

Tartu Ülikool
Sotsiaalteaduste valdkond
Psühholoogia instituut

Nele-Liis Merilo

**FaceReader programmi võimekus inimemotsioonide tuvastamisel ja seosed
videomodellide ekstravertsusega**

Uurimistöö

Juhendajad: Kairi Kreegipuu, PhD

Liina Juuse, MA

Läbiv pealkiri: Ekstravertsus ja FaceReaderi emotsioonide tuvastamine

Tartu 2025

FaceReader programmi võimekus inimemotsioonide tuvastamisel ja seosed videomodellide ekstraversusega

Lühikokkuvõte

Uurimistöös keskendusin FaceReaderi programmi täpsuse hindamisele inimemotsioonide tuvastamisel videoklippidest ning seosele videomodellide ekstraversuse ja FaceReaderi tuvastamistäpsuse vahel. Varasemad uuringud on näidanud, et ekstravertse isiksusejoone ja väljendatud ekspressiivsete näoväljenduste vahel esinevad seosed. FaceReader analüüsis 99 modelli (38 meest) poolt esitatud kolme emotsiooni – rõõmu, vastikust ja hirmu – väljendavaid 294 videoklippi. Programm määras protsentuaalselt videoklippides tuvastatud baasemotsioonide esinemise määra, mille alusel lõin binaarse muutuja tabamuse ja möödalasu eristamiseks. Analüüsisin pakutud emotsioonide tuvastamistäpsust, koostas in sagedustabeli FaceReaderi hinnangute kohta kolme emotsiooni lõikes ning kasutasin logistilist regressioonanalüüsi, et uurida, kas modelli ekstraversuse skoor ennustab FaceReaderi emotsioonide tuvastamise täpsust. FaceReaderi üldine täpsus kõigi kolme emotsiooni tuvastamisel oli 37,4%. Programm oli märgatavalt edukam rõõmu tuvastamisel (96,9%) kui vastikuse (11,3%) ja hirmu (4%) tuvastamisel. Logistilise regressioonanalüüsi tulemusel ei osutunud modelli ekstraversuse skoor statistiliselt oluliseks täpsuse ennustamiseks. Uurimistöö pakub väärtuslikku ülevaadet uude emotsioonituvastustarkvara toimimisest, tuues esile selle kitsaskohad ja arenguvõimalused. Samuti viitab töö vajadusele uurida põhjalikumalt seoseid analüüsitavate modellide ekstraversuse ja teiste isiksuseomaduste ning emotsioonide tuvastamise täpsuse vahel.

Märksõnad: emotsioonituvastusprogramm, FaceReader, ekstraversus, emotsioonide äratundmine

The accuracy of FaceReader in detecting human emotion and its relationship with video models' extraversion

Abstract

This study focused on evaluating the accuracy of the FaceReader software in detecting human emotions from video clips and examining the relationship between video models' extraversion scores and the accuracy of FaceReader's emotion recognition. Previous studies have indicated a link between extraversion as a personality trait and expressed facial expressiveness. FaceReader analyzed 294 video clips featuring 99 models (38 men) expressing three emotions: joy, disgust, and fear. The software calculated the percentage of basic emotions detected in the facial expressions within each video clip, based on which I created a binary variable to distinguish between hits and misses. I analyzed the accuracy of the suggested emotions, created frequency tables to illustrate FaceReader's predictions for all three emotions, and conducted a logistic regression analysis to investigate whether the extraversion scores of the models predicted FaceReader's accuracy in detecting the correct emotion, considering the three expressed emotions (joy, disgust, fear). FaceReader's overall accuracy across the three emotions was 37.4%. The software was significantly more successful in detecting joy (96.9%) compared to disgust (11.3%) and fear (4%). The logistic regression analysis showed that the extraversion score of the models was not a statistically significant predictor of FaceReader's accuracy. This study provides valuable insights into the functioning of this relatively novel emotion recognition software, highlighting its limitations and opportunities for improvement. Additionally, it points to the need for further research into the relationships between extraversion and other personality traits of the analyzed models and the accuracy of emotion detection.

Keywords: emotion detection program, FaceReader, extraversion, emotion recognition

Tänapäeva digitaliseerivas ühiskonnas kasvab emotsioonide tuvastamisele keskenduva tarkvara tähtsus ja osakaal erinevates valdkondades jõudsalt, olgu selleks siis psühholoogia, inimese ja arvuti vaheline suhtlus või hoopis julgeolek. Käesolev uurimistöö uurib tänapäevase mitmetes valdkondades kasutuseloleva masinõppeprogrammi FaceReader võimekust inimemotsioonide tuvastamisel, keskendudes seostele videos esinevate inimeste ekstraversusega. FaceReaderit rakendatakse üle maailma näiteks erinevates psühholoogiaalastes uuringutes (vaatlused, käitumiskspereimendid, psühhofüsioloogilised testid), tarbimiskäitumist ja -eelistusi hindavates uuringutes, neuroteaduses, toodete või teenuste kasutajakogemuse parandamisel ning sotsiaalmeediauuringutes (Noldus, 2025). Emotsioonituvastustarkvarade potentsiaalne kasutusala on väga lai ning paljusid neist arendatakse juba praegu innukalt: näiteks saavad tööandjad emotsioonituvastusprogramme kasutada uute töötajate värbamisel (jälgides nende stressitaset intervjuu ajal või väljendatud emotsioonide siirust), tervishoiutöötajad patsientide enesetunde monitoorimisel (emotsioonituvastusprogramm võiks anda märku muutustest patsiendi tujus või seisundis) või brändijuhid oma toodete turundamisel, jälgides emotsioonituvastusprogrammide abiga klientide reaktsioone näiteks nende reklaamide nägemisel või toote kasutamisel (Smith jt, 2024).

Emotsioonid, näoväljendused ja nende tähtsus inimsuhtluses

Emotsioonid ning afektiivsed seisundid arvatakse olevat kujunenud evolutsiooni käigus, kuna neil on adaptiivne väärtus, mis aitab inimestel elus toime tulla (Ekman, 1992). Alates sellest, kui Charles Darwin avaldas 1872. aastal teose “Emotsioonide väljendumine inimesel ja loomadel”, on näoväljendustel arvatud olevat oluline roll emotsionaalse kogemuse signaliseerimisel (Dupré jt, 2020). Darwini teooria kohaselt on näoilmed suuresti instinktiivsed jäänukid meie loomadest eelkäijate käitumisest, mis kunagi täitsid ellujäämisega seotud eesmärgid ning andsid liigikaaslaste ees konkurentsieelise (Dupré jt, 2020). Ehkki Darwin ei postuleerinud olemuslikku seost emotsioonide ja näoväljenduste vahel, põhineb tema töodel hilisem baasemotsioonide teooria, mis väidab, et eksisteerib limiteeritud arv emotsioone (näiteks rõõm, kurbus, viha, hirm, üllatus ja vastikus), mida iseloomustavad just sellele emotsioonile iseloomulikud väljendusviisid (Dupré jt, 2020).

Näoilmetes nähakse niisiis võimalust nende põhiemotsioonide tuvastamiseks, kuna need koosnevad spetsiifilistest näolihaste liigutuste konfiguratsioonidest, mida paljud teadlased arvavad olevat prototüüpsed, kaasasündinud ja universaalsed (Dupré jt, 2020). Emotsioonide

mõistmine on vajalik adumaks, kuidas inimesed maailma tajuvad ja sellega suhestuvad. Need annavad meile (nii liigisest kui liigiülest) informatsiooni asetleidnud sündmuste, nendega kaasnevate reaktsioonide ja neile tõenäoliselt järgneva käitumise kohta (Ekman, 1992). Näiteks kui näeme vastikust väljendava näoilmega inimest, võime oletada, et antud isiku reaktsioon on vastuseks millelegi ebameeldivale või vastumeelsele, et isik toob suurema tõenäosusega kuuldavale vastikust kui näiteks naudingut või rahulolu väljendava häälsuse, ning et suure tõenäosusega pöörab isik vastikust tekitanud stiimulist eemale või kaugeneb sellest (Ekman, 1992).

Lisaks on emotsioonid hädavajalikud interpersoonalsete suhete kujunemisel ja regulatsioonis (Ekman, 1992). Näiteks on emotsioone väljendavad näoilmed seotud kiindumussuhete kujunemise ja reguleerimisega (nii imikueas kui ka paarisuhete loomisel), aga ka agressiooni reguleerimisega (Ekman, 1992). Ekman toob välja, et on uurinud Möbiuse sündroomi ehk kaasasündinud näonärvide arenguanomaaliaga inimesi, kes on raporteerinud suuri raskuseid suhete loomisel ja säilitamisel teiste inimestega, kuna näolihaste halvatus tõttu puudub neil võime end miimika ja näoilmete kaudu väljendada (Ekman, 1992).

Ekmani käsitluse kohaselt saab eristada kuut baasemotsiooni, milleks on viha, hirm, rõõm, kurbus, vastikus ja üllatus, kuna nendega kaasnevad distinktiivsed ja universaalsed näoväljendused, ehkki üllatuse puhul on tõendid selgesti eristatavatest näoväljendustest mõnevõrra ebakindlamad kui teiste baasemotsioonide puhul (Ekman, 1992). Niisiis on paljude emotsioonide tähenduslikkus paljude teadlaste meelest suuresti universaalne ja aitab meil mõista iseenda ja teiste käitumist. Võib öelda, et emotsioonid on efektiivse suhtluse ja sotsiaalse läbikäimise aluseks (Mayer jt, 2008). Siiski tasub mainida, et mitte kõik teadlased pole emotsioonide universaalsuse teemal üksmeelel – mõned uuringud näiteks on leidnud, et emotsioone väljendavaid näoilmeid ei saa pidada universaalseteks, näiteks võrreldes Ida ja Lääne kultuure (Jack jt, 2009; Jack jt, 2012).

