

Tartu Ülikool
Sotsiaalteaduste valdkond
Psühholoogia instituut

Kaisa Viisileht

**Emotsionaalsete näoväljenduste äratundmise võrdlus inimhindajate ja FaceReader
masinõppe tarkvara vahel**

Uurimistöö

Juhendajad: Liina Juuse, MA

Kairi Kreegipuu, PhD

Läbiv pealkiri: Inim- ja masinhinnangud emotsioonide äratundmisel

Tartu 2025

Emotsionaalsete näoväljenduste äratundmise võrdlus inimhindajate ja FaceReader masinõppe tarkvara vahel

Lühikokkuvõte

Antud töö eesmärgiks oli uurida, kui täpselt suudab masinõppe tarkvara FaceReader tuvastada Ekmani baasemotsioone ning võrrelda selle tuvastamistäpsust inimhindajate tulemustega. Uuringus osales 144 inimhindajat, kellest olid mehed 19,4% ja 79,9% naised. Inimhindajatele ja FaceReaderile esitati 1-sekundilisi videoklippe, mis kujutasid rõõmu, hirmu ning vastikust, ning eesmärk oli stiimulmaterjali hinnata kuue baasemotsiooni (rõõm, hirm, vastikus, üllatus, viha ja kurbus) põhjal. Tulemused näitasid, et inimesed vastasid õigesti 54,4% juhtudest, samas kui masinõppe täpsus oli 37,4%. Korduvmõõtmiste ANOVA analüüs näitas, et inimhindajad suutsid edukalt ära tunda rõõmu ja vastikust, kuid mitte hirmu. FaceReader tuvastas edukalt rõõmu, kuid vastikust ja hirmu mitte. Sõltumatute rühmade ANOVA tulemuste kohaselt olid inimhindajad FaceReaderist statistiliselt olulisel määral täpsemad hirmu ja vastikuse, kuid mitte rõõmu tuvastamisel. Käesoleva töö tulemused on kooskõlas varasemate uurimustega, viidates, et mittevalideeritud ning dünaamilise stiimulmaterjali hindamisel on inimesed masinõpetest sageli täpsemad.

Märksõnad: masinõppe, FaceReader, emotsioonituvastus, näoväljendus, inimhindajad

Comparison of emotional facial expression recognition between humans and FaceReader machine learning software

Abstract

The aim of this work was to find out how accurately the machine learning software FaceReader can detect Ekman's basic emotions and to compare FaceReader with human evaluations in recognizing basic emotions. A total of 144 human evaluators, of whom 19.4% were men and 79.9% women. They were presented with 1-second video clips of happiness, fear and disgust, to which they had to find the answer among the 6 basic emotions. Results showed that humans (54.4%) gave more correct answers than machine learning (37.4%). A repeated-measures ANOVA revealed that people correctly identified happiness and disgust, but not fear. Machine learning correctly recognized happiness, but not disgust or fear. According to an ANOVA with independent groups, human and FaceReader responses differed significantly for fear and disgust, but not for happiness. This work is consistent with previous research and can suggest that humans are more likely to be more accurate than machine learning when evaluating non-validated and dynamic stimulus material.

Keywords: machine learning, FaceReader, emotion recognition, facial expression, human evaluators

Emotsioonid mängivad meie elus olulist ja suunavat rolli – ühelt poolt toetavad need ellujäämist, teiselt poolt annavad elule värviküllust. Neid saab väljendada ja ära tunda erinevate meediumite kaudu (näiteks näoilmed, hääletoon, žestid, kehahoiakud; Keltner jt, 2019b). Käesolev töö keskendub emotsioonide edastamisele näoilmete kaudu.

Emotsioonide äratundmise olulisust rõhutatakse peamiselt suhtlussituatsioonides, kus need toimivad sotsiaalse navigatsiooni vahendina. Samuti on emotsioonid tihti liikumapanevaks jõuks, aidates tegutseda ja olukordi kujundada (Frijda, 1986; Frijda jt, 1989). Nii on näiteks on hirmu väljenduse eesmärgiks hoiatada ümbritsevaid võimaliku ohu eest. Kuigi hirmu kogemine kutsub tajujas esile vältimiskäitumist, tekitab see teistes pigem lähenemiskäitumist, mis suunab koostööle ohtlikus olukorras (Marsh jt, 2005). Samas pole kõigi emotsioonide näoväljenduste eesmärkide osas veel konsensust saavutatud – näiteks üllatuse väljenduse funktsioon on teaduslikult jätkuvalt arutluse all (Shariff & Tracy, 2011). Edukaks suhtlemiseks ja sotsiaalseks toimetulekuks on kokkuvõttes hädavajalik oskus interpreteerida teiste näoilmeid ja mõista nende emotsionaalset tähendust.

Tänapäeva ühiskonnas on emotsioonide mõistmisele lisandunud inimeste kõrval ka tehisintellekti poolt kasutatavad vahendid. Need mängivad üha suuremat rolli ning seega on oluline uurida, kui hästi masinõppe meetodid suudavad emotsioone tuvastada ja tõlgendada. Mitmed suurfirmad, nagu Google, Meta Platforms ja Microsoft, kasutavad tehisintellekti mitmekülgsetes rakendustes, sealhulgas tarkades kodudes, intelligentsetes vestlussüsteemides ja tervishoius (Zhang et al., 2020). Emotsioonide äratundmise potentsiaal ulatub ka vaimse tervise valdkonda, kus sügavõppel (*deep learning*) põhinevaid mudeleid nähakse lootustandva vahendina näiteks bipolaarse häire diagnoosimisel, raviplaani koostamisel ja ravi kulgemise ennustamisel (Librenza-Garcia et al., 2017). Browning jt (2019) tõestasid, et masinõppe suudab emotsionaalsete kallete ja tajutud sümptomite põhjal antidepressantide ravi tulemuslikkust prognoosida oluliselt kiiremini, aidates patsientidel vältida ooteaega ning valida enne ravimite toimimist sobivaim sekkumine. Selliste edasiarengute tõttu on vajalik uurida, kui usaldusväärset ja täpselt tehisintellekt suudab emotsioone tuvastada ning kas selle tulemused on võrreldavad või isegi paremad inimeste hindamistäpsusest. See aitab mõista, kuidas masinõppe saab täiendavalt toetada emotsioonide äratundmist nii teadusuuringutes kui ka rakenduslikult.

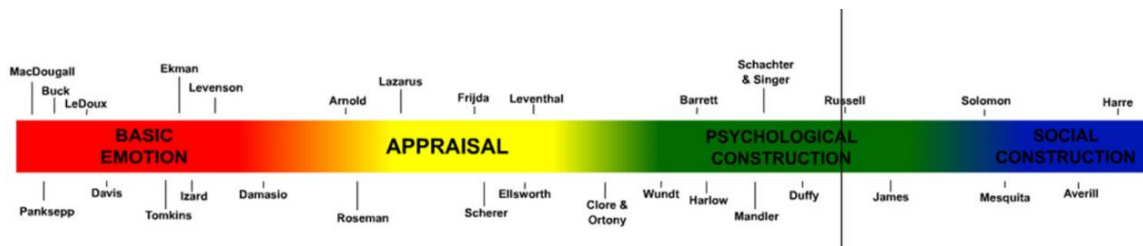
Emotsiooniteooriad

Emotsioon on psühhofüsioloogiline nähtus, mille ühtses definitsioonis pole teadlased senini kokku leppinud (Zhang et al., 2020). Siiski ollakse samal meelel, et emotsioon on psühholoogiliste seisundite kogum ning selle kogemisele on omased subjektiivne kogemus, väljenduslik käitumine ja muutused füsioloogias (Gross & Barrett, 2011).

Sarnaselt on ka emotsioone käsitlevaid teooriaid sadu (Strongman, 2003), mille hulgast konkureerivad enim baasemotsioonide ja konstruktivistlikud teooriad. Gross ja Barrett (2011) on emotsiooniteooriad jaganud ka nelja võimalikku klastrisse, kus teooriaid on võimalik kujutada kontiinumil ning vasakult paremale liikudes kasvab kultuuri ja kognitsioonide mõju (joonis 1). Tõenäoliselt pole ükski emotsiooniteooria kõige tõesem ega välista teiste õigsust, vaid need toimivad üksteise puudujääke täiendades (Scherer, 2000).

Joonis 1

Nelja võimaliku emotsiooniteooria klastri kontiinum (Gross & Barrett, 2011).



Märkus. Vasakult paremale: *basic emotion* (baasemotsioonid), *appraisal* (hinnangu teooria), *psychological construction* (psühholoogiline konstruktivism) ning *social construction* (sotsiaalne konstruktivism).

Baasemotsioonide teooria, millest lähtun käesolevas töös, on olnud emotsioonide uurimise valdkonnas väga mõjukas (Keltner et al., 2019a). Selle kohaselt on olemas kindel hulk emotsioone, mis selgelt erinevad üksteisest olemuse ja funktsiooni poolest. Iga baasemotsioon saab tekkida ajus kindlast närvingest, mis käivitab eelnevalt mainitud emotsiooni kogemisele omased tunnused nagu näiteks väljenduslik käitumine või füsioloogilised muutused (Gross & Barrett, 2011). Kuigi erinevad baasemotsioonide teooriad rõhutavad erinevaid kirteeriume defineerimaks baasemotsiooni mõistet (Ortony, 2021), on mitteverbaalselt emotsioonide väljendamisel keskeltläbi järgmised viis omadust (Keltner et al., 2019b). (1) Emotsioonid on lühiajalised ja igale subjektiivsele kogemusele vastab teatud äratuntav käitumismuster; (2) emotsioonid annavad aimu emotsionaalsest hetkeseisundist, kavatsustest ja motivatsioonidest ning

olukorrahinnangust; (3) emotsioonide esile kutsumine ja äratundmine sarnaneb kultuuriti; (4) emotsioonid on evolutsiooniliselt ühised imetajate seas; (5) emotsioonidele järgnevad kindlad füsioloogilised muutused. Lisaks loetakse baasemotsioonide omaduseks niinimetatud baasilisust ehk neid pole võimalik jagada eraldi emotsioonideks, vaid nendest moodustuvad kognitiivsete funktsioonidega koostöös kompleksemotsioonid (Celeghin et al., 2017).

Ekmani teooria ei ole ainus käsitlus baasemotsioonidest (Arnold, 1960; Tomkins, 1962; Frijda, 1986; Panksepp & Watt, 2011; Plutchik, 1960). Tegelikult on erinevate baasemotsioonide kombinatsioonide põhjal oletatud, et universaalset ja kindlat hulka baasilisi emotsioone ei pruugi olemas olla (Ortony, 2021). Lisaks on kriitikat pälvinud asjaolu, et need teooriad ei arvesta piisavalt igapäevaelu keerukamat emotsioonide väljendamist ja infotöötlust äratundmisel (Keltner jt, 2019b). Kuigi baasemotsioonidega seostatakse kindlaid näoilmeid, ilmneb märkimisväärne varieeruvus nii näoväljendustes endis kui ka nende tähenduses, mis sõltub kontekstist (Aviezer jt, 2008; Ortony, 2021).

Emotsioonitöötlus

Inimese emotsioonitaju

Inimeste võimekus näoilmetest baasemotsioone tuvastada varieerub mitmete tingimuste alusel. Negatiivse valentsiga emotsioonide (näiteks hirm, viha ja vastikus) eristamine on keerulisem ning neid kiputakse tihti omavahel segi ajama (Dores jt, 2020). Näiteks hirmu aetakse sageli segamini üllatusega (Zhao jt, 2017), samas kui viha ja vastikuse sarnane näoilme põhjustab nende omavahelise segiajamise (Hendel jt, 2023). Samuti on leitud, et hirmu ja vastikuse tuvastamine on lihtsam, kui näoilmeid väljendab samasse etnilisse gruppi kuuluv isik (Elfenbein & Ambady, 2002). Positiivse valentsiga emotsioonid, nagu rõõm, on kergemini äratuntavad kui negatiivse valentsiga (De Araujo jt, 2022; Skiendziel jt, 2019; Ivanova & Borzunov, 2020; Lévêque jt, 2022; Lewinski jt, 2014; Dores jt, 2020; Elfenbein & Ambady, 2002).

