

Automatische stressdetectie met behulp van Virtual Reality

Een volgende stap in de behandeling van tinnitus



Algemene informatie

Naam: Joyce Rijs

Functie: Technical Medicine masterstudente TU Delft, Erasmus MC en LUMC

E-mail: j.rijs@student.tudelft.nl

Gegevens stage

Bedrijf: NeuroVR Bv.

Startdatum: 13-3-2023

Einddatum: 12-5-2023

Begeleiding

Hoofdbegeleider: Marcel Geraeds, eigenaar NeuroVR

LDE begeleider: Dr. Ir. Jan de Laat

Met speciale dank aan House of Secrets en CBoost.

Dit verslag is geschreven als overdrachtsdocument voor een volgende Technical Medicine stagiair(e) bij NeuroVR. Het is daarom in het Nederlands en er is uitgegaan van enige voorkennis met betrekking tot machine learning. Het verslag is niet bedoeld voor wetenschappelijke publicatie. Bij vragen mag er altijd contact opgenomen worden!

1. Samenvatting

Introductie: NeuroVR heeft als doel tinnitus te verlichten met behulp van VR. Stress is een veroorzaker van tinnitus en kan bestaande klachten ook heviger maken. Een volgende stap binnen NeuroVR is dan ook om stressdetectie in VR toe te voegen. Wanneer stress wordt herkend kunnen er stress-reducerende oefeningen worden aangeboden om de tinnitusklachten te verlichten. Het doel van dit project was om te onderzoeken of automatische stressdetectie met behulp van VR-parameters mogelijk is.

Methode: In VR is het mogelijk om lichaams-, gezichts- en oogparameters te meten. Er is data verzameld van gezonde proefpersonen terwijl zij in rust en stress een spel spelen in VR. Er is een statistische analyse uitgevoerd op de verzamelde data om de VR-parameters te vinden die significant verschillen tussen de groepen. Vervolgens is er een machine learning algoritme ontwikkeld om automatisch stress te herkennen.

Resultaten: In totaal hebben 20 proefpersonen deelgenomen aan het onderzoek. Enkele lichaams- en oogparameters bleken significant verschillend ($p < 0.05$) tussen beide groepen. Er is een Random Forest (RF) ontwikkeld voor stressdetectie. Na cross-validatie zijn de accuraatheid, sensitiviteit en specificiteit van het model respectievelijk 92%, 95% en 89%.

Discussie: Concluderend is uit dit onderzoek gebleken dat het mogelijk is om met een machine learning model stress te detecteren met behulp van VR. Er wordt geadviseerd om vervolgonderzoek uit te voeren met een grotere, meer heterogene groep proefpersonen met tinnitus om het model aan te passen aan de doelgroep.

2. Introductie

Patiënten met tinnitus horen geluiden die er eigenlijk niet zijn, zoals fluiten, piepen of ruisen. Tinnitus komt voor bij ongeveer twee miljoen Nederlanders. Hiervan heeft tien procent er veel last van, bij tienduizenden mensen leidt dit zelfs tot psychosociale problemen. Tinnitus kan ontstaan door stress, een langdurige blootstelling aan hard geluid en afwijkingen aan het gehoororgaan, de hersenen of het kaakgewricht (1). Er zijn enkele behandelingen beschikbaar voor tinnitus patiënten, maar geen enkele behandeling kan de tinnitus volledig uitzetten (2). NeuroVR heeft als doel tinnitus te verlichten met behulp van VR. In de VR-wereld zijn er mogelijkheden om het brein te leren het tinnitus-signaal los te laten.

Stress is niet alleen een veroorzaker van tinnitus, maar kan de bestaande tinnitus ook heviger maken. Tevens wordt binnen NeuroVR gedacht dat het brein in stress minder ontvankelijk is om te leren de tinnitus los te laten. Een volgende stap binnen NeuroVR is dan ook om stressdetectie in VR toe te voegen. Wanneer stress wordt herkend zullen er in de VR-wereld ontspanningsoefeningen worden aangeboden zodat de stress en daarmee de tinnitusklachten worden verlicht (3).

Het doel van dit onderzoek was om data te verzamelen voor het ontwikkelen van een machine learning algoritme voor stressdetectie met behulp van VR. Vanwege de korte tijdsduur van dit project is gekozen voor een try-out met gezonde proefpersonen om te onderzoeken of het überhaupt mogelijk is om stress te detecteren met behulp van VR. Indien de resultaten van dit onderzoek positief blijken, zal er een groter onderzoek opgezet worden om data te verzamelen van tinnitus patiënten.

Bij stress wordt het sympathische zenuwstelsel geactiveerd. Deze activatie zorgt voor een verhoging van de hartslag, elektrodermale activiteit en ademhalingsfrequentie. Omdat er minder activiteit is van het parasympathische zenuwstelsel verlaagt de hartslagvariabiliteit. Verder is er een verhoogd niveau van cortisol in het bloed te detecteren (4,5). Deze parameters zijn echter niet direct te meten met VR-sets. Er zouden losse sensoren of invasieve bloedmetingen nodig zijn, wat niet gewenst is. Met de VR-set beschikbaar via NeuroVR (Oculus Quest Pro) is het echter mogelijk om verschillende lichaams-, gezichts- en oogparameters te meten. Er zijn meerdere studies geweest die op basis van deze parameters een stressdetectie algoritme hebben ontwikkeld (6–8). Deze studies zijn uitgevoerd met videobeelden van deelnemers. In de huidige studie werd gepoogd deze resultaten te reproduceren met VR-parameters, zodat er geen videobeelden van de deelnemers gemaakt hoeven te worden.

3. Methode

Om een antwoord te vinden op de onderzoeksvraag is er een onderzoek opgezet om data te verzamelen van gezonde proefpersonen terwijl zij in rust en stress een spel spelen in VR. Er is een statistische analyse uitgevoerd op de verzamelde data om de parameters te vinden die significant verschillen tussen de twee groepen. Vervolgens is er een machine learning algoritme ontwikkeld om automatisch stress te herkennen.

3.1 Data verzameling

3.1.1 VR-parameters

Het is met de Oculus Quest Pro mogelijk om drie types parameters te meten tijdens het spelen van een VR-spel (9).

3.1.1.1 Lichaamsparameters

De VR-set kan detecteren waar het hoofd en de handen zich bevinden in de ruimte. De posities en rotaties kunnen meerdere keren per seconden opgeslagen worden. Vanuit deze posities en rotaties is het mogelijk om bijvoorbeeld de snelheid en de versnelling van een persoon af te leiden. Uit eerder onderzoek blijkt dat gestreste personen vaker en sneller met het hoofd bewegen dan personen in rust (10).

