
Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Irányítástechnika és Informatika Tanszék

# Arcfelismerő rendszerek támadása természetes módszerekkel

**TDK dolgozat**

Készítette:

Kófiás Kristóf

Konzulens:

dr. Szemenyei Márton

2022

# Kivonat

A mélytanulásos neurális hálók eredményesen használhatóak emberek azonosítására. A mai kor technológiájával már képesek vagyunk egy videófelvételen, egyetlen referencia kép alapján azonosítani személyeket, így széles körű megfigyelési rendszerek hozhatóak létre. Ezek használata viszont jogilag nem egyértelműen szabályzott és etikailag is megkérdőjelezhető.

Korábbi kutatások bebizonyították, hogy mind digitális, mind való életbeli módszerekkel eredményesen lehet ilyen rendszerek működését megzavarni. Ezek a módszerek azonban gyakran speciális szaktudást igényelnek, így a hétköznapi emberek számára nem elérhetőek.

A dolgozatban olyan szaktudást nem igénylő és természetes hatást keltő módszerekkel kísérletezek, amelyekkel az arcfelismerő rendszerek megtéveszthetők. Ilyen a smink, a szemüveg és a maszk viselete. Ezek olyan, a valóságban is használható módszerek, amik nem keltenek feltűnést.

A vizsgált módszereket az ArcFace arcfelismerő modell segítségével értékeltem ki. A tesztben 4 alany vett részt, akikről videó felvételeket készítettem. A kiértékelés során megvizsgáltam, hogy az arcfelismerő rendszer százalékban mért pontossága milyen mértékben csökken az egyes technikák alkalmazása esetén.

# Abstract

Deep learning neural networks can be effectively used to identify people. With today's technology, we are already able to identify people based on a single reference image in a video recording, thus wide-range surveillance systems can be created. However, their use is not clearly regulated by law and may also be ethically questionable.

Previous research has proven that both digital and real-life methods can effectively attack such systems. However, these methods often require special expertise, so they are not available to ordinary people.

In this paper, I experiment with natural-looking methods that do not require special knowledge, and are capable of fooling facial recognition systems. These include wearing makeup, glasses and masks. Also, these are methods that can also be used in real-world settings and are inconspicuous.

I evaluated the investigated methods using the ArcFace face recognition model. Four subjects participated in the test, of whom I made video recordings. During the evaluation, I examined to what extent the accuracy of the face recognition system, measured in percentages, decreases when using these techniques.

# 1. Bevezetés

A napjainkban már kiterjedt biztonsági kamera-rendszerek és a dinamikusan fejlődő mély neurális hálók eredményeként már csak szándék kérdése egy kiterjedt arcfelismerő hálózat kiépítése, amely képes lenne emberek milliót beleegyezésük nélkül követni. Az interneten szinte mindenkiről szerepelnek olyan védelem nélküli képek amelyek alkalmasak arra, hogy olyan adatbázist lehessen belőlük építeni, amely szükséges egy ilyen rendszer működtetéséhez. Ezekből a képekből egyszerűen kinyerhetőek egy személy azonosításához szükséges biometriai azonosítók, melyek ezután alkalmasak az egyén megfigyelésére nyilvános tereken is.

Az elmúlt évtizedben a mély neurális hálók átvették az arcfelismerés hagyományos módszereinek helyét, amik minden eddiginél nagyobb pontossággal azonosítanak személyeket. A Google által fejlesztett FaceNet már 99.63% pontosságot ért el az arcfelismerés legfontosabb benchmark tesztjén, a Labeled Faces in the Wild (LFW) adatbázison [1, 2]. Ez már az emberi pontosságot is meghaladja, így eredményesen és nagy hatékonysággal alkalmazható nagy adathalmazokon [3]. Az ilyen neurális hálók módosított képosztályozó konvolúciós hálózatok, amelyek utolsó softmax rétege eltávolításra került, így egy, a modell architektúrájától függő méretű vektor lesz a kimenete, amely alkalmas arra, hogy egy személyt egyedileg reprezentáljon. Ennek köszönhetően nincs szükség a felismerendő célszemélyek arcát a modell tanításakor ismerni, tehát univerzálisan használható egy arcot reprezentáló egyedi azonosító előállításához.

Ilyen rendszerek számos személyiségi-jogi és etikai kérdéseket is felvetnek. A megfigyelés abban az esetben is befolyásolja az embereket, ha nem tesznek törvénybe ütköző cselekedeteket, vagy nem is terveznek ilyet. Ilyen esetben az emberek hajlamosak saját viselkedésüket cenzúrázni, és ezt akár tudat alatt is, az őket körülvevő csoport normáihoz igazítani. Mindezek a hatások hozzájárulhatnak szorongás előidézéséhez, mely jelentősen befolyásolja az egyén kifejezőképességét és általános jólétét [4].

Fontos problémaforrás még a biometrikus azonosítók megváltoztathatatlanságának kérdése is. Az arc elválaszthatatlan az embertől és az emberi identitás fontos része. A hagyományos azonosítási módszerek biometrikusra cserélése az emberi testet egy egyszerű eszközzé, adattá degradálja, amelyet nem lehet megváltoztatni, ezzel megfosztva az embert a saját teste és adatai feletti ellenőrzéstől. Ennek következménye, hogy az arcból képzett adat már elválasztható az embertől, így az egyén személyes identitása mások által is

birtokolható, használható, tehát megszűnik személyesnek lenni [5].

Az arcfelismerő neurális hálók tanítására használt legnépszerűbb adatbázisok nagy része olyan képekből áll, amelyeknek felhasználásához a képen szereplő személyek nem járultak hozzá. Bár a képek jogilag szabadon felhasználhatóak, az adatbázisban szereplő személyek nem tudtak hozzájárulni a biometrikai azonosítóik felhasználásához. Ezeket az adatbázisokat később állami, kereskedelmi és katonai kutatásokhoz is felhasználták, a személyes adatokat termékké alakítva [6].

Az ilyen adatok kinyerését szinte lehetetlen megakadályozni, így csak az adatok felhasználását tudjuk befolyásolni. A megfigyelés hatásai elkerülhetőek, az arc olyan jellegű módosításával, ami egy arcfelismerő rendszer számára az arcot felismerhetetlenné teszi, ezzel biztonságérzetet nyújtva az egyénnek. Az arcfelismerő neurális modellek működésének ismeretével lehetséges olyan smink generálása, amely jelentősen csökkenti a felismerés sikerességét, azonban ez a módszer speciális szaktudást igényel, amely nem általánosan hozzáférhető.

A dolgozatban a fenti problémák orvoslására olyan módszereket vizsgálok, amelyek nem igényelnek ilyen szaktudást, és bárki számára hozzáférhetőek. Többféle ilyen módszer létezik: az arc bizonyos részeinek mindennapos kiegészítőkkel, mint például napszemüveggel, vagy maszkkal letakarása; smink használata, vagy az arcszőrzet változtatása. Mindezek a módszerek az embereket is bizonyos szinten megzavarhatják, így megvizsgálom, hogy ezek az eszközök alkalmasak-e a személyiségi-jogok védelmére és azt is vizsgálom, hogy milyen módon zavarják ezek egy arcfelismerő modell működését.

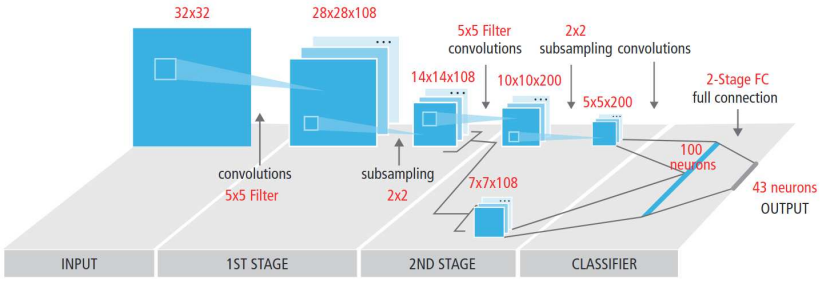
A vizsgálataim alapján ezek a módszerek csak korlátozottan alkalmasak erre a feladatra. Egy, az interneten szabadon elérhető előtanított modell esetében is a kiinduló 53,77% sikerességi rátát mindössze 1,28%, 23,15%, 50,18%, 45,55% értékekre sikerült csökkenti sorban a maszk viselésével, napszemüveg viselésével, smink használatával, valamint a szemöldök eltávolításával.

# 2. Korábbi munkák

A neurális hálók, ahogy a nevük is sugallja neuronokból épülnek fel, amelyek a valódi, biológiai idegsejtekhez hasonlóan kapcsolatban állnak egymással, így hálózatot képezve. A biológiai megfelelőik, az axonjaik segítségével kapcsolódnak más idegsejtek dendritjeihez, ezeken átadva az elektromos kisüléseket. Amikor a beérkező elektromos kisülés elér egy bizonyos szintet, az idegsejt egy választ küld axonjain keresztül a hozzá kapcsolódó más sejteknek.

Ezzel analóg módon működnek a mesterséges neurális hálók is. A neuronok a bemeneteiken érkező értékeket összegzik, majd a hozzájuk rendelt aktivációs függvény alapján a kimeneteiken valamilyen értéket továbbítanak a háló következő rétegének. A bemeneti rétegen beérkező adatok a háló rejtett rétegein áthaladva jutnak el az utolsó réteghez, amelyen így megjelenik a bemenethez tartozó eredmény. Az ilyen hálózatokat teljesen kapcsolt hálózatoknak nevezzük.

A konvolúciós neurális hálók (CNN) egy ilyen módon felépített háló speciális változata, amely egy, vagy több konvolúciós rétegből áll. Ezeket a hálózat végén néhány teljesen kapcsolt réteg követ. Ez a módszer az idegrendszer sejtszintű vizsgálatával szemben, a látás mechanizmusát követi. Először a bemenetet egészben, majd egyre kisebb részleteiben elemzi, így az adatok részeinek térbeli elhelyezkedését is figyelembe veszi. Ezért elsősorban képek elemzésére a legalkalmasabb, de más felhasználási területei is vannak [7].




