Datan merkitys kasvojentunnistuksessa

Turun yliopisto
Tietotekniikan laitos
LuK-tutkielma
Tietojenkäsittelytiede
Toukokuu 2025
Jaakko Nurminen

Ohjaaja:
Jari Björne

TURUN YLIOPISTO
Tietotekniikan laitos

Jaakko Nurminen: Datan merkitys kasvojentunnistuksessa

LuK-tutkielma, 29 s.
Tietojenkäsittelytiede
Toukokuu 2025

Kasvojentunnistus on kehittynyt merkittävästi perinteisistä tilastollisista menetelmistä syväoppimiseen perustuvien neuroverkkojen myötä. Erityisesti konvoluutioneuroverkot mahdollistavat entistä tarkemman kasvojen piirteiden tunnistamisen, mutta ne edellyttävät suurta määrää monipuolista ja laadukasta koulutusdataa.

Tässä kandidaatintutkielmassa tarkastellaan koulutusdatan roolia kasvojentunnistusmallien kehittämisessä sekä sitä, miten datan rakenne, laatu ja demografinen edustavuus vaikuttavat mallien toimivuuteen, yleistämiskykyyn ja oikeudenmukaisuuteen. Tutkielma on toteutettu kirjallisuuskatsauksena, ja se vastaa neljään tutkimuskysymykseen, jotka käsittelevät kasvojentunnistusmallien rakennetta, oppimisprosessia, teknologian haasteita sekä datan aiheuttamia virheitä ja vinoumia.

Tarkastelussa hyödynnetään ajankohtaisia tutkimuksia ja esimerkkejä, kuten virheellisten tunnistusten seurauksia viranomaistoiminnassa. Tutkimuksessa havaitaan, että kasvojentunnistusjärjestelmien oikeudenmukaisuus on suoraan sidoksissa koulutusdatan kattavuuteen ja edustavuuteen. Vaikka synteettinen data tarjoaa mahdollisuuksia edustavuuden laajentamiseen, sen käyttöön liittyy myös teknisiä ja eettisiä haasteita.

Johtopäätöksissä korostetaan tarvetta eettisemmälle datan keruulle, suuremmalle läpinäkyvyydelle sekä yhteiskunnalliselle keskustelulle biometrisen datan käytöstä. Teknisten ja eettisten haasteiden tunnistaminen on olennainen osa kasvojentunnistusteknologian kehittämistä luotettavampaan ja käyttökelpoisempaan suuntaan.

Asiasanat: kasvojentunnistus, konvoluutioneuroverkko, koulutusdata, data-etiikka

Sisällys

Kuvat

Taulukot

1 Johdanto

Kasvojentunnistus (engl. face recognition) on konenäön sovellusalue, jossa pyritään tunnistamaan digitaalisessa kuvassa esiintyvä henkilö analysoimalla tämän kasvonpiirteitä. Kasvojentunnistuksen tavoitteena on identifioida tietty ihminen, kun taas kasvojen havaitsemisessa (engl. face detection) tavoitteena on vain havaita kuvassa esiintyvät kasvot [1]. Kasvojentunnistuksella on monta erilaista käyttökohdetta ja sovellusaluetta. Yksilölle kasvojentunnistus voi olla tuttua esimerkiksi puhelimen lukitusnäytön avaamisen yhteydestä, sosiaalisesta mediasta, turvakamerajärjestelmistä, lentokenttien passintarkastuksesta tai kuvakirjastojen henkilöidentunnistamisominaisuuksista.

Tekoäly on kehittynyt valtavasti viime vuosina grafiikkasuorittimien kehityksen ja suurten data-aineistojen saatavuuden johdosta. Kuvantunnistuksessa erityisesti syväoppiminen ja konvoluutioneuroverkot ovat mullistaneet erilaiset kuvantunnistustehtävät. Syväoppivat neuroverkot tarjoavat huomattavasti tarkemman kuvantunnistuksen perinteisiin sääntöpohjaisiin algoritmeihin verrattuna, mutta vaativat kuitenkin paljon koulutusdataa ja laskentatehoa. [2]

Yksi keskeisimmistä haasteista kasvojentunnistusmalleja kehitettäessä on tarve riittävän suurelle ja laadukkaalle data-aineistolle [2][3]. Koneoppimismallien kehittämiseen käytettävä koulutusdata koostuu yleisesti tuhansista valokuvista, joihin on mahdollisesti liitetty tietoa kasvojen asennosta, valaistuksesta, kuvan taustasta, ka-

meran laadusta, kuvassa olevan henkilön sukupuolesta tai kasvoilla olevasta ilmeestä. Toimenpidettä, jossa kuvaan liitetään tietoa, kutsutaan data-annotaatioksi. [1]

Puutteellinen data-annotaatio tai liian suppea koulutusaineisto voivat heikentää koneoppimismallin kykyä yleistää oppimaansa uuteen, koulutusaineistoon kuulumattomaan dataan. Tämä saattaa johtaa mallin heikompaan suorituskykyyn tietyillä väestöryhmillä, erilaisissa valaistusolosuhteissa tai muissa vaihtuvissa olosuhteissa. Heikko suorituskyky aiheuttaa vastavuoroisesti virheitä tai epätarkkoja tuloksia. Tällaiset puutteet heikentävät mallin luotettavuutta ja oikeudenmukaisuutta käytännön sovelluksissa. [3][4]

Tämän tutkielman tavoitteena on tarkastella, millainen rooli koulutusdatalla on kasvojentunnistusmallien kehittämisessä, ja miten datan laatu vaikuttaa mallien tarkkuuteen. Lisäksi tutkielmassa perehdytään datan keräämiseen liittyviin haasteisiin, kasvojentunnistuksen virheiden ja puolueellisuuden vähentämiseen sekä kasvojentunnistusteknologioiden kehittämiseen liittyviin eettisiin kysymyksiin. Tutkielma on toteutettu kirjallisuuskatsauksena, ja sen tarkoituksena on muodostaa ajankohtainen kokonaiskuva siitä, kuinka koulutusdatan hallinta ja eettiset haasteet vaikuttavat kasvojentunnistusmenetelmien kehittämiseen. Tutkielman tutkimuskysymykset ovat seuraavat:

- TK1: Millainen on kasvojentunnistusmallin rakenne ja kuinka malli oppii sille annetusta koulutusdatasta?
- TK2: Millaisia haasteita liittyy syväoppimispohjaiseen kasvojentunnistukseen?
- TK3: Millä tavoin kasvojentunnistusjärjestelmissä esiintyy virheitä ja vinoumia, sekä mistä nämä ongelmat johtuvat?
- TK4: Miten koulutusdatan laatu ja koostumus vaikuttavat kasvojentunnistusjärjestelmien tarkkuuteen ja oikeudenmukaisuuteen?

Tutkielmassa käytetty aineisto on haettu IEEE- ja ACM-hakukannoista sekä Google Scholarista. Aineiston rajaamiseksi hakulausekkeissa on hyödynnetty koneoppimiseen ja kasvojentunnistukseen liittyviä avainsanoja, kuten *facial recognition*, *data annotation* ja *dataset bias*.

Hakutuloksia karsittiin aluksi niiden otsikoiden ja tiivistelmien perusteella, jotta voitiin tunnistaa aihepiirin kannalta oleellisimmat julkaisut. Tämän jälkeen lopullisesti aineistossa käytetyt julkaisut valikoitiin niiden sisällön analyysin perusteella, jolloin varmistettiin, että ne käsittelevät tutkielman aihetta riittävän syvällisesti ja vastaavat asetettuihin tutkimuskysymyksiin. Hakukannoissa käytettyjen hakulausekkeiden, niiden tuottamien hakutulosten ja näistä valittujen julkaisujen määrät on koottu tarkemmin taulukkoon 1.1.

Tutkielman rakenne koostuu johdannon, yhteenvedon ja johtopäätösten lisäksi kolmesta käsittelyluvusta. Luvussa 2 määritellään kasvojentunnistukseen liittyvät keskeiset käsitteet ja esitellään konvoluutioneuroverkkojen yleinen tekninen rakenne. Lisäksi luvussa tarkastellaan, kuinka syväoppivat neuroverkot oppivat käsittelemään niille syötettyä koulutusdataa. Luvussa 3 keskitytään kasvojentunnistusmallien virheisiin ja vinoumiin. Siinä tarkastellaan näiden ilmiöiden taustalla vaikuttavia tekijöitä sekä niiden vaikutuksia mallien oikeudenmukaisuuteen ja luotettavuuteen. Luku 4 käsittelee puolestaan koulutusdatan roolia kasvojentunnistuksessa, mukaan lukien datan keräämiseen ja annotointiin liittyvät haasteet sekä eettiset näkökulmat. Tutkielma päättyy yhteenvetoon ja johtopäätöksiin, joissa vastataan tutkimuskysymyksiin sekä pohditaan kasvojentunnistuksen kehityksen tulevaisuuden suuntaviivoja koulutusdatan näkökulmasta.

Taulukko 1.1: Hakukannoissa käytetyt hakulausekkeet ja hakutulokset

Hakukanta	**Hakulauseke**	**Yhteensä**	**Otsikko**	**Valittu**
ACM	(”privacy concerns” OR ”data protection”) AND (”facial recognition” OR ”face recognition”) AND (”annotation ethics”)	33	5	1
ACM	”facial recognition” AND ”methods” OR ”techniques” AND ”computer vision” OR ”convolutional neural network”	211	8	1
ACM	(”face recognition dataset” OR ”facial recognition dataset”) AND (”bias” OR ”fairness” OR ”ethical issues”)	48	7	2
IEEE	(”All Metadata”:facial recognition) AND (”All Metadata”:data annotation) AND (”All Metadata”:performance)	67	14	1
IEEE	(”face recognition” OR ”facial recognition”) AND (”data quality” OR ”label quality”) AND (”model generalization”)	5445	13	1
Google Scholar	”face* recognition” AND (”annotation methods” OR ”annotation techniques”)	97	13	4

Hakukanta: Käytetty tietokanta (esim. Google Scholar, IEEE, ACM). **Hakulauseke**: Hakulause, jolla haut tehtiin. **Yhteensä**: Hakutulosten kokonaismäärä. **Otsikko**: Otsikon perusteella valitut julkaisut. **Valittu**: Lopullisesti tutkielmaan valitut julkaisut.