Erinevused võivad ilmned ka staatilise ja dünaamilise stiimulmaterjali kasutamisel. Emotsioonide kogemine ja töötlemine jaguneb üldiselt kaheks peamiseks protsessiks: alt-üles ja ülevalt-alla töötlus (Adolphs, 2002). Alt-üles töötlus keskendub stiimulite omaduste eristamisele näojoonte põhjal, pakkudes esmast sensoorse informatsiooni analüüsi. Ülevalt-alla töötlus seevastu hõlmab mälust pärinevate varasemate teadmiste kaasamist, hinnangute andmist ja stiimulite väärtustamist.

Igapäevaelus toimivad emotsionaalsed näoväljendused dünaamiliste signaalidena, mis koosnevad väikestest näolihaste liigutustest (Kilts jt, 2003). Erinevate emotsioonide töötlemine võib aktiveerida erinevaid ajupiirkondi – näiteks on vastikusega seotud insula ja basaalganglionid (Holtmann jt, 2020). Samuti aktiveeruvad staatiliste ja dünaamiliste näoväljenduste töötlemise ajal ajus mõnevõrra erinevad piirkonnad (Kilts jt, 2003). See tähendab, et on võimalik, et inimestel on videomaterjali puhul emotsioone märksa lihtsam ära tunda, kuna liikumine pakub rohkem informatsiooni ja on inimestele loomulikum tajumisprotsess (Krumhuber jt, 2013). Lewinski jt (2014) leidsid, et pildi- ja videomaterjalidel põhineva emotsioonide äratundmise täpsuses oli vaid väike erinevus: vastavalt 82% ja 87%. Siiski võiksid kehvema kvaliteediga materjali puhul olla erinevused videote kasuks suuremad (Krumhuber jt, 2013). Staatiliste ja dünaamiliste stiimulite võrdlemine, eriti masinõppe ja inimeste hinnangute osas, vajab seega täiendavat uurimist.

Masinõppe emotsioonide äratundjana

Masinõppe meetodeid emotsioonide tuvastamisel saab jagada kahte rühma: traditsioonilised või sügavõppel põhinevad lähenemised (Ko, 2018). Traditsioonilised meetodid kasutavad käsitsi loodud funktsioone ning tuvastavad esmalt näo, eraldavad tunnused, misjärel liigitavad näoilme. Siia kuuluvad näiteks tugivektormasin (*support vector machine* ehk SVM), AdaBoost, Random Forest jne. Sügavõppe meetodid on aga muutunud üha levinumaks meetodiks emotsioonide tuvastamisel ning emotsioonide äratundmise uurimine on oluliselt arenenud sügavõppe kasutamisega (De Araujo et al., 2022), need põhinevad sügavatel närvivõrkudel (*Deep Neural Network* ehk DNN) ning vajavad usaldusväärseks tuvastamiseks suuremat hulka treenimismaterjali (Saganowski, 2022). Siia alla käivad näiteks *recurrent neural networks* (RNNs), *convolutional neural networks* (CNNs), Residual Networks, StresNet, *Long Short-Term Memory* (LSTM) jne. Sügavad närvivõrgud jäljendavad bioloogilisi neuroneid ning sünapside kaudu kas tugevdatakse või nõrgendatakse signaalne, misjärel läheb info edasi neuronisse (Sze, 2017). Sügavad närvivõrgud suudavad erinevalt traditsioonilistest meetoditest õppida ilma juhendamiseta. Sügavõppel põhinevad programmid on emotsioonide tuvastamisel sageli täpsemad kui traditsioonilised (Ko, 2018). Kuivõrd emotsioone tuvastavate algoritmide loomine ja treenimine on väga ajakulukas ning vaearikas protsess, on emotsioonide uurimisel üha populaarsemaks muutunud ka kommertstarkvarade rakendamine (De Araujo et al., 2022).

Masinõppe algoritmide seas on emotsioonide tuvastamismäär erinevates uurimustes tugevalt varieerunud, näiteks 26%-97% (Dupré et al., 2020; Küntzler et al., 2021; Lévêque et al., 2022). Sarnaselt inimestele esineb masinõpetel enim raskusi negatiivse valentsiga emotsioonide äratundmisel. Keeruline on tuvastada kurbust (Ivanova & Borzunov, 2020). Masinõpped ajavad sageli segi viha kurbuse ja hirmuga ning vastikust vihaga (De Araujo et al., 2022). Viha ja kurbust on loetud ka neutraalsuseks (Küntzler et al., 2021). Neutraalsust on pakutud vastuseks juhtudel, kui näoväljendust ei suudeta klassifitseerida või pilt on tehtud loomulikus keskkonnas (Lévêque et al., 2022; Küntzler et al., 2021). See-eest kõrgete tuvastamismäärade põhjal võib väita, et kõige kergemini äratuntav emotsioon kaldub olema samuti sarnaselt inimestele rõõm (De Araujo et al., 2022; Skiendziel et al., 2019; Ivanova & Borzunov, 2020; Lévêque et al., 2022; Lewinski et al., 2014; Dores et al., 2020; Elfenbein & Ambady, 2002). Samal teemal on Eestis varasemalt kirjutanud oma magistritöö Hans Univer (2022), kuid sealne masinõppe algoritm usaldusväärselt emotsioone tuvastada ei suutnud tõenäoliselt väikse hulga treenimismaterjali tõttu. Seepärast polnud võimalik teha järeldusi masinõppe suutlikkuse kohta.

Kumb on emotsioonide tuvastamises parem – masinõpe või inimene?

Selle kohta, kumb hindaja kokkuvõttes emotsioonide äratundmises võimekam on, leidub erinevaid tulemusi. Mitmetes võrdlevates uurimustes ei jää inimeste ja masina äraarvamise protsent üksteisest ka oluliselt kaugele (Küntzler et al., 2021). Näiteks Lewinski ja kolleegide (2014) töös saavutati 89% (masinõpe) ja 87% videoklippide ning 88% (masinõpe) ja 82% piltmaterjalide puhul. Samas Dupré jt (2020) leidsid tuvastusmäärad kaheksa erineva masinõppe keskmisena 54% ning inimeste puhul 72%. Õigeid vastuseid võivad muuhulgas mõjutada pildi selgus, kvaliteet, näoilme intensiivsus (De Araujo et al., 2022), poseerimise olukord kui ka algoritmi omapära (Lévêque et al., 2022; Saganowski, 2022). Loomulikke emotsionaalseid näoväljendusi tuvastavad kehvemini ka inimesed (Lévêque et al., 2022; Goodfellow et al., 2015). Sellegipoolest, kui pildid on kehvema kvaliteediga või tehtud spontaanselt (ebaselgelt poseeritud ja loomulikumas keskkonnas), on inimesed emotsioonide tuvastamisel tunduvalt paremad (Lévêque et al., 2022; Yitzhak et al., 2017). Märkimist vajab ka asjaolu, et paljud masinõpped on emotsioonide tuvastamiseks treenitud standardiseeritud pilt- ja videomaterjalidel, mida kasutavad stiimulitena ka paljud uurimused. Stöckli ja kolleegide (2017) sõnul võib see selgitada kõrgeid tuvastusmäärasid mitmetes töödes.

Masinõppe ja inimeste hinnangute võrdlemine võimaldab heita valgust ka mõlema infotöötlusprotsessidele. Lévêque jt (2022) leidsid, et inimeste ja masinõppe vigade vahel on nõrk kuni mõõdukas positiivne korrelatsioon. Näiteks algoritmi DeepFace Pearsoni korrelatsioonikordaja oli 0,24 ja algoritmi RMN puhul 0,5.

FaceReader

Masinõppet esindab minu töös Hollandi firma Noldus kommertstarkvara programm FaceReader (versioon 9.0), mis suudab erinevalt paljudest teistest programmidest nii pildi- kui videomaterjalide abil inimeste nägudest tuvastada Ekmani baasemotsioone (rõõmu, kurbust, viha, üllatust, hirmu, vastikust) ja neutraalsust (Loijens & Krips, *s.a.*). Programm mõistab ka väiksemaid miimika ühikuid näos, valentsi, erutust, video-modellide isiksuseomadusi kui ka verevoolu intensiivsust pulsi muutumisel. FaceReaderi kasutamise eeliseks on ka selle pikaajaline treenimine, millega alustati juba 2005. aastal ning selle treenimiseks on kasutatud üle 20 000 pildi. FaceReader põhineb Deep Face'i meetodil ehk teeb emotsioone kindlaks tehisnärvivõrkude abil. Selleks et näoilmet klassifitseerida, peab programm esmalt leidma näo. Teisena toimub näo modelleerimine ja 468 punkti sünteesimine sügavate närvivõrkude põhjal, misjärel viiakse läbi peakomponentide analüüs.

Programm võib olenevalt video-modellist kalduda teatud emotsioone eelistama, seda mõju soovitatakse kalibreerimise teel vähendada (Loijens & Krips, *s.a.*). Tulemused esitatakse tabelis ning vastavalt iga emotsiooni olemasolule esitatakse arvvärtus vahemikus 0-1 (0 - ei tuvastatud üldse ja 1 - tuvastati täielikult), kuna emotsioonide väljendamine on sageli palju komplekssem, mistõttu võib väljenduses täheldada ka erinevaid emotsioone. Positiivse valentsiga emotsioon on ainult rõõm ning negatiivsega kurbus, viha, hirm. Valentsi tuvastamiseks lahutatakse rõõmu väärtusest negatiivse valentsiga emotsioonide väärtus. Erutust hinnatakse 20 *action unit*i ehk väiksemate näolihaste liigutuste abil FACS süsteemi (Coan, J. A., & Allen, 2007) alusel. Selle kohaselt on olemas 44 väiksemat näolihaste liigutust, mida kombineerides saame kokku panna kindlatele baasemotsioonidele omased näoliigutused (Ekman et al., 2002). Selle põhjal tehtud hinnangud esitatakse sarnaselt vahemikus 0-1. FaceReader ei erista teeseldud emotsiooni loomulikust. Kultuuriliste erinevuste mõjude vähendamiseks on programmi treenitud erinevast rassist modellidega.

FaceReaderi efektiivsust on hinnatud mitmes uurimuses, kus on kasutatud erinevaid programmi versioone. Lewinski jt (2014) leidsid, et FaceReaderi (versiooni 6.0) täpsus piltidelt emotsioonide tuvastamisel oli 88%, samas kui inimestel oli vastav näitaja

85%. Programmil esines enim raskusi viha tuvastamisega. Skiendziel jt (2019) uurimuses saavutas FaceReader versiooniga 7 kalibreerimata dünaamiliste näoilmete puhul keskmise tuvastustäpsuse 79%, kusjuures kõige keerulisem oli hirmu äratundmine. Seda aeti sageli segamini üllatusega. Rõõmu tuvastas programm seevastu õigesti 100%. Autorid märkisid, et FaceReader võib olla tundlikum videomaterjalide suhtes, mis võib osaliselt seletada selle tingimuse madalamat täpsust. Küntzler jt (2021) võrdlesid kolme emotsioonide tuvastamise programmi ning leidsid, et FaceReader (versioon 8.0) oli neist kõige täpsem – see tuvastas emotsioone 97% täpsusega. Siiski oli programmi nõrkuseks neutraalsete ja üllatuse näoilmete tuvastamine. Kõige vähem tundis programm ära neutraalset ning üllatuse näoilmet ning mittestandardiseeritud piltide puhul ei suutnud FaceReader enam usaldusväärset emotsioone tuvastada.