1. ábra. Egy konvolúciós neurális hálózat felépítése. Forrás: [7]

Egy konvolúciós réteg általánosan $H \times W \times C_{in} \times C_{out}$ dimenziójú, ahol a $H$ és $W$, sorban réteg magassága és szélessége, $C_{in}$ és $C_{out}$ a réteg bemeneti és kimeneti csatornáinak száma. Ahogy a hálóban egyre mélyebbre jutunk,

úgy növekszik a csatornák száma, ezzel együtt csökken a réteg kiterjedése. Az utolsó konvolúciós rétegben végül egy $1 \times 1 \times n$ méretű vektort kapunk, ami a bemenetünket azonosító feature vektor lesz. Ezt használhatjuk teljesen kapcsolt rétegekben osztályozási feladatokra.

A konvolúciós neurális hálók az elmúlt évtizedben váltak népszerűvé, mivel a korábbi megoldásoknál jobban alkalmasak képfeldolgozási eljárásokra. Teljesen kapcsolt hálózatokban nagy méretű képek feldolgozása, a sok paraméter miatt nehézkes. Egy ilyen hálózat nem csak nagy számítási kapacitást, de sok tárhelyet is igényel. Ezzel szemben a konvolúciós hálózatok kevesebb paramétert tartalmaznak, így tanításuk gyorsabb, kevesebb memóriát igényelnek, valamint kevésbé érinti őket túlillesztés problémája.

Az egyik leggyakrabban használt konvolúciós architektúra a ResNet. Ez az architektúra a rétegeket reziduális blokkokba szervezi, amelyek bemenete és kimenete között egy kapcsolatot létesít. Ezzel elkerülhetőek az eltűnő, vagy felrobbanó gradiensek miatti konvergencia problémák. Így a konvolúciós hálózatok mélysége a korábbi 30-ról akár 100-200 mélységűre is növelhető [8].

## 2.1. Arcfelismerő rendszerek

Az arcfelismerés az objektumazonosítás egyik speciális formája. Nehézségét az okozza, hogy minden embernek különböző arca van, így azokat mind külön osztályként kellene kezelni. Ez azonban feltételezi, hogy már a tanítás során ismerjük a felismerendő személyekhez tartozó képeket. Ez viszont csak nagyon speciális környezetben alkalmazható módszer, például akkor, ha tudjuk, hogy a rendszerünkhöz később sem fogunk új személyeket hozzáadni. A másik problémát pedig a tanító adatok száma okozza. A megfelelően pontos eredményekhez ugyanis minden alanyról több képre lenne szükség. Speciális környezetben ez szintén megoldható, viszont amennyiben minden személyről csak egy képpel rendelkezünk, például az útlevélen található fotóval, akkor ez a megoldás nem alkalmazható.

Ezen problémákat úgy lehet orvosolni, hogy a modellünket minél általánosabban tanítjuk meg *arcok* azonosítására. Tehát nem konkrét személyt kell felismernie, hanem csak egy azonosító adatot kell generálnia. Ezt az adatot tudjuk ezután a modellen kívül az azonosításra használni. Két ilyen leíró adatot összehasonlítva megkapjuk, hogy azonos személyhez tartoznak-e, vagy sem. Az összehasonlítás elvégezhető a két vektor távolságának kiszámításával, vagy egy erre specializált neurális hálóval is. A távolságszámítás legelterjedtebb módszere a koszinusz távolság használata, melynek képlete

$d_{cos}(x, y) = 1 - (\vec{x} \cdot \vec{y})$. A felismeréshez használt adatbázis alapján - amelyben benne van az összes lehetséges alany - kiszámolható egy határérték, ami alapján, két arcot reprezentáló vektorról eldönthető, hogy egy személyhez tartoznak-e.

Az arcfelismerő modelleknek több szempontból is nehéz feladat az emberek azonosítása. A hagyományos objektumazonosítás során is felmerülnek olyan problémák, mint a tárgyak elhelyezkedésének különbözősége, a megvilágítás minősége, az objektumosztályon belüli tárgyak általános különbözősége. Az emberi arc azonban ennél is több kihívást rejt, mivel ezeken kívül az arcmimika is változik, valamint az arcot eltakarhatják kiegészítők, haj, vagy arcszőrzet is, amelyek emberi szemszögből az arc fontos elemei is lehetnek.

Egy arcról emberi szemmel is könnyen megmondhatjuk, hogy az egy arc, mivel mindegyik hasonló felépítést követ: ovális, esetleg kerek forma, két szem, orr, száj és mindezek elhelyezdéseinek arányai is bizonyos határokon belül mozognak. Azonban ennek ellenére is minden arc egyedi, amelyet annak bizonyos részeinek kis eltérései okoznak. Ez sok lehetőséget biztosít az azonosító adatot létrehozó modell megzavarására.

A modell belső felépítését ismerve olyan mintát lehet generálni amit valamilyen módon az arcra helyezve, nem csak a felvételen szereplő személy kilétét lehet elrejteni (*dodging attack*), hanem akár egy másik embert is meg lehet személyesíteni (*impersonation attack*).

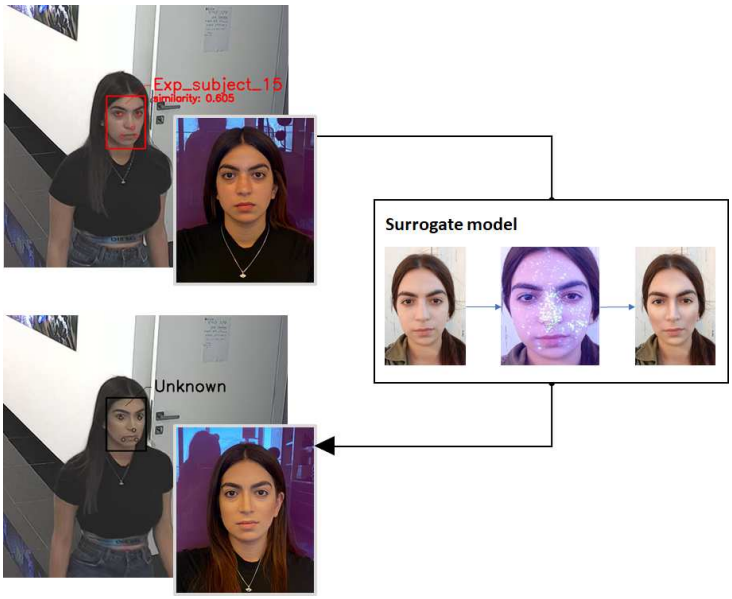
Erre mutattak módszert Sharif és tsai. [9], akik egy papírszemüvegre nyomtattak egy olyan mintát ami ezt lehetővé tette. Bár a módszer a kísérleteikben 100%-os hatékonyságot mutatott mindkét esetben, speciális szaktudást igényel és az ily módon létrehozott szemüveg feltűnő is. A modellt fekete dobozként értelmezve is sikerült a módszert eredményesen alkalmazniuk, azonban ehhez is szükséges a modell használatához való hozzáférés.

Kevésbé feltűnő módszert használtak a Dodging Attack Using Carefully Crafted Natural Makeup [10] szerzői, akik olyan sminket generáltak, amely természetesnek tűnik, mégis képes az azonosítást megakadályozni. A tesztben 20 alany - 10 férfi és 10 nő - vett részt, akiknek egy folyosón, két kamera előtt kellett elsétálniuk. A kamerák elhelyezkedése és dőlésszöge a biztonsági kamerákéhoz hasonló, így hasonlít valós gyakorlati alkalmazásokhoz.

Az ArcFace [11] mély neurális háló alapú azonosító modell használták, amely az LFW adatbázison 99,83% pontosságot ér el. Az így kinyert, az arcot azonosító vektorokat hasonlították össze, az adatbázisban szereplő adatokkal. Az adatbázisban a 20 alany mellett a teljes LFW dataset is szerepelt. Két arcot akkor tekintettek azonosnak ha a leíró adataik koszinusz távolsá-

ga kisebb volt, mint 0,42. Az arcdetektáló modell egy MTCNN (Multi-task Cascaded Convolutional Network) [12] volt.

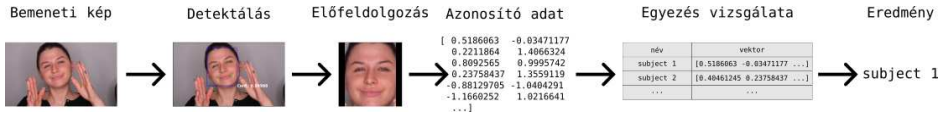
Ezzel a módszerrel a kezdeti 47,57%-os sikerességi rátát, 1,22%-ra csökkentették. A dolgozat későbbi részében rámutatok arra, hogy két arc azonosságához használt határérték, ehhez az adatbázishoz túl alacsony, abban sokkal nagyobb tűrés található, amivel jelentősen növelni lehet a kezdeti sikerességi rátát.




