

Álmoság figyelő rendszer

Drowsiness detection system

VERES Tivadar¹, Dr. SZÁNTÓ Zoltán²

¹Sapientia EMTE, Számítógépes irányítási rendszerek, veres.tivadar@student.ms.sapientia.ro

²Sapientia EMTE, Egyetemi adjunktus, zoltan.szanto@ms.sapientia.ro

Abstract

Car accidents caused by fatigue have a major impact on our daily lives. The aim of this paper is to describe a solution to help prevent such accidents. The paper presents a non-invasive external monitoring solution that uses the front-facing camera of the user's mobile phone to take pictures while driving. The images are uploaded to a server to monitor drowsiness and the results are sent back to the phone.

Keywords: microsleap, fatigue, facial landmarks, mobile phone, image processing

Kivonat

A fáradtság okozta autóbalesetek jelentős hatással vannak mindennapjainkra. A dolgozat célja egy megoldás leírása, amely segíti az ilyen balesetek megelőzését. A dolgozatban egy nem-invazív, kívülről megfigyelő megoldás kerül bemutatásra, amely a felhasználó mobiltelefonjának előlapi kameráját használja képek készítéséhez vezetés közben. A képeket felküldve egy szerveren végezzük az álmoság figyelését, majd az eredményt visszaküldjük a telefonra.

Kulcsszavak: mikroalvás, fáradtság, arc tájékozódási pont, mobiltelefon, képfeldolgozás

1. BEVEZETÉS

Az álmosággal kapcsolatos autóbalesetek továbbra is jelentős hatással vannak az út biztonságra¹. Ezeket a baleseteket el lehetne kerülni, ha időben figyelmeztetnénk a vezetőket, amint elkezdenek álmosak lenni. A probléma megoldására számos rendszer létezik, amelyek különböző szemszögekből közelítik meg a megoldást. Egyes autógyártók vezetési stílus érzékelés alapján próbálják figyelmeztetni az álmos vezetőket. Mások az autókhoz készítették kiegészítő berendezéseket, amelyek szenzorok és kamerák segítségével ismerik fel az álmos vezető viselkedését. A dolgozat egy egyszerűbb, kézenfekvőbb megoldást mutat be, amely egy telefonos applikáció keretein belül bárki használhatja a jármű lecserélése és módosítása nélkül.

2. BIBLIOGRÁFIAI TANULMÁNY

Az éberség figyelésére számos megoldás létezik. Az irodalom szerint az álmoság érzékelő rendszereket három kategóriába lehet sorolni azoknak az intézkedések alapján, amelyeket az álmosági jelek észlelésére használnak: biológiai, jármű alapú és kép-alapú rendszerek.

Az első kategóriában a biológiai alapú mérések a test fiziológiai jeleinek figyelésén alapulnak, ideértve az Elektroenkefalográfia (EEG), Elektrokardiográfia (ECG), Elektromiográfia (EMG), Elektrookulográfia (EOG) jeleket és vérnyomást.

¹ www.bankrate.com/insurance/car/drowsy-driving-statistics/



1. ábra. MediaPipe arcjellemző modellje

Ebben a típusú rendszerben az álmoság azonosítása a jel eltérése alapján történik az alap állapot jellemzőitől és az új jel analízise révén, hogy az álmoságra utal-e vagy sem.

A második kategóriában a jármű alapú mérések a jármű mozgási mintázatainak változásaira támaszkodnak különböző érzékelők által, amelyek különböző jármű- és út jellemzők mérésére vannak felszerelve. A jármű alapú rendszerek a jármű változásainak vagy abnormális viselkedésnek elemzésével következtetik le az álmoság szintjét, ideértve például a kormánykerék szögét, sebességét vagy az eltérést a sávból.

A harmadik kategória a kép-alapú méréseket foglalja magában, amelyek főként a vezető arcán és fején megjelenő álmosági jeleken alapulnak. Ezek a rendszerek az álmoságot a vezető fejmozgásából és arckifejezéseiből mint: szemek, szemöldökök, száj mozgásából határozzák meg [1].