3.1.1.2 Gezichtsparameters

Eerdere studies laten zien dat stress op het menselijk gezicht onder andere te herkennen is aan de frequentie van wenkbrauwbewegingen, het aantal oogknippers en symmetrie van lipbewegingen (8,11). Met de VR-set zijn 63 gezichtsuitdrukkingen te meten, waaronder de beweging van de wenkbrauwen, oogleden en lippen. Deze uitdrukkingen krijgen allemaal een getal tussen 0 en 1,

waarbij 0 betekent dat de uitdrukking niet aanwezig is en 1 een maximale beweging betekent. In het geval van een oogknipper betekent 0 bijvoorbeeld dat het oog open is, 0.5 dat het oog half gesloten is en 1 dat het oog volledig gesloten is. Deze gezichtsparameters kunnen ook meerdere keren per seconden opgeslagen worden, waardoor het bijvoorbeeld mogelijk is om de snelheid, frequentie en symmetrie te berekenen.

3.1.1.3 Oogparameters

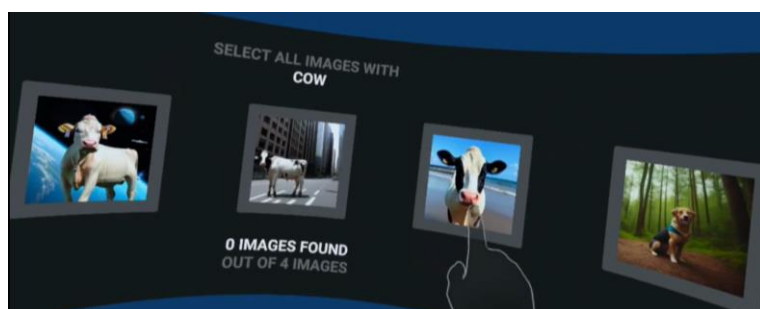
Tot slot is het mogelijk om te detecteren waar iemand naar kijkt tijdens het spelen van een VR-spel. Eerder onderzoek suggereert dat er een verschil is tussen de verdeling van de blik bij gestreste personen ten opzichte van personen in rust (8,12).

3.1.2 Proefpersonen

Voor het verzamelen van data is een groep van 20 proefpersonen zonder tinnitus benaderd. Er is een onderzoeksprotocol geschreven (Bijlage 1) en voorgelegd aan Dr.ir. Jan de Laat. Omdat er sprake was van een try-out met gezonde proefpersonen zonder consequenties ten aanzien van gezondheid was er geen oordeel nodig vanuit de medisch-ethische commissie. De proefpersonen werden vooraf geïnformeerd over hoe het onderzoek zou verlopen. Er werd ook aangegeven dat proefpersonen te allen tijde konden stoppen met het onderzoek door de VR-bril af te zetten.

3.1.3 Opzet onderzoek

Proefpersonen hebben het rust- en stressprotocol beide ondergaan. De helft van de proefpersonen heeft eerst rust en dan stress ondergaan, en de andere helft andersom (gerandomiseerd). Er werd rust of stress opgewekt, waarna de deelnemers een spel werd aangeboden waarbij VR-parameters werden gemeten. Het spel dat aan de proefpersonen werd aangeboden was een vorm van een *Completely Automated Public Turing test to tell Computers and Humans Apart* (Captcha). De opdracht in dit spel is om bepaalde objecten te herkennen in plaatjes en deze aan te klikken (zie Figuur 1). Dit spel werkt intuïtief en wekt geen stress op. De proefpersonen zagen negen plaatjes om zich heen en kregen bijvoorbeeld de opdracht om de plaatjes met een koe aan te klikken. Na het aanklikken van de vier correcte plaatjes werd een volgende opdracht aangeboden.



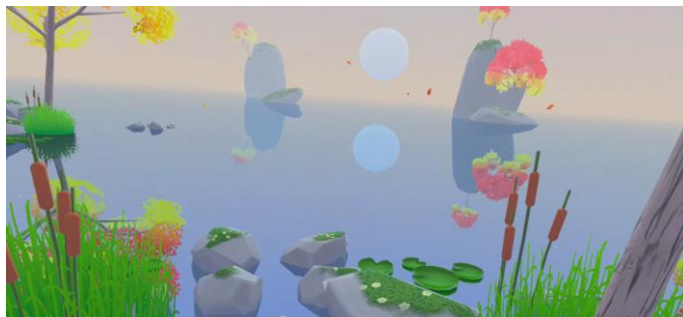
Figuur 1: Een 2D-weergave van Captcha in VR.

Tussen de protocollen kregen de proefpersonen een half uur rust. De VR-wereld is ontwikkeld in samenwerking met Sjoerd van Acker van House of Secrets.

3.1.3.1 Rust protocol

- De proefpersoon kreeg de VR-bril op.
- De proefpersoon zat vijf minuten in een rustige, ontspannen VR omgeving (zie Figuur 2).

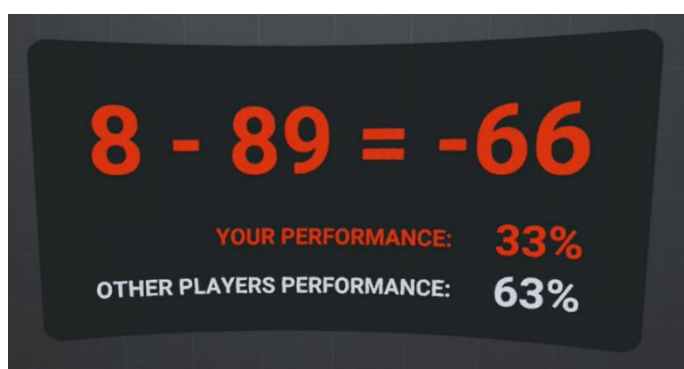
- Captcha werd gestart in VR:
 - De proefpersoon mocht tijdens de eerste opdracht het spel uitproberen.
 - Vanaf de tweede opdracht werd het spel nog drie minuten gespeeld en werden er gedurende drie minuten parameters opgenomen.
 - Het spel en de meting stopten na drie minuten.



Figuur 2: Een 2D-weergave van de VR-rustomgeving.

3.1.3.2 Stress protocol

- De proefpersoon kreeg de VR-bril op.
- De proefpersoon deed vijf minuten een stress-inductietest in VR. Deze test is ontwikkeld en gevalideerd door Rodrigues et al. (13) en bestaat uit moeilijke rekenvragen die binnen enkele seconden opgelost moeten worden (Figuur 3). Het doel van deze test was om stress op te wekken. Tijdens deze test werden er nog geen parameters gemeten. Naar verwachting waren de proefpersonen na de stress-inductietest nog zeker tien minuten in een staat van stress, zoals onderzocht voor een vergelijkbare stress-inductietest in VR (14). Hierdoor verkeerden ze tijdens Captcha nog in een stress-staat.
- Captcha werd gestart in VR:
 - De proefpersoon mocht tijdens de eerste opdracht het spel uitproberen.
 - Vanaf de tweede opdracht werd het spel nog drie minuten gespeeld en werden er gedurende drie minuten parameters opgenomen.
 - Het spel en de meting stopten na drie minuten.



Figuur 3: Een 2D-weergave van de stress-inductietest in VR.

