

Deepfake – syvät valevideot

Pauliina Mattila

Haaga-Helia ammattikorkeakoulu

Amk-opinnäytetyö

2021

Tradenomin tutkinto

Tiivistelmä

Tekijä

Pauliina Mattila

Tutkinto

Tradenomi, tietojenkäsittelyn koulutusohjelma

Opinnäytetyön nimi

Deepfake – syvät valevideot

Sivu- ja liitesivumäärä

48 + 2

Valeutisista, informaationsodasta, ja tavallisten ihmisten vaikeuksista tunnistaa vaikuttamisyriä, sekä mielipiteiden ohjailusta kertovat kirjoitukset yleistyivät viime vuosina mediassa. Myös deepfake-, eli syvät valevideot nousivat pinnalle uutisaiheena ja herättivät tekijän kiinnostuksen.

Termi deepfake koostuu englannin sanoista fake, eli valheellinen tai väärä, ja deep, joka viittaa keinoälyn deep learning- eli syväoppimismetodiin. Syväoppiminen on koneoppimista, jolla kerrytetään informaatiota kerroksittain, mallintaen aivojen tapaa prosessoida tietoa. Syvien valevideoiden taustalla oleva teknologia hyödyntää koneoppimisen ohjaamattoman oppimisen alaryhmää. Ohjaamattomassa oppimisessa ohjelmalle syötettävä data on luokittelematonta, joten ohjelma luokittelee sen samankaltaisuuksiin perustuen ja tunnistaen poikkeavuuksia suuresta määrästä dataa. Syvän valevideon tekemisessä koneoppiva neuroverkko muodostaa materiaalin pohjalta pysyvien kasvopiirteiden mallin, joka upotetaan toiseen videoon.

Keskeiseksi tutkimuskysymykseksi nousi, olisiko kenen tahansa mahdollista tehdä valevideo. Tekijä lähestyi aihetta tapaustutkimuksen keinoin, tavoitteena valevideoteknologian syvälinen ymmärtäminen käytännön kokeilujen myötä. Testattavaksi valikoitui useita eri vaikeusasteen valevideon luomistekniikoita kattavan kuvan luomiseksi.

Tekniikat rajattiin avoimesta internetistä löytyviin. Opinnäytetyössä keskityttiin ennalta luotaviin kasvojen alueen valevideoihin rajaten ulos reaaliajassa tehtävä kasvojen vaihto esimerkiksi videopuhelussa. Maailmanlaajuisen Covid-19 pandemian ollessa käynnissä keväällä 2021 opinnäytteen työstämisen aikaan kokeilu rajautui miessukupuoleen, sillä työskentely lähipiirin vapaaehtoisen mieshenkilön kanssa ei lisännyt koronavirusesitystä.

Useissa kokeilluista teknologioista oli teknisiä rajoitteita, jotka vähensivät tekniikan uskottavaa käyttöä. Useimmat puhelinsovellukset mahdollistivat lähinnä hupikäytön, joskin yhtä sovellusta olisi voinut käyttää valevideon pohjana, mikäli videon jokaisen ruudun olisi käsitellyt sovelluksella erikseen. Internetissä tarjolla oleva valevideoiden luomispalvelu tarjosi uskottavan vaihtoehdon, joskin lopputuloksen kontrollointi oli käyttöliittymän avulla hankalaa. Äänen kloonaaminen ja valokuvan animoiminen oli työläs tekniikka, joka oli käytettävissä vain suoraan edestä kuvattuihin videoihin.

Käyttökelpoisin teknologia oli vaikeammasta päästä käyttää, mutta silti aloittelijan opittavissa. Ohjelman avulla oli mahdollista luoda uskottava valevideo, tosin prosessi oli aikaa vievä, ja vaati runsaasti laadukasta kuvamateriaalia lähtökohdaksi.

Asiasanat

Valemedia, koneoppiminen, tekoäly, tapaustutkimus, video

Sisällys

1	Johdanto	1
2	Syvien valevideoiden taustaa.....	4
2.1	Syvien valevideoiden teknologiaa – koneoppiminen	4
2.2	Syvien valevideoiden teknologiaa – neuroverkot.....	5
2.3	Syvien valevideoiden teknologiaa – autoenkooderi	5
2.4	Syvien valevideoiden tunnistuskeinoja	7
2.5	Syvien valevideoiden riskit.....	9
2.6	Syvien valevideoiden hyödyt	10
2.7	Syvien valevideoiden ohjelmakarttoitus	11
3	Valevideoiden luominen käytännössä	14
3.1	Tutkimuksen lähtökohdat.....	14
3.2	Puhelinsovellukset	14
3.2.1	Reface.....	15
3.2.2	iFace	16
3.2.3	Avatarify	17
3.2.4	Faceapp.....	18
3.3	DeepFakes Web	19
3.4	Kuvan animoiminen, äänen kloonaaaminen	23
3.4.1	Testissä käytettäviin ohjelmiin tutustuminen.....	23
3.4.2	Äänen kloonaus	24
3.4.3	Videon animointi äänitiedoston pohjalta	26
3.5	DeepFaceLab	27
3.5.1	Ohjelmaan tutustuminen	28
3.5.2	Ohjelman käyttäminen paikallisesti	29
3.5.3	Mallin koulutus Google Colabissa	34
3.5.4	Google Colabin käyttörajoitusten aiheuttamat työskentelyvaikeudet.....	36
3.5.5	Kuvien yhdistäminen videoksi paikallisella koneella.....	37
4	Pohdintaa	41
	Lähteet	44
	Liitteet.....	49
	Liite 1. Deepfakesweb kierrosten 1–4 tulokset.....	49
	Liite 2. DeepFaceLab-ohjelmalla tuotetun valevideon still-kuvia.....	50

1 Johdanto

Hiljattain katsoin Yle Areenasta ranskalaista agenttisarjaa, jossa venäläiset kybersotilaat tekivät käden käänteessä videoita, joissa esiintyvä henkilö ei oikeasti ollut osallistunut millään tavalla videon tekoon. Tekniikan avulla pystyi myös tekemään uskottavan videopuhelun, kohdehenkilön tajuamatta huijausta. Tuomasin kyseessä olevan vain sarjan tekijöiden omat fantasiat, kunnes törmäsin sanomalehdessä aiheeseen.

Kyseessä olivat syvät valevideot, eli deepfake-videot, joita on jo tehty hyödyntäen tekoälyä taustatekniikkana. Viime aikoina on enenevässä määrin keskusteltu valeuutisista, informaationsodasta, sekä tavallisten ihmisten vaikeuksista tunnistaa vaikuttamisyrityksiä ja mielipiteiden ohjailua. Jos pelkästään kirjallisen tekstin tunnistaminen vääristellyksi on vaikeaa, miten on videoiden laita? Voisiko jo nyt kuka tahansa tehdä valevideon, vai tarvitaanko siihen edistynyttä teknologiaa, joka ei ole kaikkien saatavilla?

Perehtyessäni aiheeseen lisää, esiin tuli tulevaisuuden uhkakuvia valtioiden kybersodan lisäksi tavallisiin ihmisiin vaikuttavista riskeistä, kuten väärennetyn intiimin materiaalin levittämisestä vahingoittamistarkoituksessa, ja huoltajuuskiistoihin vaikuttamisesta. Toisaalta myös aitoja videoita voitaisiin myöhemmin asettaa kyseenalaisiksi väittämällä niiden olevan väärennettyjä, ja siten murentaa tiedon luotettavuutta. Myös positiivisia esimerkkejä valevideoteknologian hyödyntämisestä on – esimerkkinä elokuvat, joissa tähti näyttelijöitä on nuorennettu, tai museo, jossa kuuluisa taiteilija voidaan ”herättää henkiin” kyseisen teknologian avulla.

Aihe herätti mielessäni monia kysymyksiä. Onko valevideoita ollut jo aiemmin? Missä valevideoita tehdään, ja mitkä ovat motiivit niiden tekemiseen? Mitä teknologioita syvän valevideon tekeminen vaatii, ja toisaalta, miten väärennetyn valevideon voisi tunnistaa? Miten voisi itse välttyä menemästä lankaan? Voisiko kuka tahansa tehdä valevideon?

Selvittääkseni vastaukset kysymyksiini päätin tutkia valevideoiden taustaa, käytettyjä teknologioita ja tavallisten ihmisten mahdollisuuksia tehdä itse valevideoita tutkien kirjallisia lähteitä. Näiden pohjalta päätin kokeilla useita eri teknologioita selvittääkseni mitä taitoja valevideon luominen käytännössä vaatii, ja millaista teknologiaa tällä hetkellä on käytettävissä. Tutkin myös, onko aloittelijan helppoa luoda uskottava valevideo, ja tästä tulikin opinnäytteeni keskeisin tutkimuskysymys.

Itselläni ei ollut entuudestaan kokemusta valevideoiden luomiseen tarkoitetuista ohjelmista tai videoeditointityökaluista. Vaikka olenkin opiskellut ohjelmointikieliä, en ole opiskellut näissä ohjelmissa usein käytettävää Python-ohjelmointikieltä. Linux-

käyttöjärjestelmä oli minulle tuttu, samoin sen käyttäminen komentokehotteesta graafisen käyttöliittymän lisäksi. Pilvipalveluista minulla oli vain alkeellista käyttökokemusta, mutta rajautuen vain Amazon Web Servicesiin, eli Googlen pilvipalvelua en ollut kokeillut aiemmin.

Rajasin teknologiat niihin, joita löytää avoimesta internetistä. En tutkinut esimerkiksi, millaisia valevideopalveluja pimeässä verkossa (TOR) mahdollisesti tarjotaan. Keskityin myös työssäni kasvojen alueen valevideoihin, sillä vaikka liikkeen kloonausohjelmia esimerkiksi tanssiliikkeiden kopioimiseksi on olemassa, esimerkkinä Everybody Dance Now (Chan, Ginosar, Zhou & Efros 2019), niiden mukaan ottaminen laajentaisi aihetta tarpeetomasti. Keskityin myös valevideoihin, jotka voi luoda etukäteen, enkä tässä työssä tutkinut reaaliajassa tehtävää kasvojen vaihtoa esimerkiksi videopuhelussa. Vaikka kyseisenlainen käytötapaus innoittikin opinnäytteen aiheeseen, oletin reaaliaikaisen valevideon vaativan kokemusta valevideoiden luomisesta ylipäättään. Maailmanlaajuisen Covid-19 pandemian ollessa käynnissä opinnäytteeni työstämisen aikaan, rajautui kokeeni miessukupuoleen, sillä lähipiirissäni oli vapaaehtoinen mieshenkilö, jonka kanssa työskentely ei lisännyt koronavirusaltistusta.

Valevideo vaikuttaa aiheena melko haastavalta, johtuen taustalla olevasta tekoälyteknologiasta. Aihe voi olla teknisyydessään luotaantyöntävä. Itsekin epäilin ennen testeihin ryhtymistä, riittäisivätkö rahkeeni. Toisaalta valevideoiden laajempi leviäminen tekee tarpeelliseksi ymmärtää ilmiötä ja sen taustalla olevaa teknologiaa. Tiedon vääristely voi vaikuttaa meistä jokaiseen, joten pyrin tässä työssäni myös avaamaan valevideoiden tekniikkaa helpommin ymmärrettävään muotoon.

Keskeisiä käsitteitä:**Valevideo, deepfake**

Tekoälyn avulla tehty väärennetty video, jota on ihmissilmin erittäin vaikeaa, ellei mahdollisuutta erottaa aidosta videosta (Sipilä 11.10.2020).

Keinoäly, tekoäly

Tietokoneohjelma, joka kykenee tekemään toimintoja, joita voi kuvata älykkäiksi, kuten massiivisen laskentatehon avulla suoritettavat laajojen tietoaisteiden analyysit (Aamulehti 16.1.2017).

Algoritmi

Yksityiskohtaiset ohjeet siitä, miten jokin tietty tehtävä suoritetaan. Usein esimerkkinä käytetään keittokirjan reseptiä, joka voidaan ymmärtää algoritmina, eli suoritushyöhenä.

Koneoppiminen

Käytännön esimerkkejä koneoppimisesta ovat puheentunnistus, konenäkö, hakusuositukset sekä sähköpostien suodatus (Microsoft News Center 2018). Koneoppimisessa tietokone rakentaa mallin siitä, miten tietynlaisesta alkulähtökohdasta päästään tiettyyn lopputulokseen. Ohjelmoija ei siis rakenna logiikkaa, jolla lähtökohdasta A päästään lopputulokseen B, vaan tietokone mallintaa eli tavallaan oppii datan perusteella logiikan. (Arjoranta 2019.)

Syväoppiminen

Syväoppiminen on koneoppimista, jolla pyritään mallintamaan aivojen tapaa prosessoida tietoa. Tässä menetelmässä neuroverkko lisää attribuutteja ja informaatiota kerros kerrokselta vastaanottamaansa tietoon, pyrkien muodostamaan datasta yksityiskohtaisen näkemyksen. (Microsoft News Center, 2018.) Prosessi muistuttaa ihmisen oppimista, jossa riittävän usein erilaisia koiria nähnyt rakentaa mallin päässään siitä, millaisia koirat ylipäättään ovat.

Neuroverkko, neural network

Neuroverkko jäljittelee ihmisen aivojen toimintaa, eli sitä, miten joukko neuroneita, eli tiedonkäsittelijöitä, ja synapseja, eli viestinvälittäjiä yhteen kytkettynä kykenevät hyvinkin monimutkaisiin toimintoihin. Neuronit prosessoivat tietoa toisistaan riippumatta, kykenevät sekä tiedon tallentamiseen, että prosessointiin, ja mukauttavat ajan kuluessa tapaansa reagoida saamiinsa signaaleihin. (Reaktor & Helsingin Yliopisto 2018.)

2 Syvien valevideoiden taustaa

Deepfake-, eli valevideot ovat hyperrealistisia keinoälyn avulla luotuja videoita, joissa kuvattu henkilö sanoo ja tekee asioita, joita ei todellisuudessa ole tapahtunut. Kyseiset videot leviävät sosiaalisen median alustoilla nopeasti, ja saavuttavat suuren määrän ihmisiä, joten niiden negatiivinen vaikutus yhteiskuntaan saattaa olla merkittävä. (Westerlund 2019, 39.)

Alkuun syviä valevideoita tehtiin aikuisviihdevideoista vaihtamalla tunnetun julkisuuden henkilön pää aikuisviihdenäyttelijän ruumiiseen. Ensimmäiset kuvatun kaltaiset videot julkaistiin vuonna 2017 Reddit-foorumille, mutta sittemmin videoita on käytetty myös geopoliittikan välineenä. (Sipilä 11.10.2020; Westerlund 2019, 40.) Ainakin neljä eri ryhmää on tunnustettu valevideomateriaalin tuottajiksi: deepfake-harrastajat, poliittiset toimijat kuten valtiot tai aktivistit, pahantahtoisiin tarkoituksiin valevideoita tehtailevat, sekä tekniikkaa hyviin tarkoituksiin käyttävät toimijat, kuten televisioyhtiöt (Westerlund 2019, 41).

Syvän valevideon tekeminen tavallisella tietokoneella saattaa olla haastavaa. Useimmat valevideot on luotu käyttäen suorituskykyisiä grafiikkakortteja tehokkailta tietokoneilla, tai hyödyntämällä pilvipalvelimien laskentavoimaa. Luotuja videoita saattaa vielä joutua muokkaamaan poistaakseen tietokoneen luomia pieniä virheitä, ja tämä prosessi vaatii ammattitaitoa. Kuitenkin kaikkien saataville on tuotu myös sovelluksia, kuten mobiiliapplikaatio Zao, jossa käyttäjät voivat vaihtaa kasvonsa tunnettujen näyttelijöiden päälle, sekä avoimen lähdekoodin ohjelmia, jotka tarjoavat vähemmillä teknisillä taidoilla varustetulle henkilölle mahdollisuuden tehdä syviä valevideoita. (Sample 13.1.2020; Westerlund 2019, 40.)

2.1 Syvien valevideoiden teknologiaa – koneoppiminen

Termi deepfake tulee englannin sanoista fake, eli valheellinen tai väärä, ja deep, joka viittaa keinoälyn deep learning- eli syväoppimismetodiin (Aitamurto 13.11.2018). Tekoälyllä tarkoitetaan mitä hyvänsä ohjelmaa tai järjestelmää, joka pystyy tekemään jotain älykästä. Tekoäly käsitteenä kattaa esimerkiksi kuvien ja luonnollisen kielen prosessoinnin, sekä koneoppimisen. Tekoäly voidaan jakaa vahvaan ja heikkoon, jolloin heikolla tarkoitetaan ohjelmaa, jonka on mahdollista toimia sille etukäteen määritellyssä rajoitetussa ympäristössä. Vahvalla tekoälyllä on kyky ajatella, ratkoa ongelmia, tehdä päätöksiä, oppia sekä kommunikoida. (Arjoranta 2019.)

Koneoppiminen voidaan jakaa kolmeen eri ryhmään, joita ovat ohjattu oppiminen, ohjaimaton oppiminen, sekä vahvistusoppiminen. Ohjatussa oppimisessa aineistona käytetään

valmiiksi luokiteltua materiaalia, joka syötetään koneoppimisohjelmalle, tavoitteena opettaa ohjelma luokitteluun sille annetun datan perusteella uusia syötteitä. Ohjaamattomassa oppimisessä ohjelmalle syötettävää dataa ei ole luokiteltu, joten ohjelma jakaa dataa luokkiin perustuen samankaltaisuuksiin ja tunnistaa poikkeavuuksia suuresta määrästä dataa. Ohjaamattoman oppimisen etuna on, että oppimista ei ohjaa luokittelun painotukset. Vahvistusoppimisessa ohjelma saa positiivista tai negatiivista palautetta toimintansa seurauksena, ja pyrkii saamaan mahdollisimman paljon positiivista palautetta. (Lanham 2021, luku 2; Vaiste 2017.)