2 Kasvojentunnistuksen menetelmät

2.1 Perusteet ja koneoppiminen

Teknologian kehittyessä kasvojentunnistus on muuttunut yksinkertaisista, osin manuaalisista menetelmistä yhä kehittyneempiin, automaattisesti oppiviin järjestelmiin. Aiemmin kasvojen tunnistaminen perustui ennalta määriteltyihin piirteisiin ja tilastollisiin malleihin, mutta viime vuosina syväoppimisen ja konvoluutioneuroverkkojen (engl. convolutional neural network, CNN) kehitys on mullistanut alan. Ensimmäiset esimerkit toimivasta kasvojentunnistusteknologiasta ovat peräisin vuodelta 1991, jolloin tutkijat Alex Pentland ja Matthew Turk esittelivät Eigenfaces-algoritmin. [5] Tämä menetelmä perustuu koulutuskuvista laskettuun keskimääräiseen kasvokuvaan ja algoritmin määrittelemiin kasvojen pääpiirteisiin, jotka esitetään lineaarisessa muodossa. Eigenfaces-algoritmilla tehtävässä kasvojentunnistuksessa uusi kasvokuva sijoitetaan kuva-avaruuteen, ja luokitellaan vertaamalla sitä lähimpään tunnettuun kasvokuvaan [6]. Muita perinteisiä kasvojentunnistusmenetelmiä ovat esimerkiksi lineaarinen diskriminanttianalyysi (engl. linear discriminant analysis, LDA) ja paikalliset binäärikuviohistogrammit (engl. local binary pattern histogram, LBPH), jotka olivat laajasti käytössä ennen syväoppimispohjaisten mallien yleistymistä [6].

Toisin kuin Eigenfaces, lineaarinen diskriminanttianalyysi toimii luokittelemalla kuvat luokkiin ja laskemalla kullekin luokalle oman keskiarvon. Täten algoritmi

maksimoi eri luokkien välisen hajonnan ja minimoi luokkien sisäisen hajonnan laskemalla kullekin luokalle erillisen keskiarvon hyödyntämällä hajontamatriiseja. Kasvojentunnistuksessa lineaarinen diskriminanttianalyysi hyödyntää näitä matriiseja ja soveltaa ominaisarvoja sekä -vektoreita erottaakseen luokat toisistaan. Lineaarisen diskriminanttianalyysin etuna Eigenfaces-menetelmään on sen parempi suorituskyky vaihtelevissa valaistusolosuhteissa, ilmeiden muutoksissa ja kasvojen asennon vaihteluissa. [6]

Paikalliset binäärikuviohistogrammit kehitettiin alun perin tekstuuriluokitteluun, mutta myöhemmin menetelmää sovellettiin myös kasvojentunnistukseen. LBP-algoritmi toimii vertaamalla kuvan pienimpiä yksittäisiä osia, eli yhdestä tai useammasta luvusta koostuvaa pikseliä, sitä ympäröiviin pikseleihin. Pikselien vertailulla algoritmi muodostaa binääriluvun, jossa keskimmäistä pikseliä suuremmat tai yhtä suuret arvot merkitään ykkösiksi ja pienemmät nolliksi. Näin syntyneet paikallispiirteet yhdistetään histogrammiksi, jota voidaan verrata muihin kuviin erilaisten vertailutekniikoiden, kuten euklidisen metriikan, khiin neliö -testin tai piirteiden itseisarvojen avulla. [6][7]

Kasvojentunnistuksen perinteiset menetelmät koostuvat tyypillisesti kolmesta päävaiheesta: kasvojen havaitsemisesta, piirteiden erottelusta ja lopullisesta luokittelusta [8]. Näissä menetelmissä tunnistus perustuu usein tilastollisiin malleihin ja ennalta määriteltyihin piirteisiin, kuten kasvojen muotoon, ihon tekstuuriin tai tiettyihin kasvojen rakenteellisiin ominaisuuksiin. Eigenfaces-menetelmä hyödyntää matriisilaskentaa kasvojen pääkomponenttien erotteluun, kun taas paikalliset binäärikuviohistogrammit keskittyvät tekstuuripiirteisiin [6]. Vaikka nämä menetelmät saavuttavat kohtuullisen tarkkuuden tietyissä olosuhteissa, niiden tehokkuus heikkenee valaistuksen muutosten, kasvonilmeiden vaihteluiden ja muiden häiriötekijöiden esiintyessä [7][9].

Koneoppimisen myötä kasvojentunnistus on kehittynyt merkittävästi, ja uudet menetelmät ovat syrjäyttäneet perinteiset algoritmit niiden rajoitteiden vuoksi. Erityisen tehokkaiksi menetelmiksi on todettu varsinkin syväoppivat monikerroksiset neuroverkot, joiden kehitys on saanut inspiraationsa ihmisaivojen hermosolujen muodostamista verkostoista, ja jotka etäisesti jäljittelevät näiden rakennetta [2][3].

Syväoppiminen on koneoppimisen osa-alue, joka keskittyy monimutkaisten ja haastavien ongelmien ratkaisemiseen – erityisesti sellaisten, joihin perinteiset koneoppimismenetelmät eivät helposti sovellu. Koneoppiminen yleisesti tarkoittaa ohjelmistojen kykyä oppia suurista datamääristä, parantaa suorituskykyään kokemuksen myötä ja tehdä päätöksiä erilaisten (usein tilastollisten) menetelmien avulla ilman tarkkaa manuaalista ohjelmointia. Syväoppiminen vie tämän vielä pidemmälle hyödyntämällä syviä, monikerroksisia neuroverkkoja [2].

Syväoppimismallit, kuten konvoluutioneuroverkot, residuaaliset neuroverkot ja siamilaiset neuroverkot, eivät perustu etukäteen määriteltyihin piirteisiin, vaan ne oppivat itse erottelemaan kasvojen ominaisuuksia monikerroksisten neuroverkkojen avulla. Tämä mahdollistaa abstraktimpien ja informatiivisempien kasvojen representaatioiden luomisen, mikä parantaa tunnistustarkkuutta merkittävästi. Syväoppimispohjaisilla menetelmillä kasvojentunnistusjärjestelmät pystyvät saavuttamaan hyvin korkean tarkkuuden, joka on lähellä tai jopa parempi kuin ihmisen suorituskyky vastaavissa tunnistustehtävissä. [8]

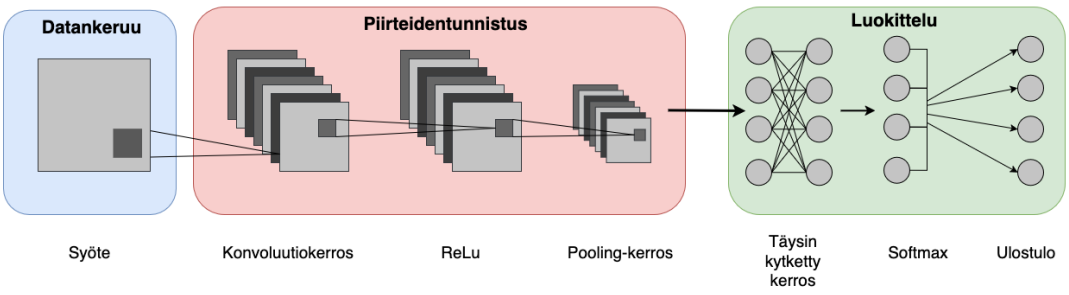
Konvoluutioneuroverkot ovat yksi syväoppimisen merkittävimmistä innovaatioista kasvojentunnistuksessa. Toisin kuin perinteiset menetelmät, konvoluutioneuroverkot käyttävät kerroksellista rakennetta, jossa dataa käsitellään hierarkkisesti yhä abstraktimmilla tasoilla. Tämä mahdollistaa kasvojen entistä tarkemman tunnistamisen myös vaihtelevissa olosuhteissa. Seuraavassa aliluvussa käsitellään tarkemmin: tyypillisen konvoluutioneuroverkon arkkitehtuuria, mitä kerroksia konvoluutioneu-

roverkon rakenteeseen saattaa kuulua ja mikä on kunkin kerroksen tehtävä osana kokonaisuutta.

