In the last decade, machine vision has become part of the everyday life of ordinary people. Smartphones have advanced image manipulation capabilities, social media use image recognition algorithms to sort and filter visual content, and games, narratives and art increasingly represent and use machine vision techniques such as facial recognition algorithms, eye-tracking and virtual reality.

The ubiquity of machine vision in ordinary peoples’ lives marks a qualitative shift where once theoretical questions are now immediately relevant to the lived experience of ordinary people.

MACHINE VISION will develop a theory of how everyday machine vision affects the way ordinary people understand themselves and their world through 1) analyses of digital art, games and narratives that use machine vision as theme or interface, and 2) ethnographic studies of users of consumer-grade machine vision apps in social media and personal communication. Three main research questions address 1) new kinds of agency and subjectivity; 2) visual data as malleable; 3) values and biases.

MACHINE VISION fills a research gap on the cultural, aesthetic and ethical effects of machine vision. Current research on machine vision is skewed, with extensive computer science research and rapid development and adaptation of new technologies. Cultural research primarily focuses on systemic issues (e.g. surveillance) and professional use (e.g. scientific imaging). Aesthetic theories (e.g. in cinema theory) are valuable but mostly address 20th century technologies. Analyses of current technologies are fragmented and lack a cohesive theory or model.

MACHINE VISION challenges existing research and develops new empirical analyses and a cohesive theory of everyday machine vision. This project is a needed leap in visual aesthetic research. MACHINE VISION will also impact technical R&D on machine vision, enabling the design of technologies that are ethical, just and democratic.
Section a: Extended Synopsis of the scientific proposal
Snapchat’s selfie lenses are one of many examples of everyday machine vision. Snapchat was launched in 2011, and selfie lenses were introduced less than two years ago, but have become popular with extreme speed: in all probability, more than 10% of all Norwegians play with selfie lenses every single day. Selfie lenses are fun. They also normalise biometrics, a form of machine vision that is otherwise used to identify and control individuals (Rettberg 2017). When you launch the Snapchat app on your phone, it opens to the camera, showing you the image of your own face. If you hold down a finger on the image of your face, a biometric grid is superimposed on your face, referencing a visual rhetoric more commonly used in representations of identification, surveillance and control. Icons at the bottom of the screen show the different lenses you can apply, and by swiping through them, you can instantly see your face altered in a dozen ways or more. The facial recognition algorithms underlying Snapchat’s selfie lens technology allow the distortions and masks to perfectly align with the user’s own eyes, nose and mouth, so you can move with the lens, seeing for a moment what it would be like to have rabbit ears and whiskers, or divinely smooth skin, large eyes and a flower crown. Pass a phone to a child or adult and let them play with the selfie lenses: they will almost always laugh with delight.

MACHINE VISION investigates how ordinary people interact with machine vision (the registration, analysis and representation of visual information by machines and algorithms) in their everyday lives. While existing research has tended to emphasise system-level use of machine vision, for instance in scientific imaging or surveillance, I will instead examine its domestication and the playful nature of individuals’ interaction with machine vision. MACHINE VISION is founded upon the assumption that art and aesthetic expression deeply engage with changes in perceptual technologies, and that the study of art, narratives and games will be a productive source for insights about our relationships with machine vision, and so my main empirical data will be digital art, games and narratives that thematise and simulate machine vision, using digital methods and network analysis to gain an overview of a corpus, followed by close readings of a selection of works. This will be complemented by ethnographic fieldwork studying how ordinary people use machine vision apps and technologies for self-representation and communication in social media, and by a thorough survey of the history and theories of pre-digital machine vision. My overall objective is to develop a theory that explains how the algorithmic machine vision of the 21st century affects the way ordinary people see themselves and understand the world. Further objectives are specified under the individual research tasks on page 4-5.