2.2 Syvien valevideoiden teknologiaa – neuroverkot

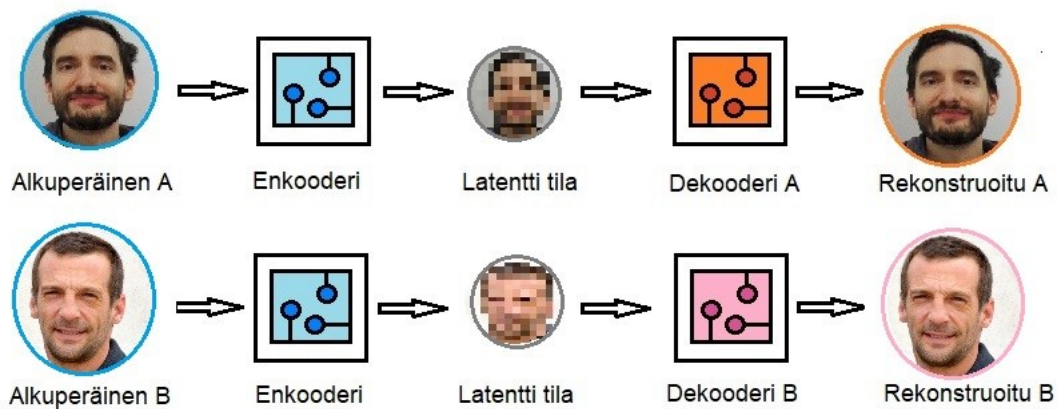
Syväoppimisessä jäljitellään ihmisaivojen tapaa prosessoida tietoa. Tällöin käytetään keinotekoisia neuroverkkoja, joissa kerros kerrokselta kasvatetaan järjestelmän informaatiota opittavasta aiheesta. (Microsoft News Center 2018.) Vaikka syväoppiminen yhdistetään aivoihin, saattaa tämä miellelyhtymä aiheuttaa epävarmuutta ja mystifioida aihetta liikaa. Aivojen hermotoiminta on huomattavasti syväoppimista monimutkaisempaa – ennemminkin aivoista on otettu syväoppimiseen malli yhteyksien luomisesta. (Lanham 2021, luku 1.)

Neuroverkkoteknologian avulla voidaan luoda esimerkiksi kuvia ihmisistä, joita ei ole olemassa. Kyseisessä koneoppimismallissa, GANissa eli generatiivisessa kilpailevassa verkossa (Generative Adversarial Network), kaksi neuroverkkoa sparraa toisiaan kehittämään yhä parempia kuvia ensimmäisen luodessa kuvia, ja toisen pyrkiessä erottamaan aidon väärennöksestä. Usein käytetty analogia on taiteen tuntija ja taiteen väärentäjä. Taiteen tuntija, eli diskriminaattori opetetaan tunnistamaan aidot teokset, ja väärentäjä, eli generaattori oppii luomalla väärennöksiä, ja tarjoamalla niitä taiteen tuntijalle. Mikäli diskriminaattori luulee oikeaa kuvaa väärennökseksi, asettaa se itselleen korkeamman häviöprosentin. Mikäli generaattori saa väärennetyn kuvan menemään läpi aitona, se antaa diskriminaattorille korkeamman häviölukeman, ja itselleen matalamman. Molemmat järjestelmät oppivat samanaikaisesti tulemaan paremmiksi joko uusien kuvien luomisessa tai kuvien oikeiksi tai väärennöksiksi tunnistamisessa. (Lanham 2021, luku 2; Niemi 2.3.2020.)

2.3 Syvien valevideoiden teknologiaa – autoenkooderi

Syvän valevideon tekemisessä koneoppiva neuroverkko muodostaa aidon materiaalin pohjalta mallin, joka upotetaan toiseen videoon (Ahokas 2020). Tätä järjestelmää – jota kutsutaan autoenkooderiksi – voisi kuvailla kyljelleen käännetyksi tiimalasiksi, jossa toisessa päässä verkostoa on enkooderi, ja toisessa dekooderi, joiden välissä on latentti tila (Lee 2019). Autoenkooderin käyttäminen on ohjaamatonta oppimista, jossa kone pystyy purkamaan datan alempaan muotoon ja rakentamaan sen uudelleen (Lanham 2021, luku 2).

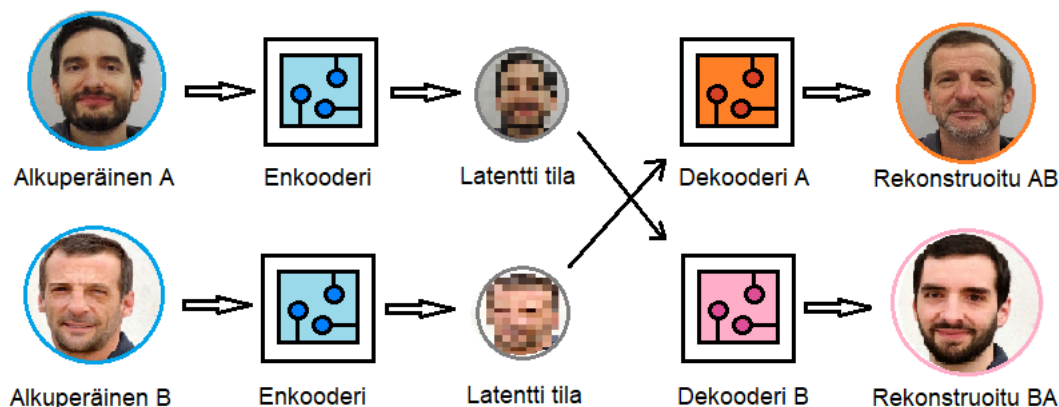
Ensin järjestelmään syötetään tuhansia kasvokuvia kahdesta henkilöstä tekoälyalgoritmin, enkooderin analysoitavaksi. Enkooderi etsii ja oppii samankaltaisuudet kuvien välillä, ja pelkistää kuvat jaettuihin yhteisiin ominaisuuksiin. (Sample 13.1.2020.) Enkooderin voisi-kin ajatella olevan häviöllinen pakkausalgoritmi, joka yrittää tiivistää mahdollisimman paljon informaatiota kasvoista rajoitettuun varastointitilaan, jota kutsutaan latentiksi tilaksi. Tärkeintä enkooderin osalta prosessissa on tunnistaa kasvoista ajan mittaan vaihtuvat ominaisuudet, kuten katseen suunta, jotta varastotilaa ei käytettäisi turhaan pysyvien ominaisuuksien, kuten silmien värin tallentamiseen. (Lee 2019.)



Kuva 1. Kuvan enkoodaus ja dekoodaus dekodeerilla (mukaillen Biard, G. 2017 CC BY-SA 3.0; Lee, T. 2019)

Seuraavaksi toinen algoritmi, dekodeeri, opetetaan löytämään henkilön pysyvät kasvopiirteet enkooderin läpikäymästä datasta. Käyttämällä kahta dekodeeria saadaan löydettyä molempien henkilöiden kasvopiirteet. (Sample 13.1.2020.) Koska molempia kasvokuvia enkoodatessa neuroverkot jakavat enkooderin parametreja, on kumpi tahansa kasvo mahdollista rekonstruoida dekodeerin avulla (Lee 2019). Kuvassa 1 on kuvattu prosessi enkoodauksesta dekodeeraukseen pyrkien luomaan lopuksi alkuperäiseen kuvaan pohjautuva rekonstruoitu kuva.

Kasvojen vaihto suoritetaan syöttämällä enkoodatut kuvat "väärälle" dekodeerille (Sample 13.1.2020). Esimerkiksi siis enkooderin pelkistämä kasvokuva henkilöstä A syötetään dekodeerille, joka on opetettu hakemaan enkooderin datasta henkilön B piirteet. Dekodeeri rekonstruoi datasta henkilö B:n kasvopiirteet henkilö A:n ilmeiden ja liikkeiden päälle. (Sample 13.1.2020.) Tämä prosessi on havainnollistettu kuvassa 2.



Kuva 2. Kuvan enkoodaus ja dekoodaus "väärällä" dekooderilla (mukaillen Biard, G. 2017 CC BY-SA 3.0; Lee, T. 2019)

Autoenkooderit käyttävät lopputuloksen arviointiin valvomatonta koneoppimista, sillä tavoitteena on tuottaa mahdollisimman identtinen jäljennös neuroverkkoon annetusta lähtömateriaalista, jolloin arviointi tapahtuu ilman ihmisen apua vertaamalla lopputulemaa annettuun alkuperäiseen syötteeseen pikseli pikseliltä. Virhepikselien, eli häviön osuus ilmoitetaan lukuna. Järjestelmä jäljittää huonolaatuisten pikselien tuottamat neuronit, ja muokkaa parametreja hiukan seuraavaa iteraatiokierrosta varten. (Lanham 20221, luku 2; Lee 2019.)

2.4 Syvien valevideoiden tunnistuskeinoja

Käynnissä on teknologinen kilpajuoksu, kissa ja hiiri -leikki, jossa pyritään tunnistamaan valevideoita samalla teknologialla, jolla ne luotiin (Sipilä 11.10.2020). Vuonna 2018 yhdysvaltalaiset tutkijat huomasivat, että valevideoissa ihmiset eivät räpytelleet silmiään normaalisti. Tekoälylle syötetyistä kuvista valtaosassa ihmisillä oli nimittäin silmät auki. Melkopian löydön jälkeen alkoi kuitenkin ilmaantua valevideoita, joissa ihmiset räpytelivät silmiään. (Sample 13.1.2020.)

Oli kyse sitten audiosta tai videosta, valevideon tunnistamiseen keinoälyn avulla tarvitaan kaksi datajoukkoa – aito ja väärennety. Väärennetyn puheen tunnistamiseksi tilastollinen malli laitetaan oppimaan datajoukkojen eroavaisuudet. Tämä tosin toimii vain, mikäli väärennöstapa on tiedossa, ja osana väärennetyn puheen datajoukkoa. Täysin uuden tekniikan avulla tehty valheellinen tuotos olisi huomattavasti vaikeampi tunnistaa, sillä "tunte mattoman tunnistaminen on pirullisen vaikeaa". (Ahokas 2020.)

Valevideon tunnistaminen eroaa valheellisen audion tunnistamisesta. Videokuva koostuu useista peräkkäisistä ruuduista, jotka nopeassa tahdissa peräkkäin esitettynä muodostavat liikkuvan kuvan. Valevideo rakennetaan myös yksi ruutu kerrallaan, mutta jokainen ruutu perustuu havaintoon alkuperäisestä kuvasta, ottamatta huomioon, millainen videon ruutusarjan aiempi ruutu on ollut. Kun yksittäiset kuvat esitetään videossa peräkkäin, ihmisilmä saattaa huomata ruutujen epäyhtenäisyyden ja näin tunnistaa videon peukaloiksi. (Ahokas 2020.)

Videoissa saattaa olla pieniä yksityiskohtia pielessä: puheen synkkaus saattaa olla heikkolaatuinen, ihonsävy vaihtelee, valaistus tai silmien iiristen heijastukset eivät ole luonnolliset tai yhtenäiset (Sample 13.1.2020). Staattisissa kuvissa tekoälyn luomat ihmiskasvot voivat olla hämäävän aidon näköiset, sillä ihmiskasvot ovat melko säännölliset, mutta kuvien taustassa on enemmän vaihtelua. Vaikka kasvoissa ei siis näkyisikään merkkejä keinoälyn tuottamista pienistä virheistä, saattaa kuvan tausta paljastaa sen feikiksi. (Niemi 2.3.2020.) Syvän valevideon erottamista aidosta voisi vaikeuttaa tekemällä kuvasta tarkoituksella rakeisen, valvontakameravideon tyyllisen, jolloin kuvassa näkyvät, yleensä syvän valevideon tunnistamista helpottavat pienet virheet eivät pistäisi silmään (Ahokas 2020).

Pitkän koulutuksen suoma medialukutaito saattaa edesauttaa valevideoiden tunnistamista sivistysvaltiossa kuten Suomessa, mutta tilanne saattaa olla toinen maissa, joissa kansa ei ole yhtä valveutunutta (Ahokas 2020). Köyhän ja kouluttamattoman, mutta älypuhelimien varustetun kansalaisen heikommat valmiudet kyseenalaistaa näkemäänsä luo otollisen pohjan poliittisen disinformaation levittämiseksi (Sipilä 11.10.2020).

Jotta syvien valevideoiden tunnistaminen ei olisi vain ihmisten oman tarkkaavaisuuden varassa, yhteisöpalvelu Facebook lanseerasi vuonna 2019 valevideoiden tunnistuskilpailun. Ajatuksena oli, että samaan tapaan kuin on helppokäyttöisiä valevideoiden luomisohjelmia, tulisi olla myös helppokäyttöisiä valevideoiden tunnistusohjelmia, jotta tunnistamisprosessia voisi automatisoida. Vaikka kilpailun tuloksena korkein tunnistamisprosentti olikin vain 65 %, se oli kuitenkin arvausta parempi, ja tulokset saattavat parantua tulevaisuudessa. (Coldewey 12.6.2021.)

Vastuullisesti toimivat yritykset, jotka tarjoavat deepfake-tekniologiapalveluita, lisäävät tuotteisiinsa merkkejä, joilla ne voi tunnistaa. Esimerkiksi Modulate-yritys tarjoaa äänen kloonauspalveluita pääasiassa peliyrityksille. Modulate lisää audioon tunnistettavan audiosormenjäljen, jota ei voi ihmiskorvin kuulla, mutta jonka voi nähdä tutkiessa ääniaallon muotoa estääkseen petokset. (Wilson 2019.)

2.5 Syvien valevideoiden riskit

Syvät valevideot uhkaavat yhteiskuntaamme, koska ne hankaloittavat journalistien työtä erottaen oikeat ja valeutiset toisistaan, uhkaavat kansallista turvallisuutta mahdollistaen propagandan levittämisen ja vaaleihin vaikuttamisen, nakertavat kansalaisten uskoa viralliseen informaatioon ja muodostavat yksityishenkilöihin sekä organisaatioihin kohdistuvia kyberturvauhkia (Westerlund 2019, 42). Valheelliset videot yritysjohtajista tai pankkiireista voisivat vaikuttaa osakekursseihin tai yrityksen maineeseen tehden yritykselle taloudellista haittaa (Venkataramakrishnan 2019).

Valevideoteknologia saattaa mahdollistaa aikuisen esiintymisen lapsena, tai vihaisen entisen puolison esiintymisen uutena kumppaniehdokkaana (Wilson 2019). Teknologian avulla on jo luotu ohjelma, joka otti lähtökohdaksi valokuvan naisesta pukeissa, ja muokkasi kuvasta vaatteet pois. Vaikka alaston vartalo ei ole kuvan naisen oikea vartalo, se on riittävän lähellä, jotta ohjelmaa voisi käyttää esimerkiksi vahingoittamistarkoituksessa naisen nöyryyttämiseen. Ohjelma lisäsi pienen vesileiman kuvan nurkkaan, mutta se oli helposti editoitavissa pois. Maksuton versio myös sensuroi kriittiset kohdat, mutta 50 dollarin maksua vastaan sai täysin alastoman version. (Cole 2019a.) Ohjelma on sittemmin otettu pois jakelusta sen saatua odotettua enemmän latauksia. Ohjelman kehittäjä kertoi olevansa teknologiaentusiasti, ja ohjelman saaneen alkunsa uteliaisuudesta ja halusta oppia. Kehittäjä kertoi tajunneensa, ettei halukaan tehdä rahaa keksimällään ohjelmalla. (Cole 2019b.)

Vaikka valevideot ovat melko uusi ilmiö, oli vuonna 2020 internetissä jo arviolta 200 000 taitavasti väärennettyä videota, ja videoiden määrän arveltiin tuolloin kaksinkertaistuvan puolen vuoden välein. Vuonna 2025 jopa 90 prosenttia videoista netissä saattaa olla väärennöksiä, jolloin ihmisten luottamus länsimaiseen demokratiaan ja tiedonvälitykseen mahdollisesti rapautuu. (Sipilä 11.10.2020.) Toki ihmisten manipulointia on aiemminkin tehty vääristämällä lähdemateriaalia, esimerkiksi lavastamalla kuvitteellisia tilanteita, editoimalla videoita esimerkiksi leikkaamalla, muokkaamalla toiston nopeutta nopeammaksi tai hitaammaksi, tai irrottamalla osan videosta asiayhteydestään. (Ahokas 2020; Sample 13.1.2020.) Tunnettu lähihistorian esimerkki videon toiston nopeutta hidastamalla tehdystä vääristelystä videosta on Yhdysvaltojen Nancy Pelosin puhe vuodelta 2019, jonka Pelosin vastustajat uskottelivat todistavan hänen kyseenalaisesta henkisestä tilastaan (Venkataramakrishnan 2019).

Vuonna 2021 Yhdysvalloissa teini-ikäisen cheerleadertiimin jäsenen äiti teki toisista tiimin jäsenistä väitetyksi valevideoteknologian avulla kuvia ja videoita, joissa nämä näyttivät juovan alkoholia ja polttavan tupakkatuotteita saadakseen kilpailevat tytöt poistettua tiimistä

sääntöjen rikkomisen perusteella. Kyseisen äidin väitettiin ottaneen kuvat tiimin jäsenten sosiaalisen median tileiltä, ja muokanneen ne valevideoteknologialla. (Vella 13.3.2021.) Myöhemmin oikeudessa syyttäjät poistivat valevideo-osuuden syytelistalta, sillä vaikka yhdestä videosta oli mahdollista löytää jälkiä valevideoteknologian käytöstä, poliisi ei voinut varmuudella osoittaa videotodisteiden olevan syviä valevideoita. Yhdestä videosta puuttui täysin pienet virheet eli artefaktit silmien tai kasvojen ympäristössä, joita usein syntyy valevideon luomisprosessissa, ja videossa oli fotorealistisia yksityiskohtia, kuten sähkötupakan savua, jota olisi erittäin vaikea luoda keinotekoisesti. Kyseinen tapaus saattaakin olla esimerkki tilanteesta, jossa aidon videon uskottavuus yritetään kumota väittämällä sen olevan valevideo. (Harwell 14.5.2021.)

