Vizijski sustav za identifikaciju osobe temeljen na analizi karakterističnih točaka lica

Posavec, Karlo

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Karlo Posavec

Zagreb, 2022.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Mentor:

doc. dr. sc. Tomislav Stipančić, dipl. ing.

Student:

Karlo Posavec

Zagreb, 2022.

Ovime izjavljujem da sam samostalno izradio ovaj rad koristeći navedenu literaturu i znanja stečena tijekom studija.

Srdačno zahvaljujem svojem mentoru doc. dr. sc. Tomislavu Stipančiću na izdvojenom vremenu, pristupačnosti te pruženoj pomoći tijekom izrade ovog rada.

Posebnu zahvalu upućujem svojoj obitelji na strpljenju, razumijevanju i podršci tijekom cijelog mojeg obrazovanja.

Također zahvaljujem svojim prijateljima i svima koji su mi tijekom diplomskog dijela studija bili podrška.

Karlo Posavec

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE
Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za diplomske radove studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment,
inženjerstvo materijala te mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum:	Prilog:
Klasa: 602-14/22-6/1	
Ur. broj: 15-1703-22-	

DIPLOMSKI ZADATAK

Student: **KARLO POSAVEC** Mat. br.: 0035213827

Naslov rada na hrvatskom jeziku: **Vizijski sustav za identifikaciju osobe temeljen na analizi karakterističnih točaka lica**

Naslov rada na engleskom jeziku: **Visual system for person identification based on the analysis of characteristic points of the face**

Opis zadatka:

Računalna analiza i prepoznavanje omogućuju računalu da prepoznaje značajke na slikama te ih koristi kod različitih primjena u sklopu stvarne okoline. U tu svrhu se često koriste vizijski sustavi pomoću kojih se vrši akvizicija i učenje informacija o specifičnim točkama ili značajkama od interesa na licu osobe. Potom se te karakteristične točke koriste kod identifikacije osobe.
U radu je potrebno izraditi cjelovito softversko rješenje za identifikaciju osobe temeljem više karakterističnih točaka na licu, koje uključuje:

- pronalaženje i praćenje lica na analiziranoj slici bez obzira na trenutnu orijentaciju te perspektivu gledanja,
- analizu lica temeljem specifičnih karakteristika kao što su proporcije, smještaj obrva, razmak zjenica i slično,
- usporedbu dobivenih karakteristika s karakteristikama lica koja su ranije pohranjene u bazu znanih osoba,
- poklapanje rezultata te identifikacija osobe.

Uz razvijenu aplikaciju potrebno je ostvariti sustav evidencije koji bilježi vrijeme kada je osoba identificirana.
Dobiveno softversko rješenje je potrebno eksperimentalno evaluirati uključivši ljudske subjekte.
U radu je potrebno navesti korištenu literaturu te eventualno dobivenu pomoć.

Zadatak zadan:
5. svibnja 2022.

Rok predaje rada:
7. srpnja 2022.

Predviđeni datum obrane:
18. srpnja do 22. srpnja 2022.

Zadatak zadao:
doc. dr. sc. Tomislav Stipančić

Predsjednica Povjerenstva:
prof. dr. sc. Biserka Runje

SADRŽAJ

POPIS SLIKA

SAŽETAK

Za ljude najvažniji atribut za prepoznavanja nekog pojedinca je lice. Prepoznavanje lica pripada biometrijskim metodama poput prepoznavanja šarenice, prepoznavanja otiska prsta i prepoznavanja rasporeda vena. Biometrija postaje sve popularnija za korištenje kod identificiranja i provjere autentičnosti, zbog svoje visoke točnosti i jedinstvenosti za pojedinca, kao i praktičnosti stečene time što ne zahtijeva nikakav nametnuti teret kao što su tokeni ili zapamćene lozinke. Međutim, takvi načini identifikacije i provjere autentičnosti moraju biti zaštićeni od novih vektora napada. Kada se prepoznavanje lica koristi u svrhu provjere autentičnosti i identificiranja, jedan potencijalni vektor napada je lažno predstavljanje, odnosno predstavljanje nekog oblika lažne imitacije lica ovlaštenog korisnika. Uzimajući ovo u obzir unutar ovog rada razvijen je novi sustav evidencije pohađanja u realnom vremenu na temelju prepoznavanja lica s mogućnošću diferencijacije između pravog i lažnog lica. Razvijeno rješenje može služiti kao zamjena za tradicionalne sustave evidencije pohađanja u organizacijama i učionicama koje karakteriziraju dugo trajanje i nezgrapnost održavanja. Samo rješenje izrađeno je u programskom jeziku Python. Na kraju ovog rada provedena je evaluacija koja je uključivala ljudske subjekte.

Ključne riječi: identifikacija osobe, prepoznavanje lica, detekcija lica, duboko učenje, računalni vid, sustav evidencije pohađanja, zaštita od lažiranja

SUMMARY

For people, the most important attribute for recognizing an individual is the face. Face recognition belongs to biometric methods such as iris recognition, fingerprint recognition and vein pattern recognition. Biometrics are becoming increasingly popular for use in identification and authentication, due to their high accuracy and uniqueness for an individual, as well as the convenience gained by not requiring any imposed burden such as tokens or remembered passwords. However, such means of identification and authentication must be protected against new attack vectors. When facial recognition is used for authentication and identification, one potential attack vector is spoofing, that is, presenting some form of fake imitation of the face of an authorized user. Taking this into account, in this paper a new real-time attendance system based on face recognition with the ability to differentiate between real and fake faces has been developed. The developed solution can be used as a replacement for the traditional attendance systems in organizations and classrooms, which are characterized by a long duration and cumbersome maintenance. The solution was developed in the Python programming language. An evaluation involving human subjects was carried out at the end of this paper.

Key words: person identification, face recognition, face detection, deep learning, computer vision, attendance system, anti-spoofing

1. UVOD

Svatko je imao iskustvo da nije prepoznao poznatu osobu zbog promjena u pozi, izrazu lica, osvjetljenju i udaljenosti. Stoga ne čudi da se sustav računalnog vida suočava s istim problemima. Unatoč desetljećima rada na računalnom vidu, znanstvenici iz cijelog svijeta još nisu u stanju parirati ljudima. Ipak, sustavi za prepoznavanje lica nisu loši. Najbolji sustavi mogu nadmašiti ljudske performanse unutar fiksnih idealnih uvjeta. No performanse se im drastično smanjuju kako se uvjeti mijenjaju.

Sustav za prepoznavanje lica prvo zahtijeva skup slika unutar baze podataka. Nakon što se lice detektira, počinje glavni zadatak sustava za prepoznavanje lica da identificira poznato ili nepoznato lice i postupi u skladu s tim. Ljudi često miješaju prepoznavanje s pojmom detekcije lica, no prepoznavanje lica označava provjeru autentičnosti danih podataka o licu na temelju pohranjenih podataka o licu u bazi podataka. Nakon što se podaci o licu poklapaju s bazom podataka, sustav je autentificiran.

Sustavi za prepoznavanje lica uvelike su se razvili tijekom posljednjih nekoliko desetljeća. Zbog ovog razvoja dolazi do povećanja algoritamske složenosti što zahtijeva dulje vrijeme i veću računalnu snagu. Mnogi algoritmi kao što su analiza glavnih komponenti, analiza neovisnih komponenti, neizrazita logika i genetski algoritmi bili su korišteni kod sustava za prepoznavanje lica. Lice je višedimenzionalno i stoga ima „prokletstvo dimenzionalnosti" [1] odnosno lice zahtijeva puno memorije i vremena za obradu. Za prevladavanje ovog problema potrebno je odrediti optimalne značajke za poboljšanje točnosti i uklanjanje šuma sa slika.

