МЕЃУНАРОДЕН СЛАВЈАНСКИ УНИВЕРЗИТЕТ

ФАКУЛТЕТ ЗА **ТЕХНИЧКИ НАУКИ И ИНФОРМАТИКА**
НАСОКА: **КОМПЈУТЕРСКИ СИСТЕМИ И МРЕЖИ**

МАГИСТЕРСКА РАБОТА

СОВРЕМЕНА БИОСЕНЗОРСКА ТЕХНОЛОГИЈА И НЕЈЗИНА ПРИМЕНА
анализа на говор и аудио- вештачка интелигенција

Ментор:
проф. д-р Сашо Манасов

Изработил:
Марко Филиповски

Јуни, 2024
Св. Николе

С О Д Р Ж И Н А

ВОВЕД

1. Софтвер како медицински уред ---------- 3
2. Анализа на говор во потрошувачка неврологија ---------- 15
3. Трансформирање на конверзациска вештачка интелигенција ---------- 22
4. Роботите учат емпатија ---------- 34
5. Анализа на емоционални карактеристики во говор на политичари –--------- 41
6. Емоционална вештачка интелигенција за пациенти ---------- 42
7. Аудио вештачка интелигенција за автономни автомобили ---------- 44
8. Софтвер за препознавање изрази на лице ---------- 47
9. Реализирање на практично истражување ---------- 64
10. Анализа на резултати од биосензорско следење на очи ---------- 66
11. Споредување на поединечни вредности на испитаници ---------- 82

Заклучни согледувања ---------- 92

Користена литература ---------- 94

ВОВЕД

Податоци за човековото однесување (Behavioral data) денес се нова нормалност. Тоа е напредна иновативна технологија која ги забрзува научните истражувања за препознавање емоции (emotion recognition technology). Софтверот за биосензорски истражувања е извонредно корисна алатка за кодирање и опишување на однесувањето на испитаници и нивните емоционални реакции на точен и квантитативен начин. Оваа напредна технологија обезбедува мултисензорна интеграција (multisensor integration) во една единствена компјутеризирана платформа, така што лесно може да се синхронизираат повеќе текови на податоци (multiple data streams): Следење на очи (Eye Tracking), електроенцефалографија (EEG), електродермична активност EDA (Electrodermal Activity) и анализа на изрази на лице (Facial Expression Analysis) преку екранска поставка на стимули во скоро каква било форма, од слики, видеа и веб-страници, до поставки во реалниот свет што нé опкружува.

Постојните традиционалните техники овозможуваат квалитативно и квантитативно мерење на когнитивни и емоционални искуства само како вербално изразени на свесно ниво за време на интервју или анкетирање. Наместо тоа, со користење на техники за снимање на мозокот, изрази на лице, следење на очи, елекродермална активност, отчукување на срце, процес на дишење, човечки говор и други, може да се детектираат несвесни состојби поврзани со процеси кои имаат клучно влијание врз човековото однесување, интегрирајќи ги и оние сознанија што може се добиени со вербални или писмени декларации. Затоа, биосензорските истражувања се важно и револуционерно поле на научните истражувања, и се дефинирани како нејзина “нова димензија“. Денес биосензорските истражувања добиваат силно внимание во корпоративниот свет.

Голем дел од однесувањето на луѓето се водени од несвесни процеси. Деценски истражувања покажаа дека голем дел од нашата ментална, емоционална и когнитивна обработка се случува на потсвесно ниво. Ова може да се подобри со користење апликативен интеграбилен софтвер кој им дава на истражувачите пристап до емоционални, когнитивни и податоци за однесувањето. Ваков интеграбилен софтвер со холистички пристап, интегрира и синхронизира повеќе биосен-

зорски системи кои овозможуваат различен увид на човечки психофизиолошки појави.

Вистинската јачина на биосензорите во истражувачки студии е комбинација на софистицирани алатки што претставува холистички поглед на мисли, чувства и однесување на испитаници. Пример: користење на биосензорски систем за следење на очи (eye tracking) заедно со снимање на мозочна активност со EEG (електроенцефалографија) обезбедува податоци за реакцијата на мозокот во реално време на одредени стимули на компјутерски екран или настани од реалниот свет. Вакви софистицирани мерења исто така можат да откријат колку е мотивиран, ангажиран и колкав обем на работа е ангажиран во когниција, какво когнитивно оптоварување (cognitive workload) доживува одреден испитаник. Секоја од овие метрики го комплетира прегледот на податоци за следење на очи на различни начини. Нивото на когнитивно оптоварување може да ни каже за менталната, когнитивната и емоционалната состојба на испитаникот при извршување на одредена активност (решавање сложена задача или донесување одлука).

1. Софтвер како медицински уред

Пропулзивниот развој на компјутерските науки за медицински потреби во светот, овозможи да се развие софтверско решение кое воведува нови стандарди и услуги во системот за здравствена заштита. Тоа е софтверско решение наречено **SaMD (Software as Medical Device)** кое може да извршува една или повеќе медицински функции без користење на уред. Апликациите SaMD се наменети за **лекување или спречување на болести** и обично **се користат со немедицински хардвер поврзан на виртуелна мрежа**.

Критериуми за разликување на апликација SaMD од која било друга:

- SaMD може да врши медицински активности и директно да учествува во процеси на лекување.
- SaMD може да се лансира на генерички (немедицински) хардверски уреди: паметни телефони, компјутери, таблети, паметни часовници итн.

- SaMD не е наменет за управување со хардверскиот медицински уред и е самостоен медицински производ.
- SaMD може да се користи во соработка со хардверски медицински помагала и да се интегрира во сеопфатен клинички екосистем.

Што е Дигитална терапија (DTX - Digital Therapeutics)?

Дигитална терапија (DTX) обезбедува медицински интервенции директно на пациенти користејќи софтвер заснован на докази, клинички оценет за лекување, управување и спречување на широк спектар на болести и нарушувања. DTX исполнува стандарди и регулаторен надзор како и традиционални медицински третмани.

DTX ги елиминира празнините во здравствена грижа (заштита) со:

- Спроведување терапии со помош на паметни телефони, таблети и слични технологии.
- Зголемување на пристап на пациенти до клинички безбедни и афективни терапии.
- Намалување на стигма поврзана со испорака на одредени традиционални терапии преку нудење на удобност и приватност дома.
- Проширување на способноста на лекарите да се грижат за пациентите.
- Обезбедување терапии на различни јазици, како што се: англиски, шпански, арапски, германски и француски.
- Обезбедување значајни резултати и сознанија за персонализирани цели и исходи за пациенти и нивни клинички нарушувања (состојби).

DTX се регулирани со меѓународни стандарди: **FDA** (USA) и **DiGA** (German и Europe).

Дефиниции

Digital health (**дигитално здравје**) вклучува технологии, платформи и системи кои им користат на луѓето во нивниот начин на живот, благосостојба и цели

поврзани со настани, снимање, складирање или пренос на здравствени податоци и/или поддршка на науката за подобар живот и клинички операции.

Digital medicine (**дигитална медицина**) вклучува софтвер заснован на докази и/или хардверски производи кои бараат и/или интервенираат во служба на човековото здравје.

Digital therapeutic (**DTX**) (**дигитална терапија**) се продукти кои обезбедуваат терапевтски интервенции засновани на докази за да се помогне во справување или лекување на медицинско нарушување или болест.

Digital health

Снимање, меморирање и прикажување податоци и информации

- Технологии со кои се соочува корисникот.
- Компјутерска здравствена технологија - Health information Technology (HIT).
- Информации за здравјето на клиенти.

Пренос на податоци и информации

- Далечинско здравство (Telehealth).
- Софтвер за поддршка на процес на донесување одлуки (Decision support software).
- Поддршка на големо опкружување (Enterprise support).
- Алатки за медицинска администрација и управување со медицинска (клиничка) нега.

Мерни инструменти (продукти)

- Дигитална дијагностика.
- Дигитални биомаркери.
- Електронски клинички резултати (електронски проценки на клинички резултати).
- Далечинско следење на пациенти (Remote patient monitoring).
- Софтвер за поддршка на процес на донесување одлуки (Decision support software).

Производи за мерење и интервенции

- Дигитален придружник;
- Дигитални производ;
    - (1) Мерење и интервенирање;
    - (2) Не се бара човечка интервенција за да се служи на примарна цел.

Софтвер кој обезбедува терапевтска интервенција

- **Третирање на болест**. Дигитални терапевти кои обезбедуваат медицинска интервенција **за лекување** на болест.
- **Управување со болеста**. Дигитални терапевти кои обезбедуваат медицинска интервенција **за управување** со болест.
- **Подобрување на здравствена функција**. Дигитални терапевти кои обезбедуваат медицинска интервенција **за подобрување на здравствена функција и/или спречување** на болест.

Кои проблеми ги решава?

DTX се користи за превенција, интервенција и управување со проблеми и состојби на ментално здравје, вклучувајќи, но не ограничувајќи се на:

- ADD / ADHD;
- ADDICTION (зависност);
- ANXIETY (анксиозност);
- DEMENTIA (деменција);
- Спортски перформанси;
- Когнитивни перформанси;
- Енергија и мотивација;
- Главоболки и мигрена;
- Медитација и релаксација;
- Подигнување на расположение;
- Само-програмирање;
- Спиење и несоница;

- Учење и истражување;
- Друго…

DTx е достапно во секое време, каде било, на кој било уред..

SaMD може да работи како самостоен софтвер, користејќи AI (вештачка интелигенција) за да состави математички алгоритми кои му даваат инструкции на процесорот на уредот **да генерира осцилатори и модулатори** за еден или повеќе трансдуктори, конфигурирани да обезбедат визуелна, аудитивна, вибраторна или електрична стимулација.

AI може да го оптимизира и прилагоди **IRP (individual rehabilitation protocol)** (индивидуален протокол за рехабилитација) за секој корисник, според одговори на однапреди дефинирани прашалници после секоја сесија. Вештачка интелигенција **може да учи** од корисничко искуство и да ги користи одговори од прашалници за да може да ги менува аудио и видео записите, така што корисникот добива персонализиран **IRP**.

AI користи математички алгоритми за аудио, видео, вибрирачки и електрични трансдуктори користејќи податоци од технологија SMART / CHI / BCI како ВЛЕЗНИ ПРОМЕНЛИВИ.

AI може дополнително да го оптимизира и прилагоди IRP за корисникот, според одговори на прашалници по секоја сесија.

AI може дополнително да ги калибрира математичките алгоритми за аудио и видео претворувачи според тип на избран трансдуцер.

- AUDIO: Head Phones / Speakers (Stereo of Surround);
- VIDEO: LCD / LED / AudioStrobe Glasses;
- TACTILE: Vibroshakers / Zapper / Torus.

DTX може да се користи како **самостојна терапија или во комбинација** со повеќе конвенционални третмани како фармаколошка или лична терапија.

DTX може да се користи за управување и подобрување на резултати во бројни состојби, вклучувајќи дијабетес тип 2, Алцхајмерова болест, деменција, конгестивна срцева слабост, хронична опструктивна белодробна болест, астма, бе-

лодробна болест, дебелина, абдоминација на супстанции, ADHD, несоница, хипертензија (покачен крвен притисок), анксиозност, депресија и други.

Не постои единствена методологија што се користи во практиката на **DTX**. Многу пристапи користат методи засновани на когнитивна бихејвиорална терапија за да ги поттикнат пациентите **да направат промени во животниот стил**, засилени со гејмификација, поддршка од врсници, а во некои случаи и телездравје, како што се тренинг или психотерапија.

Од **2020** година, кога регулативата за **SAMD** и **DTX** првпат стапи во пракса, листата на **DTX** ширум светот постојано расте.

Каква технологија е DTX

Оваа технологија овозможува снимање и анализа на активности на мозочни бранови кај човекот, **стратешки водејќи кон трансформативни состојби на свеста**. Наместо да вградува едноставни бинаурални фреквенции во природен soundtrack, оваа технологијата **влијае на невролошка активност преку повеќеслоен пристап кој комбинира бројни стратегии за привлекување мозочни бранови со откритија во психологија**. Исто така користи математички алгоритми за аудио и видео трансдуктори. Алгоритмите ја подобруваат ефикасноста со интегрирање на повеќе методологии во синергетски можни аудио-визуелни матрици, со што привлекуваат повеќе динамички текови во мозокот за производство на синхрони мозочни бранови и длабоки состојби на медитација. Ова резултира со креативно трансформирачко искуство.

Интеграција на SMART технологија

Напредокот во мехатроника во последните години, Internet of Things (IOT), уреди за носење и минијатуризација на сензори и електроника, значително ги зголемија можностите на паметните сензори кои го зголемуваат опсегот на нивните апликации.

Меѓу другото, здравствената заштита и медицинските области се под силно влијание на еволуцијата на овие технологии, кои даваат значителен придонес во многу апликации во оваа област.

Всушност, паметните сензори сè повеќе нудат нови решенија за многу релевантни предизвици во здравството, како што се:

- **рано откривање на патологии,** минимално инвазивно управување;
- превенција на болести со висок товар.

Понатаму, развојот на минијатуризирани и лесни системи засновани на паметни сензори може да биде клучен играч во овозможување:

- побрз раст на ненаметливи и ненадгледувани пристапи за рехабилитација дома;
- континуирано следење на статусот на пациентите.

Технологија за интеграција на човек – компјутер- интеракција (**HCI - Human – Computer Interaction**)

Human-computer interaction (HCI) е истражување во дизајнот и користење на компјутерска технологија, која се фокусира на интерфејси помеѓу луѓето (корисници) и компјутерите.

HCI вклучува и интегрира уреди за EMG (Електромиографија), ECG (Електрокардиографија) и EDA (Електродермална активност). HCI може да интегрира други уреди како EOC (ElectroOculoGraphy) и FEA (Анализа на експресија на лице) (Facial Expresion Analysis).

EMG ги мери електричните струи кои се генерираат во мускулот за време на неговата контракција и претставуваат **невромускулни** активности.

ECG ги мери електричните струи кои се генерираат во срцето за време на неговата контракција и претставуваат **неврокардио** активности.

EDA ја мери спроводливоста на кожата за да укаже на интензитет на возбуда и стрес.

EMG, ECG and EDA сензорските сигнали може да се користат за различни апликации, вклучувајќи **интерактивен computer gaming** и **неврорехабилитација**.

Технологија за интегрирање на интерфејс мозок – компјутер (BCI - Brain-Computer Interface)

BCI, понекогаш наречен интерфејс човек-компјутер, мозок-машина (**BMI - brain-machine interface**) или **паметен мозок** (**smartbrain**), е директна комуникациска патека помеѓу електрична активност на мозокот и надворешен уред, најчесто компјутер или роботски екстремитет.

BCIs често се насочени кон истражување, мапирање, помагање, зголемување или поправка на човечките когнитивни или сензорно-моторни функции.

Тие често се конципираат како интерфејс човек-машина што ја прескокнува посредничката компонента на физичкото движење на делови од телото, иако тие исто така ја зголемуваат можноста за бришење на дискретноста на мозокот и машината.

Brain-computer interfaces честопати имаат огромен потенцијал за неврообновување и неврорехабилитација за пациенти со невролошка болест. BCI може да служи за мерење и враќање на капацитети изгубени поради невролошка повреда или болест, или за **поттикнување невропластичност**, за да се подобри учењето и повторното мапирање по невронска (мозочна) повреда.

Како работи мозокот на човекот? Мозок, ум и однесување

Човечкиот мозок е сложена структура создадена за многу човечки активности:

- Активно да бараме храна;
- Да комуницираме со нашето опкружување;
- Да разговараме со колеги, пријатели, семејство;
- Да избегнеме непријатели што реално ни се закануваат;
- Да се справиме (да ги решаваме) со покомплексни секојдневни предизвици – борба за повисока оценка, интервју за работа, пишување магистерска работа и др.

Нашиот мозок **е постојано активен** без оглед на тоа дали спиеме или сме будни, дали сме во релаксирачка прошетка или се подготвуваме за следен важен деловен состанок (сонуваме, гледаме, размислуваме, чувствуваме). Ги собира (меморира, апсорбира) сите информации, ги интегрира (компактира) и повторно ги

поврзува постоечките собрани податоци и интегрира сé во постојано човечко искуство, за нас и за нашето опкружување.

Човечкиот мозок обликува (визуелна претстава) како ја гледаме околината околу нас и и филтрира и потенцира одредени предмети и информации што се најрелевантни за секој од нас. Врз основа на нашите желби, мисли, емоции, животни искуства, нашиот мозок го придвижува и води нашето однесување , дури и потсвесно контролира одредени процеси на однесување без да го забележиме тоа.

За да може да се прави анализа на човековото однесување и кои се негови основни двигатели, потребно е детално (темелно) разбирање на сложената структура на човечкиот мозок, неговата структура и функција (како функционираат одредени мозочни блокови?).

Брзиот напредок на компјутерската технологија за снимање и мапирање на човечкиот мозок овозможи денес истражувачите да нурнат длабоко во човечкиот мозок и да добијат сознанија („да видат“) како се обликуваат нашите перцепции и интеракции со светот што нé опкружува.

Како функционира човечкиот мозок?

Електроенцефалографија (**EEG –** electroencephalography) е најнова технологија за снимање на човечкиот мозок – одредување на електрична мозочна активност во поедини региони на мозокот.

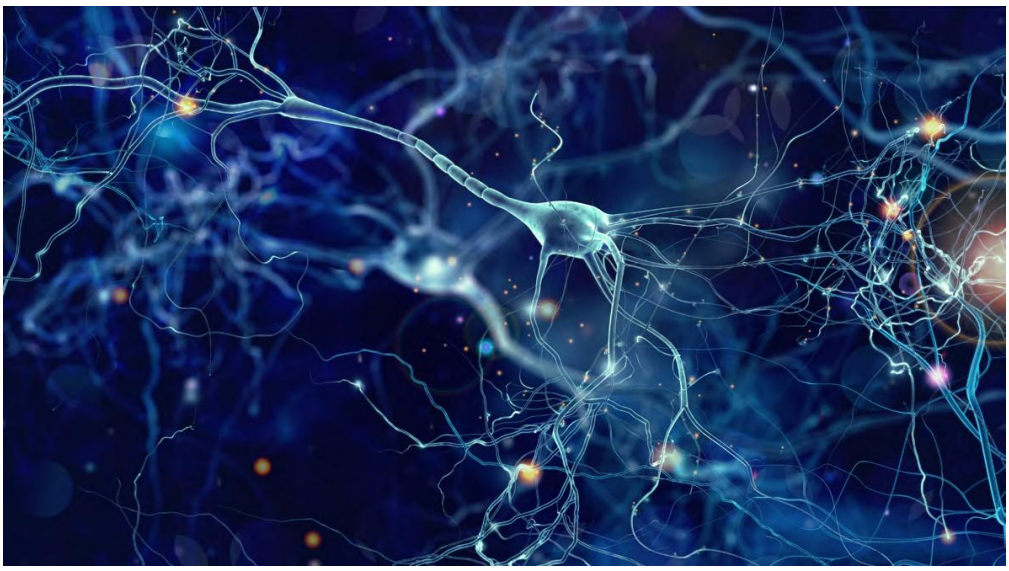


Електричната мозочна активност се евидентира со користење на електроди поставени на главата на човекот. Со ова се добиваат сознанија како различни неврони во мозочната мрежа комуницираат едни со други преку електрични импулси. Ефектите на технологијата на електроенцефалографија се непроценливи за истражување (проучување) на неврокогнитивните процеси што се наоѓаат во основата на однесувањето на човекот (human behavior).

Со електроенцефалографија се добиваат сознанија за:

- Како функционира човечкиот мозок?
- Кои региони во мозокот се активни?
- Како поедини мозочни региони комуницираат помеѓу себе?

Човечкиот мозок е составен од околу 85 милијарди неврони – мозочни клетки кои реализираат најголем дел од сите мозочни комуникации. Преносот на информации (синаптичен пренос) се реализира со помош на невротрансмитери. Вакава синаптичка активност генерира суптилно електрично поле кое трае од десетици до стотици милисекунди.



Технологијата електроенцефалографија е **неинвазивна** и пасивна техника на снимање на мозочни бранови и мозочна активност. Добиените електрични сигнали се многу мали и затоа снимените податоци се дигитализираат и се зголемуваат (засилуваат) со помош на засилувач.

Постои класификација на фреквенции добиени од мозочна активност:

- **Делта бранови** (1 -4 Hz);
- **Тета бранови** (4 – 8 Hz);
- **Алфа бранови** (8 -12 Hz);
- **Бета бранови** (13 – 25 Hz);
- **Гама бранови** (> 25 Hz).

Делта бранови се најбавни мозочни бранови со најголема амплитуда. Се присутни само во длабоко РЕМ спиење. Овозможуваат да се процени (истражи) длабочината на спиењето (посилен делтаритам значи подлабок сон). Делта фреквенциите се посилни во десната (емоционална) мозочна хемисфера и се поврзани со консолидација на човечката меморија.

Тета- брановите служат како носител на фреквенција за когнитивна обработка. Се поврзани со:

- ментално оптоварување или работна меморија;
- тежина на ментални операции;
- обработка на податоци и процес на учење;
- враќање на меморија;
- се поистакнати со зголемување на тежината на задачите.

Алфа- брановите имаат фунционална корелација со рефлектирани сензорни, мемориски и моторни функции. Имаат зголемено ниво на моќност за време на физичка и ментална релаксација со затворени очи, а се намалува при ментална или телесна активност со отворени очи,на пример при фокусирано внимание кон каков било стимул. Намалувањето на нивото на алфа бранови индицира дека нашиот мозок се подготвува да собере информации од различни сетила преку координирање на ресурсите за внимание, фокусирајќи се на она што навистина му е важно во тој значаен момент.

Алфа- брановите рефлектираат релаксирана состојба и сензорна инхибиција и се користат во одредување на нивоа на медитација кај искусни и почетници-медитатори. Исто така со нив се следи нивото на релаксација на кој било човек. Зголемено ниво на алфа- бранови значи дека лицето има подлабока релаксација

– ова се користи во сценарија за релаксација и за клиничка популација (биофидбек тренинг).

Алфа- брановите (алфа моќност) се поврзани со просторно, социјално и семантичко внимание. Расеани испитаници и лоши изведувачи покажуваат повисоко ниво на алфа моќност (Rana & Kaina, 2014)[1].

Бета- брановите (Niedermeyer & da Silva, 2012)[2] се поврзани (имаат поголема бета моќ) со состојби на:

- Активно размислување;
- Зафатено размислување;
- Вознемирувачко размислување;
- Активна концентрација.

Моќноста на бета- брановите е посилна кога човекот планира или извршува движења (особено кога се бара движење на прстите) и фокусирано внимание. Исто така човечкиот мозок ги набљудува движењата на други лица (зголемена бета-моќност) што укажува дека нашиот мозок ги имитира движењата на екстремитетите на другите. Ова укажува дека во човечкиот мозок постои „**огледален невронски систем**“ („**mirror neuron system**”) кој е координиран од бета- бранови (бета фреквенции) (Zhang et al., 2008)[3].

Во научни и медицински истражувања бета- брановите се користат за:

- Моторичка контрола – пациенти со невродегенеративни нарушувања.
- Состојби на приправност (зголемено внимание) – бета бранови се следат при стимулации со светлосни, звучни или психостимуланси кои го модифицираат нивото на будност (приправност, подготвеност за акција) и нивото на внимание.

[1] Rana & Kaina, “**Functional roles of 10 Hz alpha-band power modulating engagement and disengagement of cortical networks in a complex visual motion task”**, 2014, PLOS One.

[2] Niedermeyer & Lopes da Silva, “**Electroencephalography: Basic Principles, Clinical Applications, and Related Fields”**, 2012, Philadelphia.

[3] Zhang, Chen, Bressler &ing, “**Response preparation and inhibition: the role of the cortical sensorimotor beta rhythn”**, 2008, Neuroscience.

ГАМА- бранови рефлектираат процес на внимание. Служат како носител на фреквенција за врзување на разни сензорски впечатоци во еден предмет. Не одразуваат когнитивна обработка (Dimigen et al., 2009)[4].

2. Анализа на говор во потрошувачка неврологија

Пристапот за анализа на говор во развој на UX се преклопува со начинот на кој често, но не исклучиво се применува во рамки на наука за потрошувачка неврологија. Се гради производ или платформа и корисникот комуницира со него - таа интеракција потоа на крајот се мери како позитивна, негативна или неутрална во дискретни моменти.

Додека примена на гласовна анализа во контексти на потрошувачки невролошки науки или невромаркетинг се дискутира барем од 1980-тите, со истражувачи кои забележаа како „може да се идентификува „искрен“ одговор (‘sincere’) наспроти „услуга од усни“ (‘lip-service’). Исто така беше забележано дека тогашното истражување имало „неправилни истражувачки инструменти и резултирало со преценување на наодите“. Ова наложи потреба од подобрени пресметковни методи на анализа (сега лесно достапни) и повторно проценување на дизајни на студиите за собирање гласовни податоци. Ова доведе до преклопувања што сега ги гледаме со полето на UX истражување.

Додека испитаниците во студија за потрошувачки невролошки науки можеби нема спонтано да ги изразат своите мисли при интеракција со нов производ или реклама, анализирањето на нивните вербални одговори за време на гласно размислување, интервјуа и фокус групи обезбедува нова димензија за разбирање на нивните внатрешни мисли и чувства.

Обезбедувањето дека постои природна средина во која тие можат да зборуваат е од клучно значење за максимално искористување на гласовна анализа. Ова е, на пример, прикажано во една студија за онлајн написи за убавина и мода во кои се користат: анализа на говор и интервју за да се проценат гласовните податоци заедно со следење на очи и анализа на израз на лице.

---

[4] Dimigen et al., “**Human microsaccade-related visual brain responses”**, 2009, Journal of Neuroscience.

Може да се направи анализа на потрошувачите со анализа на говор како одговор на реклами (во комбинација со интервју или слично), но исто така важи и спротивното. Студија од 2019 година од Универзитетот Тохоку, Јапонија, покажа дека потрошувачите кои слушаат реклами за прехранбени производи имаат поголема веројатност да претпочитаат слатка или кисела храна кога гласот е на висок тон.