A képekből kinyert arcon található álmosági jelek az arcjellemző-arc tájékozódási pont (Facial Landmark) azonosító rendszerek segítségével kerülnek feldolgozásra [2]. Ezen rendszereket feloszthatjuk három csoportba: Holisztikus módszerek (Holistic methods); Korlátozott Helyi Modell módszerek (Constrained Local Model methods); illetve Regressziós alapú módszerek (Regression-based methods).

2.1. Mediapipe

A MediaPipe egy sokoldalú, platformfüggetlen könyvtár, amely valós idejű médiaközvetítési és gépi tanulási megoldások fejlesztésére szolgál. Számos előre betanított modellt tartalmaz, amelyek azonnal használhatók különböző feladatokra, ezen modellek közül a tájékozódási pont felismerést használtuk a dolgozatban².

Az arc tájékozódási pont felismerő modell három modelltől áll össze, az arc detektáló, a tájékozódási pont detektálás az arckifejezési tulajdonság kinyerő modelltől. Ezen modellek közül az utóbbi nem került felhasználásra. Az arc felismerésére a BlazeFace short-range modellt használják. A dokumentációban nem részletezik a tájékozódási pont felismerő modell működését. A könyvtár sikeres felismerés után 478 darab 3-dimenziós tájékozódási pontot térít vissza, a 1. ábrán szemléltetve, ezen pontokat használjuk fel a további számításokban

3. ELMÉLETI ÁTTEKINTÉS

3.1. Álmoság meghatározása

Az álmosági állapot egy olyan fiziológiai és pszichológiai állapot, amelyet általában aluszékonyság, fáradtság vagy alváshiány vált ki. Ebben az állapotban az egyén hajlamos lehet a koncentrációs nehézségekre, a lassú gondolkodásra és a reakcióidő lassulására. Az álmoság gyakran vezető oka az alváshiány, de lehet az eredménye olyan tényezőknek is, mint a stressz, az egészségi állapot, az alvási rendellenességek vagy bizonyos

² MediaPipe Face Landmarker: https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker

gyógyszerek mellékhatásai. Az álmoság komoly veszélyt jelenthet például vezetés közben vagy gépek kezelésekor, mivel csökkentheti a figyelmet és a reakcióképességet.

Általában az álmosági állapot megszűnik alvással vagy pihenéssel, amely lehetővé teszi az agynak és a testnek a regenerációt és az energia újratöltését.

3.2. Álmoság figyelésére alkalmas tulajdonságok

Tulajdonságok, amelyeket az arc tájékozódási pontok figyeléséből tudunk meghatározni.

3.2.1. Szem méretaránya

Képek alapján az álmoság azonosításának egyik legegyszerűbb megközelítése a szemek állapotának a figyelése, pontosabban az Eye Aspect Ratio (EAR) számolása.

$$EAR = \frac{|p_2 - p_6| + |p_3 - p_5|}{|2 \cdot |p_1 - p_4|} \quad (1)$$

Ezzel a módszerrel mérhető a szem méretaránya, hogy mennyire van nyitva vagy csukva, függetlenül a kép méretétől és a szemek helyzetétől a képen. Önmagában ezt az egy tulajdonságot figyelve, meghatározható a vizsgált személy állapota, ha több ideig ez az érték kicsi, egy bizonyos határérték alatti, akkor következtethető, hogy a vizsgált személy szemei csukva vannak.

3.2.2. Száj méretaránya

A száj méretaránya (Mouth Aspect Ratio - MAR) hasonlóan számolható a szem méretarányá-hoz és fontos szerepet játszik az álmoság meghatározásának folyamatában.

$$MAR = \frac{|62-68| + |63-67| + |64-68|}{3 \cdot |49-55|} \quad (2)$$

3.2.3. Arc pozíciója

Az arc pozíciójából meghatározható a fej mozgása, segítségével bólintást észlelhetünk amely gyakoriság függ-vényében álmosághoz vezethet, amikor a vizsgált egyén leejti a fejét. Ezt a tulajdonságot a legegyszerűbben az arc középpontjából tudjuk kiszámolni, ami megfelel az orr középpontjának. Ezt a pontot és értékeit figyelembe véve meghatározható az orr mozgása. Az archoz viszonyítva pedig meghatározható az arc pozíciója.