Sustavi za prepoznavanje lica imaju široku primjenu. Učinkovito prepoznavanje lica može biti od velike koristi kod identifikacije osoba, forenzičkoj znanosti, zdravstvu, maloprodaji, sustavima za provjeru autentičnosti, podudaranju slika osumnjičenika i pristupu korisnika sigurnosnim sustavima [2].

1.1. Motivacija

Sve organizacije trebaju sustav evidencije pohađanja kako bi ručno ili automatski bilježile prisustvo svog osoblja. Svakodnevna evidencija dolazaka učenika na nastavu neophodna je za ocjenu uspješnosti i praćenje kvalitete nastave. Prozivanje imenima ili potpisivanje na papirima su metode koje se koriste u većini organizacija, a karakterizira ih da oduzimaju vrijeme i da su nesigurne. S druge strane, većina sustava za automatsku identifikaciju ljudi temelji se na tradicionalnim metodama kao što su otisci prstiju, lozinke i skeniranje osobnih iskaznica. Međutim, sve te metode imaju nekoliko ograničenja poput zaboravljanja lozinke ili gubitka osobne iskaznice. Stoga je najprikladnija metoda za osiguranje pune sigurnosti i spremanje povijesnih zapisa sustav za prepoznavanje lica. To je brzo rastuće područje i u novije vrijeme igra važnu ulogu u sigurnosti zbog svoje točnosti kod identifikacije i provjere ljudi.

Prepoznavanje lica pokazalo se kao produktivna metoda za evidenciju pohađanja nastave. Iako je vrlo uobičajeno shvaćanje da je prepoznavanje lica riješen problem, u stvarnosti još uvijek ima puno prostora za poboljšanje u kontekstu implementacije. Razvoj općenamjenskog sustava evidencije pohađanja temeljenog na prepoznavanju lica u potpunosti ovisi o dostupnosti algoritama za detekciju i prepoznavanje lica koji mogu održati ravnotežu između brzine, točnosti i sprječavanja lažiranja na gornjoj strani referentne vrijednosti. U nekim slučajevima korištenja brzina utječe na upotrebljivost, a u nekim drugim slučajevima točnost utječe na upotrebljivost. Stoga su u ovom radu implementirani modeli detekcije i prepoznavanja koji se mogu pozabaviti svim slučajevima upotrebe .

U teorijskom dijelu objašnjene su različite vrste biometrije, metode detekcije i prepoznavanja lica kao i vrste napada lažiranjem te načini njihovog sprječavanja. Na početku praktičnog dijela opisani su korišteni programski alati potrebni za izradu samog sustava. Nakon toga rad detaljno opisuje cjelokupni postupak funkcioniranja sustava evidencije pohađanja u realnom vremenu pomoću prepoznavanja lica uz istovremeno sprječavanje lažiranja lica. Na kraju je provedena eksperimentalna verifikacija koja je uključivala ljudske subjekte te je dan zaključak.

2. TEORIJSKE OSNOVE RADA

2.1. Biometrija

Evidencija pohađanja kao vrlo značajna i korisna aktivnost administracije može postati naporna i suvišna aktivnost koja rezultira netočnošću. Tradicionalni pristup prozivki pokazuje se kao zastarjela metoda jer kod velikog broja studenata po profesoru prozivanje imena traje dugo i vođenje evidencije je vrlo zamorno.

Svaka organizacija ima svoj način uzimanja evidencije pohađanja nastave. Neke organizacije koriste pristup orijentiran na popisivanje, a druge su implementirale digitalne metode poput bilježenja u Excel tablici i tehnike provlačenja iskaznice. Međutim, ove metode pokazuju se kao zastarjele jer uzrokuju dugo čekanje u redu. U slučaju da osoba zaboravi svoju identifikacijsku iskaznicu pohađanje joj neće moći biti zabilježeno.

Nove tehnologije koje su razvijane posljednjih desetljeća učinile su poboljšanja u mnogim područjima pa tako i u ovom. U današnje vrijeme općenito se koristi sustav evidencija pohađanja uz pomoć biometrije. Jedna od najkorištenijih biometrijskih metoda je prepoznavanje lica.

Prvo zabilježeno korištenje biometrije seže u drugo stoljeće prije Krista gdje je kineski car Ts'In She ovjeravao određene pečete otiskom prsta [3]. Biometrija je mjerenje i statistička analiza jedinstvenih fizičkih karakteristika i karakteristika ponašanja ljudi. Tehnologija se uglavnom koristi za identifikaciju i kontrolu pristupa ili za identifikaciju pojedinaca koji su pod nadzorom. Osnovna premisa biometrijske autentifikacije glasi: svaka osoba se može točno identificirati prema intrinzičnim fizičkim osobinama ili osobinama ponašanja. Pojam biometrija izveden je iz grčkih riječi bio, što znači život, i metric, što znači mjeriti [4]. Biometrijski uređaji uglavnom sadrže sljedeće komponente [4]:

- **čitač ili uređaj za skeniranje** - za bilježenje biometrijskog faktora koji se provjerava;
- **softver** - za pretvaranje skeniranih biometrijskih podataka u standardizirani digitalni format i za usporedbu točaka podudaranja promatranih podataka s pohranjenim podacima;
- **baza podataka** - za sigurno pohranjivanje biometrijskih podataka za usporedbu.

Upotreba biometrije ima puno prednosti i nedostatak u pogledu njezine sigurnosti, korištenja i drugih povezanih funkcija. Biometrija je korisna iz sljedećih razloga [4]:

- teško je lažirati ili ukrasti biometrijske podatke za razliku od lozinka;
- jednostavna i prikladna metoda za korištenje;
- biometrijski podatci su uglavnom isti tijekom života;
- neprenosiva metoda;
- učinkovita jer predlošci zauzimaju manje prostora za pohranu.

Nasuprot tome, nedostatci biometrije uključuju [4]:

- Zahtijeva visoke inicijalne troškove.
- Ako sustav ne uspije uhvatiti sve biometrijske podatke, to može dovesti do neuspjeha u identifikaciji korisnika.
- Baze podataka koje sadrže biometrijske podatke još uvijek mogu biti hakirane.
- Pogreške poput lažnog odbijanja i lažnog prihvaćanja još uvijek se mogu dogoditi.
- Ako se korisnik ozlijedi, tada biometrijski sustav autentifikacije možda neće raditi. Na primjer, ako korisnik opeče ruku, tada ga skener otiska prsta možda neće moći identificirati.

S razvojem novih i naprednijih algoritama te minijaturizacijom komponenata biometrija se počinje koristiti u sve više primjena. Najtipičniji slučajevi upotrebe biometrijskih tehnologija su: provedba zakona i javna sigurnost (identifikacija kriminalaca/osumnjičenika), vojska (identifikacija neprijatelja/saveznika), kontrola granica, putovanja i migracija (identifikacija putnika/migranta/putnika), civilna identifikacija (identifikacija građanina/stanovnika/birača), zdravstvena skrb i subvencije (identifikacija pacijenta/korisnika/zdravstvenog radnika), fizički i logički pristup (identifikacija vlasnika/korisnika/zaposlenika/izvođača/partnera) te komercijalne aplikacije (identifikacija potrošača/kupca) [3].

S godinama sve više razlikovnih ljudskih karakteristika je implementirano za identifikaciju te svrstano pod biometrijske tehnologije. Trenutne vrste biometrijskih tehnologija temeljene su na prepoznavanju [5]:

- uha;
- šarenice oka;

- mrežnice oka:
- vena sklere oka;
- lica;
- geometrije prsta;
- otiska prsta (uključujući otisak dlana);
- hoda;
- geometrije šake;
- otkucaja srca;
- načina tipkanja;
- mirisa osobe;
- potpisa;
- rasporeda vena;
- govornika.