2. ábra. Az ábrán látható, hogy a korábban felismert alanyt, a smink alkalmazása után, már nem ismeri fel a rendszer. (Forrás: [10])

# 3. Technikai háttér

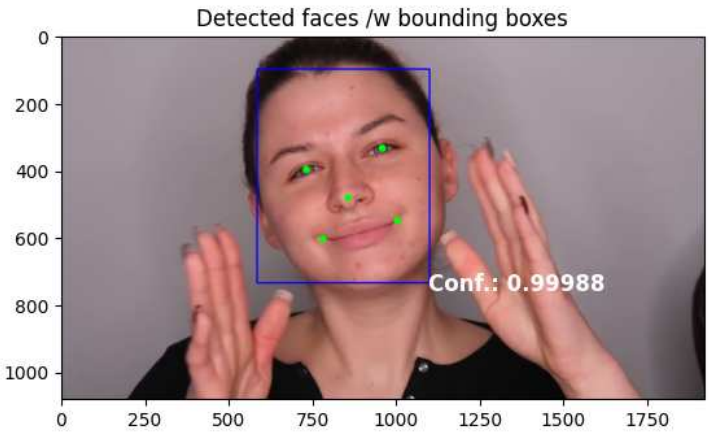
Az arcfelismerés folyamata többlépcsős. A valós környezetben használt rendszerek esetében általában egy videófolyamon kell az arcot megtalálni, majd azonosítani. A videóból kivágott képkockán az arc azonban bárhol elhelyezkedhet, különféle szögekből látszódhat. Ezért az arcfelismerési csővezeték első eleme az arc detektálása, pozíciójának meghatározása. A megtalált arcot, ezután az arc azonosító adatainak kinyerésére szolgáló modell bementének megfelelően kell átalakítani. Az átalakított képből megkapható az arcot azonosító egyedi adat, amelyet ezután az adatbázisban található személyek adataival kell összehasonlítani.




3. ábra. Az arcfelismerési csővezeték. A bementi kép forrása: Youtube

## 3.1. Detektálás

A detektálás (detection) célja egy képen egy arc helyzetének és azon belül a szem, orr és száj pozícióinak meghatározása. Az így kapott adatok segítségével az arc, az arcfelismerő modell bemenetének megfelelően körbevágható és igazítható.




4. ábra. A detektáló modell megadja az arcot befoglaló téglalapot és a szemek, az orr, és a száj pozícióit. A bemeneti kép forrása: Youtube

A hagyományos, nem mélytanulás alapú módszerek egyike a Viola-Jones objektumdetektáló, ami elsősorban az arcdetektálás problémájára lett kifejlesztve. Működésének elve azon alapul, hogy az emberi arcok, bár nagyon különböznek egymástól, jól elkülöníthető jellemzői vannak. Ilyen például az orr, és a szemek helyzete. Az algoritmus előnye, hogy gyors, viszont a mélytanulásos módszereknél pontatlanabb [13].

Ennél pontosabb eredményt adnak a konvolúciós hálózatokra épülő modellek. A legelterjedtebb módszer egy Multi-task Cascaded Convolutional Network (MTCNN)[12] alapú modell használata. Az ilyen architektúrák több konvolúciós hálózatot tartalmaznak, mindegyiket más-más feladatra specializálva. Így egy hálózatból megkapjuk az arc helyzetét, és az arc főbb pontjainak helyzetét amire az arcfelismerési csővezeték következő lépésében is szükség van.

A detektálás eredménye a bemeneti képen látható arcok mindegyikéhez egy, az arcot befoglaló téglalap, és a szem, orr és száj helyzete.

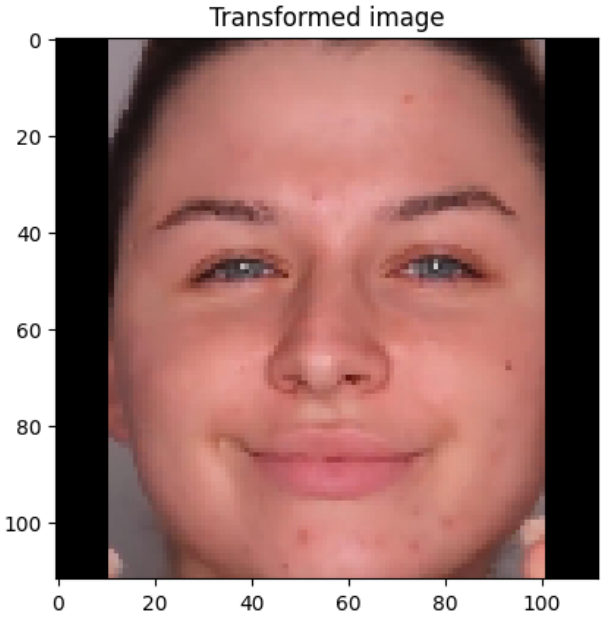
## 3.2. Előfeldolgozás

Az előfeldolgozás (preprocessing) lényege, hogy a detektált arcot olyan formátumba konvertáljuk át, ami megfelelő az arcfelismerő modell bemenetének és egyezik a tanítás során alkalmazott módszerrel.

A két szemnek egymáshoz képest, egy vízszintes vonalban kell elhelyez-

kednie. Ez azért fontos, mert az arcot befoglaló téglalapba így fér bele megfelelően az arc. Így a modellnek átadott kép a lehető legkevesebb nem-arc pixelt fog tartalmazni.

A kép mérete, meg kell, hogy egyezzen a modell bemenetének dimenzióival. Mivel az arcok arányai jelentősen eltérnek egymástól, de a modell bemenete meghatározott dimenziójú, így szükséges a képet valamilyen módon átalakítani. Vagy az arcokat megnyújtjuk, ezzel megváltoztatva az arányokat, vagy az arányok megtartása érdekében a képet fekete pixelekkel egészítjük ki a szélén.




5. ábra. Az arc főbb pontjai alapján elforgatott, majd az arc arányait megtartva, 112x112 méretűre, a széleken kiegészítve átméretezett kép. A bemeneti kép forrása: Youtube

## 3.3. Azonosító jellemzők kinyerése

Az arc gépi felismeréséhez először az arcokból numerikus adatot kell kinyerni (feature extraction). Ez tipikusan egy többdimenziós vektor.

A mélytanulásos modellek az osztályozó modellekkel megegyezően tanítják: az adatbázisokat train, test, validate egységekre bontják, majd $n$ darab

osztály felismerésére tanítják be őket, ahol ezek az osztályok különböző személyek lesznek. Ezáltal egy olyan modell keletkezik, amely képes a tanítás során felhasznált emberi arcokat felismerni. Végül a modell utolsó softmax rétegét eltávolítják, így a modell kimenete egy, a használt architektúrától függő méretű vektor lesz. Ezzel a vektorral fogjuk azonosítani a tanítás során nem ismert arcokat is.

## 3.4. Egyezés vizsgálata

A képeket reprezentáló numerikus adatokat (többdimenziós vektorokat) azok távolságai alapján tudjuk összehasonlítani. Ehhez a k-Nearest-Neighbours (kNN) módszert a leghatékonyabb implementálni. A keresett $Target$ vektort és az adatbázisban szereplő összes $Reference$ vektort összehasonlítjuk valamilyen $\phi(T, R)$ metódussal, ahol $\phi(T, R)$ valamilyen távolságszámító algoritmus, eredménye pedig egy olyan érték, amely minél alacsonyabb, annál közelebb van egymáshoz a két adat.

Két vektor összehasonlítására több ilyen távolságszámító algoritmus is használható.

A Manhattan távolság ami minden dimenzióban a pontok távolságának összege:

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

Az euklidészi távolság a Pitagorasz tételen alapul. A két vektor különségének négyzetének gyökével egyenlő:

$$d(x, y) = \sum_{i=1}^{n} \sqrt{(x_i - y_i)^2}$$

A koszinusz távolság pedig a két vektor skaláris szorzata:

$$d(x, y) = 1 - (\vec{x} \cdot \vec{y})$$

Mérések alapján a koszinusz szorzat bizonyult a legpontosabbnak és legstabilabbnak, ezért általában ez terjedt el ilyen alkalmazásokban [14].

# 4. Javasolt módszerek

Jelen dologzatban a személyazonosság elrejtésének olyan módszereit vizsgálom, melyek szaktudás nélkül is bárki számára elérhetőek, valamint csak hétköznapi eszközök szükségesek hozzájuk.

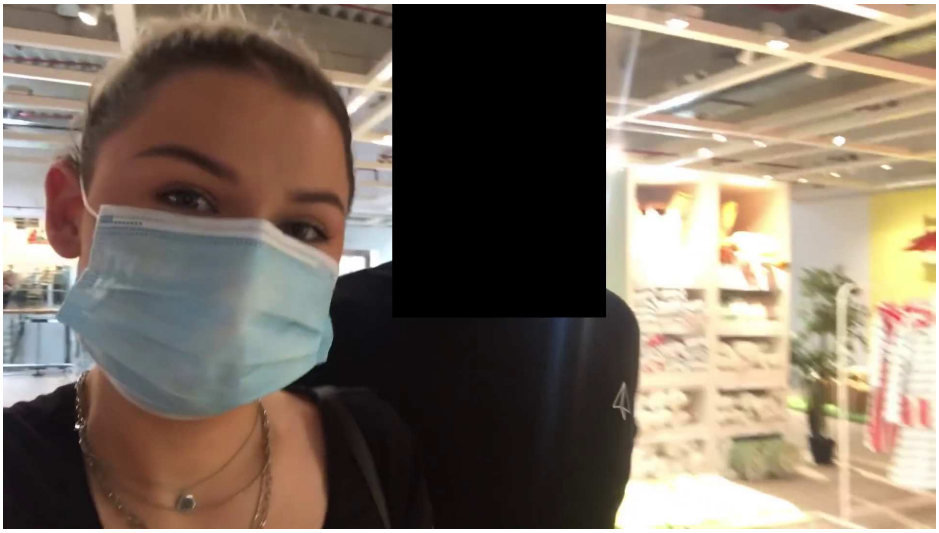
Ezeknek a módszereknek egy részét az arcfelismerés első, a referencia kép felvételének fázisában, más részüket, működés közben rögzített kép készítésekor lehet használni.

Korábbi kutatások azt mutatják, hogy az emberek számára egy arc azonosításának markerei fontossági sorrendben, a szemöldök, szem, száj és az orr [15]. A következőkben bemutatott módszerek ezeket a régiókat módosítják. A dolgozat következő fejezetében ezeket a módszereket vizsgálom egy neurális modell esetében.

## 4.1. Maszk

Az egyik legkézenfekvőbb eszköz személyazonosságunk elrejtésére a maszk. Egy maszk az arc nagy részét, főbb pontjai közül az áll, a száj és az orr környékét is eltakarja. Fontos viszont, hogy a szemet és a szemöldököt nem takarja. Utóbbi az emberek számára az egyik legfontosabb azonosítása pont az arcon.