Тие откриле дека „значително високиот“- тон на гласот веројатно ќе поттикне специфична предност за слатка храна. Ова покажува уште еден пат преку кој невромаркетарите можат дополнително да ја разберат и подобрат резонанцата на производите што помага кои производи да се донесат на пазарот.

Заклучок

Иако анализа на говор е релативно воспоставена методологија во рамки на дијагностичко истражување, таа е во голема мерка недоволно истражена и недоволно искористена алатка во комплетот метрики за човековото однесување. Во минатото, ова се должи на макотрпна работа потребна за класифицирање и декодирање на говорни сигнали и недостаток на робусни достапни методи. Меѓутоа, сега, со усовршување на софтверот и напредок во вештачката интелигенција, можеме да го видиме вистинското ветување што оваа методологија го носи за отворање и истражување на овој централен сигнал за човечкото однесување.

Говор и вештачка интелигенција

AudEERING и iMotions соработуваат за да го подобрат истражувањето за однесувањето на луѓето

Водечкиот светски снабдувач на софтвер за човечки увиди (емоционални и физиолошки) iMotions интегрира технологија за анализа на говор од германската компанија за вештачка интелигенција audEERING

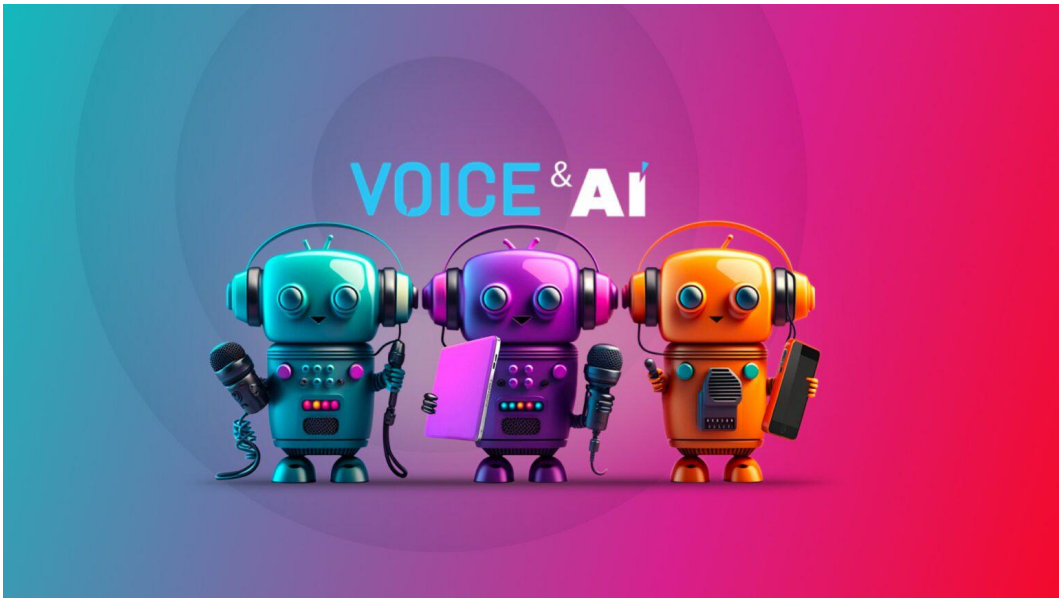
Двете компании имаат значително искуство во идентификување и анализа на човековото изразување што е можно попрецизно со цел да се извлечат заклучоци за човековото однесување, емоции и биомаркери.

Широкиот спектар на апликации во науката и индустријата ја нагласува важноста на соработката - интердисциплинарни истражувања на човековото однесу-

вање, за да ги прошири областите на мерење и анализа во истражување на науката за човекот.

Со интегрирана говорна компонента на вештачка интелигенција, мултимодалниот софтверски пакет добива друга димензија што обезбедува уште една значајна компонента за увид во човечкото однесување.

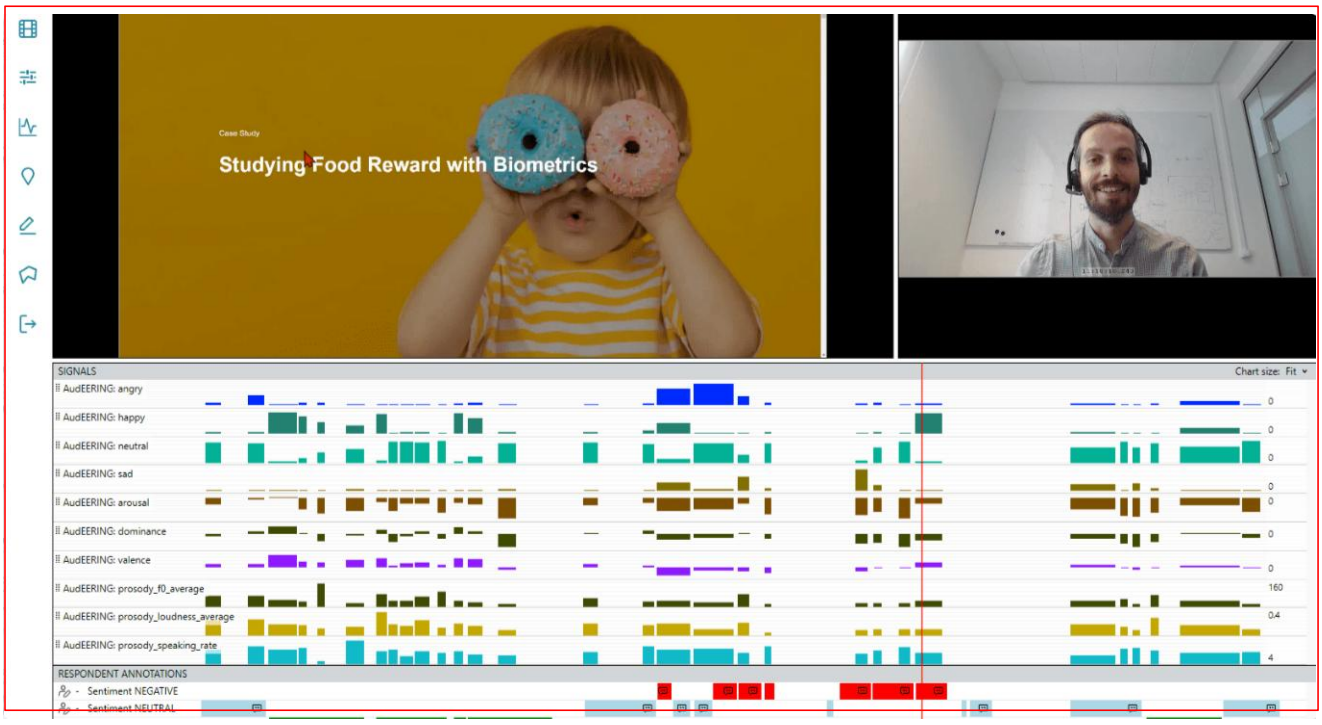
Соработката има за цел да ги искористи силните страни на двете компании за да понуди иновативни решенија за анализа и снимање на емоции и биометриски податоци. Со интегрирање на напредна технологија за аудио анализа **devAIce®** на audEERING во сеопфатната биометриска платформа за истражување на iMotions (заедно со веќе развиени биосензорски системи), научните истражувачи, маркетинг истражувачите и други професионалци можат да добијат подлабок увид во емотивните искуства на луѓето. Важен показател за тоа е човечкиот говор.



Германската компанија audEERING е специјализирана за развој на напредни технологии за вештачка интелигенција за анализа на говорниот јазик што препознава карактеристики од гласот, што овозможуваат извлекување заклучоци за емоционалното изразување, меѓу други параметри.

Со својот производ iMotions Lab, провајдерот и развивач на софтвер iMotions им овозможува на истражувачите да собираат, анализираат и користат различни модалитети на податоци, како што се: говор, видео, текст, биосигнали и

други влезни стимули паралелно за нивните истражувачки пристапи. Интегрира и синхронизира различни физиолошки сензори кои обезбедуваат различни човечки научни наоди, вклучително и следење на очите, спроводливост на кожата, EEG, EKG и EMG, како и анализа на емоции и изрази на лицето базирани на камера. Како една од петте подружници на водечкиот светски добавувач на вештачка интелигенција за човечки увид Smart Eye Group, iMotions работи во ексклузивна партнерска мрежа на специјалисти од областа на човековата наука со вештачка интелигенција, вклучително и Affectiva.



Доловување на емоции во широк опсег на апликации

Комбинација на патентирани алгоритми за аудио анализа на audEERING со биометриските мерења на iMotions овозможува детално собирање на емоции во широк опсег на области на примена. Од **развој на производ** до **рекламирање** и **истражување на пазарот** до **клиничка дијагностика**. Оваа соработка отвора нови можности за подобро разбирање и анализа на човечките емоции.

Со користење на технологијата за гласовна анализа на audEERING, истражувачите, меѓу другото, можат да пристапат до нивниот мултидимензионален модел на емоции. Ова ја мери валентноста, активирањето и доминацијата.

- **Валентноста** се однесува на позитивноста или негативноста на емоционалното изразување;
- **активирањето** на нивото на енергија;
- **доминација** на доминантен емоционален израз.

Човечкиот глас додава друга димензија на разбирање на човечките емоции кои го поттикнуваат нивното однесување и донесување одлуки. Бидејќи iMotions има веќе развиено мултимодална алатка за истражување на човековото однесување, се вклучува оваа димензија за да ги направи маркерите базирани на глас достапни за истражувачите кои веќе работат со сигнали собрани од очите, лицето, мозокот, срцето или кожата. Затоа е потребно постојано да се подобрува квалитетот на методологии за научни истражувања. Тоа дава значаен придонес за подобрување на научните процеси и воведува нова ера за анализа на човековото однесување.

Технологијата за вештачка интелигенција на audEERING веќе се користи надвор од научните истражувања: уредот се користи во различни сектори од приватниот сектор, на пример:

- во развојот на емпатични роботи за секторот за грижа на клиенти (медицинска роботика);
- за подобрување на искуството на клиентите при интеракција со центри за повици.

Со својот производ iMotions им овозможува на истражувачите да собираат, анализираат и користат различни модалитети на податоци, како што се говор, видео, текст, биосигнали (и други влезни стимули) и тоа паралелно за нивните истражувачки пристапи. Интегрира и синхронизира различни физиолошки сензори кои обезбедуваат различни човечки научни наоди, вклучително и следење на очи, спроводливост на кожата, **EEG** (електроенцефалографија), **EKG** (електрокардиографија) и **EMG** (електромиографија), како и анализа на емоции и изрази на лице базирани на камера.

Анализа на ментално здравје

Менталното здравје е сè поважна тема во денешните општества кои брзо се движат. Младите се силно погодени од барањата со кои се соочуваат во текот на нивното образование и во социјалните групи. Според Германското здружение на психотерапевти, **20%** од сите млади во Германија имаат проблеми со менталното здравје.

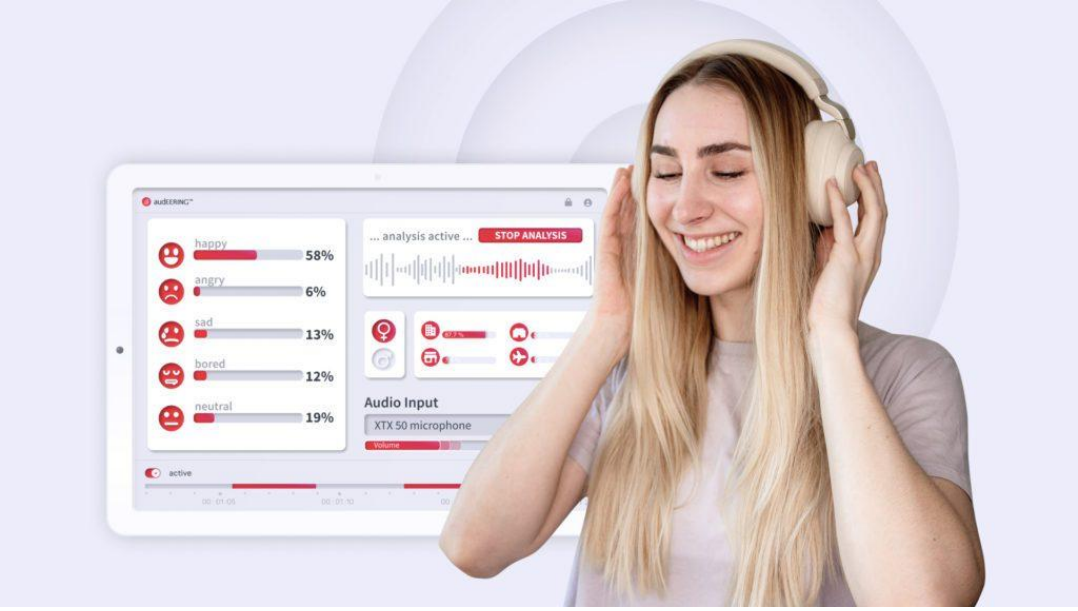


Апликација за следење на менталното здравје – MyMoodCoach

Многу е важно да се развијат нови, иновативни методи за да им се помогне на тие млади луѓе и да им се обезбедат алатки за следење на нивното ментално здравје и да побараат стручна помош доколку е потребно. Со истражувачкиот проект EcoWeb, млади луѓе на возраст меѓу 16 и 22 години ќе бидат регрутирани за да научат повеќе за **расположението**, **емоциите** и **менталното здравје** и да тестираат дали мобилните апликации за самопомош можат да промовираат емоционална благосостојба. Апликативното решение ги обединува најновите истражувања за самоследење, техники за самопомош и когнитивна бихејвиорална терапија за младите да можат да научат:

- за сопствените емоции;
- да развијат емоционална отпорност;

- да изградат благосостојба.



Придонесот на audEERING во овој проект е неговата позната технологија за откривање емоции. Врз основа на модели за машинско учење, корисниците добиваат проценка за звукот на нивниот глас во однос на доминација и возбудување. За да се направи ова, audEERING обезбедува рамка, која ги извлекува акустичните карактеристики од примероци на говор што се зборуваат корисници на телефон. Понатаму, софтверот на audEERING ги проценува емоционалните состојби од психолошки прашалници врз основа на креиран алгоритам.

Затворање на валентниот јаз во препознавање на емоции

2021 година е значајна година за истражувања кои работат на препознавање емоции од говор. Користејќи го напредокот на transformer-based architectures, за прв пат се изградија модели кои предвидуваат валентност (т.е. колку негативно или позитивно звучи одреден говорен исказ) со слична висока прецизност како возбудувањето (т.е. колку возбудливо звучи одреден говорен исказ). Затворањето на валентниот јаз е голем предизвик во **препознавање на говорни емоции** (SER-speech emotion recognition). И за разлика од претходните пристапи, овие модели ги изведуваат своите предвидувања директно од необработено аудио и не бараат транскрипции на текст.

3. Трансформирање на конверзациска вештачка интелигенција: Воведување ERA (Емоционално одговорен асистент) (Emotionally Responsive Assistant)

Во ера на брз технолошки напредок, асистенти со вештачка интелигенција (AI assistants) станаа составен дел од нашиот секојдневен живот. Од одговарање на наши прашања до обезбедување насоки, па дури и контрола на нашите паметни домови. Овие виртуелни придружници ни го направија животот полесен и поефикасен. Сепак, отсекогаш недостасуваше еден клучен аспект од овие интеракции: **способноста да се разбере нашето емоционално изразување**.

Интегрирање на Voice AI во дизајн на разговор

audEERING GmbH, пионер во областа на афективно пресметување, е подготвен да ја промени интеракцијата човек-компјутер со своето револуционерно демо прикажување за интеграција на Voice AI во **разговорни агенти за вештачка интелигенција** (AI agents). Во светот на емоционално интелигентна вештачка интелигенција, оваа иновација ќе ги револуционизира нашите интеракции со виртуелни асистенти.

Зошто ни се потребни асистенти за вештачка интелигенција?

Во денешниот свет со брзо темпо, времето е скапоцена работа. Асистентите со вештачка интелигенција нудат удобен и ефикасен начин за управување со нашите задачи, одговарање на наши прашања и насочување на нашите дневни рутини. Тие дејствуваат како наш значаен помагач, заштедувајќи ни драгоцено време и напор. Досега, овие асистенти беа ограничени во разбирање на човечките емоции, што доведе до значителен јаз во комуникацијата.

Проблем со постојни асистенти за вештачка интелигенција

Додека постојните асистенти за вештачка интелигенција се одлични во обезбедувањето точни и релевантни информации, тие често пати не успеваат да ги разберат нијансите на човечките емоции кога комуницираме со нив. Ова ограничување резултира со помалку персонализирано и помалку емпатично искуство. Нашите емоции играат фундаментална улога во комуникацијата, влијаат на начинот на кој се изразуваме и контекстот на нашите барања. Без ова емоционално

разбирање, асистентите со вештачка интелигенција не се свесни за нашите основни чувства, што доведува до пропуштени можности за приспособени одговори и поддршка.

Воведување на демо ERA - Емоционално одговорен асистент

Препознавајќи го значењето на **емоционалната свесност** во интеракции човек-компјутер, audEERING GmbH разви иновативно решение што го премостува емоционалниот јаз во разговорна вештачка интелигенција. Оваа демонстрација ја прикажува беспрекорната интеграција на Voice AI, овозможувајќи им на асистентите со вештачка интелигенција:

- точно да ги препознаат;
- да одговорат на нашите емоции.

ERA се напојува со најнова технологија за вештачка интелигенција – приклучок за Unity и Unreal - за да го анализира гласот во реално време на уредот за да ги извлече емоциите од гласот. Анализата се заснова на димензии на **возбуда** и **валентност**. Резултатите од емоциите се миксираат со резултатите од ASR (Автоматско препознавање говор - Automatic Speech Recognition) од Google и се испраќаат до **ChatGPT** на **OpenAI**. Текстуалниот одговор се трансформира во говор користејќи ја технологијата на **ElevenLabs**.

Подобрување на интеракции со емоционална интелигенција

Со интеграцијата на **Voice AI**, асистентите со вештачка интелигенција сега можат да го согледаат емоционалниот израз на корисникот и соодветно да ги приспособат нивните одговори. Да замислиме дека имаме помошник со вештачка интелигенција кој може да ја почувствува нашата фрустрација во текот на бурен и напорен ден и да одговори со смирувачки предлози или да понуди сочувствување. Понатаму, Voice AI може да овозможи персонализирани препораки специфични за секој корисник, како што е пуштање на наша омилена музика кога се чувствуваме навредени или предлагање активности што се во согласност со нашата моментална емоционална состојба.

Овој пробив ја подобрува функционалноста на асистентите со вештачка интелигенција и отвора нови можности за апликации во различни домени. Од поддршка за ментално здравје до услуги на клиенти и пошироко, потенцијалните апликации на емоционална вештачка интелигенција (emotionally intelligent AI) се огромни и ветувачки.

Каква е иднината на агенти за разговорна (конверзациска) вештачка интелигенција

Демонстрацијата на ERA на audEERING која ја прикажува интеграцијата на **Voice AI** во конверзациски агенти за вештачка интелигенција означува значајна пресвртница во доменот на **афективно пресметување** (одредување на емоции). Овозможувајќи им на асистенти со вештачка интелигенција да сфатат и да одговорат на човечко емоционално изразување, оваа иновација ги револуционизира нашите интеракции со виртуелните придружници. Додека ја прифаќаме емоционално интелигентната вештачка интелигенција, можеме да очекуваме иднина во која нашите виртуелни асистенти не само што ќе ги задоволат нашите потреби, туку, исто така, ќе нé разберат и ќе сочувствуваат со нашите емоции, подобрувајќи ја нашата севкупна благосостојба и задоволство.

Web API уреди на audEERING

Вештачка интелигенција



Трансформирање на конверзациска вештачка интелигенција: воведување на ERA (Емоционално асистент)

Прв чекор: Емоции

devAIce® интегрира аудио анализа во софтвер и хардвер. Тоа е основна технологија на audEERING. Во аудио анализа комбинира различни функции како што се:

- откривање емоции;
- атрибути на говорникот;
- детекција на сцена;
- глас;
- музичка активност.

Тоа им овозможува на клиентите да ја користат технологијата за вештачка интелигенција на audEERING во нивна деловна апликација на едноставен начин за да создадат искуство погодно за програмери со одлична корисничка употребливост.

Човечка интелигенција зад Voice AI

Миленко Сапоња (сопственик на производот Voice AI), е еден од software developers на audEERING со повеќе од шест години искуство во различни оддели за истражување и развој во компанијата. Со својата специјализација за роботика, тој исто така може да биде клучен учесник во проектот **#ERIK**, кој има за цел да развие иновативен концепт **за терапија на деца со аутизам**. За ова, **#robot-pepper** беше прилагоден со основната AI-технологијата **devAIce®** на **audEERING**.

Процес на развој на технологија за вештачка интелигенција

Како изгледа процес на развој на технологија за вештачка интелигенција? За сопственикот на производот Миленко и неговиот тим за уреди, процесот на развој се врти околу интеграција на сите различни модели на кои работи внатрешниот истражувачки тим и го трансформира во производ што клиентите можат да го користат.

Можности на производот и потенцијални апликации

Производот Voice AI им овозможува на клиентите лесно да ја користат технологија audEERING, со едноставни зборови да ја анализираат секоја аудио датотека на човечкиот глас, на пример:

- препознавање емоции;
- детекција на пол, возраст;
- интонација на зборување;
- брзина на зборување и така натаму.

Само врз основа на акустика може да ни каже каде сме. На пример, ако сме во канцеларија или сме во автомобил. Можни се и одредени настани како плачење на бебе, ѕвонење на телефонот, музика во позадина.

Примери за можни случаи на користење

Интегрирањето на уредот во центри за повици (call centers) е фасцинантна можност за откривање на тонот на гласот. Ќе се дознае колку се задоволни клиентите или колку се пријателски настроени самите агенти.

Исто така, на полето на истражување на пазарот околу општото мислење и **задоволството на клиентите** кон одреден производ, deAIce и воопшто Voice AI технологијата создава огромна иновација.

Понатаму, апликацијата оди во **IOT** уреди и во секоја основна интеракција човек-машина, како што е интеракција со роботи и автомобили.

devAIce® им овозможува на корисниците користење на најнова технологија за вештачка интелигенција на едноставен начин, така што буквално е потребна еден линиски код за да се добие препознавање на акустични (говорни) емоции и не се потребни некои GPU сервери за пишување, туку може да работи и на едноставен физички уред.

Ние луѓето сè уште не знаеме многу за гласот, кој е најважен. Го користиме само за секојдневна комуникација, но може да се користи за многу повеќе од тоа. На пример, кога имаме настинка тоа е лесно забележливо во нашиот глас, но оваа мерлива информација може да се користи за многу повеќе. Исто така, не се лесно забележливи состојби како:

- детекција на депресија;
- откривање на ниво на стрес;
- Алцхајмер и Паркинсон.

Денес можностите на гласот (говорот) не се доволно истражени, но од друга страна, сигурно во иднина тие ќе бидат многу поистражени, анализирани и многу поактуелни.

Во основа, со SDK решение, клиентот може да ја активира (стартува) лиценцата на својот физички уред и затоа нема потреба од Интернет конекција и голема пресметувачка моќност. Интернет конекција е потребна само за барање за лиценца.

Корисниците не мора да плаќаат за ажурирање на секоја лиценца. Може да се продолжи да се користи било која верзија со која моментално се работи. Корисниците на devAIce SDK можат експлицитно да се одлучат за да ја добијат новата верзија. Но, се препорачува да се користат ажурирања на најнова верзија за да се има корист од најнови додатоци, подобрувања и поправки на моделот.

Втор чекор: Моделирање емоции

Дефиниција и улога на емоцијата, чувството, расположението и афектот

Постојат многу теории и модели кои се обидуваат да ги објаснат чувствата и емоциите. Тие се обидуваат да ги објаснат емоциите од различни перспективи: невробиолошки, когнитивни, психолошки.

Афективен пристап за да биде поефективен

Прв чекор во моделирање емоции е да одлучиме дали треба да ги моделираме **дискретно** или **континуирано**. Дискретното моделирање е лесно да се разбере. Бинарното е на пример најлесно: „Среќно“ наспроти „Несреќно“, „Тажно наспроти „Нетажно“. Но, сите знаеме дека тоа не е реалноста? Можеме да бидеме повеќе од ова, да бидеме радосни, среќни, возбудени итн. Затоа, треба да се проучува **континуирано моделирање**.

Сега, треба да одлучиме за димензија или димензии. Димензија е нешто што може да ги разликува емоциите подобро од било што друго. Интуитивно е прашањето: **Како се перципираат емоциите**?

Валентност и возбуда за да се разликува емоционалното изразување

Валентноста ги дели емоциите на позитивни и негативни. На пример, тагата е негативна, а среќата е позитивна. Депресивен човек е повеќе негативен отколку тажен, а возбудениот е повеќе позитивен отколку среќен. **Валентноста** еднодимензионален простор.

За да се направи разлика помеѓу две негативни или позитивни емоции, потребна е втора димензија.