Käesoleva uurimuse erisused

Võrreldes varasemate uuringutega ei kasuta ma käesolevas töös treenitud video-moodelle, vaid emotsioone esitavad eelneva ettevalmistuseta inimesed, mis võib suuremaid üldistusvõimalusi tagada. Lisaks esitan stiimulid ehk emotsiooni väljendavad materjalid videolõikudena piltide asemel, mille kasutamist on hetkeseisuga vähem uuritud. Lisaks pole selle töö videomaterjale veel ühelegi masinõppele varem analüüsimeks antud.

Käesoleva töö eesmärk, uurimisküsimused ja hüpoteesid

Käesoleva töö eesmärk on uurida, kui täpselt suudab masinõppe tarkvara FaceReader (versioon 9.0) tuvastada baasemotsioone ning võrrelda FaceReaderi tulemusi inimeste hinnangutega baasemotsioonide äratundmises. Lähtudes varasemast kirjandusest, püstitatakse järgmised hüpoteesid:

- H1. FaceReader on võimeline väljendatud baasemotsioone tuvastama vähemalt 70% täpsusega.
- H2. FaceReader on võimeline hinnatavatest baasemotsioonidest kõige suurema täpsusega ära tundma rõõmu.
- H3. FaceReader tuvastab teistest baasemotsioonidest kõige väiksema täpsusega hirmu.

Lisaks püstitatakse masinõppe ja inimese hinnangute võrdlemiseks järgmised eksploratiivsed uurimisküsimused:

- UK1. Kes on emotsioonide tuvastamisel täpsem - FaceReader või inimesed?

UK2. Missuguseid erinevusi esineb emotsioonide tuvastamises FaceReaderi ja inimeste vahel?

Meetod

Valim

Video-modellid

Juuse (2020) töös osales 119 isikut, kelle hulgast valiti 100 videomodelli. Nende kohta koguti veebikeskkonnas KAEMUS hinnanguid emotsioonide väljendamisele. Ühe katseisiku emotsiooni ei suutnud FaceReader välja lugeda, mistõttu lõplik videomodellide valim koosnes 99 isikust. Sellest valimist moodustasid mehed 38,4% ning naised 61,6%. Vanusevahemik oli 18-49 aastat ning keskmine vanus 24,91 ($SD=6,41$). Eesti keelt kõnelejaid isikuid oli 77,8%, vene keelt 16,2% ning inglise keelt 6,1%.

Inimestest hindajad

Inimhinnangud pärinevad Rebase (2023) tööst, kus originaalvalimis kasutati 144 katseisiku vastustest vaid 121, sest kõik osalejad ei teinud veebikatset lõpuni. Käesoleva töö jaoks oli piisav ka ainult ühele videoklipile hinnangu andmine. Seega koosnes siinne valim 144 inimesest. Inimhindajate vanusevahemik oli 14-77 aastat ning keskmine 31 aastat ($SD=13,89$). Meessoost osalejad moodustasid 19,4% ja naissoost osalejad 79,9%. Üks katseisik oma sugu avaldada ei soovinud. Katseisikuid kutsuti uuringust osa võtma Tartu Ülikooli e-kirja loendite ning sotsiaalmeedia kaudu ehk tegemist oli mugavusvalimiga.

Stiimulmaterjal

1-sekundilised videoklipid, mida selle uurimistöös katseisikud analüüsisin, valmistas enda magistritöö tarbeks Juuse (2020). Täpsemalt toimus suurem andmete kogumine uurimisprojekti „Recognizing expressed and hidden emotions by face and electrophysiological signals from the brain” raames. Töö valmimist toetas Eesti Teadusagentuuri grant PRG1151 “Tähelepanueelne informatsioonitöötlus ajus: seosed seisundite, püsitunnuste ja käitumisega” (01.01.2021-31.12.2025).

Kokku loodi neli katsetingimust, millest käesolevas uurimistöös kasutan vaid esimest ehk emotsioonide väljendamise tingimust, kus paluti osalejatel näoväljenduste abil peegeldada ekraanil esitatud miimilisi emotsioone (viha, vastikust, rõõmu, kurbust, üllatust ja hirmu) ja neutraalsust. Sealne stiimulmaterjal pärines Ekmani näopiltide

andmebaasist (JACFEE) (Matsumoto & Ekman, 1988). Teised katsetingimused ei puuduta käesolevat tööd ning nende ülesehituse kohta on võimalik lähemalt lugeda Juuse (2020) magistritööst.

Inimhinnangud kogus Rebane (2023), näidates neile veebikeskkonnas KAEMUS Juuse (2020) 1-sekundilisi videoklippe rõõmust, hirmust ja vastikusest. Katseisikutel paluti hinnata, millist emotsiooni ning millisel määral video-modell seda väljendas. Võimalik oli valida kõigi kuue baasemotsiooni vahel. Lähemalt on võimalik selle katse kohta lugeda Juuse (2020) ja Rebase (2023) töödest.

Antud töös analüüsisin seega kõikidest baasemotsioonidest vaid rõõmu, hirmu ning vastikust. Selle tulenes oli soovist vältida katseisikute ülekoormamist. Rõõmu, hirmu ja vastikust kaldutakse omavahel kõige harvem segi ajama (Calvo & Nummenmaa, 2015). Selle töö tarbeks andsid katseisikud seega hinnanguid kokku 294 videoklipile: rõõm 98, hirm 99 ja vastikus 97. Enne esitamist hindasid eelnevalt 6-sekundiliste klippide emotsiooni alg-, kõrg- ning lõpppunkti viis sõltumatut isikut, mille põhjal lõigati välja lõplik 1-2 sekundi osa.

Facereader analüüsis klippe igast baasemotsioonist kui ka neutraalsusest, aga inimestega võrdlemise eesmärgil kasutasin ainult rõõmu, hirmu ning vastikust. Antud töös hindas programm kokku 294 klippi: rõõmu 98, hirmu 99 ja vastikust 97 korda. FaceReader annab iga pildi või videoklipi kohta hinnangu skaalal 0-1 vastavalt sellele, kui palju ta konkreetsetes materjalis kõiki kuute baasemotsiooni ja neutraalsust tuvastas. Siin töös võtsin täpsuse leidmisel arvesse kõrgeima hinnanguga baasemotsiooni, mida FaceReader videolõigule pakkus. Selle hinnangu abil tehti kindlaks, kas vastus osutus õigeks ja vastas videos esitatud emotsioonile.

Protseduur

FaceReaderile (vt sissejuhatuse alapeatükk FaceReader) sisestati videoklipid, misjärel hindas programm 0-1 skaalal, kui suurel määral igas klipis kuute baasemotsiooni tuvastas. Andmeid töödeldi Tartu Ülikooli neuroturunduse laboris ajavahemikus juuni-august 2024. aastal. Videomaterjalid sisestati 6-sekundiliselt ehk kogupikkuses. Inimhindajatega samaväärse materjali analüüsimiseks eraldati sellest ajatempli abil 1-2 sekundi osa.

Autori panus

Mina võrdlen varem kogutud inimhinnangute ja FaceReaderi vastuste põhjal mõlema täpsusprotsente ning tuvastusmäärasid skaalal 0-1. Sama materjali põhjal uurib ka Merilo (2025) FaceReaderi võimekust emotsioonide tuvastamisel ning kuidas seostub

sellega videomodellide ekstraversus. Selles töös tugineb ta sellele, kas suurima väärtusega emotsioon vastab videoklipis esitletule. Siin töös analüüsin lisaks täppi ja mööda pandud vastustele ka teisi FaceReaderi poolt pakutud vastuseid. Käesolev töö aitab antud valdkonda panustada eelkõige uute ja rohkem dünaamiliste (videoklipid piltide asemel) stiimulite kasutamisega, mille abil saab ka FaceReader programmi testida ning samal ajal kõrvutada inimhinnangutega.

Eetilised aspektid

Katseisikuid informeeriti varasemalt eesmärkidest, protseduurikäigust ning said anda oma nõusoleku enne osalemist. Katses osalemine oli vabatahtlik, igal hetkel oli võimalik teha pause kui ka osalemisest loobuda. Osalejate isikutega soetud andmed hävitati uuringu lõppedes. Tartu Ülikooli eetikakomiteelt oli luba läbi viia ka stiimulite loomist Juuse magistritöö raames (2020). Kuna mõlemas töös on katseisikud andnud loa enda andmeid edasistes uuringutes kasutada, tohin ka mina neid enda töös rakendada eeldusel, et kasutan neid vaid uurimistöö tarbeks ning kannan hoolt selle eest, et need kuhugi ei lekiks. Lisaks olen nii uurimuse kulgemise kui tulemuste edastamisel läbipaistev ning esitan need sisuliselt muutmata ja ausal kujul, et nii mina kui teised lugejad saaksid teha võimalikult tõesed järeldused antud teemal.

Andmeanalüüs

Andmeanalüüsid viisin läbi JASP 0.17.3.0 tarkvaraprogrammis.

Masinhinnanguteks kasutasin FaceReader (versioon 9.0) komertstarkvara (Noldus Information Technology, Wageningen, Holland). FaceReaderi puhul hinnati videoklippe skaalal 0-1 vastavalt sellele, kui suurel määral kuute baasemotsiooni ning neutraalset näoväljendust see ära tundis. Kõigi analüüside puhul oli olulisuse nivooks seatud 0,05.

Andmete tõlgendamiseks viisin läbi kolm statistilist testi. Et hinnata, milliseid emotsioone ja kas statistiliselt teistest erinevalt suudeti eristada, viisin läbi korduvmõõtmiste dispersioonanalüüsi (ANOVA) nii inimeste (analüüs 1) kui ka masinõppe vastuste (analüüs 2) põhjal. Korduvmõõtmiste ANOVA eelduseks on andmete sfäärilisus, mis eeldab, et erinevate faktortasemete hajuvused on sarnased. Kõigil juhtudel oli see eeldus siiski rikutud, mida kinnitas positiivne tulemus Mauchly testis. Tulemuste tõlgendamiseks kasutasin seega Greenhouse-Geisseri korrektsiooni.

Masinõppe ja inimeste vastuste kõrvutamiseks viisin läbi sõltumatute gruppidega kahefaktoriaalse ANOVA kolme emotsiooni lõikes (analüüs 3). Täpsemalt keskendusin interaktsioonidele ehk sellele, kas inimhindajad ja FaceReader hindasid

emotsioonivastuseid sarnaselt või erinevalt. Sõltuv muutuja oli vähemalt intervallskaalal ning sõltumatud muutujad kategoriaalsed, kuid täidetud ei olnud hajuvuste sarnasuse (kontrollitud Levene'i testiga, $p < 0,001$) ning normaaljaotuslikkuse (kontrollitud kvartiil-kvartiil diagrammi, asümmeetriakordaja (3,78), ekstsessi (2,16) väärtuste ning Shapiro-Wilki testi ($p < 0,001$) kaudu) eeldused. Otsustasin jätkata ANOVA läbiviimisega, sest uurimuse eesmärk põhines interaktsioonide analüüsimisel ning mitteparameetrilised testid seda ei võimalda. Ka JASP-is puuduvad nende rikkumistega koosmõjude uurimiseks teised võimalused. Seega tuleb parameetrilise testi tulemustesse ettevaatusega suhtuda.

Tulemused

Inimeste suutlikkus emotsioone tuvastada

Inimhindajad suutsid üldiselt kolme baasemotsiooni (rõõmu, vastikust ja hirmu) tuvastada 54,4% täpsusega (tabel 1). Kõige täpsemini suudeti kolmest baasemotsioonist õigesti ära tunda rõõmu (82,7%). Kõige keerulisem oli inimestel ära tunda hirmu (19,2%). Vastikust suudeti ära tunda 61,9% täpsusega.