Een uitgebreidere beschrijving over de stappen genomen tijdens de uitvoering van het onderzoek is te vinden in Bijlage 2.

3.1.4 Verzamelde data

De leeftijd en het geslacht zijn per proefpersoon genoteerd. Er is voor beide keren dat de proefpersoon Captcha heeft gespeeld een Excel-bestand opgeslagen met informatie over de verschillende VR-parameters. Deze parameters zijn 70 keer per seconde gemeten gedurende drie minuten. Naast de parameters omschreven in kopje 3.1.1 is er ook opgeslagen hoe vaak proefpersonen correcte en foute plaatjes hebben aangeklikt tijdens het spel.

3.2 Dataverwerking

De dataverwerking is gedaan in Python scripts ontwikkeld in Visual Studio Code versie 1.75.0. GitHub is gebruikt voor versiebeheer. De scripts kunnen hier worden gevonden:

<https://github.com/joycerijs/NeuroVR>

3.2.1 Voorbewerking data

De metingen van de VR-parameters over drie minuten zijn samengevat in verschillende variabelen. Van de hoofd- en handpositie en hoofd-, hand- en oogrotatie zijn de volgende variabelen berekend:

- Standaard deviatie
- Gemiddelde snelheid
- Standaard deviatie van de snelheid
- Gemiddelde versnelling

Van de gezichtsparemeters zijn de volgende variabelen berekend:

- Standaard deviaties
- Gemiddelde snelheid
- Standaard deviatie van de snelheid
- Frequentie
- Symmetrie (indien parameter voor links en rechts aanwezig was)

De aantallen correcte en foute antwoorden tijdens het spel werden ook als variabelen toegevoegd. Verdere specifieke voorbewerkingsstappen zijn beschreven in het Python script.

3.2.2 Statistische analyse

Om te onderzoeken welke variabelen significant verschillen tussen de rust- en stressdata is er een gepaarde t-toets uitgevoerd voor iedere variabele. Een t-toets is een parametrische test die de gemiddeldes tussen twee groepen vergelijkt (15). Een p-waarde < 0.05 werd als significant beschouwd.

3.2.3 Model

Vervolgens is er een machine learning model ontwikkeld om stress te detecteren. Hiervoor is het Scikit-learn pakket in Python gebruikt (16). Het ontwikkelde model is een Random Forest (RF). Dit model bestaat uit een collectie van beslisbomen (17). Een RF kan omgaan met complexe data, eventuele uitschieters en ruis. Ook is het voor een RF niet nodig om de data te schalen. Een RF heeft een natuurlijke manier om de variabelen te selecteren die van belang zijn omdat iedere beslisboom wordt gemaakt met een andere steekproef uit de data en de variabelen. Het aantal beslisbomen is

gebaseerd op het meest optimale resultaat van de leercurves (zie 3.2.5). Voor de overige RF-parameters zijn standaardwaardes gebruikt.

De data zijn opgesplitst in een train- en testset. De trainset bestond uit 95% van de data en is gebruikt om het model te laten leren. De testset (5% van de data) is gebruikt om het model te evalueren. Er is gekozen voor een cross-validatie waarbij de data iedere keer in een andere train- en testset werden opgesplitst (18). Op deze manier is de prestatie van het model op ongeziene data geëvalueerd.

Het model is geëvalueerd op basis van accuraatheid (hoeveel procent van de testset is correct geclassificeerd?) en een contingentietabel waarbij de aantallen terecht-positieven (tp), terecht-negatieven (tn), fout-positieven (fp) en fout-negatieven (fn) zijn weergegeven. Vanuit deze tabel zijn de sensitiviteit en specificiteit van het model berekend. De sensitiviteit staat voor het vermogen van het model om stress aan te tonen bij gestreste proefpersonen (Formule 1), de specificiteit staat voor het vermogen rust aan te tonen bij proefpersonen in rust (Formule 2). Een hoge sensitiviteit wordt in dit onderzoek als belangrijker beschouwd dan een hoge specificiteit, omdat het belangrijk is om stress niet te missen. Dit heeft mogelijk grotere consequenties dan het foutief classificeren van een persoon in rust.

$$(1) \text{ Sensitiviteit} = \frac{tp}{(tp+fn)}$$

$$(2) \text{ Specificiteit} = \frac{tn}{(tn+fp)}$$

Er is gekozen om voor deze try-out alleen een RF-model te ontwikkelen, omdat er aangenomen werd dat dit een passend model voor het huidige probleem zou zijn. Indien dit model onvoldoende resultaat zou geven, zouden er ook andere type modellen worden getest. Een onvoldoende resultaat werd gedefinieerd als minder dan 75% accuraatheid, aangezien eerder ontwikkelde modellen op basis van videobeelden tussen 75% en 91% accuraat waren (6–8).

3.2.5 Leercurves

Een mogelijk probleem bij machine learning modellen is dat het model te veel meebeweegt met de spreiding van de trainset en daardoor niet goed kan generaliseren op ongeziene data. Dit wordt *overfitting* genoemd. Ook is het tegenovergestelde juist mogelijk: het model beweegt dan niet genoeg mee met de spreiding in de trainset en is *underfit*. Om te onderzoeken of er sprake was van één van deze fenomenen in het ontwikkelde RF-model zijn er leercurves gemaakt. Leercurves laten de prestatie van het model zien op basis van het aantal voorbeelden in de trainset (19). De trainset is opgesplitst in een train- en cross-validatieset. Het model heeft een goede *fit* wanneer de cross-validatiescore naar een stabiel punt gaat dat dicht bij de train score ligt. De leercurves zijn gebruikt om te bepalen welk aantal beslisbomen de beste *modelfit* gaf. De waardes 50, 100, 150 en 200 zijn getest. Het aantal beslisbomen wordt een hyperparameter genoemd. Het kleinste aantal beslisbomen waarbij een goede *modelfit* werd gevonden werd gekozen als meest optimale hyperparameter en gebruikt voor het model.

3.2.4 Sub-analyses

Er zijn verschillende sub-analyses uitgevoerd. Zo is onderzocht of het model betere resultaten gaf indien de VR-parameters werden samengevat over kleinere stukken data (30, 60 en 90 seconden). Iedere dataset van een proefpersoon werd opgedeeld in verschillende *windows*. Bij het opknippen in 30 seconden waren er bijvoorbeeld zes *windows* per dataset. De variabelen beschreven in 3.2.1

waren vervolgens per *window* berekend. Om *bias* te voorkomen, zijn de *windows* van dezelfde personen bij elkaar gehouden in de train- en testsets. De hypothese was dat het model bij het opknippen in *windows* beter zou presteren omdat het dan van een groter aantal voorbeelden kon leren in de trainset.