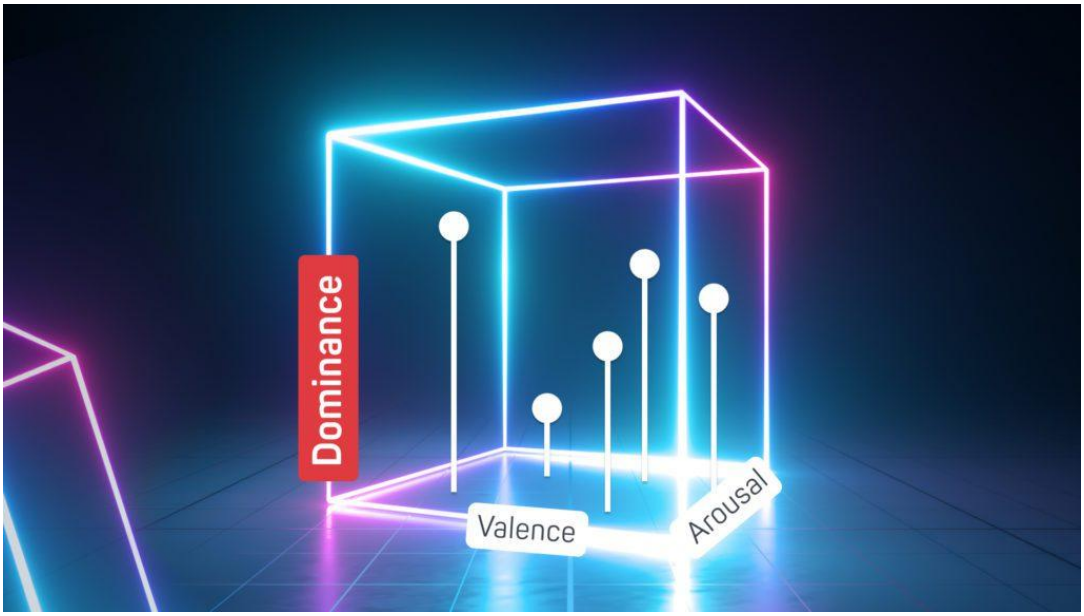
- Како тие се разликуваат кога ќе се земе предвид нивниот израз?

Кога некој е возбуден, личноста е повознемирена (more aroused) и погласна. Кога некој е среќен, тој/таа е помалку возбуден. Истото е и со гневот и тагата. Кога сум лут, викам (има и ладен гнев) и имам негативни емоции, а кога сум тажен, речиси тивко плачам и имам негативни емоции. Значи, добар кандидат за втората димензија е **возбуда**. Колку повеќе сум возбуден, толку е поинтензивна емоцијата. Значи, сега имаме дводимензионален простор.

Доминација како трета димензија

Пример: кога се **плашиме** и **не фаќа паника**, доживуваме **негативна емоција**, а исто така **викаме**, што значи дека имаме и **голема возбуда**. Слично на тоа, кога сме **лути**, исто така доживуваме негативна емоција со висока возбуда. Сепак, тие воопшто не се слични. Затоа овие две димензии не се доволни. Значи, која е разлика помеѓу **паника** и **гнев**? Тоа е **контрола над ситуацијата**. Тоа е на-шата **доминација**. Значи, третата димензија ќе биде **контрола** или **доминација**.

Конечно, имаме **3D простор** што можеме да го користиме за моделирање на емоции.



Слика. 3D простор за моделирање емоции

Секојдневен процес: Слушање и идентификување

Пример: експеримент. Да ги затвориме очите и да слушаме нечиј глас. Можеме да видиме дека користејќи ги овие димензии, можеме да препознаеме речиси секоја емоција и да најдеме простор за неа.



На ваков начин можеме математички да ги претставиме нашите емоции. Тоа е **основа да ги научиме компјутерите да го гледаат светот вака**. Следен

чекор е да навлеземе подлабоко во техничката страна и да видиме како функционира машинското учење.

Трет чекор: Машинско учење и процеси на вештачка интелигенција

Што се машинско учење и вештачка интелигенција и како можеме да ги користиме за да создадеме агент за вештачка интелигенција?

Подучувањето на компјутерите со користење на исти правила како и луѓето се нарекува машинско учење. Наместо да им даваме правила, им дозволуваме самите да ги пронајдат правилата и шемите. Машинското учење е еден од основните процеси во создавањето на вештачка интелигенција (ВИ), што нѐ носи до следната дефиниција.

Транспарентна технологија

Вештачка интелигенција е кога компјутерски системи можат да извршуваат задачи кои вообичаено бараат човечка интелигенција, како што се:

- визуелна перцепција;
- говор;
- препознавање емоции.

Систем создаден со овој метод понекогаш се нарекува „**Агент за вештачка интелигенција**“ (“**AI Agent**”) и наша одговорност е да го „**обучиме**“ овој агент. За да го направиме ова, компјутерите треба да видат примери за учење. Ист процес се случува кога ги тренираме нашите деца. Овие примери се нарекуваат **податоци**. Собираме многу податоци и рачно ги нотираме. Тоа значи дека имаме аудио- викање (говорење) од 3 секунди и текст кој вели дека лицето **70%** има лутина (емоција лутина). И имаме рафално смеење од 2 секунди и текст кој вели дека лицето **90%** е среќно (емоција среќа). Сепак, мора да се погрижиме да останеме објективни. Затоа дозволуваме голема група луѓе со различни перцепции (на пример, културни и родови варијации) да ги коментираат овие аудио делови и откако ќе го просечеме резултатот и ќе се осигураме дека овие луѓе се согласуваат за нешто, го храниме компјутерот со нив.

Податоци како храна за алгоритам

Ние генерално ги делиме нашите податоци во 3 групи:

1. Податоци за обука;

2. Податоци за развој;

3. Тест- податоци.

Прво, ја користиме базата на податоци за обука за да ги обучиме моделите. Моделот е **невронска мрежа** која ја има улогата на мозокот на нашиот агент со вештачка интелигенција. Тоа е фаза во која го учиме нашето **бебе- компјутер**. Пробуваме различни методи како GAN (Generative Adversarial Networks) (Генеративни советнички мрежи) или CNN (Convolutional Neural Networks) и разновидни алгоритми додека не најдеме задоволителен резултат. Потоа, го користиме **вториот** сет на податоци кој обично е најголем што се развива на него. Тоа е **тинејџерски период** кога вештачката интелигенција пробува сè за себе. Тоа е некоја форма на **игралиште на животот за нашата вештачка интелигенција**. Вештачката интелигенција пробува илјадници различни променливи додека не стигне до фаза во која работите веќе не се менуваат толку многу.

Сега кога нашата тинејџерска вештачка интелигенција порасна доволно, ние како родители **треба да го тестираме** пред да може да излезе во неговиот **возрасен живот** и да биде корисна алатка за луѓето.

Перформанси на човечко ниво (Human-Level-Performance)

Креираме “**test- set**”. Тоа е збир на податоци што нашата вештачка интелигенција никогаш претходно не ги видела. Ниту во обуката ниту во фазата на развој. Колку е добра таа во овој тест, дефинира колку точна и интелигентна стана нашата вештачка интелигенција. Вообичаено, ние **го дефинираме овој праг според перформанси на човечко ниво**. Тоа значи дека очекуваме нашата вештачка интелигенција да биде што е можно попрецизна за перцепција на човечко ниво. За некои задачи како автоматско препознавање на говор, имаме прилично јасна основна вистина (што кажува личноста)и можеме да се стремиме да достигнеме

100%, но станува незгодно за некои други задачи како што е препознавање емоции (луѓето не се согласуваат едни со други).

За такви задачи, сепак ја земаме предвид објективноста и очекуваме вештачката интелигенција да работи што е можно поблиску до нас. Ако кажам реченица со релативно радосен глас, а потоа ја снимаме и дозволиме да ја слушаат 4 случајни луѓе на оваа планета, на 3 може да им биде радосна и на 1 неутрална. Во меѓувреме, верувам дека бев радосен, но само 75% од луѓето ме доживуваат како радосен, а бидејќи емоциите се нешто што го изразуваме, мораме секогаш да ја имаме предвид перцепцијата. Со други зборови, емоционалната перцепција на човечко ниво, во овој случај, е 75% и очекуваме нашата вештачка интелигенција да работи што е можно поблиску до ова ниво.

Така може да се обучат компјутерите да ги разбираат човечките емоции.

Технологија

audEERING's ERA demo прикажува интеграцијата на **Voice AI** во разговорни агенти за вештачка интелигенција и претставува значајна пресвртница во доменот на афективното пресметување (откривање ниво на емоции).



Овозможувајќи им на асистенти со вештачка интелигенција да сфатат и да одговорат на човечко емоционално изразување, оваа иновација ги револуционизира нашите интеракции со виртуелни придружници.



Слика. devAIce® Web API 4.2 & 4.3 Updates – Exciting Additions to audEERING's Web API

Најнови подобрувања во верзиите на Web API на devAIce®. Овие ажурирања носат мноштво нови функции и подобрувања за да ги поттикнат програмерите да го искористат целосниот потенцијал на оваа напредна технологија.



Слика. devAIce® SDK 3.8 and 3.9 Updates – New Powerful Module

devAIce® SDK 3.8 и 3.9, носат значителни подобрувања и нови функции за да ги зајакнат истражувачките можности на вештачката интелигенција.

devAIce Web API 4.0.0

Автентикација е со едноставен API клуч кој се пренесува како HTTP заглавие во секое барање, што е добро поддржано од вообичаените алатки за развој на веб.

Во моментов постојат официјални стабилни клиенти за следните јазици: Python, C#, Java и JavaScript.



4. Роботите учат емпатија

Лидерот на германскиот пазар за анализа на говор базирана на вештачка интелигенција **audEERING** развива хуманоидни роботи со способност за емпатија заедно со Hanson Robotics, еден од водечките светски експерти за роботи. **Преку гласовна анализа на audEERING, роботите препознаваат емоции**. Ова за првпат овозможува **социјална интеракција меѓу луѓето и роботите на највисоко ниво**. Се очекува препознавањето на емоции да овозможи и подобро користење на роботите во секојдневниот живот, особено во здравствениот сектор и здравствена заштита.

Hanson Robotics е познат во светот по своите реалистични роботи. Изгледот и социјалните вештини се моделирани според луѓето. До каде стигнал раз-

војот на машини слични на човекот, покажува најпознатиот робот и прв робот со **официјално државјанство, Софија**. Таа веќе разговараше со тогашната германска канцеларка Ангела Меркел во 2018 година и со Обединетите нации, демонстрирајќи како луѓето и машините можат да комуницираат.



Соработката помеѓу Hanson Robotics и auDEERING, исто така, има за цел дополнително да ги развие социјалните вештини на Sophia Hanson (auDEERING и Hanson Robotics развиваат роботи со највисока социјална компетентност за секојдневна употреба). Во иднина, Софија ќе препознава емоции во разговор и како резултат ќе може емпатично да одговори. Во активности на медицински сестри и други области погодени од недостигот на квалификувани работници, такви роботи опремени со социјална вештачка интелигенција можат да помогнат во иднина. Преку поддршка за дијагностика со вештачка интелигенција, автоматската анализа на говор ја подобрува превентивната грижа и терапија за пациентите и ги заменува скапите тестови на крвта, меѓу другото. На пример, audEERING неодамна беше во можност да покаже во едногодишно истражување со пациенти со MS (мултиплекс склероза) дека индикатори поврзани со болеста на **мултиплекс склероза**, како **замор** и **депресија** може да се откријат преку кратки, редовно изведени тестови за говор. Притоа, технологијата за вештачка интелигенција на aud-

EERING, која е развиена според етички принципи, ја анализира содржината на она што е кажано само врз основа на **акустика** и **фонетика**, што овозможува најдобра можна заштита на податоци.

Емпатија преку препознавање на емоции

За да им овозможи на роботите да комуницираат со луѓето, Hanson Robotics веќе анализира човечки **изрази на лице** (невербални), **контакт со очи** и **говор** (вербален) и работи со **емоционален израз на лице** на роботи слични на човекот. Со патентирана технологија за анализа на глас и говор на auDEERING, роботите сега ги препознаваат емоциите на говорникот и **можат да се стават во емоционалната состојба на другата личност**. На пример, ако роботот го анализира гневот кај колега на некој човек преку гласот, тој би можел да реагира мирно и смирувачки, а ако открие радост, би можел да реагира особено еуфорично. Технологијата од audEERING за гласовна вештачка интелигенција на тој начин го затвора преостанатиот јаз во паралингвистичките информации (висина/ pitch на глас, тон на глас) и овозможува социјална интеракција човек-машина од таков вид што претходно беше невозможна.

Анализата на гласот и говорот во согласност со приватноста на audEERING, која веќе се користи на меѓународно ниво од GN/Jabra и други водечки компании, ги подобрува перцептивните способности на роботите. Во комбинација со постоечките модели на вештачка интелигенција, како што се: **препознавање говор базирано на текст** (ASR- AI models such as text-based speech recognition) или **визуелна анализа**, роботите исто така **можат да согледаат иронија или противречности во човечкиот израз**, на пример.

Д-р David Hanson, основач и извршен директор на Hanson Robotics вели:

„Емоциите се основа за сите меѓучовечки интеракции. Со препознавање на емоции, ние сме голем чекор поблиску до нашата цел за природна интеракција помеѓу

луѓето и роботите. За нас беше важно анализата на податоците да се врши според највисоки етички стандарди и дека вештачката интелигенција придонесува роботите да работат за благосостојба на општеството. Технологијата на audEERING нè убеди во овој поглед“.

Dagmar Schuller, експерт за вештачка интелигенција и управен директор на audEERING е исто така позитивен во врска со соработката:

„Емпатијата е најважен фактор за успех на нашето време. Особено во областа на зголемување на автоматизација и оптимизација на процеси, емпатијата и подобро разбирање се поважни од кога било за нашата благосостојба. Со нашата технологија во врска со социјалната роботика, особено сакаме да ја подобриме оваа способност за емпатија во вештачката интелигенција насочена кон човекот и да создадеме додадена вредност за сите луѓе“.

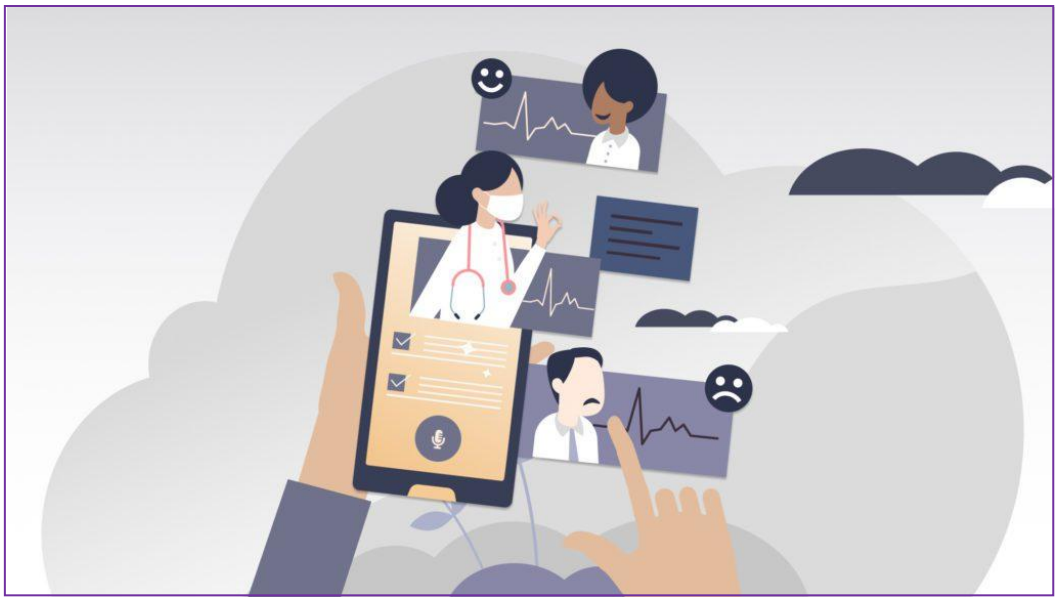
Најприродна форма на интерперсонална комуникација е аналогниот разговор. Во него соговорниците разменуваат информации – вербално и невербално (изрази на лице, говор на тело) и паралингвистички (глас, prosody). Досега не беше можно да се реплицираат сите овие нивои во разговор со робот, така што не можеше да се случи целосна социјална интеракција помеѓу луѓето и машините. Патентираната технологија на германскиот пионер на вештачката интелигенција audEERING ќе го промени тоа со препознавање на емоции.

Откривање (детекција) на Corona од глас (говор)

Многу болести може да се откријат од гласот бидејќи тие влијаат на продукција на говор. Ова го вклучува и СОВИД-19 бидејќи е респираторно заболување. Како болеста може да се открие преку гласот?

Voice biomarkers за COVID-19

Научни студии покажаа дека кашлањето, кивањето и други промени во гласот се сигурни биомаркери за откривање на СОВИД-19. Тоа значи дека болеста влијае на респираторниот систем. Со многу релевантни податоци од болни и здрави лица, модел на вештачка интелигенција може да се обучи да разликува дали некое лице е заразено со вирус или не.



Портал AI SoundLab за собирање гласовни податоци за COVID-19

Порталот **AI SoundLab** е развиен за собирање релевантни податоци за обука на вештачка интелигенција за откривање на COVID-19. Се собираат податоци од:

- Говор на лице кое чита текст;
- Звуци на кашлање;
- Звуци од кивање;
- Звуци од дишење: вдишување и издишување.

Платформата овозможува да се следат симптомите во подолг временски период. Ако некој добие позитивен тест со PCR или брз антиген тест, се замолува повторно да ја направи студијата за да се види како влијае на неговиот глас.

Предност на гласовен тест за паметни телефони

Врз основа на податоци што се собрани со платформата **AI SoundLab**, потоа може да се создаде апликација, која може да му овозможи на секој корисник на паметен телефон да прави кратки гласовни снимки и да добие индикација дали има COVID-19. Во случај на голема веројатност да бидат позитивни, корисниците треба да го потврдат својот резултат на пример со PCR- тест.

Откривање болести од анализа на говор со Вештачка Интелигенција

Што може да направи вештачката интелигенција? Аудио вештачката интелигенција навистина може да ја придвижи медицината. Болести како Паркинсонова депресија и СОВИД-19 имаат една заедничка работа: тие влијаат на начинот на кој звучи гласот на една личност.



Клучот е во аудио- вештачка интелигенција

Кои се овие промени и како може да се откријат за да се даде индикација дали некое лице е заразено или не? Со COVID-19 тоа е едноставно и сложено во исто време. Пред сè, тоа е респираторно заболување, што значи типични симптоми се кашлање и кивање. Сепак, најверојатно има многу асимптоматски заразени лица, кои речиси и да не покажуваат симптоми или не ги препознаваат како

такви, туку ја шират болеста. Понатаму, обична настинка или грип би можеле да звучат слично? Ако е така, како може сигурно (reliably) да се откријат симптомите?

Прецизност од над 80% за откривање на типични симптоми

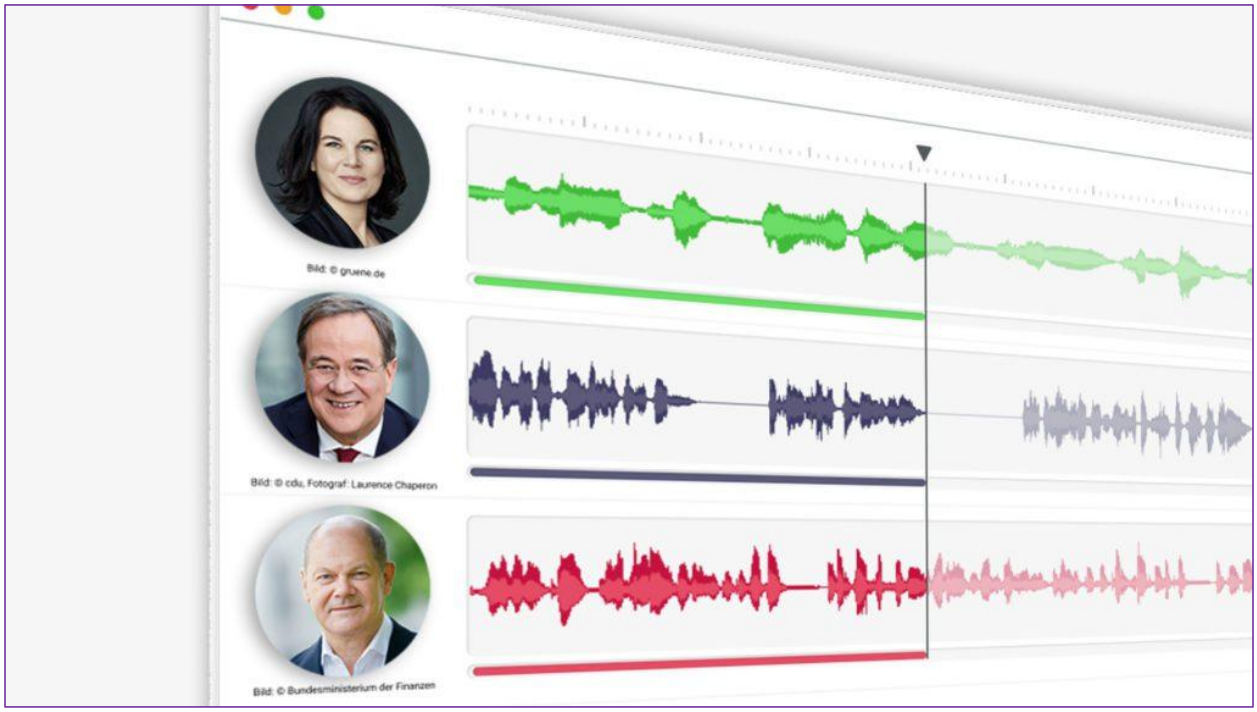
Одговорот е дека аудио- вештачката интелигенција може да открие мали разлики во гласот на една личност, ако моделот зад механизмот за откривање е обучен со илјадници збирки податоци. Врз основа на примероци од глас од заразени и здрави луѓе, моделот со вештачка интелигенција може да научи обрасци кои се типични за COVID-19 и да открие инфекција со прецизност од над **80%**.



Доколку постои голема веројатност евидентираните симптоми да се предизвикани од КОВИД-19, ќе се препорача дополнително тестирање од лекар. На ваков начин бројот на негативни резултати од тестот може драстично да се намали. Апликацијата може на пример да даде индикации дека многу луѓе имаат само обична настинка и дека не е потребен тест за корона. Ова му помага на здравствениот систем бидејќи неговите капацитети можат да се искористат за оние кои веќе покажуваат голема веројатност да бидат заразени со корона вирусот. Понатаму, решението е добро за животната средина бидејќи го намалува пластичниот отпад од тест- стикови и други медицински материјали.

5. Анализа на емоционални карактеристики во говор на политичари со користење на вештчка интелигенција

Германските политичари Annalena Baerbock, Armin Laschet и Olaf Scholz беа анализирани од Аудио вештачката интелигенција на audEERING додека се кандидираа за канцелар. Анализите на audEERING не се однесуваа на содржината на говорите, туку на акустични, лингвистички и емоционални карактеристики. Овие се идентификувани со користење на научни процедури и наградуваната технологија за вештачка интелигенција devAIce™ и openSMILE ™ на audEERING.



Акустична анализа: openSMILE ™

Дел од карактеристиките што беа анализирани за акустичната анализа беа **паузите**, **ритамот** и **артикулацијата**. Говорењето не само што несвесно се обликува врз основа на овие карактеристики, туку може да биде и свесно дизајнирано. Начинот на зборување е важен дел од доставувањето информации и водење на публиката низ нив во контекст на говорот.

Паузи за време на зборување

Говорите на Armin Laschet покажуваат најголем број на паузи. Голем број на паузи може да изгледа харизматично и доминантно и да има позитивен ефект врз разбирливоста. Во споредба со другите кандидати, таа дополнително гради кратки фрази помеѓу паузите, што и помага на публиката да ја следи.

Карактеристики на слогови, брзина на зборување и ритам

Во однос на параметар на ритам, анализата покажува резултати за слогови во секунда. Високата стапка на говор во комбинација со други параметри како што е **артикулација** е поврзана со **харизма** и **доминација**. Со пет слога во секунда, Олаф Шолц зборува малку побрзо од останатите кандидати. Сепак, девијацијата на брзина е најниска во говорите на Олаф Шолц и највисока во говорите на Baerbock. Оваа динамична карактеристика помага да се плени публиката и да се води низ информациите. На крајот на краиштата, Олаф Шолц исто така го менува својот ритам на зборување помалку од Baerbock и Laschet. Монотоното говорење може да биде заморно за публиката.

Прецизност на артикулација

Во споредба со другите кандидати, Annalena Baerbock зборува со највисока артикулациска прецизност Олаф Шолц со најниска. Ниска прецизност на артикулацијата е во нејзиниот начин на дејствување, чин на балансирање помеѓу **смиреност** и **замор**. **Прецизната артикулација** е поврзана со **енергетска напнатост на телото**, што пак има позитивен ефект врз тоа како личноста ќе биде перципирана.

Анализа на емоции: возбуда и доминација со користење на ВИ

Во областа на анализа на емоции се испитува акустичната експресија на емоционална активација и доминација. Тука Baerbock е најактивна во нејзиното емотивно изразување и на тој начин може да сигнализира вклученост (инволвираност). Нејзините говори се и најдоминантни. Ова ја прави да изгледа најпосветена и најемпатична во споредба со другите кандидати.

6. Емоционална вештачка интелигенција за комфорт (удобност) на пациент

Човечката интеракција помеѓу лекар и негов пациент е еден од најважните процеси кога станува збор за дијагноза и третман на која било болест.

Комуникацијата е клучна за секој третман

Научно е докажано дека комуникацијата, како и гестови и изрази на лице, имаат влијание врз благосостојбата и процесот на заздравување на пациентот.

Поради тоа, многу е важно да се обезбеди можност на здравствените работници да прават грешки во безбедна средина и да се учат од нив.



VR (виртуелна реалност) решение како безбедна средина за обука

VR може да ја даде оваа безбедна средина. Компанијата The Simulation Crew разви решение за обука засновано на VR, кое вклучува **откривање емоции** преку audEERING. Преку VR очила и VR-контролер, корисникот има чувство дека навистина стои пред пациент и се грижи за него или неа.

Обука за секаков тип на ситуација

Доколку системот покажува дека пациентот доживува стрес, корисникот има можност да примени техники за ослободување од стрес за индивидуално да го смири пациентот. Корисниците исто така можат да испробаат различни техники со различни типови на пациенти, на пример, за пациенти кои се особено срамежливи или лути. Затоа, системот на секој корисник му нуди совршена опција за подобрување на грижата за пациентите без никакви ризици.

7. Аудио- вештачка интелигенција за автономни автомобили

Пазарот за автономно возење се предвидува да порасне од 5,7 милијарди во 2018 година на 60 милијарди американски долари во 2030 година. Ова го покажува потенцијалот што оваа технологија ќе го има во иднина и е една од причините зошто германската влада обезбеди средства за проектот SEMULIN.



