

Summary – Struggling with Algorithmic Seeing (Gabriel Pereira)

ENGLISH

This thesis is concerned with the burgeoning forms of machine seeing represented by Computer Vision: an umbrella term for contemporary applications of Machine Learning and AI that carry out different forms of image recognition/analysis and pattern extraction to create actionable intelligence from images and videos (e.g., summarization, categorization, moderation). My motivation in this study is investigating and critiquing how Computer Vision operates, what kind of power structures it enables, and how it may become infrastructural in our everyday lives. I aim to do this in order to explore how it is in fact not a neutral operation, but one full of uncertainties, contentions, and politics which allow us to imagine potential alternatives. Differently than current discussions around the possibility of an ethical AI or of “solving bias,” I side with scholars who argue for embracing the complexity of such systems to understand their intrinsic forms of biases and limitations. This means that, rather than believing in Computer Vision’s potential to be “fair” and “transparent,” I foreground its politics and “conditions of emergence,” how they come to exist and work in the ways they do. I thus situate this thesis within the emerging interdisciplinary field of critical algorithm studies, which suggests to study algorithms not only as technical issues, but also as social, political, and economic.

I contribute to critical algorithm studies with an exploration of Computer Vision through methodologies from qualitative research, practice-based research, and cultural critique. As such, this thesis engages with the infrastructures of Computer Vision, its situated uses and practices, as well as the narratives and imaginaries that support its operations. As an article-based thesis, it is composed of two complementary halves: *Kappe* and *Practices*.

The *Kappe* serves as an overview of the thesis, outlining the theoretical and methodological framework and contributions, as well as drawing the conceptual contributions from the collection of articles in *Practices*. In this sense, I ask two sets of research questions, summarized as: 1) How does Computer Vision get constructed as hegemonic? 2) How may we develop counter-hegemonic practices towards hegemonic CV? These research questions evolved out of my trajectory as a researcher, as I made sense of my experience studying Computer Vision through the *Practices*.

Therefore, the first aim of the *Kappe* is to address the dominant paradigm of algorithmic seeing, which I term “hegemonic CV.” I develop this conceptual notion by bridging a critical theoretical framework around “hegemony” and “common sense” with the empirical and practical exploration of Computer Vision that I have conducted. The framework builds upon the classic definitions offered by Italian theorist Antonio Gramsci, while also benefitting from the diverse scholarship that has been inspired or developed from it (for example Raymond Williams, Stuart Hall, Laclau & Mouffe, Paulo Freire, as well as contemporary scholars of technology and society). Hegemonic CV allows for understanding the struggle behind algorithmic seeing and the dynamics around its dominant development. Different facets of hegemonic CV are discussed by drawing on the empirical work from the *Practices*, particularly how Computer Vision is persuasively promoted as inevitable, efficient, and neutral, while becoming an infrastructural and imperceptible part of everyday life. The notion of hegemony is also useful for understanding the surveillance dynamics that computer vision enables (i.e. forms of commercial, governmental, and military surveillance). Moreover, it allows to engage with Computer Vision’s “colonizing impulse”: how it implies changes in how visibilities are constructed in society more broadly.

The second aim of the *Kappe* is to engage with contestations over hegemonic CV, demonstrating the possibility and importance of antagonistic practices which refuse hegemonic formations, particularly

through interventionist practice-based research drawing from art and activism. Identifying and discussing hegemonic CV thus serves not only as a critique, but also to offer the foundations for speculating on what algorithmic seeing could otherwise be. Rather than heeding to the dominant narrative that poses Computer Vision as an inevitable reality, this thesis explores what forms of antagonistic practice can question, subvert, and ultimately suggest alternatives. This antagonistic practice, which I operationalize through an artistic “critical technical practice,” means not only exposing the hegemonic narratives and operations of CV, but also creating cracks for thinking and doing otherwise through “refusing.”