Tulevaisuudessa deepfake-videoiden uskottavuuden parantumista suurempi riski saattaa-kin olla aitojen videoiden uskottavuuden kärsiminen, sillä niiden voidaan väittää olevan myös valevideoita (Ahokas 2020). Tätä vastaan voisi yrittää taistella varustuen esimerkiksi kamerat ominaisuudella, joka lisää vesileiman, jolla kuvan voisi todistaa oikeaksi (Sipilä 11.10.2020). Toisaalta valeuutinen saattaa ehtiä levitä laajasti ennen vesileiman tutkimista ja tuotoksen todistamista aidoksi tai vääräksi (Wilson 2019). On myös ehdotettu lohkoketjuteknologian valjastamista pitämään kirjaa videoiden, kuvien ja audion alkuperästä, jotta näiden alkuperä ja muunnelmat voitaisiin tarkistaa (Sample 13.1.2020).

2.6 Syvien valevideoiden hyödyt

Deepfake-tekniikka ei ole aina pahasta. Käyttökelpoisia kohteita on esimerkiksi kloonata äänensä sairaudelle menettäneen henkilön ääni. Teknologia saattaa myös rikastaa museoelämystä, kuten Floridassa Dali-museossa, jossa deepfake-Salvador Dali kertoo taiteestaan ja esiintyy vierailijoiden selfiekuville (Lee, D. 2019). Toinen esimerkki jo edesmenneen kuuluisuuden herättämisestä henkiin on chatbot, eli tekoälyä hyödyntävä automaattinen asiakaspalvelija, jolle on annettu Albert Einsteinin olemus ja ääni (Lomas 2021).

Elokuville tekniikkaa voidaan käyttää esimerkiksi ”herättämällä henkiin” jo edesmennyt näyttelijä, tai parantaen dubbauksen laatua. (Sample 13.1.2020.) Myös animaatioissa riskinä on suositun ääninäyttelijän äänen menettäminen, esimerkiksi näyttelijän menehtyessä, joten äänen kloonaminen deepfake-teknologialla voisi vähentää tätä riskiä. Ongelmana on tosin, että tämänhetkinen teknologia ei kunnolla osaa tuoda kloonattuun ääneen tunnetilojen nyansseja. (Katwala 3.10.2021.)

Videopelimaailma voi myös hyötyä deepfake-teknologiasta. Esimerkiksi jalkapallovideopelissä kommentaattorit voisivat puhua pelin pelaajasta tämän nimellä, tai videopelihahmon

nimen voisi muuttaa mieleiseksi, ja saada kaikki pelin muut hahmot viittaamaan puheessaan hahmoon uudella nimellä. Videopelin äänimaailmaa voisi teknologian avulla muuttaa helposti, jotta dialogi ei olisi riippuvaista vain näyttelijöiden studiossa tuottamista lauseista. (Katwala 3.10.2021.) Modulate-niminen yritys käyttää äänen deepfake-kloonusteknologiaa, jonka avulla pelaaja voi puhua valitsemansa pelihahmon äänellä mahdollistaen immersivisemmän peliavataarin ja pelaamiskokemuksen ryhmäpeleissä (Wilson 2019).

2.7 Syvien valevideoiden ohjelmakartoitus

Kartoitin erilaisia valevideoiden luomisohjelmia. Lista ei ole kaikenkattava, vaikka pyrinkin mahdollisimman monta sisällyttämään listaukseen.

Zao on kiinalainen deepfake-videoiden luomiseen kehitetty ilmainen sovellus, joka muuttaa yksittäisen valokuvan ihmisestä osaksi elokuvaa. Sovellus on saatavilla ainoastaan Kiinassa. Zaon käyttäjän tulee ottaa valokuva itsestään puhelimen etukameralla, joten sovellusta ei voi käyttää valevideon luomiseen kuvattavan henkilön tietämättä siitä. Sovellus arvioi sille syötetyn valokuvan sopivuutta jo olemassa olevaan elokuvapätkien kirjastoon, ja antaa käyttäjälle ehdotuksia. (Gilbey 4.9.2019.) Sovellus on tunnettu esimerkki valevideoista, mutta sen rajoitettu saatavuus rajoittaa Zaon kokeilua käytännössä.

Länsimainen vastine Zaolle on iPhonelle ja Android-puhelimille saatavilla oleva Reface-sovellus. Zaon tavoin sillä on mahdollista upottaa omat kasvonsa osaksi tunnettua elokuvaa. Sovellus on Zaon tavoin tarkoitettu enemmänkin hupikäyttöön kuin pahantahtoiisiin tarkoitukseen. (Kärkkäinen 20.10.2020.)

Faceapp on tunnettu etenkin ominaisuudestaan, jolla voi vanhentaa tai nuorentaa käyttäjän kuvaa. Sovellusta voi käyttää ilmaiseksi matkapuhelimella joko Androidilla tai iPhonella, mutta osa ominaisuuksista vaatii maksua. (Peckham 2020.) Faceapp on aiheuttanut polemiikkia, sillä sovelluksen kehitystiimi juontaa juurensa Venäjälle, joten sovellukseen ladattujen kuvien yksityisyys on huolestuttanut käyttäjiä laajasti. Sovelluksen tietoja ei kuitenkaan käsitellä Venäjällä. (Mahdawi 2019.)

Avatarify on iPhonelle tehty sovellus, joka ottaa lähtökohdaksi ihmisen kasvokuvan, ja käyttämällä puhelimen selfiekameraa on mahdollista animoida ja kontrolloida kasvokuvassa esiintyvän ihmisen ilmeitä ja eleitä. Avatarifyn käyttöehdot kieltävät teknologian käyttämisen pahantahtoiisiin tarkoitukseen, mutta ei käytännössä valvo säännön noudattamista. Kuvia ei ladata pilvipalvelimille, vaan ne pysyvät käyttäjän puhelimesta. Avatarify

lisää myös näkymättömän vesileiman ohjelmalla luotuihin videoihin, mutta sen voi poistaa maksusta. (Fowler 25.3.2021.)

Jiggy on puhelimelle ladattava ohjelma, jolla voi vaihtaa kasvokuvan erilaisiin lyhyisiin videoihin, keskittyen liikkeen animointiin. Ohjelman avulla voi liittää kasvokuvaan erilaisia vartaloita, vaihtaa kuvan taustaa ja esimerkiksi tanssiliikkeitä luodakseen uniikin videon. (Sandusky 2021.) Ohjelma on saatavilla sekä Android, että iPhone-versiona.

DeepFaceLab on ohjelma, joka käyttää hyväkseen GAN-teknologiaa, jossa kaksi neuroverkkoa kilpailee tuottaakseen mahdollisimman uskottavan lopputuloksen. Ohjelman voi ladata Github-repositoriosta, eli julkaisuarkistosta, ja sitä voi käyttää tietokoneella joko Linux tai Windows-ympäristössä. (Chanon 2021.) DeepFacelabin kehittäjät olivat huolissaan syväoppimisteknologian käytöstä pahantahtoisin tarkoituksiin, ja kokivat avoimen lähdekoodin deepfakeprojektin olevan paras tapa opettaa sekä akateemikot että tavalliset kansalaiset tunnistamaan valevideot, sekä ymmärtämään kuinka teknologia toimii (Perov ym. 2020).

Faceswap on DeepFaceLabin tapaan tietokoneella käytettävä deepfake-ohjelma. Faceswapilla voi tehdä syvän valevideon alusta loppuun pilkkomalla ensin videot kehyksiin, tunnistamalla niistä kasvot, kouluttaen neuroverkon kuvilla, ja lopulta tuottamalla valevideon, jossa yhden henkilön kasvopiirteet, ja toisen eleet ja ilmeet on yhdistetty. (Lee 2019.)

Faceswap-GAN on jo hieman vanhentunut työkalu, jota ei aktiivisesti ole päivitetty viime vuosina (Lanham 2021, luku 9). Ohjelman voi ladata Github-repositoriosta, eli julkaisuarkistosta, ja siitä on olemassa Google Colab -versio.

CelebAMask-HQ on ohjelma, jolla voi muokata ja kopioida kasvojen piirteitä. Vaikka ohjelmalla ei voikaan tehdä täyttä kasvojen vaihtoa, sitä voi käyttää muiden työkalujen rinnalla kasvojen muokkaamisen apuvälineenä (Lanham 2021, luku 9).

StarGAN version 2 on ohjelma, jolla voi muokata tai vaihtaa kasvoja perustuen useisiin ohjelman oppimiin attribuutteihin (Lanham 2021, luku 9). Ohjelman Github-repositoriossa, eli julkaisuarkistossa, on esimerkkejä muokkauksista, joita ovat muun muassa hiusten tai ihon värin vaihto, sukupuolen tai hiustyylin vaihtaminen.

FSGAN on ohjelma, joka vaatii hyvää tietopohjaa kasvojen vaihtamisprosessista, joten se ei ole sopivin aloittelijoille. Edistyneemmille käyttäjille ohjelma tarjoaa hyvän työkalun, ja mahdollisuuden käyttää sitä paikallisesti omalla koneella Github-repositoriosta ladattuna, tai Google Colab-palvelussa. (Lanham 2021, luku 9.)

DeepfakesWeb on nettisivusto, joka tarjoaa maksua vastaan valevideoiden luomista palveluna. Kyseessä on XaaS, eli Anything as a Service -mallin laajennos valevideoiden luomiseen. XaaS viittaa mihin tahansa ohjelmistoon tai palveluun, joka ei vaadi fyysistä asennamista, ja joka toteutetaan internetin yli jaettavana palveluna. Palvelun ostaja lataa palveluun videot ja kuvat, ja sivusto luo kahden dollarin tuntihinnoittelulla valevideon. (Marshall 2019; Viira 2017.)

Teknisesti osaava henkilö voi yhdistellä muutamaa teknologiaa luodakseen valevideon. Wil Kwan yhdisti videossaan kaksi Github-palvelusta löytämäänsä valmista projektia luodakseen valevideon. Toinen projekti ohjelmakoodeineen (Real-Time-Voice-Cloning) mahdollisti äänen kloonauksen sekä tuottamisen tekstistä puheeksi, ja toisella (First Order Motion Model for Image Animation) pystyi animoimaan valokuvan yhdistäen siihen huulisynkkauksen. (Kwan 2020.)

Kuvia voi animoida yhdistäen niihin huulisynkkauksen myös Wav2Lip Model -ohjelmalla. Ohjelmaa voi ajaa Google Colab-palvelussa, ja se löytyy myös Github-repositoriosta. Ohjelma tekee automaattisesti valokuvan huulisynkkauksen perustuen sille syötettyyn wav-äänitiedostoon. (Sandusky 2021.)

3 Valevideoiden luominen käytännössä

3.1 Tutkimuksen lähtökohdat

Testatakseni tutkimuskysymystä, voisiko kuka tahansa luoda uskottavia syviä valevideoita, päätin kokeilla useita eri teknologioita nähdäkseni, miten kukin niistä suoriutuu, ja kuinka käyttäjäystävällisiä ne ovat. Valitsin testiin kartoittamistani vaihtoehtoista useita puhelinsovelluksia, yhden tietokoneelle ladattavan sekä pilvipalvelimilla toimivan sovelluksen, sekä tekniikan, jossa animoidaan kuva videoksi ja yhdistetään siihen kloonattu ääni tekstistä puheeksi-syntetisaattorilla. Tavoitteenani oli testata sekä helppoja, keskivertoja, että teknisesti vaativampia sovelluksia saadakseni kattavan kuvan tekniikan käyttömahdollisuuksista ja vaatimuksista ohjelmien käyttäjän tekniselle taitotasolle.

Saaranen-Kauppinen ja Puusniekan (2006) mukaan valittua lähestymistapaa voi luonnehtia tapaustutkimukseksi, jossa tutkitaan rajattua kokonaisuutta, esimerkiksi joukkoa tapauksia, joiden tutkimisessa prosessit ovat tarkastelun kohteena. Saaranen-Kauppinen ja Puusniekan mukaan tapaustutkimus valitaan usein menetelmäksi, kun on tavoitteena ymmärtää tutkittavaa ilmiötä syvällisesti. Vaikka tapaustutkimuksen perusteella ei Saaranen-Kauppinen ja Puusniekan mukaan voi tehdä yleistyksiä, tulosten merkitystä ja oikeellisuutta voi lisätä perusteellisen kuvauksen luominen aineistosta ja sen analyysistä. Tapauksen kokonaisvaltainen ymmärtäminen onkin tapaustutkimuksen tavoite, ja ymmärtäminen on yleistettävyyttä tärkeämpää. Tapaustutkimus sopii tutkimusstrategiaksi etenkin, kun tavoitteena on saada vastauksia kysymyksiin miten, miksi ja kuinka. (Leinonen 2020.) Vaikka tapaustutkimus on menetelmänä laadullinen, aiheeseen perehtymisen lopputuloksena syntyy useita videoita toiminnallisen tutkimuksen tapaan. Videot eivät kuitenkaan ole tämän opinnäytteen prosessissa tärkein päämäärä, vaan pikemminkin niiden tekemisen avulla rakentuva aiheen syvälinen ymmärtäminen.

3.2 Puhelinsovellukset

Kokeilin useita puhelinsovelluksia selvittääkseni, voiko jollain niistä tuottaa uskottavia valevideoita. Ennen testejä oletin sovellusten olevan helppokäyttöisimpiä, mutta minulla oli epäilykseni tulosten uskottavuudesta. Hyvin pian kävi selväksi, että suurin osa puhelinsovelluksista on tarkoitettu hupailuun, eikä riski niiden käyttämiseen pahoihin tarkoituksiin ole suuri. Lähes kaikissa kokeilemistani puhelinsovelluksista tarjolla oli rajallinen valikoima kehittäjien valitsemia elokuvakatkelmia, tai muita lyhyitä videoita, joihin kasvokuvan saattoi upottaa. Valevideoita ei siis voisi luoda minkä tahansa videon pohjalta näissä sovelluksissa, ja tarjolla olevat videokatkelmat olivat sävyltään joko humoristisia tai neutraaleja. Käytin kaikkiin puhelinsovelluksiin ottamaani kuvaa 3 pohjana.



Kuva 3. Pohjakuva puhelinsovelluksia testattaessa

3.2.1 Reface

Refacen asensin puhelimeen, jossa on Android-käyttöjärjestelmä. Refacella oli mahdollista valita useista elokuvakatkelmista mieleisensä, ja istuttaa kasvokuvan pohjalta henkilön lyhyeen katkelmaan elokuvasta. Lopputulos ei onnistunut häivyttämään taustalla olevan näyttelijän piirteitä, ja tuloksena oli yhdistelmä näyttelijän, ja valokuvan henkilön kasvonpiirteitä. Ohjelma ei siis onnistunut erottamaan täysin vaihtuvia ilmeitä ja henkilön pysyviä piirteitä, jotta eleiden päälle voisi istuttaa kasvokuvan kaikkine pysyvine piirteineen. Lopputuloksena ei ollut tunnistettava valokuvan henkilö ikään kuin esiintymässä elokuvassa, kuten kuvasta 4 voi havaita.



Kuva 4. Videon vertailu, yllä alkuperäinen elokuva, alla Refacella tuotettu kuva (Favreau 2008)

3.2.2 iFace

Toinen Refacen tyyppinen ohjelma oli iFace, jonka voi ladata iPhoneille. Lopputulos oli epäuskottava - joskin hiukan yhdennäköisempi kuin Refacella - eivätkä kasvonpiirteet olleet yhtenäisesti tunnistettavat, vaan taustalla olevan näyttelijän piirteet dominoivat, kuten voi huomata kuvassa 5 vasemman puolen kuvista.



Kuva 5. Videon vertailu, yllä alkuperäinen elokuva, alla iFacella tuotettu kuva (Favreau 2008)

Uskottavuutta lisäsi hieman, jos onnistui löytämään videokatkelman, jossa näyttelijän kasvojen pituus ja leveys täsmäsivät valokuvan henkilöön. Tällöin lopputulos oli lähempänä valokuvan henkilöä, joskaan ei täysin yhdennäköinen, kuten kuvasta 6 voi todeta.



Kuva 6. Videon vertailu, yllä alkuperäinen elokuva, alla iFacella tuotettu kuva (Raimi 2002)

3.2.3 Avatarify

Asensin iPhoneille Avatarifyn, johon ladatun kasvokuvan voi animoida. Valittavana on useita erilaisia animointeja, esimerkiksi elokuvista, mainoksista tai nettimeemeistä otettuja lyhyitä katkelmia. Pohjalla oleva valokuva animoidaan tekoälyn avulla liikuttamaan esimerkiksi suuta ja kulmakarvoja pohjalla olevan videon mallin mukaan, ja taustalle lisätään myös katkelman ääni. Osassa videoista animointi ei ollut kovin onnistunut, ja animointi oli ilmiselvää, mutta joissain videoissa ilmehtiminen vaikutti luonnollisemmalta.



Kuva 7. Avatarify, alkuperäinen kuva ja tulokset, video 1

Otteita onnistuneimmasta Avatarifyn luomasta videosta on kuvassa 7. Kuvasta huomaa, ettei henkilölle ole lisätty yläetuhampaita ollenkaan, vaikka toisessa kuvassa vasemmalta katsoja olettaisi niistä näkyvän vilauksen.



Kuva 8. Avatarify, alkuperäinen kuva ja tulokset, video 2

Kuvassa 8 on otteita videosta, jossa Avatarify ei onnistunut. Tarkemmin ruutuja tutkiessa voi huomata, miten esimerkiksi ylähuulen sekä alaleuan parran linja muuttuu. Tästäkin videosta yläetuhampaat puuttuvat. Kasvojen koko myös vaihtelee useiden ruutujen osalta, mikä luo epäuskottavan vaikutelman.

