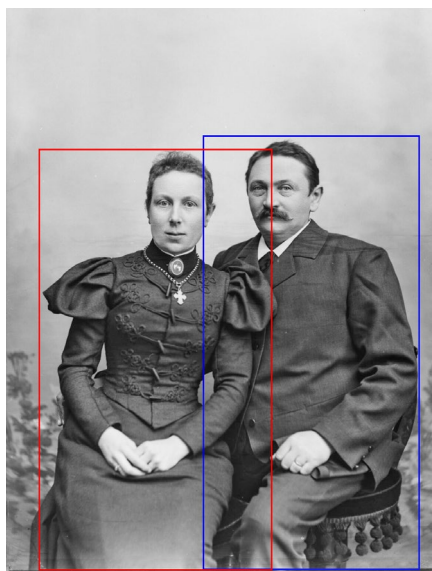
b) Størrelserne af de
10 klynger i kvinde-
mand-gruppen.



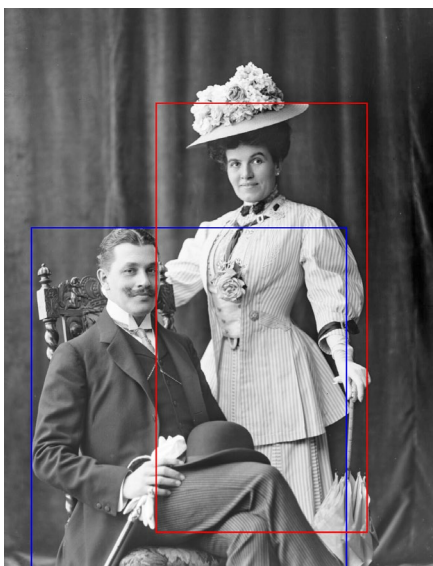
Ill. 13: Klyngerne i
kvinde-mand-grup-
pen.



a) Elfelt, 1893: klynge 0.



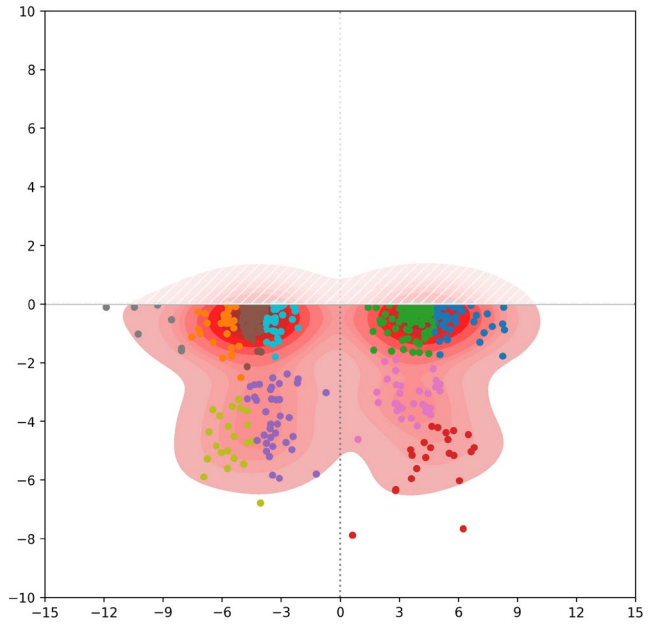
b) Elfelt, 1898: klynge 2.



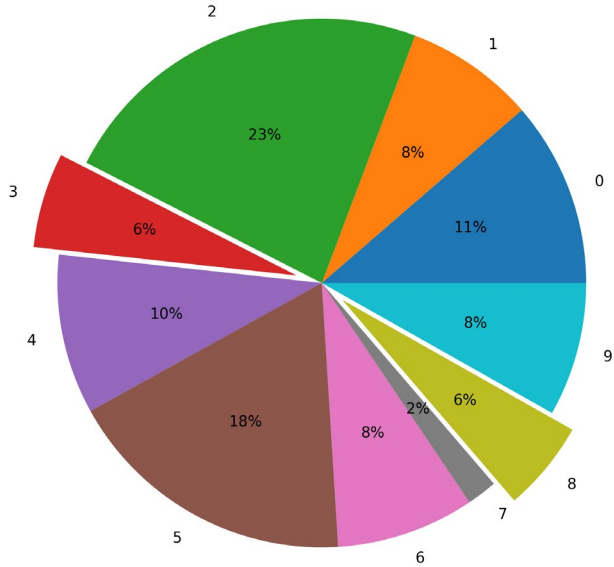
c) Elfelt, 1906: klynge 5.