2.2 Konvoluutioneuroverkkojen arkkitehtuuri

Konvoluutioneuroverkko on syväoppiva neuroverkko, joka soveltuu erityisen hyvin kaksiulotteisten kuvien analysointiin. Neuroverkot koostuvat useista eri kerroksista, joista jokainen suorittaa niille ominaisia operaatioita tulkitakseen kuvan sisällön. Neuroverkon ensimmäistä kerrosta, johon analysoitavat kuvat syötetään, kutsutaan syötekerrokseksi. Vastaavasti viimeistä kerrosta, joka tuottaa lopullisen ennusteen, kutsutaan ulostulokerrokseksi. Näiden välissä sijaitsevat piilokerrokset. Syvällä neuroverkolla viitataankin neuroverkkoon, jossa on enemmän kuin kolme piilokerrosta. [6]

Syväoppivalla konvoluutioneuroverkolla ei ole yhtä vakiintunutta rakennetta, vaan sen arkkitehtuuri määräytyy ratkaistavan ongelman ja käyttötarkoituksen mukaan [6]. Tästä huolimatta konvoluutioneuroverkot noudattavat tyypillisesti samaa perusrakennetta, jossa verkko voidaan jakaa kahteen päävaiheeseen [2]: **piirteiden tunnistukseen** ja **luokitteluun**. Kuvassa 2.1 on havainnollistettu tyypillinen konvoluutioneuroverkon rakenne. [2]



Kuva 2.1: Konvoluutioneuroverkon tyypillinen arkkitehtuuri [2]

Konvoluutiokerros on keskeinen osa konvoluutioneuroverkossa. Sen tehtävänä on tunnistaa kuvasta paikallisia piirteitä, kuten reunoja, kulmia ja kuvioita. Kasvojentunnistuksessa nämä piirteet voivat tarkoittaa esimerkiksi silmien ääriviivoja, nenän muotoa tai suun sijaintia. Piirteiden tunnistus tapahtuu suodattimien eli konvoluutioikkunoiden avulla, jotka skannaavat syötettä liukumalla tämän yli, samalla poimien siitä merkityksellisiä yksityiskohtia. Konvoluutioikkunoiden toiminta muistuttaa paikallisten binäärikuviohistogrammien toimintaa sillä, että kuvassa olevien pikselien arvoja yhdistetään matemaattisten operaatioiden avulla.[2][6][9]

Konvoluutiokerroksen tuottamat arvot syötetään tämän jälkeen **aktivaatiokerrokselle**, joka lisää malliin epälineaarisuutta. Epälineaarisuus on välttämätöntä monimutkaisten tehtävien, kuten kasvojentunnistuksen kannalta, sillä kasvojentunnistus ei ole myöskään lineaarinen ongelma. Yleisesti käytetty ReLU-aktivaatiofunktio (Rectified Linear Unit) muuttaa negatiiviset arvot nolliksi ja välittää positiiviset arvot sellaisenaan eteenpäin. Arvojen epälineaaristaminen auttaa verkkoa oppimaan tehokkaammin, nopeuttaa laskentaa, mutta säilyttää tärkeät kuvadatan piirteet jatkokäsittelyä varten. [2][6]

Pooling-kerros puolestaan pienentää datan kokoa yhdistämällä sen merkityksellisimpiä osia. Yleisimpiä pooling-menetelmiä ovat maksimi-pooling, jossa valitaan suurin arvo tietyltä alueelta, ja keskiarvo-pooling, jossa lasketaan alueen keskiarvo. Pooling-kerros vähentää laskennallista kuormaa, tekee verkosta sietokykyisemmän pienille muutoksille syötteessä ja nopeuttaa merkittävästi verkon tekemää datankäsittelyä. Usein konvoluutio- ja pooling-kerroksia toistetaan useita kertoja, jotta verkko pystyy tunnistamaan yhä monimutkaisempia piirteitä. [2][6][9]

Luokitteluvaiheessa nämä piirteet siirtyvät **täysin kytketylle kerrokselle**, jossa ne muunnetaan lopulliseksi ennusteeksi syötteenä annetun datan sisällöstä. Täysin kytketty kerros rakentuu useista kerroksista neuroneja, jotka ovat kaikki yhdistetty toisiinsa. Täysin kytketty kerros mahdollistaa kasvojentunnistusmallin op-

pimisen syötteestä, ja sen toimintaa käsitellään tarkemmin aliluvussa 2.3. Kerros muuttaa sille syötetyt piirteet vektorimuotoon, joihin käytetään lopuksi Softmax-aktivointifunktiota, joka tuottaa todennäköisyysjakauman eri luokkien välillä. Näin verkko pystyy tunnistamaan, mihin luokkaan syöte todennäköisimmin kuuluu. [2]

2.3 Syväoppivan neuroverkon kouluttaminen

Jotta syväoppiva neuroverkko pystyisi suoriutumaan kasvojentunnistuksesta tai vastaavasta tehtävästä, se täytyy ensin kouluttaa esimerkkidatan avulla. Koulutuksessa neuroverkko oppii säätämään jokaisen neuronin sisäistä painokerrointa ja vakiotermiä (engl. bias). Neuroni laskee syötteelle arvon käyttäen kaavaa 2.1, missä Y on neuronin ulostulo, x_i on sisääntulon arvo, w_i on painokerroin ja b on vakiotermi [3].

$$Y = \sum_i (x_i \cdot w_i) + b \tag{2.1}$$

Koulutus etenee epookkeina, eli jaksoina, joiden aikana koko koulutusaineisto käydään läpi useita kertoja. Jokaisessa epookissa verkko tekee ennusteita syötteiden perusteella, ja näitä verrataan todellisiin vastauksiin tappiofunktion avulla. Tappiofunktio mittaa, kuinka kaukana verkon tuottamat ennusteet ovat oikeista vastauksista. Mitä pienempi tappiofunktion arvo, sitä parempi ennuste. [2]

Tappiofunktion arvoa pyritään minimoimaan takaisinkytkennän (engl. backpropagation) ja optimointialgoritmien, kuten stokastisen gradienttilaskeuman (engl. stochastic gradient descent, SGD), avulla [3]. Takaisinkytkentä on menetelmä, jolla verkko laskee virheen vaikutuksen taaksepäin jokaiselle kerrokselle ja säätää painoja sen mukaisesti [10]. SGD puolestaan on algoritmi, joka päivittää painoja asteittain pienten tietojoukkojen avulla, pyrkien kohti mahdollisimman pientä tappiofunktion arvoa [11]. Näiden menetelmien ansiosta verkko oppii vähitellen paremmin suoriutumaan annetusta tehtävästä. Koulutus jatkuu, kunnes tappiofunktion arvo vakiintuu

eli konvergoituu, mikä viittaa siihen, että verkko on oppinut tehtävänsä riittävällä tarkkuudella [3].

Keskeinen haaste neuroverkon kouluttamisessa on varmistaa, että malli oppii yleistettäviä piirteitä sen sijaan, että se omaksuu koulutusdatan yksityiskohdat liian tarkasti. Tilannetta, jossa malli oppii tunnistamaan data-aineiston yksittäiset kuvat, mutta ei kuitenkaan pysty yleistämään uuteen dataan, kutsutaan ylisovittamiseksi. Ylisovittaminen on yleistä erityisesti silloin, kun malli on hyvin monimutkainen suhteessa koulutusaineiston määrään tai sen kouluttamiseen käytetään liikaa epookkeja. Vastakohtaisesti alisovittaminen kuvaa tilannetta, jossa neuroverkko ei opi edes koulutusdatan rakenteita. Tämä voi johtua esimerkiksi siitä, että neuroverkon arkkitehtuuri on liian yksinkertainen tai koulutusepookkeja on liian vähän. [12]

Tässä luvussa on käsitelty sitä, kuinka neuroverkot koulutetaan ja miten kyseiset neuroverkkomallit oppivat tunnistamaan datasta piirteitä muuttamalla sisäisiä painojaan. Näin neuroverkot pystyvät tekemään yhä tarkempia ennusteita. Neuroverkkojen oppiminen perustuu täysin sille syötettyyn dataan. Malli ei ainoastaan opi tunnistamaan koulutusdatan yleisiä rakenteita, vaan myös sen sisäiset vinoumat voivat siirtyä koulutettavaan malliin. Seuraavassa luvussa käsitellään tarkemmin, miten data-aineistoissa olevat vinoumat vaikuttavat neuroverkon toimintaan ja millaisia seurauksia kasvojentunnistusjärjestelmän vinoumilla saattaa olla.

3 Kasvojentunnistusmallien virheet ja vinoumat

Vaikka kasvojentunnistusteknologiat ovat kehittyneet viime vuosina merkittävästi ja pystyvät jo monissa tapauksissa suoriutumaan tunnistustehtävistä erittäin tarkasti, ei teknologia ole vielä ongelmatonta tai kaikilta osin täysin luotettavaa. Erityisen hyvin kasvojentunnistusjärjestelmät toimivat hallitussa ympäristössä, jossa tunnistuksen olosuhteet ovat otolliset ja henkilön kasvot vastaavat mallin koulutusdataa [13]. Reaalimaailman tilanteet kuitenkin harvoin vastaavat laboratorio-olosuhteita. Kasvojentunnistusjärjestelmät voivat yhä tehdä virheitä, jotka saattavat aiheuttaa yksilöille merkittäviäkin seurauksia.