Ook is er onderzocht welk resultaat het model gaf indien er niet naar drie minuten aan data werd gekeken, maar alleen naar het eerste x-aantal seconden. Proefpersonen hebben in dit experiment drie minuten een spel gespeeld. Voor de implementatie is het voordelig om de proefpersonen zo kort mogelijk het spel te laten spelen, waarbij het model nog steeds voldoende resultaat geeft. De accuraatheid van het model over de verschillende aantallen seconden is weergegeven in een grafiek.

Als laatste is het model opnieuw getraind met alleen de hoofd- en handvariabelen. Dit is gedaan omdat de oudere versie van de Oculus-bril geen oog- en gezichtsparameters kan meten. Indien het model met alleen hoofd- en handvariabelen al voldoende resultaat zou geven, zou in een volgend experiment mogelijk de oudere bril gebruikt kunnen worden.

4. Resultaten

4.1 Proefpersonen

In totaal hebben 20 proefpersonen deelgenomen aan het onderzoek, waaronder 9 vrouwen. De gemiddelde leeftijd van de proefpersonen was 25 jaar. Alle proefpersonen hebben beide protocollen ondergaan. Bij proefpersoon 15 zijn de rustdata niet goed opgeslagen. Er zijn uiteindelijk dus 20 Excel-bestanden met stressdata en 19 met rustdata opgeslagen.

4.2 Statistische analyse

14 variabelen bleken significant verschillend tussen beide groepen (P -waarde < 0.05). Een tabel met de significante variabelen en bijbehorende P -waardes is te vinden in Tabel 1 in Bijlage 3. Onder de significante variabelen vallen hand-, hoofd- en oogvariabelen, maar geen enkele gezichtsvariabele. Ook het aantal correcte en foute antwoorden bleek niet significant te verschillen.

4.3 Model

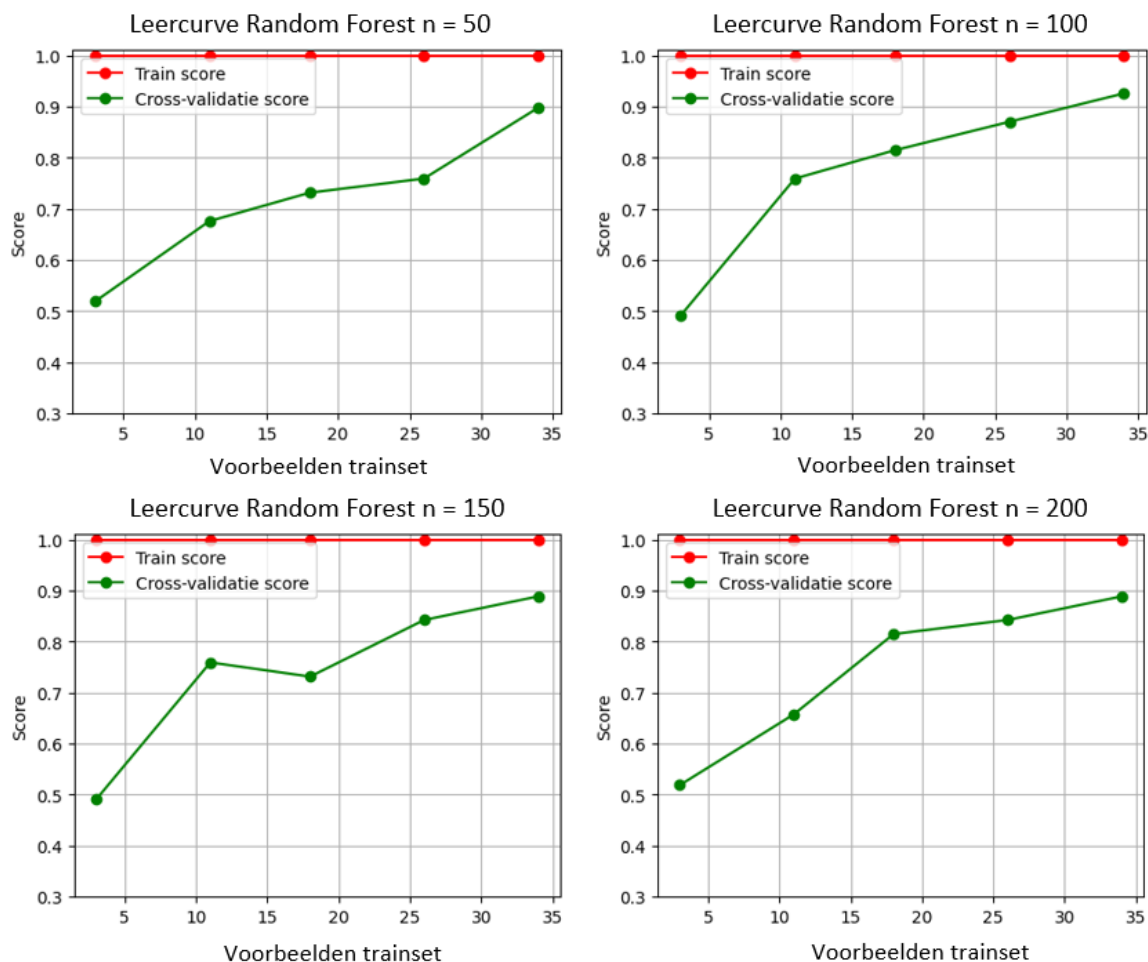
Er is een RF ontwikkeld met de waarde 100 voor het aantal beslisbomen (zie 4.4). Per trainset was er een klein verschil in welke variabelen significant verschillend waren. Dit waren altijd hand-, hoofd- en oogvariabelen, daarom is ervoor gekozen de RF te trainen met alle hand-, hoofd- en oogvariabelen samengevat over drie minuten. Het model heeft gemiddeld 92% van de testset correct geclassificeerd na cross-validatie. Tabel 2 beschrijft de gemaakte classificaties van het model in een contingentietabel na cross-validatie. De sensitiviteit is 95% en de specificiteit is 89%.

	Stress aanwezig	Rust aanwezig
Classificatie stress	19 (tp)	2 (fp)
Classificatie rust	1 (fn)	17 (tn)

Tabel 2: Classificaties van het machine learning model na cross-validatie. *tp* = terecht-positieven, *tn* = terecht-negatieven, *fp* = fout-positieven, *fn* = fout-negatieven

4.4 Leercurves

Er zijn leercurves ontwikkeld waarbij voor het aantal beslisbomen de waarden 50, 100, 150 en 200 werden gedefinieerd (Figuur 4). Een score van 1 geeft een 100% correcte voorspelling aan. Op de x-as staat het aantal voorbeelden waarop het model is getraind. Er is te zien dat de cross-validatiescore stijgt naarmate het model is getraind met meer voorbeelden. Bij een waarde van 100 beslisbomen werd als de beste *modelfit* beschouwd, omdat de cross-validatiescore het dichtst bij de train score komt.