III. 14: De kvinde-mand-portrætter, som ligger tættest på klyngecentrene i klynge-
nerne 0, 2 og 5.

a) 10 klynger i kvinde-kvinde-gruppen.



b) Størrelserne af de 10 klynger i kvinde-kvinde-gruppen.



Ill. 15: Klyngerne i kvinde-kvinde-gruppen.

dette antal, benytter vi en generel metode (i feltet kaldet ‘albuemetoden’) til at bestemme det optimale antal klynger, og for både kvinde-mand- og kvinde-kvinde-datasættene viser dette antal sig at være 10.

Ill. 13 viser, hvordan portræterne i kvinde-mand-gruppen fordeler sig på de 10 klynger. Det fremgår nu endnu tydeligere, hvordan nogle af klyngerne, nærmere bestemt de største klynger, 2, 8 og 9, er tæt samlede omkring øjne i samme højde som mandens. Noget tilsvarende gør sig gældende for kvinde-kvinde-gruppen (ill. 15), hvor der dog er lidt mindre forskel i størrelserne mellem 9 af de 10 klynger.

Klyngedannelserne gør det muligt at udpege de portrætter, der ligger tættest på klyngecentrene, fx viser ill. 14 de mest centrale kvinde-mand-portrætter i klyngerne 0, 2 og 5.

Ligesom klynger kan hjælpe os med at identificere centrale datapunkter til videre analyse, kan metoden også anvendes til at illustrere, hvordan man kan forstå lignende portrætter ud fra den metrik for opstilling, som vi her benytter, nemlig afstand mellem øjenpar.

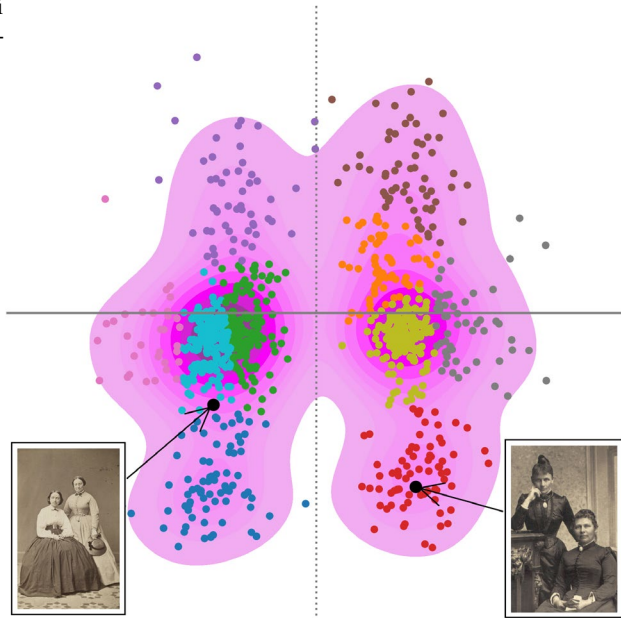
Bryllupsbilledtesen

De to eksempler fra Søndergaards artikel kan statistisk placeres i forhold til de klynger, som vi i det foregående har inddelt kvinde-mand- og kvinde-kvinde-portræterne i (se ill. 15). Søndergaards første eksempel med Marie Topsøe og Augusta Fenger falder i klynge 3, hvis man opfatter det som et kvinde-kvinde-portræt og ligeledes i klynge 3, hvis man opfatter det som et kvinde-mand-portræt, hvor den højeste kvinde (Marie Topsøe) indtager rollen som mand i vores klassifikation. Og tilsvarende for portrættet af Natalie Zahle og Pouline Govl, som tilordnes klynge 8 som kvinde-kvinde-portræt og klynge 9, hvis det fortolkes som kvinde-mand-portræt med den stående Zahle i rollen som mand.