Emotsioone võib näha ka signaalidena - rõõm näiteks sisaldab endas signaali soovist teistega ühineda, kurbus jällegi annab märku mingisugusest kaotusest ja soovist saada lohutust ja tuge või olla üksi (Mayer jt, 2008). Emotsioonide uurimine on oluline, sest emotsioonide regulatsiooni seostatakse sotsioemotsionaalse kompetentsuse ja vaimse tervisega nii laste kui täiskasvanute puhul, ning just emotsioonide, nende tekkimise ja arengu paremini mõistmine aitab

arendada välja uusi ja tõhusamaid meetodeid ning terapeutilisi sekkumisi nende reguleerimiseks (Pollak jt, 2019).

Ekstravertsuse seosed emotsioonide äratundmisega

Kõik inimesed ei pruugi väljendada emotsioone sarnaselt. Mitmed varasemad uuringud on leidnud, et loomult ekstravertsemad inimesed kipuvad olema ekspressiivsemad ehk emotsionaalselt väljendusrikkamad ja ilmekamad (Kring jt, 1994; Riggio & Riggio, 2002). Nii Rebane (2023) kui Veltmann (2024) leidsid, et kui näidata hulgale katseisikutele klippe videomodellidest väljendamas erinevaid emotsioone, siis kõrgema ekstravertsusega inimeste väljendatud emotsioone tunti statistiliselt olulisel määral rohkem õigesti ära.

Ekstravertsuse ja väljendusrikkuse seoste uurimisel saab kasutada ACT-testi (*Affective Communication Test*; Friedman jt, 1980), mis mõõdab individuaalseid erinevusi mitteverbaalses suhtluses, täpsemalt indiviidide soovi ja võimekust mitteverbaalselt ilmekalt väljenduda (Tucker & Friedman, 1993). On leitud, et ekstravertsus korreleerub nii meestel kui naistel tugevalt ACT-testi skooriga (Tucker & Friedman, 1993). Ilmekalt väljendumise oskust võiks lugeda ekstravertsuse üheks aspektiks ehk et väljendusrikkuse näol on tegemist laia ja üldise ekstravertsuse konstrukti teatud osa rekontseptualiseeringuga, mitte täielikult uue konstruktiga (Tucker & Friedman, 1993).

Friedman ja kolleegid viisid 1980. aastal läbi uuringu, kus katseisikutel paluti pärast isiksusetesti täitmist kaamera ees läbi etteantud lausete esitamise erinevaid baasemotsioone näidelda. Neist salvestatud videoklipid anti vaatamiseks hindajatest katseisikute rühmale, kes pidid kuue etteantud emotsiooni (rõõm, kurbus, viha, üllatus, hirm, vastikus) või neutraalse ilme seast valima, millist emotsioon või näoilmet videos esinev isik väljendab (Friedman jt, 1980). Selgus, et ehkki ekstravertsus näis mõnevõrra korreleeruvat võimekusega emotsioone etendada, siis hoopis tugevam seos esines kitsama ekshibitsiooni (*Exhibition*) konstruktiga, mis pärineb Murray (1938) baasvajaduste süsteemist ja viitab vajadusele olla dramaatiline ja tähelepanu keskpunktis (Friedman jt, 1980). Uuringu autorid järeldasid sellest, et inimestele, kes on osavad emotsioonide etendamises ehk kelle näideldud emotsioone tunti paremini ära, meeldib olla teiste inimeste seltskonnas, nad on impulsiivsed ja vähem vaoshoitud (nagu ekstraverdidki), kuid eriti naudivad ja ihkavad olla dramaatilised ning esineda publikule (Friedman jt, 1980).

Ekstravertsuse uurimine on oluline, kuna see aitab paremini mõista, kuidas isiksuseomadused mõjutavad emotsionaalsete väljenduste äratundmist ja tõlgendamist. See on

eriti tähtis emotsioonituvastusprogrammide arendamisel, et tagada nende täpsus ja usaldusväärsus erinevate inimtüüpide ja -käitumiste kontekstis.

Emotsioonide tuvastamisele keskenduv tarkvara

Termin “afektiivne tarkvaraarendus”, mis viitab emotsioonidega seotud tarkvarale, defineeriti juba mõned aastakümned tagasi, 1995. aastal (Garcia-Garcia jt, 2017). Just viimastel aastatel on aga eriti jõudsalt arendatud erinevaid masinõppeprogramme, mis keskenduvad inimeste emotsioonide tuvastamisele.

Enamik ekspressiivse käitumise analüüsimisele keskenduval programme tugineb eelnevalt mainitud baasemotsioonide teooriale – programmid kategoriseerivad videost või pildi pealt tuvastatud emotsiooni kindla arvu etteantud emotsioonide seast, milleks ongi üldjuhul kuus baasemotsiooni ehk rõõm, kurbus, viha, hirm, üllatus ja vastikus (Dupré jt, 2020). Selline süsteem põhineb eeldusel, et emotsionaalsed väljendused vastavad prototüüpsetele näoliigutuste muustritele (Dupré jt, 2020).

Emotsioonide tuvastamist on uuritud peamiselt läbi kolme liiki andmete: näideldud emotsioonid, loomulikud spontaansed emotsioonid ja stiimuliga esile kutsutud emotsioonid (Caridakis jt, 2007). Parimad tulemused saadakse näideldud emotsioonide põhjal, kuna need sisaldavad endas ilmekaamaid ja stereotüüpsemaid emotsionaalseid väljendusi. Tundub loogiline, et masinõppeprogramm, mida on treenitud videotega näideldud emotsioonidest, ei pruugi spontaansed emotsioone nii täpselt tuvastada, kuna näideldud ja tõelisi näoväljendusi vahendavad kaks eraldiseisvat juhteteed, mis saavad alguse aju eri osadest (Littlewort jt, 2007). Tahtlikud näolihaste liigutused saavad alguse kortikaalsest motoorsest juhteteest, samas kui spontaansed näoilmed pärinevad aju subkortikaalsetest piirkondadest (Littlewort jt, 2007). Kortikaalne süsteem innerveerib tugevamalt pigem teatud lihaseid näo alaosas, subkortikaalne süsteem aga näo ülaosas (Littlewort jt, 2007). Kahest erinevast juhteteest alguse saanud näoilmed erinevad teineteisest dünaamika poolest ning rakendavad erinevaid näolihaseid. Subkortikaalsest juhteteest pärinevad näoilmed on üldjuhul sünkroniseeritud, sujuvad, sümmeetrilised, järjepidevad ja meenutavad reflekse. Kortikaalsest juhteteest alguse saanud näoilmed kipuvad olema vähem sujuvad ning dünaamika poolest varieeruvamad (Littlewort jt, 2007). See annab alust oletada, et teeseldud ja tõelist emotsiooni väljendavad näoilmed on erinevad, mistõttu näideldud emotsioone tuvastama õpetatud masinõppeprogramm ei pruugi edukalt tuvastada spontaanselt esinenud emotsioone ja vastupidi.

Emotsioonide tuvastamisele spetsialiseerunud tarkvara eesmärgiks on analüüsida näoilmeid, häält, žeste, teksti ja füsioloogilist seisundit, et automaatselt emotsioone klassifitseerida (Garcia-Garcia jt, 2017). Siiski on automatiseeritud programmi välja töötamine, mis emotsioonide tuvastamisega hästi hakkama saaks, kujunenud tarkvaraarendajatele parajaks väljakutseks. Probleeme võib esineda mitmes etapis: esmalt teatud pildi või videoklipi osast inimnäo tuvastamine, teisalt näoilme kohta informatsiooni hankimine ning viimaks ilme liigitamine (näiteks emotsioonikategoriatesse; Pantic & Rothkrantz, 2000).

Et edukalt inimemotsioone tuvastava masinõppeprogrammi välja töötamine pole osutunud kergeks ülesandeks, võiks selgitada ka asjaoluga, et inimestel endilgi pole alati lihtne kellegi teise näoilmet tõlgendada ja nad kipuvad selles eksima. John Bassili 1978. aasta uuringu kohaselt suudab treenitud hindaja näo, mis väljendab ühte kuuest baasemotsioonist, korrektselt klassifitseerida keskmiselt 87% tõenäosusega. Tulemus oleneb mitmest faktorist: kuivõrd on hindaja hinnatava näoga ja iseloomuga tuttav, hindaja üldine kogemus eri näoilmete äratundmisel, näole suunatav tähelepanu ning mittevisuaalsed vihjed (näiteks kontekst, milles nägu esineb; Bassili, 1978). Seega näib liialt optimistlik oodata masinõppeprogrammidelt suurepäraseid tulemusi, kui ka inimesed ise pole teiste emotsioonide lugemises täiuslikud.

FaceReader programm

Üheks inimemotsioonide tuvastamisele keskenduvaks programmiks on FaceReader tarkvara (FaceReader 9.0, Noldus Information Technology, Wageningen, Holland), mis võimaldab analüüsida inimese näoväljendusi kas reaajas või eelsalvestatud videomaterjali põhjal. FaceReader põhineb Paul Ekmani FAC-süsteemil (*Facial Action Coding System*) ning on treenitud emotsiooni liigitama ühe Ekmani välja pakutud põhiemotsiooni (rõõm, kurbus, viha, üllatus, hirm, vastikus) alla. FaceReader võib liigitada emotsiooni ka “neutraalseks”. Andmeid saab sünkroniseerida ka näiteks füsioloogiliste andmete või silmaliigutustega (Dupré jt, 2020). FaceReaderit on treenitud üle 20 000 pildiga, et tehisklik neuraalne võrgustik oleks võimeline tuvastama mustreid näos (Loijens & Krips, n.d.).