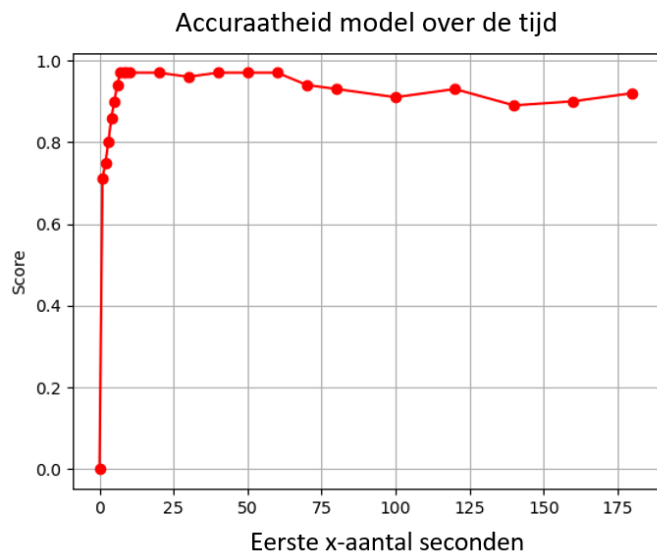
Figuur 4: Leercurves met verschillende aantallen beslisbomen in de RF ($n = 50, 100, 150$ en 200).

4.5 Sub-analyses

Als eerste is naar de prestatie van het model gekeken wanneer de data werden opgeknipt in kleinere stukken dan drie minuten, waardoor er meer trainvoorbeelden beschikbaar waren. De prestatie van het model werd geleidelijk minder naarmate de data werden opgeknipt in kleine stukken (van 92% gemiddelde accuraatheid voor drie minuten naar 68% voor 30 seconden). Er is daarom gekozen om de variabelen niet over kleinere stukken samen te vatten.

Als tweede is onderzocht hoeveel seconden aan data er minimaal nodig was voor een accurate modelvoorspelling. De resultaten zijn te vinden in Figuur 5. Een score van 1 geeft een 100% correcte voorspelling aan. Uit dit figuur blijkt dat het model vanaf enkele seconden al een accurate

classificatie kan maken. De gemiddelde accuraatheid stijgt tot 97% op zeven seconden en blijft daarna stabiel. Na een minuut daalt de gemiddelde accuraatheid langzaam van 97% bij 60 seconden naar 92% bij 180 seconden.



Figuur 5: De gemiddelde accuraatheid van het model op basis van eerste x-aantal seconden data.

Als laatste is het model nog een keer getraind met alleen de hoofd- en handvariabelen. De oogvariabelen zijn in dit model weggelaten. De resultaten zijn vergelijkbaar met het eerder ontwikkelde model, opnieuw is gemiddeld 92% van de testset correct geïdentificeerd na cross-validatie.

5. Discussie

In dit onderzoek zijn VR-data verzameld van proefpersonen in rust en stress voor de ontwikkeling van een machine learning model. Hand-, hoofd- en oogvariabelen bleken significant verschillend tussen de twee groepen, maar gezichtsvariabelen niet. Bij eerder uitgevoerde onderzoeken waar gezichtsvariabelen significant verschilden werden videobeelden gemaakt tijdens de stress-inductietesten. Deze bestonden uit het geven van een speech, het voorlezen van tekst of het bekijken van emotie-opwekkende video's (8,11). In het huidige onderzoek werden data opgenomen tijdens het spel volgend op de stress-inductietest. Tijdens het spelen van het spel hoefden proefpersonen hun gezicht niet te bewegen. Er werd geen emotie opgewekt en de proefpersonen hoefden niet te praten. Wellicht is dit een verklaring voor het feit dat gezichtsvariabelen niet verschilden tussen de twee groepen. Een ander opvallend resultaat is dat de standaard deviaties van hand-, hoofd- en oogbewegingen bij rust groter waren dan bij stress. Een mogelijke verklaring hiervoor is dat gestreste mensen het spel meer doelgericht spelen, waardoor er minder variatie in de data te zien is.

Met de hand-, hoofd- en oogvariabelen is een RF-model ontwikkeld dat in 92% van de gevallen een correcte voorspelling maakt. Het model heeft een vergelijkbare, al dan niet hogere accuraatheid dan machine learning modellen ontwikkeld op basis van videobeelden (tussen 75-91%) (6-8). De optimale hyperparameter voor het aantal beslisbomen in de RF is gebaseerd op leercurves. De cross-validatie score was nog niet volledig geconvergeerd naar de train score in deze curves, dit betekent

dat het model mogelijk nog beter zou presteren na het toevoegen van meer voorbeelden in de trainset (19). Hiervoor zou data verzameld moeten worden van een grotere groep proefpersonen. Vanwege het try-out karakter en de relatief korte tijdsduur van dit onderzoek is er één type model getest. Mogelijk kunnen andere modeltypes nog betere resultaten geven.

Wanneer de oog- en gezichtsvariabelen worden weggelaten bij het ontwikkelen van het model blijven de resultaten vergelijkbaar. Het lijkt dus mogelijk om deze test in een vervolgonderzoek ook met de vorige versie van de Oculus VR-bril uit te voeren.

Opvallend is dat het model minder presteert wanneer de data worden opgeknipt in kleinere *windows*, terwijl het goed presteert wanneer alleen de eerste zeven seconden aan data worden meegenomen. Een mogelijke verklaring hiervoor is dat het stressniveau bij proefpersonen na de stress-inductietest langzaam daalt. Het ontwikkelde model classificeert momenteel 'stress' of 'geen stress'. Een vervolgstap zou kunnen zijn om een regressiemodel te ontwikkelen dat een stresspercentage kan detecteren. Voor het labelen van de train- en testdata zouden het eerste aantal seconden na de stress-inductietest 100% stress zijn, dit percentage zou dan langzaam afnemen gedurende het spel. Op deze manier kan rekening worden gehouden met de mogelijke afname van stress.

Het uiteindelijke doel van dit project is om chronische stress te detecteren bij tinnitus patiënten met behulp van VR. Vanwege de korte tijdsduur van de stage is ervoor gekozen om acute stress bij een kleine, homogene groep gezonde proefpersonen te detecteren. Vanwege het positieve resultaat van deze try-out wordt geadviseerd om in vervolgonderzoek een grotere groep proefpersonen met tinnitus te testen van verschillende leeftijden, waarbij er chronische in plaats van acute stress wordt gedetecteerd. Ook kan er dan beter gecontroleerd worden voor factoren die mogelijk van invloed zijn voor het meten van de VR-parameters, zoals cafeïne-inname, tijd van de dag en de grootte van het hoofd. Een verdere vervolgstap is om het effect van stressdetectie en daaropvolgende stressreductie op de tinnitusklachten te onderzoeken.

De automatische herkenning van stress is niet alleen zinvol voor het helpen van tinnitus patiënten, er zijn meer mogelijke toepassingen te bedenken. Een voorbeeld is de detectie van stress bij verstandelijk gehandicapten of demente ouderen die hun stressniveau niet goed zelf kunnen aangeven. Chronische stress kan niet alleen leiden tot een verergering van tinnitusklachten, maar ook tot verschillende inflammatoire, metabolische en neurodegeneratieve ziekten (20). Door het vroeg herkennen van chronische stress kan de ontwikkeling van deze ziekten mogelijk voorkomen worden, wat een verlaging in zorgkosten teweeg kan brengen (21).