3.2.4 Faceapp

Faceapp-ohjelma oli ladattavissa joko iPhonelle tai Androidille. Kyseisellä ohjelmalla voi muokata kasvokuvastaan vanhemman tai nuoremman näköisen, kokeilla eri hius- tai partatyyliä tai jopa tutkia, miltä näyttäisi, mikäli omat kasvopiirteet muokkaisi vastakkaisen sukupuolen mukaiseksi. Ohjelma teki melko uskottavia muutoksia, ja tuotti kokeilemistani sovelluksista laadukkaimmat kuvat esimerkiksi terävyyden ja yksityiskohtien suhteen. Kuvat olivat stillkuvia, joten toisaalta vertailu videoihin on epätarkoituksenmukaista.



Kuva 9. Alkuperäisten ja Faceappilla tuotettujen deepfake-kuvien vertailu, vasemmanpuoleiset alkuperäiset, oikealla ohjelman luomat (mukaillen Biard, G. 2017 CC BY-SA 3.0)



Kuva 10. Alkuperäisten ja Faceappilla tuotettujen deepfake-kuvien vertailu, vasemmanpuoleiset alkuperäiset, oikealla ohjelman luomat (mukaillen Biard, G. 2017 CC BY-SA 3.0)

Vertaillessa kuvia 9 ja 10, on selvää, että alkuperäiset kuvat vaikuttavat lopputulokseen. Vaihdoin vasemmanpuoleisimman kuvan tutkiakseni, miten erilaiset kuvat samasta henkilöstä vaikuttavat ohjelman luomaan kuvaan. Kuvien 9 ja 10 sinipaitaiset deepfake-versiot

eivät muistuta kasvonpiirteiltään toisiaan. Kuvan 9 sinipaitaisen deepfake-version voisi arvoida muistuttavan jonkin verran alkuperäisen kuvan harmaapaitaista henkilöä, mutta kuvan 10 deepfake-versio ei juuri muistuta alkuperäistä henkilöä. Faceapp-sovellus siirsi melko hyvin henkilöiden kulmakarvojen, silmien ja nenän piirteet, mutta leukalinja ja kasvojen leveys-pituussuhde pohjautuu selvästi alkuperäiseen kuvaan vaikuttaen negatiivisesti lopputuloksen uskottavuuteen.

3.3 DeepFakes Web

Kokeillakseni monipuolisesti erilaisia tapoja tuottaa valevideoita, päätin testata myös AAS-mallin sovellusta, eli valevideo palveluna, jonka tarjosi DeepFakes Web. Palvelun käyttäminen ei vaadi syvällistä osaamista valevideoiden tekemisen taustateknologiasta. Yksinkertaisimmillaan palveluun rekisteröidytään, ostetaan pilvipalvelinaikaa, ladataan lähde- ja kohdevideo, ja lopulta annetaan palvelun tuottaa valevideo annetuista materiaaleista. Palvelu mahdollistaa valevideon tekemisen kenelle tahansa, joka on valmis uhraamaan projektiin rahaa. Videoiden muokkaaminen vaatii jonkin verran teknistä osaamista, mutta helppokäyttöisiä työkaluja on löydettävissä.

Aloitin kuvaamalla lähdehenkilön kanssa videon, johon istutettaisiin kohdehenkilöksi ranskalainen näyttelijä Mathieu Kassovitz, josta vertailua varten kasvokuva esillä kuvassa 11. Video kuvattiin suoraan edestä, ja editoin sen myöhemmin riittävän lyhyeksi palveluun lataamista varten. Seuraavaksi etsin internetistä hyvälaatuisia videoita näyttelijästä. Tärkeää oli löytää videoita, joissa näyttelijä olisi kuvattu suoraan edestä, videossa ei esiintyisi muita henkilöitä, ja kuva olisi riittävän laaja. Nämä kriteerit oli lueteltu DeepFakes Webin videoiden lisäyssivustolla.



Kuva 11. Kuva näyttelijästä, eli kohdehenkilöstä (Biard, G. 2017 CC BY-SA 3.0)

Löydettyäni sopivat videot, latsin ne Youtube-palvelusta 4K Video Downloader ohjelmalla voidakseni muokata niitä. VideoPad-videoeditorin avulla muokkasin videoista pois osuudet, joissa oli muita ihmisiä, kuvakulma oli väärä, tai näyttelijän kasvot peittyneet. Videoille oli määritelty myös maksimitiedostokoko 50 megatavua, joten videot tuli muokata alle maksimikoon. DeepFakes Webiin oli mahdollista lisätä myös lähde- ja kohdevideoiden

lisäksi 1500 kpl kuvia näyttelijästä, joten loppuista löytämistäni videoista muokkasin jpeg-kuvia samalla VideoPad-videoeditorilla. Lopulta lisäsin palveluun videoiden lisäksi maksimimäärän kuvia näyttelijästä, ja aloitin valevideon tekemisen. Aloitin minimimäärällä pilvipalveluaikaa, joka oli 300 minuuttia ja maksoi 15 dollaria. Sain sähköpostiini ilmoituksen valevideon luomisprosessin valmistuttua. Ensimmäisen kierroksen lopputulos ei ollut vielä kovinkaan uskottava, lähinnä kasvojen keskikohta näytti hieman sumealta, ja joitain näyttelijän kasvonpiirteitä saattoi erottaa, kuten voi todeta kuvasta 12. Taustalla oleva lähdehenkilö erottui vielä selvästi kasvonpiirteineen. Pilviaikaa oli vielä jäljellä 100 minuuttia.



Kuva 12. Deepfakesweb alkuperäiset kuvat yllä ja alla valevideot ensimmäisen kierroksen jälkeen

Toiselle kierrokselle valitsin vaihtoehdon käyttää aiempaa mallia uuden videon pohjana. Ostin uudelleen 300 minuuttia pilvipalvelinaikaa. Lisäsin palveluun uudelleen kohdevideon, sekä uuden videon lisäksi kuvamateriaalia näyttelijästä. Tässä vaiheessa palvelu ilmoitti häviön määräksi lähdemateriaalissa 0.0324 ja kohdemateriaalissa 0.03511. Mitä alhaisemman arvon saa, sitä parempi on lopputulos DeepFakes Webin ohjeiden mukaan. Toisen kierroksen valmistuttua video oli edelleen keskeltä utuinen, mutta joitain näyttelijän piirteitä pystyi jo erottamaan terävämpinä kuin ensimmäisessä videossa, kuten voi todeta kuvasta 13. Kontrasti lähde- ja kohdehenkilön välillä oli vielä selkeä, eikä lopputulos ollut uskottava.



Kuva 13. Deepfakes web alkuperäiset kuvat yllä ja alla valevideot toisen kierroksen jälkeen

Kolmannelle kierrokselle lisäsin kuvamateriaalia lähdehenkilöstä muokkaamalla toisesta videosta kuvasarjan, ja lataamalla kuvat videon lisäksi palveluun. Lisäsin myös kohde näyttelijästä uuden videon, sekä kuvamateriaalia. Tässä vaiheessa häviölukemat olivat lähdehenkilöllä 0.03364 ja kohteella eli näyttelijällä 0.02571. Videon valmistuttua utuisuus oli vielä jäljellä, mutta verrattuna aiempiin videoihin, etenkin alue kulmakarvojen välissä näytti hieman terävämmältä, joskaan ei yhtä terävältä, kuin video, johon kasvot upotettiin. Kolmannen iteraation tulokset voi nähdä kuvasta 14.



Kuva 14. Deepfakesweb alkuperäiset kuvat yllä ja alla valevideot kolmannen kierroksen jälkeen

Neljäs iteraatio käynnistyi jälleen uudella videolla ja kuvamateriaalilla näyttelijästä, sekä lähdehenkilön videolla ja kuvilla. Häviölukemat olivat lähdehenkilöllä 0.02997 ja kohteella eli näyttelijällä 0.02939. Pilvipalvelinaikaa ei tarvinnut ostaa lisää, sillä aiemmilta kierroksilta oli jäänyt riittävästi uuden videon luomiseen aiemmin koulutetun mallin pohjalta.



Kuva 15. Deepfakesweb alkuperäiset kuvat yllä ja alla valevideot neljännen kierroksen jälkeen

Kuvassa 15 voi huomata keskimmäisestä deepfake-kuvasta kulmakarvojen välisen rypyn hieman silottuneen verrattuna kahden edellisen kierroksen kuvaan. Kaikkien iteraatioiden kuvat ovat liitteessä 1, josta näkee allekkain jokaisen vaiheen vertailukuvat. Muutokset kierrosten välillä ovat hyvin pieniä, ja ylipäätään Deepfakes Web on ainoastaan muokannut kasvojen keskikohtaa, ja jättänyt leukalinjan ja korvat ennalleen. Ero on selvästi huomattavissa, sillä alkuperäinen video erottuu tarkempana, ja retusoitu kohta on sumeampi. En jatkanut iterointeja pidemmälle, sillä erot kierrosten välissä olivat minimaalisia, ja itse asiassa ensimmäisen ja viimeisen kierroksen kuvat olivat melko yhtenäiset etenkin kulmakarvojen välisen rypyn osalta. Mikäli kierrosten tuottamien kuvien laatu olisi selvästi eronnut toisistaan, ja kuvat parantuneet silminnähden, olisin kokenut jatkamisen mielekkäämmäksi.

3.4 Kuvan animoiminen, äänen kloonaaminen

Tutkiessani erilaisia tapoja syvän valevideon luomiseksi löysin videon, jossa Will Kwan (2020) käytti kopioituja ohjelmakoodeja animoidakseen valokuvan henkilön puhumaan, yhdistäen videoon äänenkloonaussovelluksen, joka pystyi kloonaamaan äänen jo alle puolen minuutin mittaisesta katkelmasta. Tekniikka vaikutti melko helppokäyttöiseltä, joten päätin kokeilla sitä valevideon tuottamiseen.

3.4.1 Testissä käytettäviin ohjelmiin tutustuminen

Will Kwan käytti äänen kloonaamiseen Real-Time Voice Cloning -ohjelmaa (Corentin 2019). Ohjelma on Jemine Corentinin pro gradu -työtään varten kehittämä työkalu, joka luo sille annetusta lyhyestä audiokatkelmasta mallin, jota käyttää tekstistä puheeksi syntetisaattorin syötteenä, mahdollistaen äänen kloonauksen pikavauhtia aiempien työläiden ja aikaa vievien mallin koulutusten sijasta. Corentin on julkaissut ohjelmansa avoimena lähdekoodina Github-repositoriossa.

Valokuvan animointiin Kwan käytti First Order Motion Model -tutkimuksen (Siarohin, Lathuilière, Tulyakov, Ricci & Sebe 2019) Github-julkaisuarkistosta löytyvää Jupyter Notebook -työkalulla tehtyä demoa. First Order Motion Model mahdollistaa liikkeen kloonaamisen. Kun ohjelman tekoälymalli on koulutettu saman kategorian asioilla (esimerkiksi kasvot, tai ihmisvartalot), voi mallia käyttää mihin tahansa saman kategorian asian animoimiseen. Malli tarvitsee taustalle ohjaavan videon, josta liike kopioidaan, ja kohdekuvan, joka animoidaan. Tästä ohjelmasta on mahdollista ladata omalle koneelleen paikallinen versio, tai käyttää ohjelman ajamiseen Google Colabia.

Google Colab on palvelu, jossa voi kirjoittaa ja suorittaa Python-ohjelmakoodia, sekä luoda, ladata ja jakaa Jupyter Notebook -muistioita esimerkiksi Google Drivella tai Github-

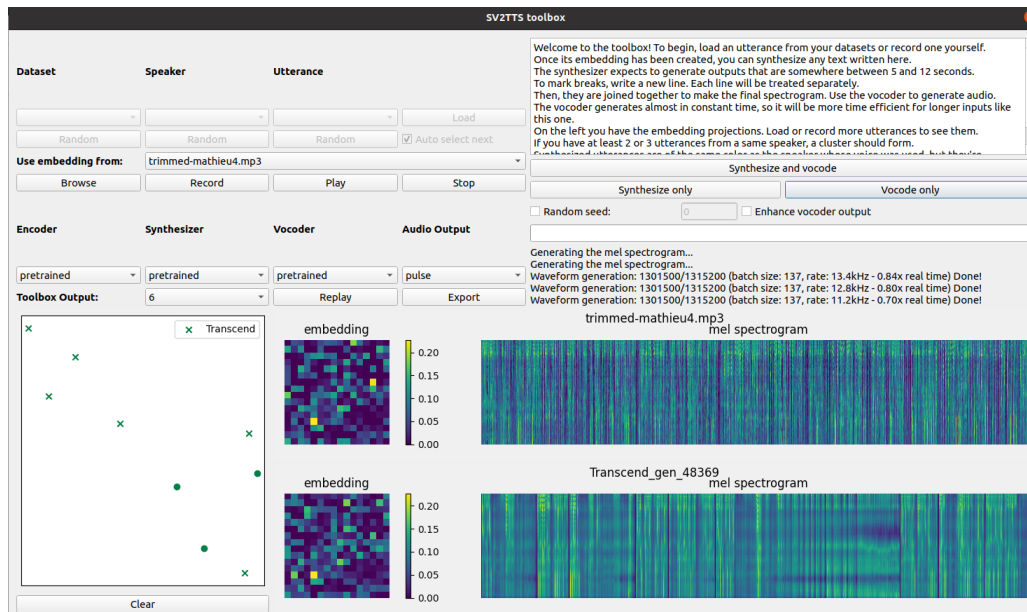
palvelussa. Jupyter Notebook on Python-kielen ohjelmointityökalu, jota voi käyttää selaimella. Google Colabissa on mahdollista käyttää ilmaiseksi pilvipalvelua tehokkaalla grafiikkasuorittimella ja hyödyntää esiasennettuja datakirjastoja kuten NumPy, Matplotlib sekä koneoppimiskirjastoja kuten Keras, PyTorch ja Tensorflow. Jupyter Notebookilla teksti, kuva ja ohjelmakoodi yhdistyy muodostaen tarinamaisen käsikirjoituksen, ja työkalu mahdollistaakin nopean ideoiden kokeilun ja tuotosten jakamisen muille. (Garbade 2021; Harhanen 2019.)

Koska Kwanin videosta (2020) ei käynyt ilmi yksityiskohtia esimerkiksi siitä, miten käyttää kuvan animoimisen taustalla omaa videota, ja kuinka voi yhdistää äänen kuvaan, päätin asiaa tutkittuani ostaa Udemy -internetkurssikirjastosta videokurssin (Nelson 2020), jossa tekniikat käydään havainnollisesti läpi. Kurssi opasti minua myös Google Colabin käyttöön.

3.4.2 Äänen kloonauk

Aloitin testin lataamalla tietokoneelle Real-Time Voice Cloning -ohjelman. Muodostin oppimani pohjalta työnkulun, jossa ensimmäisenä askeleena oli luoda kloonattu äänitiedosto. Äänitiedoston luomisen jälkeen olisi mahdollista kuvata ohjaava video, jolla yrittää huulisynkronoida äänitiedoston sanat mahdollisimman hyvin. Kun ohjaava video olisi tehty, olisi mahdollista yrittää kasvojen liikkeiden kopioimista haluamaansa kuvaan, ja lopulta yhdistää animoitu kuva äänitiedostoon. Tutkittuani aihetta lisää löysin myös Wav2Lip-ohjelman (Prajwal, Mukhopadhyay, Namboodiri & Jawahar 2020), jolla äänitiedoston pohjalta valokuvan voi suoraan animoida liikuttamaan huulia äänen mukaisesti. Tämä poistaa prosessista työvaiheet, jossa kuvataan itse ohjaava video äänitiedostoon huulisynkronoiden, ja valokuvan animoinnin sen pohjalta. Päätin siis muokata Will Kwanin prosessia hieman, ja edetä äänitiedoston luomisesta suoraan huulisynkkauksen animointiin Wav2Lipillä.

Real-Time Voice Cloning -ohjelman lataamiseen oli selkeät ohjeet projektin Github-repositoriossa. Käytin työhön Linux Ubuntu -käyttöjärjestelmää, ja latasin PyTorch-koneoppimiskirjaston sekä ffmpeg-ohjelman äänen ja videon editointiin. Lisäksi latasin muut tarvittavat vaatimukset-tiedostossa mainitut ohjelmat, sekä mallitiedostot.