Ground-breaking Nature of the Research: Visual technologies have undergone radical changes in the last decades and are still changing rapidly, creating a pressing need for new theories that build upon and go well beyond earlier work to understand a world where machine vision is not only ubiquitous, but where the majority of the population have access to machine vision as creators and publishers, not just consumers. Although there is a great deal of technological research and development on machine vision and computer vision in computer science, there is a large research gap on how this cultural shift to seeing with and through technology affects us as a society and as individuals. My research is unique in its focus on the personal use of machine vision, and in my three research questions I focus on shifts in agency, on the perceived objectivity of images, and on the values embedded in machine vision. This is high-risk research, because of its cross-disciplinary methodology (combining aesthetic analysis with digital methods and ethnographic research), novel object of study and innovative combination of theory and empirical analysis, but it is also extremely high-gain. 10 years from now, humanities scholars will use MACHINE VISION’s findings to analyse digital visual technologies with the precision we now only can apply to older media, and computer scientists will use its insights to develop technology that supports a democratic, ethical and just society.

1 A survey in late 2016 showed that 36% of Norwegian residents over 18 use Snapchat every day (Ipsos 2016), and according to Snap Inc., over a third of Snapchat’s daily users play with selfie lenses every day (Snap Inc. 2017).
Research Questions: I have three main research questions, addressing agency, objectivity and values:

RQ1) Which kinds of agency develop and which are limited in the interaction between individual users and machine vision? I hypothesise that posthumanist theories of distributed agency/subjectivity between humans and machines (Hayles 1999; Braidotti 2013; Flanagan 2014; Nayar 2014) can be applied to ordinary peoples’ interactions with machine vision, but this has not yet been tested. Playing with selfie lenses or a digital camera app with filters and photo editing software gives ordinary users a level of control over the way that they are represented that was previously only available to expert users and artists, and yet there are many ways in which their creativity and agency are constrained. Games, art works and narratives provide particularly rich material to understand constraints upon agency, with their richly imagined dystopias but also through player/viewer/user interfaces that simulate machine vision. Ethnographic research will provide interpretations of user’s experiences of their interactions with and through machine vision. In addition to posthumanist theories, I will also consider other approaches to agency, such as Actor-Network Theory (Latour 2007; Law and Hassard 1999), postphenomenology (Ihde 1990; Rosenberger and Verbeek 2015) and the photography theories of Flusser (2000), Virilio (1994) and others.

RQ2) Does the experience of being able to manipulate the visual as data through everyday interaction with machine vision lead us to see the world and ourselves as malleable? Many scholars have argued that seeing the world as data leads us to believe in the objectivity or truth status of data above our own perception. José van Dijck calls this “almost ideological belief in data” dataism (Dijck 2014). A high-risk element of this project is that I will explore an alternative to the currently dominant theory of dataism: my hypothesis is that the malleability and alternative visualisation possibilities of visual representations support a post-truth society where we do not trust visual or ultimately, perhaps, our own perceptions, which are increasingly enmeshed with machine vision. The idea of vision as a more reliable sense than others (“seeing is believing”) is connected to our centuries-long assumption that vision is passive. This relates to long-standing discussions on the status of representation in the philosophy of science (Hacking 1983) as well as in influential writings on art and culture (Debord 1967; Baudrillard 1981). What has changed since their work is not simply the immense increase in visual display and visual communication, but more importantly that machine vision is part of the visual vernacular today. We all create and manipulate images using machines, and the experience of altering or choosing between equally valid visual representations (for instance by considering which filter to apply to an image before sharing it) is becoming common for ordinary people and not only experts. Machine vision is no longer left to specialists, and this means ordinary people have an embodied knowledge that makes today’s experience of the malleability of reality qualitatively very different from postmodern theoretical understandings. If so, this has great impact on our understanding of current discussions about post-factual and post-truth politics, where images frequently are used to support claims.

RQ3) Which values are embedded in machine vision, and which biases are introduced or supported? Racial bias and biases against other groups are often coded into technology, intentionally or unintentionally (Winner 1980). For instance, 20th century analogue film was calibrated to show nuances in light-skinned but not dark-skinned faces (Roth 2009), and image recognition algorithms have misrecognised African faces as “gorillas”. My hypothesis is that everyday machine vision privileges certain kinds of subject. This research question therefore asks how machine vision shapes its users: What types of bodies or behaviours are supported or constrained by different kinds of personal machine vision? What kind of role does a work of digital art or a computer game require you to perform? Do different groups of people (e.g. refugees vs university students, women vs men, young vs old, Asian vs European, Global North vs Global South) experience personal machine vision in the same way? To answer these questions, the project will draw upon research in areas such as critical algorithm studies (Gillespie and Seaver 2015), feminist and queer theory, disability studies and post-colonial theory, as well as applying narratology and ludology to understand the position of the user in interacting with relevant apps, games, narratives and artworks.

