

Ronja Reiten Fjærli

Manipulert virkelighet: Deepfakes, etiske konsekvenser av teknologien, både generelt og i populærkulturen, med spesielt fokus på serien falso amor

Bacheloroppgave i medievitenskap

Veileder: Aurora Hoel

Mai 2024

Ronja Reiten Fjærli

Manipulert virkelighet: Deepfakes, etiske konsekvenser av teknologien, både generelt og i populærkulturen, med spesielt fokus på serien falso amor

Bacheloroppgave i medievitenskap
Veileder: Aurora Hoel
Mai 2024

Norges teknisk-naturvitenskapelige universitet
Det humanistiske fakultet
Institutt for kunst- og medievitenskap



Kunnskap for en bedre verden

INNHALDSFORTEGNELSE

1. INTRODUKSJON	3
2. PROBLEMSTILLING OG METODE.....	5
3. TEORI	6
4. BRUKEN AV DEEPFAKE I MEDIEKULTUREN.....	7
5. REALITY-TV OG INTRODUKSJON TIL FALSO AMOR.....	11
6. FALSO AMOR	13
6.1 ANALYSE AV BRUKEN AV DEEPFAKE I FALSO AMOR.....	13
6.2 FØRSTE EPISODE: DEEPFAKE TEKNOLOGI	14
6.3 BADEBASSENGSCENEN	16
6.4 MARS THE WHITE ROOM	18
7. AVSLUTTENDE DISKUSJON	22
8. KILDER	24

1. INTRODUKSJON

I en verden der teknologi stadig utfordrer grenser, har fremveksten av deepfake skapt både reaksjoner og bekymringer rundt de potensielle etiske konsekvensene, som kan forekomme ved bruken av denne teknologien. Deepfake er en teknologisk utvikling som tok verden med storm. Det er en avansert teknologi innenfor kunstig intelligens (KI) som gjør det mulig å lage overbevisende forfalskede visuelt medieinnhold, og lydinnhold. Deepfakes er en form for digital manipulasjon som er laget for å skape realistiske videoer, bilder, eller lydopptak av personer som utfører handlinger eller sier noe de faktisk aldri har gjort eller sagt. Bruken av denne type teknologi har nå funnet vei inn i populærkulturen, som den spanske Netflix serien Falso Amor. Dette er hva min bacheloroppgave kommer til å handle om, så mitt spørsmål blir: Hva er deepfakes, og hvilke etiske konsekvenser kan forekomme rundt denne teknologien, med spesielt fokus på populærkulturen. Grunnen for at jeg har valgt dette temaet for min bacheloroppgave er fordi dette er en teknologi som er forholdsvis ganske ny. Og det er en teknologi ikke så mange har kunnskap om fra før. Jeg tenker det er viktig at flere forstår og vet om denne teknologien siden den kan være farlig uten forståelse og kunnskap. Deepfake teknologi har tatt mer og mer plass i samfunnet, og det har blitt mer og mer avansert. En nylig hendelse som tok sted på et landstreff i Stavanger da musikkgruppen Ballenciaga startet sin opptreden med å vise en video som da var en deepfake av kong Harald som sier:

«I dag er dere russ på landstreff Stavanger, nordmenn er jenter som er glad i jenter, gutter som er glad i gutter og russ som er glad i kokain, horer og Ballin. Og sist gang jeg traff guttene i Ballin, tok jeg kongelige linjer med de på Skaugum. De kan virkelig feste. Landstreff Stavanger, vær klar for å klikke, danse, og gjøre alt vi gjør daglig på slottet. Og når dere hører denne lyden.. vær klar for å gå inn i moshpiten. Husk å gjør meg stolt, kong Harald ut».

(Solberg et al., 2024)



Figur 1: Utklipp av videoen som ble tatt på landstreffet i Stavanger.

Dette eksempelet understreker behovet for å utforske og forstå konsekvensene av denne teknologien. Eksempelet viser til hvordan denne teknologien kan misbrukes til å skape forfalskninger som kan skape forvirring, manipulere oppfatninger, og til og med true tilliten til offentlige figurer og autoriteter. I dette tilfellet ble det kongen vår i Norge. Denne deepfaken av kongen kan skape forvirring rundt om kongen faktisk har vært med på å lage denne videoen. De som ikke har kunnskap om teknologien, kan godt tro på at han har vært med på dette. Og dette skaper problemer for kongehuset, siden det kan svekke den tilliten nordmenn har til kongehuset, for hvordan kan vi ha en konge som synes det er greit å ta kokain? Eksempler som dette har gjort at interessen min for å skrive om de etiske konsekvensene rundt deepfake har økt.

2. PROBLEMSTILLING OG METODE

Oppgaven kommer til å utforske de etiske konsekvensene av deepfake teknologien og dens rolle i populærkulturen, med spesiell vekt på Netflix serien falso amor. Sentralt i oppgaven står spørsmål rundt manipulasjon, etikk, og tillit. Oppgaven kommer til å handle om hva deepfakes er, dens historie og dens utvikling. Videre kommer jeg til å skrive om reality-tv, med spesielt fokus på reality-serien falso amor. Jeg kommer til å gjøre en analyse av første episode i serien. Jeg kommer til å analysere starten av episoden der vi blir introdusert for deltakerne, en scene som viser et eksempel på hvordan de tar i bruk deepfake i serien, og til slutt en scene i «the white room». Scenen i «the white room» er den scenen jeg kommer til å ha mest fokus på i min analyse. I denne scenen så vet ikke deltakerne at klippene de ser kan være falske, så her får jeg analysert ekte reaksjoner mennesker har mot deepfake. Jeg får også sett på hva de mener om teknologien når deltakerne får vite etterpå, at showet tar i bruk deepfake-teknologi. Jeg kommer til å gjøre en kvalitativ forskningsmetode, der jeg kommer til å gjøre en kvalitativ innholdsanalyse av første episode i serien. Grunnen for at jeg har valgt denne forskningsmetoden er fordi den passer best i forhold til min empiri som da er den første episoden i serien. Ved å gjøre dette får jeg sett ekte reaksjoner rundt deepfake teknologien, som gjør at jeg kan peke på etiske dilemmaer som skapes av deepfake teknologien, ved å se på hva mennesker tenker og mener om teknologien.