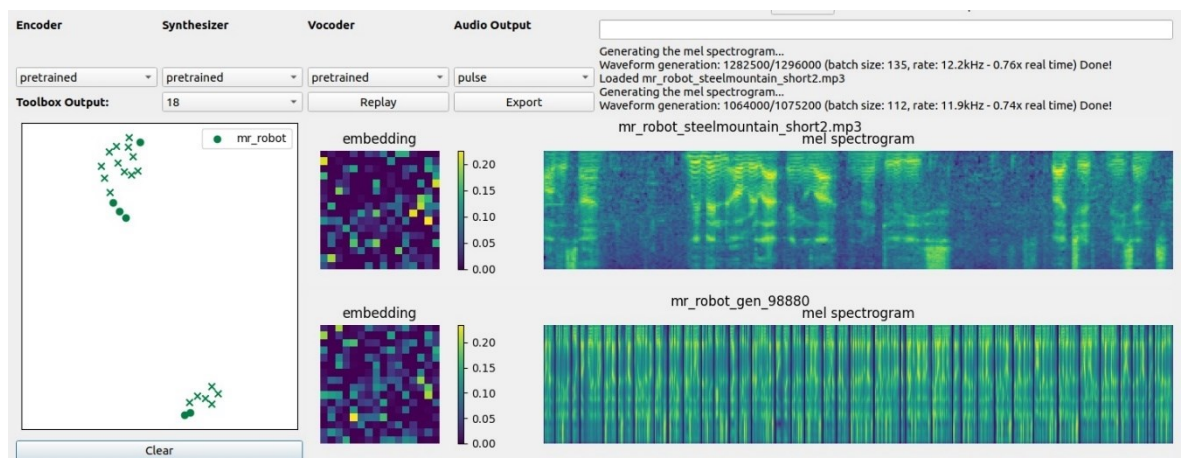
Kuva 16. Real-Time Voice-Cloning -ohjelman käyttöliittymä

Kuvassa 16 näkyy ohjelman käyttöliittymä. Ohjelmalla voi selata ja valita omia äänitiedostoja Browse-nappia painamalla, tai äänittää omaa ääntä painamalla Record-nappia. Painamalla Synthesize and vocode ohjelma kloonaa äänen, ja tuottaa äänen tekstistä puheeksi syntetisaattorilla pohjautuen yläoikealla tekstilaatikkoon kirjoitettuun käsikirjoitukseen. Ohjelma tekee taukoja puheeseen rivinvaihdolla, ja koska se on koulutettu äänikirjoihin perustuvalla materiaalilla, puhe ei kuulosta täysin luonnolliselta. Alavasemmalla näkyvään laatikkoon ilmestyy kustakin omasta äänitiedostosta pallon muotoinen symboli, ja kustakin kloonatusta puhekatkelmasta rasti. Symbolien pitäisi muodostaa rykelmiä, joissa samasta lähteestä peräisin olevat äänikatkelmat ryhmittyvät lähekkäin, ja esimerkiksi kloonattujen äänikatkelmien rastien pitäisi asettua lähelle lähdeäänikatkelmien pallorykelmiä.

Ensimmäisissä testeissäni symbolit olivat kuitenkin hajallaan. Tämä saattoi johtua siitä, että testasin äänittämistä itse, ja vaikka puhuinkin englantia, en puhunut sitä natiivitasolla. Ohjelma on kuitenkin koulutettu amerikan-englannin aksentilla puhutulla äänikirjamateriaalilla, joten se pystyy parhaiten kopioimaan vastaavaa materiaalia. Ohjelman repositorion kysymys-vastauspalstalla oli useita kysymyksiä aiheeseen liittyen, ja kehittäjä vastasi ohjelman toimivan parhaiten natiivienglannilla. Mikäli haluaisi ohjelman toimivan toisella kielellä tai aksentilla, tulisi sille tarjota runsaasti hyvälaatuista koulutusmateriaalia kyseisellä kielellä.

Päätinkin seuraavaa kokeilua varten testata natiivitasolla amerikanenglantia sisältävän äänitiedoston avulla tehtävää kloonausta, ja tämä testi onnistuikin paremmin. Sekä alkuperäiset lähteet, että kloonatut äänikatkelmat muodostivat rykelmiä, ja ääni myös kuulosti jonkin verran alkuperäiseltä. Ongelmia aiheutti lähinnä se, että ohjelma ei kopioi

esimerkiksi alkuperäisen äänilähteen puheen rytmiä, vaan noudattaa koulutusmateriaalin puherytmiä, jolloin ääni ei kuulosta aivan puhujaltaan. Myös jotkin sanat saattoivat jäädä heikosti lausutuiksi, tai kokonaan välistä, ja äänessä oli jonkin verran artefakteja, kuten metallinen kaiku, jotka paljastivat helposti äänen olevan jollain tapaa muokattu. Kokeilin useilla eri audiokatkelmilla, pituuden vaihdellessa viidestä kahteenkymmeneen sekuntiin, ja lopulta onnistuin saamaan äänen kuulostamaan melko luonnolliselta. Päätin käyttää tätä audiokatkelmaa pohjana videon huulisynkkauksen animointiin. Kuvassa 17 on kuva-kaappaus lopullisen kokeilun, ja sitä edeltävien äänikatkelmien ryhmittymisestä. Kuvan vasemmalla sijaitsevassa laatikossa rastilla merkityt ovat ohjelmalla luotuja kloonattuja audiokatkelmia, ja pallolla merkityt ovat alkuperäisiä lähteenä käytettyjä äänikatkelmia.



Kuva 17. Real-Time-Voice-Cloning -ohjelman kuvaajat

3.4.3 Videon animointi äänitiedoston pohjalta

Kun äänitiedosto oli valmiina, kokeilin Wav2Lip-ohjelmaa. Katsoin Youtubesta What Make Art -nimisen kanavan Make Art Talk -ohjevideon ohjelman käyttämisestä, ja ohjeita seuraten siirryin ohjelman Github-repositorion linkistä Google Colab-tiedostoon, jossa kohta kohdalta play-nappeja painellen äänitiedoston pystyi yhdistämään joko videoon, tai valokuvaan, joka oli muokattu videoksi. Kokeilin molempia tapoja, ja vaikka huulisynkkaus vaikutti pätevältä, valokuvasta oli helppo huomata, että kuvan henkilö on animoitu. Valokuvien suhteen uskottavamman lopputuloksen tuotti, mikäli kuvan henkilö ei hymyillyt leveästi suu auki, sillä silloin ohjelma ei osannut luoda animaatiota suun sulkemiseksi. Normaalisti ihmisillä hymyillen puhuessakin suu on hetkittäin kiinni, joten video oli selvästi epäluonnollinen.

Videon huulisynkkaus yhdistyi valokuvaa paremmin. Videossa oli tärkeää huomata, että jokaisessa ruudussa oli oltava animoitava henkilö, eli tyhjiä ruutuja ei saanut olla, ja videon tuli olla äänitiedoston mittainen. Tämän ohjelman käytössä tarvittiin siis hieman

videoeditointitaitoja, mutta muokkaukset olivat melko yksinkertaisia, ja vaadittavat toimet mielestäni helposti opittavissa verkkohakuja ja ohjeita hyödyntämällä.

Testieni perusteella huulien animointi toimi parhaiten, mikäli kuvan henkilö katsoi suoraan kameraan, tai ainoastaan hieman viistoon, mutta mikäli henkilö katsoi suoraan sivulle, huulien animoinnin kohdistaminen epäonnistui. Myös kasvoja kallistaessa animointi kohdistui väärään kohtaan kasvoja piirtäen siihen mustan epämääräisen muotoisen kolon. Esimerkkejä epäonnistuneista kohdista puheen animoinnissa on kuvassa 18.



Kuva 18. Sivuprofiilissa ja kasvoja kallistaessa Wav2lip epäonnistuu animoimaan suun auki luontevasti, still-kuvat videosta

Tässä menetelmässä oli useita rajoitteita, jotka vaikuttivat lopputuloksen uskottavuuteen. Menetelmässä täytyi yhdistää useita eri ohjelmia, joskin ohjelmat olivat melko helppokäyttöisiä. Tulos oli kaikkein uskottavin, jos animointi pohjautui videoon ja videon henkilön kasvat olivat kameraan päin, eivätkä liikkuneet juurikaan sivuille, ylös, alas tai kallistuen. Äänen suhteen kloonaaminen onnistui parhaiten natiivin tasolla amerikanenglantia puhuvan henkilön kohdalla, mutta silti puheen rytmi kuulosti ohjelman koulutukseen käytetyn äänikirjan ääneen lukemiselta, eikä täysin luonnolliselta puheelta.

3.5 DeepFaceLab

Kahdesta samankaltaisesta valevideoiden luomisohjelmasta valitsin DeepFaceLabin kokeiluun Faceswapin sijasta, koska löysin sen käytöstä huomattavan paljon materiaalia. Päätin kokeilla DeepFaceLabia ensin, ja taustalla oli ajatus siirtyä Faceswapiin, mikäli en saisi ohjelmaa toimimaan. DeepFaceLabin käyttöä puolsi myös se seikka, että ohjelmaa voi ajaa joko omalla koneella paikallisesti, tai Google Colab -palvelussa, tai yhdistäen molempia. Koska käytössäni olevan tietokoneen grafiikkasuoritin ei ollut parhain mahdollinen tähän projektiin, oli järkevää ajaa ohjelmaa paikallisesti silloin kun tarvitsin hienosäätöön graafista käyttöliittymää, ja Google Colabia mallin kouluttamisprosessissa, jossa tehokas grafiikkasuoritin oli tärkeässä roolissa.

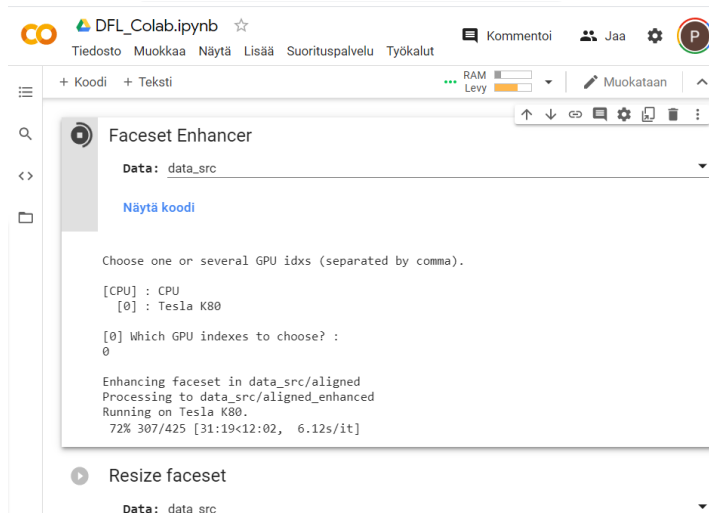
DeepFaceLab kehitettiin helpottamaan valevideoiden torjuntaa tarjoamalla helppokäyttöisen käyttöliittymän hyvälaatuisten valevideoiden luomiselle. Avoin lähdekoodi ja jatkuva ohjelman kehitystyö luovat hyvän alustan suurelle yleisölle kokeilla itse valevideoiden luomista, jotta niiden tunnistaminen ja taustalla olevan teknologian ymmärtäminen helpottuisivat.

3.5.1 Ohjelmaan tutustuminen

Löysin internetiä tutkimalla Youtube-palvelusta videotutoriaaleja ja verkkosivuja ohjeineen, joita tutkimalla sain yleiskuvan työn kulusta. Seurasin ensimmäisessä kokeilussa Sergio Virahondan (2021) ohjeita työkalun käyttämiseen Googlen Colab-palvelussa.

Virahondan (2021) ohjeita seuraten loin ensiksi paikalliselle koneelleni workspace-nimisen kansion, johon toin kohde- ja lähdevideot nimillä `data_src.mp4` ja `data_dst.mp4`. Käyttämällä näitä tiedostonimiä DeepFaceLab-ohjelma löytäisi oikeat tiedostot. Videon nimellä `data_src.mp4` tuli sisältää kasvokuvia henkilöstä, jonka kasvot kloonataan, ja videoon nimellä `data_dst.mp4` sijoitetaan kloonatut kasvot. Käytin samoja videoita näyttelijä Mathieu Kassowitzista (kuva 11), ja kohdevideosta (kuvat 12-15) johon kasvot upotetaan, joita olin jo käyttänyt luodessani videota DeepFakesWeb-palvelulla.

Pakkasin tiedostot workspace-nimiseen kansioon ohjeiden mukaisesti ja tallensin ne oman Google Driven juureen. Sitten avasin linkistä DeepFaceLabin Google Colab -muistion ja kopioin sen omaan Driveeni, josta siirsin sen omaan kansioonsa. Lähdin osio kerrallaan ajamaan ohjelmakoodia pitämällä ohjeiden mukaiset oletusasetukset. Kuten kuvasta 19 näkee, ohjelmakoodia ajetaan painamalla kolmionmuotoista nappia, ja ohjelmaa suoritettaessa on näkyvillä neliönmuotoinen nappi, josta ohjelman suorittaminen pysäytetään, ja joka on käyttöliittymänä intuitiivinen ja tuttu mistä tahansa kaukosäätimestä. Mikäli ohjelma tarvitsee käyttäjän syötteitä, ilmestyy ruudulle pieni laatikko, ja helpommillaan laatikkoon voi painaa enter-näppäintä, jolloin ohjelma käyttää oletusasetuksia, jotka näkyvät laatikon selitteen alussa hakasuluissa.



Kuva 19. Google Colab -ohjelmakoodin ajaminen

Yritin alkuun ajaa kaiken pilvessä, alkaen videoiden pilkkomisesta ruutuihin, ja päättyen lopullisen videon luomiseen, mutta jo alkuvaiheessa ajaessani ohjelmaa, joka erottelee kasvot videosta, huomasin, ettei minulla ole juuri kontrollia tai näkymää siihen, mitä ohjelma tekee. Päätinkin vaihtaa lähestymistapaa, ja ladata DeepFaceLabin paikallisesti tietokoneelleni projektin Github-repositoriosta, ja useita ohjeita yhdistäen luoda oman työkulkuni yhdistäen sekä paikallisella koneella tehtävän työn, että Google Colabin tehokkaan pilvipalvelun hyödyntämisen saavuttaakseni laadukkaan lopputuloksen.

3.5.2 Ohjelman käyttäminen paikallisesti

Aloitin lataamalla DeepFaceLab-ohjelman projektin Github-repositoriosta paikalliselle Windows-tietokoneelle, jossa oli NVIDIA:n grafiikkasuoritin, jonka CUDA-ohjelmointirajapintaa ohjelma käyttää graafisen prosessointiyksikön hyödyntämiseksi CPU-suorittimen sijasta. Purettuani pakatun kansion ja asennettuani projektin, ohjelman näkymä oli melko yksinkertainen, kuten kuvasta 20 voi todeta.

_internal	04/09/2021 23.11	File folder	
workspace	04/09/2021 23.11	File folder	
1) clear workspace	04/09/2021 23.11	Windows Batch File	1 KB
2) extract images from video data_src	04/09/2021 23.11	Windows Batch File	1 KB
3) cut video (drop video on me)	04/09/2021 23.11	Windows Batch File	1 KB
3) extract images from video data_dst FU...	04/09/2021 23.11	Windows Batch File	1 KB
3.optional) denoise data_dst images	04/09/2021 23.11	Windows Batch File	1 KB
4) data_src faceset extract MANUAL	04/09/2021 23.11	Windows Batch File	1 KB
4) data_src faceset extract	04/09/2021 23.11	Windows Batch File	1 KB
4.1) data_src view aligned result	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src sort	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util add landmarks debug i...	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset enhance	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset metadata restore	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset metadata save	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset pack	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset resize	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util faceset unpack	04/09/2021 23.11	Windows Batch File	1 KB
4.2) data_src util recover original filename	04/09/2021 23.11	Windows Batch File	1 KB
5) data_dst faceset extract + manual fix	04/09/2021 23.11	Windows Batch File	1 KB
5) data_dst faceset extract MANUAL	04/09/2021 23.11	Windows Batch File	1 KB
5) data_dst faceset extract	04/09/2021 23.11	Windows Batch File	1 KB
5) data_dst faceset MANUAL RE-EXTRAC...	04/09/2021 23.11	Windows Batch File	1 KB

Kuva 20. DeepFaceLab-ohjelman käyttöliittymä

Kuvassa 20 ylimpänä näkyy `_internal`-kansio, jossa on DeepFaceLabin kirjastot ja komentosarjat. `workspace`-kansio sisältää työtilan, jossa on videot tai kuvat ja ohjelman rakentama malli henkilöistä. Kansioden alla on batch-komentotiedostoja, joilla ohjelmaa ajetaan. Tiedostot on numeroitu, ja käyttö etenee numerojärjestyksessä. Komentotiedoston voi avata tuplaklikkaamalla, jolloin aukeaa näkymä komentokehoteeseen. Komentokehoteessa ohjelma pyytää antamaan asetukset, joista oletusasetus on merkattu kysymyksen eteen hakasulkeissa, ja vaihtoehdot kysymyksen perässä, joista kysymysmerkkiä käytettäessä voi avata ohjeistuksen kyseisestä komennosta. Enteriä painamalla komento suoritetaan, ja ohjelma avaa ikkunan tai suorittaa koodin taustalla. Ohjelma ilmoittaa taustajon päättymisestä kehottamalla painamaan mitä tahansa näppäintä poistuakseen komentokehotteesta.

DeepFaceLabia käytetään sijoittamalla `workspace`-kansioon (kuvassa 18) `data_dst`-video eli video, johon kloonatut kasvot upotetaan, ja `data_src`-video, jossa on kloonattava henkilö. Model-kansioon ohjelma rakentaa opettelemansa mallin kasvoista. Mikäli haluaa käyttää useita videoita kloonattavasta henkilöstä saadakseen mahdollisimman monipuolisesti kuvakulmia, ilmeitä ja valaistusolosuhteita auttaakseen tekoälyn oppimista, on helppointa editoida ja yhdistää videot jo ennen ohjelman käyttämistä yhdeksi videoksi, josta DeepFaceLabin avulla voi eritellä ruudut. Parhaita videoita yleisestä internetistä esimerkiksi julkisuuden henkilöstä ovat haastatteluvideot, joissa kuva keskittyy henkilön kasvoihin, joissa ei ole juuri muita henkilöitä (joskin muita henkilöitä sisältävät ruudut voi DeepFaceLabin avulla melko helposti poistaa), ja jotka ovat mahdollisimman hyvälaatuisia eli teräviä.