Подобра комуникација помеѓу возачот и автомобилот

Целта на проектот е да се развие самоподдржлив, природен интерфејс човек-машина (HMI- human-machine interface) за автоматско возење користејќи мултимодални влезни и излезни модалитети. Тоа се меѓу други:

- изрази на лице;
- гестови;
- погледи;
- говор.

Во врска со околината на возачот (на пример во автомобил) ова е холистички пристап кон HMI кој е прилагоден на потребите на човечките сетила. Овој систем ќе ја подобри интеракцијата помеѓу автомобилот и возачот и ќе создаде подобро корисничко искуство во автономните автомобили.

Откривање на емоции во автомобил

Вештачка интелигенција ќе биде составен дел на овој систем. audEERING ја обезбедува својата добро позната Audio AI за откривање емоции врз основа на глас (говор). Основа за ова се **машинско учење** и **афективно пресметување**. Со подобрување на комуникацијата помеѓу возачот и автомобилот, автономното возење ќе стане поприродно и поудобно.

Комбинација на различни сензори

Додека audEERING се фокусира на анализа на гласот на возачот, другите партнери во конзорциумот како **Blickshift** обезбедува **технологија за следење на очи** или **Infineon** придонесува за **откривање гестови**. Комбинирање на влез од овие различни сензори ќе овозможи откривање на положбата (состојбата) и расположението на возачот. Интелигентниот систем во автомобилот потоа може да реагира соодветно и да дава индивидуализирани препораки.

Конзорциум на светски лидери во автомобилскиот сектор. Проектот е финансиран од германското сојузно министерство за економски прашања и енергетика. Fraunhofer IIS е лидер на конзорциум од осум партнери, меѓу кои и Infineon и Mercedes-Benz.

Системот ја олеснува интеракцијата, го зголемува корисничкото искуство и прифаќање на автономно возење во сите области. Методите за мерење на задоволството на корисниците развиени во текот на проектот служат како основа за слично дизајнирани проекти.

Мотивација и предизвик

Корисничкиот интерфејс игра главна улога во автоматското возење бидејќи, поради зголемена сложеност и барања, мора да поддржува повеќе функции, да обработува информации и да биде многу лесен за користење.

Во моментов има ограничувања за природна интеракција помеѓу возачот, патникот и возилото, особено кога станува збор за префрлување помеѓу различни режими (изрази на лице, говор, осветлување, звучници итн.) или нивна комбинација.

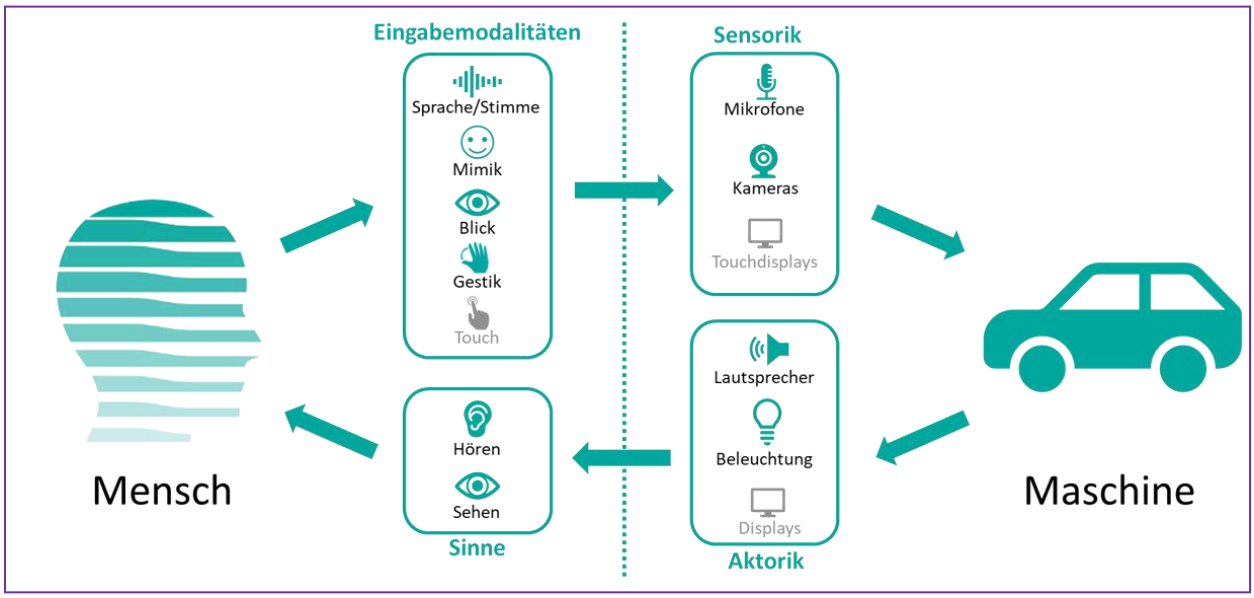
За да се овозможи оваа човечко-центрирана интеракција, мора да се земат предвид и различните модалитети и контекстуални информации и значајно да се

поврзат заедно. Посебен предизвик лежи во правилно идентификување на точна намера на корисникот да дејствува и да генерира дејствија за системот.

Интерфејс човек-машина со интелигентна интерпретација на сензор и фузија на податоци

Со цел да се развие човек-центриран човек-машина интерфејс со прилагодена системска архитектура, земајќи го предвид целокупниот контекст, се истражуваат сите достапни модалитети за да може интелигентно да се толкуваат и поврзат собраните податоци од сензорот.

За да се направи ова, се користиат воспоставени технологии, како што е софтвер за анализа SHORE ® за препознавање емоции базирани на видео и други вградени сензори, за однапред да се обработат податоците и интелигентно да се толкуваат. Користејќи процедури засновани на правила и поддржани од вештачка интелигенција и нивна мултимодална имплементација, се поврзуваат податоците меѓу себе.



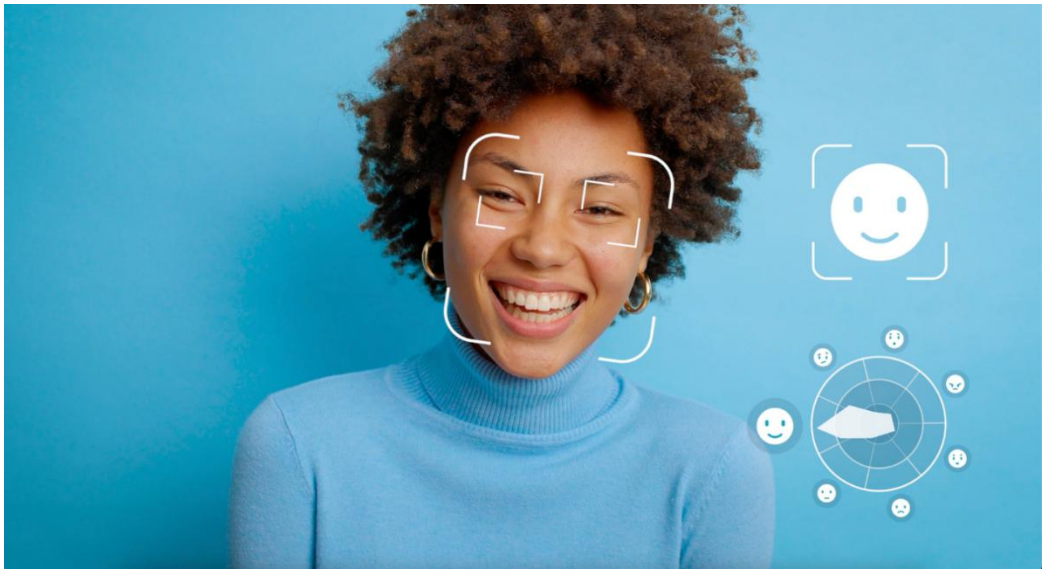
Со помош на интелигентна интерпретација и фузија на сензорот, може:

- да се дадат изјави за состојбата на корисникот;
- да се идентификува неговата намера да дејствува;
- да се изведат можни реакции.

Нови пристапи, како што е интерактивно учење, имаат за цел да обезбедат системот постојано да се прилагодува на потребите на соодветниот корисник.

8. Софтвер за препознавање изрази на лице SHORE ® – Емпатија за нашата технологија

Изрази на лице на човек кажуваат повеќе од зборовите. Тоа е алатка и **превозно средство за нашите чувства**, **свесни** и **несвесни**. Нашите изрази на лицето може да ги потврдат или контрадикторат вербалните изјави.



Со цел овие несвесни или контрадикторни емоционални изрази да бидат препознатливи на лицето, анализа на изрази на лице поддржана со софтвер може да биде важна компонента во нашиот систем, особено во области на примена во кои се потребни автентични, спонтани реакции на корисникот.

Анализа на емоции и откривање на отчукувања на срце директно во прелистувач (browser)

Благодарение на WebAssembly, SHORE ® сега работи директно во прелистувачот: без приклучок, на сите уреди и независни од платформата. SHORE ® може лесно да се интегрира во апликации базирани на прелистувач, како што се

алатки за истражување. Податоците се обработуваат само на користениот уред и се препраќаат како чисти метаподатоци. Единствен неопходен инструмент е **веб камера** или **камера за мобилен телефон**.

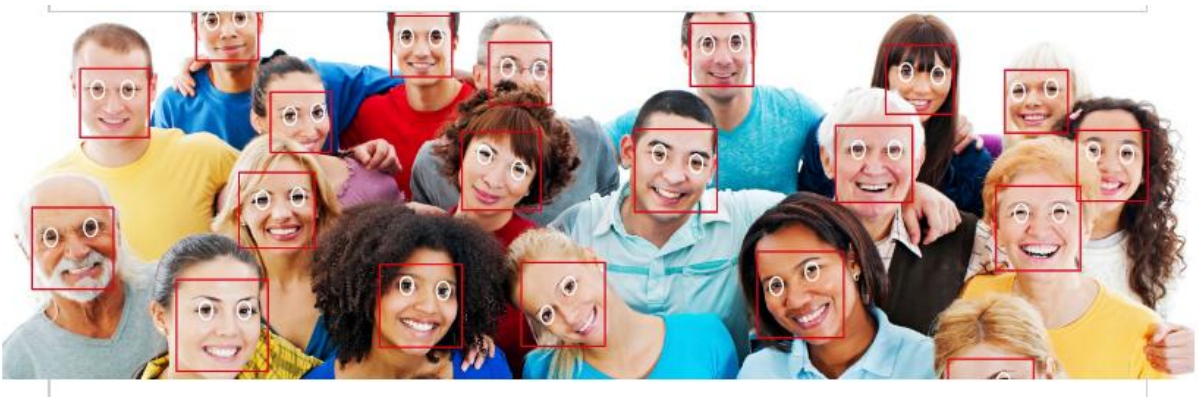
SHORE ® препознавање лице: Брзо, робусно и во реално време

Заштита на податоци и обработка на податоци

Го користиме провајдерот YouTube за да интегрираме видеа. Како и повеќе веб-локации, YouTube користи cookies за да собира информации за посетители на неговата веб-локација. Стартувањето на видеото може да предизвика операции за обработка на податоци. **SHORE ®** нема никакво влијание врз тоа. Дополнителни информации за заштитата на податоците на YouTube може да се најдат во нивната декларација за заштита на податоците.



SHORE ® ги евидентира **основните емоции** (среќа, лутина, тага, изненадување, исплашеност, загрозеност, неутралност) (happy, angry, sad, surprised, frightened, disgusted, neutral) како и **валентност**, односно дали преовладувачката емоција е позитивна или негативна. SHORE ® исто така ја проценува **возраста** и **полот**. Може да се генерираат сигурни резултати и за RGB и за инфрацрвени камери.



Оваа технологија овозможува да се прави анализа на израз на лице за одредени лица кои се во групен стимул. Се селектира лицето на одреден човек и се детектира каква емоција изразува во одреден момент.



Слика. Одредување каква емоција изразува секој поединечен човек во групна слика

SHORE ® бесконтактно го снима **пулсот** на човекот. Се што ни треба е стандардна веб- камера. Во врска со **препознавање емоции**, SHORE ® е уште подобро способен да дава изјави за **психофизиолошката состојба** на една личност.

SHORE ® – области на примена

1. **Развој на производ ориентиран кон целни групи со SHORE ®**

- Дали е пријатно чувството за производот?
- Работи ли ударната линија во комерцијалата?
- Како се прима производот?

Овие и многу други прашања се разјаснети во истражување на пазарот и во тестирање на производи и UX. За ова се користат различни алатки, како што се анкети на клиенти врз основа на прашалници.

Проблем: Непосредни, неизвалкани (непристрасни) реакции на испитаниците и емоционалните ефекти често не се земаат предвид.

За да им понудиме на нашите клиенти најдобра можна услуга или најодличен производ, треба точно да знаеме што сакаат клиентите и како тие соодветно да ја прилагодат нашата понуда.



SHORE ® е добар додаток и вистински градежен блок за нашата алатка за истражување на пазарот

Со имплицитно снимање на несвесните реакции на клиентите, добиваме научно засновани квантитативни податоци за емоционалното влијание на нашиот производ - брзо и без заштита на податоците.

Благодарение на WebAssembly, SHORE ® работи во прелистувачот и може да се интегрира во стандардни апликации на прелистувачите, како што е потребно

за секое истражување. Анкети со функционалности SHORE ® може да се промовираат дома или во движење, на компјутер, таблет или мобилен телефон – целосно флексибилни и приспособени на нашето истражување.

2. Силни емоции за силни игри



Софтвер за препознавање емоции SHORE® – нова димензија надвор од контролер

Тестови за употребливост на игра

Во развој на одредена игра, последната фаза пред објавување на играта е особено возбудлива: Дали играта ги буди токму емоциите што програмерите ги имале на ум во одреден момент?

Токму тука влегува SHORE®.

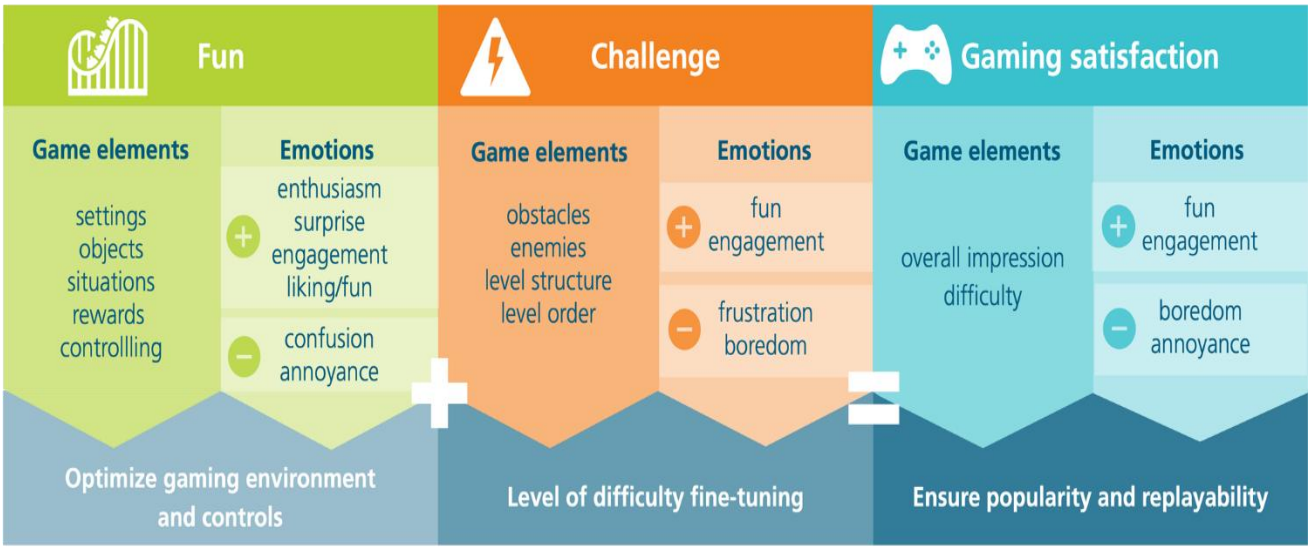
Тестови за употребливост со SHORE® покажуваат колку емотивно вклучени се испитаниците - вистински, во реално време и во согласност со прописи за заштита на податоци. Овој софтвер за анализа на емоции го зема предвид:

- емоционалното влијание на играта;
- ги идентификува факторите за успех;
- дава предлози за можни прилагодувања.

На тој начин може да се направат потребни оптимизации пред почетокот на продажбата и ќе се зголемат шансите за успех на одредена игра.

Емоционално играње со SHORE®

Преку анализата на слики базирана на вештачка интелигенција, SHORE® може да ја препознае емоционалната состојба на играчот во реално време и дали е во позитивно или негативно расположение (препознавање на валентност).



Програмерите на видео игри можат да ја прилагодат играта на емоционалната состојба на играчот со интегрирање на различни параметри. Реакциите на играчот го одредуваат понатамошниот тек на играта, а со тоа и развојот на силен гејм-плеј.

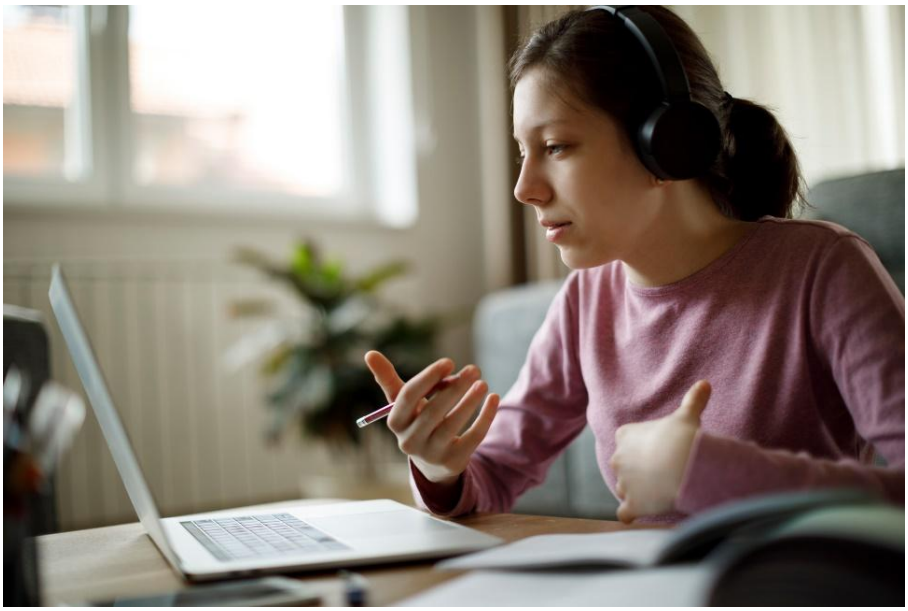
1. Оптимизирано е-учење со SHORE ®

Индивидуално прилагодена содржина за учење промовира самонасочено, успешно учење.

Е-учење со SHORE ® – за подобрени процеси на учење

Во областа на дигитално учење, особено е важно содржината да биде подготвена на начин насочен кон корисникот со цел да се избегне чувство на **преоптоварување** или **досадување** на учениците.

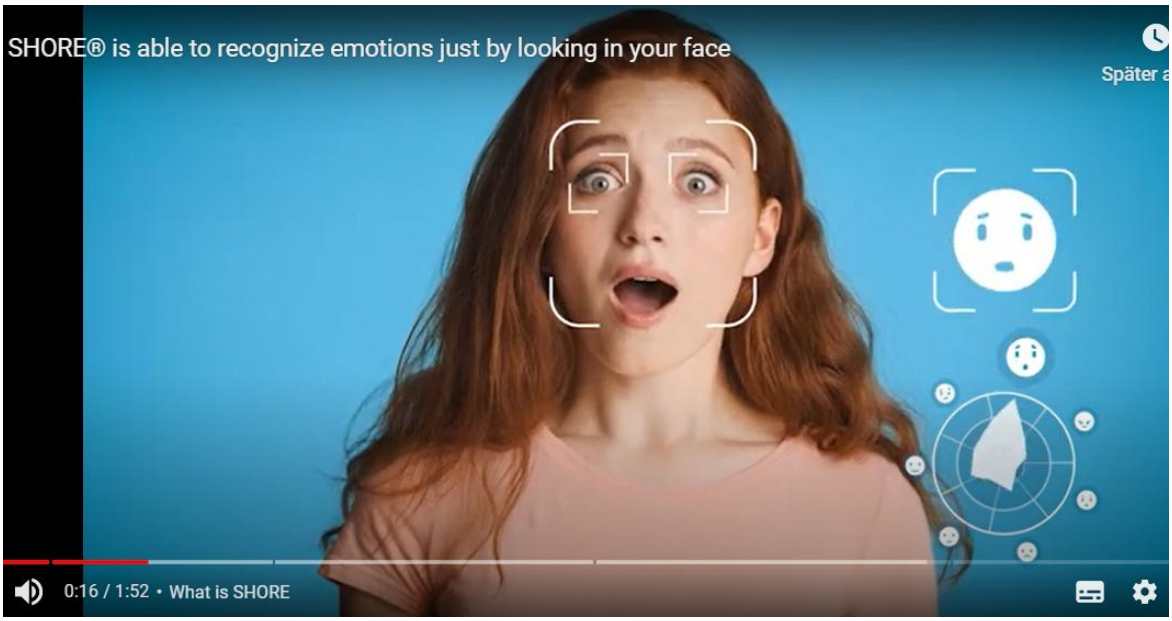
Непознати теми може да покренат прашања за разбирање. Од друга страна, ако содржината е дефинирана со премногу низок праг, **вниманието** страда.



Визија: Програма за учење препознава во реално време дека содржината за учење предизвикува **проблеми** или дека корисникот е **расеан**.

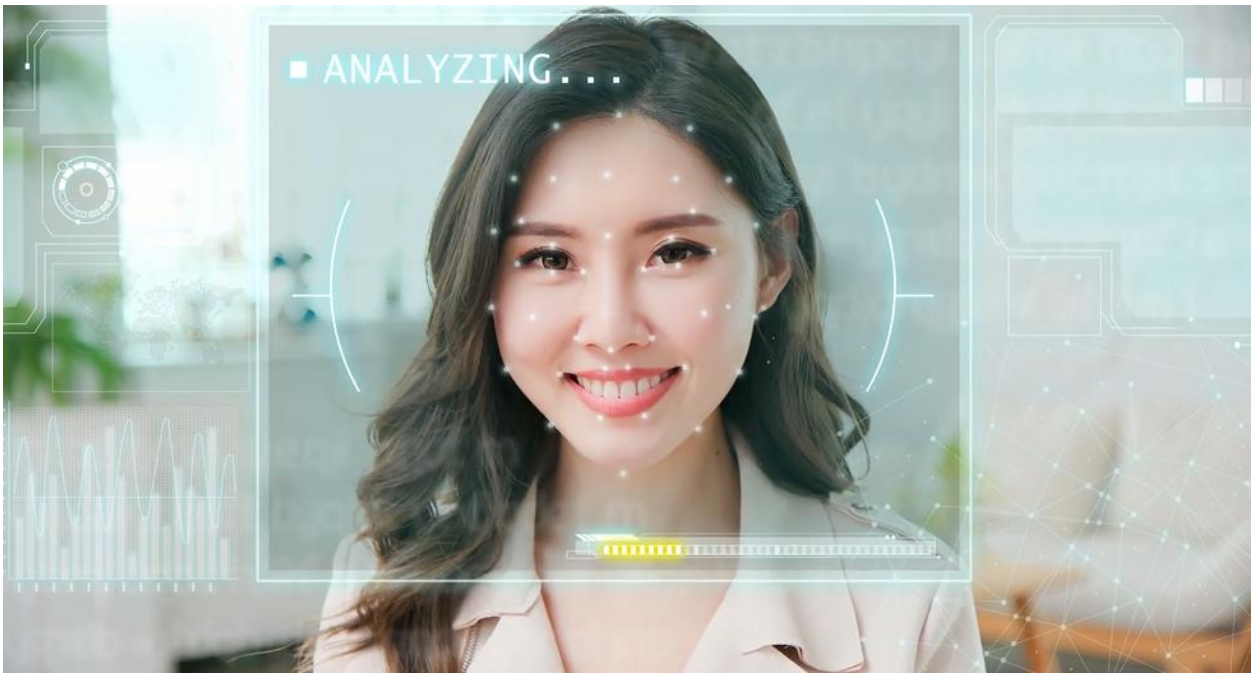
Преку паралелна анализа и евалуација на реакции на корисници, содржина за учење што не е разбрана може автоматски да се повтори, да се понуди помош или да се прескокнат поединечни програмски точки.

Со користење на аватар, програма за учење може директно и индивидуално да комуницира со корисникот и лично да одговори на неговата реакција.

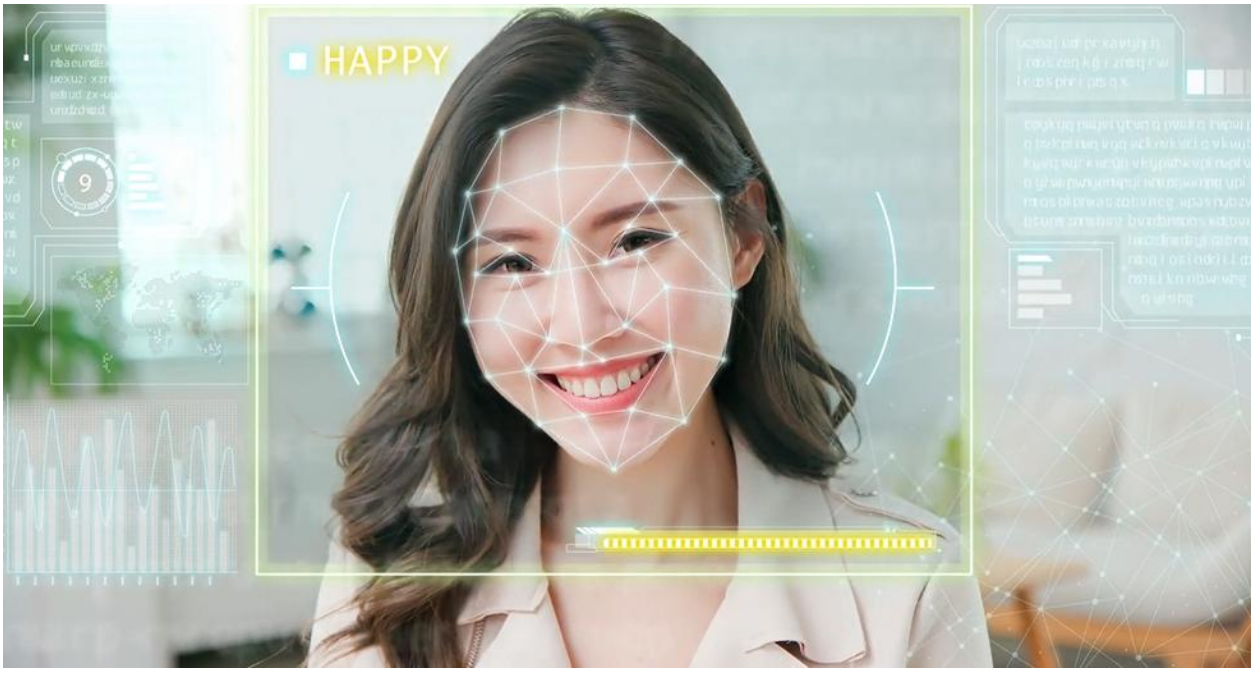


Оваа иновативна технологија може да препознае емоции само со следење и анализа на израз на лице.