Concluderend is uit dit onderzoek gebleken dat het mogelijk is om met een machine learning model stress te detecteren met behulp van VR. Dit is het eerste onderzoek dat VR-parameters in plaats van videobeelden gebruikt voor stressdetectie. Er wordt geadviseerd om vervolgonderzoek uit te voeren met een grotere, meer heterogene groep proefpersonen met tinnitus om het model aan te passen aan de doelgroep. De uitkomsten van dit onderzoek kunnen door NeuroVR worden gebruikt voor het verder ontwikkelen van een volledige behandelmethodede voor tinnitus in VR.

6. Referenties

1. NVKNO. Richtlijn Tinnitus. 2016;
2. Mazurek B, Hesse G, Dobel C, Kratzsch V, Lahmann C, Sattel H. Chronic Tinnitus: Diagnosis and Treatment. *Dtsch Arztebl Int.* 2022 Apr 1;119(13):219–25.
3. Roland LT, Lenze EJ, Hardin FM, Kallogjeri D, Nicklaus J, Wineland AM, et al. Effects of mindfulness based stress reduction therapy on subjective bother and neural connectivity in chronic tinnitus. *Otolaryngol Head Neck Surg.* 2015 May;152(5):919–26.
4. Ulrich-Lai YM, Herman JP. Neural regulation of endocrine and autonomic stress responses. *Nat Rev Neurosci.* 2009 Jun;10(6):397–409.
5. Barry RJ, Furedy JJ. Stimulus intensity and novelty interact in elicitation of the phasic electrodermal orienting response. *International Journal of Psychophysiology.* 1993 May;14(3):249–54.
6. Dinges DF, Rider RL, Dorrian J, McGlinchey EL, Rogers NL, Cizman Z, et al. Optical computer recognition of facial expressions associated with stress induced by performance demands. *Aviat Space Environ Med.* 2005 Jun;76(6 Suppl):B172-82.
7. Gao H, Yuce A, Thiran JP. Detecting emotional stress from facial expressions for driving safety. In: 2014 IEEE International Conference on Image Processing (ICIP). IEEE; 2014. p. 5961–5.
8. Giannakakis G, Pediaditis M, Manousos D, Kazantzaki E, Chiarugi F, Simos PG, et al. Stress and anxiety detection using facial cues from videos. *Biomed Signal Process Control.* 2017 Jan 1;31:89–101.
9. Meta Quest. Movement SDK for Unity. Available from: <https://developer.oculus.com/documentation/unity/move-overview/>.
10. Liao W, Zhang W, Zhu Z, Ji Q. A Real-Time Human Stress Monitoring System Using Dynamic Bayesian Network. 2005.
11. Metaxas D, Venkataraman S, Vogler C. Image-Based Stress Recognition Using a Model-Based Dynamic Face Tracking System. In 2004. p. 813–21.
12. Paletta L, Pszeida M, Nauschnegg B, Haspl T, Marton R. Stress Measurement in Multi-tasking Decision Processes Using Executive Functions Analysis. In 2020. p. 344–56.
13. Rodrigues J, Studer E, Streuber S, Sandi C. IMVEST, an immersive multimodal virtual environment stress test for humans that adjusts challenge to individual's performance. *Neurobiol Stress.* 2021 Nov 1;15.
14. Shibani Y, Diemer J, Brandl S, Zack R, Mühlberger A, Wüst S. Trier Social Stress Test in vivo and in virtual reality: Dissociation of response domains. *International Journal of Psychophysiology.* 2016 Dec;110:47–55.
15. Mishra P, Singh U, Pandey C, Mishra P, Pandey G. Application of student's t-test, analysis of variance, and covariance. *Ann Card Anaesth.* 2019;22(4):407.
16. Pedregosa F, Michel V, Grisel O, Blondel M, Prettenhofer P, Weiss R, et al. Scikit-learn: Machine Learning in Python [Internet]. Vol. 12, *Journal of Machine Learning Research.* 2011. Available from: <http://scikit-learn.sourceforge.net>.
17. Breiman L. Random Forests. Vol. 45. 2001.
18. Berrar D. Cross-Validation. In: *Encyclopedia of Bioinformatics and Computational Biology.* Elsevier; 2019. p. 542–5.
19. Viering T, Loog M. The Shape of Learning Curves: A Review. *IEEE Trans Pattern Anal Mach Intell.* 2022;1–20.

20. Mariotti A. The effects of chronic stress on health: new insights into the molecular mechanisms of brain-body communication. *Future Sci OA*. 2015 Nov;1(3):FSO23.
21. de Weerdt AC, Gouwens S, Koopmanschap MA, van der Meer A, Kommer GJ. Aandoeningen op basis van zorguitgaven. 2022.
22. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Front Hum Neurosci*. 2020;14:96.

Bijlagen

Bijlage 1: Onderzoeksprotocol

Plan

Titel: Dataverzameling voor het ontwikkelen van een stressdetectie algoritme in Virtual Reality (VR)

Auteurs en uitvoerders: Marcel Geraeds, Jan de Laat, Joyce Rijs

Datum: april en mei 2023

Contactpersoon: Joyce Rijs, j.rijs@student.tudelft.nl, 0646693559

Aanleiding

Patiënten met tinnitus horen geluiden die er eigenlijk niet zijn, zoals fluiten, piepen of ruisen. Tinnitus komt voor bij ongeveer twee miljoen Nederlanders. Hiervan heeft tien procent er veel last van, bij tienduizenden mensen leidt dit zelfs tot psychosociale problemen. Tinnitus kan ontstaan door een langdurige blootstelling aan hard geluid, stress en afwijkingen aan het gehoororgaan, de hersenen of het kaakgewricht (1). Er zijn enkele behandelingen beschikbaar voor tinnitus patiënten, maar geen enkele behandeling kan de tinnitus helemaal uitzetten (2). NeuroVR heeft als doel tinnitus te verlichten met behulp van VR. In de VR-wereld zijn er mogelijkheden om het brein te leren het tinnitus signaal los te laten. Een volgende stap binnen dit project is om stressdetectie in VR toe te voegen. Stress is niet alleen een veroorzaker van tinnitus, maar kan de bestaande tinnitus ook heviger maken. Daarom is het doel van dit onderzoek om data te verzamelen voor het ontwikkelen van een algoritme dat kan herkennen wanneer patiënten gestrest zijn m.b.v. VR. Wanneer stress wordt herkend zal er in de VR-wereld een ontspanningsoefening worden aangeboden zodat de stress en daarmee de tinnitusklachten worden verlicht (3).