Tabel 1

Sagedustabel inimhindajate õigetest vastustest kolmele baasemotsioonile ainult kõrgeima hinnangu põhjal.

Baasemotsioon	Õigesti vastatud		Valesti vastatud	
	sagedus	%	sagedus	%
rõõm	81	82,7	17	17,4
hirm	19	19,2	80	80,8
vastikus	60	61,9	37	38,1
kokku	160	54,4	134	45,6

Korduvmõõtmiste ANOVA näitas, et statistiliselt oluliselt suudeti eristada rõõmu ($F(1,4) = 218,53$; $p < 0,001$, $\eta^2 = 0,69$). Täpsemate erinevuste leidmiseks viisin läbi korrigeeritud holmi järeltesti, mille kohaselt suudeti rõõmu eristada kõikidest teistest baasemotsioonidest (lisa 1) ning kõige kõrgema tuvastusmääraga. Seega pakuti keskmiselt rõõmu klippidele enim vastuseks rõõmu ($M = 0,69$; $SD = 0,31$; $N = 98$) (tabel 2) ning kõige vähem viha ($M = 0,04$; $SD = 0,09$; $N = 98$).

Hirmu suudeti eristada stasisiliselt oluliselt ($F(3,2) = 16,1$; $p < 0,001$, $\eta^2 = 0,14$). Korrigeeritud holmi järeltesti kohaselt suudeti statistiliselt oluliselt eristada hirmu rõõmust ja vihast, kuid mitte üllatusest, kurbusest ega vastikusest (lisa 1, tabel 2). Hirmu

klippidele pakkusid inimesed keskmiselt kõige rohkem vastuseks üllatust ($M = 0,29$; $SD = 0,28$; $N = 99$) (tabel 2) ning kõige vähem statistiliselt oluliste hulgast rõõmu ($M = 0,06$; $SD = 0,13$; $N = 99$).

Vastikust suudeti sarnaselt rõõmule eristada statistiliselt oluliselt kõikidest baasemotsioonidest korrigeeritud holmi järeltesti andmetel ($F(1,9) = 77,52$; $p < 0,001$, $\eta^2 = 0,45$) (lisa 1, tabel 3). Vastikusele pakkusid inimhindajad enim vastuseks vastikust ($M = 0,54$; $SD = 0,34$; $N = 97$) (tabel 2) ning kõige vähem üllatust ($M = 0,05$; $SD = 0,05$; $N = 97$).

Tabel 2

Kui suurel määral tuvastasid inimhindajad keskmiselt kolme emotsiooni lõikes kuut baasemotsiooni skaalal 0-1.

Baasemotsiooni klipp	vastikus	rõõm	kurbus	viha	üllatus	hirm
rõõm	0,05 (0,08)	0,69 (0,31)	0,11 (0,19)	0,04 (0,09)	0,08 (0,06)	0,04 (0,06)
hirm	0,14 (0,19)	0,06 (0,13)	0,18 (0,22)	0,09 (0,13)	0,29 (0,28)	0,23 (0,17)
vastikus	0,54 (0,34)	0,08 (0,17)	0,15 (0,21)	0,13 (0,14)	0,05 (0,05)	0,06 (0,07)

Märkus. Paksus kirjas on tähistatud õige hinnang (vastavus eesmärk-emotsioonile). Kaldkirjas on esitatud vastava videoklipi kõige kõrgem hinnang. Sulgudes on esitatud standardhälve.

FaceReader tarkvara võime emotsioone tuvastada

FaceReader oli võimeline üldiselt kolme baasemotsiooni tuvastama 37,4% täpsusega (tabel 3). Kõige rohkem suutis FaceReader kolmest baasemotsioonist õigesti ära tunda rõõmu (96,9%). Kõige ebatäpsem oli programm hirmu tuvastamisel (4%). Vastikuse tuvastamismäär jäi 11,3%.

Tabel 3

Sagedustabel FaceReaderi õigetest vastustest kolmele baasemotsioonile ainult kõrgeima hinnangu põhjal.

Baasemotsioon	Õigesti vastatud		Valesti vastatud	
	sagedus	%	sagedus	%
rõõm	95	96,9	3	3,1
hirm	4	4	95	96
vastikus	11	11,3	86	88,7
kokku	110	37,4	184	62,6

Baasemotsioonide hindamistäpsuse analüüsimiseks FaceReaderi kaudu viidi läbi ühefaktoriline korduvmõõtmiste ANOVA, mis näitas, et rõõmu suutis Facereader eristada statistiliselt oluliselt ($F(1,2) = 541,1; p < 0,001, \eta^2 = 0,85$). Holmi järeltest näitas, et masinõpe suutis rõõmu eristada kõikidest teistest baasemotsioonidest (lisa 1, tabel 4). Rõõmu klippidele pakuti enim vastuseks ka rõõmu ($M = 0,68; SD = 0,25; N = 98$) (tabel 4) ning kõige vähem hirmu ($M = 0,00; SD = 0,01; N = 98$).

Hirmu suutis FaceReader statistiliselt oluliselt eristada ($F(2,4) = 47,5; p < 0,001, \eta^2 = 0,33$). Rõõmust ja üllatusest suutis hirmu eristada, kuid mitte vihast, kurbusest ega vastikusest (lisa 1, tabel 5). Hirmu klippidele pakkus FaceReader keskmiselt kõige rohkem vastuseks rõõmu ($M = 0,26; SD = 0,24; N = 99$) (tabel 4) ning kõige vähem statistiliselt oluliste hulgast üllatust ($M = 0,14; SD = 0,16; N = 99$).

Programm suutis ka vastikust eristada statistiliselt oluliselt ($F(1,8) = 64,39; p < 0,001, \eta^2 = 0,4$). Holmi järeltesti kohaselt suudeti vastikust eristada rõõmust ja hirmust, kuid mitte kurbusest, vihast ega üllatusest (lisa 1, tabel 6). Vastikusele pakkus FaceReader enim vastuseks rõõmu ($M = 0,32; SD = 0,27; N = 97$) (tabel 4) ning kõige vähem statistiliselt oluliste hulgast hirmu ($M = 0,01; SD = 0,02; N = 97$).

Tabel 4

Kui suurel määral tuvastas FaceReader keskmiselt kolme emotsiooni lõikes kuut baasemotsiooni skaalal 0-1.

Baasemotsiooni klipp	vastikus	rõõm	kurbus	viha	üllatus	hirm
rõõm	0,01 (0,03)	0,68 (0,25)	0,01 (0,03)	0,04 (0,06)	0,04 (0,06)	0,00 (0,01)
hirm	0,02 (0,03)	0,26 (0,24)	0,02 (0,03)	0,09 (0,11)	0,14 (0,16)	0,03 (0,07)
vastikus	0,09 (0,12)	0,32 (0,27)	0,3 (0,05)	0,12 (0,14)	0,04 (0,06)	0,01 (0,02)

Märkus. Paksus kirjas on tähistatud õige hinnang (vastavus eesmärk-emotsioonile). Kaldkirjas on esitatud vastava videoklipi kõige kõrgem hinnang. Sulgudes on esitatud standardhälve.

Masinõppe ja inimeste võrdlus

Masinõppe ning inimeste väljendatud emotsioonide äratundmistäpsuse võrdlemiseks viidi läbi mitmefaktoriline sõltumatute rühmade ANOVA muutujatega emotsioon (6) x meetod (2) x tingimus (3). Emotsioonide all mõeldi kuut võimalikku valikut baasemotsioonide hulgast (rõõm, hirm, vastikus, kurbus, üllatus ning viha), meetodi all masina või inimhinnangut ning tingimuse all eeldatavat eesmärk-emotsiooni (kas väljendatav emotsioon pidi väljendama rõõmu, hirmu või vastikust). Tulemused on

esitatud tabelis 5. Kõik kolmetasandilise interaktsiooni (emotsioon (6) x meetod (2) x tingimus (3)) järeldest tulemused on esitatud lisas 2.

Tabel 5

Masinõppe ning inimeste väljendatud emotsioonide äratundmistäpsust võrdlev dispersioonanalüüs (ANOVA).

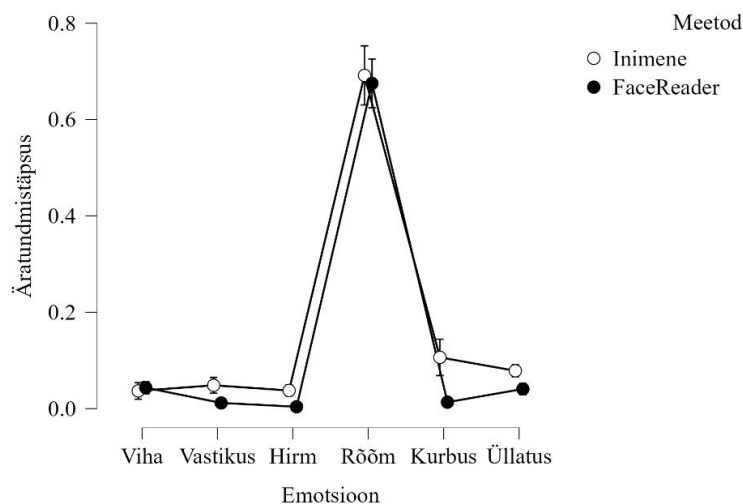
	<i>df</i>	<i>F</i> -statistik	<i>p</i> -väärtus	η^2
Emotsioon	5	273,74	< 0,001	0,178
Meetod	1	121,91	< 0,001	0,016
Tingimus	2	5,10	0,006	0,001
Emotsioon * Meetod	5	83,73	< 0,001	0,054
Emotsioon * Tingimus	10	195,47	< 0,001	0,254
Meetod * Tingimus	2	5,10	0,006	0,001
Emotsioon * Meetod * Tingimus	10	31,76	< 0,001	0,041

Märkused. *df* = vabadusaste; η^2 = efektiivsus. Jääkide vabadusaste (*df*₂) on 3492. ANOVA läbiviimiseks on kasutatud III tüüpi ruutude summa meetodit.

Huvialusteks võrdlusteks olid kolmetasandilise interaktsiooni tulemused. Rõõmu tingimuses ilmnes, et inimhinnangute ja masina täpsuse vahel ei esinenud statistiliselt olulist erinevust ($t(3492) = 0,75$; $p = 1$; $d = 0,11$). Nii inimesed kui FaceReader andsid rõõmu väljendavate videoklippide hinnanguks sarnaseid täpseid vastuseid (inimhindajad $M = 0,69$; $SD = 0,31$; FaceReader $M = 0,68$; $SD = 0,25$) (joonis 2).

Joonis 2

Võrdlus inimese ja FaceReaderi vahel rõõmu äratundmisel.



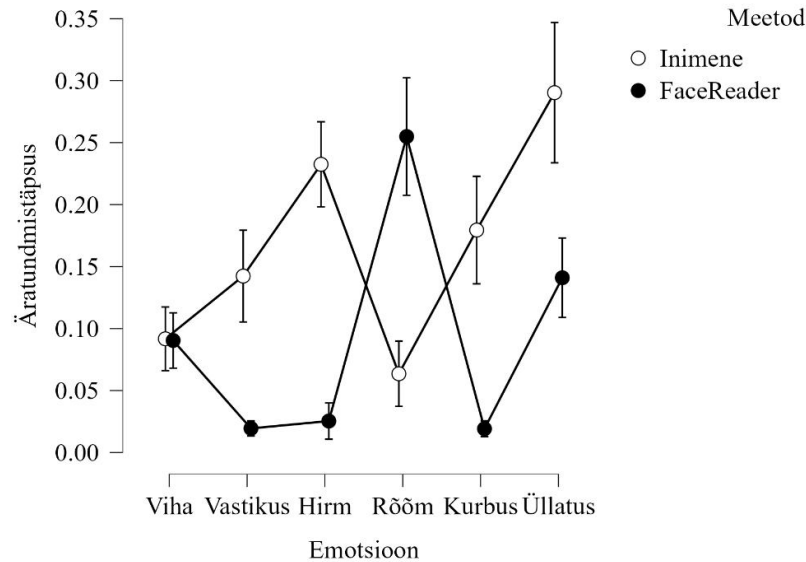
Märkus. Äratundmistäpsus tähendab siin 0-1 skaalal kuue baasemotsiooni nägemist klippides.