Name	Date modified	Type	Size
data_dst	04/09/2021 23.11	File folder	
data_src	04/09/2021 23.11	File folder	
model	04/09/2021 23.11	File folder	
data_dst	04/09/2021 23.11	MP4 File	104 473 KB
data_src	04/09/2021 23.11	MP4 File	13 454 KB

Kuva 21. DeepFaceLab workspace-kansion sisältö

Aloitin ohjelman käytön sijoittamalla lähde- ja kohdevideot workspace-kansioon ja ajamalla komentotiedoston 2) extract images from video data_src ja 3) extract images from video data_dst FULL FPS erotellakseni videoista yksittäiset kuvat eli ruudut. Valitsin face type, eli kasvotyyppiä wf eli whole face, jolloin ohjelma käyttää kasvoja leuasta hiusrajaan asti rakentaessaan mallia kasvoista.

```

C:\windows\system32\cmd.exe
[y] Continue extraction? ( y/n ?:help ) : y
Choose one or several GPU idxs (separated by comma).

[CPU] : CPU
[0] : NVIDIA GeForce GTX 1650
[1] : AMD Radeon(TM) Graphics

[0,1] Which GPU indexes to choose? :
0,1

[wf] Face type ( f/wf/head ?:help ) :
wf
[512] Image size ( 256-2048 ?:help ) :
512
[90] Jpeg quality ( 1-100 ?:help ) :
90
Extracting faces...
Running on AMD Radeon(TM) Graphics
Running on NVIDIA GeForce GTX 1650
100%#####| 128/128 [01:22<00:00, 1.54it/s]
-----
Images found:      128
Faces detected:    128
-----
Done.
Press any key to continue . . .

```

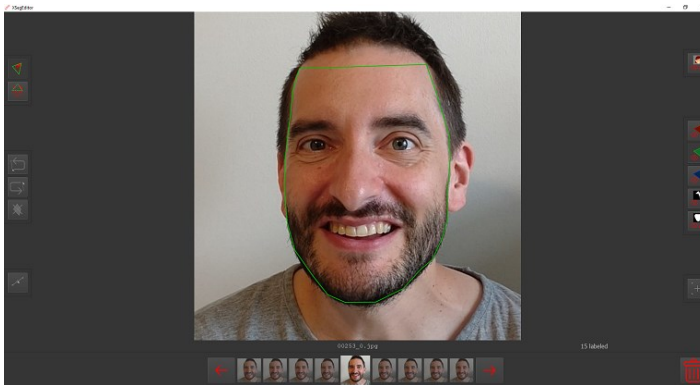
Kuva 22. DeepFaceLab kasvojen erottelu kuvista

Kun videot on jaettu kuvasarjoiksi, kuvista täytyy löytää kasvot ja suoristaa kasvojen tunnuspiirteet, kuten silmät samalle tasolle. DeepFaceLab tarjoaa työkalun myös tähän tarkoitukseen komentosarjassa 4) data_src faceset extract ja 5) data_dst faceset extract käyttäen kasvojen löytämiseen S3FD-algoritmia (kuva 22). Mikäli lopussa images found ja faces detected -lukemat, eli löydettyjen kuvien ja löydettyjen kasvojen määrät eroavat toisistaan, on ohjelma löytänyt videosta useita henkilöitä, tai kuvia, joissa ei ole henkilöitä, jolloin nämä täytyy erikseen käydä läpi ja poistaa jättääkseen aineistoon vain kuvia oikeasta henkilöstä. Kuvat hyvä käydä läpi, ja jättää jäljelle kloonattavan henkilön osalta vain kuvat, joissa laatu on paras, ja joissa on monipuolisesti eri kuvakulmia, ilmeitä ja valaistusolosuhteita.

Kun kasvot on erotettu kuvista, ja suoristettu, komentosarjalla 4.1) data_src view aligned result aukeaa ikkuna, jossa voi Windowsin resurssienhallintaa nopeammin selata kuvia, ja valita poistettavat. Kuvat voi myös lajitella komentosarjalla 4.2) data_src sort, esimerkiksi

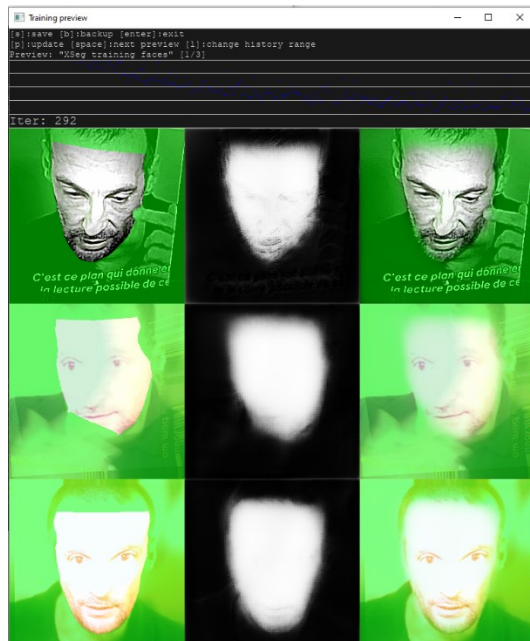
histogrammin mukaan, sumentuneiden kuvien mukaan tai kasvojen suunnan mukaan, jolloin huonolaatuiset kuvat on koottu yhteen ja helpommin poistettavissa. Kasvojen erotellun ja suoristuksen seurauksena ohjelma luo workspace-kansioon data_dst ja data_src-kansioihin aligned-kansion, jossa on jatkoa varten tarpeelliset kuvat. Aligned-kansion ulkopuolella olevia kuvia ei jatkossa ennen uudelleen videoksi yhdistämistä tarvita, ja etenkin pilveen siirtäessä ne kannattaa tilan säästämiseksi jättää odottamaan jatkokäyttöä omalle koneelle.

Seuraava askel on segmentointi, eli osiin jakaminen. Kyseisessä prosessissa esimerkiksi kasvoilla olevat paksut hiussuortuvat, silmälasit, ynnä muut esteet, erotellaan. Kasvoille piirretään DeepFaceLabin komentosarjalla 5.XSeg) data_dst mask for XSeg trainer – edit ja 5.XSeg) data_src mask for XSeg trainer – edit useaan kuvaan monikulmio, eli maski. Whole face -mallin osalta piirretään maski leuasta hiusrajaan kasvojen reunoja seuraten, korvat ulos jättäen ja kasvojen esteet pois erotellen (kuva 23). Tämä kertoo tekoälylle, mitä tulee ottaa kasvoihin, ja mitä jätetään ulos, jottei esimerkiksi paksut hiussuortuvat tai henkilön sormet joissain kloonattavan henkilön kuvissa siirtyisi irrallisina objekteina lopulliseen kloonausmalliin. Kuvajoukon koosta ja kuvakulmista ja ilmeistä riippuen 20–200 monikulmioilla merkattua maskimallia saattaa riittää, ja ohjelma opettelee tämän aineiston perusteella tekemään maskin lopuille kuville.

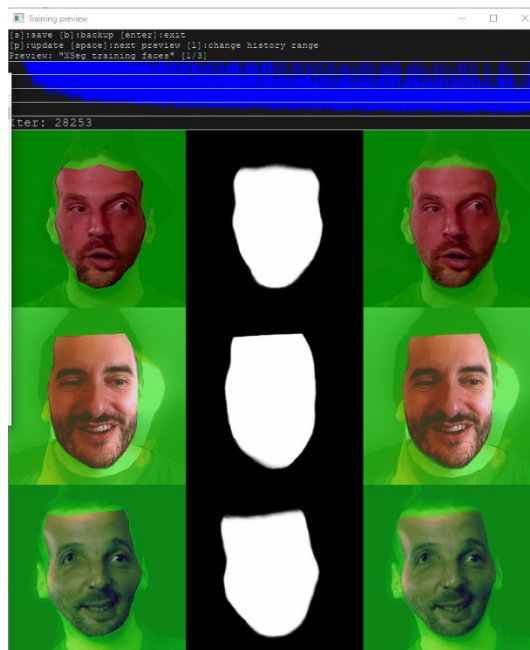


Kuva 23. XSeg maskimonikulmion määrittäminen

Kun XSeg-maski on määritelty, ajetaan maskin koulutusohjelma komentosarjalla XSeg) train.bat. Koulutuksen aikana voi nähdä esikatselukuvan suoristetuista kasvoista ja maskista, ja mikäli maskin reunat koulutukset aikana vääristyvät ottaen mukaan jonkin objektin, jota ei maskin määrittelyn aikana ole rajattu pois, voi koulutuksen keskeyttää, määrittellä objektin rajattavaksi pois maskista, ja jatkaa koulutusta. Iteraatioiden kertyessä esikatselukuvan maskin rajat tarkentuvat (kuvat 24 ja 25).



Kuva 24. XSeg-maski koulutuksen alussa



Kuva 25. XSeg-maski koulutuksen lopussa

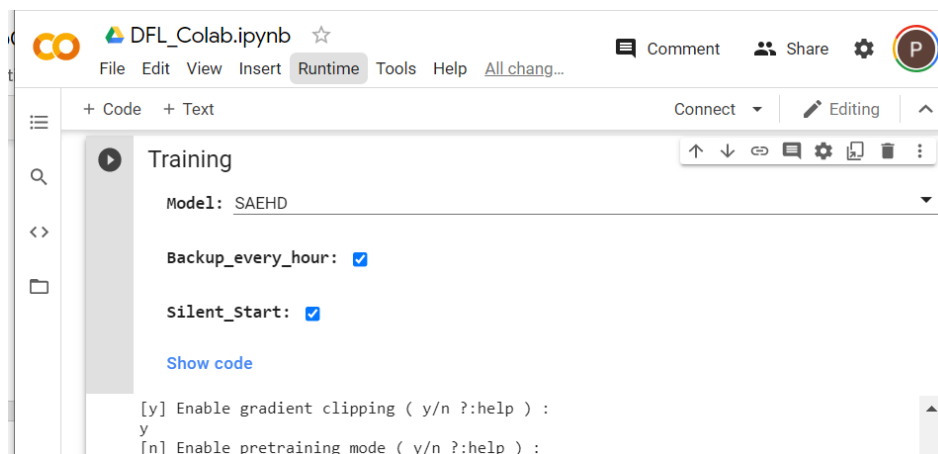
Iteraatioiden määrälle ei ole määritelty mitään tiettyä lukemaa, joka on riittävä, kuten ei monille muillekaan kohdille DeepFaceLabin luomisprosessissa. Tärkeää on seurata lopputulosta esikatseluista, ja silmämääräisesti päättää, mikä on omasta mielestä riittävä määrä iteraatioita.

Kun olin tyytyväinen maskin tarkkuuteen, lopetin XSeg-koulutuksen, ja ajoin komentosarjan `5.XSeg Generic) data_dst whole_face mask – apply` ja `5.XSeg Generic) data_src whole_face mask – apply` tallentaakseni XSeg-mallin kuviin.

3.5.3 Mallin koulutus Google Colabissa

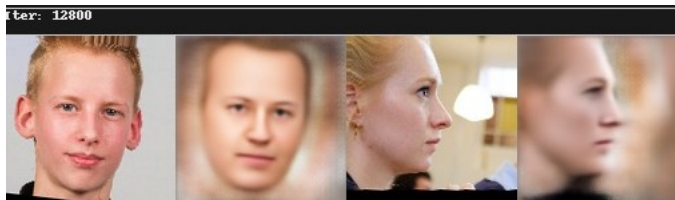
Seuraavaksi oli vuorossa mallin koulutus, jonka päätin tehdä pilvessä, sillä tietokoneessa olisi hyvä olla tehokas grafiikkasuoritin, jotta työ sujuisi joutuisasti. Aloitin lataamalla Google Drive -kansioni juureen workspace-kansioni, johon sisällytin aligned-kansion, eli suoristetut kuvat sekä kohdehenkilöstä, että kohdevideosta, että model-kansion, jossa oli tiedot XSeg-maskista. Myöhemmin model-kansioon tallentuisi myös varmuuskopiot mallin koulutuksesta.

Seuraavaksi siirryin Google Colab-näkymään, ja yhdistin asetuksista DeepFaceLab-notebookin graafiseen suorittimeen. Tarkistin suorittimen tyyppin ajamalla komennon Check GPU, ja sain tulokseksi Tesla K80. Valitsin alkuasetuksista DeepFaceLab-ohjelman asentamisen yhteydessä ladata käytettäväkseni FFHQ-aineiston (Flickr-Faces-HQ), jossa on 70 000 hyvälaatuista PNG-tiedostomuodon kuvaa 1024x1024 resoluutiolla. Kyseinen aineisto on ladattu alkujaan Flickr-palvelusta, ja sisältää huomattavasti toisistaan eroavia kuvia ihonvärin, iän, kuvan taustan, ja koristeiden, kuten esimerkiksi hattujen, silmä- ja aurinkolasien suhteen. Tausta-ajatuksenani oli kouluttaa ensin tekoälyä tunnistamaan ihmiskasvot kuvista ylipäätään FFHQ-aineiston avulla, ja yleisen mallin kouluttamisen jälkeen edetä kouluttamaan mallille kohdehenkilön kasvot. Muutamat tutkimistani lähteistä painottivat, että tällä tekniikalla lopputuloksesta tulee laadukkaampi kuin kouluttamalla mallia vain kohdehenkilöllä. Toin workspace-kansion ohjelman käyttöön ajamalla Manage workspace -kohdasta komennon Import from Drive, eli toin ohjelman käyttöön workspace-kansion Google Drive -kansioistani. Lopulta pääsin ajamaan mallille määrittämäni asetukset kohtaan Train model, jota kuva 26 havainnollistaa. Valitsin malliksi jo käyttämäni SAEHD, ja hyväksyin oletusasetukset, hyväksyen viimeisen valinnan Enable pretrain mode, jotta ohjelma esikouluttaa mallia FFHQ-aineistolla opettaakseen mallille ihmiskasvot ylipäätään.



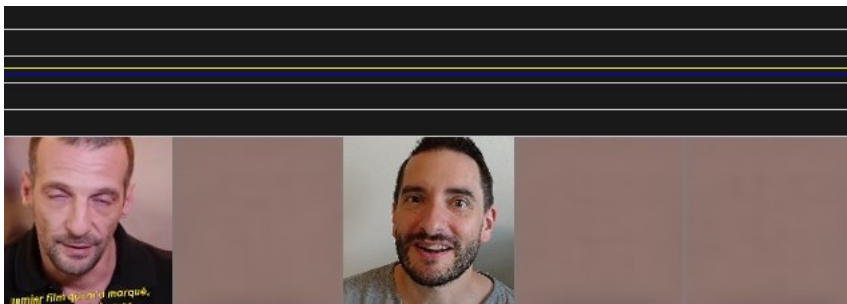
Kuva 26. Mallin koulutus FFHQ-aineistolla pretraining mode -asetuksella

Lopetin pretrain-mallin ajamisen hieman alle 13 000:n iteraation jälkeen. Luin muutamista lähteistäni, että liika esikoulutus vaihtelevilla ihmiskasvoilla kohdehenkilön sijasta saattaisi tehdä mallista vähemmän kohteen näköisen, joten lopetin nähdessäni esikatselukuvasta tekoälyn osaavan melko hyvin sijoittaa ihmisten kasvonpiirteet kohdilleen, kuten kuvasta 27 voi todeta.



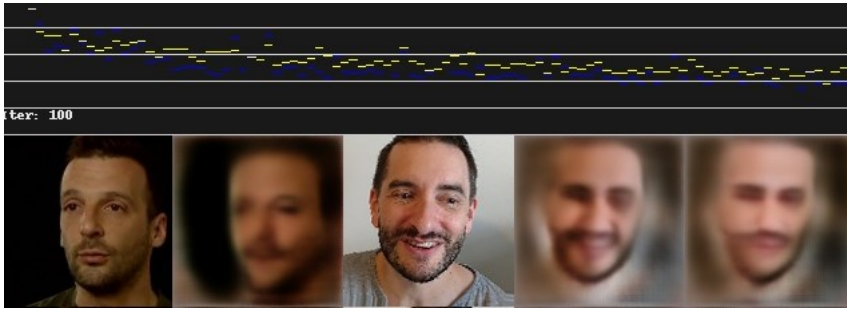
Kuva 27. Esikatselukuvat mallista, vasemmalla alkuperäinen, oikealla tekoälyn luoma

Seuraava askel oli muuttaa asetuksista pretraining mode pois päältä, ja alkaa kouluttaa mallia kohdehenkilön ja kohdevideon kuvilla. Alun esikatselukuvasta (kuva 28) voi hyvin todeta mallin kouluttamisen alkavan alusta. Yläpuolella häviölukemaa osoittava graafi on tyhjä, ja alkuperäisten kuvien vieressä tekoälyn luomat versiot ovat sumeat.



Kuva 28. Ensimmäinen esikatselukuva mallin koulutuksesta kohdehenkilöllä ja -videolla

Annettuani ohjelman ajaa jonkin aikaa, tekoälyn luoma kuva alkoi hahmottua, kuten kuvassa 29, jossa kasvot ovat jo näkyvillä hieman sumeana. Häviögraafi alkaa hahmottua, ja iteraatiolukema näkyy kuvassa vasemmalla graafin alla. Esikatselukuvassa vasemmanpuoleisin on alkuperäinen kuva kohdehenkilöstä, seuraava tekoälyn luoma kuva. Seuraavaksi on keskellä alkuperäinen kuva kohdevideoista, ja tekoälyn luoma kuva sen oikealla puolella. Viimeisin kuva oikealla on tekoälyn luoma kuva kohdehenkilöstä istutettuna kohdevideoon.