Achtergrond en eerdere studies

Bij stress wordt het sympathische zenuwstelsel geactiveerd. Deze activatie zorgt voor een verhoging van de hartslag, elektrodermale activiteit en ademhalingsfrequentie. Omdat er minder activiteit is van het parasympatische zenuwstelsel verlaagt de hartslagvariabiliteit. Verder is er een verhoogd niveau van cortisol in het bloed te detecteren (4,5). Deze parameters zijn echter niet direct te meten met VR-sets. Er zouden losse sensoren of invasieve bloedmetingen nodig zijn, wat niet gewenst is. Via de VR-set beschikbaar via NeuroVR (Meta Quest Pro) is het echter mogelijk om verschillende gezichtsparameters te meten, zoals gezichtsuitdrukkingen en oogbewegingen. Ook is het mogelijk om de beweging van het hoofd en de handen te meten. Eerdere studies laten zien dat stress op het menselijk gezicht te herkennen is aan de frequentie van oogbewegingen, het aantal oogknippers en bewegingen van de lippen, de wangen en het hoofd (8). Er zijn meerdere studies geweest die op basis van deze parameters een stressdetectie algoritme hebben ontwikkeld (6,7,13). Het blijkt dus mogelijk om met gezichts- en bewegingsparameters een stressdetectie algoritme te ontwikkelen. Deze studies zijn uitgevoerd met videobeelden van deelnemers. In de huidige studie wordt gepoogd deze resultaten te reproduceren met gezichts- en bewegingsparameters gemeten met VR, zodat er geen videobeelden van de deelnemers gemaakt hoeven te worden.

Relevantie

Tot op heden is er geen behandelmethodede die alle oorzaken van tinnitus aanpakt. NeuroVR heeft als doel om met VR ervoor te zorgen dat tinnitusklachten verminderen. Het detecteren en verlichten van stress is hier een belangrijk onderdeel van. Wanneer een volledige behandeling mogelijk is met

slechts een VR-set kunnen patiënten er zelf thuis mee aan de slag. Dit leidt ertoe dat de behandeling niet arbeidsintensief is wat betreft de hulp die door zorgverleners gegeven moet worden en uiteindelijk over de hele wereld toegankelijk kan worden.

Plan van aanpak

Voor het ontwikkelen van het algoritme zal een groep van 15-20 proefpersonen benaderd worden. We zullen hierbij vrienden en familie zonder tinnitus aanspreken. De proefpersonen worden vooraf geïnformeerd over hoe het onderzoek zal verlopen en hoe lang het zal duren. Er wordt ook aangegeven dat proefpersonen te allen tijde kunnen stoppen met het onderzoek door de VR-bril af te zetten.

Voor het ontwikkelen van het algoritme zijn er data van personen in zowel rust als stress nodig. Deelnemers zullen het rust- en stressprotocol beide ondergaan. De helft van de proefpersonen zal eerst rust en dan stress ondergaan, en de andere helft andersom (gerandomiseerd). Er zal rust of stress opgewekt worden, waarna de deelnemers een game wordt aangeboden waarbij VR-parameters worden gemeten. Tussen de protocollen krijgen de proefpersonen een half uur rust.

Rust protocol

- Proefpersoon krijgt VR-bril op.
- Proefpersoon zit een aantal minuten in een rustige, ontspannen VR omgeving.
- Spel* wordt gestart in VR:
 - o Proefpersoon krijgt uitleg over hoe het spel werkt en mag het even uitproberen.
 - o Als de proefpersoon het spel begrijpt klikt hij/zij op 'start'. De game wordt dan drie minuten gespeeld, en er worden drie minuten parameters opgenomen.
 - o Het spel en de meting stoppen na drie minuten.

Stress protocol

- Proefpersoon krijgt VR-bril op.
- Proefpersoon doet 5 minuten een stress-inductietest in VR. Deze test is ontwikkeld en gevalideerd door Rodrigues et al. (13) en bestaat uit moeilijke rekenvragen die snel opgelost moeten worden. Het doel van deze test is om stress op te wekken. Tijdens deze test worden er nog geen parameters gemeten.
- Spel* wordt gestart in VR:
 - o Proefpersoon krijgt uitleg over hoe het spel werkt en mag het even uitproberen.
 - o Als de proefpersoon het spel begrijpt klikt hij/zij op 'start'. De game wordt dan drie minuten gespeeld, en er worden drie minuten parameters opgenomen.
 - o Het spel en de meting stoppen na drie minuten.

*Het spel dat aan de proefpersonen wordt aangeboden is een vorm van een *Completely Automated Public Turing test to tell Computers and Humans Apart* (Captcha). De opdracht in dit spel is om bepaalde objecten te herkennen in plaatjes en deze aan te klikken. Dit spel werkt intuïtief en wekt geen stress op.

Naar verwachting zijn de proefpersonen na de stress-inductietest nog zeker 10 minuten in een staat van stress, zoals onderzocht voor een andere stress-inductietest in VR (14). Hierdoor zullen ze tijdens

Captcha nog in een stress-staat verkeren. Onze hypothese is dat gestreste personen tijdens dit spel onder andere sneller bewegen en vaker met hun ogen knipperen.

Wanneer de proefpersonen beide protocollen hebben doorlopen, is er genoeg data beschikbaar voor het ontwikkelen van een machine learning algoritme dat kan detecteren of iemand gestrest is tijdens het spelen van de aangeboden game. Als het uit deze try-out mogelijk blijkt om stress te detecteren in VR zal er uiteindelijk een groter onderzoek opgezet worden met tinnitus patiënten. Het ontwikkelde algoritme kan dan uiteindelijk worden gebruikt voor stressdetectie bij tinnitus patiënten met behulp van VR.

Projectomschrijving

Tijdsduur: 1/2 weken

Deelnemers: 15-20 vrienden/familieleden die uit vrije wil mee willen werken aan dit onderzoek

Beoogde uitvoerders: Marcel Geraeds en Joyce Rijs

Expertise: Eigenaar NeuroVR (Marcel Geraeds) en masterstudente Technische Geneeskunde (Joyce Rijs)

Locatie: In een rustige ruimte in Delft of Den Haag

Doel

Zoals omschreven in het plan van aanpak zullen er in deze try-out data worden verzameld voor het ontwikkelen van een machine learning algoritme voor stressdetectie in VR. Als de resultaten uit dit onderzoek positief zijn, zal er een groter onderzoek opgezet worden waarbij data van tinnituspatiënten van alle leeftijden verzameld worden. Deze data zullen ervoor zorgen dat het algoritme wordt aanpast aan de uiteindelijke doelgroep voordat het wordt geïmplementeerd. Het algoritme kan uiteindelijk detecteren wanneer tinnitus patiënten gestrest zijn. Wanneer stress wordt herkend zal er in de VR-wereld een ontspanningsoefening worden aangeboden zodat de stress en daarmee de tinnitusklachten worden verlicht.