Биосензорот **прво** го анализира изразот на лицето.



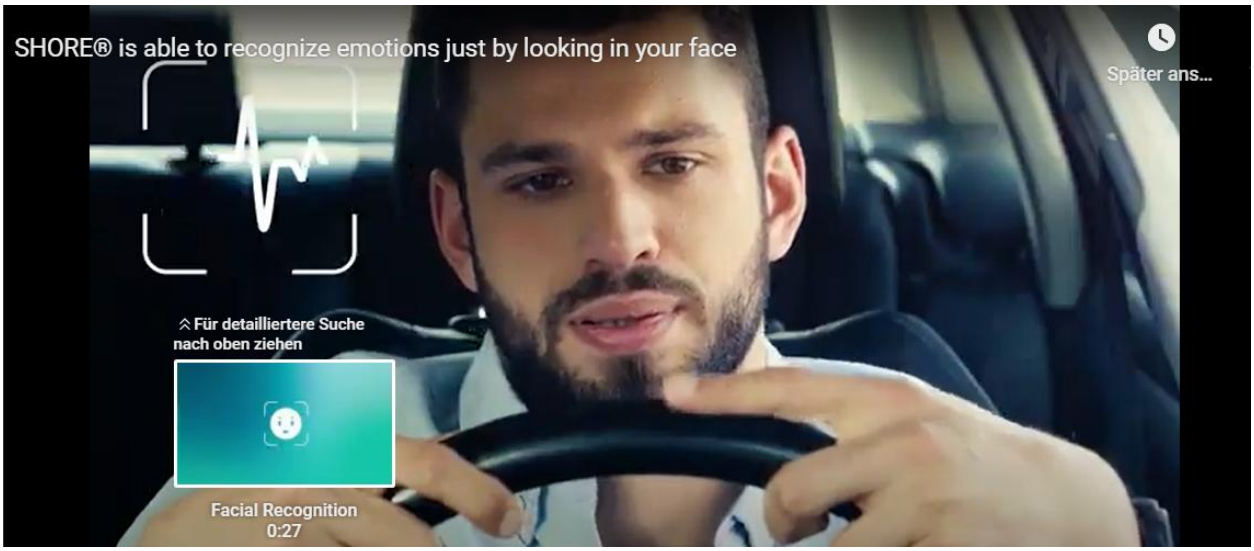
Втор чекор: Прикажува каква емоција е изразена.



2. Бесконтактно откривање статус на возач со SHORE ®

Според истражување на **ABI Research**, камери за набљудување на возачи, со прогноза од 17,5 милиони продадени единици за 2026 година, се смета за клучна технологија за системи за откривање статусот на возачот што ги бараат регулативите на ЕУ од 2024 година и за полуавтономно возење.

Како што автомобилската индустрија се буди со безбедносни предизвици за управување со човечки интеракции со автономни системи, ABI Research идентификува **системи за следење на возач**и базирани на камера (DMS- driver monitoring systems) како главни овозможувачи на безбедно, полуавтономно возење.

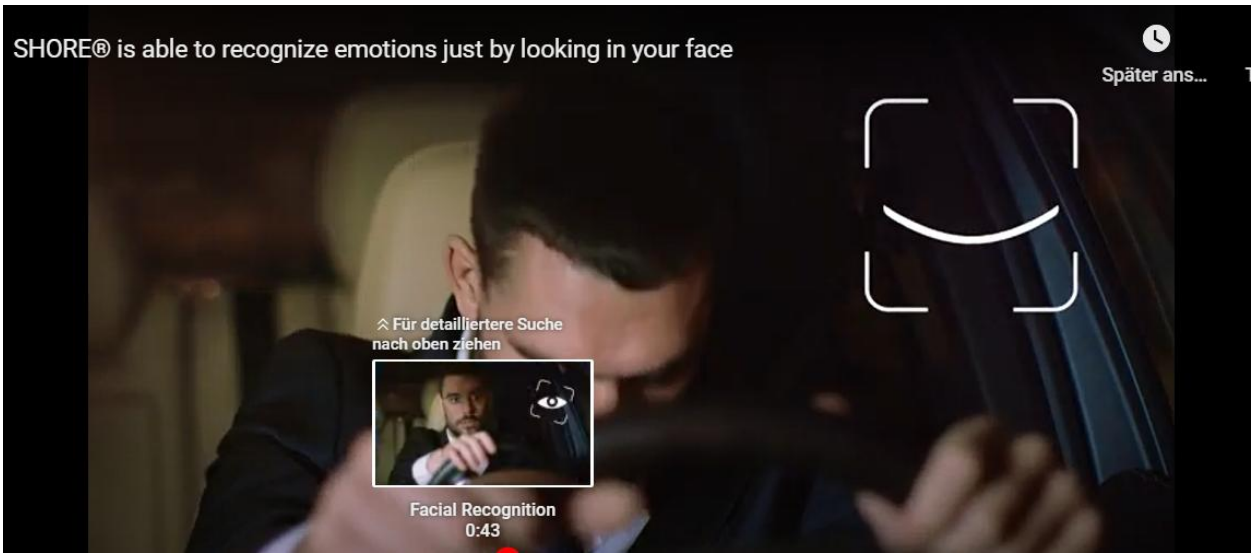


Голем број лансирани полуавтономни системи од OEM како Mercedes-Benz, Nissan и Tesla ја истакнаа важноста на робусен интерфејс за човечка машина, или HMI (human machine interface), во сценарија за кои е потребен автоматизиран систем да работи во тандем со човек-возач. Употребата на DMS базирана на камера за да се обезбеди хост- автономен систем со сеопфатно разбирање за биометриска достапност на возачот, е основа на безбеден, полуавтономен HMI.

Биометриска достапност ја опфаќа:

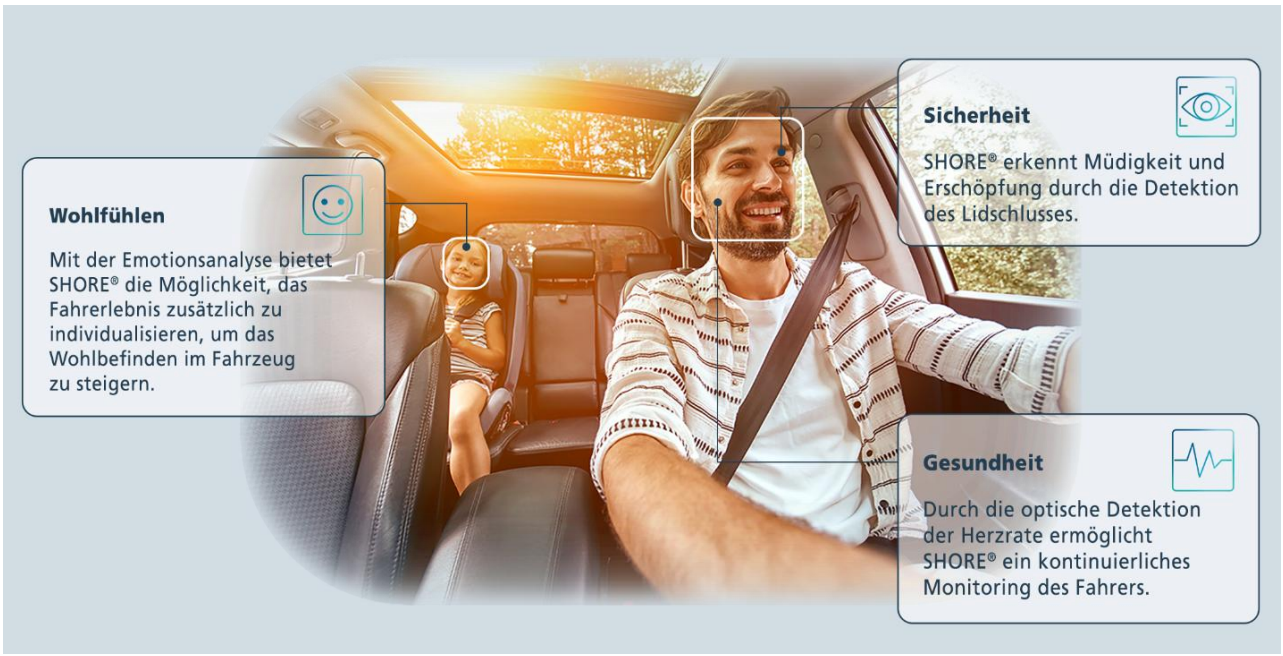
- будноста на возачот;
- негов или нејзин ангажман во процес на возење, па дури и

- способност на возачот рачно да комуницира со системот по потреба.



Клуч за овозможување на таков холистички систем за следење на возач (DMS) е користење на внатрешни камери:**стереоскопски** или **time-of-flight** - за да се идентификуваат и следат:

- црти (изрази) на лице;
- насока на поглед (следење на очи);
- положба на горниот дел од телото.



ABI Research идентификува голем број компании за анализа на вид активни на овој простор, вклучувајќи ги EDGE3 Technologies, FotoNation, Jungo Connectivity и Gestigon.

Софтверската библиотека SHORE ® обезбедува податоци базирани на камера за состојбата на возачот и на другите патници, што го поддржува безбедносниот аспект и може позитивно да ја подобри благосостојбата во возилото.

SHORE® обезбедува постојано висок квалитет на податоци и во симулатор за возење и директно во возилото. Во реални услови се развиваат и адаптации на софтверска библиотека специфична за клиенти.

9. Откривање на емоции од говор

Емоционалното изразување се пренесува преку гласот. Во нашиот секојдневен живот, ние автоматски вршиме аудитивна анализа кога зборуваме со некого. Моделот на емоции во devAIce® дозволува секој хардвер или софтвер да се фокусира на изразување на гласните емоции и да ги изведе димензиите на емоциите од гласовната анализа.

Откривање на гласовна активност

Глас или бучава во позадина? devAIce® робусно открива присуство и отсуство на говор во реално време со минимална латентност.



Карактеристики на говорник

Кој зборува? devAIce® може да ни даде увид во атрибути на говорник како возраст или пол.



Откривање на акустична сцена

Каде се наоѓа говорникот? devAIce® може да разликува 3 акустични класи: внатрешен, отворен, транспорт – за да анализира конкретна акустична сцена како: улица, парк, јавна зграда, автобус, лифт, тоалет и други.



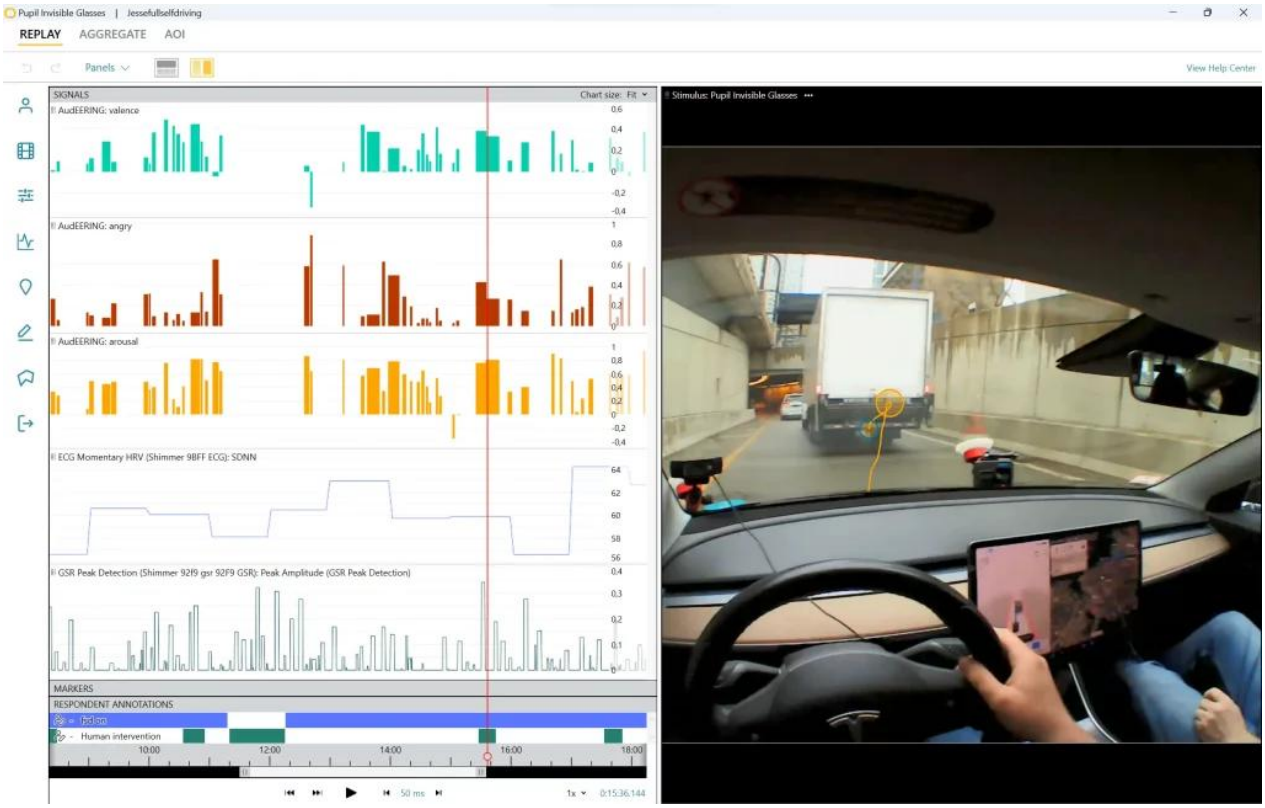
Подобрување на истражувањето со прецизни гласовни увиди

Анализата на говорот (гласот) ни овозможува да извлечеме информации надвор од лингвистиката на гласовите, како што се: **емоции** и **возбуда**, **пол**, **возраст**, **брзина на зборување** и **карактеристики на прозодија**.

Како работи?

Секое видео или снимање на некој што зборува може да се обработи со модулот за гласовна анализа. Ова значи дека не само што можеме да започнеме да го користиме ова за ново истражување, туку можеме да се вратиме и повторно да ги анализираме старите податоци - што ќе ни овозможи да откриеме нови наоди или да ја зголемиме валидноста на нашиот сет на податоци со додавање на уште една димензија на мерење.

Откако аудиото ќе се обработи локално, можеме да избереме кои сигнали да се прикажуваат на временска линија, овозможувајќи брза визуелизација на податоците.



Во партнерство на iMotions (лидер во креирање биосензорска технологија) со audEERING се креира исклучителна интеграција за врвни истражувања. Моделите devAIce® од audEERING доаѓаат со модулот за гласовна анализа и целата обработка работи локално на поединечен сопствен хардвер, обезбедувајќи приватност и безбедност.





Слика. Приказ на податоци за одредени испитаници при анализа на израз на лице

Области на примена

Клиничко истражување

Клиничкото истражување со децении користеше гласовна анализа за да направи разлика помеѓу клиничката популација. На пример, студиите идентификуваа гласовни биомаркери за дизартричен говор, депресија, Алцхајмерова болест и Паркинсонова болест

Социјални и бихејвиорални истражувања

Гласовната анализа вообичаено се применува во општествените и бихејвиоралните науки за подобро да се разбере како нашиот глас ја одразува нашата личност, влијае на луѓето околу нас или како различни фактори влијаат на тоа како ги перципираме другите гласови.

UX и компјутерски науки



Гласовната анализа може значително да го подобри нашето корисничко искуство и разбирање за гласовните асистенти, безбедноста, психологијата и маркетингот. Помагање на дизајнерите и инженерите да создадат подобри искуства.

Комбинирање со Speech-to-Text за анализа на чувства

Со автоматска транскрипција на говорот и анализа на содржината се анализираат чувства врз основа на употребени зборови. iMotions нуди говор во текст преку Assembly AI integration за да се добиваат податоци директно во iMotions за лесна анализа. Ова е моќна комбинација со гласовна анализа и овозможува да се добие целосна слика за емоционалната состојба на лицето што зборува и што кажува.

Биологијата на нашиот глас и што можеме да научиме од него

Нашиот глас е составен од композит од звучни бранови кои се состојат од многу различни фреквенции. Звучните бранови доаѓаат од нашите гласни жици кои произведуваат вибрации преку два мускули наречени гласни набори (vocal folds). Кога зборуваме, наборите се спојуваат додека воздухот од нашите бели дробови се движи низ нив. Воздухот што дува низ наборите ги тера да вибрираат, со што се произведува звук.

Звучните бранови доаѓаат од нашиот гласен тракт (vocal tract) и се произведуваат од нашиот грклан (larynx), каде што мускулите ја контролираат тензијата на два гласни набори (vocal folds), кои често се нарекуваат и гласни жици (vocal cords).

Постојат два стандардни начини за категоризација на звуците на нашите гласови:

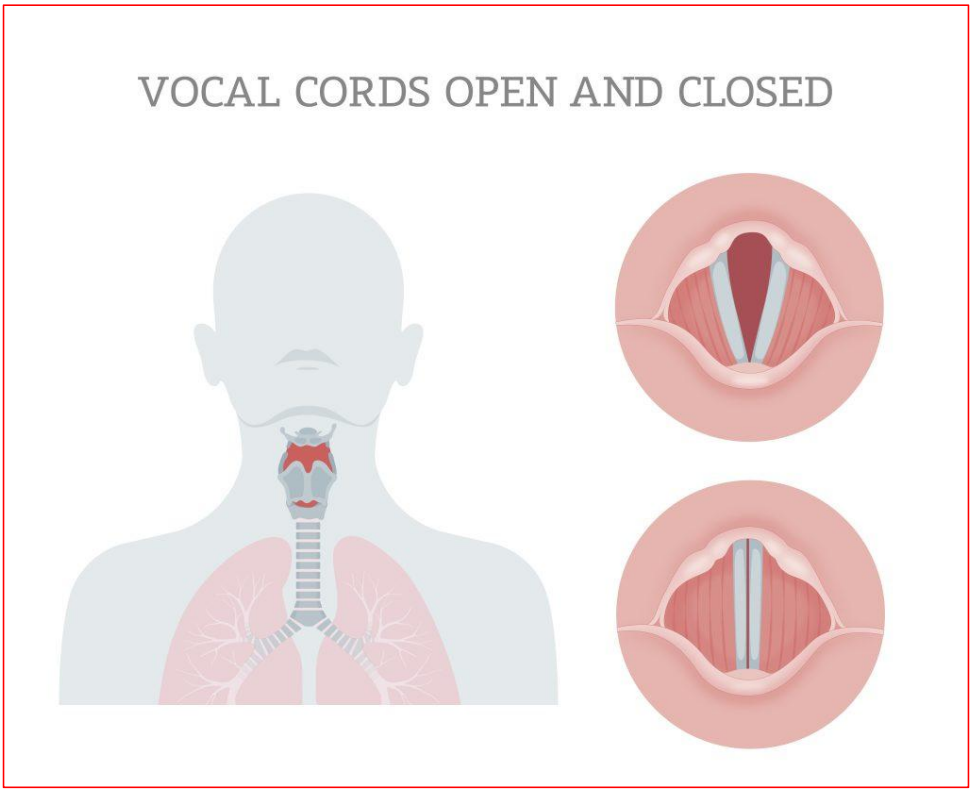
1. **Хармоники** (Harmonics): Хармониките се призвук кои се создаваат од вибрациите на изворот на звук (нашите гласовни набори) и се одредуваат според обликот на гласните набори.
2. **Форманти** (Formants): Се клучни резонантни фреквенции создадени од големината и обликот на гласниот тракт. Тие создаваат типични уникатни звуци на говорни фонеми (speech phonemes) како што се /**aax**/ и /**eex**/. Кога земаме часови по глас или пеење, директно го тренираме нашиот вокален тракт да произведува различни формати кои ќе го модифицираат оригиналниот извор на звук.

Така, хармониката и форматите се одредени од нашата биологија, но и двете можат да се модифицираат преку обука.

Од својствата како што се хармониците и форматите ги заклучуваме вистинските звучни карактеристики што ги анализираме и интерпретираме при гласовната анализа.

Со напредокот на вештачката интелигенција, сега имаме достапни алгоритми кои лесно и прецизно можат да ги идентификуваат и мерат нашите гласови.

Овие алгоритми се способни да го разликуваат звукот на човечкиот глас од, на пример. сообраќајна бучава или инструментална музика.



9. РЕАЛИЗИРАЊЕ НА ПРАКТИЧНО ИСТРАЖУВАЊЕ: користење биосензор за следење на очи

Оваа истражување за мојот магистерски труд го реализирав во лабораторија за биометриски истражувања (невромаркетинг и невропсихологија) на Меѓународен Славјански универзитет во Св. Николе. За таа цел користев биосензорски систем за следење на очи и биосензорски систем за анализа на израз на лице. Истажувањето го реализирав во период од 15 Март до 10 Април 2024 година. Анализите и пресметките ги реализирав на софистициран софтвер од Данската компанија iMotions. Користев екрански базирани биосензори, со стимули кои ги поставив на компјутерски екран.

Цел на истражувањето

За ова невромаркетинг истражување користев стимул (паштета-конзерва со прикажана цена на чинење). Беа понудени неколку различни видови производи кои имаат иста цена на чинење, а еден производ е потенциран (препорачан – подвлечен со жолто) за купување (да се види како испитаникот ќе реагира кога се препорачува еден производ како најдобар за купување). Истражувањето го реализирав со два методи на истражување:

- **Анкетен прашалник**. Традиционален метод на истражување каде испитаникот заокружува (на отпечатен лист хартија) кој производ најмногу му се допаѓа.
- **Биосензорско истражување**, мерење на потсвесни одговори на секој испитаник. Се користи биосензор за следење на очи со кој се добиваат повеќе метрики при опсервација на екрански стимул од испитаници:
    - Фокусирано внимание (ниво на интерес) со креирање на топлински мапи;
    - области на интерес каде на одредени бафер- зони се прикажуваат поголем број на параметри поврзани со преференциите на испитаниците;

- низа од фиксации кои покажуваат како и каде се движел погледот на секој испитаник;
- дилатација на зениците на очите на испитаниците (поотворени зеници се индикатор на зголемено ниво на возбуда при следење екрански стимул);
- фреквенција на трепкање која е индикатор за ниво на концентрација, фокус и поспаност.



Слика. Стимул (слика) за опсервација со екрански биосензор за следење на очи

10. Анализа на резултатите од Биосензско следење на очи: поединечни и збирни анализи

Пред да се започне со биосензорско снимање на секој испитаник, потребно е да се реализира процес на **Калибрација** (за секој испитаник посебно).



Слика 1. Топлинска мапа на испитаник 2

На слика 1 е прикажана топлинска мапа на испитаник 2 каде може да се забележи дека овој испитаник најголем интерес покажал за производот број 2 (неговото фокусирано внимание било насочено кој овој производ), што значи дека овој производ најмногу му се допаднал и би го купил. Овој испитаник и на анкетниот прашалник го заокружил истиот производ што покажува дека не постои дискрепанца (отстапување) во двата методи на истражување.

На слика 2 е прикажана топлинска мапа (spotlight стил) за истиот испитаник каде што нивото на осветленост е индикатот (пропорционален) за нивото на фокусирано внимание (ниво на интерес).

На слика 3 е прикажана топлинска мапа каде површината на сенка е пропорционален индикатор за нивото на интерес на испитаникот.



Слика 2. Топлинска мапа (spotlight), испитаник 2



Слика 3. Топлинска мапа (shadow), испитаник 2

На слика 4 е прикажана топлинска мапа на испитаник 2 каде се прикажани поединечни (грозд) фиксации (сакади) кои покажуваат поединечни дискретни погледи. Колку повеќе фиксации се регистрирани (темносини точки), тоа е индикација на поголемо ниво на интерес и фокус. Овде се забележува дека овој испитаник најголем интерес има за производ 2, а многу помал интерес за останатите производи.



Слика 4. Точки на визуелизација, испитаник 2

На слика 5 е прикажана топлинска мапа на испитаник 5. Може да се забележи дека овој испитаник нема производ кој најмногу му се допаѓа. Неговото фокусирано внимание (индицирано преку топлински мапи) и дистрибуирано на повеќе производи. Биосензорското мерење покажува дека испитаникот најголем интерес покажал за производот 4, додека во анкетниот прашалник го заокружил производот 3. Ова покажува дискрепанца (отстапување) кога ги споредуваме двата методи на истражување. Од научен аспект може да се даде следниот заклучок:

- Биосензорското мерење е поврзано со потсвесни реакции на испитаникот. Тоа е поврзано со негови емоционални преференци и

одлуките ги носи со БРЗО размислување. Ова значи дека ако се бара испитаникот да донесе брз одговор, тогаш тој ќе се одлучи за производ 4.

- Анкетното истражување е поврзано со рационално размислување и донесување одлуки. Овој процес се одвива побавно. Ова значи дека ако на овој испитник му оставиме подолго време да денесе одлука, тој би се одлучил за производ 3.