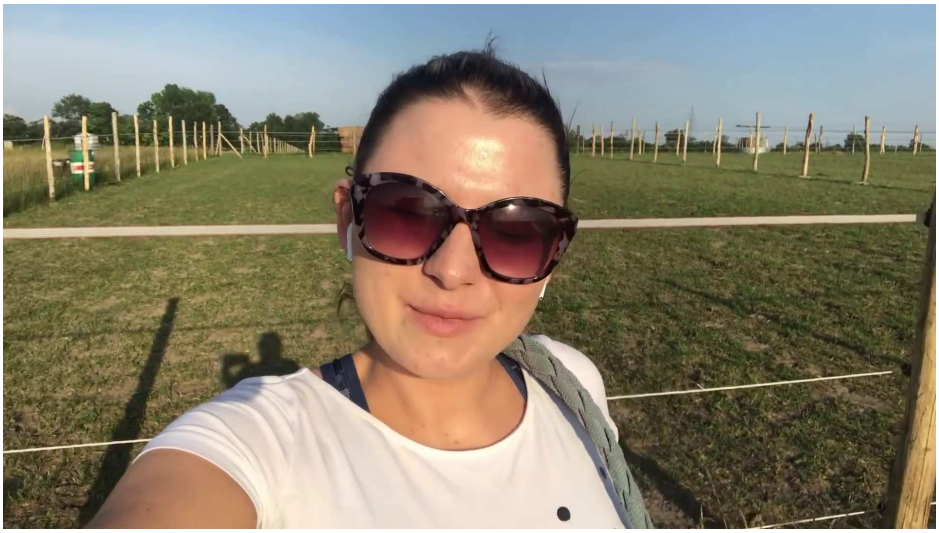
A COVID-19 pandémia óta, már általánosan elfogadott a maszk viselete nyilvános tereken, így bár feltűnést kelt, de nem tűnik gyanúsnak. A maszkra mérete és elhelyezkedése miatt mintákat, akár egy másik ember arcának képét is el lehet helyezni. Ez a módszer viszont nem alkalmazható a referencia kép elkészítéséhez, mivel az arcot részlegesen, vagy egészen takaró eszközöket a képek elkészítése előtt el kell távolítani [16].



6. ábra. Az arc nagy részét eltakaró maszkot viselő személy. A kép forrása: Youtube

## 4.2. Napszemüveg

A napszemüveg ez szemben az arc felső részét, így a szemet és szemöldököt is takarja. Elterjedt módszer a szem kitakarása olyan esetekben is, amikor egy fotón kell elrejteni a rajta szereplő személy kilétét. Ezért ez a módszer is hatékony egy emberi szemlélő ellen, viszont szintén nem lehet a referencia kép elkészítéséhez használni.



7. ábra. A szemet és szemöldököt eltakaró napszemüveget viselő személy. A kép forrása: Youtube

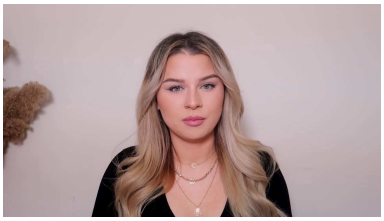
## 4.3. Szemöldök eltávolítása

A szem eltakarása nélkül magát a szemöldököt is el lehet tüntetni. Korábbi kutatások azt találták, hogy csak a szemöldök eltávolítása jelentősen csökkenti az arcfelismerés hatékonyságát, lényegesen jobban, mintha a szemet távolítanánk el [17]. Kis mértékben feltűnést kelthet, azonban semmiképpen sem gyanús, ha valakin nincs szemöldök. Ez a módszer a referencia kép készítésekor és működés közben is használható, azonban csak az egyiknél. Olyan esetekben, amikor a referencia képet digitálisan kell leadni, digitálisan is eltávolítható a szemöldök.

8. ábra. A képről digitálisan eltávolított szemöldökű személy. A kép forrása: Youtube

## 4.4. Smink

A smink egy olyan eszköz, amit nagyon sokféleképpen lehet felhasználni. Szinte észrevehetetlen, vagy, akár az arcot teljesen megváltoztató formában is alkalmazható. A nagyon extrém változatai nyilván feltűnést keltenek, de a hétköznapi, természetes sminkkel is jelentősen lehet módosítani, akár az arc formáján és a bőr színén is a sminkeszközök megfelelő használatával. Komplexebb alkalmazásai igényelnek bizonyos fokú szakértelmet, azonban egyszerűbb formái nem. A szemöldök eltávolításához hasonlóan szintén használható az azonosítás akármelyik fázisában, de akár mindkettőben is.



(a) Természetes smink Forrás: Youtube

(b) Extrém smink Forrás: Youtube

9. ábra. Az arcot alig megváltoztató természetes smink (a) és a szinte felismerhetetlenségig eltorzító extrém smink (b)

# 5. Kísérleti eredmények

Nitzan Guetta és tsai. tanulmányához [10] hasonlóan itt is az ArcFace modellt fogom használni, pontosan ugyanazt az előtanított változatát, amit ők is használtak. Ez az ArcFace ResNet100 alapú változata, az *LResNet100E-IR, ArcFace@ms1m-refine-v2.* [1] A távolságmetrika hasonlóan a koszinusz távolság lesz.

Így a felismerés elkerülésének hétköznapi módszereinek sikeressége összhasonlítható a speciális szaktudást igénylő módszerrel.

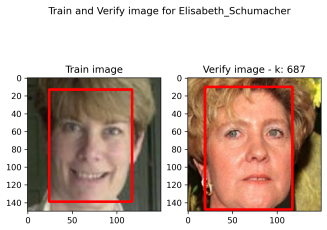
## 5.1. Baseline pontosság meghatározása

A letöltött előtanított modell minsőgének ellenőrzésére először meghatározom annak pontosságát. A használt tanító adatoktól függően az ilyen modellek pontossága jelentősen eltérhet egymástól, ezért használat előtt mindenképpen érdemes ellenőrizni azokat.

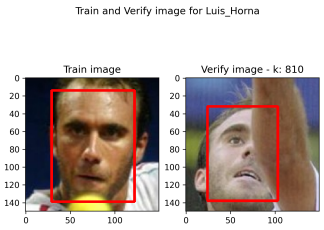
A pontosság meghatározáshoz az LFW adatbázist használom. Az adatbázis 13233 képet tartalmaz 5749 személyről. Azonban mindössze 1680 személyről van kettő vagy több kép, így csak ezeket lehet használni ezekhez a tesztekhez Szükség van egy referencia képre, amiről az arcfelismerő rendszer adatbázisába eltárolhatjuk az arcot azonosító vektort, majd szükség van egy célpont képre is, amit ehhez az adatbázishoz hasonlítunk.

Az adatbázis tartalmaz hibákat is, amelyet az adatbázis publikálása után találtak meg, ezért néhány személyt kihagytam a vizsgálatból. Így 1666 kép párt tartalmaz az adatbázis.

[1] https://github.com/deepinsight/insightface/tree/master/model_zoo



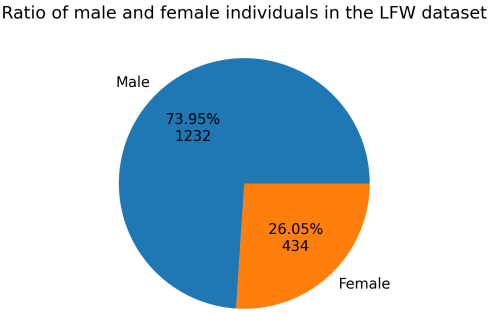

(a) A két képen különböző személyek láthatóak. A bal oldalon Elizabeth Schumacher, a jobb oldalon Elisabeth Schumacher




(b) A jobb oldali kép nagy része ki van takarva, így ez is hibás eredményre fog vezetni

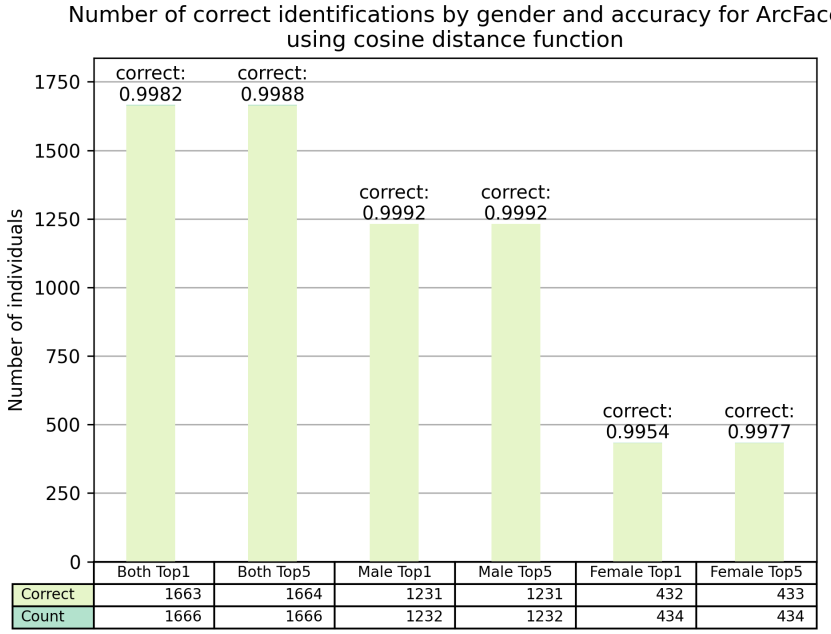
10. ábra. Példák az LFW adatbázis hibáira.

Az adatbázisban nem egyenlő arányban szerepelnek női és férfi arcok, ezért fontos, hogy az eredményeket külön is értékeljük, mivel jelentős eltérés tapsztalható a két csoport számossága között. Érdemes megjegyezni, hogy a modell tanítására használt MS1MV2 adatbázis is jelentősen több férfi arcot tartalmaz, mint nőit. Az alanyok számának kiegyenlítése a két csoport között, azonban nem feltétlenül jelenti a felismerés sikerességének kiegyenlítődését [18].