FaceReader tegutseb näoilme tuvastamisel mitmes osas. Esmalt tuleb video- või pildimaterjalilt tuvastada nägu - programm kasutab esmalt populaarset Viola-Jones algoritmi, et tuvastada näo olemasolu, ning seejärel luuakse täpsem näo piirjooni määratlev “mask” (Noldus Information Technology, n.d.) Järgmisena loob programm näost detailsema 3D-mudeli, lähtudes 491 punktist näol, seejuures näiteks punktidest kulmudel, silmadel ja suul, mis on teatud

emotsioonide puhul tugevateks indikaatoriteks (Noldus Information Technology, n.d.). Programm võtab arvesse ka näo tekstuuri, näiteks informatsiooni naha kortsumise ja valgusemuudatuste kohta, mis samamoodi sisaldavad vihjeid näol oleva ilme kohta (Noldus Information Technology, n.d.).

2014. aastal viisid Lewinski, den Uyl ja Butler läbi uuringu, kus testisid tollal uusimat (6.0) versiooni FaceReader programmist. FaceReaderile esitati 207 pilti WSEFEP pildikogumikust (*Warsaw Set of Emotional Facial Expression Pictures*) ning 154 pilti ADFES pildikogumikust (*Amsterdam Dynamic Facial Expression Set*). Iga pilt väljendas ühte kuuest baasemotsioonist (rõõm, kurbus, vastikus, hirm, viha või üllatus) või neutraalset ilmet (Lewinski jt, 2014). Esimesel pildikogumiku puhul tuvastas FaceReader õige emotsiooni 88% piltidelt ning teise kogumiku puhul 89% piltidelt (Lewinski jt, 2014). Näo tuvastamisel FaceReaderil erilisi probleeme ei esinenud - esimese pildikogumi puhul ei õnnestunud programmil pildilt nägu tuvastada 0,95% juhtudel ning teise pildikogumi puhul 3,77% juhtudel (Lewinski jt, 2014). Mõlemal juhul tuvastas FaceReader kõige edukamalt rõõmu (äratundmisskoor 96%), kõige kehvemini aga viha (äratundmisskoor 76%). Programm oli naismodellide nägudest emotsioonide välja lugemisel veidi edukam kui meesmodellide puhul (äratundmisskoorid vastavalt 89% ja 86%; Lewinski jt, 2014). Inimhindajate emotsioonide äratundmisskoor oli nende kahe andmekogumi puhul keskmiselt 85%, niisiis tegid autorid antud uuringutulemuste põhjal järelduse, et FaceReader on emotsioonide ära tundmises sama võimekas kui inimene (Lewinski jt, 2014).

2020. aastal jõudsid Dupré ja kolleegid aga mõnevõrra erinevate tulemusteni. Uurimisrühm andis kaheksale emotsioonide automaatseks klassifitseerimiseks loodud programmile, mille hulka kuulus ka FaceReader (antud uuringus versioon 7.0), analüüsimiseks 937 stiimulvideot, mis kõik väljendasid ühte kuuest baasemotsioonist. Videod pärinesid kahest andmebaasist, kus ühe andmebaasi (*BU-4DFE*) videod väljendasid spontaanseid emotsioone, teise andmebaasi (*UT Dallas*) videod aga näideldud emotsioone (Dupré jt, 2020). Kõik esimese andmebaasi videoklipid (näideldud) nii algasid kui lõppesid neutraalse näoilmega ja emotsiooni väljendav osa klipist kestis keskmiselt 4 sekundit (Dupré jt, 2020). Teise andmebaasi videoklipid (spontaansed) võisid alata ja lõppeda nii neutraalse kui ka emotsiooni väljendava näoilmega ning emotsiooni väljendav osa klipist kestis keskmiselt 6 sekundit (Dupré jt, 2020). Ehkki FaceReader oli 57,3% täpsusskoori juures edukuselt teine programm, jäi ta emotsioonide korrektsetes

tuvastamises siiski oluliselt alla inimhindajatele, kelle keskmine täpsusskoor oli 72,5% (Dupré jt, 2020). Kui näideldud emotsioonide tuvastamisel ei olnud FaceReader edukuselt inimhindajatest oluliselt kehvem, siis spontaansete emotsioonide hindamisel olid inimhindajad kõigist uuritud masinõppeprogrammidest peajagu üle (Dupré jt, 2020). Kuna varasemadki uuringud on leidnud, et spontaansed emotsiooniväljendused erinevad teadlikult näideldud emotsiooniväljendustest näiteks kestuse ja intensiivsuse poolest, oletasid ka antud uuringu läbiviijad, et spontaansete emotsioonide kehvem tuvastamine tuleneb tõenäoliselt videomodellide vähesemast ekspressiivsusest emotsioonide väljendamisel (Dupré jt, 2020).

Hetkel on kasutamiseks saadaval FaceReaderi versioon 9.1, mis Nolduse esindajate väitel tuvastas valideerimisel kõiki baasemotsioone peale kurbuse 100-protsendilise täpsusega – kurbuse puhul oli täpsuseks 95,8% ning 4,2% kordadest liigitas programm antud emotsiooni viha alla (Loijens & Krips, n.d.). Ka ei tohiks FaceReaderi jaoks väidetavalt vahet olla, kas tegemist on spontaanse ja reaalselt tajutud või käsu peale näideldud emotsiooniga – programm tuvastab emotsiooni edukalt mõlemal juhul (Loijens & Krips, n.d.).

Univer (2022) on oma magistritöö raames võrrelnud masinõppeprogrammi ja inimeste võimekust tuvastada baasemotsioone samadelt stiimulvideotelt, mida antud uurimistööski kasutatakse. Univeri uuringus osalenud katseisikute (inimhindajad) keskmine võimekus emotsioonide äratundmisel oli 43,9%. Kõige paremini tunti ära rõõmu (77,4%; Univer, 2022). Need tulemused jäävad märkimisväärselt alla eelmainitud välismaistes uuringutes osalenud katseisikute skooridele (85% ja 72,5% inimhindajate keskmine täpsusskoor varasemates uuringutes) (Lewinski jt, 2014; Dupré jt, 2020). Masinõppeprogramm, millele Univer keskendus, oli aga teine, ning kogutud andmeid polnud lõpuks võimalik esitada, kuna need jäid märkimisväärselt alla varasemate tööde tulemustele (Univer, 2022).

Nagu varasemalt välja toodud, viitavad uuringud, et ekstravertsema iseloomuga inimesed kipuvad olema emotsioonide väljendamisel ilmekamad (Kring jt, 1994; Riggio & Riggio, 2002). Et masinõppeprogrammidel on lihtsam tuvastada videomodellide emotsioone, kui need on esitatud ekspressiivsemalt, võiks siit loogiliselt järelduda, et FaceReader programm võiks – nagu inimesedki – ekstravertsemate videomodellide emotsioone paremini ära tunda. Varasemad leiud annavad alust oletada, et nende näoilmed on keskmiselt väljendusrikkamad.

Käesoleva uurimistöö eesmärk ja hüpoteesid

Käesoleva uurimistöö eesmärk on uurida kaasaegse masinõppeprogrammi võimekust inimemotsioonide tuvastamisel FaceReader programmi näitel, keskendudes seostele videomodellide ekstraversusega. Uurimisküsimus kõlab järgmiselt: milline on FaceReader programmi võimekus videoklippidest inimemotsioonide tuvastamisel ja kas sellel on seos videomodelli ekstraversusega? Varasemale kirjandusele ja leidudele tuginedes püstitasin järgnevad hüpoteesid:

H1. FaceReader programmi algoritm suudab väljendatud baasemotsioone tuvastada vähemalt 75% täpsusega.

H2. FaceReader programm tuvastab kõrge ekstraversusega videomodellide väljendatud emotsioone rohkem õigesti kui madala ekstraversusega videomodellide väljendatud emotsioone statistiliselt olulisel määral.

Meetod

Valim

Uurimistöös kasutatavad andmed pärinevad kahelt valimilt ning autori originaalpanus on FaceReader programmi emotsioonituvastustulemuste analüüsimine. Videomodellide valim (119 osalenut) koguti algselt Juuse (2020) magistr töö raames, kes värbas vabatahtlikud osalejad ülikooli meililistist ning isiklikest sotsiaalmeediakanalistest. Veltrmann (2024) kasutas enda uurimistöös neist 100 osalenu andmeid. FaceReader programm ei suutnud ühe osaleja emotsioone tuvastada ning selles uuringus kasutatav lõplik valim oli seega 99 osalejat (Joonis 1). 61 videomodellidest olid naised ja 38 mehed. Modellide keskmine vanus oli 24,9 eluaastat ($SD = 6,41$), noorim modell oli 18 ja vanim 49 aasta vanune. 77 osalejat täitsid katse eesti keeles, 16 vene keeles ning 6 inglise keeles. 5 katseisiku kohta puudub isiksusetesti tulemus ehk neid ei arvestatud isiksuseanalüüsis.

Rebase (2023) uurimistöö lõpliku valimi näol oli tegemist mugavusvalimiga, kus katseisikud leiti enamjaolt autori tuttavate kaudu ning kuhu kuulus 121 veebikatse täielikult lõpetanud osalejat. Antud töö raames kasutati hinnanguid ka osalejailt, kes andsid kasvõi ühele näoväljendusele hinnangu (osalejad, kes katset täielikult ei lõpetanud). Antud uuringusse kaasati seega 144 osalejat, kellest 115 olid naised ja 28 mehed. Üks katseisik ei soovinud sugu avaldada. Katseisikute keskmiseks vanuseks oli 31 eluaastat ($SD = 13,89$), noorim katseisik oli

14 ja vanim 77 aasta vanune. 88 katseisikut vastasid, et õpivad parajasti, 74 töötas parajasti (võisid samaaegselt ka õppida) ning 10 katseisikut vastas, et tegelevad parajasti millegi muuga.