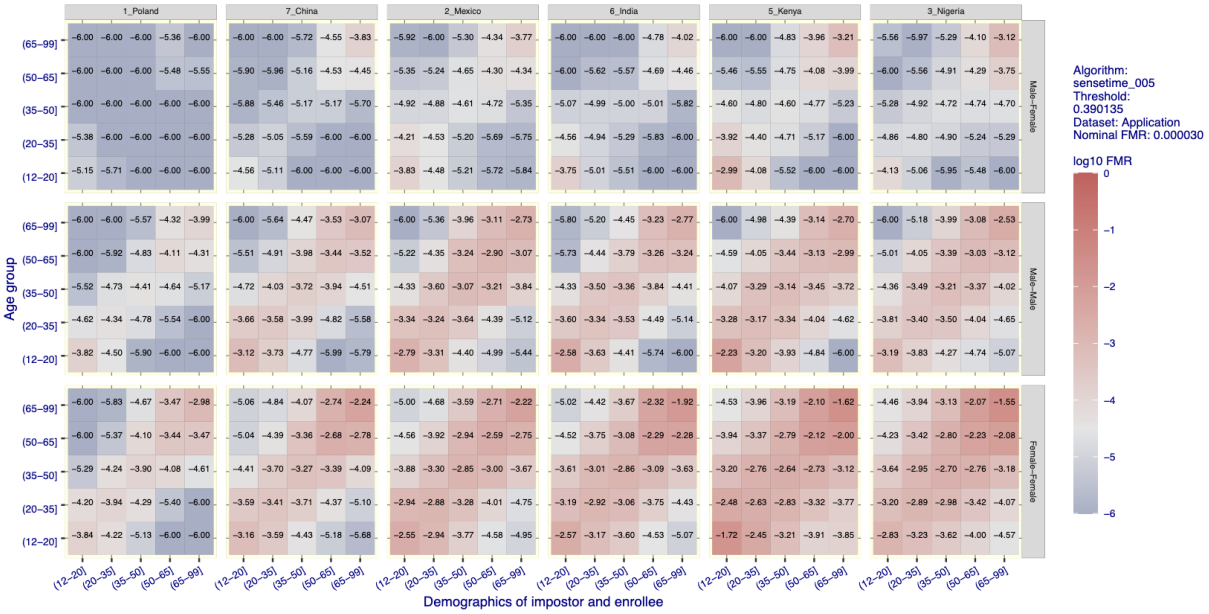
Kasvojentunnistusjärjestelmien tekemiä virheitä voidaan jakaa karkeasti kahteen eri kategoriaan: virheellisiin positiivisiin ja virheellisiin negatiivisiin. Virheellisillä positiivisilla tarkoitetaan tilanteita, joissa kasvojentunnistusjärjestelmä yhdistää kasvot tietokannassa olevaan identiteettiin, mutta kyseinen identiteetti on väärä. Virheellisillä negatiivisilla puolestaan viitataan tilanteisiin, joissa kasvoja ei osata yhdistää identiteettiin, vaikka kyseinen henkilö löytyy järjestelmän tietokannasta. [14]

Virheiden seuraukset vaihtelevat suuresti eri käyttökontekstien välillä, mikä asettaa erilaisia vaatimuksia kasvojentunnistusjärjestelmien kehitykselle. Arkipäiväinen tunnistusvirhe näkyy esimerkiksi silloin, kun älypuhelimen kasvojentunnistus ei avaa

lukitusnäyttöä ensimmäisellä yrittämällä – tilanne on lähinnä harmillinen, mutta harvoin vakava. Tällaisissa sovelluksissa järjestelmä voidaan kehittää ensisijaisesti minimoimaan virheelliset positiiviset tunnistukset, vaikka se merkitsisikin hieman enemmän virheellisiä negatiivisia [14]. Sen sijaan lentokentän passintarkastuksessa tunnistusvirhe voi aiheuttaa matkustajalle stressiä tai viivästyksiä, ja järjestelmältä edellytetään tasapainoa sujuvuuden ja turvallisuuden välillä. Vakavimmillaan virheet poliisin kasvojentunnistusjärjestelmässä voivat johtaa syyttömän henkilön epäilyyn tai pidätykseen, jolloin väärien positiivisten tunnistusten minimointi on ensisijaisen tärkeää. Näissä tapauksissa järjestelmän kehittämisessä on erityisen tärkeää huomioida eettiset ja oikeudelliset näkökulmat sekä suunnitella tunnistusalgoritmit niin, että ne kestävät kriittistä tarkastelua ja mahdollistavat läpinäkyvän virheanalyysin.

Kasvojentunnistusteknologian käyttö viranomaistoiminnassa on herättänyt eettisiä huolenaiheita erityisesti järjestelmien ennakoimattomien ja rakenteellisten vinoumien vuoksi. Mengjun Tao, Richard Jiang ja Carolyn Downs korostavat tutkimuksessaan [15], että Iso-Britanniassa ainakin 14 poliisilaitosta on ottanut käyttöön rikosten ennustamiseen tarkoitettuja ohjelmistoja, joiden on havaittu altistavan etnisiä vähemmistöjä ja matalatuloisia asuinalueita suhteettomalle valvonnalle.

Eräs havainnollistava esimerkki kasvojentunnistusteknologian ongelmallisesta käytöstä on Robert Williamsin tapaus Detroitissa, Yhdysvalloissa 2020. Williams pidätettiin virheellisesti, kun poliisin käytössä ollut kasvojentunnistusjärjestelmä yhdisti hänet virheellisesti epäillyn kasvoihin. Järjestelmän tekemä virhe perustui sen puutteelliseen koulutusaineistoon, joka koostui enimmäkseen valkoihoisten henkilöiden kuvista. Data-aineiston heikko edustavuus heikensi kasvojentunnistusalgoritmin kykyä erottaa tummaihoisia henkilöitä toisistaan. Williamsin tapaus ei kuitenkaan ole ainutlaatuinen, sillä virheellinen kasvojentunnistus johti Yhdysvalloissa vuosina 2019-2020 ainakin kahteen muuhun perusteettomaan pidätykseen. [16]



Taulukko 3.1: NISTin Face Recognition Vendor Test -tutkimuksesta (2022) poimittu taulukko esittää virheellisten osumien todennäköisyyksiä (engl. False Match Rate, FMR) eri syntymämaiden ja ikäryhmien välillä kiinteällä kynnysarvolla. Arviot perustuvat eri henkilöiden kasvokuvien vertailuun. Ylimmällä rivillä vertailu on tehty sukupuolten välillä, keskellä miesten ja alimmalla naisten välillä. Solujen tekstit ja värit kuvaavat logaritmista virheosuuden arvoa ($\log_{10}$ FMR), jossa kahden yksikön ero tarkoittaa 100-kertaista eroa virheosumien määrässä. [18]

Vaikka Yhdysvaltain standardointi- ja teknologiainstituutti (NIST) on tutkimuksissaan vuonna 2018 todennut virhemarginaalien olevan parhailla kasvojentunnistusalgoritmeilla vain 0,2 %, voivat jo pienet epätarkkuudet suuren mittakaavan käyttökontekstissa johtaa huolestuttavaan määrään virheellisiä tunnistuksia [17]. Raposon (2023) mukaan vain 0,01 %:n virheprosentti Heathrow'n lentoasemalla tarkoittaisi yli 20 väärin tunnistettua henkilöä päivässä [17]. Taulukossa 3.1 on havainnollistettu NISTin vastaavaa tutkimusta vuodelta 2022, joka osoittaa, ettei kasvojentunnistus ole tasapuolista kaikissa väestöryhmissä [18].

NIST toteaa kuitenkin vuoden 2019 tutkimuksessaan, että lähes kaikissa heidän analysoimistaan 189:stä eri kasvojentunnistusalgoritmista sisälsi jonkinlaisen vinouman mustia tai aasialaisia kasvoja kohtaan. Tutkimuksesta kävi myös ilmi, et-

tä virheellisiä positiivisia esiintyi useammin naisilla kuin miehillä algoritmista riippumatta. Tämä johtuu usein epätasapainoisista koulutusdatan kokoelmista, joissa valkoihoiset miehet ovat yliedustettuina. Kasvojentunnistusjärjestelmien vinoumien korjaaminen on myös hankalaa algoritmien ”mustaa laatikkoa” muistuttavan luonteen vuoksi. Järjestelmän toiminnan läpinäkymättömyys, jopa sen kehittäjille, vaikeuttaa vinoumien tunnistamista ja tekee niiden korjaamisesta haastavaa. [17]

Edelliset esimerkit osoittavat, että kasvojentunnistusjärjestelmien virheet ja vinoumat eivät synny pelkästään kasvojentunnistusjärjestelmien algoritmisista ongelmista, vaan niiden taustalla vaikuttavat usein syvemmät ongelmat koulutusdatassa. Mallien suorituskyky ja oikeudenmukaisuus ovat tiiviisti sidoksissa siihen, millaisesta datasta ne oppivat: miten data on kerätty, ketkä siinä ovat edustettuina ja millä tavoin eri ryhmät esitetään. Seuraavassa luvussa syvennytään koulutusdatan merkitykseen kasvojentunnistusmallien kehityksessä. Tarkastelun kohteena ovat erityisesti datassa esiintyvät vinoumat, niiden vaikutukset mallien toimintaan ja se, miten nämä ongelmat voivat siirtyä huomaamatta osaksi teknologian tekemää päätöksentekoa.

4 Koulutusdata kasvojentunnistuksessa

4.1 Data-annotoinnin haasteet

Kasvojentunnistusmallien suorituskyky on vahvasti sidoksissa niiden koulutuksessa käytettyyn dataan. Kasvojentunnistuksessa hyödynnettävä koulutusdata koostuu tyypillisesti kaksi- tai kolmiulotteisista kasvokuvista. Kun data-aineistoja koostetaan, yksittäiseen kuvaan eli esimerkkiin liitetään semanttista lisätietoa prosessissa, jota kutsutaan data-annotaatioksi. Sen tarkoituksena on tehdä datasta ymmärrettävämpää koneoppimisalgoritmeille ja siten parantaa niiden kykyä oppia ja tehdä tarkempia ennusteita [19]. Kasvojentunnistuksessa koulutusdatana toimiviin kasvokuviin voidaan liittää tietoa henkilön sukupuolesta, etnisestä taustasta, pään asennosta, nimestä, iästä tai muun muassa siitä, missä kohtaa kuvaa kasvot sijaitsevat, sekä onko henkilöllä silmälasit, parta tai kasvomaski [20].

Yleisimpiä datasta johtuvia haasteita kasvojentunnistuksessa ovat pään asennon ja valaistuksen vaihtelut [9][21]. Kortylewski ym. [22] osoittavat, että koneoppimismallin on nähtävä eri kasvonasentoja kattavasti oppiakseen tunnistamaan kasvot niiden asennon vaihtelusta huolimatta. Mikäli data-aineisto ei ole kasvonasentojen osalta riittävän monipuolinen, malli voi virheellisesti tulkita saman henkilön eri kulmista otetuista kuvista eri henkilöksi.