State of the Art
My interest in machine vision began with my work on selfies and self-tracking for my book Seeing Ourselves Through Technology: How We Use Selfies, Blogs and Wearable Devices to See and Shape Ourselves (2014b). At first I saw selfies primarily as a parallel to the self-representations and self-narrations I had studied previously in blogs and electronic literature (2014a; 2017b; 2009; Walker 2005a), but as I analysed selfie apps and “beautifying” filters, as well as the sharing and liking ecosystems selfies are embedded in, I realised that the algorithmic manipulation of images that is inherent to digital technology is deeply entwined with our self-understanding. Building upon this work, I wrote a chapter on “Biometric Citizens” that further examines the connections between selfies and machine vision (Rettberg 2017a), and the way these are embedded within large-scale systems of power and surveillance. I also developed my interest in artistic explorations of the topic through two conference papers discussing art works that...
interrogate machine vision (Rettberg 2016a; Rettberg 2016b), building upon my earlier work on digital art (Walker 2003), games (Corneliussen and Rettberg 2008), and electronic literature (Walker 2004; 2005a).

Digital cameras are one of the key machine vision technologies ordinary users engage with, and the history and theories of photography and cinema provide valuable perspectives on my three research questions. Algorithmic images and machine vision are only the latest of a succession of visual technologies ranging from the mathematical models of Alberti’s perspective to the camera obscura, magic lanterns, kaleidoscopes, photography, stereoscopes, cinema, video, x-rays, magnetoscopes, virtual reality headsets, drones, airport security systems, home surveillance systems and satellite mapping. Artists and philosophers have long argued that the way we see the world, the way we are able to see the world, is deeply linked to the technologies and techniques we use to see it. “Imagine an eye unruled by man-made laws of perspective, an eye unprejudiced by compositional logic,” film-maker Stan Brakhage wrote in 1962. We have heard these ways of seeing since birth. We base our self-understanding on the media technologies of our time, as Kittler noted when he pointed out that Plato described the human soul as a wax tablet, while we after the introduction of cinema imagine that just before death, we see our lives play before our eyes like a film (Kittler 2009, 49). It is likewise well established that major shifts in representational technologies have profound effects on human culture. The introduction of linear perspective in the 15th century changed our relationship to religion and made enlightenment technologies possible (Edgerton 2009), and the introduction of print changed the way we share and create knowledge (Eisenstein 1979).

Histories of visual technology that focus on art, cinema and photography generally tell the story as leading up to a specific technology, such as perspective (Kemp 1992; Edgerton 2009), biometrics (Gates 2011) or digital color (Kane 2014), thus leaving out points that are key for machine vision in a broader sense, or, they stop short too early, as is the case with Kittler’s Optical Media, written in 1999 (2009). Other relevant research includes theoretical work on the visual (Vertov 1984 orig. 1923; Virilio 1994; Flusser 2000; Friedberg 2006), on new media studies and digital culture (Manovich 1993; Kane 2014; Uricchio 2011; Losh 2015), on research on scientific imaging and visualisation within science and technology studies (STS) and media aesthetics (Coopmans et al. 2014; Carusi and Hoel 2014; Hausken 2007; Hausken 2015) as well as research on surveillance images and biometrics (Introna and Wood 2004; Gates 2011; Greene 2015). I will explore contemporary philosophical approaches to machine vision (e.g. Agamben 2011), and will draw upon research on the interaction between visual technologies and literature (Danius 2002). Research on algorithmic bias and the inscription of certain expectations for identity in technologies (Roth 2009; Chun 2011; Gray 2014) will also be important. Within internet studies and digital anthropology there is a growing interest in selfies, frequently discussed in terms of performance and identity (Burns 2015; Tiidenberg 2015; Tiidenberg 2014; Seko and Lewis 2016).

What is lacking is a comprehensive theory that combines the aesthetic insights of visual theorists with contemporary cultural expression as it thematises and simulates machine vision, and with empirical findings of how actual users experience everyday machine vision.