3. TEORI

Siden jeg studerer deepfake fra et medieteoretisk standpunkt så har jeg valgt å se på tre ulike teoretiske tilnærminger rundt denne type teknologi. Jeg kommer til å basere meg på nyere medieteorier som medieestetikk der jeg kommer til å ta i bruk Liv Hausken sine tilnærminger rundt deepfake teknologien. Jeg kommer til å ta i bruk teksten hennes «photorealism versus photography. AI-generated depiction in the age of visual disinformation». Liv Hausken har utforsket deepfake teknologien fra et medieestetisk perspektiv. Hennes tilnærming går ut på å analysere hvordan denne teknologien utfordrer tradisjonelle oppfatninger av autensiteten til bilder og den estetiske dimensjonen ved manipulert medieinnhold. (Hausken, 2024) Jeg kommer også til å se på Lev Manovich sin teori om nyere medier. Jeg kommer til å ta i bruk boken hans «The language of new media» der jeg kommer til å fokusere på begrepet fotorealisme, som «defineres som muligheten til å simulere et hvilket som helst objekt på en slik måte at datamaskinbildet ikke kan skilles fra dets bilde» (Manovich, 2001, s.199). Ved å inkludere denne teorien i min oppgave så hjelper det meg med å forstå hvordan deepfake-teknologien gjør det mulig å lage visuelt innhold som er så realistisk at det forveksles med ekte visuelt innhold. Dette blir viktig for å forstå risikoer og utfordringer knyttet til deepfakes, samt de etiske konsekvensene som forekommer av bruken. Til slutt så kommer jeg til å se på teorier rundt reality-tv, der Arild Fetveit sine tilnærminger blir sentralt. Jeg kommer til å ta i bruk teksten hans «Reality TV in the digital era: a paradox in visual culture». Her kommer Fetveit med interessante perspektiver når det gjelder reality-tv og dens påvirkning på samfunnet. Fetveit har utforsket hvordan reality-tv påvirker seernes oppfatning av virkeligheten, samt hvordan denne sjangeren formidler normer og verdier. Fetveit diskuterer hvordan reality-tv kan være med på å forvrengte folks oppfatninger av virkeligheten. (Fetveit, 2004, s. 788-789) ved å skape kunstige konflikter og dramatiske situasjoner for underholdningens sin skyld. Ved å dra inn hans tilnærminger når det gjelder reality-tv og spesielt med tanke på serien falso amor får jeg undersøkt hvordan tilnærmingene til Fetveit og deepfake muligens forsterker hverandre.

4. BRUKEN AV DEEPFAKE I MEDIEKULTUREN

Deepfake er en avansert teknologi innenfor kunstig intelligens (KI) som gjør det mulig å lage overbevisende forfalskede visuelt medieinnhold, og lydinnhold. «Deepfakes oppsto i 2017, og da var denne teknologien ny, men baserte programvaren som skapte disse bildene på en teknologi som hadde blitt produsert noen år tidligere» (Langguth, 2021, s.5).

«Deepfakes fungerer gjennom en kombinasjon av «deep learning techniques» og nevralt nettverk for å analysere store mengder data og deretter manipulerer ansikter og stemmer til å bli realistisk. Denne teknikken bruker ofte generative modeller, blant annet Generative adversarial networks (GANs). GANs består av to komponenter: en «generator» som lager forfalskede medieelementer og en «diskriminator» som evaluerer hvor godt generatoren gjør det. Denne prosessen mellom disse komponentene fører til stadig mer realistiske resultater.» (Fraga-Lamas, 2020, s. 34).

I Lev Manovich sin bok «the language of new media» kommer han med interessante tilnærminger for å forstå hvordan digital teknologi former moderne kultur og kunstneriske uttrykk. I boken så diskuterer han rundt hvordan digital teknologi, spesielt datamaskinbasert bildebehandling endre måten vi skaper, og tolker medieinnhold.

Dette blir relevant når vi ser på deepfake teknologien. Manovich sine ideer om «mediespesifisitet», som refererer til de unike egenskapene, språket og estetikken til medieformer, kan anvendes ved deepfake (Manovich, 2001, s. 237). Deepfakes demonstrerer hvordan digitale medier kan manipulere virkeligheten på en måte som man ikke kunne se for seg tidligere. Ved å kombinere algoritmer og maskinlæringsteknikker kan deepfakes skape overbevisende manipulert visuelt innhold og lydopptak som utfordrer vår oppfatning av virkeligheten. I boken så diskuterer Manovich også ideen om lagdeling i digitale medier, som referer til hvordan digitale objekter ofte består av flere lag med informasjon og manipulasjon. Deepfakes utnytter jo denne lagdelingen ved å manipulere medieinnhold på ulike nivåer, noe som resulterer i realistiske forfalskninger som kan være vanskelig å oppdage (Manovich, 2001, s. 30-35). Et begrep som Manovich skriver om i sin bok som er veldig relevant i forhold til deepfake er fotorealisme. Fotorealisme referer til digitale bilder som vil gjengi

virkeligheten på en troverdig måte. (Manovich, 2001, s. 199). Deepfakes tar dette konseptet til nye høyder ved å produsere innhold som er så realistiske at de kan forveksles med det ekte. Et eksempel på deepfake er da Jordan Peele lagde en deepfake video av den tidligere amerikanske presidenten Barack Obama. Peele lagde denne deepfaken med hensikt i å varsle folk om den potensielle faren ved deepfake. I videoen ser man Obama som sitter foran et kamera og leverte en kunstig generert tale som var designet av Peele. Det mange reagerte på ved denne videoen var Obama sitt ansikt og stemme. Det var nøyaktig replikert så det var vanskelig å skille den fra en ekte opptreden av den tidligere presidenten. Peeles deepfake av Obama er et godt eksempel på hvor avansert og realistisk denne typen teknologi har blitt. Videoen var med på å få folk til å være mer kritisk til digitale medieinnhold, og forstå mulige konsekvenser ved misbruk av den. (Giansiracusa, 2021, s.50)



Figur 2 viser til et eksempel på deepfake som Jordan Peele lagde av Barack Obama (Buzzfeedvideo, 2018).