Epäselvät ja epäjohdonmukaiset esimerkit voivat vaikuttaa mallien oppimiseen sekä heikentää niiden tuottamien ennusteiden luotettavuutta [23]. Kasvojentunnistuksessa tämä tarkoittaa esimerkiksi kohinaa sisältäviä kuvia, joissa esiintyy sumentumista, ylivalottuneisuutta, värimuutoksia tai joiden resoluutio on matala [24]. Tiettyjen vääristymien vaikutus koneoppimismallin kykyyn tunnistaa kasvot on suurempi kuin toisten. Henkilöstä riippuen koneoppimismallin kyky tunnistaa kasvot voi vaihdella sen mukaan, mihin kasvojen alueeseen kuvassa oleva vääristymä kohdistuu. Koska kasvojen ilmeikkyyden ja tunnistettavuuden kannalta erityisen merkittäviä alueita ovat silmät, nenä ja suu [24], näihin kohdistuvat vääristymät voivat vaikuttaa mallin suorituskykyyn erityisen voimakkaasti. Lisäksi, jos koulutusdataan sisältyy huomattava määrä virheellisiä esimerkkejä, malli voi oppia vääriä piirteitä, mikä heikentää sen kykyä yleistää todellisiin tunnistustehtäviin.

Data-annotointi ei kuitenkaan ole vain tekninen prosessi, vaan siihen sisältyy myös merkittäviä inhimillisiä ja sosiaalisia ulottuvuuksia [25]. Annotointityötä tekevät yleisesti yritysten sisäiset annotaattorit, alihankkijat tai joukkouttamalla (engl. crowdsourcing) hankitut resurssit. Vaikka automaattisia ja puoliautomaattisia työkaluja on kehitetty avustamaan prosessia, ihmistyön rooli on edelleen keskeinen. Annotointityö on vahvasti hierarkkista, ja se tehdään tiukasti asiakasyritysten tai yritysjohdon antamien ohjeiden mukaisesti vastaamaan heidän tavoitteita ja tarpeita [26]. Tunnistettavan henkilön kasvojen ikä, sukupuoli ja etninen tausta ovat yleisiä annotoinnin luokkia, mitkä ovat kuitenkin harvoin universaalisti määriteltävissä. Nämä luokat heijastavat usein tilaajatahojen käsityksiä ja tavoitteita. Annotointityössä korostuu symbolinen valta, jossa tilaaja määrittää datalle ”oikean” merkityksen, ja tämä merkitys siirtyy koulutusdatan kautta osaksi koneoppimismallien päätöksentekoa [26]. Luokitusten taustalla olevia oletuksia ei useinkaan dokumentoida avoimesti, mikä voi johtaa läpinäkymättömiin ennusteisiin ja mallin ennalta-arvaamattomaan käyttäytymiseen.

4.2 Aineiston keruu, yksityisyys ja lainsäädäntö

Kasvojentunnistusmallin kouluttamiseen käytetty data vaikuttaa siihen, kuinka hyvin malli yleistää oppimansa myös koulutusaineiston ulkopuoliseen dataan [27]. Jotta malli pystyy tunnistamaan kasvot luotettavasti ja oikeudenmukaisesti kaikissa tilanteissa, on tärkeää, että koulutusaineisto kattaa monipuolisesti kaikki sen toiminnalle olennaiset muuttujat. On kuitenkin haastavaa määrittää, missä määrin tiettyä muuttujaa tulisi data-aineistossa korostaa, jotta aineisto tukisi mallin yleistyvyyttä parhaalla mahdollisella tavalla [1].

Laajat ja monipuoliset data-aineistot, joissa erilaiset demografiset, fyysiset ja kontekstuaaliset tekijät ovat kattavasti edustettuina, ovat kuitenkin harvinaisia, sillä niiden kokoaminen on sekä teknisesti että eettisesti haastavaa [28]. Valtaosa saatavilla olevista aineistoista on koottu haravoimalla julkisia verkkosivuja (engl. web scraping) ja yhdistämällä aiempia data-aineistoja [27]. Koska dataa kerätään valtava määrä eri lähteistä, on miltei mahdotonta pitää kirjaa kuvassa esiintyvän henkilön suostumuksesta data-aineistossa esiintymiseen [27]. Lähteet, joista tiedonharavointirobotit hakevat tietoa, eivät myöskään yleisesti sisällä mekanismeja sen seuraamiseen, kuka lataa dataa verkkosivuilta ja mihin tarkoitukseen sitä käytetään [27].

Se kuinka biometristä dataa, kuten kasvokuvia, saa kerätä, vaihtelee suuresti eri maiden ja niiden tietosuojalakien välillä. **Euroopan unionin** yleisen tietosuojaasetuksen (engl. General Data Protection Regulation, GDPR) mukaan biometrinen data, kuten kasvonpiirteet, on erityisen suojattua henkilötietoa, jonka käsittely edellyttää selkeää ja vapaaehtoista suostumusta aineistossa esiintyvältä henkilöltä [29]. Yksityisyyden suoja ja anonymisointi ovatkin olleet ongelmallisia avoimissa kasvojentunnistusaineistoissa. Useaa tunnettua aineistoa, kuten MegaFacea, MS-Celeb-1M:ää ja VGGFace2:ta, on jouduttu rajoittamaan oikeudellisten ja eettisten ongelmien vuoksi [30].

Yhdysvalloissa sääntely on hajanaisempaa ja vaihtelee merkittävästi eri osavaltioiden välillä. Yhdysvalloissa ei ole liittovaltiotasolla yhtä tiukkaa sääntelyä kuin Euroopassa, mutta useat osavaltiot ovat säätäneet omia lakejaan biometrisen datan käsittelyyn. Yksi keskeisimmistä osavaltiotason säädöksistä on Illinois'n vuonna 2008 voimaan tullut Biometric Information Privacy Act (BIPA), jota pidetään edelleen yhtenä tiukimmista biometrisen datan sääntelylaeista Yhdysvalloissa. BIPAn vaikutuksesta myös Texasin, Washingtonin ja Kalifornian osavaltiot ovat säätäneet omia lakejaan, jotka koskevat pääasiassa yrityksiä ja yksityisiä toimijoita. Vaikka liittovaltiotasolla ei toistaiseksi ole kattavaa sääntelykehystä, on asiasta käyty keskustelua esimerkiksi Facial Recognition Technology Warrant Act -lakiesityksen muodossa, joka edellyttäisi viranomaisia hankkimaan tuomioistuimen luvan ennen kasvojentunnistusteknologian käyttöä valvontatarkoituksiin. [31]

Kiinassa kasvojentunnistusteknologian sääntely on kehittynyt vähitellen käytännön kokemusten ohjaamana. Vuosien ajan kasvojentunnistusta ovat soveltaneet laajasti niin viranomaiset, yritykset kuin julkiset laitokset, usein ilman selkeitä oikeudellisia rajoja. Vaikka yksittäisiä lakeja henkilötietojen suojaamiseksi on säädetty jo 2000-luvun alusta lähtien, vasta vuonna 2021 voimaan tullut Personal Information Protection Law (PIPL) muodosti kattavamman sääntelykehyksen. PIPL korostaa erityisesti biometrisen tiedon, kuten kasvokuvien, arkaluontoisuutta ja edellyttää erillistä suostumusta muuhun kuin julkisen turvallisuuden ylläpitämiseen. Toisin kuin Yhdysvalloissa, Kiinan lähestymistapaa leimaa kuitenkin vahva valtionohjaus ja utilitaristinen näkökulma teknologian hyödyntämiseen, jossa yksilön tietosuoja voi väistyä yhteiskunnallisen turvallisuuden ja tehokkuuden edistämisen tieltä. [31]

4.3 Synteettinen data

Datan keräämiseen ja käyttöön liittyviin ongelmiin on ehdotettu synteettisen datan käyttöä. Synteettisen datan perusidea on lupaava: sen sijaan että kasvojentun-

nistusmalleja koulutettaisiin aidoilla ihmiskasvoilla, data voitaisiin luoda erilaisten tietokoneohjelmien avulla. Tällöin keinotekoinen aineisto muistuttaisi ihmiskasvoja, mutta siinä ei esiintyisi todellisia henkilöitä. Synteettisen datan käyttö tarjoaa mahdollisuuksia parantaa data-aineistojen monimuotoisuutta ja edustavuutta. Kuitenkin uusien data-aineistojen luomisessa tai olemassa olevien aineistojen laajentamisessa piilee riski siitä, että monimuotoisuuden lisääminen jää vain näennäiseksi eikä aidosti korjaa vinoumia alkuperäisessä datassa. [28]

Viimeaikaiset edistysaskeleet generatiivisessa tekoälyssä ovat mahdollistaneet kasvojentunnistusmallien koulutusaineistojen laajentamisen ja anonymisoinnin luomalla realistisen näköisiä kasvoja. Synteettisen datan generointi tarjoaa useita etuja perinteisiin aineistoihin verrattuna ja esittää ratkaisuja aiemmin esiteltyihin dataan liittyviin ongelmiin. Generatiiviset mallit voivat auttaa yksityisyydensuojaan liittyvissä haasteissa ja mahdollistaa suurten sekä kohdennettujen aineistojen luomisen, joissa aineistoon luodaan variaatiota eri kasvojen asentoihin, valaistukseen ja tunnistettavien kasvojen ikään [32]. Generatiiviset mallit voivat myös auttaa poistamaan peitteitä, kuten aurinkolaseja, huiveja tai kasvomaskeja, sekä vähentämään kuvissa esiintyvää kohinaa tai muita teknisiä häiriöitä [32].