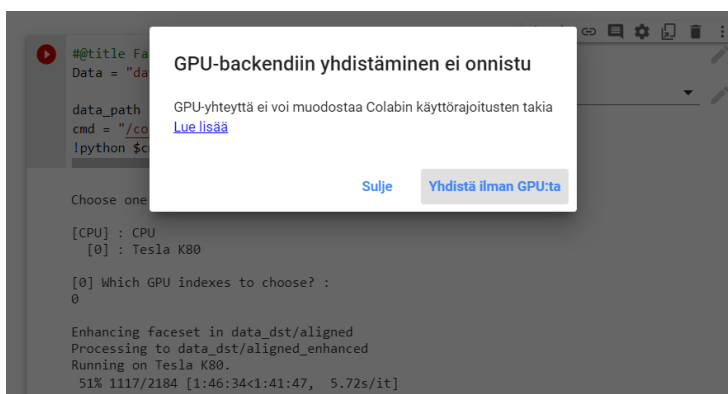


Kuva 29. Esikatselukuva mallista, iteraatioita 100

3.5.4 Google Colabın käyttörajoitusten aiheuttamat työskentelyvaikeudet

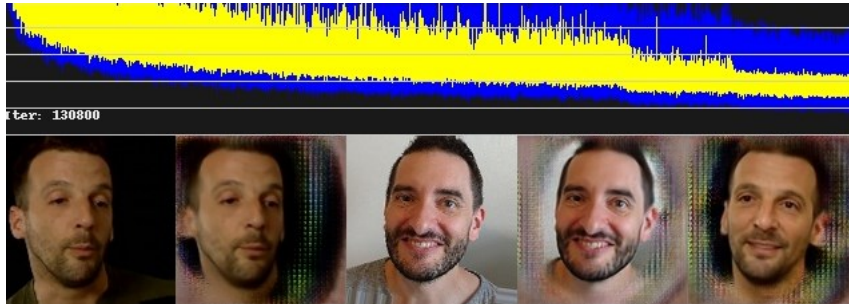
Alkaessani kouluttamaan mallia olin suunnitellut jättäväni koneen ajamaan yön yli mallin koulutusta. DeepFacelabin Google Colab-ohjelmakoodissa oli myös mahdollisuus ajaa skripti, eli komentosarja, jonka tarkoitus oli pitää sessio käynnissä, sillä Google tarkistaa aika-ajoin, käyttääkö ohjelmaa ihminen. Ikävä kyllä Google oli parantanut robotintunnistustaan luomalla CAPTCHA-kyselyn, joka tuli ruudulle epäsäännöllisin väliajoin. Mikäli ei riittävän pian raksittanut ruutua ja vakuuttanut, ettei ole robotti, ohjelma katkaisi yhteyden grafiikkasuorittimeen (kuva 30).

Colab saattoi katkaista yhteyden myös satunnaisesti, pyrkien jakamaan yhteyksiä käyttäjien välillä, ja tästä mahdollisuudesta on maininta myös käyttöehdoissa. Yhteyden katketua yllättäen varmuuskopion tekeminen viimeisimmästä ohjelman vaiheesta ei enää onnistunut, vaan käytettävissä oli ainoastaan ohjelman luoma varmuuskopio, jonka voi asettaa enintään yhden tunnin välein otettavaksi. Mikäli tunnin rajapyykki oli lähellä, ja ohjelman suoritus katkaistiin Googlen toimesta, tuo vajaan tunnin kuluessa tapahtunut edistys mallin kouluttamisessa menetettiin. Nämä seikat hidastivat mallin koulutustyötä huomattavasti, sillä session tultua katkaistuksi oli pääsääntöisesti odotettava seuraavaan päivään, jotta ohjelman ajamista voi jatkaa. Kiersin tätä hieman ottamalla käyttöön jo aiemmin toista tarkoitusta varten luomani ylimääräisen Google-tilin, ja lataamalla varmuuskopion sinne aloitin ohjelman ajamisen siitä mihin toisella Google-tilillä jäin. Varmuuskopiot olivat tosin useiden gigatavujen kokoisia, ja latausprosessikin vei aikaa.



Kuva 30. Google Colab yhteyden katkaisu

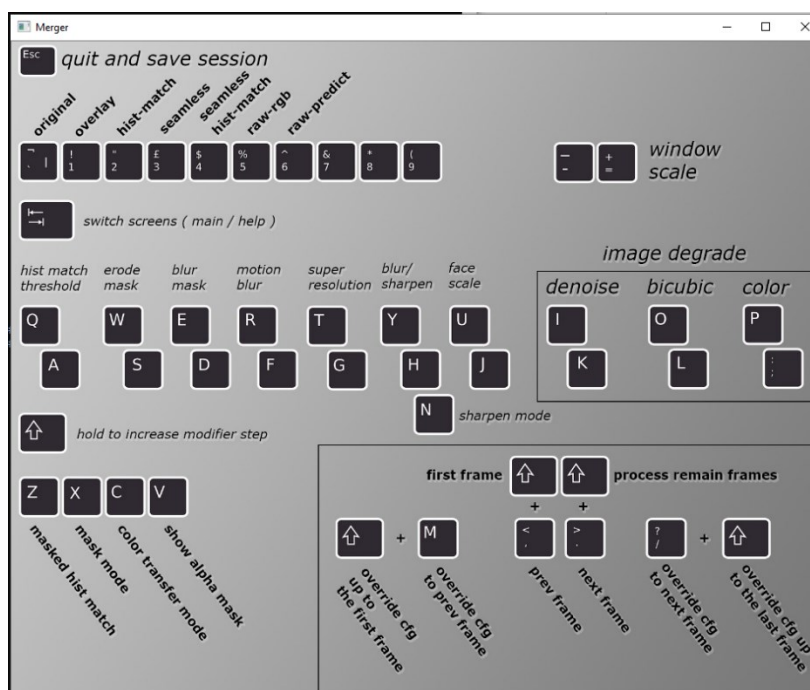
Jatkettuani työskentelyä yllä mainituista rajoitteista huolimatta, hieman yli 130 000 iteraation jälkeen esikatselukuvasta saattoi erottaa jo kasvojen ryppyjäkin, ja malli näytti mielestäni hyvältä (kuva 31). Häviölukema ei näyttänyt enää juurikaan laskevan, joten päätin yrittää videoksi yhdistämistä tässä vaiheessa.



Kuva 31. Esikatselukuva mallista, iteraatioita 130 800

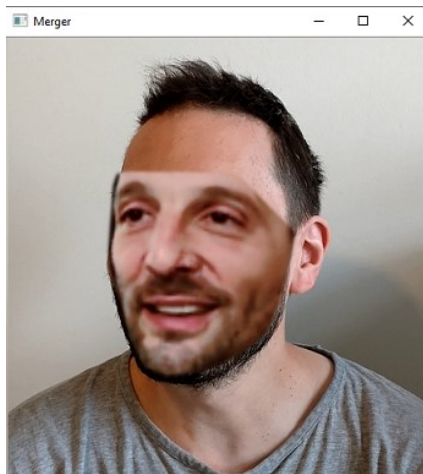
3.5.5 Kuvien yhdistäminen videoksi paikallisella koneella

Päätin tehdä kuvien yhdistämisen videoksi paikallisella koneella, jotta minulla olisi enemmän kontrollia lopputulokseen. DeepFaceLabin paikallinen versio tarjoaa videon luomiseen työkalun, jossa voi interaktiivisesti muokata mallin yhdistämisessä kuvia. Yhdistämistyökalulla voi kuva kavalta päättää esimerkiksi, kuinka suuri on liitettävä kloonattu kasvo suhteessa alkuperäiseen, kuinka sumeat rajat ovat ja miten laajasti kloonattua maskia käytetään suhteessa alla olevaan kuvaan. Käyttöliittymänä on näppäimistö, ja esikatselukuvan ja komento-ohjeikkunan välillä voi liikkua tabulaattoria painamalla (kuva 32).

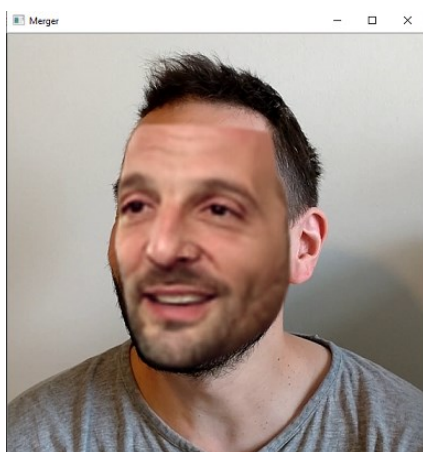


Kuva 32. Kuvien yhdistämisen komennot

Tässä yhdistämisvaiheessa huomasin käytännössä, kuinka vaikeaa oli saada alla oleva videokuva ja siihen istutettu tekoälyn luoma kasvomaski yhdistymään muodostaen huomattomasti lopputuloksen. Kasvojen skaalaaminen oikeaan kokoon oli vaikeaa arvioida, ja maskin rajat näkyivät selvästi kuvassa. Ihonväri poikkesi alkuperäisen videon, ja kloonattujen kasvojen suhteen, ja maskin saaminen oikean kokoiseksi, jotta se peittäisi alla olevan alkuperäisen kuvan vaati tasapainottelua. Huomasin, että suuressa osassa valitsemistani mallin koulutuskuviissa ollut musta tausta piirtyi esiin, mikäli suurensin maskin osuutta kuvassa, tehden kasvoihin mustat rajat (kuva 33). Kuvassa 34 maski on pienempi kuin sen alla olevat kohdevideon kasvat, joten rajan alkuperäisen ja kloonatun kasvon välillä näkee selvästi.



Kuva 33. Suurempi maski lisää kasvoihin mustat ääriiviivat mallin koulutuskuviin taustasta



Kuva 34. Liian pieni maski, jolloin kloonatun ja alkuperäisen kuvan raja näkyy selvästi

Huomasin myös videon kuvia käsitellessäni useita kohtia, joita voisi myöhemmissä versioissa parantaa. Olin tehnyt XSeg-maskin otsan osalta suurpiirteisesti, kuten opastukseksi katsomassani videossa, mutta tällä tavoin määritelty maski tekee selvän rajan otsaan. Olisi ollut parempi tehdä maski hiusrajaa selvemmin myötäillen, jotta myös lopullinen

maski seuraisi tarkemmin henkilön hiusrajaa. Olisin voinut myös valita alkuperäiseen kloonattavien kasvojen koulutuskuvajoukkoon vain kuvia, joissa tausta on vaalea, kuten kohdevideoissa, jotta maskin raja ei piirtyisi niin selvänä esiin. Lisäksi opin, että miesten parta on hankala kasvojen kloonauksessa, ja parasta olisikin, jos kohdevideoissa henkilöllä olisi parta ajettu, jotta se ei aiheuttaisi ongelmia ja piirtyisi taustalta selvästi esiin kuvia yhdistäessä. Eräässä lähteessä myös mainittiin, että naisten osalta meikin tulisi olla yhdenmukainen kloonattavien kasvojen aineistossa, samoin hiustyylin ja -värin tulisi pysyä samanlaisena.

Vaikka mallin koulutusvaiheen pienet esikatselukuvat vaikuttivat tarkoilta, videon yhdistämisvaiheessa huomasin kasvojen olevan selvästi utuisemmat kuin odotin. Yritin eri asetuksia yhdistelemällä saada kuvaa tarkemmaksi, mutta en onnistunut niin hyvin kuin halusin. Useat lähteet mainitsevatkin, että olisi hyvä jälkikäsitellä videota ja maskia sen päällä videoeditorilla kuten Adoben After Effects tai DaVinci Resolve, jotta lopputuloksesta saisi paremman. Yksi - joskin työläs - vaihtoehto olisi myös käsitellä kaikki kuvat ruutu ruudulta jollain valokuvanmuokkausohjelmalla. Utuiseen tai epätarkkaan kuvaan saattaa vaikuttaa myös mallin koulutuskuvat, sillä etenkin jos kohdehenkilön kuvat ottaa videoista, on useissa ruuduissa liike-epätarkkuutta. Lähtöaineiston huono laatu saattaa vaikuttaa niin, että lopullinen tulos ei ole aivan tarkka, vaikka iteraatioita jatkaisi pitkään.

Itse päätin vielä kouluttaa mallia yli 200 000:n iteraation, jonka jälkeen kokeilin uudelleen kuvien yhdistämistä videoksi. Tällä kertaa pääsin parempaan lopputulokseen, ja mielestäni etenkin pienessä koossa epätarkkuuteen ei kiinnittänyt niin paljon huomiota. Kuvassa 34 on paras lopputulos, johon pääsin. Vasemmalla on kuvien yhdistämisen asetukset, keskellä videoon istutetut kloonatut kasvat, ja oikealla vastaavan ruudun alkuperäinen kuva vertailtavaksi.



Kuva 35. Kuvien yhdistäminen videoksi, asetukset, lopputulos ja vertailukuva

Mielestäni lopputulos oli melko hyvä, ja video vaikutti melko uskottavalta. Valevideoksi sen tunnistaa esimerkiksi hampaiden epätarkkuudesta, sekä vasemman korvan vieressä olevasta vaaleasta varjosta, jota ei alkuperäisessä videossa ole. Lisäkuvia videosta on liitteessä 2.

4 Pohdintaa

Opinnäytetyössäni kokeilin useita eri tapoja tuottaa valevideoita, tavoitteenani selvittää onko kenen tahansa mahdollista luoda uskottava valevideo. Kokeilujeni, sekä aihetta tutkiessani näkemäni aineiston perusteella uskottavan valevideon luominen vaikuttaa olevan mahdollista kenelle tahansa, joskin työlästä opetella. Käyttökelpoisimmaksi kokeilussani osoittautui DeepFaceLab, jonka käytön hallitsemalla uskottavien valevideoiden tuottaminen onnistuu. Ohjelma pystyy luomaan valevideoita myös siinä tapauksessa, että kohdevideon henkilö ei katso suoraan kameraan, eikä siis aseta usean muun valevideoiden luomisohjelman tapaan rajoitteita kohdevideon suhteen. Vaatii kuitenkin useita kokeiluja, ja sinnikästä opettelua ottaa haltuun sekä valevideoiden luomisohjelma, että videoiden jälkikäsitteilyohjelma. Myös laadukas kuvamateriaali ohjelman kouluttamiseksi on avainasemassa uskottavan videon luomisessa. DeepFaceLab on hyvin laajalle levinnyt, ja sillä on luotu useita taidokkaita valevideoita.

Internetissä hyvä esimerkki uskottavasta valevideosta on DeepTomCruise, jonka luonut Chris Ume on sitä mieltä, että samanlaiseen työnjälkeen eivät tavallisen ihmisen taidot riittäisi, sillä hän on ammatiltaan visuaalisten tehosteiden taiteilija (Kahn 5.3.2021). Kahnin haastattelussa Ume paljastaa saaneensa avukseen Miles Fisherin, joka on taitava Tom Cruisen imitoija. Tekoälyn koulutusprosessi vei Umelta kaksi kuukautta kaikkineen, joka jälkeen malli oli riittävän valmis liitettäväksi lähes mihin tahansa videoon. Videoon liittäminen veisi vielä kahdesta kolmeen kokonaista päivää. Yhden videon jälkikäsitteilyyn meni 24 tuntia, sillä hän käsitteli jokaisen ruudun erikseen. (Kahn 5.3.2021.) Umen luomat valevideot ovat kuitenkin osoitus siitä, että riittävillä resursseilla – ajan, käytettävissä olevan teknologian, ja teknisen kyvykkyyden suhteen – on mahdollista tehdä ällistyttävän realistisia valevideoita, joskaan tällä hetkellä vasta-alkaja ei kykenisi yhtä hyvään työn laatuun.

Vaikka Chris Umen videota eivät valevideon tunnistusohjelmat tunnistaneetkaan, Ume on sitä mieltä, että se johtuu tunnistusohjelmien kouluttamisesta vanhemmalla, heikompi-
soisella materiaalilla (Kahn 5.3.2021). Osaltaan siis valevideoiden tunnistamisessa on tärkeää tuottaa riittävän laadukasta materiaalia automaattisille tunnistusohjelmille teknologian edistyessä. Mikäli tunnistusteknologia laahaa jäljessä kehityksen kärjestä, ei automaattista valevideoiden tunnistusohjelmaa voi pitää uskottavana.

Tässä opinnäytteen rajatussa aikataulussa en ehtinyt kokeilla, onko mahdollista luoda reaaliajassa valevideopuhelu, vaikka löysin taustatutkimusta tehdessäni ohjelmakoodin siihenkin tarkoitukseen DeepFaceLabin luojilta. Rajasin myös liikkeenkloonausohjelman pois tästä opinnäytteestä. Mikäli aikaa olisi jäänyt enemmän kokopäivätyöltä ja lapsiperheen arjelta, olisin kokeillut useita eri versioita DeepFaceLab-ohjelmalla, pyrkien

parantamaan lopputulosta kokeilu kokeilulta. Olisi myös ollut hyvä, jos aikaa olisi jäänyt enemmän opetella videoeditointiohjelmien käyttöä.

Aihe oli itselleni todella mielenkiintoinen, ja etenkin Covid-19:n aiheuttamien aikataulu-
muutosten (lasten sairastelu ja karanteenit, sekä oma sairastelu) aiheuttamien paineiden
keskellä, luoviessani vaativan työharjoittelun, opintojen viimeistelyn sekä perhe-elämän
yhdistämisen ristivedossa aiheen kiehtovuus auttoi jaksamaan vaikeuksista huolimatta.
Uskoisin jatkavani kokeiluja aikataulujen salliessa, sillä koin tämän myös hauskaksi pro-
jektiksi.

Useissa kokeilemissani valevideoiden luomisohjelmissa käytin apunani Google Colabia.
Omien kokemusten perusteella Google Colabia oli helppo käyttää, sillä parhaimmillaan
koodia testatakseen sitä ei tarvitse muokata ollenkaan, vaan ajaa vain järjestyksessä hii-
rellä klikkaillen moduulin kerrallaan. Tällöin ei tarvitse myöskään osata esimerkiksi ohjel-
mointikieltä kuten Python, saati osata käyttää esimerkiksi Linux-käyttöjärjestelmän komen-
tokehotetta, ja riittävien ohjeiden avulla kuka tahansa voi kokeilla ohjelman ajamista.
Google Colab poistaa myös taloudellisen kynnyksen kokeilla tekoälyohjelmaa pilvipalve-
lussa, sillä ohjelma on ilmainen, ja tarjoaa käyttöön tehokkaan grafiikkasuorittimen. Jollei
Colabia tai mahdollisesti vastaavaa internetin yli tarjottavaa pilvipalvelua käytä, on vaihto-
ehtoina ostaa itselleen hintava grafiikkasuoritin, tai ostaa pilvipalvelua joltain palveluntarjo-
ajalta, jolloin myös käyttöympäristöjen asennus vaatii asiantuntemusta. Mikäli jo omistaa
tietokoneen, jossa on tehokas grafiikkasuoritin esimerkiksi tietokonepelien pelaamista var-
ten, on kynnyksensä valevideotekniikan kokeiluun matala.