11. ábra. A férfi és női személyek aránya az LFW adatbázisban. Mivel a két csoport mérete között jelentős eltérés van, így a továbbiakban nemek szerint lebontva is vizsgálom a pontosságot.

Az adatbázisban minden két képpel rendelkező személyhez tartozó első kép szolgál referencia képként, a második pedig a célpont képként, amihez az adatbázisból meg kell találni a hozzá tartozó nevet. A vektorok távolsága szerint sorba rendezett eredmények között, az első helyen szereplő címke egyezése a célpont nevével számít sikeres felismerésnek. Vizsgálom továbbá a Top5 sikerességet is, amikor elengendő a helyes címkének az első öt hely között szerepelnie.




12. ábra. Az ArcFace pontossága *koszinusz* távolság-számító algoritmust használva. Mindkét nemre összegezve 99,82%-os Top1 eredményt lehet vele elérni az LFW adatbázis hibák nélküli, redukált változatán.

## 5.2. Határérték számítása

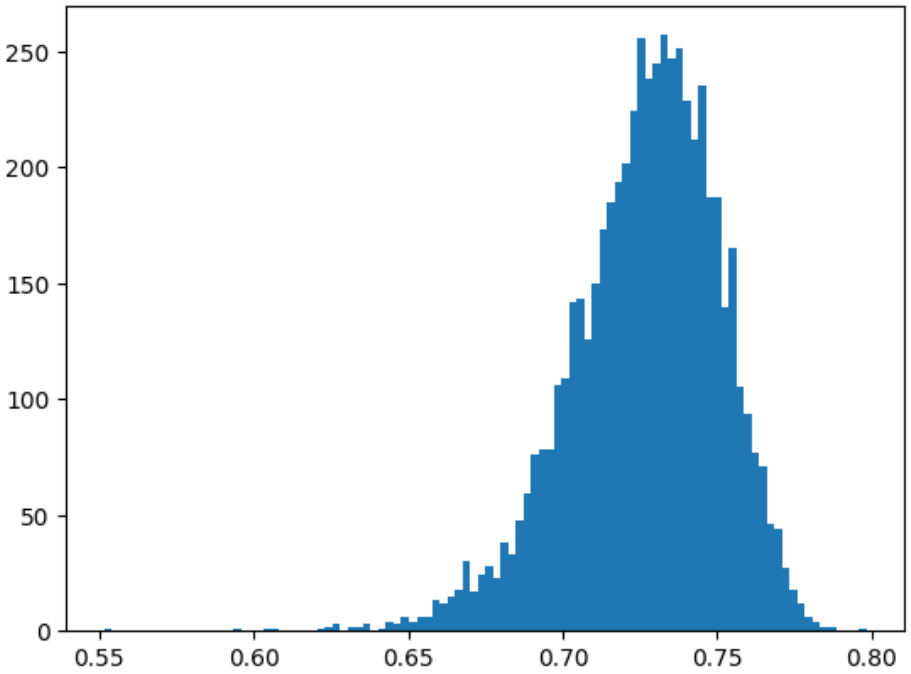
Ahhoz, hogy két arcról eldönthessük, hogy azonos személyhez tartoznak, azonosító vektoraik távolságának egy bizonyos határértéken belül kell maradniuk. Intuitív módon is belátható, hogy két véletlenszerűen választott arcpár között akár jelentős, vagy akár csak minimális különbség is lehet. Ha az adatbázisban nem is szerepel a keresett személy, akkor is lesz egy, az ő arcát reprezentáló vektorhoz legközelebbi találat, egy olyan személy akinek arca

a legjobban hasonlít az övére. Ezt hívjuk hamis pozitív találatnak, amikor eredményünk látszólag igaz, azonban a valóság szerint nem az. Ezt a problémát egy határérték megállapításával orvosolhatjuk. Minden, a határértéknél alacsonyabb távolságot véve helyes találatnak, minimalizálhatjuk a hamis pozitív találatok számát.

Legyen $d(x, y)$ valamilyen távolságszámító algoritmus és $\varepsilon$ a határték. Ekkor az azonosság sikeressége a következőképpen írható le:

$$Success = \begin{cases} True, & \text{if } d(x,y) < \varepsilon, \\ False, & \text{if } d(x,y) > \varepsilon \end{cases} \tag{1}$$

A határértéket minden adatbázisra egyedileg érdemes kiszámolni, mivel így pontosabban meg tudjuk azt határozni. A kiszámításhoz azt számoljuk ki, hogy, egy az adatbázisban, nem szereplő archoz, mi a legközelebbi, az adatbázisban szereplő arc távolsága. Az adatbázisból eltávolítjuk az $i$-edik arcot, majd kiszámoljuk a távolságát az adatbázisban lévő többi archoz képest. Az így kapott legalacsonyabb távolságok minimuma lesz az adatbázishoz tartozó határérték, mivel ez az érték alatti távolságok biztosan egy archoz tartoznak. Természetesen az adatbázisban nem megtalálható új arcokhoz ennél kisebb távolság is lehetséges, azonban kellően nagy mintával ez a hiba is minimalizálható.




13. ábra. A kiszámított legalacsonyabb távolságok hisztogramja.

| min | mean | max |
|---|---|---|
| **0.5516** | 0.7267 | 0.7983 |

A hisztogrammon látható, hogy egy kiugró adat is van, ezért ezt, és ez az érték nélküli nagyobb határértéket is tesztelem.

A kiugró adat nélküli értékek:

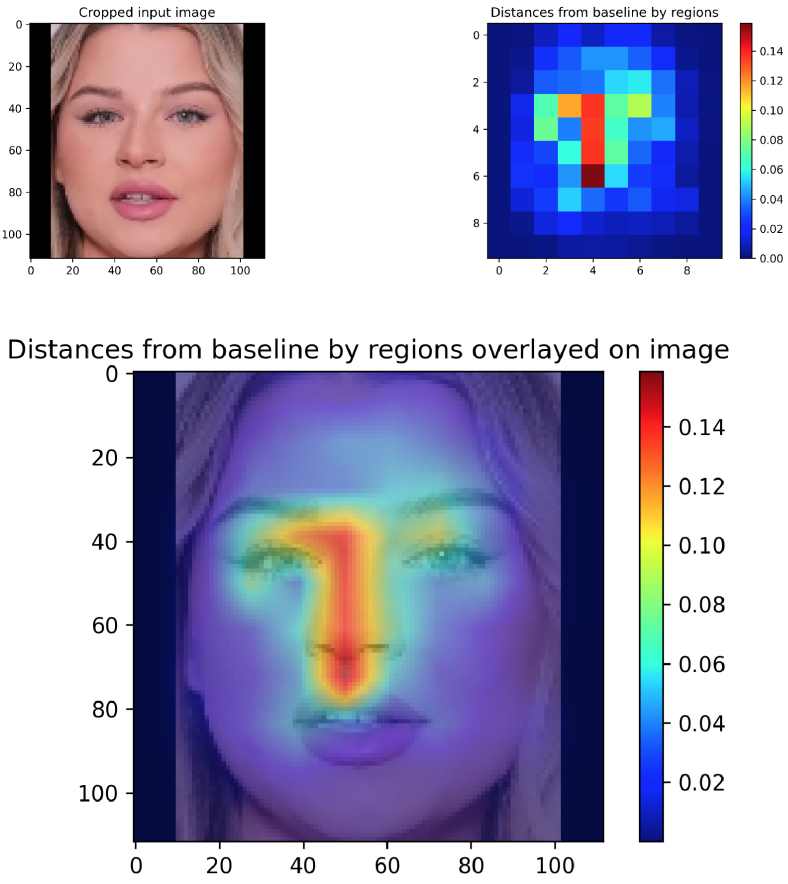
| min | mean | max |
|---|---|---|
| **0.5947** | 0.7267 | 0.7983 |

## 5.3. Kitakarás vizsgálata

Annak megállapítására, hogy az arc mely részeinek megváltozására a legérzékenyebbek a modellek, az arcot részenként le kell takarni, majd megvizsgálni, hogy a letakarással keletkezett kép milyen távolságra van az eredetitől. Az adatokból egy olyan heatmap generálható, amin megvizsgálhatjuk, hogy milyen részek letakarásai okozzák a legnagyobb eltérést. Ez alapján eldönthető,

hogy a vizsgálandó módszerek közül melyek azok, amelyek a legígéretesebbnek bizonyulnak.

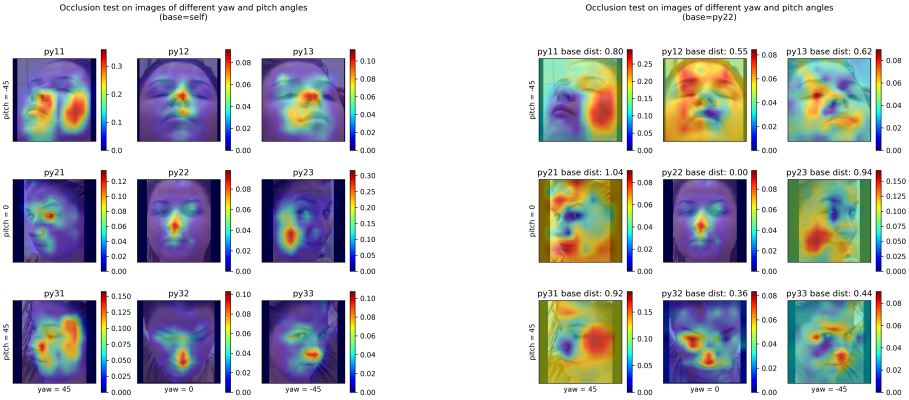
A tesztben 10 sorra és 10 oszlopra osztom a képet és az így keletkezett téglalapokat sorban kitakarom.