На сликите 5, 6, 7, 8 и 9 се прикажани топлински мапи за испитаник 5 за различни стилови (осветленост, сенки и грозд од фиксации). Овие мапи потенцираат различни индикатори од кои визуелно може да се види кои се преференциите на испитаникот и на кои сегменти (делови) од секој производ го задржал неговиот поглед (имал зголемено ниво на интерес). т.е може да се забележи кои детали побудиле кај него зголемен фокус на интерес.



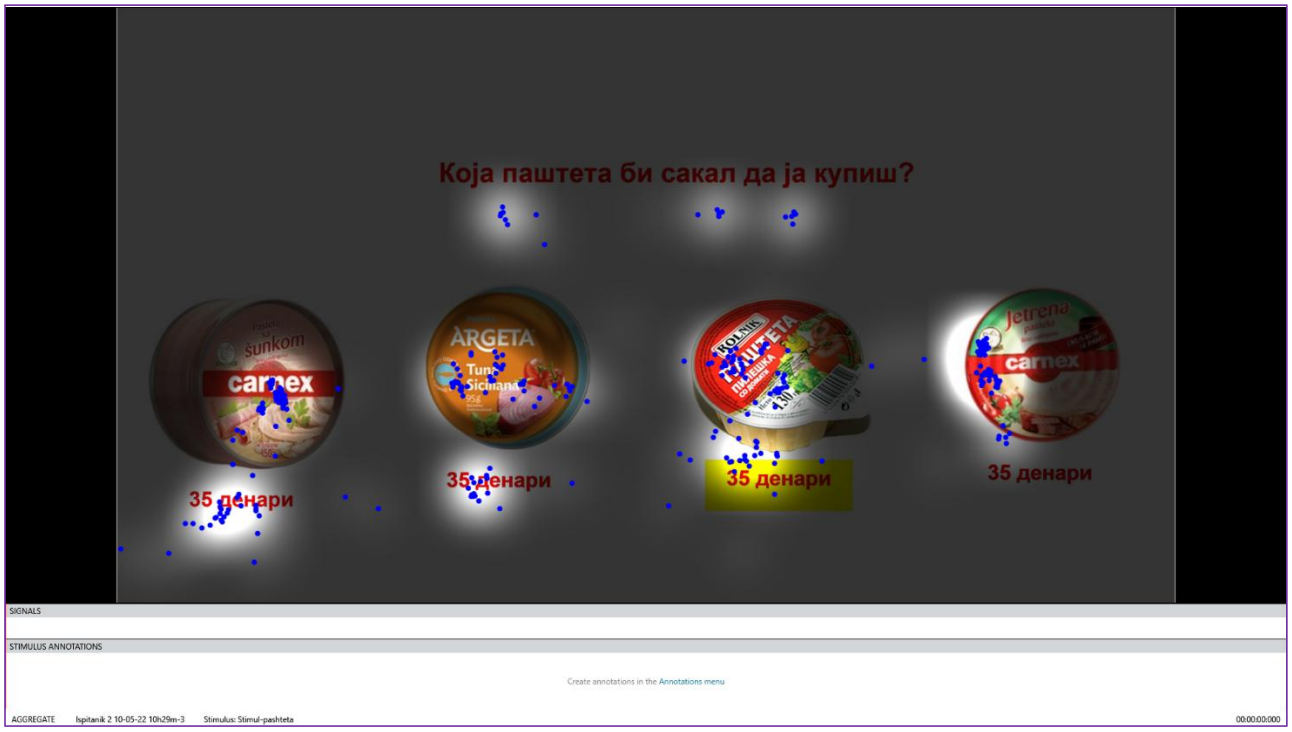
Слика 5. Топлинска мапа, испитаник 5



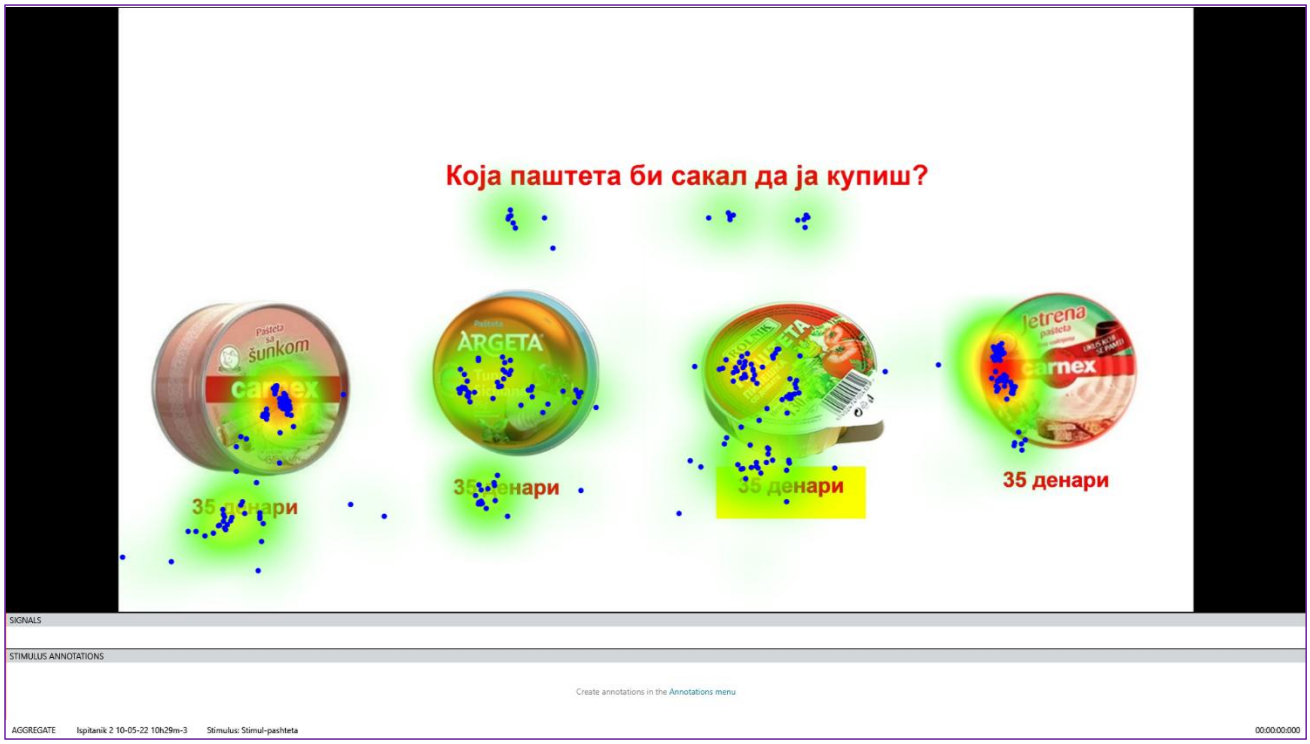
Слика 6. Топлинска мапа (shadow), испитаник 5



Слика 7. Топлинска мапа (spotlight), испитаник 5



Слика 8. Топлинска мапа (spotlight), испитаник 5, со прикажани групирани погледи на очите



Слика 9: Топлинска мапа на испитаник 5, со групиран поглед

а) Топлински мапи (Heat maps) – Збирни анализи



Слика 10. Топлинска мапа, збирна (од сите испитаници)

На слика 10 е прикажана збирна топлинска мапа (од сите испитаници заедно). Може да се забележи дека испитаниците најголем интерес покажале за производ 2, а значајно помал интерес за останатите производи. Топлински збирни мапи во различни стилови се прикажани на слика 11, слика 12 и слика 13. Овие слики се одличен индикатор (ниво на осветленост, површина на сенка, грозд на поединечни фиксации на поглед) за да се донесе заклучок каде испитаниците најмногу го фокусирале своето внимание, кој производ кај нив побудува најголем интерес и кој производ би сакале да го купат. Вакви збирни биосензорски анализи се одличен индикатор за менаџерите во маркетинг за да ја градат својата маркетинг стратегија.

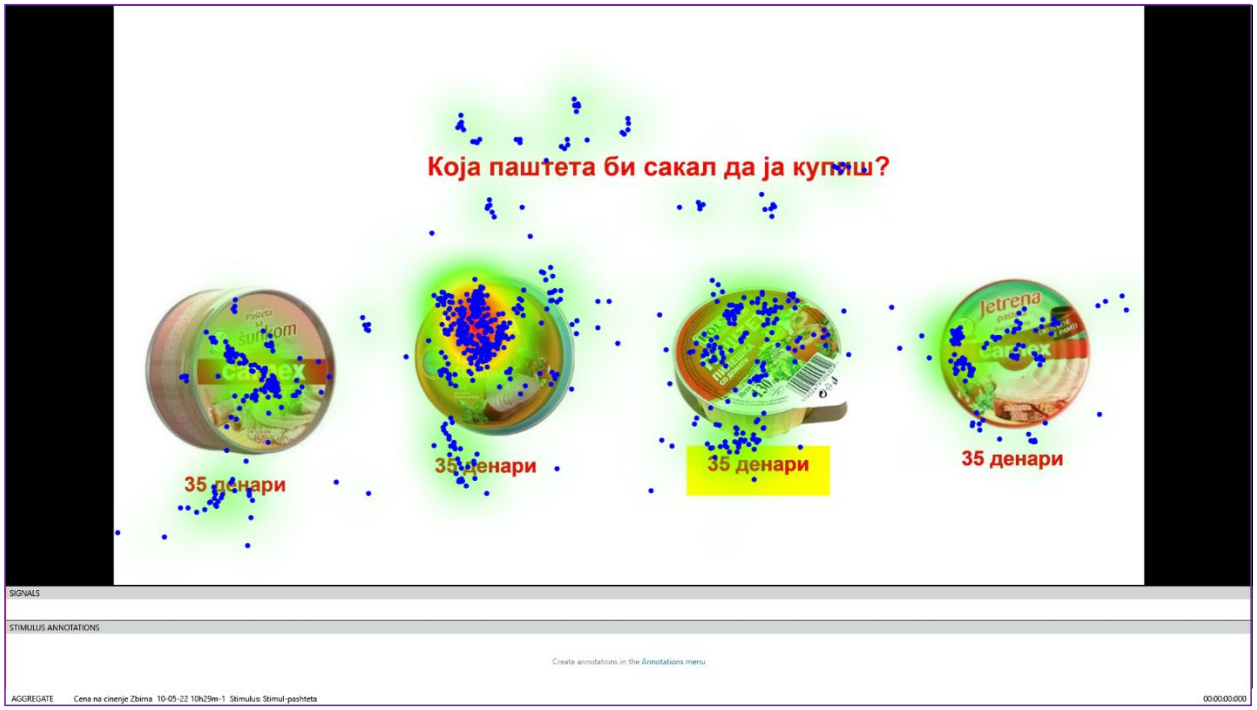
Сумирајќи ги резултатите од анкетното истражување од сите испитаници (прикажано во споредбена табела на крајот од магистерскиот труд), се забележува дека и во оваа збирно истражување испитаниците го избрале производ 2 (со најголем процент на избор).



Слика 11. Топлинска мапа, збирна (од сите испитаници)



Слика 12. Топлинска мапа, збирна (од сите испитаници)



Слика 13. Топлинска мапа, збирна (од сите испитаници), со групирани погледи на очите

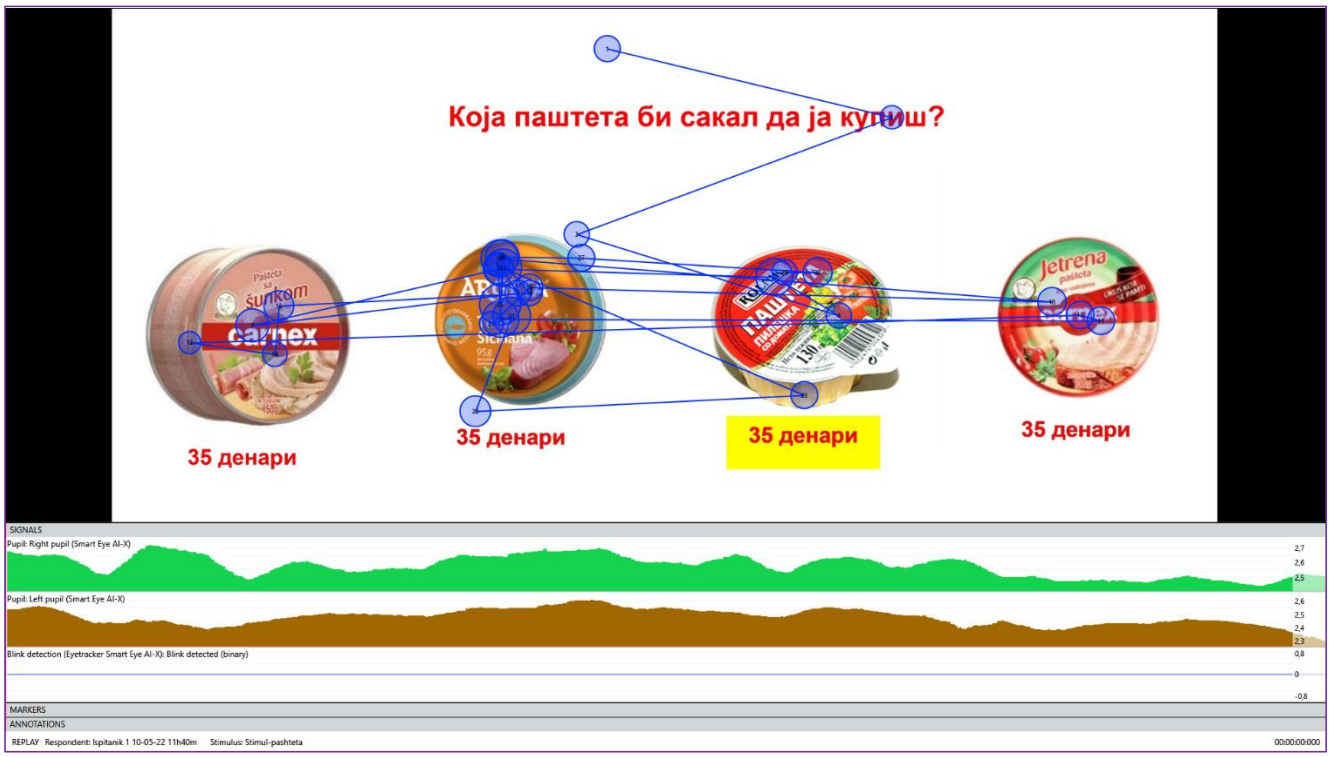


Слика 14. Топлинска мапа, збирна (од сите испитаници) во 4-та секунда од следењето на очите

На слика 14 е прикажана интересна анализа при сумирање на збирни резултати. Е прикажан динамичкирежим: Во точно време (во четвртата секунда на следење на стимулот,од вкупно време на следење во траење од 10 секунди) на следење на стимулот колку испитаници погледот го задржале на одреден производ. Може да се забележи дека во тој момент двајца испитаници го гледале производ 1, а еден испитаник се фокусирал на цената на чинење на производот 3.

b) Низи на фиксации (Fixation sequences)

Низа на фиксации овозможува за секој испитаник да се одреде: Каде гледал и колку долго гледал одреден детал од екранскиот стимул и дава детален увид во редоследот на фокусирано внимание (погледи на очите).

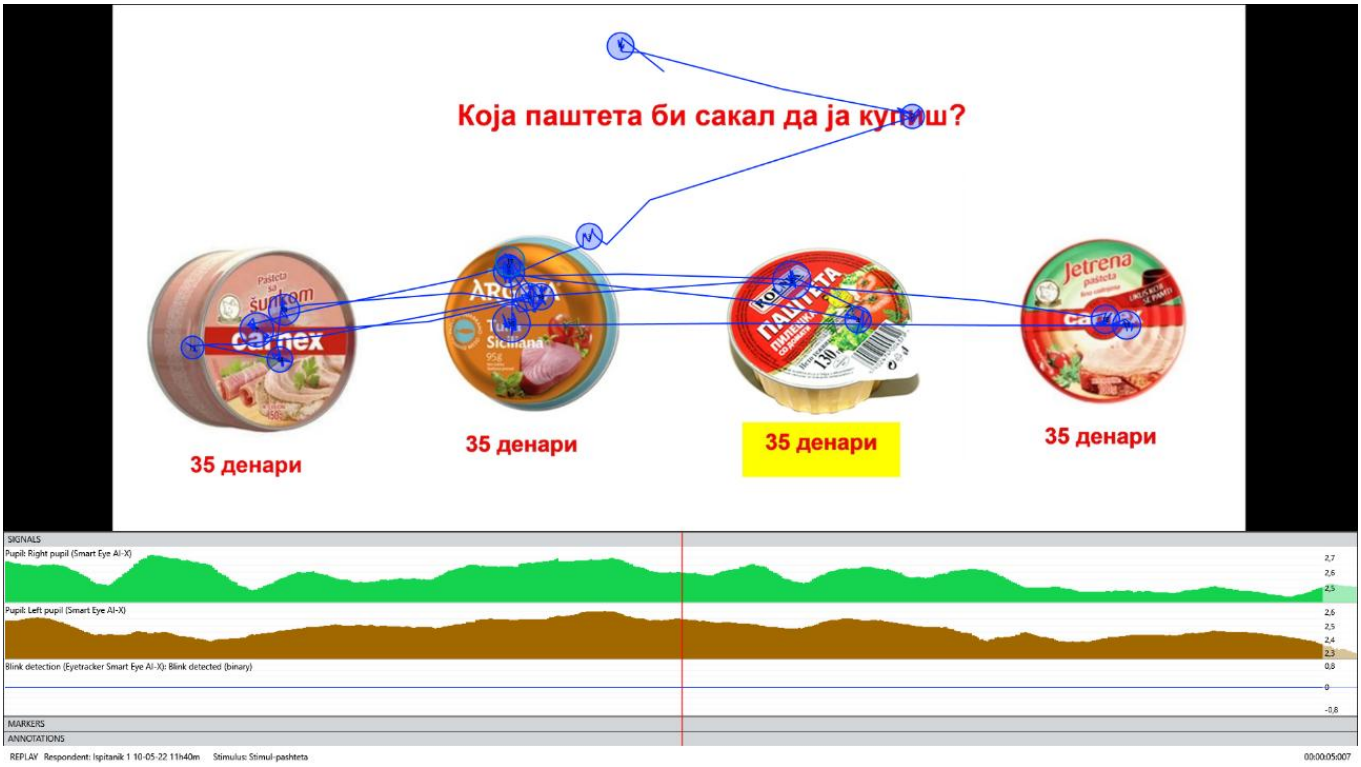


Слика 15: Низа на фиксации на испитаник **2**

Кај испитаник 2, покрај низа на фиксации е прикажана и дилатацијата на зеницата на левото и десното око. Вредностите на зениците континуирано се прикажуваат за време на целиот период на опсервација на стимулот (10 секунди). Пого-

лема дилатација (отворена зеница) покажува зголемено ниво на возбуда на испитаникот.

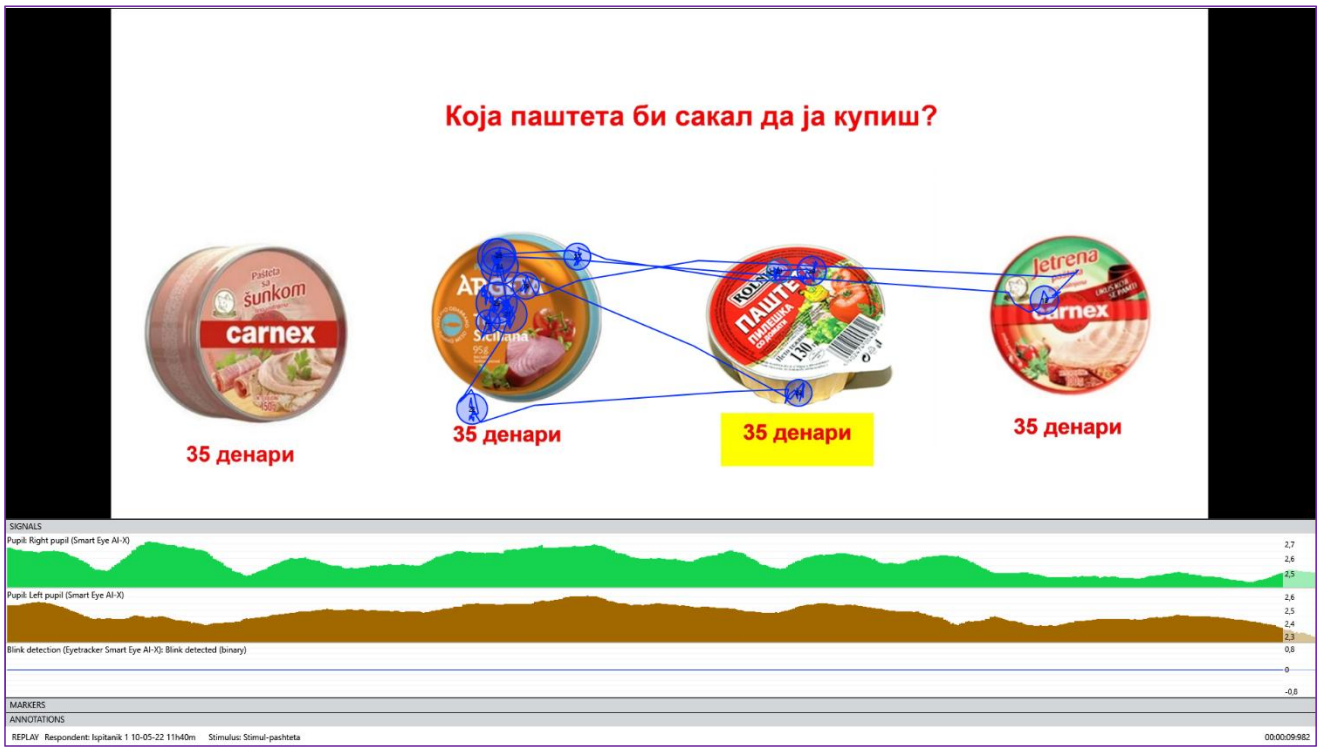
На слика 16 се прикажани низа од фиксации во период од првите 5 секунди од следењето на стимулот на компјутерски екран. Софтверското решение овозможува да се прави динамичка реконструкција на целокупниот период на следење на екранскиот стимул, со забавено следење, што овозможува да се прави детална анализа на погледите на секој испитаник. Во секој момент може да се одреди каде гледал испитаникот, колку време се задржал во гледање на одреден сегмент од стимулот, колкави се дилатациите на двете зеници на очите и фреквенцијата на трепкање (колку трепкања и во кој момент се случиле).



Слика 16: Низа на фиксации на испитаник 2 (од 0 до 5 секунди)

На слика 17 е прикажана низа на фиксации на испитаник во втората половина од следење на екранскиот стимул (од 5 до 10 секунда). Се забележува дека во овој период дилатациите на зениците се значајно намалени (позатворени зеници, нема зголемена возбуда). Ова е поради фактот што испитаникот го нашол сво-

јот омилен производ и се фокусирал на негово следење (што го покажува и распоредот на фиксации кои се сконцентрирани на производ 2).



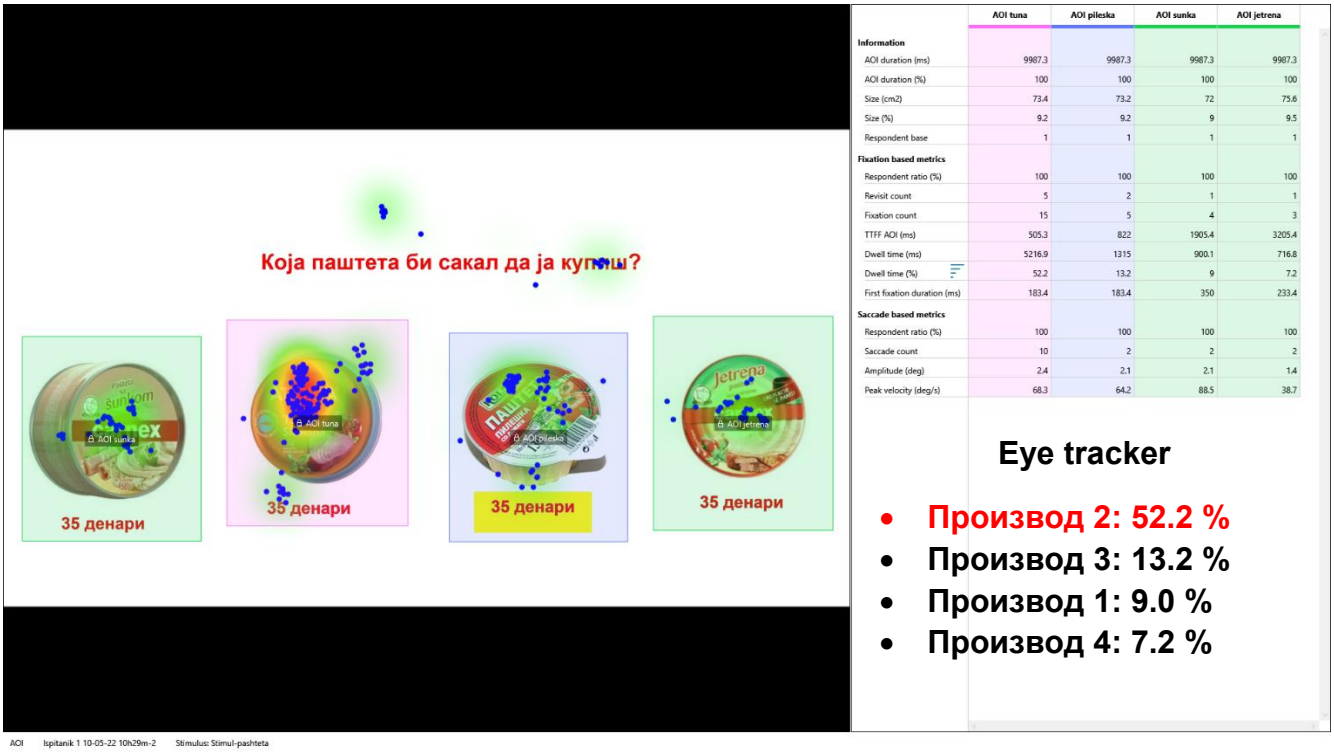
Слика 17: Низа на фиксации на испитаник 2 (од 5 до 10 секунди)

Од оваа визуелна анализа може да се заклучи дека, испитаникот 2 во периодот од петтата до десеттата секунда од своето набљудување, фокусираното внимание го насочува кон вториот и третиот производ (особено кон вториот производ кој побудува најголем интерес кај него).

c) **Области на интерес** (Areas of interest): **Поединечни** анализи

Области на интерес претставуваат кориснички дефинирани **подрегиони** на прикажан екрански стимул. Се прикажуваат поголем број на метрики за одреден подрегион од стимулот (бафер зона), што овозможува да се прават проценки (и споредувања, и статистички и математички анализи) за перформансите на повеќе специфични области кои се произволно избрани (селектирани).