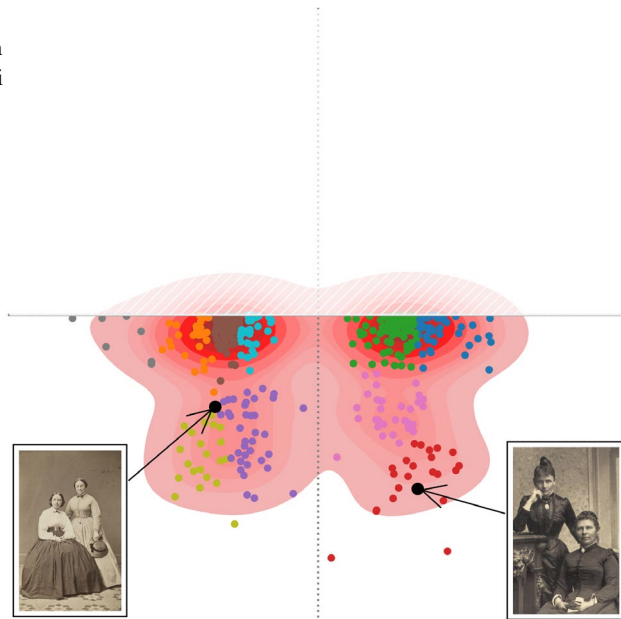
Selve klyngetilordningen siger naturligvis ingenting – den er ikke en fortolkning, men en statistisk udpegning af, hvilke andre datapunkter der ligner de to portrætter. Lad os nøjes med at analysere eksempel 1.

- Når vi anskuer eksempel 1 som det kvinde-kvinde-portræt, det jo er, ser vi, at det ligger ganske centralt i klyngen af røde punkter, som udgør klynge 3 i ill. 15a.
- Klynge 3 er i sig selv ikke så stor (se ill. 15b), så observationerne er belagt med en vis usikkerhed.

a) Eksemplerne ligger i klyngerne 3 og 9 i kvinde-kvinde-gruppen.



b) I kvinde-mand-gruppen er billederne placeret i klynge 3 og 8.



Ill. 16. Søndergaards to eksempler placeret i deres statistisk tilskrevne klynger.

- Hvis vi ønsker en mere kvalitativ tilgang, kan vi betragte de tre nærmeste naboer fra vores datasæt (se ill. 17). Disse viser umiddelbart opstillingsmæssige ligheder med eksempel 1 (hvilket jo er, hvad hele vores metrik handler om), men de viser også lidt af den variation, der stadig er inden for lignende opstillinger. I eksempel 1 står personerne uden overlap mellem kroppene, men med en hånd bag et hoved (sml. hhv. b og b, c, d). I et af portræterne (c) bærer kvinderne hat, og deres påklædninger ligner ikke lige eksempel 1. Disse oplysninger om forgrund/baggrund, positur og påklædning fanger vores metrik ikke, så der er brug for yderligere undersøgelser.

Konklusionen på denne fortolkning er, at eksempel 1 ikke er atypisk for en gruppe kvinde-kvinde-portrætter, selvom det heller ikke er den mest udbredte opstilling. Ved hjælp af vores simple metrik kan vi udtrække andre portrætter, som deler opstilling med eksempel 1, men som ved nærmere inspektion viser noget af den variation, som man kan undersøge videre ved klassiske eller kombinerede metoder.