14. ábra. Az eredeti képre rávetített occlusion map. A bemeneti kép forrása: Youtube

A valóságban azonban fontos, hogy az arc csak nagyon ritkán néz a kamerával szemben, így más szögekből is meg kell vizsgálni. Mivel a referencia kép minden esetben szemből készül, ezért ehhez képest is érdemes a távolságokat mérni. Azért, hogy ebben az esetben is egyértelmű legyen, hogy milyen változást okoz a kitakarás, és milyet a póz változása, a képeken csak a kitakarás

okozta különbséget jelenítem meg. Olyan módon, hogy a végső eredményből, a póz okozta távolságot kivonom.




(a) A távolságot minden képnél **saját magához képest** számítva.

(b) A távolságot minden képnél a **középső (py22) képhez képest** számítva. Az eredményből kivontam a teljes, kitakarás nélküli kép távolságát, hogy csak az eltérések látszódjanak. A kitakarás nélküli képek távolsága a referencia képhez képest a címekben olvasható.

15. ábra. A kitakarás több szögből megvizsgálva. A bemeneti képek forrása: Youtube

Az eredményekből leolvasható, hogy szemből az arcon elsősorban az orr takarására okozza a legnagyobb eltérést, szemben az emberi szemlélő számára fontos szem és szemöldök takarásával. Oldalról azonban már az arc oldalsó része és a szem is hangsúlyos.

Mindezek a kitakarási kísérletek elvégezhetőek a használandó módszerekkel is, így a maszkkal, napszemüveggel, sminkkel, és a szemöldök eltávolításával.

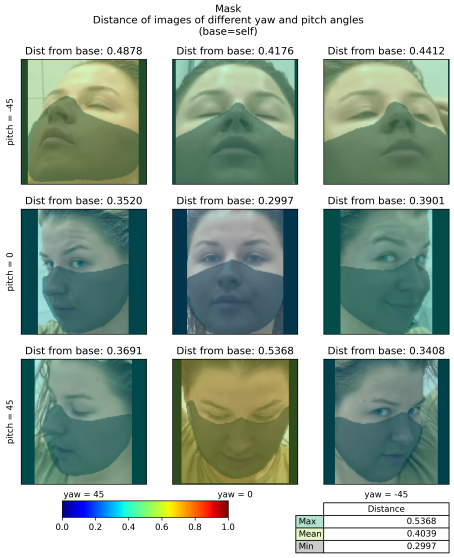
A kísérletekhez egy korábban nem használt adatbázist, a MR Face Similarity Dataset-et (MRFSD) használom, amely egy személyt tartalmaz különböző pózokban, smink nélkül, sminkben, maszkban és napszemüvegben is. A képek forrásául szolgáló videók Creative Commons BY 3.0 licensz alatt érhetőek el. Az A jelű függelékben minden képhez elérhető a forrás és a készítő

neve.

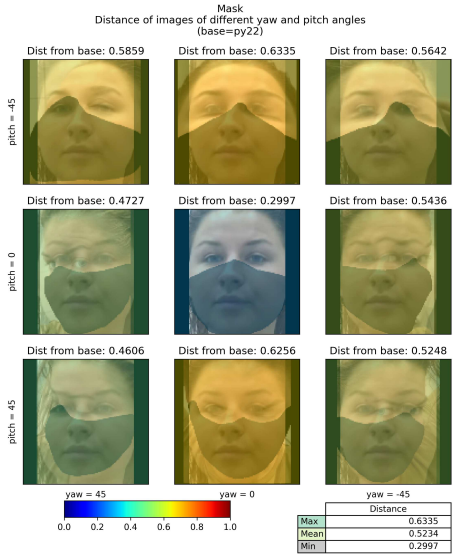
### 5.3.1. Maszk

A maszkot a kísérleti személy képeire digitálisan is rá lehet vetíteni, így biztosak lehetünk benne, hogy a változást mindössze a maszk okozza, és nem a fényviszonyok, póz és a kép minőségének változása. A maszkokat egyéb referenciák alapján, kézzel rajzolom fel Gimp segítségével.

Az eredményeken látható, hogy annak ellenére, hogy a maszk az arc nagy részét eltakarja, a két arc távolsága mégsem annyira nagy. Ez annak is köszönhető, hogy a képen minden más pixel változatlan, így a háttér is.




(a) A távolságot minden képnél **saját magához képest** számítva.



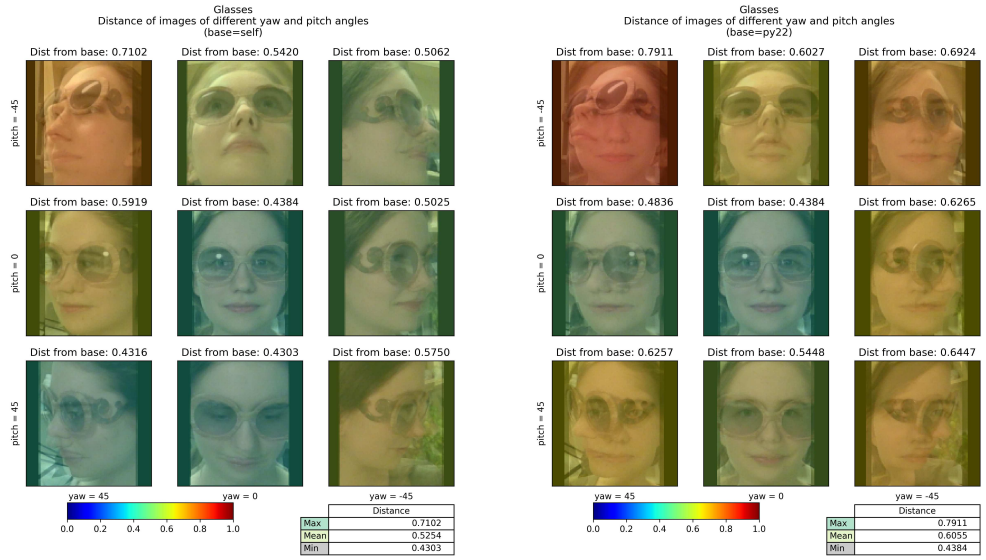

(b) A távolságot minden képnél a **középső (py22) képhez képest** számítva.

16. ábra. Egy digitálisan felhelyezett maszk több szögből megvizsgálva. A képek színe a távolság mértékét jelzi. A bemeneti képek forrása: Youtube

### 5.3.2. Napszemüveg

Megfelelő tesztadat és a napszemüveg digitális reprodukálásának nehézsége miatt ezt a tesztet valós környezetben tesztelem.

Az eredményeken látható, hogy a valós környezet ellenére sem haladja meg jelentősen a képek távolsága a megállapított határértéket.




(a) A távolságot minden képnél **saját magához képest** számítva.

(b) A távolságot minden képnél a **középső (py22) képhez képest** számítva.

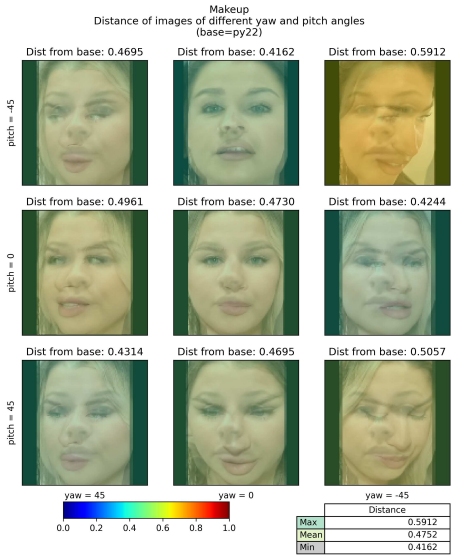
17. ábra. Napszemüveg viselete több szögből megvizsgálva. A képek színe a távolság mértékét jelzi.

### 5.3.3. Smink

A smink esetében különböző forrásból származó adatokat használok.

Az eredmények, az előzőekhez hasonlóan szintén, a megállapított határértékeknél alacsonyabb értékek. Érdemes azonban megfigyelni, hogy a referencia képhez (py22) hasonlított eredmények alacsonyabbak, mint a saját pózukhoz hasonlított képeknél. Ennek oka lehet az eltérő forrásokból adódó különségek is.


(a) A távolságot minden képnél **saját magához képest** számítva.




(b) A távolságot minden képnél a **középső (py22) képhez képest** számítva.

18. ábra. Smink viselete több szögből megvizsgálva. A képek színe a távolság mértékét jelzi. A bemeneti képek forrása: Youtube

### 5.3.4. Szemöldök eltávolítása

A szemöldököt digitálisan távolítottam el, így a különbségek biztosan csak ebből következnek.

Az eredményeken a referencia képhez képest látható nagy eltérés. Ezeknél a képeknél a referenciához képesti távolság a határértéket jóval meghaladja. Ezért ez a módszer alkalmasnak tűnik a valós környzetben való alkalmazáshoz.


(a) A távolságot minden képnél **saját magához képest** számítva.


(b) A távolságot minden képnél a **középső (py22) képhez képest** számítva.

19. ábra. Digitálisan eltávolított szemöldök több szögből megvizsgálva. A képek színe a távolság mértékét jelzi. A bemeneti képek forrása: Youtube

## 5.4. Valós környzeti teszt

Annak eldöntésére, hogy mely módszerek a legalkalmasabbak egy arcfelismerő rendszer általi azonosítás megakadályozására, valós környezetbeli teszteket is végre kell hajtani. Valós környezetben, a biztonsági kamerákat, vagy védett objektumok körül, vagy ha általános megfigyelés a cél, forgalmas helyeken helyezik el. Ilyen például egy utca, tér, vagy épületeken belül egy folyosó. A következő tesztekben az utóbbi lehetőséget vizsgálom. A tesztkörnyezet Nitzan Guetta és tsai. tanulmányához [10] hasonlóan egy körülbelül 6 méter hosszú folyosó, amire a kamera enyhén oldalról és 2 méter magasból lát rá. Mivel egy hosszabb folyosón tesztelés nem volt megoldott, ezért a felvett videót, egyszer teljes felbontásban, majd feleakkora felbontásban is megvizsgálom, ezzel szimulálva egy kétszer hosszabb folyosót. A fényt nagyjából 5000K színhőmérsékletű lámpák biztosítják.