На слика 18 е прикажана област на интерес за испитаник 2 со селектирани четири (4) бафер- зони (околу секој производ).



	AOI tuna	AOI pileska	AOI sunka	AOI jetrena
Information				
AOI duration (ms)	9987.3	9987.3	9987.3	9987.3
AOI duration (%)	100	100	100	100
Size (cm2)	73.4	73.2	72	75.6
Size (%)	9.2	9.2	9	9.5
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	5	2	1	1
Fixation count	15	5	4	3
TTFF AOI (ms)	505.3	822	1905.4	3205.4
Dwell time (ms)	5216.9	1315	900.1	716.8
Dwell time (%)	52.2	13.2	9	7.2
First fixation duration (ms)	183.4	183.4	350	233.4
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	10	2	2	2
Amplitude (deg)	2.4	2.1	2.1	1.4
Peak velocity (deg/s)	68.3	64.2	88.5	38.7

AOI Ispitanik 1 10-05-22 10h29m-2 Stimulus: Stimul-pashteta

Слика 18. Област на интерес, испитаник 2

Во табелата на слика 18 се прикажани повеќе значајни метрики кои покажуваат кој производ најмногу му се допаднал на испитаникот. Една важна метрика е **Dwell time (%)**. Оваа метрика покажува колкав процент од времето на набљудување на стимулот испитаникот потрошил за следење на одреден производ (опфат опфатен со соодветна бафер површина).

Се забележува дека најголем процент има за **производот 2** (**52.2%**). Додека многу помал процент има за останатите производи:

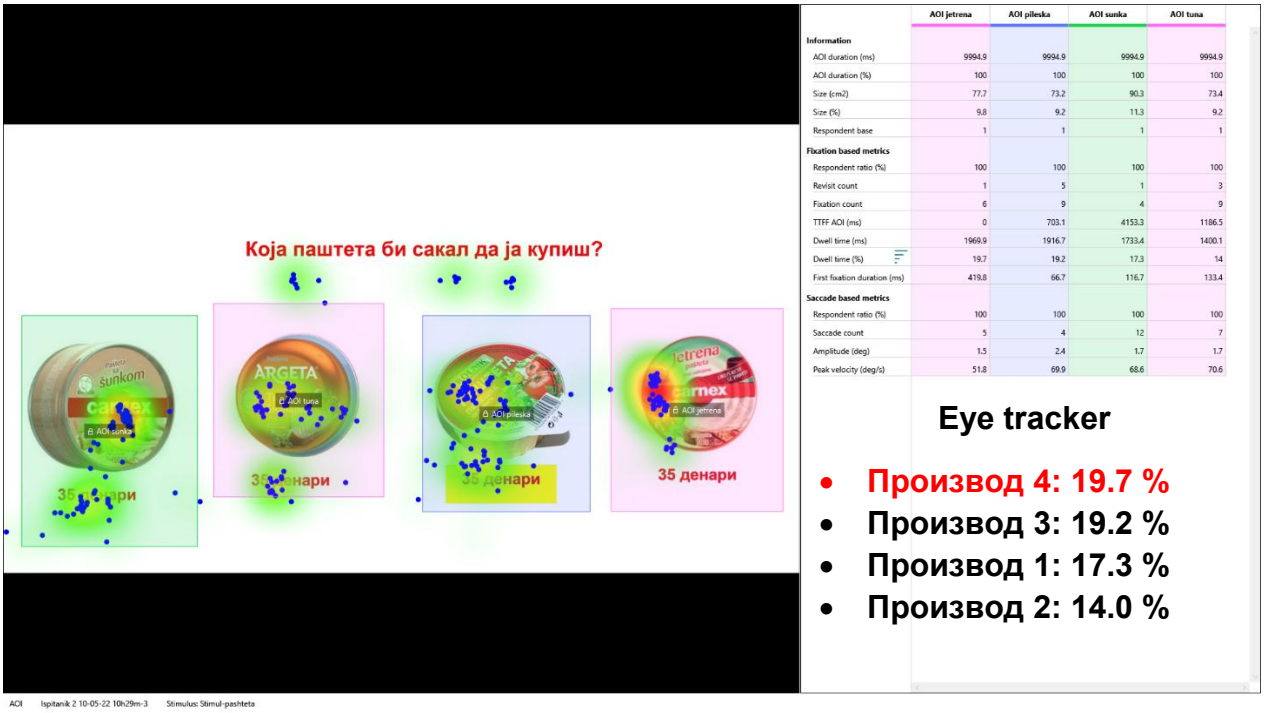
- **Производ 2: 52.2 %;**
- **Производ 3: 13.2 %;**
- **Производ 1: 9.0 %;**
- **Производ 4: 7.2 %;**

Овој испитаник на анкетниот прашалник го заокружил (избрал) производ 2.

На слика 19 е прикажана област на интерес за испитаник 5. Овој испитаник нема одреден производ што најмногу му се допаѓа. Со **Dwell time (%)** се добиени следните метрики во проценти:

- **Производ 4: 19.7 %;**
- **Производ 3: 19.2 %;**
- **Производ 1: 17.3 %;**
- **Производ 2: 14.0 %;-**

Според биосензорско мерење (потсвесно размислување) овој испитаник го избрал производ 4. Во Анкетно истражување го заокружил (избрал) производ 3. Ова е дискрепанца (отстапување), но за овие два производи процентуалните вредности се многу близа.



	AOI jetrena	AOI pileska	AOI sunka	AOI tuna
Information				
AOI duration (ms)	9994.9	9994.9	9994.9	9994.9
AOI duration (%)	100	100	100	100
Size (cm2)	77.7	73.2	90.3	73.4
Size (%)	9.8	9.2	11.3	9.2
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	1	5	1	3
Fixation count	6	9	4	9
TTFF AOI (ms)	0	703.1	4153.3	1186.5
Dwell time (ms)	1969.9	1916.7	1733.4	1400.1
Dwell time (%)	19.7	19.2	17.3	14
First fixation duration (ms)	419.8	66.7	116.7	133.4
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	5	4	12	7
Amplitude (deg)	1.5	2.4	1.7	1.7
Peak velocity (deg/s)	51.8	69.9	68.6	70.6

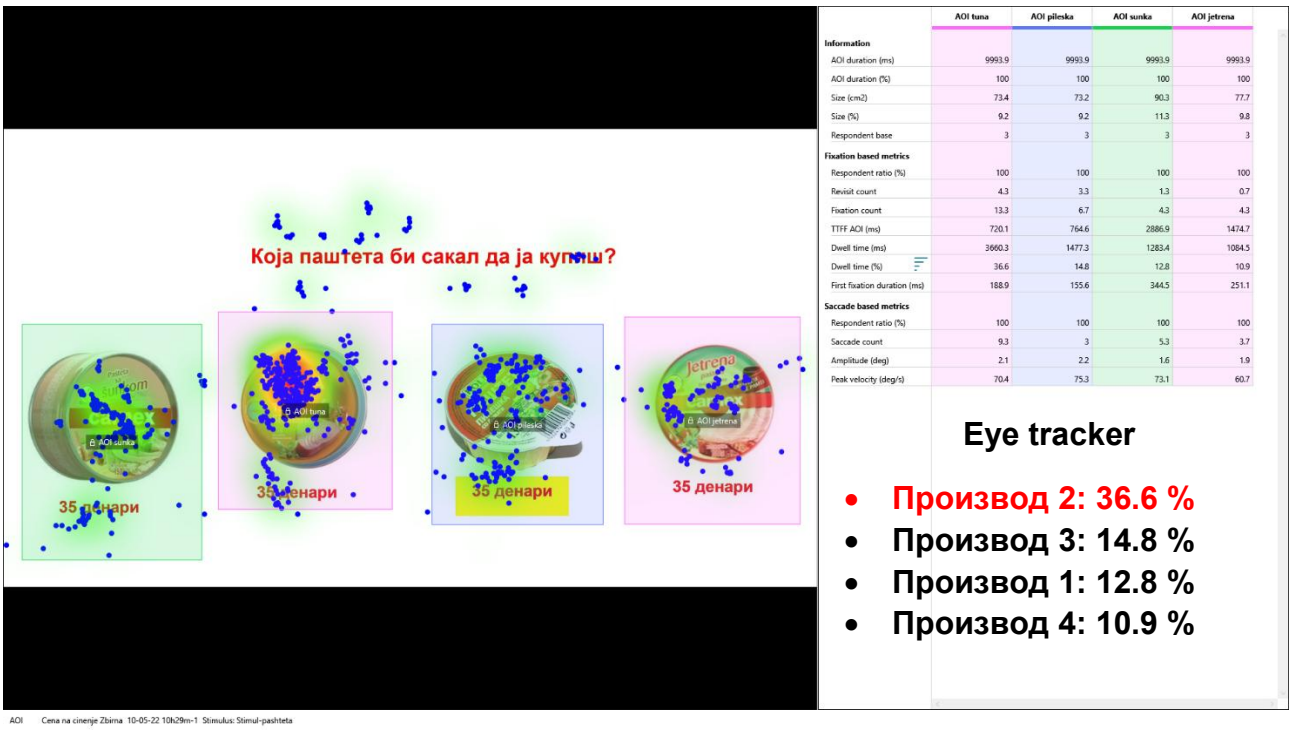
Слика 19. Област на интерес, испитаник 5

е) Области на интерес (Areas of interest): **збирни** анализи

На слика 20 е прикажана збирна анализа (од сите испитаници заедно) со прикажување област од интерес. Креирани се четири (4) бафер- зони околу сите четири производи. Со **Dwell time (%)** се добиени следните метрики во проценти:

- **Производ 2: 36.6 %;**
- **Производ 3: 14.8 %;**
- **Производ 1: 12.8 %;**
- **Производ 4: 10.9 %;**

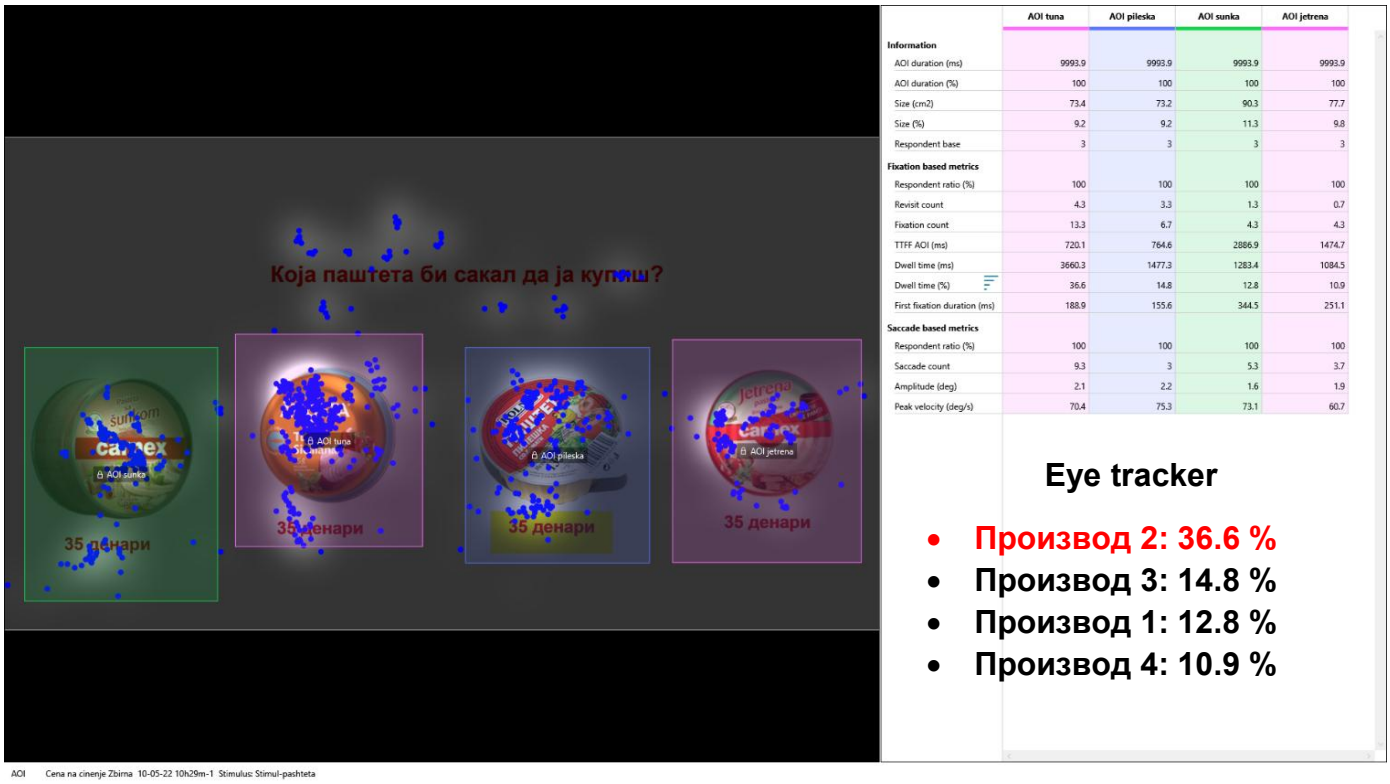
Се забележува дека испитаниците најголем интерес покажале за производ 2. Значајно помал интерес покажале за другите производи (конзерви-паштета). Исто така на сликата се прикажани и грозд фиксации собрани од сите испитаници (сини точки). И овие фиксации недвосмислено покажуваат дека испитаниците најмногу го задржале погледот на производ 2.



Слика 20. Област на интерес, збирна анализа со топлинска мапа (од сите испитаници)

На слика 21 е прикажана збирна област на интерес (од сите испитаници заедно) за четири (4) области на интерес. На сликата е прикажана топлинска мапа каде индикатор на фокусираност е осветленоста. Исто така се прикажани и збирни фиксации (грозд) од фокусиран поглед на очи од испитаниците. Прикажани се по-голем број на метрики во табелата за секоја бафер- зона (т.е за секој поединечен

производ). Може да се забележи дека **производ 2** има најголема осветленост и најмногу сини точки на фиксација.



	AOI tuna	AOI pileska	AOI sunka	AOI jetrena
Information				
AOI duration (ms)	9993.9	9993.9	9993.9	9993.9
AOI duration (%)	100	100	100	100
Size (cm2)	73.4	73.2	90.3	77.7
Size (%)	9.2	9.2	11.3	9.8
Respondent base	3	3	3	3
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	4.3	3.3	1.3	0.7
Fixation count	13.3	6.7	4.3	4.3
TTFF AOI (ms)	720.1	764.6	2886.9	1474.7
Dwell time (ms)	3660.3	1477.3	1283.4	1084.5
Dwell time (%)	36.6	14.8	12.8	10.9
First fixation duration (ms)	188.9	155.6	344.5	251.1
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	9.3	3	5.3	3.7
Amplitude (deg)	2.1	2.2	1.6	1.9
Peak velocity (deg/s)	70.4	75.3	73.1	60.7

Слика 21. Област на интерес, збирна анализа со точки на поглед (од сите испитаници)

11. Споредување на поединечни вредности за сите испитаници

Стимул			
Испитаник	**Анкетен прашалник**	**Биометриско мерење со Eye Tracker**	**Дискрепанца (2-3)**
Испитаник **1**	**Паштета 1**	**Паштета 1**	**нема**
Испитаник **2**	**Паштета 2**	**Паштета 2**	**нема**
Испитаник **3**	**Паштета 2**	**Паштета 2**	**нема**
Испитаник **4**	**Паштета 1**	**Паштета 2**	**Има: 1/2**
Испитаник **5**	**Паштета 3**	**Паштета 4**	**Има: 3/4**
Испитаник **6**	**Паштета 4**	**Паштета 1**	**Има: 4/1**
Испитаник **7**	**Паштета 2**	**Паштета 2**	**нема**
Испитаник **8**	**Паштета 2**	**Паштета 2**	**нема**
Испитаник **9**	**Паштета 1**	**Паштета 3**	**Има: 1/3**
Испитаник **10**	**Паштета 3**	**Паштета 3**	**нема**
Испитаник **11**	**Паштета 2**	**Паштета 2**	**нема**
Испитаник **12**	**Паштета 3**	**Паштета 3**	**нема**
Збирни резултати	**Паштета 2: 5/12 (41.66.0%)** **Паштета 3: 3/12 (25.0%)** **Паштета 1: 3/12 (25.0%)** **Паштета 4: 1/12 (8.33%)**	**Паштета 1: 2/12 (16.66%)** **Паштета 2: 6/12 (50.0%)** **Паштета 3: 3/12 (25.0%)** **Паштета 4: 1/12 (8.33%)**	

Кај одредени испитаници (4, 5, 6 и 9) има дискрепанца (отстапување) кај двата вида на истражување (анкетно и биосензорско). Ова отстапување се појавува поради постоење можна пристрасност кај анкетното истражување – **пристрасност** која се јавува поради различни причини кај испитаниците. Кај биосензорското ис-

тражување очите се прозорец на мозокот – го рефлектираат она што навистина го мисли мозокот во даден момент.

Второ истражување

За потребите на мојот магистерски труд, во лабораторијата на Меѓународен Славјански универзитет го реализирав и следното биосензорско истражување со користење биосензор за следење на очи. Користев екрански стимул: Избор на свежо месо за конзумирање. Ги добив следните резултати:

Го користев следниот стимул:



На слика 1 е прикажана топлинска мапа на испитаник 1 каде се забележува дека изборот е недвосмислен на вториот понуден производ.



Слика 1. Топлинска мапа со фиксации, испитаник 1



Слика 2. Топлинска мапа со фиксации, испитаник 4

На слика 2 е прикажана топлинска мапа на испитаник 4 каде изборот е на производ 4.

На слика 3 е прикажана топлинска мапа на испитаник 5. Овој испитаник го избрал производ 2 (тоа го покажува неговото фокусирано внимание и бројот и густина на фиксации).

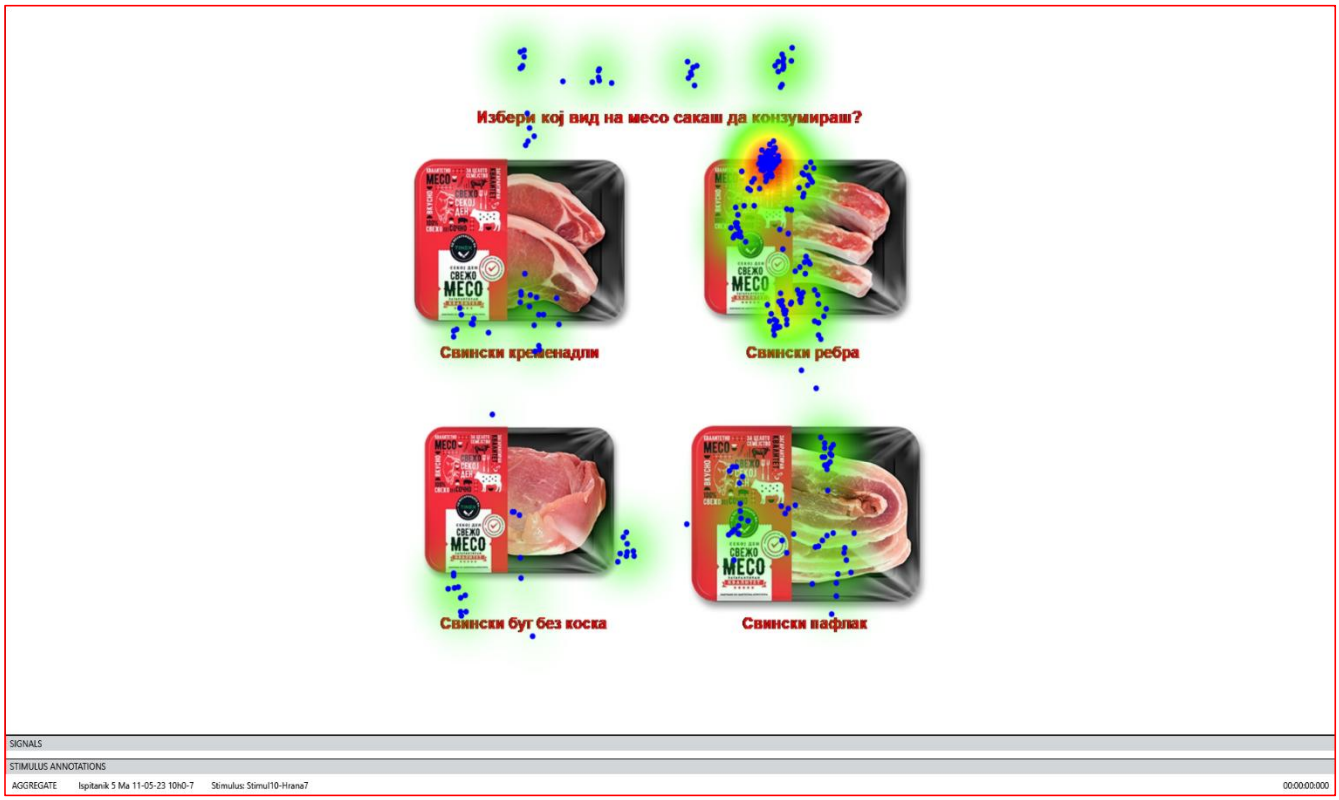
На слика 4 е прикажана топлинска мапа на испитаник 6. Овој испитаник има два производи кои му се допаѓаат. Тоа се **производ 2** и **производ 4**.

На слика 5 е прикажана топлинска мапа на испитаник 6 каде индикација на фокусирано внимание е осветленоста.

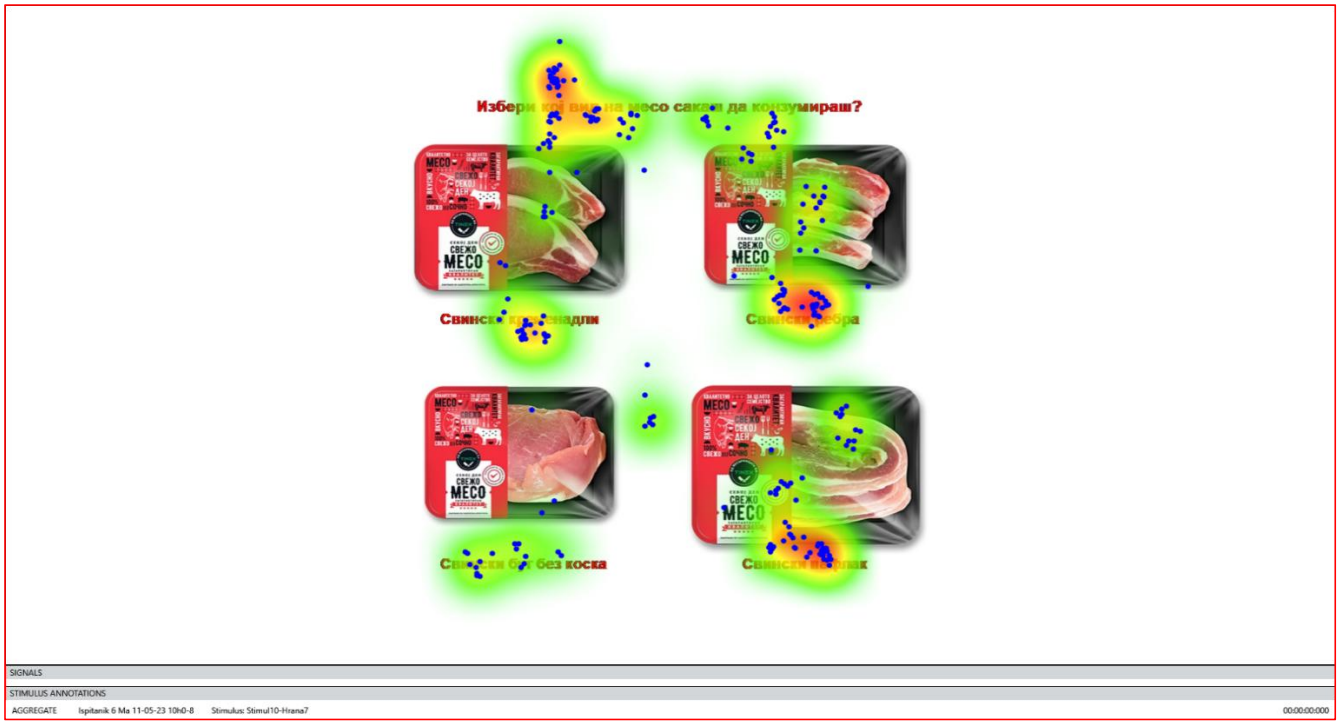
На слика 6 е прикажана топлинска мапа на испитаник 7 кој има свој омилен производ – тоа е производ 4.

На слика 7 е прикажана топлинска мапа на испитаник 8 кој има свој омилен производ – тоа е производот 1.

На слика 8 е прикажана топлинска мапа (осветленост) на испитаник 10, кој за свој омилен го избрал производ 1.