The second part of the thesis, *Practices*, contains five research articles and three practice-based outputs, each with their own approaches and contributions. The first article (co-authored with Christoph Raetzsch) focuses on Automated License Plate Recognition (ALPR) cameras, systems that automatically detect and transform license plates of passing cars into data and mechanisms of control. The article particularly identifies the way Computer Vision is constructed as a banal form of surveillance. The second article concerns the automated memory-making created by the Apple Memories application. Building upon Mackenzie & Munster’s notion of “platform seeing,” my argument is that infrastructures of Computer Vision are important means for the consolidation of platform and infrastructure power. The third article (co-authored with Bruno Moreschi) focuses on a practice-based experiment, using Computer Vision from Big Tech companies to read artworks in the Van Abbemuseum (Netherlands). Both the article and the short film “Recoding Art” focus on the “glitches” in the way these algorithms saw the artworks, defamiliarizing these “ways of machine seeing” and reverse-engineering their ways of understanding art, while also using it as a way to see differently. The fourth article (co-authored with Sarah Schorr and Carlos de Oliveira) presents a practice-based research on the algorithmic analysis of color, which resulted in the net art “Algorithmic Sea.” The interactive essay both defamiliarizes the computational operation of color and imagines what would be specu-

lative ways algorithms could see color that are more fluid, focusing on ambiguity, movement, and self-determination. The fifth article is directly engaged with exploring radical forms of antagonistic practice to Computer Vision. The article sketches a framework around “refusing,” framing this form of response as a counter-hegemonic practice – be it the refusal to work or other, more creative, responses. Refusal is also the theme of the short film “Future Movement Future – REJECTED” (co-directed with Bruno Moreschi), which speculates on how Computer Vision could be regulated (and its consequences).

As I conclude in the *Kappe*, the main benefit of centering the struggles of (counter-)hegemony is, thus, how Computer Vision is made not into a totalitarian form of power. There are many levers through which we can move it, many ways our antagonisms can take place, many fissures in each layer of its sedimented ways of seeing. Recognizing the dominant frame, instead of closing off alternative possibilities, serves to highlight what is being left out, what visibilities are being ignored, and who could benefit from alternative formations. It allows for a more complex analysis at a moment when algorithmic systems tend to be framed as merely a matter of technical issues, of increasing diversity in data sets, creating ethical guideline checkboxes, or even anxiety-filled discussions about surveillance capitalism. Rather than heeding to reformist calls which see algorithms of Computer Vision as mere technical “black boxes,” it allows us to understand them in practice as a contested and fallible system which we can (and should) question.

Overall, the contributions of this thesis to the emergent field of critical algorithm studies can be summarized as: 1) The concept of hegemonic CV, and how it allows for a more complex understanding of Computer Vision and its politics through the bridging between critical theory and empirical/practice-based research; 2) A conceptualization of a counter-hegemonic antagonistic practice centered on building contention through “refusing”; 3) An analytically rich analysis of Computer Vision in practice, as offered by the five articles and three practice-based outputs in *Practices*, which engage with diverse objects

and arenas; 4) A methodological contribution, as this thesis shows how research on algorithmic seeing can act interdisciplinarily, bridging social research and practice-based methodologies.

DANSK

Denne afhandling omhandler de spirende former for *machine seeing* repræsenteret ved *Computer Vision*: en fællebetegnelse for nutidige applikationer af Machine Learning og AI, der udfører forskellige former for billedgenkendelse/analyse og mønstrekstraktion for at skabe praktisk intelligens fra billeder og videoer (f.eks. opsummering, kategorisering, mådehold). Min motivation i denne undersøgelse viser, hvordan Computer Vision fungerer - hvordan den faktisk ikke er neutral, men fuld af usikkerheder, stridigheder, politik osv., samtidig med at den genopfinder den ved at engagere sig med potentielle alternativer. I modsætning til aktuelle diskussioner omkring muligheden for “ethical AI” eller “fixing bias” tager jeg parti med de forskere, der argumenterer for at omfavne kompleksiteten af sådanne systemer for at forstå deres iboende former for skævheder og begrænsninger. Det betyder, at jeg i stedet for at tro på Computer Visions potentiale til at være “fair” og “gennemsigtig” fremhæver politikken og “betingelserne for fremkomst”, samt hvordan de opstår og fungerer på de måder, de gør. Jeg placerer således denne tese inden for det nye tværfaglige felt af *critical algorithm studies*, som foreslår ikke kun at studere algoritmer som tekniske anliggender, men også som sociale, politiske og økonomiske.