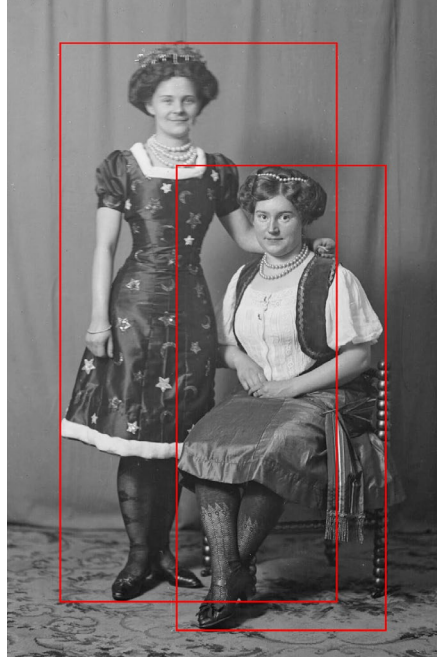
Hvis vi kontrafaktuelt anskuer eksempel 1 som et kvinde-mand-portræt, kan vi godt sige noget om, hvor meget det ligner andre kvinde-mand-portrætter i vores datasæt (igen mht. øjenafstand). Ligesom ovenfor viser illustrationen, at eksempel 1 placerer sig centralt i den tilordnede klynge 3 af kvinde-mand-portrætter. Opstillingen er altså ikke en 'outlier' i klyngen, men klynge 3 er en af de mindre klynger. Når vi betragter de nærmeste naboer, er det måske sværere end ovenfor at uddrage direkte paralleller og forskelle, men det falder i øjnene, at Marie Topsøe har sin højre hånd hævet til kinden. Sådan en positur ser vi ikke i de andre nærmeste naboer, men vi har heller ikke søgt at analysere positurer (endnu).

Kan man se det på fotografiet?

Vi har i det foregående vist, at vi med en kvantitativ tilgang til et forholdsvist omfattende corpus på ca. 26.000 billeder har fundet frem til 379 kvinde-kvinde-billeder (en manuel gennemgang viser, at datasættet med kvinde-kvinde-billeder kun indeholder få falske positiver i form af kvinde-barn-billeder). At finde frem til de op mod 400 billeder ville før i tiden have været en meget langsommelig proces. Blot det, at der findes en forholdsvis stor mængde billeder, gør, at vi kan analysere dem statistisk. Alt i alt giver de kvantitative resultater mulighed for at



a) Eksempel 1.



b) Elfelt, 1912.



c) Elfelt, 1900.

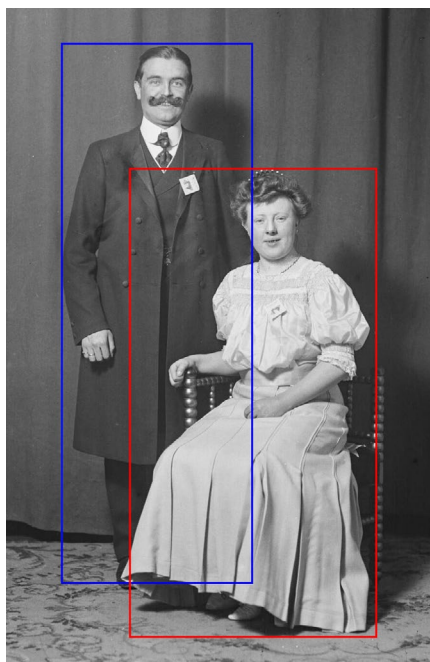


d) Elfelt, 1903.

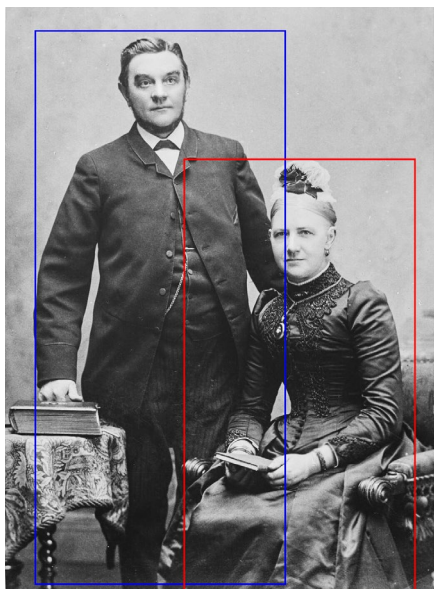
Ill. 17. Eksempel 1 og de tre billeder fra kvinde-kvinde-klynge 3, der ligger tættest på eksempel 1 ift. opstilling.