Слика 3. Топлинска мапа со фиксации, испитаник 5



Слика 4. Топлинска мапа со фиксации, испитаник 6



Слика 5. Spotlight мапа, испитаник 6



Слика 6. Топлинска мапа со фиксации, испитаник 7



Слика 7. Топлинска мапа со фиксации, испитаник 8

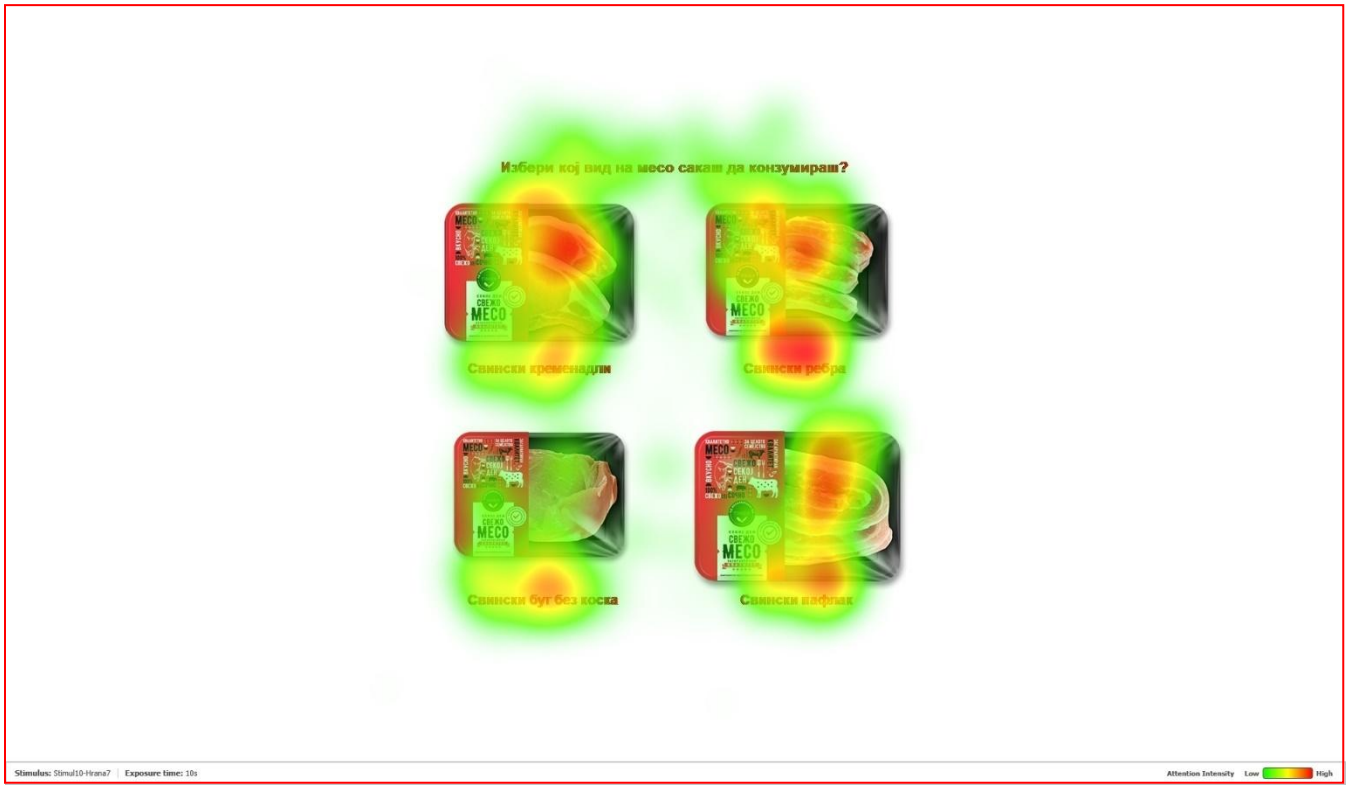


Слика 8. Spotlight мапа, испитаник 10

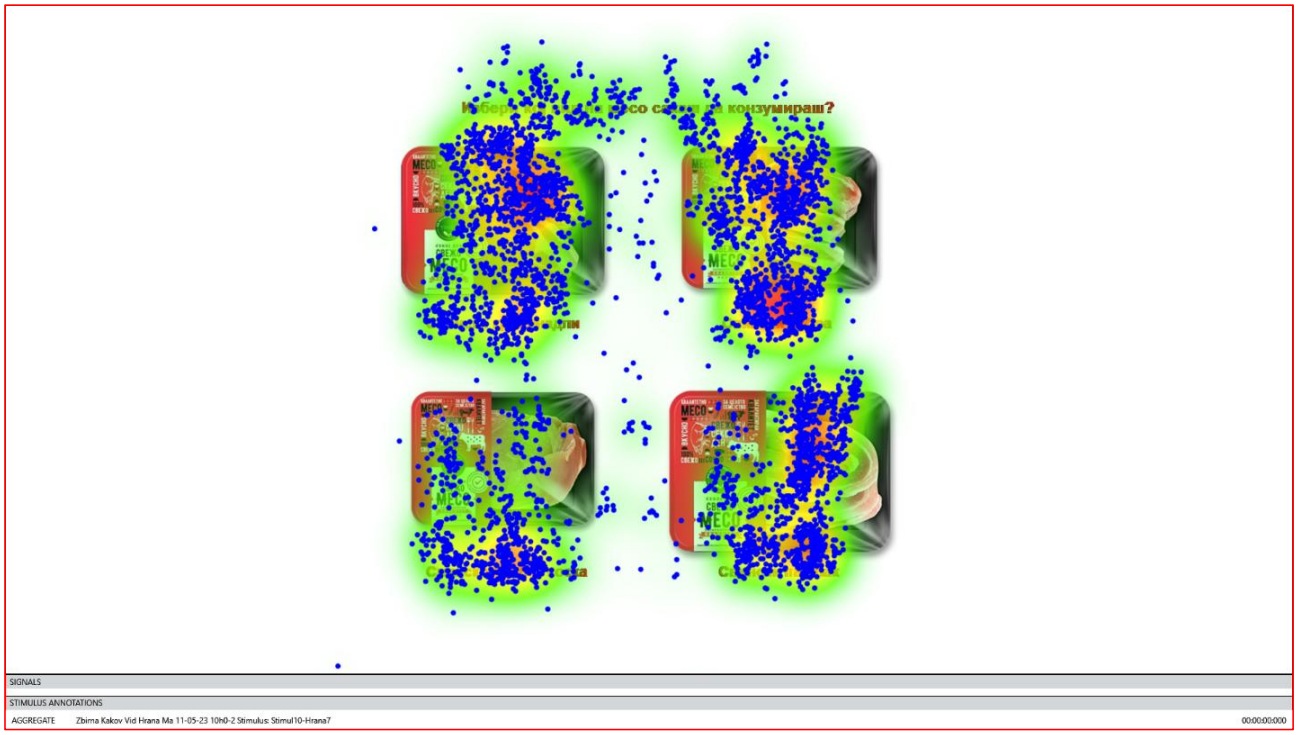
Топлинска мапа (Heat maps): Збирни анализи

На слика 9 е дадена збирна топлинска мапа од сите испитаници заедно. Може да се забележи на кои производи испитаниците најмногу го фокусирале своето внимание.

На слика 10 е прикажана иста збирна топлинска мапа, но во овој визуелен приказ се инкорпорирани и збирни фиксации (грозд фиксации) од сите испитаници за временски период на следење на екрански стимул во траење од 10 секунди.

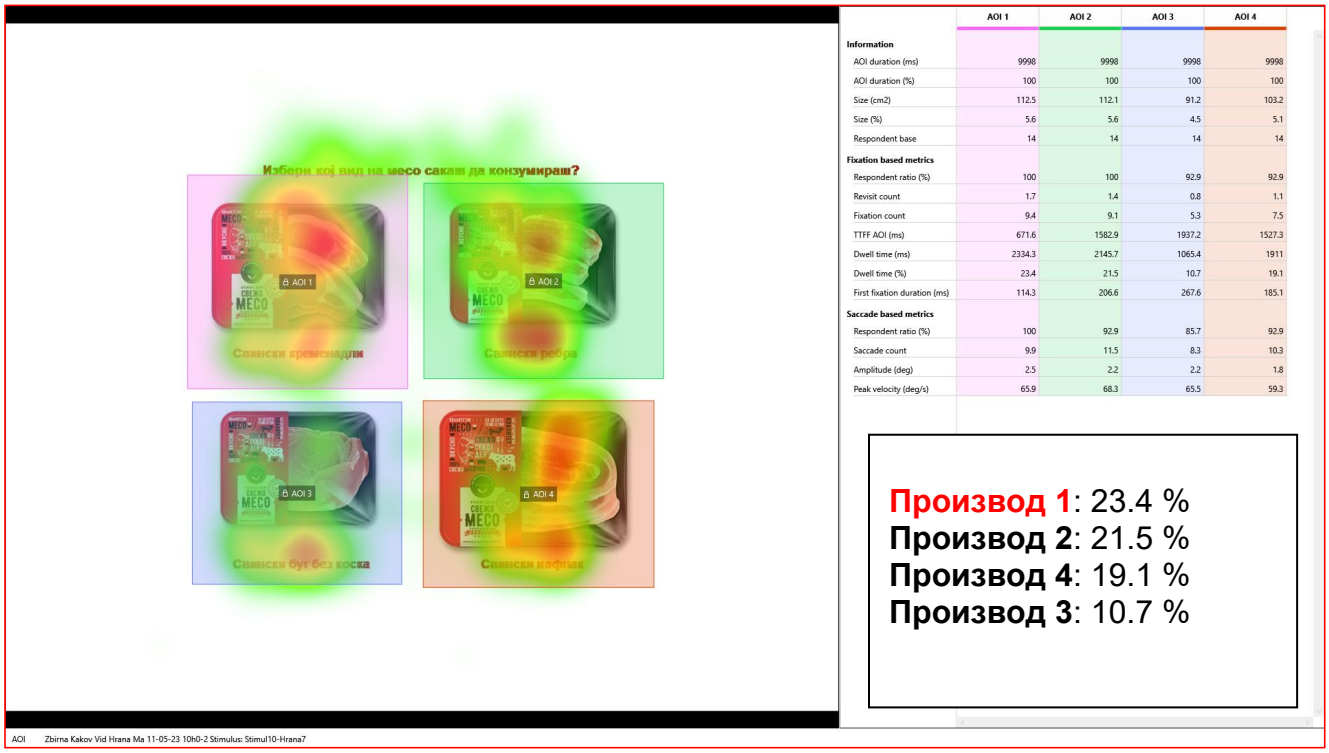


Слика 9. Топлинска мапа, збирна



Слика 10. Топлинска мапа со фиксации, збирна

Области на интерес (Areas of interest): **збирни анализи**



	AOI 1	AOI 2	AOI 3	AOI 4
Information				
AOI duration (ms)	9998	9998	9998	9998
AOI duration (%)	100	100	100	100
Size (cm2)	112.5	112.1	91.2	103.2
Size (%)	5.6	5.6	4.5	5.1
Respondent base	14	14	14	14
Fixation based metrics				
Respondent ratio (%)	100	100	92.9	92.9
Revisit count	1.7	1.4	0.8	1.1
Fixation count	9.4	9.1	5.3	7.5
TTFF AOI (ms)	671.6	1582.9	1937.2	1527.3
Dwell time (ms)	2334.3	2145.7	1065.4	1911
Dwell time (%)	23.4	21.5	10.7	19.1
First fixation duration (ms)	114.3	206.6	267.6	185.1
Saccade based metrics				
Respondent ratio (%)	100	92.9	85.7	92.9
Saccade count	9.9	11.5	8.3	10.3
Amplitude (deg)	2.5	2.2	2.2	1.8
Peak velocity (deg/s)	65.9	68.3	65.5	59.3

Слика 11. Област на интерес, збирна

На слика 11 е прикажана збирна област на интерес за сите испитаници заедно. Со **Dwell time (%)** се добиени следните метрики во проценти:

- **Производ 1**: 23.4 %;
- **Производ 2**: 21.5 %;
- **Производ 4**: 19.1 %;
- **Производ 3**: 10.7 %;

Се забележува дека испитаниците најголем интерес покажале за производ 1. Нешто помал интерес покажале за производ 2 и производ 4, а најмал интерес покажале за производ 3..

Споредување на поединечни вредност за сите испитаници

Стимул			
Испитаник	**Анкетен прашалник**	**Биометриско мерење со Eye Tracker**	**Дискрепанца (2-3)**
Испитаник **1**	**Храна 1**	**Храна 1**	**нема**
Испитаник **2**	**Храна 2**	**Храна 2**	**нема**
Испитаник **3**	**Храна 3**	**Храна 3**	**нема**
Испитаник **4**	**Храна 2**	**Храна 2**	**нема**
Испитаник **5**	**Храна 4**	**Храна 4**	**нема**
Испитаник **6**	**Храна 4**	**Храна 4**	**нема**
Испитаник **7**	**Храна 3**	**Храна 3**	**нема**
Испитаник **8**	**Храна 3**	**Храна 3**	**нема**
Испитаник **9**	**Храна 2**	**Храна 2**	**нема**
Испитаник **10**	**Храна 2**	**Храна 1**	**има: 2 и 1**
Збирни резултати	**Храна 1: 1/10 (10.0%)** **Храна 2: 4/10 (40.0%)** **Храна 3: 3/10 (30.0%)** **Храна 4: 2/10 (20.0%)**	**Храна 1: 2/10 (20.0%)** **Храна 2: 3/10 (30.0%)** **Храна 3: 3/10 (30.0%)** **Храна 4: 2/10 (20.0%)**	

Од табелата се забележува дека дискрепанца (отстапување) во двата методи на истражување има само кај испитаник 10.

ЗАКЛУЧНИ СОГЛЕДУВАЊА

Голем дел од човековото однесување е водено од несвесни процеси. Деценискиистражувања покажаа дека голем дел од нашата ментална обработка се случува на потсвесно ниво. Ова може да се подобри со користење апликативен интеграбилен софтвер кој им дава на истражувачите пристап до емоционални, когнитивни и бихејвиорални (однесување) податоци. Ваков софтвер интегрира и синхронизира повеќе БИОМЕТРИСКИ СЕНЗОРИ кои овозможуваат различен увид на човечки психо-физиолошки појави.

Податоци за човековото однесување (Behavioral data) денес се нова нормалност. Тоа е напредна иновативна технологија која ги забрзува научните истражувања за препознавање емоции (emotion recognition technology). Софтверот за биосензорски истражувања е извонредно корисна алатка за кодирање и опишување на однесувањето на испитаници и нивните емоционални реакции на точен и квантитативен начин. Оваа напредна технологија обезбедува мултисензорна интеграција (multisensor integration) во една единствена компјутеризирана платформа, така што лесно може да се синхронизираат повеќе текови на податоци (multiple data streams): Следење на очи (Eye Tracking), електроенцефалографија (EEG), електродермична активност EDA (Electrodermal Activity) и анализа на изрази на лице (Facial Expression Analysis) преку екранска поставка на стимули во речиси каква било каква форма, од слики, видеа и веб-страници, до поставки во реалниот свет што нé опкружува.

Постојните традиционалните техники овозможуваат квалитативно и квантитативно мерење на когнитивни и емоционални искуства само како вербално изразени на свесно ниво за време на интервју или анкетирање. Наместо тоа, со користење на техники за снимање на мозокот, изрази на лице, следење на очи, елекродермална активност, отчукување на срце, процес на дишење, човечки говор и други, може да се детектираат несвесни состојби поврзани со процеси кои имаат клучно влијание врз човековото однесување, интегрирајќи ги и оние сознанија што може се добиени со вербални или писмени декларации. Затоа, биосензорските истражувања се важно и револуционерно поле на научните истражувања, и се дефини-

рани како нејзина “нова димензија“. Денес биосензорските истражувања добиваат силно внимание во корпоративниот свет.

На пример, испитаниците не се секогаш искрени и врз нивните перцепции може да влијае како е поставено прашањето. Она што излегува од нашите усти не е секогаш совршена изведба на она што се случува во нашите мозоци.

Постојат различни причини зошто пристапот на прашање-одговор има свои недостатоци. Таквите методи се склони кон пристрасност, со испитаници кои не сакаат или не можат да го кажат она што го мисли нивниот ум. Дури и кога испитаниците се обидуваат да бидат целосно отворени и искрени, во игра се потсвесни пристрасности (subconscious biases) и разни други фактори кои можат да го променат нивниот одговор. Со обезбедување метод за мерење на потсвесните мисли и чувства на испитаник, овие фактори можат да се квантифицираат, споредуваат и искористат за да се направат заклучоци што можат да ги поддржат или побијат оригиналните хипотези. Секој биосензор обезбедува различен и колаборативен елемент во тоа разбирање.

Традиционалните методи и биосензорните методи можат да се надополнуваат одговарајќи на различни типови прашања во врска со психолошкото и физиолошкото патување на една личност во однос на одреден стимул или искуство.

Користена литература

1. Banai, I. P., Banai, B., & Bovan, K. (2017). Vocal characteristics of presidential candidates can predict the outcome of actual elections. *Evolution and Human Behavior*, 38(3), 309-314.
2. Bocklet, T., Steidl, S., Nöth, E., & Skodda, S. (2013, September). Automatic evaluation of parkinson's speech-acoustic, prosodic and voice related cues. In *Interspeech* (pp. 1149-1153).
3. Burkhardt, F., Wagner, J., Wierstorf, H., Eyben, F., & Schuller, B. (2023). Speech-based Age and Gender Prediction with Transformers. *arXiv preprint arXiv:2306.16962*.
4. Eyben, F. (2015). *Real-time speech and music classification by large audio feature space extraction*. Springer.
5. Eyben, F., Wöllmer, M., & Schuller, B. (2010, October). Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia* (pp. 1459-1462).
6. Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., André, E., Busso, C., ... & Truong,
7. K. P. (2015). The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE transactions on affective computing*, 7(2), 190-202.
8. Fan, M., Lin, J., Chung, C., & Truong, K. N. (2019). Concurrent think-aloud verbalizations and usability problems. *ACM Transactions on Computer-Human Interaction (TOCHI), 26(5),* 1-35.
9. Fan, M., Zhao, Q., & Tibdewal, V. (2021, May). Older adults' think-aloud verbalizations and speech features for identifying user experience problems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).
10. Gregory Jr, S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of personality and social psychology*, 70(6), 1231.
11. Han, E. Y., Yun, J. Y., Chong, H. J., & Choi, K. G. (2018). Individual therapeutic singing program for vocal quality and depression in Parkinson's disease. *Journal of movement disorders*, 11(3), 121.
12. Laukka, P., Linnman, C., Åhs, F., Pissiota, A., Frans, Ö., Faria, V., ... & Furmark, T. (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior*, 32, 195-214.

13. Meilán, J. J., Martínez-Sánchez, F., Carro, J., Sánchez, J. A., & Pérez, E. (2012). Acoustic markers associated with impairment in language processing in Alzheimer's disease. *The Spanish journal of psychology*, 15(2), 487-494.

14. Neubauer, C., Chollet, M., Mozgai, S., Dennison, M., Khooshabeh, P., & Scherer, S. (2017, November). The relationship between task-induced stress, vocal changes, and physiological state during a dyadic team task. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction* (pp. 426-432).

15. Niebuhr, O., Voße, J., & Brem, A. (2016). What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. *Computers in Human Behavior*, 64, 366-382.

16. Narendra, N. P., & Alku, P. (2018, September). Dysarthric Speech Classification Using Glottal Features Computed from Non-words, Words and Sentences. In *Interspeech* (pp. 3403-3407).

17. Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, 26(5), 388-397.

18. Schucker, B., & Jacobs Jr, D. R. (1977). Assessment of behavioral risk for coronary disease by voice characteristics. *Psychosomatic Medicine*, 39(4), 219-228.

19. Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J.K., Baird, A., Elkins, A., Zhang, Y., Coutinho, E., Evanini, K. (2016) The INTERSPEECH 2016 Computational Paralinguistics Challenge: Deception, Sincerity & Native Language. Proc. *Interspeech* 2016, 2001-2005, doi: 10.21437/Interspeech.2016-129.

20. Scherer, S., Hammal, Z., Yang, Y., Morency, L. P., & Cohn, J. F. (2014). Dyadic Behavior Analysis in Depression Severity Assessment Interviews. *Proceedings of the ... ACM International Conference on Multimodal Interaction. ICMI (Conference), 2014*, 112–119. https://doi.org/10.1145/2663204.2663238.

21. Stel, M., Dijk, E. van, Smith, P. K., Dijk, W. W. van, & Djalal, F. M. (2012). Lowering the Pitch of Your Voice Makes You Feel More Powerful and Think More Abstractly. *Social Psychological and Personality Science*, 3(4), 497–502. https://doi.org/10.1177/1948550611427610.

22. Taguchi, T., Tachikawa, H., Nemoto, K., Suzuki, M., Nagano, T., Tachibana, R., Nishimura, M., & Arai, T. (2018). Major depressive disorder discrimination using vocal acoustic features. *Journal of affective disorders*, 225, 214–220. https://doi.org/10.1016/j.jad.2017.08.038.

23. Wagner, J., Triantafyllopoulos, A., Wierstorf, H., Schmitt, M., Burkhardt, F., Eyben, F., & Schuller, B. W. (2023). Dawn of the transformer era in speech emotion recognition: closing the valence gap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*

24. Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems 6(2), 156-166 (2005). doi: 10.1109/TITS.2005.848368.

25. Zontone, P., Affanni, A., Bernardini, R., Piras, A., Rinaldo, R., Formaggia, F., Minen, D., Minen, M., Savorgnan, C.: Car Driver's Sympathetic Reaction Detection Through Electrodermal Activity and Electrocardiogram Measurements". IEEE Transactions on Biomedical engineering, vol. 49(12), 3413-3424 (2020). doi: 10.1109/TBME.2020.2987168.
26. Lemonnier, S., Bremond, R., Baccino, T.: Gaze behavior when approaching an intersection: Dwell time distribution and comparison with a quantitative prediction. Transportation Research Part F: Traffic Psychology and Behaviour 35, 60-74 (2015). doi: 10.1016/j.trf.2015.10.015.
27. Jabon, M., Bailenson, J., Pontikakis, E., Takayama, L., Nass, C.: Facial expression analysis for predicting unsafe driving behavior. IEEE Pervasive Comput 10(4), 84-95 (2011). doi: 10.1109/MPRV.2010.46.
28. Mehler B., Reimer, B., Coughlin, J.F., Dusek, J.A.: Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Transp Res Rec 2138(1), 6-12 (2009). doi: 10.3141/2138-02.
29. Collet, C., Clarion, A., Morel, M., Chapon, A., Petit, C.: Physiological and behavioral changes associated to the management of secondary tasks while driving. App Ergon 40(6), 1041-1046 (2009).
30. Wierwille, W.W., Eggemeier, F.T.: Recommendations for mental workload measurement in a test and evaluation environment. Hum Factors 35(2), 263-281 (1992). doi: 10.1177/001872089303500205.
31. Evans, A., Boonstra, F. M. C., Perera, T., Vogel, A. P., Kolbe, S. C., Butzkueven, H., … van der Walt, A. (2018). What speech can tell us: A systematic review of dysarthria characteristics in Multiple Sclerosis. Autoimmunity Reviews, 17(12), 1202–1209. https://doi.org/10.1016/j.autrev.2018.06.010.
32. Skodda, S., Grönheit, W., Lukas, C., Bellenberg, B., Von Hein, S. M., Hoffmann, R., & Saft, C. (2016). Two different phenomena in basic motor speech performance in premanifest Huntington disease. Neurology, 86(14), 1329–1335. https://doi.org/10.1212/WNL.0000000000002550.
33. Martínez-Sánchez, F., Meilán, J. J. G., Carro, J., & Ivanova, O. (2018). A Prototype for the Voice Analysis Diagnosis of Alzheimer's Disease. Journal of Alzheimer's Disease, 64(2), 473–481. https://doi.org/10.3233/JAD-180037.
34. Schuller, B., Friedmann, F., & Eyben, F. (2013). Automatic recognition of physiological parameters in the human voice: Heart rate and skin conductance. In ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings (pp. 7219–7223). https://doi.org/10.1109/ICASSP.2013.6639064.
35. Koyuncu, E., Çam, P., Altınok, N., Çallı, D. E., Duman, T. Y., & Özgirgin, N. (2016). Speech and language therapy for aphasia following subacute stroke.

Neural Regeneration Research, 11(10), 1591–1594. https://doi.org/10.4103/1673-5374.193237.
36. Moore, R. K.: Appropriate voices for artefacts: some key insights. In 1st Int. Workshop on Vocal Interactivity in-andbetween Humans, Animals and Robots (VIHAR-2017). 2017.
37. Mori, M.: The Uncanny Valley. Energy, 7(4), pp. 33–35, 1970.
38. Wilson, S. and Moore, R. K.: Robot, alien and cartoon voices: implications for speech-enabled systems In 1st Int.
Workshop on Vocal Interactivity in-and-between Hu- mans, Animals and Robots (VIHAR-2017). 2017.
39. Nass, C. and Brave, S.: Wired for Speech. MIT Press 2005.
40. C. Saint-Georges, M. Chetouani, R. Cassel, F. Apicella, A. Mahdhaoui, F. Muratori, M.-C. Laznik, and D. Cohen, “Motherese in interaction: at the cross-road of emotion and cognition? (a systematic review),” PLoS ONE, vol. 8, no. 10, pp. 1–17, 2013. 593.
41. M. Spinelli, M. Fasolo, and J. Mesman, “Does prosody make the difference? a meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes,” Developmental Review, vol. 44, pp. 1–18, 2017.
42. C. Kitamura and D. Burnham, “Pitch and communicative intent in mother’s speech: Adjustments for age and sex in the first year.” Infancy, vol. 4, pp. 85–110, 2003.
43. D. Burnham, U. Vollmer-Conna, and C. Kitamura, “What’s new, Pussicat? on talking to babies and animals,” Science, vol. 296, p. 1435, 2002.
44. L. J. Trainor, C. M. Austin, and R. N. Desjardins, “Is infantdirected speech prosody a result of the vocal expression of emotion?” Psychological Science, vol. 11, no. 3, 2000.
45. C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, The measurement of meaning. University of Illinois press, 1957, no. 47.
46. K. Scherer, R. Banse, H. Wallbott, and T. Goldbeck, “Vocal cues in emotion encoding and decoding,” Motivation and emotion , vol. 15, no. 2, pp. 123–148, 1991.