Deepfakes er jo et fenomen som har skapt både fasinasjon og frykt i dagens digitale samfunn. Med dens avanserte algoritmer og maskinlæringsteknikker så har dette fenomenet revolusjonert vår evne til å manipulere digitale medier, og potensialet for misbruk har økt betraktelig. (Giansiracusa, 2021, s. 34) Deepfakes er et kontroversielt fenomen som truer både personvern og integritet. Fenomenet er med på å vekke en rekke spørsmål rundt hva som er ekte eller falskt. Deepfakes har evnen til å skape overbevisende falske videoer og lydopptak som gjør at vi ikke lenger kan stole på det vi ser og hører på nett. Dette er med på å fremme spørsmål om hva sannhet egentlig har å si i en verden der det manipulerede innholdet som blir skapt kan være like troverdig som det som faktisk er ekte. Det er også med på å utfordre vårt syn på troverdige kilder og hvordan vi kan stole på informasjon i en tid der manipulasjon er så godt utviklet. Når det gjelder det etiske så reiser deepfake teknologien bekymringer knyttet til personvern, tillit, og misbruk. Det med at deepfakes kan brukes for å skade enkeltpersoner, spre falsk informasjon eller manipulerer meninger til enkeltpersoner kan ha alvorlige konsekvenser for samfunnet generelt.

Som nevnt tidligere så oppsto deepfakes i 2017 men det er en lang historie rundt manipulert visuelt innhold. I boken *Fatamorgana-Pressefotografiets virkelighet og læsernes av Bruno Ingemann* skriver han om historien til bildemanipulasjon. I boken diskuterer Ingemann rundt hvordan

«pressefotografier skal være realistiske, troverdige, sannferdige og vise tingene som de er. Og at i bokens kontekst; at pressefotografiet skal være sett av en objektiv, neutral, og realistisk øye. Kameraet er en garanti for at dette skal være tilfellet, men det alminnelige pressefotografi utfordres gjennom den digitale bildekultur». (Ingemann, 1998, s. 10)

Ingemann skriver at man ikke lenger kan være sikker på det man ser. Dette har jo blitt ekstremt relevant den dag i dag, med tanke på all ny teknologi som kommer. Videre formulerer Ingemann at

«et fotografi som ser realistisk ut, kan ha blitt manipulert gjennom en konstruksjon av flere individuelle fotografier som er slått sammen til et bilde. Man kan ikke lenger være sikker på at personene man ser på et bilde faktisk har vært på samme sted, på samme tid og har blitt fotografert sammen. Personene kan være fotografert på forskjellige steder og til forskjellige tider. Og derfor er bildet manipulert.» (Ingemann, 1998, s. 10)

Ingemann kommer frem med et eksempel på dette i boken. På figuren nedenfor ser man et bilde fra en reklamefilm. I bildet ser man Elton John og Louis Armstrong som spiller sammen i Coca Colas reklamefilm. Elton John ble spilt inn i 1991 mens Louis Armstrong er hentet fra en scene i High society fra 1956.



Figur 3 er et eksempel som viser til hvor langt tilbake visuell manipulasjon har blitt tatt i bruk. (Ingemann, 1998, s. 11).

Dette eksempelet kan også kalles shallowfakes som er manipulasjon av medieinnhold som ikke innebærer avansert teknologi eller dyptgående kunstig intelligens. Dette kan inkludere enkle redigeringsmetoder som å klippe og lime ansikter på videoer, endre stemmer på lydopptak eller å lage falske bilder ved hjelp av enkle verktøy og apper. Shallowfakes inkluderer ikke like avansert teknologi så det er noe som er tilgjengelig for hvem som helst. (Giansiracusa, 2021, s. 44-45). Et eksempel på shallowfakes som har gått enda lengere tilbake enn eksempelet med Elton John og Louis Armstrong er da «hodet til Abraham Lincoln ble satt på kroppen til John C. Calhoun som var sterk tilhenger av slaverisystemet i USA. Dette ble gjort for å gi Lincoln en mer heroisk positur som kunne gagne Lincoln videre» (Giansiracusa, 2021, s. 43). Dette eksempelet går tilbake til 1860 som også viser hvor lang historie visuell manipulasjon har.

5. REALITY-TV OG INTRODUKSJON TIL FALSO AMOR

Reality-tv har blitt en stor del av den moderne tv underholdningen vi ser i dag. Disse programmene gir seerne et glimt inn i virkelige menneskers liv. Men hva er reality tv, når oppsto det, og hvordan oppsto reality tv? Og ikke minst hvilke spørsmål vekker det når det gjelder fakta og fiksjon? Reality sjangeren går inn for å skildre mennesker i situasjoner med fokus på hverdagsliv, personlige konflikter eller konkurranser. I motsetning til serier der skuespillere går inn i roller basert på et skrevet manus så involverer reality-tv ekte mennesker som står ovenfor utfordringer, konkurranser eller rett og slett blir observert i sitt daglige liv. (Kavka, 2012, s.3)

Reality sjangeren har fokus på dramaet og humoren rundt menneskers liv med en tilnærming som skal være at seeren blir «fluen på veggen». Reality tv gir seerne et innblikk i virkeligheten, men det er også viktig å poengtere at reality programmer ofte inneholder elementer av fiksjon og iscenesettelse. Produksjonen og tv-nettverk kan manipulere situasjoner, redigere opptak og til og med oppmuntre deltakerne i programmet til å oppføre seg på bestemte måter for å øke dramatikken og underholdningsverdien. Dette fører til at seerne får et skjevt bilde av virkeligheten og de menneskene som er med i programmene. Og gjør at reality sjangeren, ikke lenger blir en «reality».