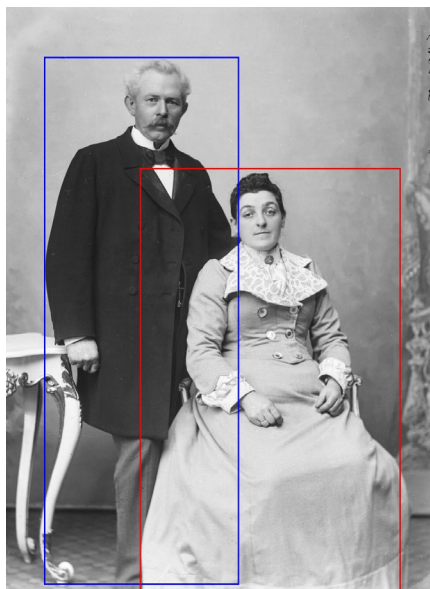
a) Eksempel 1.



b) Elfelt, 1908.



c) Elfelt, 1907.



d) Elfelt, 1901.

Ill. 18. Eksempel 1 og de tre billeder fra kvinde-mand-klynge 3, der ligger tættest på eksempel 1 ift. opstilling. Vi har set bort fra et 'outlier'-billede, som forestiller en dreng og en pige.

analysere forskellene mellem kvinde-mand-billederne og kvinde-kvindebillederne for på den måde at kunne sige noget mere om, hvordan personerne er placeret i billedet.

I forhold til spørgsmålet om, hvorvidt man ved at se på fotografierne af kvinde-kvinde og kvinde-mand og ikke blot på bryllupsbilleder vil kunne sige noget om kvindernes status som par, er resultatet negativt. Fotografier af mand-kvinde har en lighed med fotografierne af Govl og Zahle samt Topsøe og Fenger, men det samme gælder et stort antal kvinde-kvinde-fotografier, hvorom vi fra metadata til Elfelt-samlingen ved, at ca. en fjerdedel af vores søgesæt på 379 er mor-datter- og søster-søster-portrætter.

At så tvivl om Søndergaards påstand om lighed med bryllupsfotografier og parportrætter er imidlertid ikke det samme som at forkaste, at fotografierne har en betydning for bestemmelsen af relationen mellem kvinderne som mere end en veninderelation. Hvis man fortsætter med at bruge Søndergaards metode med at se *på* fotografiet, så vil man skulle finde et nyt visuelt spor at gå efter, noget, som kan binde fotografierne sammen.

De amerikanske samlere Neal Treadwell og Hugh Nini har i deres samling af fotografier af mænds kærlighed til mænd flere billeder af to mænd under en paraply – mange i solskin og i et tilfælde i en situation, hvor to mænd ‘vies’ under en paraply. Det er oplagt at se som et symbol, som kan vinde i styrke for hvert fotografi, der findes kvantitativt, og i det tilfælde vil en vintagedetektor anvendt på et virkelig stort kildemateriale være en gevinst. Den kvalitative analyse vil imidlertid også altid være nødvendig, for i sidste ende er der behov for andet kildemateriale, der vidner om en romantisk situation. Faktum er imidlertid, at det er en meget lille del af den fotografiske kultur, der består i at ‘signalere’ med symboler, og at langt de fleste fotografier er langt mere flertydige.

Det gælder også, når vi bevæger os til fotografier, hvor det er de afbildedes fysiske nærhed, der fortolkes, noget, der på sin vis er beslægtet med Søndergaards tese, fordi den er en del af bryllupstesen. Langt oftere er det mere kropslige tegn som omfavnelser eller kys, der fortolkes, og det er stærke indikatorer, men kræver også kritisk stillingtagen. Treadwell og Nini har en stor samling af pasfotos af to mænd, som sidder tæt og er fjollede, og som de har publiceret i bogen *Loving*.