3.3. Zavarótényezők

3.3.1. Kép minősége

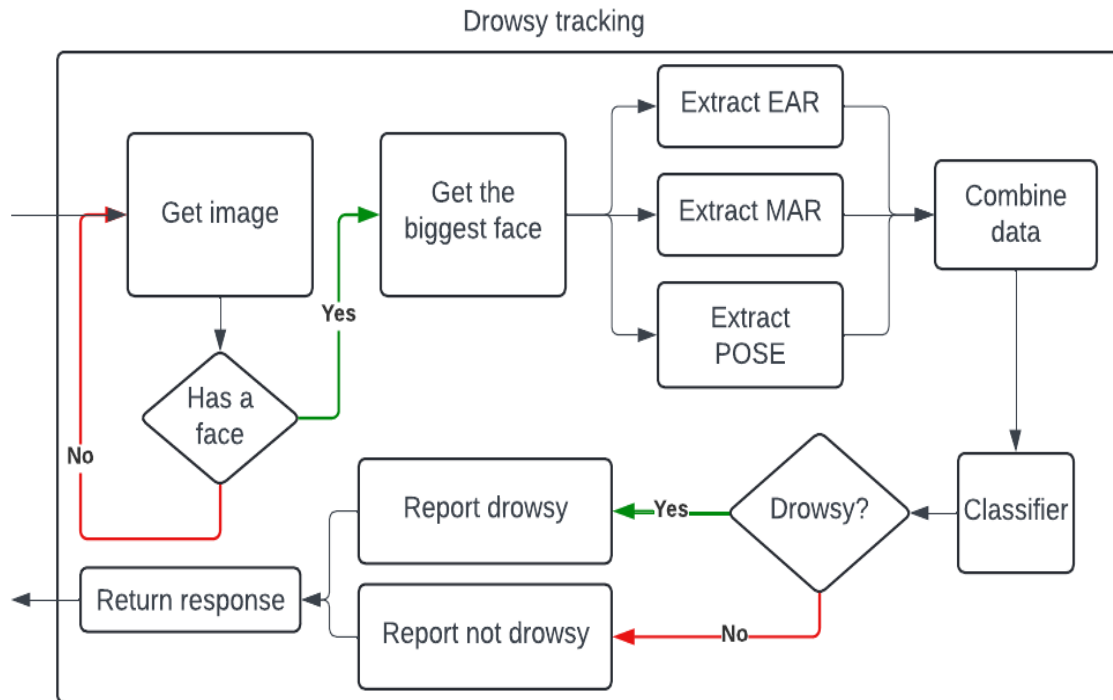
A kép minősége nagy mértékben befolyásolja a felismerés minőségét. Alacsony felbontású képeknél megtör-ténhet hogy az arc tájékozódási pontok felrajzolása nem jár sikerrel, vagy zajos lesz. Magas felbontású képek feldolgozása pedig megterhelheti a rendszert, amely nem elvárt viselkedést eredményezhet. Ezért fontos hogy a tanítás és a működés alatt konzisztens képminőséggel dolgozzunk.

3.3.2. Kamera pozíciója

Fontos, hogy a kamera amely vételezi a figyelni kívánt személyt kis szögben térjen el az arcától, hogy a felismert arc tájékozódási pontok minél pontosabban legyenek felrajzolva az arcára. Nem megfelelő felrajzolás esetén zaj kerül be a rendszerbe és a megfigyelés veszít a pontosságából.

3.3.3. Eltakarás és fényviszonyok

A megfigyelt személy arcán megtörténhet, hogy eltakaró tényezők vannak, mint például napszemüveg, maszk, vagy egyéb tartozékok. Az ilyen esetekben a modern arc tájékozódási pont felismerő algoritmusok már elég jó pontossággal dolgoznak. Viszont megtörténhet, hogy egyes tulajdonságok még így sem felismerhetők, ezért fontos, hogy többet is figyeljünk egyszerre. A fényviszony változás esetében is teljesen a tájékozódási pont felrajzó algoritmusra hagyatkozunk.