2.1.1. Prepoznavanje otiska prsta

Sustavi za prepoznavanje otisaka prsta analiziraju lokacije „minucija" – završetaka i bifurkacija tarnih grebena na jastučiću prsta. Često se također koriste dodatne informacije, kao što je broj grebena između točaka minucije. Postoji niz metoda za hvatanje otisaka prstiju: optičko hvatanje koristi vidljivu svjetlost, kapacitivni senzori koriste električnu struju koja se provodi kroz prst, a ultrazvuk koristi visokofrekventne zvučne valove. Kontaktni senzori (Slika 2.1.) snimaju sliku uzorka grebena u kontaktu sa senzorom, dok beskontaktni senzori učinkovito snimaju sliku (2D ili 3D, ovisno o sustavu) uzorka grebena dok se prst drži na udaljenosti od senzora [6].

Prepoznavanje otiska prsta može pružiti koristan stupanj diskriminacije među prijavljenim korisnicima. Međutim, može biti teško izmjeriti apsolutnu razinu sigurnosti, osobito ako je sustav podvrgnut zlonamjernim lažnim napadima u izazovnom operativnom okruženju. Ostvarene stope pogrešaka ovisit će o nizu čimbenika, uključujući okruženje postavljenog sustava i sastav korisničke populacije. To može otežati određivanje potrebne

izvedbe. Pitanja prihvaćanja od strane korisnika ne bi se smjela zanemariti pri uvođenju bilo kojeg biometrijskog sustava. To naročito vrijedi za otiske prstiju jer se oni ponekad povezuju s provođenjem zakona i kriminalom, što može dovesti do sprječavanja njihovog prihvaćanja u nekim situacijama [6].

Slika 2.1. Prikaz otključavanja automobila kontaktnim senzorom otiska prsta [7]

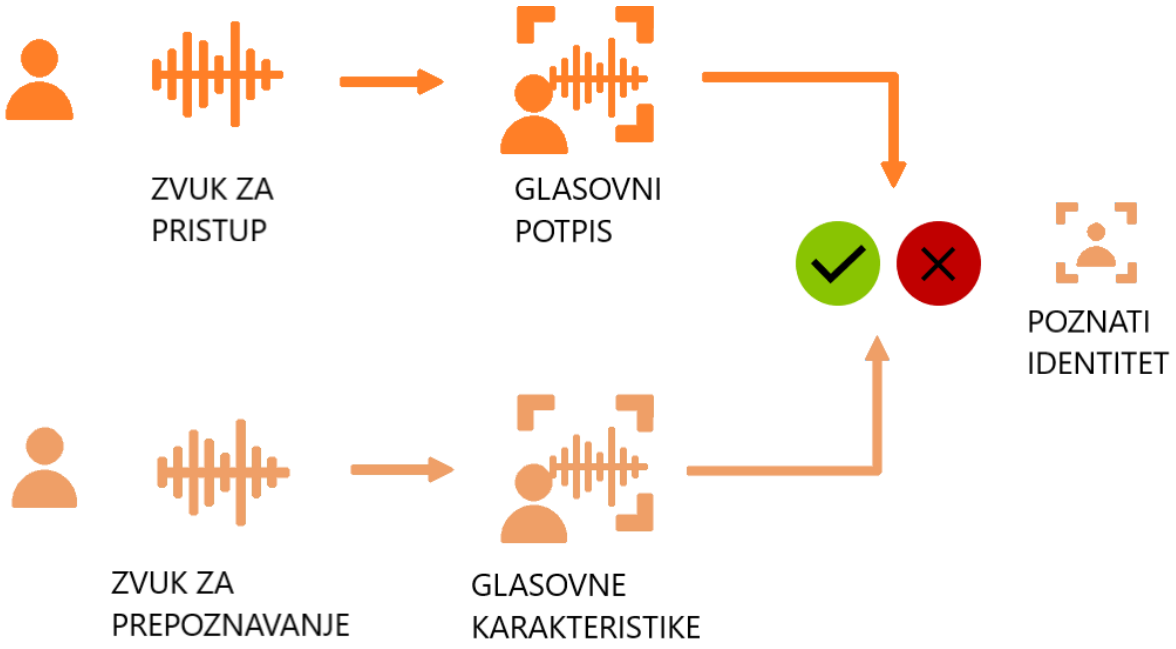
2.1.2. Prepoznavanje govornika

Prepoznavanje govornika temelji se na zvuku glasa. Prepoznavanje govornika ima dva oblika [6]:

1. **Ovisno o tekstu** – subjekt izgovara određenu lozinku.
2. **Neovisno o tekstu** – subjekt govori na nesputan način.

Kao i sve biometrijske metode, prepoznavanje govornika kombinira fizičke i bihevioralne komponente. Koristi se akustičnim karakteristikama govora pojedinca na koje utječu biološke značajke kao što su oblik vokalnog trakta i grkljana. Osim ovih fizičkih osobina, svaki je govornik razvio dodatne karakteristike, uključujući naglasak, ritam, intonaciju, izbor

vokabulara i tako dalje. Sustavi za prepoznavanje govornika uzorkuju te karakteristike i grade prepoznatljiv model za svakog 'poznatog' govornika. Na temelju tog principa (Slika 2.2.) se vrši identifikacija osobe. Prepoznavanje govornika ne treba brkati s povezanom nebiometrijskom tehnologijom prepoznavanja govora koja je korištena pri prepoznavanju riječi za diktiranje ili automatiziranje instrukcija koje se daju putem mobitela [6].



Slika 2.2. Prikaz identifikacije osobe na temelju prepoznavanja govornika

2.1.3. Prepoznavanje šarenice oka

Sustavi za prepoznavanje šarenice oka (Slika 2.3.) snimaju slike oka pomoću infracrvenog svjetla. Melanin šarenice oka proziran je pod infracrvenim osvjetljenjem, što omogućuje otkrivanje detalja šarenice oka bez obzira na boju očiju. Algoritmi za prepoznavanje lociraju granice šarenice i zatim obrađuju dio slike na kojem se nalazi šarenica oka kako bi pružili jasan i sažet prikaz uzorka šarenice oka pojedinca. Ovakav prikaz nudi vrlo visoku razinu razlikovanja između pojedinaca unutar populacije [6].

Prepoznavanje šarenice relativno je sigurna biometrijska metoda prepoznavanja ljudi. To je dijelom zbog visoke razine razlikovanja među populacijom koja se može postići.

Prepoznavanje šarenice može biti prilično otporno protiv prezentacijskih napada za sustave s odgovarajućom provjerom živosti, zbog poteškoća u dobivanju i predstavljanju potrebnih slika visoke kvalitete [6].

Slika 2.3. Prikaz upotrebe sustava za prepoznavanje šarenice oka [8]

2.1.4. Prepoznavanje rasporeda vena

Potkožne krvne žile ljudskog tijela čine poseban raspored za svaku osobu. Rasporedi vena mogu se uhvatiti osvjetljavanjem područja tijela infracrvenim svjetlom i fotografiranjem reflektiranog svjetla. U nekim sustavima, snimaju se fotografije infracrvenog svjetla koje se prenosi kroz tjelesno tkivo koje se snima. Krvne žile apsorbiraju infracrveno svjetlo više od okolnog tkiva i izgledaju tamnije na dobivenoj slici [6].