Synteettistä dataa pystytään tuottamaan generatiivisten vastakkaisverkkojen (engl. Generative Adversarial Network, GAN) ja diffuusiomallien avulla [28]. Nämä mallit perustuvat koneoppimisen menetelmiin, joissa ne oppivat luomaan uusia esimerkkejä koulutusaineistonsa perusteella. GAN-malleissa on kaksi neuroverkkoa: generaattori ja erottelija. Generaattori pyrkii tuottamaan mahdollisimman realistisia kuvia, kun taas erottelija arvioi, ovatko kuvat aitoja vai synteettisiä [33]. Tämän vastakkainasettelun myötä generaattori paranee vähitellen tuottamaan yhä uskottavampia kasvoja. Diffuusiomallit puolestaan rakentavat kuvia askel kerrallaan poistamalla kohinaa satunnaisesta syötteestä, mikä mahdollistaa hienojakoisemman kontrollin kuvan sisällöstä [34]. Molempien mallityyppien kyvykkyys luoda visuaalisesti us-

kottavaa ja muokattavissa olevaa kasvodataa tekee niistä erityisen käyttökelpoisia synteettisen datan tuotannossa kasvojentunnistuksen tarpeisiin.

Generatiivisten mallien tuottaman datan laatua voidaan puolestaan arvioida FIQA-menetelmillä (Face Image Quality Assessment), jotka mittaavat yksittäisten kasvokuvien laatua erityisesti kasvojentunnistuksen näkökulmasta. FIQA-menetelmät eivät keskity pelkästään kuvan tekniseen laatuun, kuten resoluutioon tai valaistukseen, vaan myös siihen, kuinka hyvin kasvokuva soveltuu tunnistamiseen tai vertailuun. Tämä on olennaista, sillä vaikka synteettinen kuva näyttäisi visuaalisesti aidolta, se ei välttämättä tarjoa mallille hyödyllistä informaatiota ilman riittävää tunnistettavuutta tai kasvonpiirteiden selkeyttä. FIQA-arviointien avulla voidaan siis karsia huonolaatuisia synteettisiä kuvia ja parantaa kokonaisaineiston laatua. Lisäksi FIQA-menetelmiä hyödynnetään käytännön kasvojentunnistusjärjestelmissä valitsemaan videomateriaalista ne ajanhetket, joissa kasvojen laatu on riittävän korkea luotettavaa tunnistusta varten. FIQA-menetelmät voivat siis tukea kasvojentunnistusjärjestelmiä niin kehitysvaiheessa kuin myös niiden suorituksen aikana. [35]

Synteettisen datan käyttäminen kasvojentunnistusmallien kouluttamisessa ei kuitenkaan ole ongelmatonta. Generatiivinen tekoäly on suhteellisen uusi ilmiö, jonka rajoitteet eivät ole vielä laajasti tunnettuja. Generatiivisia malleja voidaan käyttää luomaan näennäisesti monimuotoisia aineistoja. Vaikka data näyttäisi sisältävän riittävästi erilaisia kasvoja eri väestöryhmistä, se ei silti välttämättä perustu todelliseen edustavuuteen vaan heijastaa alkuperäisen generatiivisen mallin koulutukseen käytettyä vinoutunutta data-aineistoa [28]. Jos generatiivisen mallin kouluttamiseen käytetty data on vinoutunutta, heijastuvat samat vinoumat myös jossain määrin sen tuottamaan dataan.

Synteettisen datan käyttäminen lisäksi sumuttaa entisestään datan keruun läpinäkyvyyttä ja sen todellista alkuperää. Ellei synteettisen datan alkuperää ja käsit-

Kuva 4.1: Esimerkkikasvokuvia, jotka ovat vuotaneet generatiivisten mallien harjoitusaineistosta. Ensimmäisellä rivillä näkyvät aidot kuvat, toisella niitä vastaavat synteettiset kuvat. Kuva on peräisin julkaisusta [36].

telyvaiheita dokumentoida tarkasti, sen jäljittäminen alkuperäiseen dataan voi olla haastavaa. Synteettinen data luo lisäksi teknisiä ja hallinnollisia esteitä kasvojentunnistusmallien uudelleenkouluttamiselle, mikäli havaitaan mallin olevan vinoutunut. On mahdollista, että synteettiseen dataan on periytynyt piirteitä alkuperäisestä datasta, jolloin myös yksilön henkilötietojen poistaminen olemassa olevasta mallista tulee haastavammaksi; tuolloin sekä kasvojentunnistusmalli että synteettisen datan generoiva malli on koulutettava uudelleen. [28]

Yksityisyydensuojan näkökulmasta piirteiden periytyminen synteettiseen dataan voi olla ongelmallista. Vaikka synteettisten kasvokuvien tarkoituksena on anonymisoida alkuperäiset henkilöt, on havaittu, että generatiiviset mallit voivat tuottaa kuvia, jotka muistuttavat huolestuttavan paljon koulutusdatan yksilöitä. Generatiiviset mallit, kuten GAN- ja diffuusiopohjaiset järjestelmät, voivat oppia ja säilyttää liiallisia yksityiskohtia alkuperäisistä kasvoista. Tuloksena on synteettisiä kuvia, jotka ovat visuaalisesti lähes identtisiä tiettyjen koulutuskuvien kanssa. Synteettistä dataa ei siis voida suoraan pitää turvallisena ja yksityisenä vaihtoehtona perinteiselle datalle. Kuvassa 4.1 havainnollistetaan, kuinka synteettinen data voi vuotaa generatiivisen mallin kouluttamiseen käytettyä dataa. [36]

Vaikka synteettinen data tarjoaa merkittäviä mahdollisuuksia datan saatavuuden ja yksityisyydenhallinnan näkökulmista, sen ja autenttisen datan välillä on edelleen selkeä laatu- ja ominaisuusero. Nykyiset tutkimukset viittaavat siihen, että pelkästään synteettisellä datalla koulutetut kasvojentunnistusmallit eivät vielä saavuta samaa suorituskykyä kuin mallit, jotka on opetettu todellisilla kasvoilla. Optimaalisimmaksi lähestymistavaksi onkin osoittautunut autenttisen ja synteettisen datan yhdistäminen. Tällainen hybridikonfiguraatio kykenee parhaiten hyödyntämään kummankin lähteen vahvuudet: autenttisen datan tarkkuus ja synteettisen datan kontrolloitavuus tukevat toisiaan ja johtavat järjestelmiin, jotka ylittävät suorituskyvyssä ne perusjärjestelmät, jotka on koulutettu yksinomaan autenttisella aineistolla. [32]

5 Yhteenveto

Kasvojentunnistus on yksi konenäön tutkituimmista sovellusalueista. Vaikka nykyaikaiset mallit saavuttavat erittäin korkean tarkkuuden, teknologia kohtaa yhä merkittäviä haasteita. Erityisen vaikeaa on saada kasvojentunnistusjärjestelmä toimimaan tarkasti ja oikeudenmukaisesti tilanteissa, joissa kuvausolosuhteet muuttuvat, populaatio on demografisesti monimuotoinen tai joissa virheellinen tunnistus voi aiheuttaa merkittävää haittaa yksittäiselle henkilölle. Tämän tutkielman tavoitteena oli havainnollistaa kasvojentunnistusmallien tekemiä virheitä sekä luoda kattava ymmärrys siitä, miten koulutusaineiston laatu ja monimuotoisuus vaikuttavat mallien yleistämiskykyyn. Seuraavaksi tarkastellaan, miten saadut tulokset vastaavat esitettyihin tutkimuskysymyksiin ja mitä johtopäätöksiä niiden pohjalta voidaan tehdä.

TK1: Millainen on kasvojentunnistusmallin rakenne ja kuinka malli oppii sille annetusta koulutusdatasta? Luvussa 2 tarkasteltiin kasvojentunnistusmallien arkkitehtuuria sekä sitä, miten mallit oppivat hyödyntämään niille annettua koulutusdataa. Kasvojentunnistusmallin rakenne riippuu pitkälti sen käyttötarkoituksesta, mutta useimmat nykyaikaiset mallit pohjautuvat syviin konvoluutioneuroverkkoihin (engl. CNN, Convolutional Neural Network). Tällainen verkko voidaan yleisesti jakaa kahteen pääosaan: piirteiden tunnistamisesta vastaaviin tasoihin ja lopullista päätöksentekoa, kuten luokittelua, suorittaviin tasoihin.

Piirteiden tunnistus tapahtuu konvoluutiokerrosten avulla, joissa suodattimet etsivät kasvojen kuvista merkityksellisiä rakenteita, kuten silmien, nenän ja suun

muotoja. Näiden kerrosten yhteydessä käytetään usein aktivaatiofunktioita, kuten ReLU (Rectified Linear Unit), jotka tuovat verkkoon epälineaarisuutta ja mahdollistavat monimutkaisempien piirteiden oppimisen. Pooling-kerrokset puolestaan tiivistävät dataa säilyttäen tärkeimmät informaation osat, mikä pienentää laskennallista kuormaa ja parantaa mallin yleistämiskykyä.

Verkon loppupäässä sijaitsee täysin kytketty kerros (engl. fully connected layer), joka yhdistää oppimansa piirteet kokonaiskuvaksi ja tekee lopullisen päätöksen, esimerkiksi tunnistaen henkilön. Luokittelussa käytetään usein softmax-aktivaatiofunktiota, joka muuntaa verkon tuottaman numeerisen arvon todennäköisyysjakaumaksi eri luokkien välillä.

Mallin oppiminen tapahtuu koulutusdatan avulla, jossa syötetystä kasvokuvasta lasketaan ennuste ja sen tuottama virhe verrattuna todelliseen luokkaan. Virheen perusteella mallin painokertoimia päivitetään takaisinkytkennän (engl. backpropagation) ja optimointialgoritmien, kuten stokastisen gradienttilaskeuman (engl. SGD, stochastic gradient descent), avulla. Toistuvan harjoittelun myötä malli oppii tunnistamaan yhä tarkemmin kuvissa esiintyvät kasvonpiirteet ja yhdistämään ne oikeisiin henkilöihin tai luokkiin.