Reality tv begynte å ta form på 1940 – 50- tallet, og sjangeren har gjennomgått en betydelig transformasjon siden dens begynnelse (Kavka, 2012, s.14). Gjennom Arild Fetveits tilnærminger rundt sjangeren i teksten "Reality TV in the Digital Era: A Paradox in Visual Culture" kommer Fetveit med gode poeng når det gjelder den utviklingen sjangeren har hatt gjennom tidene. Fetveit fokuserer både på de kulturelle og teknologiske dynamikkene som har formet reality-tv over tid, og ser på hvordan disse elementene har ført til paradoksale utfall i visuell kultur. (Fetveit, 2004, s. 787) Reality-tv startet med programmer som eksperimenterte med å vise virkelige hendelser eller situasjoner, ofte uten manus. Programmer som «candid camera» og «an american family» skapte en ny form for tv-underholdning som både skapte fasinasjon og provoserte ved å blottlegge det private i offentligheten. Disse programmene skapte grunnlaget for en sjanger som ville utforske virkeligheten som underholdning. Overgangen med digitale endringer som Fetveit beskriver, har digital teknologi hatt en enorm innvirkning på reality-tv. Tilgjengeligheten av billig digital videoopptaksteknologi og bredere

distribusjonskanaler, spesielt gjennom internett har demokratisert hvem som kan produsere og distribuere innhold. (Fetveit, 2004, s. 788-789)

Dette har åpnet for en eksplosjon av reality-innhold som spenner fra profesjonelle produksjoner til brukergenererte videoer. Noe som står sentralt i teksten til Fetveits er paradokset som oppstår når reality-tv blir en dominerende form for visuell kultur. (Fetveit, 2004, s. 792-794) På den ene siden tilbyr reality-tv en form for «ekte» underholdning som appellerer til publikums ønske om autensitet og ufiltrerte menneskelige erfaringer. På den andre siden, fremmer reality-tv ofte forvrengte eller iscenesatte versjoner av virkeligheten som kan være villedende. Dette skaper et spenningsfelt mellom fakta og fiksjon, hvor seere må navigere mellom det ekte og det produserte. Sosiale medier har transformert seeropplevelsen ved å tillate publikum i å engasjere seg direkte med reality-programmer og deres deltakere. Denne interaktiviteten forsterker følelsen av deltakelse og innflytelse, men den kan også forsterke sjangerens problematiske aspekter, som manipulasjon og overvåkning. Dette er noe som er nytt innenfor sjangeren og gjøres gjennom programmer som for eksempel love island. I dag har digital innovasjon presset grensene for hva reality sjangeren kan omfatte. Reality-tv må håndtere både teknologiske og etiske utfordringer mens den fortsetter å utvikle seg og tilpasse seg en stadig mer tilkoblet og mediedrevet verden. (Fetveit, 2004, s. 794 – 796)

Reality sjangeren utfordrer ofte våre oppfatninger av fakta og fiksjon, siden man går inn i det med en tanke om det man ser er virkelig. Dette er med på å reise etiske spørsmål som berører deltakere så vel som seere. Med tanke på deltakerne, vet vi for eksempel om deltakerne som er med i en reality serie er fullt informert om hva de går inn på? og er det egentlig tilstrekkelig hensyn til hva som er best for dem og deres privatliv? Når det gjelder manipulasjon, hvor langt bør produksjonen gå når det gjelder å manipulere situasjoner med deltakere for å skape dramatik og underholdning? Hva med konsekvenser av å være med på disse reality programmene? Kan eksponeringen og offentliggjøringen ha negative konsekvenser for deres personlige og profesjonelle liv? Reality sjangeren er en mye omdiskutert sjanger, men det fortsettes å produseres nye programmer hele tiden, med nye konsepter for å holde seg med trendene, et program som viser til dette et Netflix serien falso amor.

6. FALSO AMOR

Falso Amor er en spansk reality serie som ble publisert på Netflix. Serien går ut på at det er fem kjærestepar med et ønske om å teste kjærligheten. Kjæresteparene blir splittet opp og plassert sammen med single mennesker i to luksusvillaer, som kalles venus og mars. Etter hvert som deltagerne henger sammen med de single blir deepfakes skapt. Hvert par får se ulike videoer som enten kan være ekte eller falske og de skal deretter gjette om de er det. Det paret med minst feil vinner til slutt en pengepremie på 100 000 euro. Programmet har kjennetegn som stemmer godt overens med reality sjangeren. Vi har deltakere som blir overvåket med kameraer, vi har klipp som blir redigert for å virke mer dramatisk enn det, det egentlig er, og vi har en pengepremie som deltakerne kjemper om å vinne. Noe nytt som denne reality serien tar i bruk som ikke har blitt gjort før er bruken av deepfake teknologien. (Cuarzo producciones, 2023)

6.1 ANALYSE AV BRUKEN AV DEEPFAKE I FALSO AMOR.

I denne delen av min oppgave skal jeg analysere første episode i serien Falso amor. Der jeg kommer til å ta i bruk min kunnskap om deepfake, og reality-tv. Jeg kommer til å dra inn teori fra både Liv Hausken, Lev Manovich, og Arild Fetveit i min analyse. Deres tilnærminger kommer til å hjelpe med å forstå reaksjoner og hendelser i serien fra et medieteoretisk perspektiv. Oppgaven min omhandler etiske konsekvenser rundt deepfake, så ved å analysere den første episoden av serien, får jeg sett på ekte reaksjoner og tanker rundt fenomenet. I den første episoden av falso amor vet ikke deltakerne konseptet til showet og tror at alt de ser av klipp av kjæresten er ekte, så deltakerne er ikke klar over at noe av det de har sett kan være falskt. Ved å analysere denne episoden får jeg sett ekte reaksjoner til deepfake teknologien, siden man får se reaksjoner til deltakerne når de ser klippene som er deepfakes som de tror er ekte videoer. Etter at deltakerne får vite at noen av klippene kunne vært deepfakes så former de seg meninger rundt fenomenet som gir meg innsikt i hva de mener om deepfakes som også vil være nyttig for å kunne svare på mitt hovedspørsmål i oppgaven.