Proizvodi su dizajnirani za rad s dijelovima tijela koji se lako mogu prikazati senzoru: dlanovima (Slika 2.4.), prstima, zapešćima i nadlanicom. Biometrija vena čini obitelj sličnih modaliteta - tehnologije rasporeda vena na prstima, dlanovima i zapešćima razlikuju se jedna od druge što otežava opće izjave o njihovoj izvedbi. Unutarnja priroda venskih rasporeda otežava pasivno promatranje, predstavljajući određene izazove za napadače. Razine performansi za biometriju rasporeda vena dlana i prstiju izmjerene su kao dobre u ograničenom testiranju [6].

Slika 2.4. Prikaz upotrebe sustava za prepoznavanje rasporeda vena na dlanu [9]

2.2. Prepoznavanje lica

Prepoznavanje je jedna od najkorisnijih ljudskih funkcija vizualnog sustava. Čovjek može lako otkriti i prepoznati razne objekata na prvi pogled te bez dodirivanja, ali to je težak zadatak za računalo. Kako ljudi uče, organiziraju objekte i kategorije u korisne i informativne taksonomije i povezuju ih s jezikom. Repliciranje ove sposobnosti u strojevima koji nas okružuju bi duboko utjecalo na praktične aspekte naših života, uglavnom na bolje. Svakako, ovo je najuzbudljivija i najteža zagonetka s kojom se suočavaju znanstvenici i inženjeri u ovom desetljeću.

Prepoznavanje lica je potproblem vizualnog prepoznavanja uzoraka. Ljudi cijelo vrijeme prepoznaju vizualne obrasce, a vizualne informacije dobivaju kroz oči. Te informacije mozak prepoznaje kao smislene koncepte. Za računalo, bilo da se radi o slici ili videu, to je matrica koja se sastoji od mnogo piksela. Stroj bi trebao otkriti koji koncept predstavlja određeni dio podataka u podacima. Ovo je problem grube klasifikacije u vizualnom

prepoznavanju modela. Za prepoznavanje lica potrebno je razlučiti kome lice pripada u dijelu podataka koji svi strojevi smatraju licem.

Prepoznavanje lica uključuje detekciju lica, položaj lica, prepoznavanje identiteta, pretprocesiranje slike, itd. Cilj detekcije lica je pronaći koordinatni sustav svih lica na jednoj slici. To je postupak skeniranja cijele slike koji utvrđuje je li određeno područje lice. Položaj lica koordinatni je položaj značajke lica u koordinatnom sustavu detekcije lica. U usporedbi s detekcijom lica, vrijeme izračuna algoritma za pozicioniranje lica puno je kraće. Želi se pronaći idealna transformacijska funkcija kako bi se postigao optimalan učinak prepoznavanja, ali sam proces traženja je vrlo težak.

2.2.1. Povijesni pregled prepoznavanja lica

Povijest prepoznavanja lica može se pratiti do 1960-ih. Tada je matematičar i informatičar Woodrow Wilson Bledsoe prvi razvio sustav mjerenja koji se koristio za stavljanje fotografija lica u različite klasifikacije. Zbog ovog rada Bledsoe je poznat kao neslužbeni otac tehnologije prepoznavanja lica. Agencije za provođenje zakona ubrzo su se zainteresirale za Bledsoeov rad. A od 1970-ih do 1990-ih, agencije su razvile vlastite sustave za prepoznavanje lica. Oni su bili grubi u usporedbi s današnjom tehnologijom, ali rad na tim sustavima doveo je do modernih programa za prepoznavanje lica [10].

Mnogi ističu 2001. godinu kao ključnu godinu za tehnologiju prepoznavanja lica. Tada su službenici za provođenje zakona upotrijebili prepoznavanje lica kako bi lakše identificirali ljude u gomili na Super Bowlu XXXV. Iste godine, ured šerifa okruga Pinellas na Floridi stvorio je vlastitu bazu podataka za prepoznavanje lica. Međutim, tek 2010-ih godina računala su postala dovoljno moćna da prepoznavanje lica postanu standardnija značajka. Naime, 2011. godine softver za prepoznavanje lica potvrdio je identitet terorista Osame bin Ladena. Policijska uprava u Baltimoreu 2015. koristila je prepoznavanje lica za identifikaciju prosvjednika u prosvjedima nakon što je Freddie Gray poginuo od ozljede kralježnice koju je zadobio dok se prevozio u policijskom kombiju [10].

Potrošači sada koriste prepoznavanje lica na svojim pametnim telefonima i drugim osobnim uređajima. Windows Hello i Androidov Trusted Face 2015. omogućili su ljudima da se prijave na svoje uređaje jednostavnim usmjeravanjem uređaja prema svojem licu. Appleov iPhone X predstavio je svoju tehnologiju prepoznavanja lica Face ID 2017. godine. Bilo je kontroverzi oko ove tehnologije, a kritičari su govorili da predstavlja napad na privatnost. Gradovi poput San Francisca, Oaklanda i Bostona zabranili su vlastima korištenje prepoznavanja lica. Nakon Black Lives Matter prosvjeda protiv policijske brutalnosti u ljeto 2020., nekoliko tehnoloških divova, uključujući Amazon, Microsoft i IBM, objavilo je da više neće prodavati svoju tehnologiju prepoznavanja lica agencijama za provođenje zakona [10].

2.2.2. Izazovi, prednosti i nedostatci, sklonost predrasudama i greškama te primjeri upotrebe prepoznavanja lica

Iako se izgradnja sustava prepoznavanja lica čini lakim zadatkom, to nije tako jednostavno na slikama iz stvarnog svijeta koje se snimaju bez ikakvih ograničenja. Postoji nekoliko izazova s kojima se sustav za prepoznavanje lica suočava [11]:

- **Osvjetljenje -** Drastično mijenja izgled lica, uočeno je da male promjene u uvjetima osvjetljenja značajno utječu na njegove rezultate.
- **Poza glave -** Sustavi za prepoznavanje lica vrlo su osjetljivi na pozu glave, što može rezultirati neispravnim prepoznavanjem ili nikakvim prepoznavanjem.
- **Izrazi lica -** Različiti izrazi lica iste osobe još su jedan značajan faktor koji treba uzeti u obzir. Moderni sustavi prepoznavanja lica mogu se lako nositi s ovim izazovom.
- **Niska razlučivost** – Treniranje sustava prepoznavanja lica mora se provoditi na slici dobre rezolucije, inače model neće uspjeti izdvojiti značajke.
- **Starenje –** S godinama, karakteristike ljudskog lica kao što su linije i tekstura se mijenjaju što predstavlja još jedan dodatan izazov.

Kako bi se doskočilo tim izazovima koristi se tehnika jednokratnog učenja. Jednokratno učenje je zadatak klasifikacije gdje se preko jednog, ili nekoliko, primjera vrši klasificiranje mnogih novih primjeri u budućnosti. Ovo karakterizira zadatke viđene u području prepoznavanja lica, kao što su identifikacija i autentifikacija lica, gdje ljudi moraju biti ispravno klasificirani s različitim izrazima lica, uvjetima osvjetljenja, naočalama i frizurama s obzirom

na jedan ili nekoliko predložaka fotografija. Moderni sustavi za prepoznavanje lica pristupaju problemu jednokratnog učenja za prepoznavanje lica putem učenja bogatog niskodimenzionalnog prikaza značajki, nazvanog embedding lica, koji za lica može biti lako izračunat i uspoređen kod zadataka autentifikacije i identifikacije [12].