Hirmu tingimuses ilmnemid statistiliselt olulised erinevused inimhinnangute ja masina emotsiooni äratundmistäpsuse vahel ($t(3492) = 9,32$; $p < 0,001$; $d = 1,33$). Hirmu

äratundmistäpsus oli inimeste ($M = 0,23$; $SD = 0,17$) puhul kõrgem kui FaceReaderi ($M = 0,03$; $SD = 0,07$) puhul (joonis 3).

Joonis 3

Võrdlus inimese ja FaceReaderi vahel hirmu äratundmisel.

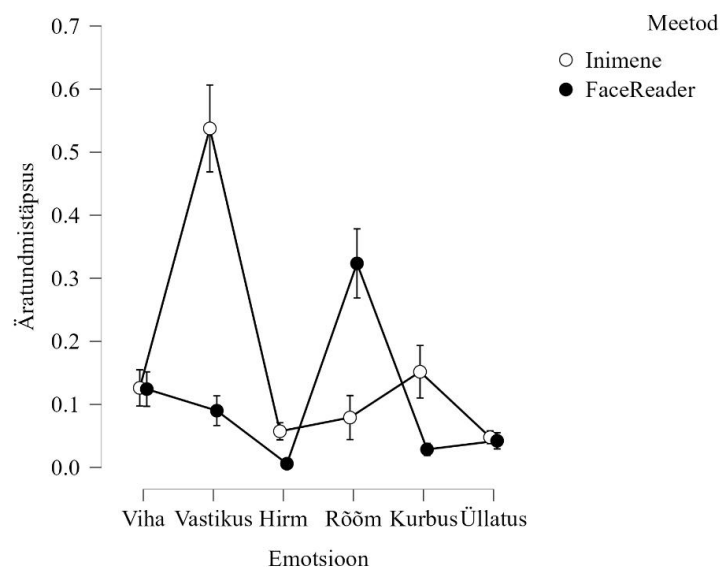


Märkus. Äratundmistäpsus tähendab siin 0-1 skaalal kuue baasemotsiooni nägemist klippides.

Ka vastikuse tingimuses ilmnese erinevused inimhinnangute ja masina emotsiooni äratundmistäpsuse vahel ($t(3492) = 19,94$; $p < 0,001$; $d = 2,86$). Vastikuse äratundmistäpsus oli inimeste ($M = 0,54$; $SD = 0,34$) puhul kõrgem kui FaceReaderi ($M = 0,09$; $SD = 0,12$) puhul (joonis 4).

Joonis 4

Võrdlus inimese ja FaceReaderi vahel vastikuse äratundmisel.



Märkus. Äratundmistäpsus tähendab siin 0-1 skaalal kuue baasemotsiooni nägemist klippides.

Arutelu

Käesoleva töö eesmärgiks oli selgitada välja, kui täpselt suudab masinõppe tarkvara FaceReader baasemotsioone tuvastada ning seejärel võrrelda FaceReaderi emotsioonide tuvastamise täpsust inimese täpsusega baasemotsioonide äratundmisel.

Tuginedes varasemale kirjandusele, püstitasin kolm hüpoteesi. Esimeseks hüpoteesiks (H1) püstitasin, et FaceReader on võimeline väljendatud baasemotsioone tuvastama vähemalt 70% täpsusega. FaceReader oli hoopis võimeline kõigi kolme baasemotsiooni üleselt emotsioone tuvastama 37,4% täpsusega (tabel 3). Samas suutis FaceReader kõige rohkem kolmest baasemotsioonist õigesti ära tunda rõõmu 96,9% täpsusega, mis kinnitab ka teist hüpoteesi (H2: FaceReader tuvastab hinnatavatest baasemotsioonidest kõige täpsemalt rõõmu). See tähendab, et esimene hüpotees leidis kinnitust vaid osaliselt. Madal üldine tuvastustäpsus pole kooskõlas mitmete FaceReaderit hindavate uurimustega, mis on saavutanud emotsionaalsete näoväljenduste tuvastamisel oluliselt kõrgemaid täpsusprotsente (Lewinski et al., 2014; Skiendziel et al., 2019;). Samas kui vaadata Küntzleri ja kolleegide (2021) tööd, siis ka seal täheldati erinevusi FaceReaderi võimekuses emotsioone ära tunda ning seda standardiseeritud ja mittestandardiseeritud stiimulmaterjali tingimustes. Masinõpete puhul on üldiselt märgatud, et kehvema pildi- või videomaterjalide kvaliteet, selgus ja näoilme intensiivsus võivad tuvastamisprotsenti väiksemaks muuta (De Araujo et al., 2022). Need võivad seletada ka selle töö tulemusi, kuna siinne videomaterjal ei sarnanenud valideeritud pilt- või videomaterjalidele. Ühest küljest võib oletada, et siinsed videoklipid on ebakvaliteetsed või kehva nurga all üles lindistatud, mistõttu võiks neist arusaamine olla raskendatud. Kui aga arvesse võtta siinses töös ilmnenu kõrgeid tuvastamisprotsente rõõmule nii inim- ning masinõppehinnangute puhul, ei saa ilmselt lõpuni tulemusi ka sellega ära seletada.

Nagu eelnevalt mainitud, suutis FaceReader kõige täpsemalt kolmest baasemotsioonist õigesti ära tunda rõõmu (96,9%). Kui vaadata tuvastamismäära skaalal 0-1 (tabel 4), siis näeme näeme kõrget ja teistest valikuvariantidest tugevalt eristuvat hinnangut (järgmisena oleks FR pakkunud viha tuvastamismääraga 0,04). Samale tulemusele, et rõõmu on masinõpetel lihtsaim ära tunda, on jõudnud ka paljud varasemad uurimused (De Araujo et al., 2022; Skiendziel et al., 2019; Ivanova & Borzunov, 2020; Lévêque et al., 2022; Lewinski et al., 2014; Dores et al., 2020; Elfenbein & Ambady, 2002). Kõige ebatäpsem oli programm seevastu hirmu tuvastamisel (4 %), mis kinnitab

kolmandat hüpoteesi (H3: FaceReader tuvastab baasemotsioonidest kõige väiksema täpsusega hirmu). Ka seda on mitmed varasemad tööd täheldanud (Ivanova & Borzunov, 2020). Üldiselt kiputakse negatiivse valentsiga emotsioone omavahel rohkem segi ajama (De Araujo et al., 2022; Küntzler et al., 2021; Ivanova & Borzunov, 2020; Skiendzieli et al., 2019), ent siin töös pakkus FaceReader lõige rohkem hirmu klippidele vastuseks rõõmu ning alles teisena alles üllatust. Peale rõõmu ja üllatuse ei suutnud masinõppe statistiliselt oluliselt ka ülejäänud emotsioone hirmust eristada. Seda võib seletada tavapärasest erineva ja mitte valideeritud stiimulmaterjali kasutamisega. Samas üllatuse pakkumist teisena võisid põhjustada sarnased näoilmed hirmu ja üllatuse puhul, sest nende väljendamisel kasutatakse sarnaseid väiksemaid näolihaseid (FACS).

Masinõppe ja inimhinnangute võrdlemisel püstitasin üheks eesmärgiks vastata ka uurimisküsimusele (UK1): Kes on emotsioonide tuvastamises täpsem - FaceReader või inimhindajad? Sarnaselt masinõpetele ja FaceReaderile on inimestel enamasti lihtsaim ära tunda rõõmu (De Araujo et al., 2022; Skiendziel et al., 2019; Ivanova & Borzunov, 2020; Lévêque et al., 2022; Lewinski et al., 2014; Dores et al., 2020; Elfenbein & Ambady, 2002), mida kinnitas ka see töö. Ka inimesed ajavad negatiivseid emotsioone omavahel rohkem segamini (Dores et al., 2020). Hirmu tuvastamine on sageli keeruline (Lévêque et al., 2022; Elfenbein & Ambady, 2002), mida täheldas ka see töö. Hirmu klippidele pakuti kõige rohkem vastuseks üllatust, mida on ka varem näidatud (Zhao et al., 2017). Vastikusele pakuti kõige rohkem vastuseks vastikust ennast, samas mitte ülekaalukalt järgnesid kurbus ja viha, mida on varasemates tööd as omavahel seostatud (Hendel et al., 2023).

Teiseks uurimisküsimuseks püstitasin küsimuse (UK2): Millised erinevused ilmnevad emotsioonide tuvastamisel FaceReaderi ja inimeste hinnangute vahel? Kui kõrvutada nii õiged ja valed vastused (tabelid 1 ja 3) kui ka hinnangud skaalal 0-1 (tabelid 2 ja 4), siis on näha, et inimene valis rohkem õigeid emotsioone vastuseks ning suutis ka emotsioone intensiivsemaks hinnata. Seega võiks öelda, et antud töös olid inimesed masinõppest täpsemad. Puudutades eelnevalt läbi käinud teemat ning vaadates uurimusi, mis on võrrelnud inimeste ja masinõppe sooritust emotsioonide tuvastamisel staatiliste vs dünaamiliste stiimulitega ning selgete vs ebakvaliteetsete materjalidega, on leitud, et inimestele videoklipid olulisi raskusi ei paku ning nende kasutamine võib viia isegi parema äratundmiseni (Krumhuber et al., 2013). Seda seepärast, et liikumise kaudu on võimalik tuvastada rohkem informatsiooni, eriti halvema kvaliteedi stiimulite puhul.

On leitud, et loomulikumaid näoväljendusi tuvastavad lisaks masinõpetele kehvemini ka inimesed (Lévêque et al., 2022; Goodfellow et al., 2015), aga kui pildid on kehvema kvaliteediga või tehtud spontaanselt (ebaselgelt poseeritud ja loomulikumas keskkonnas), kalduvad inimesed siiski emotsioone võrreldes masinõppe meetoditega paremini ära tundma (Lévêque et al., 2022; Yitzhak et al., 2017). Eelnev võib aimu anda, miks masinõppe ning inimhinnangud niivõrd erinesid ning inimeste hinnangud paremaks osutusid.

Antud töö tugevuseks on videoklippide ja mittevalideeritute stiimulite kasutamine, mis aitavad kaasa uuringu välisele valiidsusele. Kvaliteetsed piltmaterjalid on tõenäoliselt paremini arusaadavamad, ent düünamilisus ja mitte alati hea nägemisnurk on omased loomulikule keskkonnale, kus harilikult emotsioonide äratundmist vaja on. Varasemad tööd on näidanud, et videoklippide näitamine masinõpetele võib tuvastusmäära ebatäpsemaks muuta (Lévêque et al., 2022; Yitzhak et al., 2017). Siiski ei pruugi see erinevus olla oluliselt suur (Lewinski et al., 2014). Võib-olla oleksid FaceReaderi tulemused nii siin kui ka varasemates sarnase tingimusega töödes veidi täpsemad olnud, kui masinõppel oleks treenimise ajal rohkem kogemust ka kehvema kvaliteedi materjalidega. Näiteks Stöckli ja kolleegid (2017) toovad välja, et paljud masinõppe algoritmid on emotsioonide tuvastamiseks treenitud standardiseeritud selgete pilt- ja videomaterjalide abil. Selliseid stiimulmaterjale kasutatakse tihti ka masinõppe ja emotsioonituvastamise uurimustes, et nende äratundmisvõimekust hinnata.