6.2 FØRSTE EPISODE: DEEPPFAKE TEKNOLOGI

Den første episoden av falso amor kalles deepfake-teknologi. Episoden starter dramatisk med å vise programlederen for showet Raquel Sánchez Silva. Hun gir informasjon om at parene som skal være med i showet har kommet dit for å teste sin kjærlighet for hverandre. Dette er med å skape en spenning og nysgjerrighet som bygges opp rundt serien. Vi får deretter sett en rekke med klipp fra serien for å få en liten smakebit på hva vi skal begi oss ut på. Klippene inkluderer både ekte klipp og noen deepfakes. (0.50 sek, Cuarzo producciones, 2023) Blant annet deltakerne som driver og danser med hverandre, kysser, og til og med en sex scene helt til slutt. Deretter blir vi presentert for «the white room». Det er rommet deltakerne skal se klipp av kjæresten sin som enten kan være ekte eller falsk. så får vi høre om konseptet med serien, der Raquel gir informasjon om deepfake teknologien som blir tatt i bruk i et reality program for første gang. Vi får deretter sett et eksempel på en deepfake som blir brukt i serien. Så blir vi presentert med flere klipp som viser frem bruken av deepfake-teknologi i serien. (Cuarzo producciones, 2023)

Liv Hausken sin teori om manipulert virkelighet og fotorealisme er relevant her, da første del av episoden demonstrer hvordan deepfake-teknologien manipulerer virkeligheten ved å skape falske situasjoner som blir vanskelig å skille fra ekte situasjoner. Deltakerne skal få se klipp av sine kjærester, noen ekte og noen som er falske, og de blir da utsatt for en form for virkelighetsmanipulasjon uten engang å vite det. Dette kommer til å utfordre deres oppfatning av hva som er ekte og hva som er falskt, og dette er med på å legge grunnlag for den konflikten og spenningen som skal utfolde seg i løpet av serien. (Hausken, 2024, s. 2-5)

Vi får også sett flere av reaksjonene som deltakerne har mot disse deepfake klippene. Vi får sett reaksjonen til Javier som er en av deltakerne i serien. Han sier «jeg tror ikke på det», «jeg ser ikke forskjell» mens du ser tydelig at han er frustrert over det han ser på skjermen. Her er en av de første reaksjonene vi får sett deltakerne har mot deepfake teknologien. I dette tilfellet så vet ikke Javier om de han ser er ekte eller ikke, men det er frustrerende i forhold til at deepfake teknologien er såpas god at han ikke klarer å se forskjell om det er ekte eller ikke. (2.45 min, Cuarzo producciones, 2023) Javier uttrykker frustrasjon over sin manglende evne til å skille mellom ekte og falske klipp, noe som henter til hvor avansert deepfake teknologien har blitt. Dette viser til hvordan digitale medier har gjort det mulig å produsere visuelt innhold til å bli så realistisk sånn at de kan forveksles med ekte visuelt innhold. Dette er med på å

utfordre tradisjonelle oppfatninger av virkeligheten. De ulike klippene med forskjellige scener som vi kommer til å se i showet kommer frem. Der får vi se flere deltakere sine reaksjoner på de ulike klippene de får se av kjæresten sin. Noen blir sinte, andre blir frustrert, og noen begynner til og med å gråte. Vi får altså se en hel del reaksjoner til deltakerne i starten som er med på å øke spenningen for showet, som gjør at seeren får lyst til å fortsette å se på. Etter at teaseren for showet er ferdig får vi endelig møte de fem parene som skal være med på showet. Raquel presenterer showet ovenfor deltakerne, og så får vi blitt kjent med de fem parene.

De fem parene som skal være med i serien er: Manuel og Aida, Javier og Paula, Àngel og Gabriela, Rubén og Isabel, Alejandro, og Ramón. Vi får en kort presentasjon av hvert par der vi får informasjon om hvor lenge hvert par har vært sammen, og grunnlaget for deres deltakelse i serien. Etter presentasjonen av hvert par så blir de splittet til å dra i hver sin villa. Dette var starten av første episode, videre i min analyse kommer jeg til å analysere to scener som jeg mener er viktigst i denne episoden. Den første scenen jeg kommer til å analysere er badebassengscenen. Jeg har valgt denne scenen fordi den viser hvordan serien tar i bruk deepfake-teknologien og er et eksempel på hvordan en deepfake blir laget. Den andre scenen jeg har valgt å analysere er når mars villiaen er i «the white room». Jeg har valgt denne scenen siden det er her deltakerne får sett en deepfake for første gang. Her får man sett reaksjonene til deltakerne, og tanker de har om teknologien.

6.3 BADEBASSENGSCENEN



Figur 4 viser til et eksempel på hvordan serien tar i bruk deepfake-teknologi. (18.48 min, Cuarzo producciones, 2023)

Den første scenen jeg kommer til å analysere er badebassengscenen. I denne scenen får vi sett et eksempel på hvordan deepfake teknologien blir brukt i showet og viser til hva det er deltakerne får sett av hverandre. En av de single i villaen som heter Jesús utfordrer Aida og Torres at de skal komme veldig nær hverandre uten å kysse hverandre i 10 sekunder. Aida og Torres gjør dette og ligger ved bassengkanten og ser hverandre inn i øynene med ansiktene nære hverandre uten at de gjør noe mer. De andre deltakerne teller ned fra 10 mens de ser hverandre i øynene. Deretter popper det opp et redigert klipp av Aida der ansiktet hennes bli «scannet» og vedsiden av står det navn, alder, og kjønn. Deretter kommer det en voice-over av Raquel som informerer om at nå skal deepfake-teknologien endre situasjonen mellom Aida og Torres til at det blir et kyss. (19.16 min, Cuarzo producciones, 2023) Skjermen splitter seg, og vi får sett den lagde deepfaken sammenlignet med det som faktisk skjedde. Her får vi se en skuespiller som skal spille Aida som da kysser med Torres, og de har da laget en deepfake ved å sette ansiktet til Aida over denne skuespilleren sitt ansikt for å få det til å se ut som det er Aida Torres kysser. Etter at vi har blitt vist deepfaken og den ekte sitasjonen sammenlignet med hverandre, kutter scenen til når Manuel er i «the white room» og skal få se deepfaken som har blitt laget av kjæresten Aida. Manuel sier med en gang «det kan ikke være sant»

mens han ser veldig frustrert ut over det han ser på skjermen. Det klippes til ulike reaksjoner til deltakerne inne på villa Venus der noen gråter, noen er sjokkerte, og til slutt et klipp over at Manuel er oppriktig sint, og knytter knyttnevene og slår seg selv på lårene. Så er scenen over. (19.31 min, Cuarzo producciones, 2023)