Kao i sve druge tehnologije prepoznavanje lice ima i prednosti i nedostatke. Osim otključavanja pametnog telefona, prepoznavanje lica donosi i druge prednosti [13]:

- **Povećana sigurnost** - Na državnoj razini, prepoznavanje lica može pomoći u identificiranju terorista ili drugih kriminalaca. Na osobnoj razini, prepoznavanje lica može se koristiti kao sigurnosni alat za zaključavanje osobnih uređaja i za kamere za osobni nadzor.
- **Smanjen kriminal** - Prepoznavanje lica olakšava ulaženje u trag provalnicima, lopovima i uljezima. Sama spoznaja o prisutnosti sustava za prepoznavanje lica može poslužiti kao odvraćanje, posebno od sitnog kriminala. Osim fizičke sigurnosti, postoje i prednosti kibernetičke sigurnosti. Tvrtke mogu koristiti tehnologiju prepoznavanja lica kao zamjenu za lozinke za pristup računalima.
- **Uklanjanje pristranosti kod zaustavljanja i pretraživanja** - Zabrinutost javnosti zbog neopravdanih zaustavljanja i pretraga izvor je kontroverzi za policiju, tehnologija prepoznavanja lica mogla bi poboljšati proces. Upotrebom automatiziranog, a ne ljudskog procesa, tehnologija prepoznavanja lica mogla bi pomoći u smanjenju potencijalne pristranosti te smanjiti zaustavljanja i pretrage građana koji poštuju zakon.
- **Veća praktičnost** - Kako tehnologija postaje sve raširenija, kupci će moći plaćati u trgovinama svojim licem, umjesto da izvlače svoje kreditne kartice ili gotovinu. To bi moglo uštedjeti vrijeme u redovima na blagajni. Za prepoznavanje lica nije potreban kontakt kao što je slučaj s otiscima prstiju ili drugim sigurnosnim mjerama što je osobito korisni u svijetu nakon COVID-a.
- **Brža obrada** - Prepoznavanje lica omogućuje brzu i učinkovitu provjeru identiteta osobe. U eri kibernetičkih napada i naprednih alata za hakiranje, tvrtkama su potrebne sigurne i brze tehnologije.
- **Integracija s drugim tehnologijama** - Većina rješenja za prepoznavanje lica kompatibilna je s većinom sigurnosnih softvera. Zapravo, lako se integrira i s drugim biometrijskim tehnologijama.

Dok neki ljudi nemaju ništa protiv snimanja u javnosti i ne protive se korištenju prepoznavanja lica tamo gdje postoji jasna korist ili razlog, tehnologija može potaknuti intenzivne reakcije kod drugih. Neki od nedostataka ili problema uključuju [13]:

- **Nadzor** - Neki se brinu da korištenje prepoznavanja lica u kombinaciji sa sveprisutnim nadzornim kamerama i umjetnom inteligencijom stvara potencijal za masovni nadzor, što bi moglo ograničiti slobodu pojedinca. Dok tehnologija prepoznavanja lica omogućuje vladama da uđu u trag kriminalcima, također bi im mogla omogućiti da u svakom trenutku uđu u trag običnim i nedužnim ljudima.
- **Prostor za pogreške** - Podaci o prepoznavanju lica nije oslobođena grešaka, što bi moglo dovesti do toga da ljudi budu upleteni u zločine koje nisu počinili.
- **Povreda privatnosti** - Pitanje etike i privatnosti je najspornije. Poznato je da vlade pohranjuju slike građana bez njihova pristanka. Godine 2020. Europska komisija rekla je da razmatra zabranu tehnologije prepoznavanja lica u javnim prostorima do pet godina, kako bi se dalo vremena za izradu regulatornog okvira za sprječavanje zloupotreba privatnosti i etičkih zloupotreba.
- **Ogromna pohrana podataka** - Softver za prepoznavanje lica oslanja se na tehnologiju strojnog učenja, koja zahtijeva ogromne skupove podataka za „trening" kako bi dala točne rezultate. Tako veliki skupovi podataka zahtijevaju robusnu pohranu podataka. Mala i srednja poduzeća možda nemaju dovoljno resursa za pohranjivanje potrebnih podataka.

Kao što je iznad navedeno prepoznavanje lica može biti sklono pogreškama, što može implicirati ljude za zločine koje nisu počinili. Softver za prepoznavanje lica posebno je loš u prepoznavanju žena, djece, Afroamerikanaca i drugih etničkih manjina tako da ih često krivo identificira. Kada je riječ o pogreškama, postoje dva ključna pojma koja treba razumjeti. „Lažno negativno" je kada sustav za prepoznavanje lica ne uspije spojiti lice osobe sa slikom koja se zapravo nalazi u bazi podataka. Drugim riječima, sustav će greškom vratiti nula rezultata kao odgovor na upit. „Lažno pozitivno" je kada sustav za prepoznavanje lica podudara lice osobe sa slikom u bazi podataka, ali to podudaranje zapravo nije točno. To je kada policajac pošalje sliku „Joea", ali sustav pogrešno kaže policajcu da je na fotografiji „Jack" [14].

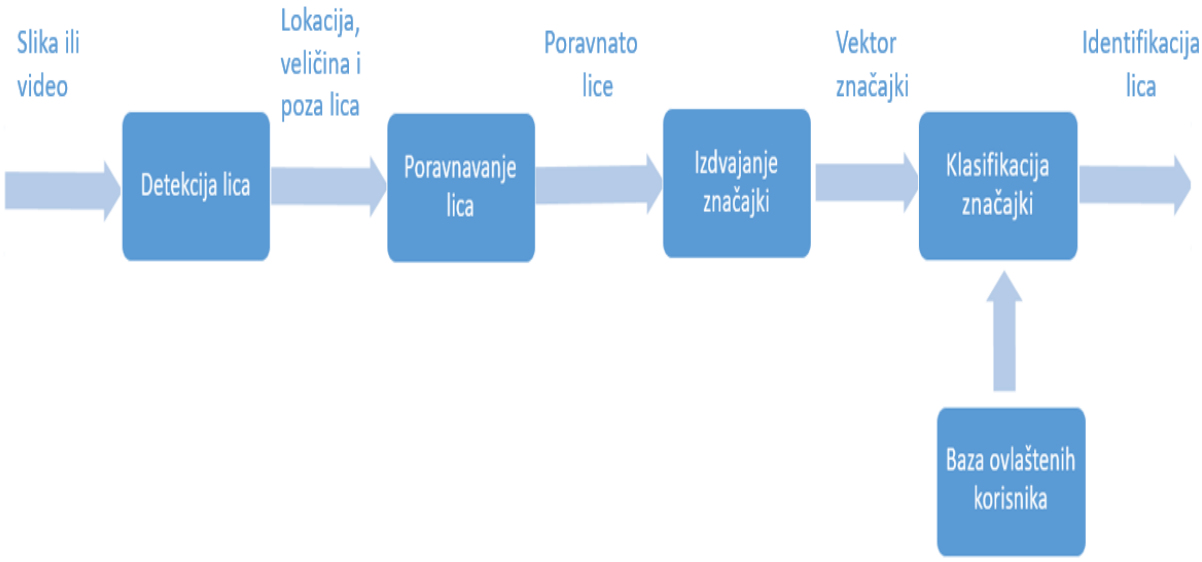
Kod istraživanja sustav za prepoznavanje lica, važno je pomno promatrati stopu „lažno pozitivnih" i „lažno negativnih" postotaka, jer gotovo uvijek postoji kompromis. Na primjer, kod korištenja prepoznavanja lica za otključavanje mobitela, bolje je ako sustav ne uspije identificirati vlasnika mobitela nekoliko puta (lažno negativno) nego da sustav pogrešno identificira druge osobe kao vlasnika mobitela i dopusti tim osobama da otključaju njegov mobitel (lažno pozitivno). Ako je rezultat pogrešne identifikacije da nevina osoba ode u zatvor (poput pogrešne identifikacije u bazi podataka), tada bi sustav trebao biti dizajniran tako da ima što je moguće manje lažno pozitivnih rezultata. [14].