Jeg bidrager til *critical algorithm studies* med en undersøgelse af Computer Vision gennem metoder fra kvalitativ forskning, praksisbaseret forskning og kulturkritik. Som sådan beskæftiger dette speciale sig med infrastrukturen i Computer Vision, dets situerede anvendelser og praksisser samt fortællinger og forestillinger, der understøtter dets funktioner. Som en artikelbaseret afhandling består den af to komplementære halvdele: *Kappe* og *Practices* (Praksisser).

Kappen fungerer som en oversigt over specialet, der skitserer de teoretiske og metodiske rammer og bidrag, samt trækker på de konceptuelle bidrag fra samlingen af artikler i *Practices*. I denne henseende stiller jeg to sæt af forskningsspørgsmål, opsummeret som: 1) Hvordan bliver Computer Vision konstrueret som hegemonisk? 2) Hvordan kan vi udvikle kontrahegemoniske metoder til hegemonisk CV? Disse forskningsspørgsmål udviklede sig under min rejse som forsker i takt med, at jeg forstod meningen med de erfaringer jeg gjorde mig ved at studere Computer Vision gennem praksis.

Derfor er det første mål med *Kappen* at adressere det dominerende paradigme for *algorithmic seeing*, som jeg kalder “hegemonic CV.” Jeg udvikler denne konceptuelle ide ved at forbinde en kritisk teoretisk ramme om “hegemoni” og “common sense” med den empiriske og praktiske undersøgelse af Computer Vision, som jeg har udført. Rammerne bygger på de klassiske definitioner, der tilbydes af den italienske teoretiker Antonio Gramsci, samtidig med at de drager fordel af den mangfoldige lærdom, der er blevet inspireret eller udviklet heraf (for eksempel Raymond Williams, Stuart Hall, Laclau & Mouffe, Paulo Freire samt nutidige forskere i teknologi og samfund). Hegemonic CV giver mulighed for at forstå kampen bag *algorithmic seeing* og dynamikkerne omkring dets dominerende udvikling. Forskellige facetter af hegemonic CV diskuteres ved at trække på det empiriske arbejde fra *Practices*, især hvordan Computer Vision overbevisende fremmes som uundgåeligt, effektivt og neutralt, samtidig med at det bliver en infrastruktur og umærkelig del af hverdagen. Begrebet hegemoni er også nyttigt for at forstå den overvågningsdynamik, som Computer Vision muliggør (dvs. til kommerciel, offentlig og militær overvågning). Desuden giver det mulighed for at interagere med Computer Visions “koloniserende impuls”: hvordan det indebærer ændringer i, hvordan synligheder er konstrueret i samfundet mere bredt.

Det andet mål med *Kappen* er at engagere sig i stridigheder om hegemonic CV samt demonstrere muligheden og betydningen af “antagonistic practices”, der nægter hegemoniske formationer, især gennem

interventionistisk, praksisbaseret forskning, der trækker på kunst og aktivisme. Identifikation og diskussion af hegemonisk CV fungerer således ikke kun som en kritik, men også som et grundlag for at spekulere i, hvad algorithmic seeing ellers kunne være. I stedet for at tage hensyn til den dominerende fortælling, der udgør Computer Vision som en uundgåelig virkelighed, undersøger denne afhandling, hvilke former for antagonistisk praksis, der kan sætte spørgsmålstejn ved, undergrave og i sidste ende foreslå alternativer. Denne antagonistiske praksis, som jeg operationaliserer gennem en kunstnerisk “critical technical practice”, betyder ikke kun at afsløre de hegemoniske fortællinger og funktioner i CV, men også at skabe revner for at tænke og gøre andet ved “at nægte” (“*refusing*”).