I denne scenen så blir vi introdusert for bruken av deepfake teknologi i serien. I begynnelsen av scenen blir deltakerne Aida og Torres utfordret til å komme nær hverandre uten å kysse. Etter dette blir vi introdusert for det tekniske ved serien gjennom en visuell illustrasjon, der vi ser Aidas fjes som blir «scannet» mens hennes personopplysninger vises ved siden av. Dette viser hvordan deepfake-teknologien opererer ved å manipulere og generere falske situasjoner basert på eksisterende data. Dermed får vi en forklaring på hvordan serien tar i bruk teknologien gjennom en dramatisk scene som demonstrerer både teknologiens potensiale og de utfordringene den stiller opp med. I badebassengscenen blir vi vitne til deepfake-teknologiens manipulasjon av virkeligheten gjennom produksjonen av en falsk situasjon mellom Aida og Torres. Dette viser til den problematiske bruken av KI-generert avbildning noe som Hausken kritiserer i sin tekst. Hun kommer frem med et eksempel fra Amnesty International der hun diskuterer etiske og sosiale konsekvenser gjennom KI-generert avbildning. Hausken advarer mot hvordan denne type teknologi kan utnyttes for å skape forvrengt og manipulert visuelt innhold som kan være med på å forvirre publikum. (Hausken, 2024, s.7-10) Deepfake-teknologien i scenen viser til akkurat dette, da den skaper et realistisk, men falskt bilde av et kyss mellom Aida og Torres. Det presenteres som ekte, noe det ikke er. Hauskens tilnærming til dette belyser over til de etiske og sosiale implikasjonene av deepfake-teknologien. Bruken av denne type teknologi er ikke bare med på å undergrave tilliten til mediene og svekke troverdigheten til visuelt innhold, men det kan også muligens føre til alvorlige konsekvenser. For eksempel spredning av falske nyheter og desinformasjon, samt krenkelse av personvern til enkeltpersoner. (Hausken, 2024, s. 9-10) Dette blir spesielt tydelig når kjæresten til Aida, Manuel reagerer med frustrasjon til kjæresten sin, noe som viser til de følelsesmessige og psykologiske konsekvensene av denne type teknologi.

6.4 MARS THE WHITE ROOM

Den neste scenen jeg har valgt å analysere er når mars villaen skal inn i «the white room» for å se klipp av kjærestene sine. Grunnen for at jeg har valgt å analysere denne scenen er fordi det er her jeg faktisk får sett reaksjonene til deltakerne rundt deepfake teknologien.

Deltakerne kommer inn sammen i «the white room» og setter seg ned på hver sin krakk. Hver deltaker sitt ansikt blir zoomet inn på og alle deltakerne ser veldig stresset ut der de sitter.

Raquel informerer deltakerne om at fremover er det i «the white room» de kommer til å få se klipp av partneren sin. Hun gir beskjed om å se på stolen fremfor dem og informerer om at det er «sannhetens sete» og at der er bare den personen som sitter i den stolen som får lov til å se klippene som blir vist. (29.15 min, Cuarzo producciones, 2023)

Den første som skal sitte i «sannhetens sete» er Javier. (30.50 min, Cuarzo producciones, 2023) Javier er stresset og håper at kjæresten Paula har hatt det gøy, men innenfor grensene de har satt mellom hverandre. En rekke klipp av Paula og en av de single kommer opp, disse klippene er ekte. Men det er et ekstra klipp på slutten som viser til Paula og Lolo som er i bassenget. (32.02) Dette klippet er en deepfake. Javier blir oppgitt og han skjuler ansiktet med hendene i fortvilelse over det han ser på klippet. Deretter sier han «hun kan ikke ha det uttrykket» «det ansiktsuttrykket oppmuntrer til å gjøre hva som helst». Han fortsetter med å uttrykke sin misnøye ved klippene han har sett av kjæresten, og ender turen sin med å si «hvorfor er vi her? for å teste oss selv, eller for å filme en voksenfilm? Kommentarene til Javier gir uttrykk for frustrasjon, og misnøye over klippene han har sett, i hvert fall over deepfaken han så av kjæresten Paula i bassenget.

Neste person som skal i «sannhetens sete» er Gabriela, hun er stresset og sier «jeg vet hvordan Àngel så ut da han møtte meg, jeg vil ikke se det her». Deretter får hun se klippene av kjæresten. Her er det ikke laget noen deepfakes, og alle klippene som er vist av Àngel er ekte. Àngel hadde en god tone med en av de single Zara, og på slutten så får man sett et klipp av Zara som kysser Àngel. Det blir kuttet av før man får sett om de faktisk kysser eller ikke. Dette er ikke deepfake teknologi, men det er en form av manipulasjon av innholdet som blir vist, siden dette får Gabriela til å tro at de kysset. Vi får ikke sett om de faktisk gjør det eller ikke. Og dette er noe som er typisk for reality sjangeren med å manipulere ved redigering. (33.17 min, Cuarzo producciones, 2023)

Neste person som skal i «sannhetens sete», er Rubén. Alle klippene som blir vist av kjæresten hans Isabel er ekte. I klippene så får vi sett at hun snakker med en av de single, og et klipp av at hun har en isbit i munnen som hun skal gi videre til en av de single. Rubén blir irritert, og skuffet over oppførelsen til kjæresten. (35.35 min, Cuarzo producciones, 2023)

Nestemann ut til «sannhetens sete» er Alejandro. Han er åpenbart stresset, og lei seg over situasjonen. Han er redd for å bli skuffet over det han kommer til å se av kjæresten Ramón. Alejandro får se en rekke klipp av kjæresten mange av de er ekte, men det er en deepfake helt til slutt som viser at Ramón slipper inn en av de single på soverommet sitt. Etter klippet er avsluttet sier Alejandro «Jeg blir gal her, Raquel. Seriøst» Alejandro blir overasket over klippene han har fått sett av kjæresten, og sier til og med «Vil jeg være med en sånn fyr resten av livet? For da skal jeg rote med alle guttene i huset». Dette viser Alejandro sin frustrasjon og fortvilelse. Før klippene ble vist så var han mer lei seg, men etter å ha fått sett klippene så har han blitt mer sint og irritert, men etter han har satt seg ned med de andre deltakerne tar han til tårene igjen. (37.03 min, Cuarzo producciones, 2023)