Prepoznavanje lica koristi se od strane raznih ljudi i organizacija te na mnogo različitih lokacija. Reprezentativni primjeri organizacija su: vlade država, proizvođači mobilnih telefona, fakulteti i škole, vjerske skupine; društvene mreže, zrakoplovne luke, trgovci na malo te marketinški stručnjaci i oglašivači [10].

2.3. Koraci sustava prepoznavanje lica

Sustavi prepoznavanja lica su kroz svoj značajan razvoj u posljednjem desetljeću kao jednu od rijetkih nepromijenjenih stvari zadržali to da se proces sastoji od četiri koraka (Slika 2.5.). Četiri koraka sustava prepoznavanja lica su [15]:

1. **Detekcija lica** - Metoda detekcije lica koristi se za pronalaženje lica prisutnih na danoj slici, izdvajanje lica ako postoje i obrezivanje lica samo za stvaranje komprimirane datoteke za daljnje izdvajanje značajki. Postoji više algoritama za izvođenje ovog zadatka u sustavu za otkrivanje/prepoznavanje lica.
2. **Poravnanje lica** - Poravnavanje lica je rana faza koja kroz razne translacije i rotacije poravnava lica kako bi se poboljšala točnost samog sustava prepoznavanja lica.
3. **Izdvajanje značajki** - Izdvajanje značajki je osnovni i najvažniji inicijalizacijski korak prepoznavanja lica. Izvlači biološke komponente lica promatrane osobe. Biološke komponente su crte lica koje razlikuju jednu osobu od druge. Postoje različite metode koje izdvajaju različite kombinacije značajki, poznatih kao čvorne točke.
4. **Klasifikacija značajki** - Završna faza tehnologije prepoznavanja lica je donošenje odluke podudaraju li se značajke lica novog uzorka s onima iz baze podataka lica ili ne. Ove klasifikacije temeljene na predlošku moguće su korištenjem različitih statističkih pristupa. Obično traje samo nekoliko sekundi.



Slika 2.5. Prikaz koraka sustava prepoznavanja lica

2.4. Detekcija lica

Detekcija lica je računalna tehnologija temeljena na umjetnoj inteligenciji (AI) koja se koristi za pronalaženje ljudskih lica na digitalnim slikama. Detekcija lica igra važnu ulogu kao prvi korak u mnogim važnim primjenama uključujući praćenje lica, analizu lica i prepoznavanje lica. Detekcija lica ima značajan utjecaj na to kako dobro će se sekvencijalne operacije izvoditi [16].

Detekcija lica koriste algoritme i strojno učenje za pronalaženje ljudskih lica unutar većih slika, koje često uključuju druge objekte koji nisu lica kao što su pejzaži, zgrade i drugi dijelovi ljudskog tijela poput stopala ili ruku. Algoritmi za detekciju lica obično počinju traženjem ljudskih očiju, jedne od značajki koje je najlakše otkriti. Algoritam nakon toga pokušava otkriti obrve, usta, nos, nosnice i šarenicu. Nakon što algoritam zaključi da je pronašao područje lica, primjenjuje dodatne testove kako bi potvrdio da je zapravo detektirao lice. Kako bi se osigurala točnost, algoritme je potrebno uvježbati na velikim skupovima podataka koji uključuju stotine tisuća pozitivnih i negativnih slika. Treningom se poboljšava sposobnost algoritama da utvrde ima li lica na slici i gdje se nalaze [16].

Metode detekcije lica mogu se podijeliti u četiri kategorije, a algoritmi detekcije lica mogu istovremeno pripadati u dvije ili više kategorija. Kategorije su kako slijedi [17]:

1. **Temeljene na znanju** - Metode temeljene na znanju ovise o skupu pravila, a temelje se na ljudskom znanju za otkrivanje lica. Ta pravila su da lice mora imati nos, oči i usta unutar određenih međusobnih udaljenosti i položaja. Veliki problem s ovim metodama je poteškoća u izgradnji odgovarajućeg skupa pravila. Metode bi mogle rezultirati s mnogo lažnih pozitiva ako su pravila preopćenita ili previše uska. Samostalno ovaj pristup nije dovoljan te se zato uglavnom kombinira s drugim metodama.
2. **Temeljene na značajkama** - Metode temeljene na značajkama temelje se na lociranju lica izdvajanjem strukturnih značajki lica. Prvo se treniraju kao klasifikator, a zatim se koriste za razlikovanje regija koje sadrže i ne sadrže lice. Ideja je prevladati granice ljudskog instinktivnog znanja o licima. Ovaj pristup podijeljen je u nekoliko koraka, pa čak i fotografije koje sadrže mnogo lica pokazuju stopu uspjeha od 94%.
3. **Poklapanja predložaka** - Metode poklapanja predložaka koriste unaprijed definirane ili parametrizirane predloške lica za lociranje ili detektiranje lica korelacijom između predložaka i ulaznih slika. Predlošci kod ljudskog lica mogu biti oči, konture lica, nos i usta. Također, model lica može se izgraditi pomoću rubova koristeći samo metodu detekcije rubova. Ovaj pristup je jednostavan za implementaciju, ali je neadekvatan za pouzdanu detekciju lica. Međutim, predloženi su deformabilni predlošci za rješavanje tih problema.
4. **Temeljene na izgledu** - Metode temeljene na izgledu ovise o skupu provjerenih slika lica za trening kako bi se pronašli modeli lica. Pristup temeljen na izgledu bolji je od drugih načina izvedbe. Metode temeljene na izgledu koriste tehnike statističke analize i strojnog učenja za pronalaženje relevantnih karakteristika slika lica. Modeli temeljeni na izgledu mogu se podijeliti na daljnje metode.

2.4.1. Viola – Jones algoritam

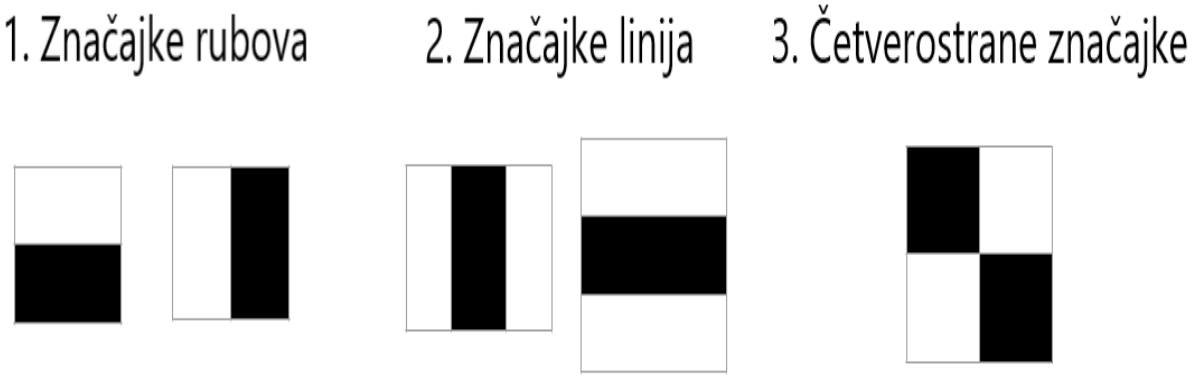
Algoritam Viola Jones nazvan je po istraživačima Paulu Violi i Michaelu Jonesu koji su predložili metodu 2001. godine u svom radu „Rapid Object Detection using a Boosted Cascade of Simple Features". Unatoč zastarjelom okviru, Viola-Jones je prilično moćan, a njegova se primjena pokazala iznimnom u detekciji lica u stvarnom vremenu. Ovaj algoritam spor je za treniranje, ali može otkriti lica u stvarnom vremenu impresivnom brzinom [18].