Tulevikus võiks eelnevast tulenevalt uurida, kas inimhinnanguid ja masinõpet võib mõjutada emotsioonide esitlemine pildi või videolõikude kujul. Otstarbekas oleks võrrelda materjale, mis pärinevad loomulikust ja igapäevasest keskkonnast stereotüüpsete ja kvaliteetsete materjalidega. Hetkel on vähe uurimusi, kus nii inimest kui masinõpet oleks neis kõrvutatud. Siin oleks varieerumist selles osas, kui dünaamiline videolõik saab olla (näiteks käesolevas 1 sekund) või mida tähendab loomulik või igapäevane keskkond ja ebaselge kvaliteet. Siiski tuleksid need teadmised kasuks nii FaceReaderi kui paljude teiste masinõpete efektiivsemal rakendamisel igapäevaellu.

Kokkuvõttes olid seega tulemused pigem ootuspärased ning kooskõlas varasemate teadmistega. Inimesed osutusid antud töös paremaks kui masinõpe, tuvastades näoilmetest emotsioone täpsemini ning suutes võrreldes masinõppega skaalal 0-1 neid ka kõrgemalt välja lugeda. Kõige täpsemalt tundsid nii inimhindajad kui FaceReader ära rõõmu ning kõige ebatäpsemalt hirmu. Antud uurimus panustas

emotsiooni tuvastamise tarkvara töö uurimisse uute stiimulite abil, võrreldes seda inimese suutlikkusega. Siin kasutatud andmeanalüüs lubab FaceReaderi vastusse põhjalikumalt keskenduda kui ainult õige ja vale täpsusprotsenti kasutades.

Kasutatud kirjandus

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current opinion in neurobiology*, 12(2), 169–177. [https://doi.org/10.1016/s0959-4388\(02\)00301-x](https://doi.org/10.1016/s0959-4388(02)00301-x)
- Arnold, M. B. (1960). *Emotion and personality*. Columbia University Press.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, Disgusted, or Afraid?: Studies on the Malleability of Emotion Perception. *Psychological Science*, 19(7), 724-732. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>
- Barrett, L. F. (2010). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion*, 12(4), 579–599. <https://doi.org/10.1080/026999398379574>
- Browning, M., Kingslake, J., Dourish, C. T., Goodwin, G. M., Harmer, C. J., & Dawson, G. R. (2019). Predicting treatment response to antidepressant medication using early changes in emotional processing. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*, 29(1), 66–75. <https://doi.org/10.1016/j.euroneuro.2018.11.1102>
- Celeghin, A., Diano, M., Bagnis, A., Viola, M., & Tamietto, M. (2017). Basic Emotions in Human Neuroscience: Neuroimaging and beyond. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01432>
- Coan, J. A., & Allen, J. J. B. (Eds.). (2007). *Handbook of emotion elicitation and assessment*. Oxford University Press.
- De Araujo, H. F., Nunes, F. L. S., & Machado-Lima, A. (2022). The impact of different facial expression intensities on the performance of pre-trained emotion recognition models. *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*. <https://doi.org/10.1145/3477314.3507268>
- Dodge, S. & Karam, L. (2017). A Study and Comparison of Human and Deep Learning Recognition Performance under Visual Distortions. *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, <https://ieeexplore.ieee.org/abstract/document/8038465>

- Dores, A. R., Barbosa, F., Queirós, C., Carvalho, I. P., & Griffiths, M. D. (2020). Recognizing Emotions through Facial Expressions: A Largescale Experimental Study. *International Journal of Environmental Research and Public Health/International Journal of Environmental Research and Public Health*, 17(20), 7420. <https://doi.org/10.3390/ijerph17207420>
- 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). <https://doi.org/10.1371/journal.pone.0231968>
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial Action Coding System. Manual and Investigator's Guide, Salt Lake City, UT: Research Nexus.
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2), 203–235. <https://doi.org/10.1037/0033-2909.128.2.203>
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2), 212–228. <https://doi.org/10.1037/0022-3514.57.2.212>
- Goodfellow, I. J., Erhan, D., Luc Carrier, P., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D. H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., Ionescu, R., Bengio, Y. (2015). Challenges in representation learning: a report on three machine learning contests. *Neural networks : the official journal of the International Neural Network Society*, 64, 59–63. <https://doi.org/10.1016/j.neunet.2014.09.005>
- Gross, J. J., & Barrett, L. F. (2011). Emotion Generation and Emotion Regulation: One or Two Depends on Your Point of View. *Emotion review : journal of the International Society for Research on Emotion*, 3(1), 8–16. <https://doi.org/10.1177/1754073910380974>
- Hendel, E., Gallant, A., Mazerolle, M., Cyr, S., & Roy-Charland, A. (2023). Exploration of visual factors in the disgust-anger confusion: the importance of the mouth. *Cognition and Emotion*, 37(4), 835–851. <https://doi.org/10.1080/02699931.2023.2212892>
- Holtmann, O., Bruchmann, M., Mönig, C., Schwindt, W., Melzer, N., Miltner, W. H. R., & Straube, T. (2020). Lateralized deficits of disgust processing after Insula-Basal

- ganglia damage. *Frontiers in Psychology*, 11.
<https://doi.org/10.3389/fpsyg.2020.01429>
- Ivanova, E., & Borzunov, G. (2020). Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions. *Procedia Computer Science*, 169, 244–248. <https://doi.org/10.1016/j.procs.2020.02.143>
- Iyortsuun, N. K., Kim, S., Jhon, M., Yang, H., & Pant, S. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare*, 11(3), 285. <https://doi.org/10.3390/healthcare11030285>
- Juuse, L. (2020). *The relationship between visual and verbal information and emotion processing: reports from EEG and self-assessment*. Magistritöö: Tartu Ülikool. Psühholoogia instituut.
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019b). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, 43(2), 133–160.
<https://doi.org/10.1007/s10919-019-00293-3>
- Keltner, D., Tracy, J. L., Sauter, D., & Cowen, A. (2019a). What basic emotion theory really says for the Twenty-First century study of emotion. *Journal of Nonverbal Behavior*, 43(2), 195–201. <https://doi.org/10.1007/s10919-019-00298-y>
- Ko, B. (2018). A brief review of facial emotion recognition based on visual information. *Sensors*, 18(2), 401. <https://doi.org/10.3390/s18020401>
- Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of Dynamic Aspects of Facial Expressions: a review. *Emotion Review*, 5(1), 41–46.
<https://doi.org/10.1177/1754073912451349>
- Küntzler, T., Höfling, T. T. A., & Alpers, G. W. (2021). Automatic facial expression recognition in standardized and non-standardized emotional expressions. *Frontiers in Psychology*, <https://doi.org/10.3389/fpsyg.2021.627561>
- Lévêque, L., Villoteau, F., Sampaio, E. V. B., & Callet, P. L. (2022). Comparing the robustness of humans and deep neural networks on facial expression recognition. *ResearchGate*.
https://www.researchgate.net/publication/366166277_Comparing_the_Robustness_of_Humans_and_Deep_Neural_Networks_on_Facial_Expression_Recognition
- 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>

- Librenza-Garcia, D., Kotzian, B. J., Yang, J., Mwangi, B., Cao, B., Lima, L. N. P., Bermudez, M. B., Boeira, M. V., Kapczinski, F., & Passos, I. C. (2017). The impact of machine learning techniques in the study of bipolar disorder: A systematic review. *Neuroscience & Biobehavioral Reviews*, 80, 538–554.
<https://doi.org/10.1016/j.neubiorev.2017.07.004>
- Loijens, L., & Krips, O. (s.a.). FaceReader Methodology Note: The Netherlands, Wageningen https://info.noldus.com/hubfs/resources/noldus-white-paper-facereader-methodology.pdf?utm_campaign
- Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The Effects of Fear and Anger Facial Expressions on Approach- and Avoidance-Related Behaviors. *Emotion*, 5(1), 119–124. <https://doi.org/10.1037/1528-3542.5.1.119>
- Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian Facial Expressions of Emotion (JACFEE)*. San Francisco, CA: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.
<https://cir.nii.ac.jp/crid/1573950399720644864?lang=en>
- Merilo, N. L. (2025). *FaceReader programmi võimekus inimemotsioonide tuvastamisel ja seosed videomodelide ekstraversusega*. Uurimistöo: Tartu Ülikool. Psühholoogia instituut.
- Ortony, A. (2021). Are all “Basic emotions” emotions? a problem for the (Basic) emotions construct. *Perspectives on Psychological Science*, 17(1), 41–61.
<https://doi.org/10.1177/1745691620985415>
- Panksepp, J., & Watt, D. (2011). What is Basic about Basic Emotions? Lasting Lessons from Affective Neuroscience. *Emotion Review*, 3(4), 387–396.
<https://doi.org/10.1177/1754073911410741>
- Plutchik, R. (1960). The Multifactor-Analytic Theory of Emotion. *The Journal of Psychology*, 50(1), 153–171. <https://doi.org/10.1080/00223980.1960.9916432>
- Rebane, E. (2023). *Suurel viisikul mõõdetud ekstraversuse seosed emotsioonide väljendamisega*. Uurimistöo: Tartu Ülikool. Psühholoogia instituut.
- Saganowski, S. (2022). Bringing Emotion Recognition Out of the Lab into Real Life: Recent Advances in Sensors and Machine Learning. *Electronics*, 11(3), 496.
<https://doi.org/10.3390/electronics11030496>
- Scherer, K. R. (2000). Psychological models of emotion. In J. C. Borod (Ed.), *The neuropsychology of emotion* (pp. 137–162). Oxford University Press.

- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for?. *Current Directions in Psychological Science*, 20(6), 395–399.
<https://doi.org/10.1177/0963721411424739>
- Skiendziel, T., Rösch, A. G., & Schultheiss, O. C. (2019). Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring. *PloS One*, 14(10), e0223905.
<https://doi.org/10.1371/journal.pone.0223905>
- Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2017). Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior Research Methods*, 50(4), 1446–1460. <https://doi.org/10.3758/s13428-017-0996-1>
- Strongman K.T. (2003). *The Psychology of Emotion: From Everyday Life to Theory*. New York: John Wiley & Sons, 5th ed.
- Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient Processing of Deep Neural Networks: A Tutorial and Survey. *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2295-2329, Dec. 2017, doi: 10.1109/JPROC.2017.2761740.
- Tomkins S. S. (1962). *Affect, imagery, consciousness*. Springer.
- Univer, H. (2022). *Masinõppe algoritmi võimekus tuvastada emotsioone näoväljenduste põhjal – kas parem või halvem kui inimene?*. Magistritöö: Tartu Ülikool. Psühholoogia instituut.
- Yitzhak, N., Giladi, N., Gurevich, T., Messinger, D. S., Prince, E. B., Martin, K., & Aviezer, H. (2017). Gently does it: Humans outperform a software classifier in recognizing subtle, nonstereotypical facial expressions. *Emotion*, 17(8), 1187–1198.
<https://doi.org/10.1037/emo0000287>
- Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59, 103–126. <https://doi.org/10.1016/j.inffus.2020.01.011>
- Zhao, K., Zhao, J., Zhang, M., Cui, Q., & Fu, X. (2017). Neural responses to rapid facial expressions of fear and surprise. *Frontiers in Psychology*.
<https://doi.org/10.3389/fpsyg.2017.00761>

Lisa 1. Inimhindajate ja FaceReaderi järetestide tulemused emotsioonide tuvastatäpsuse erinevustes kolme tingimuse lõikes

Tabel 1

Inimhindajate järetestide tulemused rõõmu tingimuses kuue baasemotsiooni tuvastatäpsuse erinevustes

		Keskliste erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	-0.643	0.036	-17.769	-4.068	< .001***
	Kurbus	-0.058	0.019	-3.102	-0.367	0.015*
	Viha	0.012	0.009	1.232	0.074	0.677
	Üllatus	-0.030	0.010	-2.879	-0.191	0.025*
	Hirm	0.011	0.009	1.232	0.068	0.677
Rõõm	Kurbus	0.585	0.048	12.238	3.701	< .001***
	Viha	0.655	0.036	18.024	4.142	< .001***
	Üllatus	0.613	0.033	18.483	3.877	< .001***
	Hirm	0.654	0.035	18.546	4.136	< .001***
Kurbus	Viha	0.070	0.020	3.528	0.441	0.004**
	Üllatus	0.028	0.020	1.385	0.176	0.677
	Hirm	0.069	0.017	4.148	0.435	< .001***
Viha	Üllatus	-0.042	0.010	-4.265	-0.265	< .001***
	Hirm	-9.325×10 ⁻⁴	0.009	-0.104	-0.006	0.917
Üllatus	Hirm	0.041	0.007	5.602	0.259	< .001***

Märkused. Coheni *d* = efektiivsus. Vabadusaste (*df*) on kõigis tingimustes 97. *P*-väärtust on kohandatud 15 võrdluseks. * *p* < .05, ** *p* < .01, *** *p* < .001.