Sistemann ut er Aida. Hun er veldig stresset for de hun skal se, hun rister med foten og viser klare signaler om at hun er stresset og ukomfortabel rundt situasjonen. Når hun blir vist klippene av kjæresten Manuel blir hun lei seg, hun begynner å gråte og tramper hardt med foten, og sier til og med «drittsekk» mot skjermen. I starten så er klippene ekte, men det nest siste klippet blir Aida vist en deepfake av kjæresten. Manuel er i bassenget med en av de single i villaen. Der blir skuespilleren som skal være Manuel kysset på nakken av den single og hun blir tatt på rumpen av skuespilleren. Når klippene er ferdig begynner Aida å gråte, deretter sier hun at hun har lyst å dra. Raquel spør Aida om hun kjenner igjen Manuel i klippene hun har blitt vist. Aida bryter sammen og sier «jeg er knust». (39.30, Cuarzo producciones, 2023)

Etter at alle deltakerne har sett hva partneren deres har gjort på den andre villaen så forteller Raquel konseptet til showet. Deltakerne trodde at de bare skulle se hva partneren hadde kommet til å gjøre hvis de var fra hverandre og med andre single mennesker. Deltakerne viste altså ikke om bruken av deepfake teknologien i showet, så når deltakerne så på klippene av partneren sin så trodde de at alt de så var ekte. Deltakerne blir sjokkert over det de hører, og begynner å tvile over det de har sett nylig av partneren sin. (41.42 min, Cuarzo producciones, 2023)

«Kysset han virkelig en annen jente?»

«Kanskje han egentlig ikke sa det som ble sagt på klippet?»

Deltakerne begynner til og med å tvile på klipp som ikke engang var deepfakes, men ekte klipp. Dette er et godt eksempel på hvordan deepfake-teknologien har blitt så realistisk at man mister virkelighetsoppfattelsen, og man lenger ikke klarer å skille mellom det falske og ekte. Når de er på vei ut av «the white room» diskuterer de rundt hva de nettopp har sett, og deepfake-teknologien. Når de er på vei ut av «the white room» diskuterer de rundt hva de nettopp har sett, og deepfake-teknologien.

Aida sier dette er sykt. (46.36 min, Cuarzo producciones, 2023)

De er identiske, jeg så Paula sier Javier (43.01min, Cuarzo producciones, 2023)

Hva om det ikke er henne med isbiten? Eller kanskje det er det? Sier Rubén (43.04 min, Cuarzo producciones, 2023)

Det var ikke ham sier Gabriela, Nei det var ikke Manuel heller sier Aida. (43.24 min, Cuarzo producciones, 2023)

Jeg er deprimert nå sier Alejandro, Nå som jeg vet at alt kan være falskt, er det som.. så blir han avbrutt, for så at han sier jeg er forvirret, jeg vet ikke om det som skjer er ekte eller falskt. (43.36 min, Cuarzo producciones, 2023)

Manuel tok ikke på rumpa hennes sier Aida (43.45 min, Cuarzo producciones, 2023)

Det er det som skremmer meg mest svarer Javier til Aida.

Deltakerne kommer tilbake til mars villaen og de single sitter i spenning for å høre hvordan det går. Gabriela begynner å forklare om at de er usikre på om det de så var falskt eller ekte siden det har blitt tatt i bruk deepfake-teknologi. Deretter sier hun: Først ville jeg dø, og nå tviler du spør en av de single for så at hun sier seg enig i at hun tviler. Senere i et intervju

forteller Gabriela at det var en liten lettelse å vite at det hun så kanskje ikke var ekte. (44.04 min, Cuarzo producciones, 2023)

I denne scenen hvor deltakerne blir introdusert for «the white room» i mars villaen blir bruken av deepfake tydelig. Denne scenen er kritisk når det gjelder å forstå hvordan serien manipulerer virkeligheten gjennom denne teknologien. Før deltakerne går inn i rommet, forsterkes spenningen og nervøsiteten deres tydelig ved deres kroppsspråk. Raquel introduserer dem til formålet med «the white room». Når deltakerne blir vist klippene av partneren sin så får man virkelig se konsekvensene ved deepfake bruken. Dette vises særlig ved reaksjonen Javier har når han blir presentert ovenfor et klipp av sin kjæreste Paula som kysser en annen deltaker. Hans frustrasjon og forvirring over det han ser, samt kommentarene han kommer med som «det kan ikke være sant», understreker virkningen deepfake-teknologien har på deltakernes oppfatning av virkeligheten. Gabrielas reaksjon er også viktig når hun uttrykker frykt for hva hun skal se og hvordan det vil påvirke hennes syn på forholdet med Ángel. Reaksjonen hennes viser den følelsesmessige byrden denne teknologien kan ha på deltakernes mentale helse og tillit til partneren deres, og hvor mye teknologien kan faktisk ødelegge mellom parene. Rubén og Alejandro's reaksjoner bidrar også til å vise den psykologiske påvirkningen av deepfake-teknologien. Rubén's irritasjon og Alejandro's fortvilelse viser effekten av å bli vist manipulert innhold som utfordrer deres oppfatning av virkeligheten og deres forhold.

I Liv Hausken sin tekst «photorealism versus photography. AI-generated depiction in the age of visual disinformation». Utforsker hun den økende bruken av KI-teknologi i å skape realistiske bilder og videoer, og hvordan dette påvirker vår oppfatning av virkeligheten. I denne scenen fra falso amor, der deepfake-teknologien blir tatt i bruk for å manipulere videoer av deltakerenes partnere, kan vi se hvordan Hauskens tilnærminger til deepfake blir relevant. I denne scenen så viser de hvordan deepfake-teknologien skaper en illusjon av virkeligheten. Ved å generere realistiske, men falske videoer basert på situasjoner som aldri har skjedd. Deltakerne blir deretter vist disse falske videoene med et inntrykk om at de er ekte, og deres reaksjoner viser hvordan denne manipulasjonen av virkeligheten kan føre til forvirring, frustrasjon og følelsesmessig belastning. (Hausken, 2024)

7. AVSLUTTENDE DISKUSJON

I denne oppgaven har jeg utforsket de etiske implikasjonene rundt deepfake-teknologien, og hvordan denne teknologien nå har tatt plass i populærkulturen, med spesielt fokus på Netflix serien falso amor. Gjennom en analyse av første episoden av serien har jeg undersøkt de etiske implikasjonene ved bruken av deepfake-teknologien. I oppgaven har jeg undersøkt hva deepfake er, dens utvikling, og dens historie.