2. ábra. Álmoság figyelés lépései

3.3.4. Használt adathalmaz

A használt adathalmaz a National Tsing Hua University driver drowsiness detection dataset (NTHUDDD)³,

amely videófelvételtől kivágott képeket tartalmaz. Az adathalmaz két fő részre van osztva, álmos és nem álmos mintákra. A minták több személyről tartalmaznak képeket. A képek különböző szituációkat fednek le, mint például szemüveges, nem szemüveges, lassú pislogás, ásítás, fejbiccentés.

4. GYAKORLATI MEGVALÓSI TÁS

Egy olyan rendszert fejlesztettünk ki, amely több tulajdonságot vizsgál a vezető arcán, az EAR, MAR, a fej pozícióját és ezen tulajdonságok alapján határozza meg a vezető állapotát egy képsorozatból. Az architektúra a [1] dolgozat alapján készült és az álmoság figyelés 2. ábrán látható lépéseket követi.

A kinyert tulajdonságokat összegyűjtjük egy FIFO adatstruktúrába, és csak akkor vonunk le következtetést, mikor már van 15 feldolgozott képünk. Ez abban segít, hogy nem csak egy kép, hanem egy képsorozat alapján döntünk, ami lehetővé teszi a vizsgált személy mozgását, gesztusát is figyelembe venni. A kinyert tulajdonságok az a két szem EAR értékének az átlaga, a száj MAR értéke, az orr pozíciójának az x koordinátája normalizálva a szélesség és az y koordinátája normalizálva a magasság által. Ez egy 4 elemű sorozat, amely belekerül egy 15 elemű adatstruktúrába, így alkotva egy 4x15 méretű adatstruktúrát.

Mivel képsorozatot használunk, amely felfogható mint egy rövid videó, az idő is belekerül az egyenletbe, így komplexebb viselkedést is meg tudunk határozni. A szem lecsukott állapotát, ha az idő függvényében is figyeljük, a hosszabb ideig lecsukott szem álmosásra utalhat. A száj nyitott állapotát, ha az idő függvényében is figyeljük, a hosszabb ideig nyitott száj ásításra utalhat.

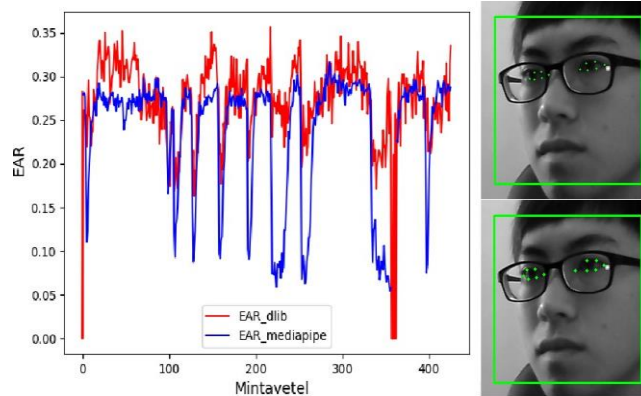
³Adathalmaz: <https://www.kaggle.com/datasets/banudeep/nthuddd2>

4.1. Tulajdonságok kinyerése

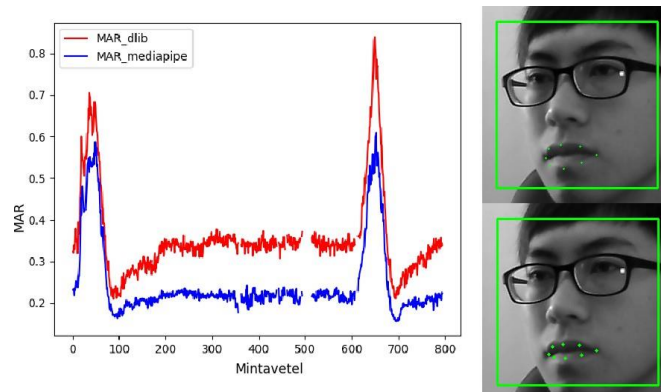
A tulajdonságok kinyerése egy fontos folyamat a rendszer működésében, és nagy mértékben befolyásolja a felismerés pontosságát. A tulajdonságok kinyerése a tanítás és a kész rendszer működése során megegyezik. Referenciaként a Mediapipe arc tájékozódási pont felrajzoló algoritmusát összehasonlítottuk a Dlib algoritmusával.