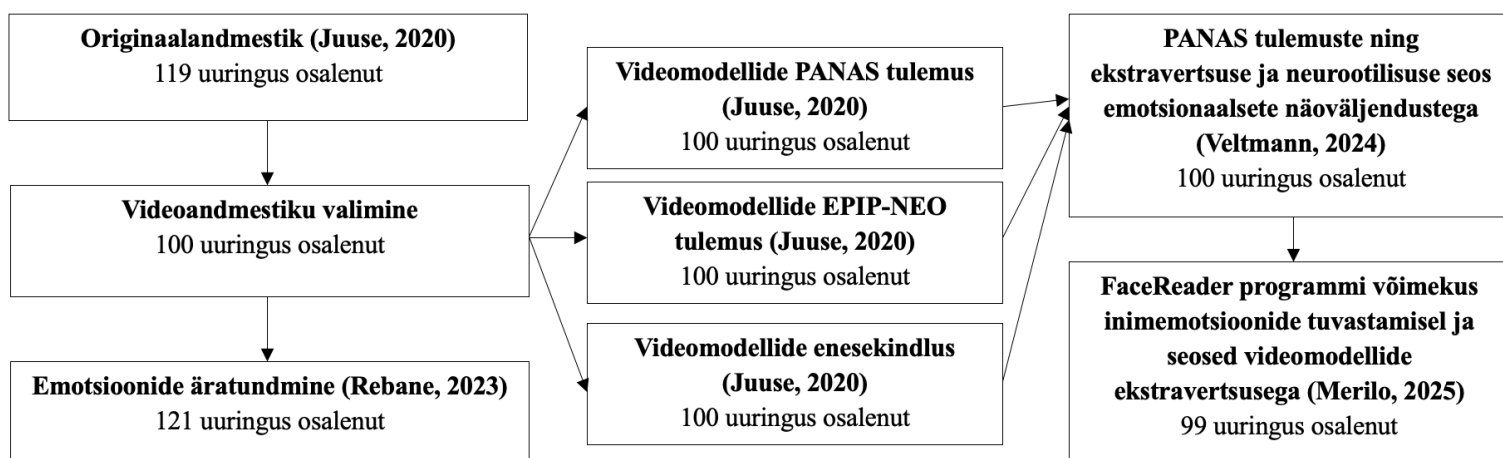
Katsematerjal

Uurimistöös kasutatavad stiimulvideod pärinevad Juuse (2020) magistritöö raames kogutud videoklippide andmebaasist. Tegemist on 6-sekundiliste videoklippidega katseisikutest, kel oli laborikatses palutud väljendada ühte kuuest baasemotsioonist (rõõm, kurbus, viha, hirm, üllatus, vastikus). Osalejatel paluti täita ka 240 enesekohasest küsimusest koosnev isiksuseküsimustik EPIP-NEO (Möttus jt, 2006), mis mõõdab suure viisiku isiksuseomadusi ning andis ülevaate videomodellide isiksuseprofiilist.

Teine osa käesolevas uurimistöös kasutatavaid andmeid pärineb Rebase (2023) uurimistööst, kus katses osalejatel tuli Tartu Ülikooli psühholoogia instituudi testimiskeskonnas KAEMUS vaadata 1-sekundilisi videoklippe ja hinnata, millist baasemotsiooni videos esinenud inimene väljendab. Katses kasutatud 1-sekundilised videolõigud kujutasid endast lühendatud versioone Juuse (2020) magistritöö raames kogutud 6-sekundilistest videotest. Juuse (2020) magistritöös oli katseisikutel palutud baasemotsioone väljendada neljas erinevas tingimuses: Ekmani näopiltide andmebaasist pärineva näoilme kopeerimine, sõnaliselt kirjeldatud emotsiooni jäljendamine, liialdatud ja varjatud väljendusviis. Rebase (2023) uurimuses kasutati neist vaid esimest, ekraanil visuaalselt esitatud Ekmani baasemotsiooni väljendamise tingimust.

Rebase (2023) valis enda uurimistöös katseisikutele esitamiseks Juuse (2020) magistritöös kogutud kuut baasemotsiooni väljendavatest videotest välja vaid kolme emotsiooni – hirm, rõõm ja vastikus – väljendavad videoklipid. Rebase (2023) tööle tuginedes tõi Veltmann (2024) välja varasema uuringu, mis leidis, et neid kolme emotsiooni aetakse omavahel vähem segamini kui teisi (Calvo & Nummenmaa, 2015). Ehkki Rebase (2023) läbi viidud katses esitati katseisikutele vaid neid kolme emotsiooni väljendavaid videoklippe, said katses osalejad emotsioonide hindamisel teha pakkumise kõigi kuue baasemotsiooni seast.

Et antud uurimistöös kogutud FaceReader programmi tulemusi emotsioonide tuvastamisel oleks võimalik Rebase (2023) uurimistöös kogutud inimhindajate tulemustega kõrvutada, anti antud uurimistöös FaceReaderile analüüsimiseks küll kõiki baasemotsioone väljendavad videoklipid, kuid neist edasi analüüsi võid visuaalse jälgendamise tingimusest saadud rõõmu, vastikust ja hirmu väljendavaid videoklippe. Uurimistöös kasutati analüüsiks 1-sekundilisi videoklippe, et tulemused oleksid paremini kõrvutatavad Rebase (2023) tööst saadud inimhinnangute väärtustega. 1-2 sekundi valik tulenes videomaterjali eelneva hindamise põhjal sõltumatutelt hindajatelt (5 hindajat), kes hindasid 6-sekundiliste klippide puhul emotsiooni alg-, kõrg- ning lõpp-punkti. Hindajate hinnangu põhjal lõigati klipid 1-sekundilisteks nende hinnatud algpunkti alusel. Varasemaid töid ning antud uurimistöö paigutumist nende konteksti kirjeldab täpsemalt alltoodud joonis (Joonis 1).



Joonis 1. Uuringus kasutatavate andmete päritolu. Modifitseeritud joonis Veltmanni (2024) tööst. Töös kasutatav muutuja (videomodeltide ekstraversus) pärineb varasematest uuringutest (Juuse, 2020). PANAS = Positive and Negative Affect Schedule, EPIP-NEO = International Personality Item Pool – Neuroticism, Extraversion and Openness (Möttus jt (2006) eestistatud NEO-PI-R test 240 väitega).

Uurimisprotseduur

Antud uurimistöö raames anti laboritingimustes FaceReader 9.0 programmile analüüsimiseks eelnevalt kirjeldatud videoklippe, mis väljendasid erinevaid baasemotsioone. Andmete analüüsimine toimus Tartu Ülikooli neuroturunduse laboris ajavahemikus juuni – august 2024. FaceReader programmile anti analüüsimiseks 6-sekundiline videoklipp ning

ajatempli abil vaadati 1-2 sekundi osa, kuna varasemalt hindasid ka inimhindajad just seda sekundit.

Eetika

Mõlemal uuringul, mille raames kogutud andmetele antud uurimistöös tuginetakse (Juuse, 2020; Rebane, 2023), oli uuringu läbiviimiseks olemas Tartu Ülikooli inimuuringute eetikakomitee kooskõlastus. Katses osalemiseks pidid osalejad täitma informeeritud nõusoleku vormi, mis tuletas meelde uuringu vabatahtlikkust ning õigust igal hetkel katse katkestada (Juuse, 2020; Rebane, 2023). Kogutud andmed olid anonüümsed ning neid kasutatakse vaid õppe- või teadustöö tarbeks üldistatud kujul (Rebane, 2023). Isikustatud andmed hävitati peale uurimisperioodi lõppu (31.03.2024) (Rebane, 2023). Mõlemas uuringus osalemise eest oli psühholoogiatudengitel võimalik saada ainetundides nõutavaid katsepunkte. Juuse (2020) uuringus said katses osalejad kompensatsiooniks ostukeskuse kinkekaardi ning Rebase (2023) uuringus osalejatel oli võimalus osaleda kinkekaardi loosimises.

Autori panus

Autor annab oma panuse emotsioonituvastustarkvara uurimisse lähemalt, keskendudes seostele ekstravertsuse ja emotsioonide väljendamise vahel. Tegemist on uudse lähenemisega, kuna varasemad uuringud on peamiselt keskendunud masinõppeprogrammide ja inimhindajate üldise täpsuse võrdlemisele, jättes kõrvale indiviidide isiksuseomadused. Käesoleva töö tulemused võivad aidata kaasa paremate ja täpsemate emotsioonituvastustarkvarade arendamisele, mis suudavad arvestada erinevate isiksusetüüpide eripäradega.

Sama materjali kasutab oma uurimistöös ka Viisileht (2025), kes keskendus varem kogutud inimhinnangute ja FaceReaderi vastuste põhjal mõlema täpsusprotsentide hindamisele ning FaceReaderi võimekusele erinevate emotsioonide eristamisel. Lisaks kasutas Viisileht (2025) oma töös protsentuaalset tuvastustäpsust ehk vaatas, kui kõrgelt FaceReader igas videoklipis kõiki emotsioone tuvastab (protsentides), samas kui mina keskendusin binaarsele muutujale (tabamus/möödalask), kus vaatasin, kas FaceReader pakkus kõrgeimalt emotsiooni, mida klipis päriselt esitatud oli (tabamus), või mõnda muud emotsiooni (möödalask).