TK2: Millaisia haasteita liittyy syväoppimispohjaiseen kasvojentunnistukseen? Syväoppimispohjaisen kasvojentunnistuksen suurimmat haasteet liittyvät niiden kouluttamiseen käytetyn datan laatuun, monipuolisuuteen ja eettiseen hankintaan. Koska syväoppivat kasvojentunnistusmallit perustavat tekemänsä ennustukset koulutusdataan, vaikuttavat datassa olevat vinoumat, annotaatiovirheet, kuvien tekninen laatu sekä valaistuksen ja kasvojen asennon vaihtelu suoraan siihen, pystyykö syväoppivamalli yleistämään oppimaansa. Tasapainotetun koulutusaineiston laatiminen on kuitenkin hyvin haastavaa, koska ei ole selkeää tapaa todeta, missä määrin eri muuttujia aineistossa tulisi esiintyä.

TK3: Millä tavoin kasvojentunnistusjärjestelmissä esiintyy virheitä ja vinoumia, sekä mistä nämä ongelmat johtuvat? Kasvojentunnistusjärjestelmissä voi esiintyä virheitä tyypillisesti kahdella eri tavalla: virheellisinä positiivisina, joissa järjestelmä yhdistää tunnistettavat kasvot väärään identiteettiin, tai virheellisinä negatiivisina, jolloin kasvoja ei osata yhdistää tietokannassa olevaan identiteettiin. Vinoumia kasvojentunnistusjärjestelmissä voi esiintyä tietyn etnisen tai demografisen ryhmän vaikeampana tunnistuksena. Tyypillisesti data-aineistoissa on yliedustettuina valkoihoisten miesten kuvat, jolloin järjestelmä tunnistaa suhteessa heikommin muihin etnisiin ryhmiin kuuluvia ihmisiä. Vinoumat voivat johtua syväoppivan neuroverkon arkkitehtonisesta rakenteesta, mutta useammin syy lienee datan laadussa ja puutteissa.

Kasvojentunnistusmallien toiminta perustuu vahvasti niiden rakenteeseen ja koulutusaineistoon. Vaikka teknologian kehitys on mahdollistanut tarkkoja ja tehokkaita ratkaisuja, siihen liittyy edelleen merkittäviä haasteita, erityisesti mallien yleistämiskykyyn ja eettisesti kestävään toteutukseen liittyen. Tutkielmassa esitetyt tutkimuskysymykset auttoivat jäsentämään aihealueen keskeisiä ongelmia ja tarkastelemaan niitä sekä teknisestä että yhteiskunnallisesta näkökulmasta. Seuraavassa luvussa esitetään näihin havaintoihin perustuvat johtopäätökset.

6 Johtopäätökset

Kasvojentunnistusjärjestelmien toimintaan liittyy useita eri virhe- ja vinoumatyyppejä, jotka voivat heikentää järjestelmien luotettavuutta. Yleisimmin esiintyvät virheet jaetaan kahteen luokkaan: virheellisiin positiivisiin, joissa järjestelmä yhdistää tunnistettavat kasvot väärään identiteettiin, tai virheellisiin negatiivisiin, joissa kasvoja ei osata yhdistää tietokannassa olevaan identiteettiin.

Virheiden ohella kasvojentunnistusjärjestelmät kärsivät systemaattisista vinoumista. Näitä esiintyy erityisesti etnisten ryhmien, sukupuolen ja iän välillä. Tutkielmassa käsitellyt tutkimukset osoittavat, että järjestelmät tekevät merkittävästi enemmän virheitä tummaihoisten ja aasialaisten henkilöiden kohdalla. Virheitä esiintyy useammin myös naisilla miehiin verrattuna. Tällaiset vinoumat ovat seurausta epätasapainoisesta data-aineistosta, jossa tietyt ryhmät, kuten valkoihoiset miehet, ovat yliedustettuina.

Vinoumien korjaaminen ei ole yksinkertaista, sillä syväoppivien neuroverkkojen perinpohjainen toiminta on vaikeasti tulkittavissa niiden ”mustaa laatikkoa” vastaavan luonteen vuoksi. Järjestelmiä ei siis voida säätää manuaalisesti, eivätkä ne toimi täysin deterministisesti. Uusien, riittävän laajojen data-aineistojen kerääminen on haastavaa taloudellisista, eettisistä ja saavutettavuuteen liittyvistä syistä. Laadukkaan kasvojentunnistusaineiston kokoaminen on valtavasta datakoosta johtuen usein ainakin osittain automatisoitua.

Tällaisen automaation toteuttaminen vaatii yrityksiltä huomattavia taloudellisia resursseja. Datan eettinen kerääminen edellyttää jokaiselta kuvassa esiintyvältä henkilöltä suostumusta, lisäksi aineiston tulisi edustaa mahdollisimman kattavasti eri väestöryhmiä. Tämä tekee keräysprosessista erittäin vaativan, ellei jopa käytännössä mahdottoman, sillä kaikkien yksilöiden kontaktointi ja luvan pyytäminen on harvoin toteutettavissa. Tämän seurauksena monet olemassa olevat kasvoaineistot on koottu julkisista lähteistä, usein ilman kuvan kohteena olevan henkilön nimenomaista lupaa.

Koska kasvokuvia voidaan pitää yksilöä yksiselitteisesti tunnistavana biometrisenä datana, tulisi jokaisella olla oikeus päättää, missä ja miten hänen kasvojaan käytetään. Tämä tarkoittaa myös sitä, että henkilön tulisi halutessaan pystyä poistattamaan omat kasvokuvansa data-aineistosta. Kasvoihin liittyvä data tulisi lisäksi käsitellä arkaluonteisena. Toisin kuin monet muut henkilötiedot, kasvot ovat jatkuvasti näkyvillä ja niitä on vaikea muuttaa tai piilottaa. Vaikka kasvot tunnistavat meidät lähes yhtä luotettavasti kuin sormenjäljet tai iirikset, suhtautuminen kasvoihin kerättävänä datana on usein huomattavasti huolettomampaa. Kasvokuvia jaetaan arjessa vapaasti esimerkiksi sosiaalisessa mediassa, mikä voi hämärtää ymmärrystä siitä, kuinka syvästi tunnistettavaa ja pysyvää tietoa kasvoista voidaan koneoppimisen keinoin muodostaa.

Tulevaisuudessa tutkimuksen tulisi entistä enemmän kohdistua datan keruuseen erityisesti yksilön näkökulmasta. On tärkeää tutkia mahdollisuuksia, joilla yksilöt voisivat vaikuttaa heistä kerättäviin henkilötietoihin, sekä tapoja, joilla yritykset voisivat läpinäkyvämmin käsitellä keräämäänsä dataa. Samalla tarvitaan laajempaa yhteiskunnallista keskustelua ja tiedottamista siitä, mihin tarkoituksiin henkilötietoja kerätään ja käytetään, sekä millaisia riskejä tähän liittyy. Tällä hetkellä yleinen tietoisuus on usein rajallinen, vaikka datan keruu laajenee yhä suurempaan osaan arkisia järjestelmiä ja palveluita.

Lopulta on syytä kysyä, onko ylipäätään mahdollista koostaa täysin neutraalia ja vinoumatonta kasvojentunnistusaineistoa. Datan keräämiseen liittyy väistämättä valtasuhteita, joissa suuret teknologiayritykset ja tutkimuslaitokset määrittävät, kuka päätyy aineistoon ja millä ehdoilla. Näillä toimijoilla on omat taloudelliset tai teknologiset intressinsä, jotka eivät välttämättä ole linjassa yksilöiden oikeuksien tai yhteiskunnallisen oikeudenmukaisuuden kanssa. Tämä asettaa kriittisiä rajoja sekä teknologian kehittämiselle että sen eettisesti kestävälle käytölle.

Lähdeluettelo

[1] G. B. Huang, M. Mattar, T. Berg ja E. Learned-Miller, "Labeled faces in the wild: A database forStudying face recognition in unconstrained environments", *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, 2008. url: https://inria.hal.science/inria-00321923.

[2] M. L. Smith, L. N. Smith ja M. F. Hansen, "The quiet revolution in machine vision - a state-of-the-art survey paper, including historical review, perspectives, and future directions", *Computers in Industry*, vol. 130, s. 103 472, 1. syyskuuta 2021, ISSN: 0166-3615. DOI: 10.1016/j.compind.2021.103472.

[3] A. A. Khan, A. A. Laghari ja S. A. Awan, "Machine learning in computer vision: A review", *EAI Endorsed Transactions on Scalable Information Systems*, vol. 8, nro 32, e4–e4, 21. huhtikuuta 2021, Number: 32, ISSN: 2032-9407. DOI: 10.4108/eai.21-4-2021.169418.

[4] Z. Khan ja Y. Fu, "One Label, One Billion Faces: Usage and Consistency of Racial Categories in Computer Vision", teoksessa *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, sarja FAccT '21, New York, NY, USA: Association for Computing Machinery, 1. maaliskuuta 2021, s. 587–597, ISBN: 978-1-4503-8309-7. DOI: 10.1145/3442188.3445920.

[5] I. Adjabi, A. Ouahabi, A. Benzaoui ja A. Taleb-Ahmed, "Past, present, and future of face recognition: A review", *Electronics*, vol. 9, nro 8, s. 1188, elo-

kuu 2020, Number: 8 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2079-9292. DOI: 10.3390/electronics9081188.

[6] S. Prabha, R. Bulchandani, R. Mishra, S. Agarwal ja S. Chauhan, "Face recognition algorithms: A review", vol. 08, nro 7, 2021. url: https://www.irjet.net/archives/V8/i7/IRJET-V8I7323.pdf.