4.1.1. EAR tulajdonság kinyerése

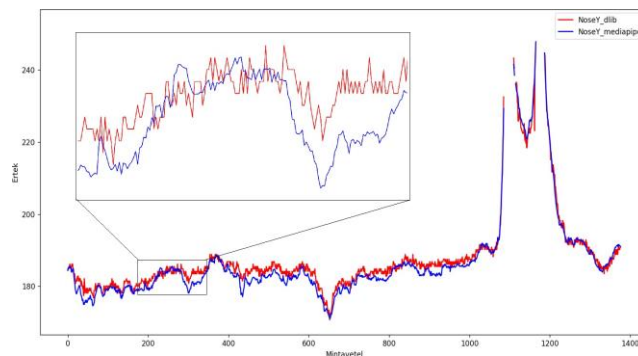
A 3. ábrán van szemléltetve a Dlib és Mediapipe könyvtárak által felrajzolt tájékozódási pontok összehasonlítása. A diagram melletti képek közül a felső felel meg a Dlib által felrajzolt tájékozódási pontoknak, az alsó a Mediapipe által



3. ábra. Dlib és Mediapipe összehasonlítása az EAR felrajzolásánál



4. ábra. Dlib és Mediapipe összehasonlítása az MAR felrajzolásánál



5. ábra. Dlib és Mediapipe összehasonlítása az orr középpont felrajzolásánál

felrajzoltaknak. A diagramon megfigyelhető két EAR érték időbeli változása. A piros vonal megfelel a Dlib és a kék vonal megfelel a Mediapipe által felrajzolt tájékozódási pontokból számolt EAR értéknek. A diagramon a hirtelen beesések a szem összecsukásának felelnek meg, ugyanis ekkor az EAR értéke lecsökken, a (1). képletnek megfelelően. Egy másik fontos észrevétel, hogy a Mediapipe által felrajzolt érték sokkal stabilabb a

Dlib által felrajzolthoz képest, kevesebb az oszcillálás, jobban látszódnak a beesések, ez azt eredményezi, hogy könnyebben felismerhető mikor a szem becsukódik. A 350. mintavétel környékén látható egy nagyobb beesés a piros vonalnál, ez azért történt mert a Dlib azokban a képekben nem ismerte fel a tájékozási pontokat és 0 érték van rendelve ahhoz az EAR értékhez.

4.1.2. MAR tulajdonság kinyerése

Hasonlóan az előző szekcióhoz a 4. ábra szemléleteti a MAR tulajdonság kinyerését. A piros vonal megfelel a Dlib és a kék vonal megfelel a Mediapipe által felismert tájékozási pontokból számolt MAR értéknek. Ezen az ábrán is megfigyelhető, hogy a 350. és az 500. mintavétel környékén nem sikeres a felismerés, ebben az esetben viszont nem kerül érték felrajzolásra. Ahol emelkedés látszik a diagramon, ott a vizsgált személy ásított és a (2). képletnek megfelelően a MAR érték növekedett. Ennél a diagramnál is látszik, hogy a Mediapipe által felismert tájékozási pontokból számolt MAR érték stabilabb. Az hogy a Mediapipe értéke átlagosan kisebb a Dlib által számolttal, az azzal magyarázható, hogy a figyelt tájékozási pontokat a Mediapipe közelebb rajzolja egymáshoz.

4.1.3. Orr középpont kinyerése

Az arc pozíciójának a kinyeréséhez az orr középpontját használjuk és annak az időbeli változását a figyelt képen. Ennél a tulajdonság kinyerésnél az orr középpontjára felrajzolt tájékozási pontot figyeljük. Az előző mérésekhez hasonlóan itt is a Mediapipe teljesít jobban. A 5. ábrán megfigyelhető a mérés az y koordináta szerint. Az 1200. mintavételnél megfigyelhető egy felugrás a vonalakban, itt a vizsgált személy kétszer leejtette a fejét. Ahol nincs érték a grafikonon, ott az arc pozíciója annyira eltért a kamera állásához képest, hogy az algoritmusok nem voltak képesek felismerni az arcot a képen.