A tesztalanyoknak ezen a folyosón kell elsétálniuk, normál arckifejezéssel, nem a kamerába nézve. A 2 férfi és 2 női tesztalanyok mind 20 év körüli életkorúak és világos bőrszínnel rendelkeznek és hozzájárulásukat adták a teszt elvégzéséhez.

| **azonosító** | **nem** | **életkor** |
|---|---|---|
| 1 | férfi | 22 |
| 2 | férfi | 22 |
| 3 | nő | 22 |
| 4 | nő | 23 |

Az első esetben mindenféle kiegészítő és smink nélkül sétálnak el a folyosón, majd külön maszkban és külön napszemüvegben. A smink felhelyezését és a szemöldök eltávolítását a róluk készült referencia képeken utómunkával végeztem el.

A tesztesetek ennek megfelelően a következőképpen alakulnak:

| teszteset neve | használt videófolyam | használt referencia |
|---|---|---|
| base | base | ref |
| mask | mask | ref |
| glasses | glasses | ref |
| makeup | base | ref_makeup |
| no-eb | base | ref_noeb |

Ahol:

- **base**: a referencia képpel teljesen megegyező arcot tartalmazó videófolyam,
- **mask**: az alany a ***base*** állapoton felül maszkot visel,
- **glasses**: az alany a ***base*** állapoton felül napszemüveget visel,
- **ref**: az igazolványképek kritériumainak megfelelő referencia kép,
- **ref_makeup**: a ***ref*** kép olyan digitálisan módosított változata amelyen az alany sminket visel,
- **ref_noeb**: a ***ref*** kép olyan digitálisan módosított változata amelyen az alanyról eltávolításra került a szemöldök

A 2-es számú alanynak világos szemöldöke van, aminek eltávolítása nem okoz szemmel látható változást, ezért esetében a szemöldököt besötétítettem.

A kamera egy iPhone 7 eszköz, 30fps@1920x1080 felbontással. A korábban említett tanulmányhoz való összehasonlíthatóság miatt, a felvételt utómunkával 10 képkocka/másodperc képfrissítési frekvenciára módosítottam és

bitsebességét 1280-1300 kbps értékre konvertáltam. Ezek az értékek jelentősen rontják a képminőséget, azonban a biztonsági kamerák minőségéhez jobban hasonlít ebben a formában.

Az arcdetektáló modell az MTCNN, az mtcnn Python csomagból.[2] A detektált arc minimum mérete 15x15 pixel.

Az arcfelismerő modell pedig a korábban említett LResNet100E-IR, ArcFace@ms1m-refine-v2.

Az arcot leíró vektorok a koszinusz távolságuk alapján kerülnek összehasonlításra. A határértékek pedig a Nitzan Guetta és tsai. tanulmányában használt 0,42, valamint az általam kiszámolt 0,5516 és 0,5947 értékek.

A rendszer adatbázisában szerepel az összes alany az LFW adatbázisból, valamint az éppen vizsgált tesztalany adatai, amit a róluk készült, az igazolványképek kritériumainak megfelelő referencia képből kerültek előállításra.

[2]https://github.com/ipazc/mtcnn

# 6. Eredmények

Az eredmények a korábban leírt módon lettek értékelve. Minden kipróbált módszerhez, az alanyok eredményeit nemek szerint együtt és külön-külön is összegeztem.

| result_type | success | | | top1 | top5 |
|---|---|---|---|---|---|
| határérték | 0,42 | 0,55 | 0,59 | - | - |
| base | 0,2741 | 0,5120 | **0,5377** | 0,6929 | 0,7372 |
| mask | 0 | 0 | **0,0128** | 0,2271 | 0,4097 |
| glasses | 0 | 0,1247 | **0,2315** | 0,5709 | 0,6957 |
| makeup | 0,2292 | 0,4883 | **0,5018** | 0,6315 | 0,7372 |
| noeb | 0,2240 | 0,3737 | **0,4555** | 0,6627 | 0,7822 |

5. táblázat. A tesztek összegzett eredményei, félkövérrel jelölve a legjobbakat minden kategóriában.

Egyértelműen látszik, hogy a legkisebb határértékkel lényegesen rosszabb eredményt kapunk, mint a magasabbakkal. Az általam megállapított 0,59 értékkel végzett vizsgálat, majdnem kétszer nagyobb sikerrátát ér el, mint a 0,42 határértékkel végzett. A továbbiakban, ezért a 0,59 értéket fogom használni.

| result_type | success | top1 | top5 |
|---|---|---|---|
| base | 0,6667 | 0,7292 | 0,8021 |
| mask | **0,0055** | **0,2155** | **0,3989** |
| glasses | 0,3152 | 0,7597 | 0,8572 |
| makeup | 0,6667 | 0,7292 | 0,8021 |
| noeb | 0,5938 | 0,7604 | 0,8542 |

6. táblázat. A férfi tesztalanyokhoz tartozó eredmények 0,59 határérték mellett, félkövérrel jelölve a legeredményesebb módszert.

| result_type | success | top1 | top5 |
|---|---|---|---|
| base | 0,4088 | 0,6567 | 0,6723 |
| mask | **0,0200** | **0,2387** | **0,4205** |
| glasses | 0,1478 | 0,3822 | 0,5341 |
| makeup | 0,3370 | 0,5338 | 0,6723 |

| result_type | success | top1 | top5 |
|---|---|---|---|
| noeb | 0,3172 | 0,5651 | 0,7103 |

7. táblázat. A női tesztalanyokhoz tartozó eredmények 0,59 határérték mellett, félkövérrel jelölve a legeredményesebb módszert.

A magasabb határérték azonban a rendszer támadását célzó módszerek hatékonyságát is jelentősen csökkenti. A kezdeti 53,77% sikerességi rátát a maszk használata 1,28%-ra csökkenti, ami megközelíti a Nitzan Guetta és tsai. által, az adversarial smink használatával elért 1,22%-os eredményt.

A többi módszer azonban csak kis mértékben csökkenti a felismerés sikerességét. A napszemüveg viselete esetén hiába okoz problémát egy emberi szemlélő számára a személy felismerése, a tesztelt előtanított modellnek ez mégsem okoz akkora problémát. 23,15% eredményével, a kezdeti sikerességet 43%-ára csökkenti.

Szintén nem váltotta be a hozzá fűzött reményeket a smink viselése és a szemöldök eltávolítása. Sorban 50,18% és 45,55% eredményükkel csak minimálisan sikerült a felismerést megakadályozni, ezért egyáltalán nem használhatóak valós környezetben.

Érdemes a sikeresség eredményeit összehasonlítani a Top1 és Top5 eredményekkel. A Top1 és Top5 eredmények azt mutatják, hogy az összes képkocka közül, hányon szerepelt a helyes címke a távolság szerint sorba rendezett találatok közül, sorban az első helyen és az első öt hely valamelyikén, tekintet nélkül a határértékre, aminél a távolságnak kisebbnek kellene lennie. Tehát ezek a számok azt mutatják meg, hogy a határérték megfelelő változtatásával, mi lenne az elméleti maximum eredmény amit el tudunk érni.

Ezek az eredmények között azt is láthatjuk, hogy a maszk használata, bár mindössze a képkockák 1%-ban teszi lehetővé a felismerést a határérték alkalmazásával, a Top1 és Top5 eredmények mégis sorban 22,71%, illetve 40,97%. Ez azt jelenti, hogy az arc nagy részét takaró maszk esetében is, a nem megerősített találatok (azaz a határértéket el nem érők) között is nagy valószínűséggel megjelenik a listában szereplő személy. Egy hosszabb videófolyamot, vagy több, különböző helyen készült videófolyamot összevetve, statisztikai módszerekkel kiszűrhető a többi találat közül a gyakrabban megjelenő személy.

A másik érdekesség pedig a szemöldök eltávolítása utáni Top5 eredmény, amely magasabb, mint az alap felismerés Top5 eredménye. Ez azt jelenti, hogy a szemöldök eltávolítása, ebben az esetben nemhogy csökkenti, de

egyenesen növeli a felismerés hatékonyságát.

A tesztalanyok adatait külön megvizsgálva az látjuk, hogy a 4. alany eredményei a megerősített (határérték alapján vizsgált) esetben a maszk és napszemüveg esetében 0%, a smink esetében 18,75%, a szemöldök eltávolítása esetében pedig 9,38%, ami jelentősen eltér a több tesztalany értékeitől. Azonban a Top1 és Top5 eredmények nem különböznek jelentősen, ami azt jelenti, hogy számára az eddigieknél is nagyobb határérték lenne szükséges.

| result_type | success | top1 | top5 |
|---|---|---|---|
| base | 0,25 | 0,7188 | 0,75 |
| mask | **0** | **0,2174** | **0,2609** |
| glasses | **0** | 0,3438 | 0,625 |
| makeup | 0,1875 | 0,5 | 0,75 |
| noeb | 0,0938 | 0,5625 | 0,8125 |

8. táblázat. A 4. alanyhoz tartozó teszteredmények

# 7. Összefoglalás

A dolgozatban kísérletet tettem arra, hogy olyan, speciális szaktudást nem igénylő módszerekkel akadályozzam meg egy arcfelismerő rendszer általi azonosítást, amelyek bárki számára hozzáférhető, hétköznapi tárgyakat és eszközöket igényelnek.

Megvizsgáltam, egy, az interneten szabadon elérhető előtanított modell pontosságát, amely az LFW adatbázison elvégzett teszt alapján, 99,82% eredményével alkalmasnak bizonyult a valós környezetbeli használatra.