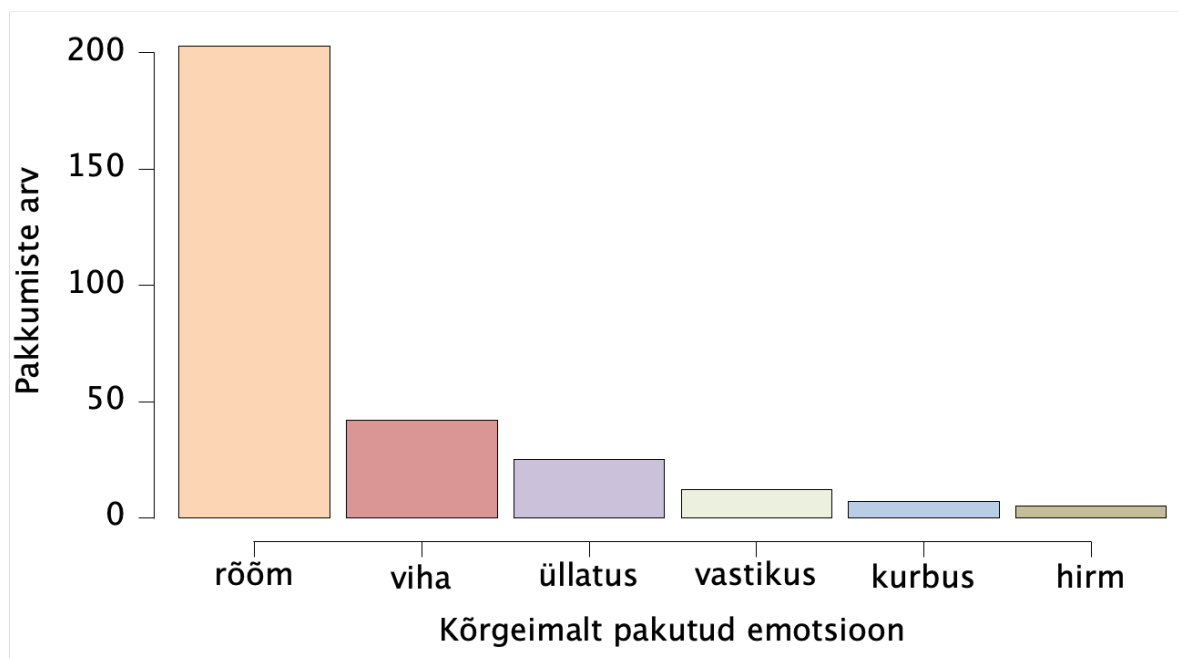
Andmeanalüüs

Antud uurimuses oli sõltuvaks muutujaks FaceReader programmi (FaceReader 9.0, Noldus Information Technology, Wageningen, Holland) täpsus emotsioonide tuvastamisel ja sõltumatuteks muutujateks olid videoklippides väljendatud emotsioonid (rõõm, vastikus, hirm) ning videomodelli ekstraversuse skoor (kogutud Juuse (2020) raames). Hüpoteeside testimiseks kasutasin JASP (versioon 0.19) tarkvara. Esimese hüpoteesi (H1: FaceReader programmi algoritm suudab väljendatud baasemotsioone tuvastada vähemalt 75% täpsusega) testimiseks vaatasin esmalt FaceReaderi täpsust kirjeldava binaarse muutuja (0 = möödalask ja 1 = tabamus) keskmist kõigi uuritavate emotsioonide lõikes, mis andis mulle FaceReaderi täpsust väljendava protsendi iga emotsiooni puhul eraldi. Lasin JASP programmil koostada ka sagedustabeli, mis andis ülevaate, mitmel korral pani FaceReader vastusega täppi ning mitmel korral eksis. Lisaks viisin läbi hii-ruut testi muutujaga, mis väljendab emotsiooni, mida FaceReader iga klipi puhul kõige kõrgemalt pakkus. Saadud sagedustabel näitab, milliseid emotsioone ja millise sagedusega FaceReader kõiki kolme emotsiooni (rõõm, vastikus, hirm) väljendavatele videoklippidele pakkus (Tabel 1). Viisin läbi ka hii-ruut testi hindamaks seost FaceReader programmi kõrgeimalt pakutud emotsiooni ja videoklipis päriselt esitatud emotsiooni vahel ning kas FaceReaderi täpsus emotsioonide tuvastamisel erineb juhuslikust pakkumisest.

Teise hüpoteesi (H2: FaceReader programm tuvastab kõrge ekstraversusega videomodellide väljendatud emotsioone rohkem õigesti kui madala ekstraversusega videomodellide väljendatud emotsioone statistiliselt olulisel määral) testimiseks viisin läbi logistilise regressioonanalüüsi FaceReader programmi tabavust näitava muutuja ja videomodellide ekstraversuse skoori vahel. See võimaldas vaadata, mil määral videomodellide ekstraversuse skoor ennustab seda, kas FaceReader määrab nende poolt väljendatud emotsiooni õigesti. Mudelisse lisasin faktorina ka näideldud emotsiooni, et arvestada FaceReaderi erinevat täpsust erinevate emotsioonide tuvastamisel. Niisiis oli sõltuvaks muutujaks mudelis FaceReaderi tabamus/möödalask, sõltumatuteks muutujateks olid modelli ekstraversuse skoor ning videoklipis väljendatud emotsioon (rõõm, vastikus, hirm). Analüüsi eeldused olid täidetud, kuna sõltuv muutuja oli dihhotoomne, vaatlused olid sõltumatud ja muutujate vahel ei esinenud multikollineaarust.

Tulemused

FaceReader programm pakkus teatud emotsioone selgelt sagedamini kui teisi. Kõige enam tuvastas programm videoklippidest protsentuaalselt kõrgeimalt rõõmu (203 korral 294 pakkumisest). Viha tuvastas programm protsentuaalselt kõrgeimalt 42 korral, üllatust 25 korral, vastikust 12 korral, kurbust 7 korral ja hirmu 5 korral. FaceReaderi pakkumisi emotsioonide lõikes iseloomustab alltoodud joonis (Joonis 2).



Joonis 2. FaceReader programmi pakkumised emotsioonide lõikes.

FaceReader pakkus emotsiooni õigesti (väljendatud emotsioonideks olid rõõm, vastikus ja hirm) 110 korral ning valesti 184 korral, andes FaceReaderi üldiseks täpsuseks kõigi uuritud kolme emotsiooni üleselt 37,4% ($SD = 0,49$). FaceReader suutis vähemalt 75% täpsusega tuvastada vaid rõõmu, tuvastades selle õigesti 96,9% kordadest. Vastikuse puhul oli õigesti tuvastamise protsent 11,3% ja hirmu puhul kõigest 4%. Sagedustabel kirjeldamaks programmi tehtud pakkumisi erinevate emotsioonide lõikes on välja toodud alltoodud tabelis (Tabel 1).

Tabel 1

FaceReader programmi pakkumiste (tabamuste ja möödalaskude) määr erinevate emotsioonide lõikes

Väljendatud emotsioon	FaceReaderi tuvastatud emotsioon						
	Rõõm	Vastikus	Hirm	Viha	Üllatus	Kurbus	Kokku
Rõõm	95	0	0	3	0	0	98
Vastikus	59	11	1	19	2	5	97
Hirm	49	1	4	20	23	2	99
Kokku	203	12	5	42	25	7	294

Märkus. Tabelis on programmi antud vastused arvudena (*counts*). Rasvases kirjas on märgitud programmi tabamused uuritava kolme emotsiooni (rõõm, vastikus, hirm) puhul.

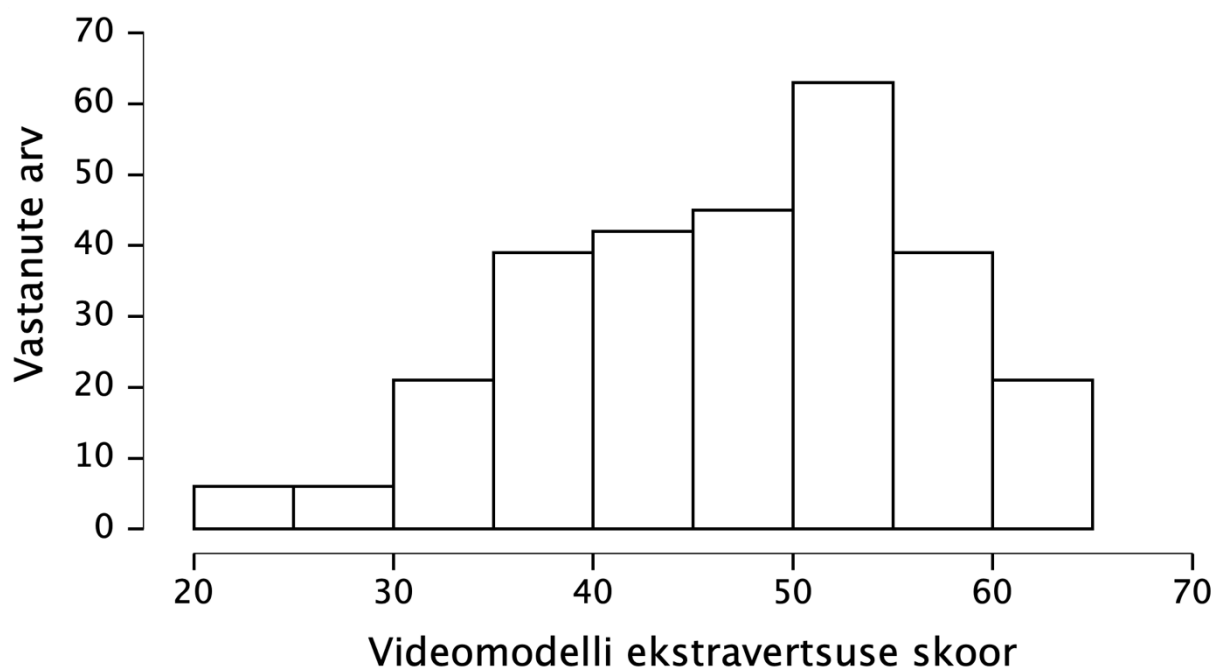
Et uurida, kas FaceReader programmi tabamuste-möödalaskude määr oli juhuslikust pakkumisest statistiliselt olulisel määral erinev, viisin läbi hii-ruut testi. Selle tulemus ($\chi^2(10) = 98,21, p < 0,001$) viitas olulistele erinevustele emotsioonide tuvastamismääras ning järeldasin, et FaceReaderi tulemused ei olnud juhusliku pakkumise tulemus. Juhuslikuks pakkumiseks pidasin kuue võimaliku emotsiooni valiku puhul 0,16 (16%).