*A Photographic History of Men in Love 1850s-1950s.*²¹ Her fortolker de en fysisk nærhed som et vidnesbyrd om mænds kærlighed til mænd, men netop her er der masser af grund til kritisk analyse, og her kommer vigtigheden af at kende til kontekst – ikke blot mændenes historie, men også fotografiets – til udtryk. Fotoautomaten er jo kendetegnet ved sin trange plads, der betinger, at man skal sidde tæt, og den bliver ofte brugt af venner, der er fulde, og som besegler deres venskab ved at tage et foto, så her må mænds kærlighed til mænd i hvert fald skulle ses i et bredere homoerotisk perspektiv, der kan omfatte venskaber.

Konteksten er nødvendig

Overordnet vil det sige, at det er en tvivlsom affære at fortolke relationer og specifikt bestemme romantiske forhold alene ud fra kropssprog og visuelle fænomener på fotografier. Det er imidlertid tilsvarende fristende at gøre det, særligt i en LGBTQ+-sammenhæng, hvor behovet for at genkende sig selv i fortiden eller finde ligesindede i nutiden er stort. Den amerikanske historiker Joan Scott bruger begrebet ‘fantasy echo’ for at beskrive historikerens drift efter at finde noget bestemt.²² Resonansen er et vilkår, men noget, der kræver en metodisk bevidsthed.²³ En indledende refleksion kan være, at fotografier af mennesker i forhold kan være helt uden romantik og ofte er det, mens fotografier af venskaber kan være meget hjertelige. Rent metodisk kan det beskrives i et falsifikatorisk perspektiv: At ét foto er et fotografi af et par, er ikke det samme som, at alle fotografier, der ser sådan ud, forestiller par. Omvendt kan et fotografi, der ikke ser sådan ud, godt forestille et par.

Fotografiet er med sin optisk-kemiske natur på én gang en kilde, der giver os en meget realistisk og samtidig meget fysisk oplevelse af fortiden, og det bliver hurtigt genstand for vores projektioner. Når vi

²¹ Neal Treadwell og Hugh Nini: *Loving. A Photographic History of Men in Love 1850s-1950s*, Milano, 2020.

²² Joan W. Scott: Fantasy Echo: History and the Construction of Identity. *Critical Enquiry* 27, no. 2, 2001, s. 284-304.

²³ På platforme som *Pinterest* og *Instagram* eksisterer der profiler med navne som “The Historical Homo” eller “Retro Lesbians” med samlinger af mange forskellige fotografier af to kvinder, der står eller sidder sammen. Nogle omfavner hinanden eller holder i hånd, nogle kysser hinanden eller ser på hinanden med et forelsket blik. På disse sites defineres de alle som fotografier af homoseksuelle med tekst, i ovenstående tilfælde profilens navn.

anvender fundne fotografier, vil der altid være et usikkerhedsrum, som kan bruges legende som i Lützens tilfælde, men også betrædes mere metodisk bevidst. Søndergaard er rodfæstet i historievidenskaben og anvender i Topsøe og Fengers tilfælde en forsigtig sprogbrug om fotografiet. Hun skriver, at det ikke er “ulig datidens bryllupsfotografier”, så udtalelsen er behæftet med forbehold. Som historiker baserer hun sin artikel på andre kilder som breve og erindringer, men også registreringer af boligforhold, økonomi osv. I den forbindelse kan man sige, at selve det, at kvinderne har ladet sig fotografere sammen, peger på en nær relation, selv om fotografiet altså ikke har en lighed med bryllupsfotografiet. Argumentet for, at der er tale om et kærlighedsforhold, kan yderligere styrkes af, at fotograferingen er sket over en årrække, som tilfældet var med lærerne Theodora Lang og Anna Höltzermann, der har gemt parbilleder fra atelierer i hele norden, hvilket formodentlig ikke er sket i mange venindeforhold.²⁴