S obzirom na sliku (ovaj algoritam radi sa slikama u sivim tonovima), algoritam gleda mnoge manje podregije i pokušava pronaći lice tražeći specifične značajke u svakoj podregiji. Treba provjeriti razne različite položaje i povećanja jer slika može sadržavati više lica različitih veličina. Viola i Jones algoritam koristi Haarove značajke za otkrivanje lica. Algoritam Viola Jones ima četiri glavna koraka, koji će biti detaljnije objašnjeni u nastavku [18]:

1. Odabir Harrovih značajki
2. Stvaranje integralne slike
3. Pokretanje AdaBoost treniranja
4. Stvaranje kaskada klasifikatora

Haarove značajke su značajke digitalne slike koje se koriste u prepoznavanju objekata. Sva ljudska lica dijele neka univerzalna svojstva ljudskog lica kao što su svojstva da je područje očiju tamnije od susjednih piksela, a područje nosa svjetlije od područja očiju. Jednostavan način da se sazna koja je regija svjetlija ili tamnija je da se zbroje vrijednosti piksela obje regije i usporede se. Zbroj vrijednosti piksela u tamnijem području bit će manji od zbroja piksela u svjetlijem području. Ako je jedna strana svjetlija od druge, to može biti rub obrve ili ponekad središnji dio može biti sjajniji od okolnih okvira, što se može protumačiti kao nos. To sve se postiže pomoću Haarovih značajki te se pomoću njih još mogu protumačiti različiti dijelovi lica. Postoje 3 vrste Haarovih značajki (Slika 2.6.) koje su Viola i Jones identificirali u svom istraživanju [18]:

1. Značajke rubova
2. Značajke linija
3. Četverostrane značajke



Slika 2.6. Prikaz Haarovih značajki

Značajke rubova i značajke linija korisne su za otkrivanje rubova odnosno linija. Četverostrane značajke koriste se za pronalaženje dijagonalnih značajki. Vrijednost značajke izračunava se kao jedan broj tj. kao zbroj vrijednosti piksela u crnom području minus zbroj vrijednosti piksela u bijelom području. Vrijednost je nula za ravnu površinu u kojoj svi pikseli imaju istu vrijednost i stoga ne pružaju korisne informacije. Budući da su ljudska lica složenih oblika s tamnijim i svjetlijim područjima, Haarove značajke daju veliki broj kada su područja u crnim i bijelim pravokutnicima vrlo različita [18].

Kao što je u prethodnom odlomku navedeno za izračunavanje vrijednosti za svaku značajku moraju se provesti izračuni na svim pikselima unutar te posebne značajke. U stvarnosti, ovi izračuni mogu biti vrlo računalno intenzivni budući da i broj piksela s povećanjem elemenata brzo raste. Integralna slika igra svoju ulogu u omogućavanju brzog izvršavanja ovih intenzivnih izračuna kako bi se moglo razumjeti odgovara li značajka kriterijima. Integralna slika (također poznata kao tablica sa zbrojenom površinom) naziv je i strukture podataka i algoritma koji se koristi za dobivanje te strukture podataka. Koristi se kao brz i učinkovit način za izračunavanja zbroja vrijednosti piksela na slici ili pravokutnom dijelu slike [18].

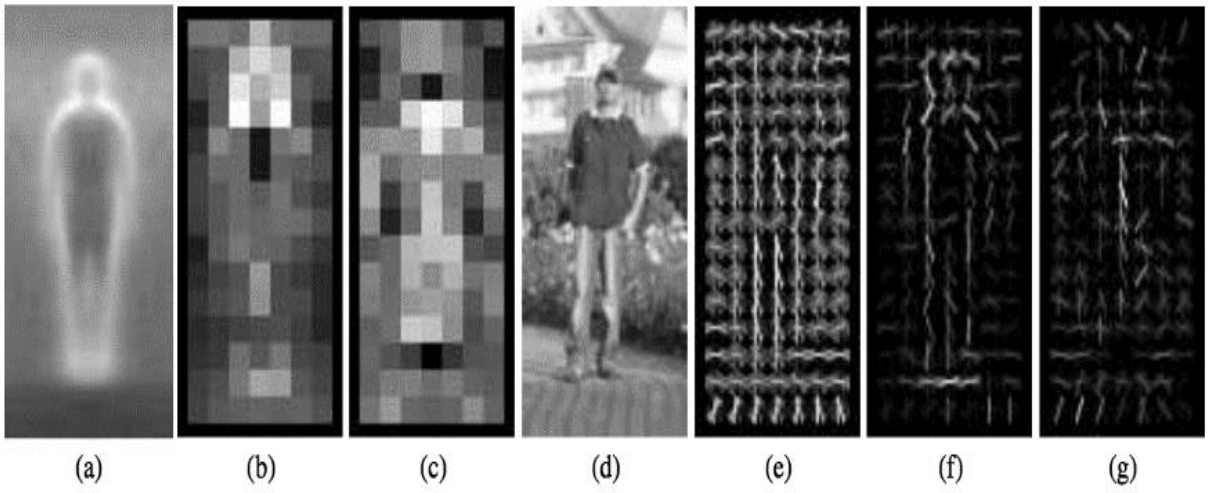
U sljedećem koraku koristi se algoritam strojnog učenja poznat kao AdaBoost. Broj značajki koje su prisutne u prozoru detektora 24×24 je gotovo 160 000, ali samo je nekoliko od tih značajki važno za identifikaciju lica. Tako se AdaBoost algoritam koristi za identifikaciju najboljih značajki u 160 000 značajki. U Viola-Jones algoritmu, svaka Haarova značajka predstavlja slabog učenika. Kako bi odredio vrstu i veličinu značajke koja ulazi u konačni klasifikator, AdaBoost provjerava izvedbu svih isporučenih klasifikatora. Kako bi se izračunala izvedba klasifikatora, ona se evaluira na svim podregijama svih slika korištenih za trening. Neke podregije proizvest će snažan odgovor u klasifikatoru. One će biti klasificirane kao pozitivne, što znači da klasifikator misli da sadrže ljudsko lice. Podregije koje ne daju snažan odgovor prema mišljenju klasifikatora ne sadrže ljudsko lice. One će biti klasificirane kao negativne. Klasifikatorima koji su se pokazali dobrim daje se veća važnost ili težina. Konačni rezultat je jaki klasifikator, koji se naziva i pojačani klasifikator, koji sadrži slabe klasifikatore s najboljom izvedbom. U konačnici algoritam postavlja minimalni prag kako bi odredio može li se nešto klasificirati kao korisna značajka ili ne [18].

Možda će AdaBoost odabrati najbolje značajke oko petine, ali izračunavanja tih značajki za svaku regiju je još uvijek dugotrajan proces. Prozorom veličine 24×24 prolazi se preko ulazne slike i mora se pronaći sadrži li neka od tih regija lice. Zadatak kaskade je brzo odbacivanje ne-lica i izbjegavanje gubitka vremena. Time se postiže brzina potrebna za detekciju lica u stvarnom vremenu. Kaskadni sustav je postavljen tako da se proces identifikacije lica dijeli na više faza. U prvoj fazi nalazi se klasifikator koji se sastoji od najboljih značajki, drugim riječima, u prvoj fazi podregija prolazi kroz najbolje značajke kao što je značajka koja identificira hrbat nosa ili ona koja identificira oči. U sljedećim fazama nalaze se sve preostale značajke. Kada podregija slike uđe u kaskadu, evaluira se prvom fazom. Ako ta faza ocijeni subregiju kao pozitivnu, što znači da misli da je lice, rezultat faze je možda. Kada podregija dobije možda, šalje se na sljedeću fazu kaskade i proces se kao takav nastavlja sve dok se ne dođe do posljednje faze. Ako svi klasifikatori odobre sliku, ona se konačno klasificira kao ljudsko lice i prikazuje se korisniku kao detekcija. Ukratko, ako prva faza daje negativnu ocjenu, slika se odmah odbacuje jer ne sadrži ljudsko lice. Ako prođe prvu fazu, ali ne prođe drugu fazu, također se odbacuje. U osnovi, slika se može odbaciti u bilo kojoj fazi klasifikatora [18].