4.2. Osztályozás

Osztályozóként egy Random Forest osztályozót választottunk, amely az NTHUDDD (lásd 3.4 fejezet) adathalmazból kinyert adatokon volt tanítva. A tanítás első lépésében feldolgozzuk az adathalmazt, kinyerjük a képekből az EAR, MAR és az arcpozíció értékeit, ami után a kinyert adatot kimentjük egy JSON fájlba. A tanítás következő lépésében kiolvassuk az adatot a JSON fájlból, felosztjuk tanító és tesztalmazra 80-20 arányban, majd tanítjuk az osztályozót. Az eredményt kiírjuk és elmentjük a modellt. A rendszer működése során az elmentett modellt felhasználva határozzuk meg a bejövő képekből a vezető állapotát.

5. TÁRGYALÁS

A következő táblázatban fel vannak tüntetve a tanítás eredményei összehasonlítva más dolgozatok eredményeivel.

Táblázat 1: Más szakdolgozatok hasonló eredményeinek összehasonlítása

Hiv.	Módszer/Használt modell	Adathalmaz	Pont.
[3]	4D CNN a szemek állapotának észlelésére	MRL Eye dataset	97.5%
[4]	CNN a szem állapotának felismerésére	Kaggle Drowsiness_dataset	97%
-	A dolgozatban bemutatott rendszer	NTHUDDD	87%
[5]	Android alkalmazás könnyű többrétegű NN-el	NTHUDDD	80%
[6]	Mobile rendszer TensorFlow-alapú algoritmussal	academic in the wild DDD	74%

6. TOVÁBBFEJLESZTÉS

A rendszer továbbfejleszthető, még rengeteg potenciál rejlik benne, mivel ehhez hasonló rendszer, amely vezetés közben akár a felhasználó mobiltelefonjáról figyelje az álmoságot még nem terjedt el széles körben a piacon.

Továbbfejlesztésként a feldolgozást el lehetne végezni a mobil eszközön, szerver használata nélkül. Kibővíteni a rendszert új adatokkal és finomhangolni.

A rendszer tervezése során megfigyeltük, hogy miközben a megfigyelt személy arcát kivágjuk a képről, a kivágott kép mérete változik és ez zajt hoz be a rendszerbe. Ezt a zajt ki lehetne küszöbölni zajsűrűséssel vagy a kivágott képek normalizálásával.

7. REFERENCES

- [1] Yaman Albadawi, Aneesa AlRedhaei, and Maen Takruri. Real-time machine learning-based driver drowsiness detection using visual features. *Journal of Imaging*, 9(5), 2023.
- [2] Yue Wu and Qiang Ji. Facial landmark detection: A literature survey. *International Journal of Computer Vision*, 127:115–142, 2019.
- [3] Israt Jahan, K. M. Aslam Uddin, Saydul Akbar Murad, M. Saef Ullah Miah, Tanvir Zaman Khan, Mehedi Masud, Sultan Aljahdali, and Anupam Kumar Bairagi. 4d: A real-time driver drowsiness detector using deep learning. *Electronics*, 12(1), 2023.
- [4] Mohammed Imran Basheer Ahmed, Halah Alabdulkarem, Fatimah Alomair, Dana Aldossary, Manar Alah- mari, Munira Alhumaidan, Shoog Alrassan, Atta Rahman, Mustafa Youldash, and Gohar Zaman. A deep- learning approach to driver drowsiness detection. *Safety*, 9(3), 2023.
- [5] Rateb Jabbar, Khalifa Al-Khalifa, Mohamed Kharbeche, Wael Alhajyaseen, Mohsen Jafari, and Shan Jiang. Real-time driver drowsiness detection for android application using deep neural networks techniques. 11 2018.
- [6] J.S. Wijnands, J. Thompson, K.A. Nice, et al. Real-time monitoring of driver drowsiness on mobile platforms using 3d neural networks. *Neural Computing and Applications*, 32:9731–9743, 2020.