Saadud andmete põhjal arvutasin välja FaceReaderi tundlikkuse indeksi (d'), mis näitab, kui hästi suudab FaceReader tuvastada spetsiifilise emotsiooni (näiteks rõõm) esinemist ning seda mitte teiste emotsioonidega (vastikus ja hirm) segi ajada. Tundlikkuse indeks (d') võtab arvesse FaceReaderi õigesti tehtud pakkumisi (nt programm pakub emotsiooniks rõõmu, kui klipis on tõepoolest esitatud rõõm) ning valesti tehtud pakkumisi (n-ö *false alarm* ehk kui programm pakub emotsiooniks rõõmu, kuid klipis esitatakse mõnda teist emotsiooni, antud juhul hirmu või vastikust). FaceReaderi d' väärtused väljendatud emotsioonide puhul on esitatud alltoodud tabelis (Tabel 2).

Tabel 2*FaceReader programmi d' väärtused kolme väljendatud emotsioonide puhul*

Väljendatud emotsioon	Õigete pakkumiste määr (<i>hit rate</i>)	Valede pakkumiste määr (<i>false alarm rate</i>)	d' väärtus
Rõõm	0,969	0,551	1,75
Vastikus	0,113	0,008	1,20
Hirm	0,045	0,008	0,73

Järgmiseks analüüsisin videomodellide ekstravertsuse skooride määra. Need jäid vahemikku 20,94-63,29 ($M = 46,87$; $SD = 9,66$) (kõrgeim võimalik skoor testis 100). Ekstravertsuse skoor jagunes normaaljaotuslikult (asümmeetriakordaja -0,42; järsakusastmekordaja -0,37). Ekstravertsuse skoori jagunemist videomodellide valimi seas illustreerib alltoodud joonis (Joonis 3).

*Joonis 3. Videomodellide ekstravertsuse skooride esinemine valimi lõikes.*

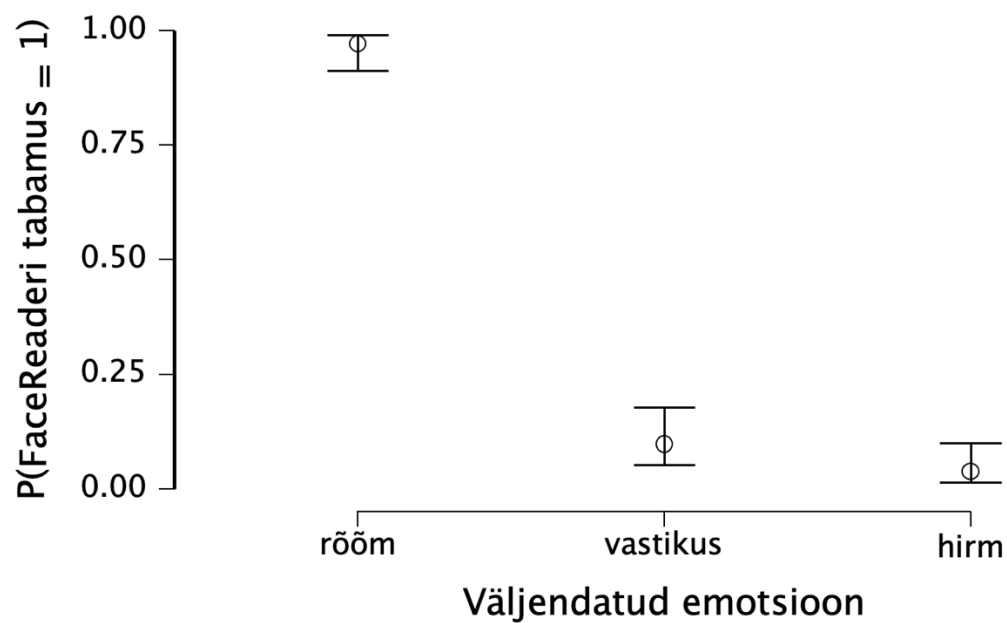
Teise hüpoteesi testimiseks (H2: FaceReader programm tuvastab kõrge ekstraversusega videomodellide väljendatud emotsioone rohkem õigesti kui madala ekstraversusega videomodellide väljendatud emotsioone statistiliselt olulisel määral) viisin läbi logistilise regressioonanalüüsi. Mudeli komponente on täpsemalt kirjeldatud alapeatükis „Andmeanalüüs“. Mudel oli statistiliselt oluline ($\chi^2(278) = 252,73, p < 0,001$) ja selgitas ära 59,2-80,8% (Cox & Snelli ja Nagelkerke) andmete varieeruvusest FaceReaderi emotsioonide tuvastamise täpsusel (tundlikkus = 86,7%, spetsiifilisus = 98,3%). Logistilise regressioonanalüüsi koefitsiente kajastab allpool toodud tabel (Tabel 3). Mudeli kohaselt olid emotsioonid (nii vastikus kui hirm) statistiliselt olulised ($p < 0,001$) ehk vastikuse ja hirmu puhul langeb oluliselt tõenäosus, et FaceReader esitatud emotsiooni õigesti tuvastab (Tabel 3, Joonis 4). Ekstraversuse skoor ei olnud mudelis statistiliselt oluline ($p = 0,054$) (Tabel 3, Joonis 5).

Tabel 3

Tabamuste ja möödalaskmiste määra ennustava logistilise regressioonanalüüsi koefitsiendid sõltuvalt ekstraversusest ja väljendatud emotsioonist

	β	SE	OR	z-skoor	Waldi test		
					Waldi statistik	df	p-väärtus
(Vabaliige)	5,867	1,450	353,200	4,047	16,379	1	<0,001
Ekstraversus	-0,050	0,026	0,951	-1,924	3,701	1	0,054
Emotsioon (vastikus)	-5,736	0,708	0,003	-8,107	65,724	1	<0,001
Emotsioon (hirm)	-6,739	0,811	0,001	-8,305	68,971	1	<0,001

Märkused. Rõõm on kodeeritud klassina 1. β = regressioonikordaja; SE = standardviga; OR = riskisuhe; df = vabadusaste.



Joonis 4. FaceReader programmi tõenäosus emotsiooni (rõõm, vastikus, hirm) korrektseks tuvastamiseks.



Joonis 5. Ekstravertsuse skoori mõju FaceReaderi täpsusele emotsioonide tuvastamisel kolme emotsiooni (rõõm, vastikus, hirm) üleselt.

Arutelu

Antud töö eesmärk oli uurida, kui täpne on FaceReader programm videoklippidest emotsioonide tuvastamisel ning kas esineb seos FaceReaderi tuvastustäpsuse ja emotsioone väljendanud modellide ekstravertsuse skoori vahel.

Andmeanalüüsi käigus selgus, et esimene hüpotees – FaceReader programmi algoritm suudab väljendatud baasemotsioone tuvastada vähemalt 75% täpsusega – sai vaid osaliselt kinnitust. FaceReader suutis vähemalt 75% täpsusega ennustada vaid rõõmu (96,9%) – vastikuse (11,3%) ja hirmu (4%) tuvastamisel oli programm oluliselt vähem edukas. Tulemus FaceReader programmi kehva täpsuse kohta läheb vastuollu varasemate uuringute ja kirjandusega. 2014. aastal läbi viidud uuring FaceReader tarkvara (versioon 6.0) testimiseks (Lewinski jt, 2014) leidis, et ühe pildikogumiku puhul tuvastas FaceReader õige emotsiooni 88% piltidelt ning teise pildikogumiku puhul 89% piltidelt, mis on oluliselt kõrgem tulemus, kui antud uurimistöö tulemustest ilmnes. Ka Lewinski ja kolleegide (2014) uuringus aga tuvastas FaceReader kõige edukamalt rõõmu (äratundmisskoor 96%). See ühtib antud uurimistöö raames saadud tulemustega, kus FaceReader tuvastas rõõmu oluliselt edukamalt kui teisi emotsioone, olles rõõmu tuvastamisel täpne 96,9% kordadest. Siiski tasub siinkohal juhtida tähelepanu, et FaceReader ka pakkus rõõmu oluliselt rohkem kui teisi emotsioone - 294 pakkumisest tervelt 203 korral panustas programm rõõmule kõrgeimalt (ehk tuvastas seda protsentuaalselt teistest baasemotsioonidest kõrgemalt). Kui ligi 69% protsendil kordadest pakub algoritm emotsiooniks rõõmu, siis tundub ka üsna tõenäoline, et nii mõnedki neist pakkumistest lähevad täppi ja videos väljendatud emotsiooniks oli tõepoolest rõõm.

Teine hüpotees – kõrge ekstravertsusega modellide väljendatud emotsioone tuntakse suuremal määral õigesti ära kui madala ekstravertsusega modellide omi – ei saanud andmeanalüüsi põhjal kinnitust. Programm ei suutnud edukamalt tuvastada kõrgema ekstravertsuse skooriga videomodellide väljendatud emotsioone ning ekstravertsus ei osutunud logistilise regressiooni mudelis statistiliselt oluliseks. Saadud tulemus ei ole kooskõlas varasema kirjandusega, mis on näidanud ekstravertsuse seost ekspressiivsemate ja ilmekamate emotsioonide väljendamisega. Näiteks on leitud, et ekstravertsemate inimeste emotsioone on lihtsam tuvastada (Kring jt, 1994; Riggio & Riggio, 2002) ning et kõrgema ekstravertsusega videomodellide väljendatud emotsioone suudetakse paremini ära tunda (Rebane, 2023; Veltmann, 2024). Seega võib erinevus antud uuringu tulemustega viidata erinevusele inimeste ja

masinõppe algoritmide emotsioonide tuvastamise viisides ehk et ekstraversus ja teised isiksuseomadused võivad mõjutada inimeste ja tehisintellekti võimekust emotsioonide tuvastamisel erinevalt.