Konklusion

Vi kan konkludere, at det at kigge på historisk billedmateriale er udfordrende på en anden måde end at se på moderne digitalt født billedmateriale og på tekstkilder. I digitale arkiver vil det være meget forskelligt, hvor meget og hvilken type metadata billeder er forsynet med, og om det giver én mulighed for at finde noget, som er relevant for det, man undersøger. Arbejdet med at lokalisere kildemateriale er en tidskrævende sorteringsproces og kræver mange arbejdstimer, hvis man vil have et godt kildegrundlag. Hvis man på den anden side vælger at gå mere kvantitativt til værks, kræver det, at man præcist definerer, hvad undersøgelsens formål er, fordi den anvendte model skal trænes til netop det. Til gengæld kan der være uventede sidegevinster, som kan bruges i andre projekter som identifikation af køn og dobbeltportrætter, viden om, hvordan par står/sidder m.m.

Med vintagedektoren får man nye muligheder: Hvis dektoren udvikles, vil den kunne håndtere fotografier tilbage til mediets opfindelse og kunne datere dem med en vis præcision på basis af genstandes udseende gennem tiderne. Kan vintagedektoren definere forskellige typer påklædning, kan det være en kilde til tidsfæstelse; omvendt kan

²⁴ Se Rikke Andreassen og Merete Pryds Helle: *De nye kvinder. Kærlighed og queers 1870-1920*, 2025.

et billede, der er tidsfæstet, bruges til at definere bestemte beklædningsgenstande.

Med AI får vi i stigende grad mulighed for at efterprøve påstande, som ikke hviler på en omfattende empirisk undersøgelse, og derfor også for at give et statistisk bud på, hvad der gemmer sig ude i 'the great unseen'. Begrebet er afledt af udtrykket 'the great unread', som refererer til den enorme mængde af litteratur og tekster, der er blevet produceret gennem historien, men som kun bliver læst af få. Med AI bliver det nemmere og finde information og lave strukturelle analyser af de store ulæste datamængder. Gennem mange år er analyser af bogdata blev brugt til at finde spor i 'the great unread', mens det er svært at anvende automatisk billedgenkendelse i arbejdet med historiske fotografier. At træne modeller kræver store mængder fotografier og en mulighed for at tilgå og processere dem, så udviklingen er også afhængig af digitalisering og tilgængeliggørelse. Vi håber at have vist, at arbejdet kræver tværfaglighed, fordi forskningsspørgsmål kræver viden om samlinger og metadata, men også en bevidsthed om, hvordan modeller skal trænes, og at udviklingen nødvendigvis er en kontinuerlig proces, hvor spørgsmål, kildemateriale og teknologi er gensidigt afhængige.

SUMMARY

Artificial intelligence offers promising new possibilities for analyzing cultural heritage collections, yet applying automated visual recognition to historical photographs presents significant methodological challenges. This article examines these challenges, both in object detection and in addressing more complex research questions related to the interpretation of visual phenomena. Using a case study based on historian Nina Søndergaard's thesis – that certain photographs of two women from around 1900 resemble wedding photographs and may thus indicate romantic relationships – we demonstrate how Søndergaard's hypothesis can be tested empirically through the application of AI, contributing to methodology within LGBTQ+ visual culture studies and visual culture more broadly.

Our corpus comprises a large collection of dated photographs (1890-1920), primarily drawn from the Royal Danish Library's Digital Collections. We applied systematic pattern identification using AI models such as Florence-2 and RetinaNet to examine body positioning. The process, however, was not a straightforward application of a tool; it involved a dynamic interplay between manual selection and automated analysis. Furthermore, we created subsets of wedding photographs based on metadata and employed both manual and automated methods to investigate how individuals positioned themselves – whether standing or seated – in wedding photographs and in paired portraits more generally. Our analysis revealed minimal visual resemblance between Søndergaard's examples and documented wedding photography conventions, highlighting the wide variety of poses employed in early twentieth-century portraiture.

These findings underscore the challenges of interpreting visual traits as direct indicators of specific emotions or life forms – an approach often adopted in cultural studies, driven by the desire to uncover what has been historically overlooked or underrepresented. Importantly, our study demonstrates how AI tools, developed through interdisciplinary collaboration between historical methodology and machine learning, can test subjective hypotheses and provide valuable analytical resources for visual culture research.