Tabel 2

Inimhindajate järetestide tulemused hirmu tingimuses kuue baasemotsiooni tuvastatäpsuse erinevustes

		Keskliste erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	0.079	0.024	3.352	0.406	0.009**
	Kurbus	-0.037	0.031	-1.210	-0.191	0.343
	Viha	0.051	0.022	2.340	0.261	0.107
	Üllatus	-0.148	0.041	-3.640	-0.763	0.004**
	Hirm	-0.090	0.027	-3.286	-0.465	0.010**
Rõõm	Kurbus	-0.116	0.026	-4.524	-0.598	< .001***
	Viha	-0.028	0.020	-1.435	-0.145	0.343
	Üllatus	-0.227	0.034	-6.683	-1.169	< .001***
	Hirm	-0.169	0.024	-6.963	-0.871	< .001***
Kurbus	Viha	0.088	0.024	3.718	0.452	0.003**
	Üllatus	-0.111	0.044	-2.498	-0.571	0.085
	Hirm	-0.053	0.032	-1.658	-0.273	0.343
Viha	Üllatus	-0.199	0.036	-5.559	-1.024	< .001***
	Hirm	-0.141	0.024	-5.747	-0.726	< .001***
Üllatus	Hirm	0.058	0.033	1.735	0.298	0.343

Märkused. Coheni *d* = efektiivsus. Vabadusaste (*df*) on kõigis tingimustes 98. *P*-väärtust on kohandatud 15 võrdluseks.

	Keskliste erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
--	--------------------	--------------	---------------------	-----------------	--------------------------

* $p < .05$, ** $p < .01$, *** $p < .001$.

Tabel 3

Inimhindajate järetestide tulemused vastikuse tingimuses kuue baasemotsiooni tuvastustäpsuse erinevustes

		Keskliste erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	0.458	0.046	10.044	2.412	< .001***
	Kurbus	0.386	0.052	7.448	2.030	< .001***
	Viha	0.411	0.042	9.757	2.165	< .001***
	Üllatus	0.490	0.038	13.025	2.577	< .001***
	Hirm	0.480	0.039	12.403	2.527	< .001***
Rõõm	Kurbus	-0.073	0.027	-2.650	-0.382	0.056
	Viha	-0.047	0.025	-1.914	-0.247	0.293
	Üllatus	0.031	0.017	1.831	0.165	0.293
	Hirm	0.022	0.019	1.148	0.115	0.508
Kurbus	Viha	0.026	0.025	1.006	0.135	0.508
	Üllatus	0.104	0.021	5.062	0.546	< .001***
	Hirm	0.094	0.019	4.909	0.497	< .001***
Viha	Üllatus	0.078	0.015	5.293	0.412	< .001***
	Hirm	0.069	0.015	4.454	0.362	< .001***
Üllatus	Hirm	-0.009	0.007	-1.444	-0.050	0.456

Märkused. Coheni *d* = efektiivsus. Vabadusaste (*df*) on kõigis tingimustes 96. *P*-väärtust on kohandatud 15 võrdluseks.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Tabel 4

FaceReaderi järetestide tulemused rõõmu tingimuses kuue baasemotsiooni tuvastustäpsuse erinevustes

		Keskliste erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	-0.663	0.026	-25.508	-6.040	< .001***
	Kurbus	-0.002	0.004	-0.431	-0.014	1.000
	Viha	-0.032	0.007	-4.722	-0.288	< .001***
	Üllatus	-0.029	0.007	-4.320	-0.264	< .001***
	Hirm	0.008	0.003	2.890	0.072	0.014*
Rõõm	Kurbus	0.662	0.027	24.708	6.026	< .001***
	Viha	0.632	0.029	21.672	5.752	< .001***
	Üllatus	0.634	0.026	24.229	5.776	< .001***
	Hirm	0.671	0.026	26.060	6.112	< .001***
Kurbus	Viha	-0.030	0.006	-4.712	-0.274	< .001***
	Üllatus	-0.027	0.007	-4.136	-0.249	< .001***
	Hirm	0.009	0.003	3.383	0.086	0.004**
Viha	Üllatus	0.003	0.009	0.306	0.024	1.000
	Hirm	0.040	0.006	6.197	0.360	< .001***
Üllatus	Hirm	0.037	0.006	6.216	0.336	< .001***

	Keskmete erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
--	-------------------	--------------	---------------------	-----------------	--------------------------

Märkused. Coheni *d* = efektisuurus. Vabadusaste (*df*) on kõigis tingimustes 97. *P*-väärtust on kohandatud 15 võrdluseks. * *p* < .05, ** *p* < .01, *** *p* < .001.

Tabel 5

FaceReaderi järelestide tulemused hirmu tingimuses kuue baasemotsiooni tuvastatäpsuse erinevustes

		Keskmete erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	-0.235	0.024	-9.709	-1.805	< .001***
	Kurbus	2.887×10 ⁻⁴	0.004	0.070	0.002	1.000
	Viha	-0.071	0.012	-5.966	-0.544	< .001***
	Üllatus	-0.122	0.017	-7.156	-0.932	< .001***
	Hirm	-0.006	0.008	-0.709	-0.045	1.000
Rõõm	Kurbus	0.236	0.025	9.478	1.807	< .001***
	Viha	0.165	0.029	5.658	1.261	< .001***
	Üllatus	0.114	0.029	3.965	0.873	< .001***
	Hirm	0.230	0.026	8.838	1.760	< .001***
Kurbus	Viha	-0.071	0.012	-6.039	-0.546	< .001***
	Üllatus	-0.122	0.017	-7.238	-0.934	< .001***
	Hirm	-0.006	0.008	-0.765	-0.048	1.000
Viha	Üllatus	-0.051	0.022	-2.349	-0.388	0.083
	Hirm	0.065	0.014	4.674	0.498	< .001***
Üllatus	Hirm	0.116	0.018	6.308	0.886	< .001***

Märkused. Coheni *d* = efektisuurus. Vabadusaste (*df*) on kõigis tingimustes 98. *P*-väärtust on kohandatud 15 võrdluseks. ** *p* < .01, *** *p* < .001

Tabel 6

FaceReaderi järelestide tulemused vastikuse tingimuses kuue baasemotsiooni tuvastatäpsuse erinevustes

		Keskmete erinevus	Standardviga	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Vastikus	Rõõm	-0.234	0.032	-7.253	-1.706	< .001***
	Kurbus	0.061	0.012	5.078	0.448	< .001***
	Viha	-0.034	0.018	-1.938	-0.250	0.111
	Üllatus	0.048	0.014	3.365	0.349	0.003**
	Hirm	0.084	0.012	6.891	0.614	< .001***
Rõõm	Kurbus	0.295	0.029	10.027	2.154	< .001***
	Viha	0.199	0.035	5.669	1.456	< .001***
	Üllatus	0.281	0.028	10.111	2.054	< .001***
	Hirm	0.318	0.028	11.367	2.320	< .001***
Kurbus	Viha	-0.096	0.015	-6.357	-0.698	< .001***
	Üllatus	-0.014	0.009	-1.552	-0.100	0.124
	Hirm	0.023	0.005	4.268	0.166	< .001***
Viha	Üllatus	0.082	0.016	5.105	0.599	< .001***
	Hirm	0.118	0.014	8.580	0.864	< .001***
Üllatus	Hirm	0.036	0.007	5.298	0.265	< .001***

Keskmete erinevus	Standardviga	t - statistik	Coheni d	P_{Holm}
-------------------	--------------	--------------------	------------	------------

Märkused. Coheni d = efektisuurus. Vabadusaste (df) on kõigis tingimustes 96. P -väärtust on kohandatud 15 võrdluseks.
* $p < .05$, ** $p < .01$, *** $p < .001$.

**Lisa 2. Masinõppe ning inimeste väljendatud emotsioonide äratundmistäpsuse
võrdlemiseks läbiviidud mitmefaktorilise sõltumatute rühmade
dispersioonanalüüsi järeldestide tulemused (Holmi korrektsiooniga)**

Analüüs on läbi viidud muutujatega emotsioon (6) x meetod (2) x tingimus (3).

Tingimus		Keskliste erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}		
Rõõm	Rõõm Inimene	Hirm Inimene	0.654	29.294	4.185	< .001***	
		Vastikus Inimene	0.643	28.811	4.116	< .001***	
		Viha Inimene	0.655	29.335	4.191	< .001***	
		Kurbus Inimene	0.585	26.215	3.745	< .001***	
		Üllatus Inimene	0.613	27.461	3.923	< .001***	
		Rõõm Masin	0.017	0.746	0.107	1.000	
		Hirm Masin	0.688	30.806	4.401	< .001***	
		Vastikus Masin	0.680	30.450	4.350	< .001***	
		Viha Masin	0.648	29.034	4.148	< .001***	
		Kurbus Masin	0.678	30.381	4.340	< .001***	
		Üllatus Masin	0.651	29.154	4.165	< .001***	
		Hirm Inimene	Vastikus Inimene	-0.011	-0.482	-	1.000
			Viha Inimene	9.325×10 ⁻⁴	0.042	0.006	1.000
			Kurbus Inimene	-0.069	-3.079	0.440	0.086
			Üllatus Inimene	-0.041	-1.833	0.262	1.000
			Rõõm Masin	-0.637	-28.547	4.078	< .001***
	Hirm Masin		0.034	1.512	0.216	1.000	
	Vastikus Masin		0.026	1.157	0.165	1.000	
	Viha Masin		-0.006	-0.260	0.037	1.000	
	Kurbus Masin		0.024	1.087	0.155	1.000	
	Üllatus Masin		-0.003	-0.139	0.020	1.000	
	Vastikus Inimene		Viha Inimene	0.012	0.524	0.075	1.000
			Kurbus Inimene	-0.058	-2.597	0.371	0.340
			Üllatus Inimene	-0.030	-1.351	0.193	1.000
			Rõõm Masin	-0.627	-28.065	4.009	< .001***
			Hirm Masin	0.045	1.995	0.285	1.000
			Vastikus Masin	0.037	1.639	0.234	1.000
		Viha Masin	0.005	0.223	0.032	1.000	
		Kurbus Masin	0.035	1.570	0.224	1.000	
	Viha Inimene	Üllatus Masin	0.008	0.343	0.049	1.000	
		Kurbus	-0.070	-3.121	-	0.076	