**Methodology**

The question of how everyday machine vision affects us is too complex to be answered through incremental efforts, and requires a large-scale, ambitious and integrative research framework. The research will be organised in five research tasks, each of which will provide a different kind of answer to the three main research questions. Task 1 provides the necessary background for the other tasks. Task 2 and 3 analyse art, games and narratives, with Task 2 developing the corpus that Task 3 is dependent on. Task 4 runs in parallel to 3, providing the ethnographic study of ordinary users’ playful interactions with machine vision. Task 5 synthesizes findings from the other tasks and develops a comprehensive theory of everyday machine vision.

**Research Task 1: Survey of History and Theories of Machine Vision**

**Objective:** Compare and synthesise existing histories and theories of machine vision, especially within scholarship on visual art and visual communication, but also considering theories from other disciplines, such as society and technology studies, surveillance studies, internet studies and computer science.

**Short Description:** This task will primarily be conducted by me, with input from the PhD candidates, who will devote a chapter of their dissertations to writing about the history and theories of machine vision within their field (art, games or narrative). The emphasis will be on examining how existing historical accounts and theories discuss issues that are key to my research questions: agency (RQ1), objectivity (RQ2) and values (RQ3).

**Research Task 2: Distant Reading of Art, Narratives and Games that Thematisate Machine Vision**

**Objective:** Gain an overview of how machine vision is represented in digital art, games and narratives, and of how it is used as interface. Use network analysis to see whether there are correlations between agency (RQ1), objectivity (RQ2) and values (RQ3) and particular genres, modes of interaction etc.
**Research Task 3: Close Readings of Selected Art, Games and Narratives**

**Objective:** Provide analyses of how agency (RQ1), objectivity (RQ2) and values (RQ3) come into play in a selection of digital art, games and narratives that thematise or simulate machine vision.

**Short Description:** Selected works from the corpus gathered in Research Task 2, will be analysed, examining the way that they represent machine vision in order to understand how they position the viewer, player or reader in relation to the machine vision. Each PhD project will look at machine vision as theme and as interface in the genre studied. For instance, the movie *Minority Report* (2002) addresses machine vision as a theme, showing consumers constantly identified and addressed by commercial software advertisements. In a dramatic scene, the protagonist has his retina removed so as to avoid biometric identification. In games like *WATCH_DOGS* (Ubisoft 2014) and *Republique* (Camoflaj 2013-16) players see and interact with the gameworld through surveillance cameras, while *Pokemon Go* (2016) uses augmented reality. Danny Cannizzaro & Samantha Gorman’s *PRY* (2014), a visual and narrative work of electronic literature, makes the reader use their fingers on the touch screen to pry open the protagonist’s eyes to access the story, while digital art installations and many games use facial recognition algorithms or eye-tracking to integrate the viewer into the artwork. The analyses will draw upon the respective disciplines and methodologies of 1) art history and visual studies, 2) electronic literature studies, comparative literature, narratology and cinema studies and 3) game studies and ludology, with particular emphasis on the three main research questions.

**Research Task 4: An Ethnographic Study of Users of Machine Vision**

**Objective:** Develop an understanding of how individual users of machine vision technologies in smartphones and social media experience agency (RQ1), the relationship between visual representations and their perception of themselves and others (RQ2), and values/bias in machine vision (RQ3).

**Short Description:** Participant observation and in depth interviews will be used to study visual self-representation and communication in social media. Ethnographers and photography theorists have studied selfies and self-representation, but have focused on issues like societal power relations (Abidin 2016; Burns 2015), human-to-human communication (Frosh 2015) and body image (Tiidenberg 2014) and not specifically on taking and sharing selfies as a form of interaction with and through machine vision technology. The researcher will do fieldwork with two distinct groups of people, where one group is young adults with relatively stable life situations, and the other a less privileged group, such as refugees, recent immigrants, or unemployed youth. This entails ethical considerations which are addressed in the ethics annex. An anthropologist or ethnologist with experience in digital anthropology (likely recruited from the Association of Internet Researchers) will conduct this part of the research, which will involve participant observation, in-depth interviews, and analysis of visual material.

**Research Task 5: Consolidate and develop integrative theory of machine vision**

**Objective:** Synthesise results from the first four research tasks to provide comprehensive answers to the three research questions, resulting in an integrative theory of machine vision.