[7] N. Singhal, V. Ganganwar, M. Yadav, A. Chauhan, M. Jakhar ja K. Sharma, "Comparative study of machine learning and deep learning algorithm for face recognition", *Jordanian Journal of Computers and Information Technology*, nro 0, s. 1, 2021, ISSN: 2413-9351. DOI: 10.5455/jjcit.71-1624859356.

[8] Q. Kuang, "Face Image Feature Extraction based on Deep Learning Algorithm", *Journal of Physics: Conference Series*, vol. 1852, nro 3, 2021. DOI: 10.1088/1742-6596/1852/3/032040.

[9] P. K. Mannepalli, D. Singh Kushwaha, S. Kalamdhar, V. Nagrale ja V. Rajpoot, "Face Recognition Based on Cascade Classifier Using Deep Learning", teoksessa *2023 1st International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP)*, maaliskuu 2023, s. 63–68. DOI: 10.1109/IHCSP56702.2023.10127172.

[10] M. Li, "Comprehensive review of backpropagation neural networks", *Academic Journal of Science and Technology*, vol. 9, nro 1, s. 150–154, 20. tammikuuta 2024, Number: 1, ISSN: 2771-3032. DOI: 10.54097/51y16r47.

[11] Y. Tian, Y. Zhang ja H. Zhang, "Recent advances in stochastic gradient descent in deep learning", *Mathematics*, vol. 11, nro 3, s. 682, tammikuu 2023, Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2227-7390. DOI: 10.3390/math11030682.

[12] S. Pothuganti, "Review on over-fitting and under-fitting problems in Machine Learning and solutions", *International Journal of Advanced Research in*

Electrical Electronics and Instrumentation Engineering, vol. 7, s. 3692–3695, syyskuu 2018. DOI: 10.15662/IJAREEIE.2018.0709015.

[13] M. O. Oloyede, G. P. Hancke ja H. C. Myburgh, "A review on face recognition systems: Recent approaches and challenges", *Multimedia Tools and Applications*, vol. 79, nro 37, s. 27 891–27 922, lokakuu 2020, ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-020-09261-2.

[14] S. Meshkinfamfard, A. Gorban ja I. Tyukin, "Tackling rare false-positives in face recognition: A case study", teoksessa *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Exeter, United Kingdom: IEEE, kesäkuu 2018, s. 1592–1598, ISBN: 978-1-5386-6614-2. DOI: 10.1109/HPCC/SmartCity/DSS.2018.00260.

[15] M. Tao, R. Jiang ja C. Downs, "Ethics of face recognition in smart cities toward trustworthy AI", teoksessa *Big Data Privacy and Security in Smart Cities*, R. Jiang, A. Bouridane, C.-T. Li et al., toim., Series Title: Advanced Sciences and Technologies for Security Applications, Cham: Springer International Publishing, 2022, s. 23–52, ISBN: 978-3-031-04423-6 978-3-031-04424-3. DOI: 10.1007/978-3-031-04424-3_2.

[16] T. J. Benedict, "The Computer Got It Wrong: Facial Recognition Technology and Establishing Probable Cause to Arrest Notes", *Washington and Lee Law Review*, vol. 79, nro 2, s. 849–898, 2022. url: https://heinonline.org/HOL/P?h=hein.journals/waslee79&i=839.

[17] V. L. Raposo, "When facial recognition does not 'recognise': Erroneous identifications and resulting liabilities", *AI & SOCIETY*, vol. 39, nro 4, s. 1857–

1869, elokuu 2024, ISSN: 0951-5666, 1435-5655. DOI: 10.1007/s00146-023-01634-z.

[18] D. L. Duewer, "Face recognition vendor test (FRVT) part 8: Summarizing demographic differentials", National Institute of Standards ja Technology, Gaithersburg, MD, NIST IR 8429, 2022, NIST IR 8429. DOI: 10.6028/NIST.IR.8429.

[19] J. Gebele, P. Brune ja S. Faußer, "Face Value: On the Impact of Annotation (In-)Consistencies and Label Ambiguity in Facial Data on Emotion Recognition", teoksessa *2022 26th International Conference on Pattern Recognition (ICPR)*, ISSN: 2831-7475, elokuu 2022, s. 2597–2604. DOI: 10.1109/ICPR56361.2022.9956230.

[20] P. Terhörst, D. Fährmann, J. N. Kolf, N. Damer, F. Kirchbuchner ja A. Kuijper, "MAAD-Face: A Massively Annotated Attribute Dataset for Face Images", *IEEE Transactions on Information Forensics and Security*, vol. 16, s. 3942–3957, 2021, Conference Name: IEEE Transactions on Information Forensics and Security, ISSN: 1556-6021. DOI: 10.1109/TIFS.2021.3096120.

[21] K. Khan, R. U. Khan, K. Ahmad, F. Ali ja K.-S. Kwak, "Face Segmentation: A Journey From Classical to Deep Learning Paradigm, Approaches, Trends, and Directions", *IEEE Access*, vol. 8, s. 58 683–58 699, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2982970. url: https://ieeexplore.ieee.org/document/9045993/?arnumber=9045993.

[22] A. Kortylewski, B. Egger, A. Schneider, T. Gerig, A. Morel-Forster ja T. Vetter, "Empirically Analyzing the Effect of Dataset Biases on Deep Face Recognition Systems", teoksessa *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, s. 2174–217 409. DOI: 10.1109/CVPRW.2018.00283.

[23] F. Wang, L. Chen, C. Li et al., ”The Devil of Face Recognition is in the Noise”, 2018. arXiv: 1807.11649 [cs.CV]. url: https://arxiv.org/abs/1807.11649.

[24] P. Majumdar, S. Mittal, R. Singh ja M. Vatsa, ”Unravelling the effect of image distortions for biased prediction of pre-trained face recognition models”, teoksessa *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada: IEEE, lokakuu 2021, s. 3779–3788, ISBN: 978-1-6654-0191-3. DOI: 10.1109/ICCVW54120.2021.00422.

[25] N. M. Barbosa ja M. Chen, ”Rehumanized Crowdsourcing: A Labeling Framework Addressing Bias and Ethics in Machine Learning”, teoksessa *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, sarja CHI ’19, New York, NY, USA: Association for Computing Machinery, 2. toukokuuta 2019, s. 1–12, ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300773.

[26] M. Miceli, M. Schuessler ja T. Yang, ”Between Subjectivity and Imposition: Power Dynamics in Data Annotation for Computer Vision”, *Proc. ACM Hum.-Comput. Interact.*, vol. 4, 115:1–115:25, CSCW2 15. lokakuuta 2020. DOI: 10.1145/3415186.

[27] M. K. Scheuerman, K. Weathington, T. Mugunthan, E. Denton ja C. Fiesler, ”From Human to Data to Dataset: Mapping the Traceability of Human Subjects in Computer Vision Datasets”, *Proc. ACM Hum.-Comput. Interact.*, vol. 7, 55:1–55:33, CSCW1 16. huhtikuuta 2023. DOI: 10.1145/3579488.

[28] C. D. Whitney ja J. Norman, ”Real Risks of Fake Data: Synthetic Data, Diversity-Washing and Consent Circumvention”, teoksessa *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, sarja FAccT ’24, New York, NY, USA: Association for Computing Machinery,

5. kesäkuuta 2024, s. 1733–1744, ISBN: 979-8-4007-0450-5. DOI: 10.1145/3630106.3659002.

[29] Confederation of European Data Protection Organisation (CEDPO), "Generative AI: The Data Protection Implications", 26. lokakuuta 2023. url: https://cedpo.eu/generative-ai-the-data-protection-implications/.

[30] F. Boutros, M. Huber, P. Siebke, T. Rieber ja N. Damer, "SFace: Privacy-friendly and accurate face recognition using synthetic data", teoksessa *2022 IEEE International Joint Conference on Biometrics (IJCB)*, Abu Dhabi, Arabiemiirikunnat: IEEE, 10. lokakuuta 2022, s. 1–11, ISBN: 978-1-6654-6394-2. DOI: 10.1109/IJCB54206.2022.10007961.

[31] W. Chen ja M. Wang, "Regulating the use of facial recognition technology across borders: A comparative case analysis of the european union, the united states, and china", *Telecommunications Policy*, vol. 47, nro 2, s. 102 482, maaliskuu 2023, ISSN: 03085961. DOI: 10.1016/j.telpol.2022.102482.

[32] P. Melzi, R. Tolosana, R. Vera-Rodriguez et al., "FRCSyn-onGoing: Benchmarking and comprehensive evaluation of real and synthetic data to improve face recognition systems", *Information Fusion*, vol. 107, s. 102 322, 1. heinäkuuta 2024, ISSN: 1566-2535. DOI: 10.1016/j.inffus.2024.102322.

[33] P. Sharma, M. Kumar, H. K. Sharma ja S. M. Biju, "Generative adversarial networks (GANs): Introduction, taxonomy, variants, limitations, and applications", *Multimedia Tools and Applications*, vol. 83, nro 41, s. 88 811–88 858, 1. joulukuuta 2024, ISSN: 1573-7721. DOI: 10.1007/s11042-024-18767-y.

[34] D. Gallon, A. Jentzen ja P. v. Wurstemberger, *An overview of diffusion models for generative artificial intelligence*, 2. joulukuuta 2024. DOI: 10.48550/arXiv.2412.01371.

[35] T. Schlett, C. Rathgeb, O. Henniger, J. Galbally, J. Fierrez ja C. Busch, ”Face Image Quality Assessment: A Literature Survey”, *ACM Computing Surveys*, vol. 54, nro 10s, s. 1–49, tammikuu 2022, ISSN: 1557-7341. DOI: 10.1145/3507901.

[36] H. O. Shahreza ja S. Marcel, *Unveiling synthetic faces: How synthetic datasets can expose real identities*, 31. lokakuuta 2024. DOI: 10.48550/arXiv.2410.24015.