Meghatároztam a tesztkörnyezetben, két arc azonosságát vizsgáló koszinusz távolságmetrikával, az LFW adatbázishoz tartozó megfelelő határértéket, amelyet 0,59 értékben állapítottam meg. Ezzel a határértékkel a valós környzeti tesztben 4 tesztalanyra 53,77%-os kezdeti eredményt értem el, a mindenféle segédeszköz használata nélküli felismerésben. A maszk, vagy napszemüveg viselete bizonyult a két legjobban teljesítő módszernek a felismerés megakadályozásában, sorban 1,28% és 23,15% eredményeikkel. Ezek olyan módszerek, amelyeket empirikusan is alkalmasnak találunk identitásunk elrejtésére, és ezt igazolják is az eredmények. Ezt követte a szemöldök eltávolításának, majd a smink használatának módszere, sorban 45,55% és 50,18% hatékonysággal. Ezek már egy emberi szemlélő számára sem jelentenek nagyobb kihívást, ez alapján nem meglepő a teszten elért - az arcfelismerő rendszer támadása szempontjából - rossz eredményük.

Megállapítottam viszont azt, hogy a használt határérték megfelelő megválasztásával, elméletileg jelentősen növelhető a hatékonyság. A maszk viselete mellett mindössze 1,28% eredményességgel teljesítő rendszer esetében is, a határérték használatának mellőzésével a Top1 és Top5 eredmények sorban 22,71% és 40,97% voltak, amelyek lehetővé teszik, hogy hoszabb videófolyamokat elemezve megállapíthassuk az azon szereplő személyek kilétét, annak ellenére, hogy a maszk az arc nagy részét eltakarja.

Ez különösen rossz hír, ha a megfigyelő rendszerek elleni védekezésként szeretnénk ilyen módszereket használni. A megállapított eredmények alapján, bármely a dolgozatban vizsgált módszert használva, ha hosszabb videófolyam áll rendelkezésre egy adott személyről, nagy pontossággal beazonosítható, vagy legalábbis közelítő javaslat adható a helyes identitásról, amelyet emberi ellenőrző már könyebben jóváhagyhat.

A dolgozat eredményei azonban további megerősítést igényelnek. A teszteket mindössze 4 alanyon végeztem el, akik mind hasonló korúak és világos

bőrűek. A kísérleteket érdemes lehet változatosabb tesztcsoporton is elvégezni, amely több korosztályt és bőrárnyalatot tartalmaz, hogy általánosabb következtetést lehessen levonni a teljes emberi populációra vonatkozóan.

A dolgozat nem tér ki az eredményeket befolyásoló pontos okokra sem, így nem vizsgálja, hogy mi az oka az egyes módszerek közötti különbségeknek. Továbbá nem vizsgálja azt az anomáliát sem, hogy a szemöldök eltávolítása, bár emberi szemlélők számára a felismerésben zavaró, az arcfelismerő rendszer esetében az elméleti maximum felismerési sikerességet - a céllal ellentétes hatást kiváltva - növeli.

# Köszönetnyilvánítás

Elsősorban szeretnék köszönetet mondani Dr. Szemenyei Mártonnak, aki konzultációival segítette a munkámat és felkeltette az érdeklődésemet a neurális hálózatok iránt.

Továbba szeretnék köszönetet mondani a tesztekben részt vevő embereknek: Bacskai Tamásnak, Kiss Máriának és Szepesi Szilvinek. Valamint a Paikka Beta - Hivatal csapatának, akikkel először kipróbálhattam egy arcfelismerő rendszert éles környezetben: Csököly Adél, Hegyi Tamás, Mile Na, Sanna Bo, Velegi Csaba, Xu Wang Dániel.

Az előzetes tesztekben használt MRFSD adatbázisért, pedig köszönet illeti Markó Rebekát, aki a videóival lehetővé tette a dolgozat elkészültét.

# Hivatkozások


[1] Gary B. Huang és tsai. *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*. Techn. jel. 07-49. University of Massachusetts, Amherst, 2007. okt.

[2] Florian Schroff, Dmitry Kalenichenko és James Philbin. „FaceNet: A unified embedding for face recognition and clustering". (2015. márc.). arXiv: 1503.03832 [cs.CV].

[3] Yaniv Taigman és tsai. „DeepFace: Closing the Gap to Human-Level Performance in Face Verification". *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014, 1701–1708. old. DOI: 10.1109/CVPR.2014.220.

[4] Margot E. Kaminski és Shane Witnov. „The Conforming Effect: First Amendment Implications of Surveillance, Beyond Chilling Speech". *University of Richmond Law Review* 49 (2015), 465–518. old.

[5] Evan Selinger és Brenda Leong. „The Ethics of Facial Recognition Technology". *SSRN Electronic Journal* (2021).

[6] Jules. Harvey Adam. LaPlace. *Exposing.ai*. 2021. URL: https://exposing.ai (elérés dátuma 2022. 10. 17.).

[7] Samer L. Hijazi, Rishi Kumar és Chris Rowen. „Using Convolutional Neural Networks for Image Recognition By". 2015.

[8] Kaiming He és tsai. „Deep Residual Learning for Image Recognition". *CoRR* abs/1512.03385 (2015). arXiv: 1512.03385. URL: http://arxiv.org/abs/1512.03385.

[9] Mahmood Sharif és tsai. „Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition". *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. CCS '16. Vienna, Austria: Association for Computing Machinery, 2016, 1528–1540. old. ISBN: 9781450341394. DOI: 10.1145/2976749.2978392. URL: https://doi.org/10.1145/2976749.2978392.

[10] Nitzan Guetta és tsai. „Dodging attack using carefully crafted natural makeup". (2021).

[11] Jiankang Deng és tsai. „ArcFace: Additive Angular Margin Loss for deep face recognition". (2018).

[12] Kaipeng Zhang és tsai. „Joint face detection and alignment using multi-task cascaded convolutional networks". (2016).

[13] Le Thanh Nguyen-Meidine és tsai. „A Comparison of CNN-based Face and Head Detectors for Real-Time Video Surveillance Applications". *CoRR* abs/1809.03336 (2018). arXiv: 1809.03336. URL: http://arxiv.org/abs/1809.03336.

[14] Sushma Niket Borade, Ratnadeep R. Deshmukh és Pukhraj Shrishrimal. „Effect of Distance Measures on the Performance of Face Recognition Using Principal Component Analysis". *Intelligent Systems Technologies and Applications*. Szerk. Stefano Berretti, Sabu M. Thampi és Praveen Ranjan Srivastava. Cham: Springer International Publishing, 2016, 569–577. old. ISBN: 978-3-319-23036-8.

[15] Pawan Sinha és tsai. „Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About". *Proceedings of the IEEE* 94.11 (2006), 1948–1962. old. DOI: 10.1109/JPROC.2006.884093.

[16] Balla József. *A biometrikus adatokat tartalmazó úti és személyazonosító okmányok biztonságnövelő hatása a határ- és közbiztonság alakulására.* 2014.

[17] Javid Sadr, Izzat Jarudi és Pawan Sinha. „The Role of Eyebrows in Face Recognition". *Perception* 32 (2003. febr.), 285–93. old. DOI: 10.1068/p5027.

[18] Vítor Albiero, Kai Zhang és Kevin W. Bowyer. „How Does Gender Balance In Training Data Affect Face Recognition Accuracy?": *CoRR* abs/2002.02934 (2020). arXiv: 2002.02934. URL: https://arxiv.org/abs/2002.02934.

# A. MRFSD adatbázis

A dolgozatban használt MRFSD adatbázis képeinek forrása. Az adatbázis letölthető a következő linken: mrfsd.zip

| filename | source_url | author | license | access_date | comment |
|---|---|---|---|---|---|
| clean_closedeyes1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| clean_htilted1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| clean1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| clean2.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses_closedeyes1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses_shadow1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses_tilted1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses_tilted2.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses1.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses2.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| glasses3.png | https://www.youtube.com/watch?v=2Uce_qZClJU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| illu1.png | https://www.youtube.com/watch?v=bOA0RCn867c | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| illu2.png | https://www.youtube.com/watch?v=bOA0RCn867c | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| illu3.png | https://www.youtube.com/watch?v=bOA0RCn867c | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| illu4.png | https://www.youtube.com/watch?v=bOA0RCn867c | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup_eyesclosed1.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup_pattern1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup_pattern2.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup1.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup2.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| lightmakeup3.png | https://www.youtube.com/watch?v=H-0MC1Aecyw | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| mask1.png | https://www.youtube.com/watch?v=alXVRge0W9A | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| mask2.png | https://www.youtube.com/watch?v=alXVRge0W9A | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| occlusion1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| sunglasses1.png | https://www.youtube.com/watch?v=alXVRge0W9A | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| tilted_lightmakeup1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| tilted1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| tilted2.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| threeeyes.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | modified |
| xmakeup1_1.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup1_2.png | https://www.youtube.com/watch?v=tTgdR1PQTH0 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup2_1.png | https://www.youtube.com/watch?v=mYBYvvbtQG8 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup2_2.png | https://www.youtube.com/watch?v=mYBYvvbtQG8 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup2_3.png | https://www.youtube.com/watch?v=mYBYvvbtQG8 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup2_4.png | https://www.youtube.com/watch?v=mYBYvvbtQG8 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| xmakeup2_5.png | https://www.youtube.com/watch?v=mYBYvvbtQG8 | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py11.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py12.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py13.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py21.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py22.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py23.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py31.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py32.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch/py33.png | https://www.youtube.com/watch?v=DTt-LN0-5EE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py11.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py12.png | https://www.youtube.com/watch?v=7avXNTpKMQE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py13.png | https://www.youtube.com/watch?v=7avXNTpKMQE | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py21.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py22.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py23.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py31.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py32.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |
| yawpitch_makeup/py33.png | https://www.youtube.com/watch?v=AFr2UlxyThU | Markó Rebeka | CC BY 3.0 | 2022.08.05. | |