**Short Description:** Building upon the historical and theoretical work, as well as the empirical analysis and distant reading, I will identify shared themes or apparently contradictory findings to develop a novel and comprehensive theory of machine vision’s effect on our understanding of self and our surroundings.
Feasibility

High-Risk Nature of the Research: This is the first attempt to develop a theory of a new technology, and is thus inherently risky. Beyond this, the cross-disciplinary methodology and the goal of incorporating prior theory are high-risk. MACHINE VISION is radically cross-disciplinary in that it combines methods from the digital humanities and network analysis with aesthetic analysis and ethnographic fieldwork. Combining ethnographic and aesthetic research is uncommon and will require careful management to ensure it is mutually productive and not simply two separate strands of research. MACHINE VISION also aims to draw upon theories from computer science, and to develop research that is valuable for these disciplines as well as for the humanities, requiring networking and an ability to understand and communicate to other disciplines. The use of network analysis in Research Task 2 is also a risk factor as this method is not yet well established in cultural analyses, and while I hope it will provide valuable answers to the research questions and a model for a new methodology, it may also give no more than a useful overview and basic information about correlation. Further, prior theory (e.g. posthumanism (RQ1) or visual theories (Research Task 1 and State of the Art)) may not support the empirical analyses. MACHINE VISION will in any case develop new theory since today’s ubiquitous machine vision is qualitatively different from previous machine vision, but while I aim to go well beyond earlier work, I do hope to build upon it. If this is not possible, I will test theories from other domains, and develop new models to explain the empirical findings.

My capability to lead the project: My track record shows that I am consistently the first or among the first to study digital cultural phenomena, including blogging (Mortensen and Walker 2002), selfies (Walker 2005b), transmedia storytelling (Walker 2005a), self-tracking (Rettberg 2014b) and Snapchat (Rettberg 2017c; Rettberg 2017). I am an excellent networker, both in social media (see my blog jilltxt.net and Twitter and Snapchat accounts @jilltxt) and in person, which has led to field-building and collaborations (Corneliussen and Rettberg 2008; Senft et al. 2014). My education and experience provide a unique multi-disciplinary position that makes me the ideal leader for a project of this scope. My MA was in comparative literature, and my early research was on digital art, electronic literature and games. My first research publication was at the Association of Computing Machinery (ACM) Hypertext conference, where my literary analysis of a hypertext fiction won the award for Best Paper by a Newcomer (Walker 1999), and where I found the mixture of computer scientists and literary critics extremely productive. Now I regularly attend the Association of Internet Researchers conferences, and have collaborated with scholars from this community (Albury et al. 2017; Warfield et al. 2016), giving me a knowledge of ethnographic approaches to understanding digital culture.

Research Environment: I lead UiB’s Digital Culture Research Group, which is a natural home for the project. I will also work with UiB’s Electronic Literature Research Group, which hosts the ELMCIP KB that is central to Research Task 2. The PhD student working on games will also participate in the Computer Games Research Network at UiB, where I am a member. My team will have full administrative support from the University of Bergen, including research training for PhD candidates and high-quality technical, informational and organisational infrastructure for all researchers. A Reference Board of top international scholars with specific expertise on aspects of the project will be established, with participants invited to annual workshops. I have lived in Bergen for many years, and will be there at least 80% of the time.

High Gain and Multi-Disciplinary Research Impact: MACHINE VISION is a potent response to exactly the kind of research challenge the humanities must rise to today: We are seeing rapid technological development and societal adoption of machine vision, with massive research efforts on the topic from computer science, and the humanities have a wealth of relevant knowledge to contribute. However, humanities research on machine vision has so far been fragmented and not sufficiently responsive to the radical technological and cultural developments of the last decades. MACHINE VISION will develop new knowledge that will give researchers the necessary new tools for interpreting the role of everyday playful interaction with machine vision in social media, games, narratives and art, thus directly affecting research in aesthetic disciplines like art history, literary studies, game studies and cinema studies, but also social science research in media studies, internet studies, anthropology, and on surveillance, human-computer communication and digital culture in general. It will also inform the technical research and development on machine and computer vision which is now largely founded in computer science and engineering disciplines. It will have a societal impact by providing the general public with an understanding of machine vision's influence on ourselves and on society that enables us to make choices about the technologies we introduce to our culture, rather than simply embracing that which is technologically possible.
Bibliography