Tulemuste analüüsimisel tuleks arvestada baasemotsioonide kategooriatega, mida osalejad ja algoritm tuvastasid. Rõõmu edukam tuvastamine võib tuleneda sellest, et sellega kaasnevad universaalsed väljendused (nt naeratamine), mida on raske teisi emotsioone signaliseerivate väljendustega segi ajada. Hirmu ja vastikuse väljendamine on aga sageli vähem ekspressiivne, mistõttu võib nende tuvastamine algoritmide jaoks keerulisem olla. Glasgow Ülikooli teadlased (Jack jt, 2014) on pakkunud välja, et baasemotsioonide kategoriseerimine võiks piirduda nelja suurema kategooriaga: rõõm, kurbus, viha/vastikus (kombineeritult) ja hirm/üllatus (kombineeritult). Seda seetõttu, et rõõmu ja kurbust signaliseerivad näoväljendused on selgelt eristatavad, samas kui teised emotsioonid jagavad signaliseerimise algfaasis sarnaseid tunnuseid – näiteks hirmu ja üllatuse puhul suured pärani silmad ning viha ja vastikuse puhul kirtsus nina. Sellised tulemused viitavad, et rõõm on üks eristuvamaid ja kergemini tuvastatavaid emotsioone ning seda ei aeta hõlpsasti segamini teiste emotsioonidega. Seevastu hirm, vastikus ja viha on omavahel sarnasemad ning nende omavaheline eristamine on keerulisem. Ka käesoleva uuringu tulemused näitavad, et FaceReader ajas sageli vastikuse segamini vihaga ning hirmu segamini üllatusega (vt Tabel 1).

Kuigi Nolduse enda andmeil on FaceReader emotsioonide tuvastamisel peaaegu eksimatu ning tuvastavat valideerimisel kõiki baasemotsioone peale kurbuse 100% täpsusega (kurbust 95,8% täpsusega; Loijens & Krips, n.d.), sarnanevad käesoleva uurimistöö tulemused pigem varasematele sõltumatutele uuringutele. 2020. aastal läbi viidud uuringus, kus võrreldi mitmeid emotsioonide klassifitseerimiseks loodud programme, saavutas FaceReader täpsusskooriks 57,3%, olles kaheksast testitud programmist edukuselt teine (Dupré jt, 2020). Samas uuringus ületasid inimhindajad FaceReaderi märkimisväärselt, saavutades keskmiseks täpsusskooriks 72,5% (Dupré jt, 2020). Ka käesoleva uurimistöö tulemused näitavad, et inimhindajad olid FaceReaderist märkimisväärselt edukamad. Rebase (2023) uurimuses tuvastasid inimhindajad siinse tööga samades stiimulvideotes rõõmu 68,7%, vastikust 53,7% ja hirmu 23,1% täpsusega. Keskmiseks emotsioonide äratundmise täpsuseks oli 48,5% (Rebane, 2023). Univeri (2022) uuringus (samad stiimulvideod) oli inimhindajate keskmine täpsus 43,9%. Kõige paremini tuvastati rõõmu (77,4%), vastikust (59,5%) ning kõige kehvemini hirmu (9,4%) (Univer, 2022).

Võrreldes FaceReaderi 37,4% keskmise täpsusega antud uurimistöös võib järeldada, et inimesed olid samade videoklippide hindamisel edukamad.

Huvitaval kombel viitas Dupré jt (2020) uuring sellele, et FaceReaderi täpsus oli inimhindajatega võrreldav näideldud emotsioonide tuvastamisel, kuid programm jäi inimestele alla spontaanseid emotsioone sisaldavate klippide hindamisel. Käesolevas töös kasutatud stiimulmaterjal koosnes just näideldud emotsioonidest, mistõttu need tulemused ei ole Dupré jt (2020) uuringuga täielikult võrreldavad. Siiski on antud uuringu tulemused kooskõlas varasemate inimhindajate tulemustega emotsioonikategooriate osas: rõõmu tuvastatakse kõige edukamalt, vastikust mõnevõrra kehvemini ja hirmu väga kehvasti (Rebane, 2023; Univer, 2022). Leidub aga ka uuringuid, mis viitavad teistsugustele tulemustele. Näiteks Bassili (1978) leidis, et inimhindajad tuvastasid näitlejatest tehtud videoklippides rõõmu hoopis kõige kehvemini (31%), vastikust kõige edukamalt (88%) ja hirmu keskmise täpsusega (69%). Sellised tulemused näitavad, et emotsioonide tuvastamine sõltub oluliselt nii stiimulmaterjalist kui ka kontekstist, mistõttu on vajalik emotsioonide äratundmise uuringute tulemuste tõlgendamisel alati arvesse võtta nii kasutatud meetodeid kui ka kasutatavate stiimulite omadusi.

Antud uurimistöö tulemused pakuvad väärtuslikku teavet FaceReader programmi täpsuse kohta emotsioonide tuvastamisel ning aitavad paremini mõista programmi tugevusi ja piiranguid. Kuigi programm suutis edukalt tuvastada rõõmu, näitab hirmu ja vastikuse tuvastamise madal täpsus, et tarkvaral on veel märkimisväärselt arenguruumi. See teadmine võib olla kasulik tarkvaraarendajatele, kes tegelevad emotsioonituvastusalgoritmide täiustamisega.

Uurimistöö ei leidnud seost ekstravertsuse, emotsioonide selgema väljendamisega ja parema tuvastamise vahel – erinevalt varasematest uuringutest, mis on seda seost inimhindajate puhul näidanud. See erinevus võib viidata vajadusele edasisteks uuringuteks, mis keskenduvad emotsioonituvastusprogrammide tõhususe ning analüüsitavate mudelite isiksuseomaduste vahelistele seostele. Varasemad uuringud on valdavalt piirdunud inimhindajatega, kuid käesolev töö tõstatab küsimuse, kas isiksuseomadused nagu ekstravertsus, mõjutavad tehisintellekti sooritust samal viisil kui inimeste puhul või hoopis erinevalt – või ei avalda need tehisintellekti sooritusele üldse olulist mõju.

Uurimuse piirangud ja edasised uurimisuunad

Antud uuringu üheks piiranguks võib pidada selle väikest valimit, mis seab piirangud saadud tulemuste üldistamisele laiemale populatsioonile. Nii videomodellide kui emotsioonide

hindajate keskmine vanus oli suhteliselt noor (vastavalt ligikaudu 25 eluaastat ja ligikaudu 31 eluaastat), seega ei pruugi tulemused olla üldistatavad ka vanematele generatsioonidele.

Teiseks piiranguks võib pidada FaceReaderile analüüsimiseks esitatud videoklippide väga lühikest kestust – kõik klipid olid vaid 1 sekund pikad. Seda tehti tulemuste paremaks kõrvutamiseks Rebase (2023) kogutud andmetega, kus ka inimhindajatele esitati 1-sekundilisi videoklippe. On aga võimalik, et 1 sekund on liiga lühike aeg, et algoritm suudaks edukalt emotsioone tuvastada. Näiteks Dupré jt (2020) uuringus, kus kasutati 4-sekundilisi (näideldud emotsioonid) ja 6-sekundilisi (spontaansed emotsioonid) klippe, oli FaceReader emotsioonide tuvastamisel oluliselt edukam kui käesolevas uuringus.

Lisaks võib piiranguna käsitleda fakti, et uuringus kasutati ainult stiimulvideoid. On võimalik, et FaceReader on emotsioonide tuvastamisel edukam staatilistelt piltidelt. Näiteks Lewinski jt (2014) uuringus, kus stiimulmaterjaliks olid pildid, saavutas FaceReader tunduvalt kõrgema täpsusskoori – ühe pildikogumiku puhul 88% ja teise puhul 89%. Need tulemused on oluliselt paremad kui käesoleva töö tulemused. Samas võib stiimulvideote kasutamist käsitleda ka tugevusena, kuna videod peegeldavad paremini olukorda, millega üldiselt elus kokku puutuda tuleb (reaalses suhtlussituatsioonis ei suhtle me ju teiste inimeste kui staatiliste piltidega), ning seega on videote kasutamisel väline valiidsus kõrgem.

Kui varasemalt on leitud seoseid videomodellide ekstraversuse ja nende väljendatud emotsioonide edukama tuvastamise vahel, siis käesolevast uuringust sellist seost ei leitud – läbiviidud analüüsides ei olnud osaleja ekstraversuse skooril statistiliselt olulist mõju FaceReaderi poolt emotsioonide äratundmisele. Arvestades antud uuringu piiranguid, on siiski oluline teostada edasisi uuringuid suuremate valimitega. Need võiksid pakkuda täpsemaid teadmisi selle kohta, kas tehisintellekti emotsioonide äratundmise puhul ilmnevad sarnased seosed väljenduse arusaadavuse ja emotsioone väljendava indiviidi isiksuseomaduste vahel (seosed, mida on täheldatud inimhindajate puhul). Sellised uuringud aitaksid süvendada meie arusaama tehisintellekti ja inimeste erinevatest infotöötlusprotsessidest emotsioonide tuvastamisel.

Edasistesse uuringutesse võiks kindlasti kaasata rohkem vanemaid inimesi, et valimid ei oleks nii noortekesksed ja saadud tulemused oleksid seega paremini üldistatavad ka vanematele põlvkondadele. Huvitav oleks uurida ka võimalikke põlvkondadevahelisi erinevusi, näiteks kas masinõppealgoritmide puhul esineb erinevusi noorema- ja vanemaealiste indiviidide

emotsioonide tuvastamises, kuna senised uuringud on läbi viidud valdavalt noortel valimitel. Näiteks antud töös viidatud Lewinski ja kolleegide uuringus pärinesid FaceReaderile analüüsimiseks antud videoklipid kahest andmebaasist, kus ühes oli modellide vanuseks 20-30 eluaastat (Olszanowski jt, 2015) ja teises 18-25 eluaastat (van der Schalk jt, 2011). Ka Dupré ja kolleegide uuringus kasutatud videoklipid pärinesid kahest noorel valimil põhinevast andmebaasist, kus üks andmebaas sisaldas videoklippe 18-45-aastastest modellidest ja teine andmebaas 18-25-aastastest modellidest (Dupré jt, 2020). See näitab, et FaceReaderi võimekust on seni testitud peamiselt noortel valimitel, mistõttu on vähem andmeid selle kohta, kuidas programm tuvastab emotsioone vanemate inimeste puhul. Vanemate inimeste näod on sageli kortsulisemad ja nende miimika võib erineda noorte omast, seega on oluline uurida, kas FaceReader ja sarnased algoritmid suudavad nende erisustega kohaneda. Spetsiaalselt treenitud andmestikud vanemaealiste emotsioonide tuvastamiseks võiksid aidata parandada algoritmide täpsust ja rakendatavust erinevates vanuserühmades.