Mogelijke risico's

VR staat erom bekend dat het bewegingsziekte kan opwekken. Hierdoor kunnen mensen misselijkheid, hoofdpijn en duizeligheid ervaren. Bewegingsziekte kan echter grotendeels worden voorkomen door de VR-omgeving goed in te richten (22). Voor de ontwikkeling van de VR-omgeving wordt samengewerkt met House of Secrets, een bedrijf met jarenlange ervaring met het ontwikkelen van VR-applicaties. We zullen ervoor zorgen dat de VR-omgeving zo wordt ingericht dat de kans op bewegingsziekte het kleinst is. Mochten proefpersonen wel enig ongemak ervaren kunnen ze de VR-bril afzetten en zal het onderzoek worden gestaakt.

Verder is het mogelijk dat de stress-inductietest bij deelnemers als te stressvol wordt ervaren. Als deelnemers zich niet meer comfortabel voelen kunnen ze te allen tijde de bril afzetten tijdens het onderzoek.

Data opslag

De opgenomen data worden geanonimiseerd opgeslagen in een database samen met de leeftijd en het geslacht van de proefpersonen. Er worden geen video's gemaakt van de deelnemers, er worden slechts getallen opgeslagen die staan voor gezichts- en bewegingsparameters (bijvoorbeeld wanneer iemand zijn wenkbrauw omhoog doet wordt er een 1 opgeslagen, anders een 0).

Bijlage 2: Protocol uitvoering onderzoek

Datum laatste aanpassing: 28-3-2023

Uitleg aan proefpersoon: Ik wil een algoritme schrijven om in VR automatisch te detecteren wanneer mensen gestrest zijn. Hiervoor hebben we data nodig van hoe mensen bewegen in VR tijdens rust en tijdens stress.

Je gaat twee keer een spel doen van drie minuten, een keer terwijl je net vijf minuten in een rustige VR omgeving hebt gezeten, en een keer terwijl je net een moeilijke rekentoets van 5 minuten in VR hebt gedaan. In totaal heb je dus 8 minuten achter elkaar de VR-bril op. Tussen de twee sets zit een half uur pauze.

De gegevens worden anoniem opgeslagen en er worden geen video's van je gemaakt.

Belangrijke instructies:

- Als je je niet meer comfortabel voelt, kun je de bril altijd afzetten
- Graag doen alsof ik niet aanwezig ben in de ruimte; dus niet tegen mij praten tijdens het spel
- Je klikt zometeen op A of B (rust of stress). Je komt dan 5 minuten in de rust of stress ruimte en gaat daarna meteen door naar het spel. Dit spel is een captcha, dat houdt in dat je 9 plaatjes te zien krijgt en je moet de plaatjes aanklikken waar bijvoorbeeld een kat op te zien is. Het is belangrijk dat je met je rechterhand de plaatjes aanklikt. Hoe snel je de plaatjes aanklikt maakt niet uit, het duurt altijd drie minuten.
- In de VR rustruimte mag je zitten, maar tijdens het spel moet je staan. Ook in de stressruimte moet je staan.
- Er zijn explosies en flitsende lichten, dus mocht je trauma's of epilepsie hebben dan kun je beter niet meedoen.
- Je kunt op dingen klikken als een touch screen, er zijn geen controllers nodig. Per spel krijg je nog uitleg over hoe het werkt.

Je start met rust/stress.

Benodigheden

- Meta Quest Pro VR-set
- Kabel voor data-overdracht naar laptop
- Laptop
- Een ruimte met minstens 2x2 meter lege ruimte
- Schoonmaakdoekjes
- Bedankjes proefpersonen

Stappenplan

1. *Connecten met wifi*
2. *Guardian maken in de ruimte*
3. Zet de bril op bij proefpersoon
4. Zorg met het menu dat hij echt goed op het hoofd zit: Apps → guide → pasvorm van headset aanpassen.
5. *Geef bril terug. Ik start het spel op.*
6. Zorgen dat ik via mijn telefoon meekijk (via Meta Quest app).

7. Bril opzetten bij proefpersoon.
8. Uitleggen dat het rode puntje betekent dat ik mee kijk, en laten zien dat er een Guardian is waar ze niet buiten kunnen.
9. *Goed centreren voor de proefpersoon met knop op controller of via app.*
10. Bij rust zorgen dat proefpersoon zit, bij stress zorgen dat hij staat.
11. Kies A of B
12. Overgang naar captcha:
 - a. Zorgen dat proefpersoon staat
 - b. Nog een keer herhalen dat ze hun rechterhand moeten gebruiken
 - c. Nog een keer herhalen dat ze niet mogen praten
 - d. Tijd en bijzonderheden noteren
13. Einde: VR bril af

Half uur pauze: bril opladen!

Einde:

1. Proefpersoon bedanken met chocola
2. Data overdracht naar laptop
3. Schoonmaken bril

Bijlage 3: Tabel statistisch significante variabelen

Naam variabele	Waarde stress (gemiddelde \pm std)	Waarde rust (gemiddelde \pm std)	P-waarde
Standaard deviatie hoofdpositie Y-richting (m)	0.01 \pm 0.01	0.06 \pm 0.02	0.000
Standaard deviatie hoofdrotatie X-richting (grad)	3.41 \pm 0.68	4.65 \pm 1.14	0.000
Gemiddelde rotatiesnelheid hoofd X-richting (grad/s)	7.54 \pm 1.34	6.31 \pm 1.26	0.005
Gemiddelde rotatiesnelheid hoofd Y-richting (grad/s)	52.74 \pm 8.6	44.11 \pm 9.98	0.006
Gemiddelde rotatiesnelheid rechterhand Y-richting (grad/s)	100.49 \pm 22.34	79.02 \pm 24.04	0.006
Gemiddelde snelheid hoofd (euclidisch, m/s)	0.6 \pm 0.1	0.5 \pm 0.13	0.009
Gemiddelde rotatiesnelheid rechterhand X-richting (grad/s)	56.69 \pm 12.98	45.09 \pm 14.92	0.012
Gemiddelde versnelling hoofd (m/s ²)	0.02 \pm 0.01	0.03 \pm 0.01	0.013
Standaard deviatie Y-positie rechterhand (m)	0.13 \pm 0.04	0.17 \pm 0.05	0.018
Standaard deviatie rotatiesnelheid rechterhand X-richting (grad/s)	101.7 \pm 32.79	78.5 \pm 28.92	0.023
Standaard deviatie oogrotatie in X-richting (grad/s)	5.02 \pm 1.08	5.66 \pm 0.7	0.031
Standaard deviatie rotatiesnelheid hoofd X-richting (grad/s)	7.76 \pm 1.75	6.62 \pm 1.47	0.032
Gemiddelde rotatiesnelheid rechterhand Z-richting (grad/s)	103.33 \pm 31.72	82.29 \pm 32.4	0.045
Gemiddelde rotatiesnelheid hoofd Z-richting (grad/s)	7.02 \pm 1.95	5.77 \pm 1.9	0.047

Tabel 1: Statistisch significante variabelen berekend met een t-toets. m = meter, grad = graden, s = seconden.