Koska lähestymistapani tutkimuskysymykseni testaamiseen oli tapaustutkimus, tärkeässä
roolissa oli lisätä ymmärrystä tutkittavasta aiheesta. Mielestäni työni vastaa hyvin kysy-
myksiin miten, miksi ja kuinka valevideoita luodaan. Koska en ollut itse aiemmin kyseisiä
teknologioita tai videoeditointityökaluja kokeillut, uskoisin oman prosessini olevan yleistet-
tävässä myös keneen tahansa, joka on kiinnostunut kokeilemaan valevideoiden luomista.
Toki koulutustaustani helpotti ohjelmien kokeilua, vaikka alussa myös itse epäröin kyky-
jäni. Päätin edetä järjestyksessä helpoimmista menetelmistä vaikeampiin luodakseni itsel-
leni onnistumisen kokemuksia, jotka toimivat kannustimina jatkaa vaikeampiin aiheisiin.
Yllätyksekseni myös lopussa kokeilemani ohjelmat olivat mielestäni suhteellisen helppo-
käyttöisiä, vaikka useissa lähteissä varoiteltiin niiden käyttämisen olevan vaativaa ja oppi-
miskäyrän jyrkkä.

Toisaalta etenkin DeepFaceLabin käyttäminen tavalla, joka tuottaa laadukkaita ja teräviä
valevideoita vaatii perehtymistä ja useita kokeiluja. Monissa löytämissäni lähteissä kerrot-
tiin myös, ettei ohjelman käyttöön ole mitään tarkkoja raja-arvoja tietyille vaiheille, vaan

kokeiluja tekemällä oppii juuri kyseiseen materiaaliin sopivat säädöt. Ohjelman käytön opettelu vaatii sinnikkyyttä kokeiluihin oppimisprosessissa, ja parhaimman tuloksen saavuttamiseksi useiden teknologioiden, kuten videoeditointityökalujen käyttötaito on tarpeen.

Omiin testeihini valikoitui poikkeuksellisen pandemia-ajan vuoksi ainoastaan miessukupuoli, mutta kiinnostavaa olisi ollut myös testata, miten naisilla esimerkiksi vaihteleva kasvomeikki vaikuttaa videon luomisprosessiin. Jatkotutkimuskohteena voisikin olla testien uusiminen tekniikan kehityttyä, sekä useiden vertailuvideoiden luominen käyttäen videoissa eri sukupuolia, etnisyyksiä, sekä vertaillen saman valevideon tuottamista useilla eri ohjelmilla, kuten DeepFaceLab ja Faceswap.

Tulevaisuuden uhkien suhteen omien kokeilujeni perusteella koen, että uskottavan valevideon luominen on vaativaa työtä, johon vaaditaan hyvä taustamateriaali kloonattavista kasvoista, riittävän samannäköinen henkilö kohdevideoon kuin kloonattava henkilö, sekä mahdollisuus joko äänen uskottavaan kloonamiseen tai taitava imitoija. Äänen uskottava kloonaminen vaatisi mittavan aineiston kloonattavan henkilön äänestä, sekä tuon aineiston translitteroinnin tekoälyn kouluttamiseen. Työ vaatii siis sekä riittäviä resursseja, että ammattitaitoa, jotta tuloksena syntynyt valevideo olisi erehdyttävän aidon oloinen. Uskottavan valevideon vaatima työmäärä jo asettaa rajoitteita siihen, mitkä tahot niitä voisivat tuottaa esimerkiksi vahingoittamistarkoituksessa. Toki tekniikka kehittyy jatkuvasti, ja ehkä jo lähitulevaisuudessa prosessi on huomattavasti helpompi.

Opin tämän opinnäytetyön tekemisen aikana paljon sekä valevideotekniikoista, tekoälystä, että pilvipalveluista, joissa valevideo-ohjelmia voi ajaa. Pehdyin myös hieman videoeditointityökaluihin, sekä kertosin ja syvensin osassa ohjelmista Linux-käyttöjärjestelmän käyttötaitojani. Vaikka varsinaista ohjelmointia en työssäni tarvinnut, oli ohjelmakoodin lukutaidosta hyötyä opetellessani Google Colabin käyttöä, sekä ajaessani kokeilemiani ohjelmia Google Colabissa. Projekti oli varmasti antoisa itselleni etenkin, koska se vaati sekä opintojeni aikana oppimien asioiden soveltamista, että uusien taitojen opettelua.

Lähteet

Aamulehti 16.1.2017. Keinoäly hoitaa tulevaisuudessa liki puolet työpaikoista. Luettavissa: <https://www.aamulehti.fi/kotimaa/art-2000007395088.html>. Luettu: 31.5.2021

Ahokas, K. 2020. Tekoäly murentaa videokuvan uskottavuutta – ”Kyllä deepfakesta tavallaan kannattaa olla huolissaan”. Luettavissa: <https://www.tivi.fi/uutiset/tekoaly-murentaa-videokuvan-uskottavuutta-kylla-deepfakesta-tavallaan-kannattaa-olla-huolissaan/706b8f5d-411a-4004-8bef-5f15d0cc379c>. Luettu: 24.4.2021

Aitamurto, T. 13.11.2018. Syvät valevideot ovat uusin ase informaatiiosodassa – ja niitä on nykyään huolestuttavan vaivaton tehdä. Luettavissa: <https://www.hs.fi/mielipide/art-2000005897195.html>. Luettu: 18.4.2021

Arjoranta, J. 2019. Tekoäly ja koneoppiminen ovat käsitteinä väärinymmärrettyjä ja siksi väärinkäytettyjä. Luettavissa: <https://bearit.fi/tekoaly-ja-koneoppiminen-ovat-kasitteina-vaarinymmarrettyja-ja-siksi-vaarinkaytettyja/>. Luettu: 25.4.2021

Biard, G. 2017. Mathieu Kassovitz Cannes 2017. Valokuva. Luettavissa: https://commons.wikimedia.org/wiki/File:Mathieu_Kassovitz_Cannes_2017.jpg. Luettu: 15.8.2021

Chan, C., Ginosar, S., Efros, A., Zhou, T. 2019. Everybody Dance Now. IEEE International Conference on Computer Vision (ICCV). Luettavissa: <https://arxiv.org/pdf/1808.07371.pdf>. Luettu: 11.9.2021

Chivers, T. 23.7.2019. What do we do about deepfake video? The Guardian. Luettavissa: <https://www.theguardian.com/technology/2019/jun/23/what-do-we-do-about-deepfake-video-ai-facebook>. Luettu: 26.5.2021

Coldewey, D. 12.6.2020. Facebook's 'Deepfake Detection Challenge' yields promising early results. Luettavissa: <https://techcrunch.com/2020/06/12/facebooks-deepfake-detection-challenge-yields-promising-early-results/>. Luettu: 9.10.2021

Cole, S. 2019a. This Horrifying App Undresses a Photo of Any Woman With a Single Click. Luettavissa: <https://www.vice.com/en/article/kzm59x/deepnude-app-creates-fake-nudes-of-any-woman>. Luettu: 14.8.2021

- Cole, S. 2019b. Creator of DeepNude, App That Undresses Photos of Women, Takes It Offline. Luettavissa: <https://www.vice.com/en/article/qv7agw/deepnude-app-that-undresses-photos-of-women-takes-it-offline>. Luettu: 14.8.2021
- Corentin, J., 2019. Real-Time Voice Cloning. Pro gradu -tutkielma. Université de Liège. Luettavissa: <https://matheo.uliege.be/handle/2268.2/6801>. Luettu: 4.9.2021
- Favreau, J. 2008. Iron Man. Paramount Pictures. Video. Katsottavissa: <https://www.youtube.com/watch?v=8hYIB38asDY>. Katsottu: 25.8.2021
- Fowler, G. 25.3.2021. Anyone with an iPhone can now make deepfakes. We aren't ready for what happens next. Luettavissa: <https://www.washingtonpost.com/technology/2021/03/25/deepfake-video-apps/>. Luettu: 3.8.2021
- Garbade, M. 2021. What is Google Colab? Luettavissa: <https://blog.education-ecosystem.com/what-is-google-colab/>. Luettu: 4.9.2021
- Gilbey, R. 4.9.2019. A 'deep fake' app will make us film stars – but will we regret our narcissism? The Guardian. Luettavissa: <https://www.theguardian.com/technology/2019/sep/04/a-deep-fake-app-will-make-us-film-stars-but-will-we-regret-our-narcissism>. Luettu: 31.7.2021
- Harhanen, M. 2019. Jupyter Notebook. Luettavissa: https://viikonvalo.fi/jupyter_notebook/. Luettu: 4.9.2021
- Harwell, D. 14.5.2021. Remember the 'deepfake cheerleader mom'? Prosecutors now admit they can't prove fake-video claims. Luettavissa: <https://www.washingtonpost.com/technology/2021/05/14/deepfake-cheer-mom-claims-dropped/>. Luettu: 14.8.2021
- Kahn, J. 5.3.2021. Deepfake master behind those viral Tom Cruise videos says the technology should be regulated. Luettavissa: <https://fortune.com/2021/03/05/tom-cruise-deepfake-creator-technology-should-be-regulated/>. Luettu: 29.9.2021
- Katwala, A. 3.10.2021. Could The Simpsons Replace Its Voice Actors With AI? Luettavissa: <https://www.wired.com/story/simpsons-voice-actors-ai-deepfakes/>. Luettu: 14.8.2021
- Kwan, W. 2020. I Cloned Billie Eilish's Voice & Face with AI. Video. Katsottavissa: <https://youtu.be/LSh2tQ6OyZs>. Katsottu: 4.8.2021

- Kärkkäinen, H. 20.10.2020. Uusi kasvot upottava sovellus leviää nopeasti Suomessa – tiedä nämä asiat ennen kuin lataat. Ilta-Sanomat. Luettavissa: <https://www.is.fi/digi-today/mobiili/art-2000006675389.html>. Luettu: 31.7.2021
- Lanham, M. 2021. Generating a New Reality: From Autoencoders and Adversarial Networks to Deepfakes. Apress. E-kirja. Luettu: 17.8.2021
- Lee, D. 2019. Deepfake Salvador Dalí takes selfies with museum visitors. Luettavissa: <https://www.theverge.com/2019/5/10/18540953/salvador-dali-lives-deepfake-museum>. Luettu: 14.8.2021
- Lee, T. 2019. I created my own deepfake—it took two weeks and cost \$552. Luettavissa: <https://arstechnica.com/science/2019/12/how-i-created-a-deepfake-of-mark-zuckerberg-and-star-treks-data/>. Luettu: 2.8.2021
- Leinonen, R. 2020. Tapaustutkimus – tutkimus tapauksesta. Luettavissa: <https://spoken.fi/tapaustutkimus/>. Luettu: 28.9.2021
- Lomas, N. 2021. AI-driven audio cloning startup gives voice to Einstein chatbot. Luettavissa: <https://techcrunch.com/2021/04/16/ai-driven-audio-cloning-startup-gives-voice-to-einstein-chatbot/>. Luettu: 14.8.2021
- Mahdawi, A. 18.7.2020. Is FaceApp an evil plot by 'the Russians' to steal your data? Not quite. Luettavissa: <https://www.theguardian.com/commentisfree/2019/jul/18/faceapp-privacy-data-russians>. Luettu: 2.8.2021
- Marshall, P. 2019. Defense against deepfakes. Luettavissa: <http://www3.gcn.com/Articles/2019/05/24/deepfake-detection.aspx>. Luettu: 3.8.2021
- Microsoft News Center 2018. Tekoälyn perusteet: koneoppiminen, työn tulevaisuus ja hyvä vai paha tekoäly. Luettavissa: <https://news.microsoft.com/fi-fi/2018/07/25/tekoalyn-perusteet-koneoppiminen-tyon-tulevaisuus-ja-hyva-vai-paha-tekoaly/>. Luettu: 25.4.2021
- Nelson, A. 2020. Computer Vision: Python Face Swap & Quick Deepfake in Colab. Udemy. E-kurssi. Katsottu: 9.8.2021

Niemi, O. 2.3.2020. Twitteriin riitelemään ilmestyi Asko Leppänen, jota ei ole olemassa – hän on seurausta suomalaisen ryhmän tutkimuksesta, joka aiheutti sensaation. Luettavissa: <https://www.hs.fi/nyt/art-2000006425258.html>. Luettu: 13.5.2021

Peckham, J. 2020. What is FaceApp? How to use the popular social media photo editor app. Luettavissa: <https://www.techradar.com/how-to/what-is-faceapp-how-to-use-the-popular-social-media-photo-editor-app>. Luettu: 2.8.2021

Perov, I., Daiheng, G., Chervoniy, N., Liu, K., Maragonda, S., Umé, C., Dpfks, M., Facenheim, C., RP, L., Jian, J., Zhang, S., Wu, P., Zhou, B., Zhang, W. 2020. DeepFaceLab: Integrated, flexible and extensible face-swapping framework. Luettavissa: <https://arxiv.org/abs/2005.05535>. Luettu 31.7.2021

Perov, I., Daiheng, G., Chervoniy, N., Liu, K., Maragonda, S., Umé, C., Dpfks, M., Facenheim, C., RP, L., Jian, J., Zhang, S., Wu, P., Zhou, B., Zhang, W. 2018. DeepFaceLab. Github-julkaisuarkisto. Luettavissa: <https://github.com/iperov/DeepFaceLab>. Luettu: 31.7.2021

Prajwal, K., Mukhopadhyay, R. Namboodiri V. & Jawahar, C. 2020. Wav2Lip. Github-julkaisuarkisto. Luettavissa: <https://github.com/Rudrabha/Wav2Lip>. Luettu: 24.10.2021

Raimi, S. 2002. Spider-Man. Columbia Pictures. Video. Katsottavissa: <https://www.youtube.com/watch?v=M-9ZgKmdsv4>. Katsottu: 28.8.2021

Reaktor & Helsingin Yliopisto 2018. Elements of AI. Neuroverkkojen periaatteet. Luettavissa: <https://course.elementsofai.com/fi/5/1>. Luettu: 16.5.2021

Roy, C. s.a. Creating Realistic Deepfakes With DeepFaceLab. Luettavissa: <https://medium.com/geekculture/realistic-deepfakes-with-deepfacelab-530e90bd29f2>. Luettu 31.7.2021

Saaranen-Kauppinen, A. & Puusniekka, A. 2006. 5.5 Tapaustutkimus. KvaliMOTV - Menetelmäopetuksen tietovaranto. Tampere. Luettavissa: <https://www.fsd.tuni.fi/menetelma-opetus>. Luettu 28.9.2021

Sample, I. 13.1.2020. AI-generated fake videos are becoming more common (and convincing). Here's why we should be worried. The Guardian. Luettavissa: <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them>. Luettu: 18.4.2021

Sandusky, M. 2021. The Best DeepFake Software – How To Make A Deepfake. Luettavissa: <https://peerthroughmedia.com/deepfake-software-testing/>. Luettu: 9.10.2021

Siarohin, A., Lathuilière, S., Tulyakov, S., Ricci, E., Sebe, N. 2019. First Order Motion Model for Image Animation. Advances in Neural Information Processing Systems. Curran Associates, s. 7137–7147.

Sipilä, A. 11.10.2020. Väärennettyjä videoita ei enää erota aidoista, ja pornosta on lyhyt askel poliittiseen sabotaasiin, varoittaa deepfake-tutkija Nina Schick. Luettavissa: <https://www.hs.fi/ulkomaat/art-2000006665981.html>. Luettu: 18.4.2021

Vaiste, J. 2017. Koneoppimisen muodot ja niiden eettiset näkökulmat. Luettavissa: <https://juhovaiste.fi/fi/koneoppimisen-muodot-ja-niiden-eettiset-n%C3%A4k%C3%B6kulmat>. Luettu: 3.8.2021

Vella, V. 13.3.2021. A Bucks County woman created 'deepfake' videos to harass rivals on her daughter's cheerleading squad, DA says. Luettavissa: <https://www.inquirer.com/news/bucks-county-raffaela-spone-cyberbullying-deepfake-20210312.html>. Luettu: 14.8.2021

Venkataramakrishnan, S. 2019. Why 'deepfake' videos mean you can no longer believe what you see. Luettavissa: <https://www.afr.com/technology/can-you-believe-your-eyes-20191028-p534wl>. Luettu: 14.8.2021

Viira, K. 20217. Tunnetko nämä 10 uuden ajan digitermiä? Luettavissa: <https://yriyksille.elisa.fi/ideat/tunnetko-nama-10-uuden-ajan-digitermia/>. Luettu: 9.10.2021

Virahonda, S. 2021. DeepFaceLab: A Pre-Packaged Alternative DIY Deep Fakes. Luettavissa: <https://www.codeproject.com/Articles/5298031/DeepFaceLab-A-Pre-Packaged-Alternative-DIY-Deep-Fa>. Luettu: 4.9.2021

Westerlund, M. 2019. The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review, 9, 11, s.39–52. Luettavissa: <https://timreview.ca/article/1282>. Luettu: 18.4.2021

Wilson, M. 2019. Startups are racing to commercialize DeepFakes's powerful, internet-breaking AI. Luettavissa: <https://www.fastcompany.com/90337767/deepfakes-started-a-war-on-whats-real-these-startups-are-racing-to-commercialize-it>. Luettu: 14.8.2021

Liitteet

Liite 1. Deepfakesweb kierrosten 1–4 tulokset



Liite 2. DeepFaceLab-ohjelmalla tuotetun valevideon still-kuvia

Vasemmalla asetukset, keskellä still-kuva valevideosta, oikealla vastaava alkuperäinen ruutu