Nagu sissejuhatuses mainitud, on emotsioonituvastusprogrammid mitmetes valdkondades kasutusel juba praegu ning neile nähakse veel mitmeid potentsiaalseid rakendusvõimalusi tulevikus (Smith jt, 2024). Just seetõttu on oluline jätkata emotsioonide tuvastamisega tegelevate algoritmide süvitsi uurimist võimalikult heterogeensetes gruppides, mis hõlmavad erinevaid vanuserühmi ja etnilisi gruppe, kuna emotsioonituvastusprogrammide laialdase kasutuselevõtu korral peab tagama, et need oleksid võimalikult täpsed ja vigadeta. Vead ja kallutatud andmetes võivad viia ebakorrektsete ja ebaõiglaste otsuste ja järeldusteni, mis võivad omakorda mõjutada näiteks värbamisprotsesse, patsientide ravi või turundusstrateegiaid, aga ka mitmeid teisi valdkondi (Smith jt, 2024). Vaid läbi põhjaliku ja pideva uurimise saab välja töötada usaldusväärseid ja täpseid emotsioonituvastuslahendusi, mis saavad pakkuda tegelikku väärtust erinevates kasutusvaldkondades.

Kokkuvõte

Käesolev uurimistöö pakub väärtuslikku teavet FaceReader programmi täpsuse kohta emotsioonide tuvastamisel, leides, et programm suutis usaldusväärselt tuvastada vaid rõõmu, samas kui vastikuse ja hirmu tuvastamisel oli täpsus madal. Lisaks ei ilmnenud statistiliselt olulist seost videomodellide ekstravertsuse skoori ja FaceReaderi poolt ennustatavate emotsioonikategooriate täpsuse vahel. Sellised tulemused võivad viidata sellele, et kuigi kõrgema ekstravertsusega inimeste emotsioone võib olla lihtsam tuvastada inimehindajate jaoks,

ei pruugi see sama kehtida tehisintellekti ja masinõppe algoritmide puhul. See on huvitav suund, mida tulevikus edasi uurida võiks.

Uurimistöö piiranguteks olid väike valim ja lühikesed videoklipid, mis võivad mõjutada tulemuste üldistatavust ja programmi täpsust. Seega oleks edasistes uuringutes soovitatav kaasata suuremaid ja mitmekesisemaid valimeid ning kasutada analüüsimisel pikemaid videoklippe, et tagada täpsemad ja usaldusväärsemad tulemused, või proovida kasutada hoopis stiimulpilte. Samuti oleks huvitav uurida, kuidas FaceReader ja sarnased algoritmid suudavad kohanduda erinevate vanuserühmadega, et parandada emotsioonituvastuse täpsust laiemas kontekstis.

Kokkuvõttes viitavad antud töö tulemused vajadusele täiustada emotsioonituvastusalgoritme, eriti selliste emotsioonide nagu hirm ja vastikus tuvastamisel. Tulemused on olulised, kuna need annavad teavet masinõppeprogrammide ja isiksuseomaduste vahelistest seostest (või nende puudumisest) ning aitavad suunata edasisi uurimusi üha rohkem kasutust leidvate emotsioonituvastusalgoritmide arendamisel ja täiustamisel.

Kasutatud kirjandus

- Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 4(3), 373–379. <https://doi.org/10.1037/0096-1523.4.3.373>
- Calvo, M. G., & Nummenmaa, L. (2015). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, 30(6), 1081–1106. <https://doi.org/10.1080/02699931.2015.1049124>
- Caridakis, G. et al. (2007). Multimodal emotion recognition from expressive faces, body gestures and speech. In: Boukis, C., Pnevmatikakis, A., Polymenakos, L. (eds) Artificial Intelligence and Innovations 2007: from Theory to Applications. AIAI 2007. *IFIP The International Federation for Information Processing*, vol 247. https://doi.org/10.1007/978-0-387-74161-1_41
- Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLOS ONE*, 15(4). doi:10.1371/journal.pone.0231968
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Friedman, H. S., Riggio, R. E., & Segall, D. O. (1980). Personality and the enactment of emotion. *Journal of Nonverbal Behavior*, 5(1), 35–48. doi:10.1007/bf00987053
- Garcia-Garcia, J. M., Penichet, V. M. R., & Lozano, M. D. (2017). Emotion detection. Proceedings of the XVIII International Conference on Human Computer Interaction - Interacción '17. <https://doi.org/10.1145/3123818.3123852>
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, 19(18), 1543–1548. <https://doi.org/10.1016/j.cub.2009.07.051>
- Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, 24(2), 187–192. <https://doi.org/10.1016/j.cub.2013.11.064>
- Jack, R. E., Garrod, O. G. B., Yu, H., & Caldara, R. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, 109(19), 7241–7244. <https://doi.org/10.1073/pnas.1200155109>

- Juuse, L. (2020). *The relationship between visual and verbal information and emotion processing: reports from EEG and self-assessment*. Magistritöö: Tartu Ülikool. Psühholoogia instituut. <http://hdl.handle.net/10062/75616>
- Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology*, 66(5), 934–949. <https://doi.org/10.1037/0022-3514.66.5.934>
- Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7(4), 227–236. <https://doi.org/10.1037/npe0000028>
- Littlewort, G. C., Bartlett, M. S., & Lee, K. (2007). Faces of pain. *Proceedings of the Ninth International Conference on Multimodal Interfaces – ICMI '07*. <https://doi.org/10.1145/1322192.1322198>
- Loijens, L., & Krips, O. (n.d.). FaceReader methodology. Noldus Information Technology. Vaadatud 10.06.2024: <https://info.noldus.com/hubfs/resources/noldus-white-paper-facereader-methodology.pdf>
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, 63(6), 503–517. <https://doi.org/10.1037/0003-066X.63.6.503>
- Murray, H. (Ed.) *Explorations in Personality*. New York: Oxford, 1938.
- Möttus, R., Pullmann, H., & Allik, J. (2006). Toward more readable big five personality inventories. *European Journal of Psychological Assessment*, 22(3), 149–157. <https://doi.org/10.1027/1015-5759.22.3.149>
- Noldus Information Technology (2021). FaceReader 9.0 [arvutitarkvara]. Vaadatud <https://www.noldus.com/facereader>
- Noldus Information Technology. (n.d.). FaceReader methodology. Vaadatud 10.06.2024: https://delarosaresearch.com/uploads/Nd_White_paper_FaceReader_Methodology_screen.pdf
- Noldus Information Technology (2025). FaceReader: Applications. Vaadatud 05.01.2025: <https://www.noldus.com/facereader/applications>

- Olszanowski, M., Pochwatko, G., Kuklinski, K., Scibor-Rylski, M., Lewinski, P., & Ohme, R. K. (2015). Warsaw set of emotional facial expression pictures: A validation study of facial display photographs. *Frontiers in Psychology*, 5(2014).
<https://doi.org/10.3389/fpsyg.2014.01516>
- Pantic, M., & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1424–1445. <https://doi.org/10.1109/34.895976>
- Pollak, S. D., Camras, L. A., & Cole, P. M. (2019). Progress in understanding the emergence of human emotion. *Developmental Psychology*, 55(9), 1801–1811.
<https://doi.org/10.1037/dev0000789>
- Rebane, E. M. (2023). Suurel viisikul mõõdetud ekstravertsuse seosed emotsioonide väljendamisega. Uurimistöö: Tartu Ülikool. <https://hdl.handle.net/10062/90925>
- Riggio, H. R., & Riggio, R. E. (2002). Emotional expressiveness, extraversion, and neuroticism: A meta-analysis. *Journal of Nonverbal Behavior*, 26, 195–218.
<https://doi.org/10.1023/A:1022117500440>
- Smith, K., Khare, A., Blanes-Vidal, V., Nadimi, E. S., & Acharya, U. R. (2024). Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information Fusion*, 102, 102019. <https://doi.org/10.1016/j.inffus.2023.102019>
- Tucker, J. S., & Friedman, H. S. (1993). Sex differences in nonverbal expressiveness: Emotional expression, personality, and impressions. *Journal of Nonverbal Behavior*, 17(2), 103–117. <https://doi.org/10.1007/bf01001959>
- Univer, H. (2022). Masinõppe algoritmi võimekus tuvastada emotsioone näoväljenduste põhjal – kas parem või halvem kui inimene? Magistritöö: Tartu Ülikool.
<https://hdl.handle.net/10062/90965>
- van der Schalk, J., Hawk, S., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, 11(4), 907–920. <https://doi.org/10.1037/a0023853>
- Veltmann, E. (2024). Positiivse ja negatiivse afekti ning ekstravertsuse ja neurootilisuse seos emotsionaalsete näoväljendustega. Uurimistöö: Tartu Ülikool.
<https://hdl.handle.net/10062/95331>

Viisileht, K. (2025). Emotsionaalsete näoväljenduste äratundmise võrdlus inimhindajate ja FaceReader masinõppe tarkvara vahel. Uurimistöö: Tartu Ülikool

Käesolevaga kinnitan, et olen korrektselt viidanud kõigile oma töös kasutatud teiste autorite poolt loodud kirjalikele töödele, lausetele, mõtetele, ideedele või andmetele.

Olen nõus oma töö avaldamisega Tartu Ülikooli digitaalarhiivis DSpace.

/Nele-Liis Merilo/

Töö valmimist toetas Eesti Teadusagentuuri grant PRG1151 “Tähelepanueelne informatsioonitöötlus ajus: seosed seisundite, püsitunnuste ja käitumisega” (01.01.2021-31.12.2025).