Tingimus		Keskmete erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
	Inimene			0.446	
	Üllatus			-	
	Inimene	-0.042	-1.875	0.268	1.000
	Rõõm Masin	-0.638	-28.589	-	< .001***
	Hirm Masin	0.033	1.471	0.210	1.000
	Vastikus Masin	0.025	1.115	0.159	1.000
	Viha Masin	-0.007	-0.301	-	1.000
	Kurbus Masin	0.023	1.046	0.149	1.000
	Üllatus Masin	-0.004	-0.181	-	1.000
Kurbus	Üllatus				
Inimene	Inimene	0.028	1.246	0.178	1.000
	Rõõm Masin	-0.569	-25.468	-	< .001***
	Hirm Masin	0.103	4.591	0.656	< .001***
	Vastikus Masin	0.095	4.236	0.605	0.001**
	Viha Masin	0.063	2.819	0.403	0.179
	Kurbus Masin	0.093	4.166	0.595	0.001**
	Üllatus Masin	0.066	2.940	0.420	0.129
Üllatus	Rõõm Masin	-0.596	-26.714	-	< .001***
Inimene	Hirm Masin	0.075	3.345	0.478	0.036*
	Vastikus Masin	0.067	2.990	0.427	0.112
	Viha Masin	0.035	1.573	0.225	1.000
	Kurbus Masin	0.065	2.920	0.417	0.134
	Üllatus Masin	0.038	1.694	0.242	1.000
Rõõm Masin	Hirm Masin	0.671	30.060	4.294	< .001***
	Vastikus Masin	0.663	29.704	4.243	< .001***
	Viha Masin	0.632	28.287	4.041	< .001***
	Kurbus Masin	0.662	29.634	4.233	< .001***
	Üllatus Masin	0.634	28.408	4.058	< .001***
Hirm Masin	Vastikus Masin	-0.008	-0.356	-	1.000
	Viha Masin	-0.040	-1.772	-	1.000
	Kurbus Masin	-0.009	-0.425	-	1.000
	Üllatus Masin	-0.037	-1.652	-	1.000
Vastikus Masin	Viha Masin	-0.032	-1.417	-	1.000
	Kurbus Masin	-0.002	-0.070	-	1.000
	Üllatus Masin	-0.029	-1.296	-	1.000
Viha Masin	Kurbus Masin	0.030	1.347	0.192	1.000
	Üllatus Masin	0.003	0.120	0.017	1.000
Kurbus	Üllatus Masin	-0.027	-1.226	-	1.000

Tingimus		Keskliste erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}	
	Masin			0.175		
Hirm	Rõõm Inimene	Hirm Inimene	-0.169	-7.606	-	< .001***
		Vastikus Inimene	-0.079	-3.547	-	0.011*
		Viha Inimene	-0.028	-1.269	-	1.000
		Kurbus Inimene	-0.116	-5.219	-	< .001***
		Üllatus Inimene	-0.227	-10.208	-	< .001***
		Rõõm Masin	-0.191	-8.614	-	< .001***
		Hirm Masin	0.038	1.717	-	1.000
		Vastikus Masin	0.044	1.984	-	0.642
		Viha Masin	-0.027	-1.208	-	1.000
		Kurbus Masin	0.044	1.997	-	0.642
	Üllatus Masin	-0.077	-3.487	-	0.014*	
	Hirm Inimene	Vastikus Inimene	0.090	4.058	-	0.002**
		Viha Inimene	0.141	6.337	-	< .001***
		Kurbus Inimene	0.053	2.387	-	0.324
		Üllatus Inimene	-0.058	-2.602	-	0.186
		Rõõm Masin	-0.022	-1.009	-	1.000
		Hirm Masin	0.207	9.323	1.325	< .001***
		Vastikus Masin	0.213	9.590	-	< .001***
		Viha Masin	0.142	6.397	-	< .001***
		Kurbus Masin	0.213	9.603	-	< .001***
Üllatus Masin		0.091	4.118	-	0.001**	
Vastikus Inimene	Viha Inimene	0.051	2.278	-	0.386	
	Kurbus Inimene	-0.037	-1.671	-	1.000	
	Üllatus Inimene	-0.148	-6.660	-	< .001***	
	Rõõm Masin	-0.113	-5.067	-	< .001***	
	Hirm Masin	0.117	5.265	-	< .001***	
	Vastikus Masin	0.123	5.531	-	< .001***	
	Viha Masin	0.052	2.339	-	0.349	
	Kurbus Masin	0.123	5.544	-	< .001***	
	Üllatus Masin	0.001	0.060	-	1.000	
	Viha Inimene	Kurbus Inimene	-0.088	-3.949	-	0.002**
Üllatus Inimene		-0.199	-8.939	-	< .001***	
Rõõm Masin		-0.163	-7.345	-	< .001***	
Hirm Masin		0.066	2.986	-	0.063	

Tingimus		Keskliste erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Kurbus Inimene	Vastikus Masin	0.072	3.253	0.462	0.029*
	Viha Masin	0.001	0.061	0.009	1.000
	Kurbus Masin	0.073	3.266	0.464	0.029*
	Üllatus Masin	-0.049	-2.218	-	0.399
	Üllatus Inimene	-0.111	-4.989	-	< .001***
	Rõõm Masin	-0.075	-3.396	-	0.019*
	Hirm Masin	0.154	6.936	0.483	< .001***
	Vastikus Masin	0.160	7.203	0.986	< .001***
	Viha Masin	0.089	4.010	1.024	0.002**
	Kurbus Masin	0.160	7.216	1.026	< .001***
	Üllatus Masin	0.038	1.731	0.246	1.000
	Rõõm Masin	0.035	1.593	0.226	1.000
	Hirm Masin	0.265	11.925	1.695	< .001***
	Vastikus Masin	0.271	12.192	1.733	< .001***
	Üllatus Inimene	Viha Masin	0.200	8.999	1.279
Kurbus Masin		0.271	12.205	1.735	< .001***
Üllatus Masin		0.149	6.720	0.955	< .001***
Hirm Masin		0.230	10.331	1.468	< .001***
Vastikus Masin		0.235	10.598	1.506	< .001***
Rõõm Masin	Viha Masin	0.165	7.406	1.053	< .001***
	Kurbus Masin	0.236	10.611	1.508	< .001***
	Üllatus Masin	0.114	5.127	0.729	< .001***
	Vastikus Masin	0.006	0.267	0.038	1.000
	Viha Masin	-0.065	-2.926	-	0.073
Hirm Masin	Kurbus Masin	0.006	0.280	0.040	1.000
	Üllatus Masin	-0.116	-5.205	-	< .001***
	Viha Masin	-0.071	-3.193	-	0.033*
	Kurbus Masin	2.887×10 ⁻⁴	0.013	0.454	1.000
	Üllatus Masin	-0.122	-5.471	-	< .001***
Vastikus Masin	Kurbus Masin	0.071	3.206	0.456	0.033*
	Üllatus Masin	-0.051	-2.279	-	0.386
	Üllatus Masin	-0.122	-5.484	-	< .001***
	Hirm Inimene	0.022	0.972	0.140	1.000
	Vastikus Inimene	-0.458	-20.428	-	< .001***
Viha Inimene	Viha Inimene	-0.047	-2.095	-	0.798

Tingimus		Keskmete erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
	Kurbus Inimene	-0.073	-3.234	- 0.464	0.037*
	Üllatus Inimene	0.031	1.393	0.200	1.000
	Rõõm Masin	-0.244	-10.890	- 1.564	< .001***
	Hirm Masin	0.073	3.263	0.468	0.035*
	Vastikus Masin	-0.011	-0.484	- 0.070	1.000
	Viha Masin	-0.045	-2.009	- 0.289	0.936
	Kurbus Masin	0.051	2.251	0.323	0.586
	Üllatus Masin	0.037	1.644	0.236	1.000
Hirm Inimene	Vastikus Inimene	-0.480	-21.401	- 3.073	< .001***
	Viha Inimene	-0.069	-3.067	- 0.440	0.063
	Kurbus Inimene	-0.094	-4.207	- 0.604	< .001***
	Üllatus Inimene	0.009	0.421	0.060	1.000
	Rõõm Masin	-0.266	-11.863	- 1.703	< .001***
	Hirm Masin	0.051	2.290	0.329	0.552
	Vastikus Masin	-0.033	-1.456	- 0.209	1.000
	Viha Masin	-0.067	-2.982	- 0.428	0.081
	Kurbus Masin	0.029	1.279	0.184	1.000
	Üllatus Masin	0.015	0.672	0.096	1.000
Vastikus Inimene	Viha Inimene	0.411	18.334	2.633	< .001***
	Kurbus Inimene	0.386	17.194	2.469	< .001***
	Üllatus Inimene	0.490	21.821	3.133	< .001***
	Rõõm Masin	0.214	9.538	1.370	< .001***
	Hirm Masin	0.532	23.691	3.402	< .001***
	Vastikus Masin	0.448	19.944	2.864	< .001***
	Viha Masin	0.413	18.419	2.645	< .001***
	Kurbus Masin	0.509	22.680	3.257	< .001***
	Üllatus Masin	0.495	22.073	3.169	< .001***
Viha Inimene	Kurbus Inimene	-0.026	-1.140	- 0.164	1.000
	Üllatus Inimene	0.078	3.488	0.501	0.016*
	Rõõm Masin	-0.197	-8.796	- 1.263	< .001***
	Hirm Masin	0.120	5.357	0.769	< .001***
	Vastikus Masin	0.036	1.611	0.231	1.000
	Viha Masin	0.002	0.085	0.012	1.000
	Kurbus Masin	0.098	4.346	0.624	< .001***
	Üllatus Masin	0.084	3.739	0.537	0.007**

Tingimus		Keskliste erinevus	<i>t</i> -statistik	Coheni <i>d</i>	<i>P</i> _{Holm}
Kurbus Inimene	Üllatus Inimene	0.104	4.627	0.664	< .001***
	Rõõm Masin	-0.172	-7.656	-	< .001***
	Hirm Masin	0.146	6.497	1.099	< .001***
	Vastikus Masin	0.062	2.750	0.933	0.162
	Viha Masin	0.027	1.225	0.395	1.000
	Kurbus Masin	0.123	5.486	0.176	< .001***
	Üllatus Masin	0.109	4.879	0.788	< .001***
Üllatus Inimene	Rõõm Masin	-0.276	-12.283	-	< .001***
	Hirm Masin	0.042	1.870	1.764	1.000
	Vastikus Masin	-0.042	-1.877	-	1.000
	Viha Masin	-0.076	-3.402	-	0.022*
	Kurbus Masin	0.019	0.858	0.489	1.000
	Üllatus Masin	0.006	0.251	0.123	1.000
	Rõõm Masin	0.318	14.153	0.036	< .001***
Hirm Masin	Vastikus Masin	0.234	10.406	2.032	< .001***
	Viha Masin	0.199	8.881	1.494	< .001***
	Kurbus Masin	0.295	13.142	1.275	< .001***
	Üllatus Masin	0.281	12.535	1.887	< .001***
	Vastikus Masin	-0.084	-3.747	1.800	0.007**
	Viha Masin	-0.118	-5.272	-	< .001***
	Kurbus Masin	-0.023	-1.011	0.538	1.000
Vastikus Masin	Üllatus Masin	-0.036	-1.618	-	1.000
	Viha Masin	-0.034	-1.525	0.232	1.000
	Kurbus Masin	0.061	2.736	-	0.219
	Üllatus Masin	0.048	2.128	0.393	0.163
Viha Masin	Kurbus Masin	0.096	4.261	0.306	0.768
	Üllatus Masin	0.082	3.653	0.612	< .001***
Kurbus Masin	Üllatus Masin	-0.014	-0.607	0.525	0.009**
				0.087	1.000

Märkused. *t* = *t*-statistik; Coheni *d* = efektiivsus. Vabadusaste (*df*) on kõigis tingimustes 3492 ja standardviga 0,22. *P*-väärtust on kohandatud 66 võrdluseks. * *p* < .05, ** *p* < .01, *** *p* < .001. Huvipakkuvad erinevused on esitatud rasvases kirjas.

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.

/Kaisa Viisileht/