Den anden del af afhandlingen, *Practices*, indeholder fem forskningsartikler og tre praksisbaserede resultater, der hver især har deres egne tilgange og bidrag. Den første artikel (medforfattet med Christoph Raetzsch) fokuserer på ALPR-kameraer (Automatisk nummerpladegenkendelse), som er systemer, der automatisk registrerer og omdanner nummerplader på passerende biler til data og kontrolmekanismer. Artiklen identificerer især den måde, hvorpå Computer Vision er konstrueret som en banal form for overvågning. Den anden artikel vedrører den automatisk hukommelsesfremstilling, der er skabt af Apple Memories-applikationen. Med udgangspunkt i Mackenzie & Munsters forestilling om “platform seeing” er mit argument, at infrastrukturer i Computer Vision er vigtige midler til konsolidering af platforms- og infrastruktursmagt. Den tredje artikel (medforfattet med Bruno Moreschi) fokuserer på et praksisbaseret eksperiment ved hjælp af Computer Vision fra Big Tech-virksomheder til at læse kunstværker i Van Abbemuseum (Holland). Både artiklen og kortfilmen “Recoding Art” fokuserer på “fejlene” i den måde, disse algoritmer så kunstværkerne på ved at fremmedgøre disse “former for machine seeing” og dekompile deres måder at forstå kunst på, mens de også bruger det som en måde at se anderledes på. Den fjerde artikel (medforfattet med Sarah Schorr og Carlos de Oliveira) præsenterer praksisbaseret forskning af den algoritmiske analyse af farve, hvilket re-

sulterede i netkunstværket “Algorithmic Sea”. Det interaktive essay fremmedgør både beregningsoperationen af farve og forestiller sig, hvilke spekulative måder, hvorpå algoritmer kunne se farve, der er mere flydende med fokus på tvetydighed, bevægelse og selvbestemmelse. Den femte artikel beskæftiger sig direkte med at undersøge radikale former for antagonistic practice til Computer Vision. Artiklen skitserer en ramme omkring “at nægte”, der indrammer denne form for svar som en kontrahegemonisk praksis – det være sig at nægte at arbejde eller andre mere kreative svar. “At nægte” er også temaet for kortfilmen “Future Movement Future – REJECTED” (medinstrueret af Bruno Moreschi), der spekulerer i, hvordan Computer Vision (og dens konsekvenser) kan reguleres.

Som jeg konkluderer i *Kappen*, er den største fordel ved at centrere stridighederne ved (kontra-)hegemoni således, hvordan Computer Vision ikke bliver til en totalitær magtform. Der er mange måder som vi kan bevæge os igennem, mange måder vores modsætninger kan finde sted på, mange sprækker i hvert lag af dets sedimenterede måder at se på. Anerkendelse af den dominerende ramme i stedet for at udelukke alternative muligheder tjener til at fremhæve, hvad der bliver udeladt, hvilken synlighed der ignoreres, og hvem der kunne have gavn af alternative formationer. Det giver mulighed for en mere kompleks analyse på et tidspunkt, hvor algoritmiske systemer har en tendens til kun at blive formuleret som spørgsmål om tekniske problemstillinger, om stigende mangfoldighed i datasæt, oprettelse af etiske retningslinje-afkrydsningsfelter eller endda angstfyldte diskussioner om overvågningskapitalisme. I stedet for at tage hensyn til reformistiske kald, der blot ser Computer Vision algoritmer som tekniske “sorte bokse”, giver det os mulighed for at forstå dem i praksis som et anfægtet og fejlbarligt system, som vi kan (og bør) stille spørgsmålstejn ved.

Samlet set kan denne afhandlings bidrag til det nye felt af critical algorithm studies opsummeres som: 1) Begrebet hegemonisk CV, og hvordan det giver mulighed for en mere kompleks forståelse af Com-

puter Vision og dets politik gennem forbindelsen mellem kritisk teori og empirisk/praksisbaseret forskning; 2) En konceptualisering af en kontrahegemonisk antagonistic practice, der er centreret om at opbygge stridigheder gennem "at nægte"; 3) En analytisk rig analyse af Computer Vision i praksis, som fremgår af de fem artikler og tre praksisbaserede resultater i *Practices*, der engagerer sig i forskellige objekter og arenaer; 4) Et metodologisk bidrag, da denne afhandling viser, hvordan forskning om algorithmic seeing kan agere tværfagligt samt bygge bro mellem social forskning og praksisbaserede metoder.