En sentral bekymring rundt deepfakes er dens potensiale til misbruk og konsekvensene som kan forekomme av dette. Jeg har kommet frem med eksempler på hvordan deepfakes kan brukes til å spre falsk informasjon, manipulere meninger, og til og med true tilliten til offentlige figurer og institusjoner. Denne teknologien har skapt mange nye muligheter og kan være positivt, men det har mye med hvem som tar i bruk teknologien, og hvordan de tar i bruk teknologien. Burde alle ha denne teknologien så lett tilgjengelig som vi har?

Teknologien kan skape konflikter, og skade på enkeltpersoner og uten noen form for regler rundt bruken, kan dette ha alvorlige konsekvenser for samfunnet.

Reality-tv har lenge utfordret seerens oppfatninger av fakta og fiksjon, og med en fremvekst av digital teknologi har sjangeren møtt nye utfordringer, men også muligheter. I serien falso amor har de valgt og tatt i bruk denne teknologien som har reist spørsmål rundt det etiske ved denne bruken av teknologien. Når jeg har undersøkt reality sjangeren har jeg sett på hvordan reality-tv allerede brytes mellom fakta og fiksjon. Med tanke på hvor manipulert reality programmer er fra før, så er elementet av deepfake-teknologien en interessant faktor å dra inn. Deepfake teknologien skaper falske situasjoner og er med på å manipulere deltakernes virkelighetsoppfattelse og tuller med følelsene deres.

Falso amor tar en risiko med tanke på hvordan de representerer teknologien og hvordan den blir forstått av seerne. Noen kan se på bruken som et kreativt valg med tanke på at de utforsker teknologiens potensial. Mens kan andre se det som en normalisering eller til og med en glorifisering av manipulasjon og destruktiv bruk av teknologi. Dette understreker behovet for nøye refleksjon rundt hvordan deepfake-teknologi blir brukt i populærkulturen. Selv om de kan bringe nye dimensjoner til underholdning, er det også viktig å tenke på de negative konsekvensene som kan oppstå, som jeg har diskutert gjennom min oppgave.

Mitt hovedspørsmål for min bacheloroppgave er hva er deepfakes, og hvilke etiske konsekvenser kan forekomme rundt denne teknologien, med spesielt fokus på populærkulturen. Gjennom denne oppgaven har jeg drøftet frem hvilke etiske konsekvenser som kan forekomme av bruken av denne teknologien. Gjennom min undersøkelse av denne teknologien, og analyse av bruken ved teknologien vil jeg si at man burde stille seg kritisk til det man ser på digitale medier, spesielt med tanke på utviklingen digitale medier har hatt ved årene. Man burde stille seg kritisk til det vi ser på nett, og det er viktig å undersøke og lære seg om endringene som skjer i den digitale verden for mediene har mye større makt enn det vi selv tror.

8. KILDER

Buzzfeedvideo. (2018, 17. april). *You won't believe what Obama says in this video!* [Video]. Youtube. https://youtu.be/cQ54GDm1eL0?si=ecH1DN_szSGHuLjj

Cuarzo producciones (produsent). (2023). *Falso Amor* (reality-serie). Netflix. <https://www.netflix.com/title/81610979>

Fetveit, A. (2004). Reality TV in the digital era: a paradox in visual culture? In *The Television studies reader* (pp. 543–556).

Fraga-Lamas, P., & Fernandez-Carames, T. M. (2020). Fake News, Disinformation, and Deepfakes: Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality. *IEEE IT Professional*. 22(2), 53-59. <https://doi.org/10.1109/MITP.2020.2977589>

Giansiracusa, N. (2021). *How Algorithms Create and Prevent Fake News: Exploring the Impacts of Social Media, Deepfakes, GPT-3, and More*. Apress. <https://doi.org/10.1007/978-1-4842-7155-1>

Hausken, L. (2024). Photorealism versus photography. AI-generated depiction in the age of visual disinformation. *Journal of aesthetics & culture*, 16(1), 1-13. <https://doi.org/10.1080/20004214.2024.2340787>

Ingemann, I. (1998). *Fatamorgana - pressefotografiets virkelighed og læsernes* (1.utg.). museum Tuscalanums forlag.

Kavka, M. (2012). *Reality TV* (1.utg). Edinburgh University Press. <https://doi.org/10.1515/9780748637249>

Langguth, J., Pogorelov, K., Brenner, S., Filkuková, P., & Schroeder, D. T. (2021). Don't Trust Your Eyes: Image Manipulation in the Age of DeepFakes. *Frontiers in Communication*, 6(632317), 1-12. <https://doi.org/10.3389/fcomm.2021.632317>

Manovich, L. (2001). *The language of new media*. MIT press

Solberg, K., Fjeld, J., Bruun, A., Wiese, D., Dorholt, I. (2024, 04. mai). Russ som liker kokain og horer. *Dagbladet*. <https://www.dagbladet.no/nyheter/russ-som-liket-kokain-og-horer/81354914>

Figur 1

Solberg, K., Fjeld, J., Bruun, A., Wiese, D., Dorholt, I. (2024, 04. mai). Russ som liker kokain og horer. *Dagbladet*. <https://www.dagbladet.no/nyheter/russ-som-liket-kokain-og-horer/81354914>

Figur 2

Buzzfeedvideo. (2018, 17. april). *You won't believe what Obama says in this video!* [Video]. Youtube. https://youtu.be/cQ54GDm1eL0?si=ecH1DN_szSGHuLjj

Figur 3

Ingemann, I. (1998). *Fatamorgana - pressefotografiets virkelighet og læsernes* (1.utg.). museum Tuscalanums forlag.

Figur 4

Cuarzo producciones (produsent). (2023). *Falso Amor* (reality-serie). Netflix. <https://www.netflix.com/title/81610979>