2.4.2. Histogram orijentiranih gradijenata za detekciju ljudi

Ovo je naširoko korišten model detekcije lica, temeljen na HoG značajkama i SVM-u objavljenim 2005. u radu „Histograms of oriented gradients for human detection". HOG ili histogram orijentiranih gradijenata deskriptor je značajki koji se često koristi za izdvajanje značajki iz slikovnih podataka (Slika 2.7.). To je najbrža metoda koja se može koristiti na procesoru te može raditi na frontalnim i ne-frontalnim slikama. Ali nije u stanju detektirati male slike i rukovati okluzijama. Također, često isključuje neke dijelove brade i čela tijekom detekcije [15].

Gradijent slike u osnovi je promjena vrijednosti piksela u x i y smjerovima slike. Matematički, to je linearni zbroj x i y derivacija slike. Ova vrsta tehnike čini kamen temeljac obrade slike i računalnog vida te se najčešće koristi u filtrima detektora rubova. Gradijenti matematički označavaju promjenu, stoga će gradijent slike označavati promjenu slike u određenom smjeru. Stoga svaki vektor gradijenta ima dvije informacije vezane uz njega: njegovu veličinu i smjer [19].



Slika 2.7. Prikaz HOG značajki ulazne slike (d) [19]

Izvorni deskriptor histogram orijentiranih gradijenata (HOG) dizajniran je za detekciju pješaka. Skup podataka korišten u ovom radu izvorni je MIT-ov skup podataka o pješacima koji ima više od 1800 označenih slika ljudi s više od 100 različitih poza i pozadina. Zadatak rada bio je detektirati pješake s velikom točnošću. Osnovna ideja je izdvojiti sve relevantne značajke oblika pješaka i upotrijebiti ih za trening klasifikatora, značajka slike dio je informacije koja je relevantna za zadatak klasifikacije i može uključivati točke, rubove ili čak objekte. Značajke mogu biti niske razine i mogu se grupirati zajedno kako bi opisale objekte više razine na slici. Nakon izdvajanja ovih značajki oblika, binarni klasifikator može se trenirati za klasificiranje prikazuje li dani skup značajki pješaka ili ne. Ovo je dio koji se odnosi na izdvajanje značajki. Da bi se znalo koje značajke treba odabrati sa slike za treniranje klasifikatora, prvo se moraju opisati dvije vrste značajki na slici. Na slici postoje dvije vrste značajki: lokalne i globalne. Globalne značajke opisuju sliku ili objekt u cjelini i koriste se za generalizaciju cijele slike. To uključuje prikaze kontura, deskriptore oblika ili značajke teksture. Nasuprot tome lokalni deskriptori, opisuju dijelove slike kao što su tekstura, kutovi itd. Deskriptori poput SURF, FREAK ili BRISK popularni su među lokalnim deskriptorima, dok je HOG globalni deskriptor [19].

Temeljna ideja je karakterizirati oblik i izgled lokalnog objekta distribucijom gradijenata lokalnog intenziteta ili smjerova rubova jer to rezultira dobrim rješenjem, čak i u onim slučajevima kad nema detaljnog znanja o odgovarajućim položajima rubova ili gradijenta. Uzimajući sliku, HOG koristi klizni prozor veličine 64x128 za otkrivanje pješaka. Kako bi izračunao skup globalnih značajki određenog pješaka, HOG radi na pronalaženju gradijenata

lokalnog intenziteta radeći na ćelijama 8x8. HOG sada izračunava vektor gradijenta bez predznaka u svakoj od tih 8x8 ćelija generirajući vektor gradijenta veličine 64, vektori gradijenta bez predznaka su između 0 i 180 umjesto 0 i 360. Sada dolazi histogramski dio HOG-a. Histogram ima 9 spremnika od po 20 stupnjeva za sve moguće stupnjeve od 0 do 180. Sva 64 vektora gradijenta doprinose spremniku, no njihov je doprinos podijeljen. Doprinos veličine svakog vektora dijeli se između susjednih spremnika ovisno o kutu. Stoga će jači gradijentni vektori imati veći utjecaj na histogram. Nakon izračunavanja histograma svih 8x8 ćelija na slici (pomicanje za 4 piksela kroz prozor detekcije), normaliziraju se histogrami. Jedan od problema s korištenjem gradijentnih vektora je taj što su vrlo reaktivni na osvjetljenje i promjene kontrasta. Kako bi se to uzelo u obzir, može se normalizirati sliku. HOG normalizira sliku u blokovima. Da bi to napravio, uzima 4 susjedna histograma (4x9= 36 komponenti), a zatim dijeli svaki od njih ukupnom veličinom. Korištenjem blokiranja preklapanja određena količina ćelija se ponovno pojavljuje tako se onda može početi generalizirati i između blokova [19].

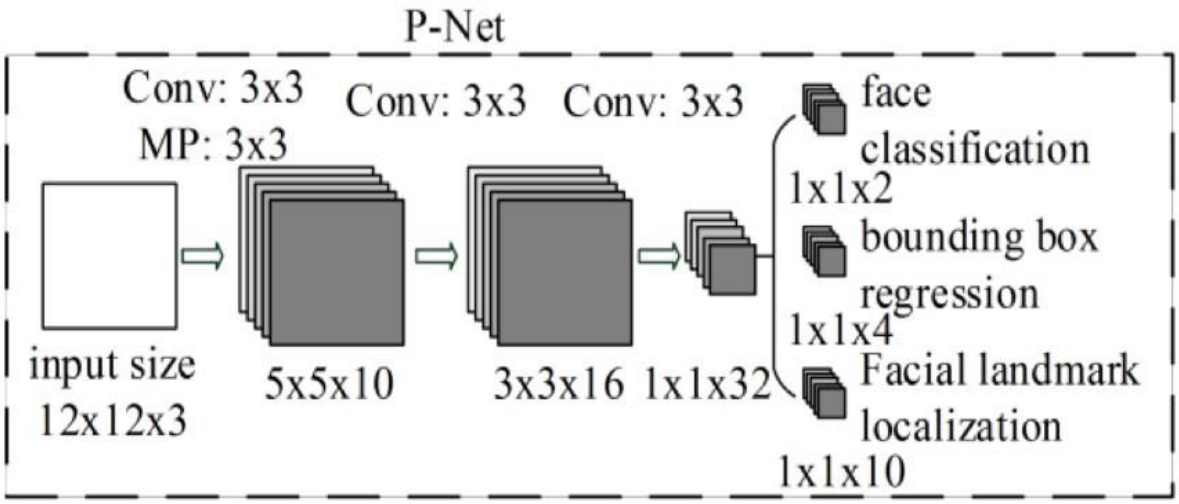
Provedbom korak u gornjem odlomku završava se s generiranjem vektora značajki za danu sliku. Nakon dobivanja ovog vektora značajki za danu sliku, metoda potpornih vektora (SVM) se trenira za otkrivanje je li ovaj vektor značajki predstavlja pješaka ili ne. SVM se ne može koristiti na izvornoj slici zato što je izvorna slika vrlo sirova. Ima svakakvih stvari poput drveća, šeširа i cesta koje nisu korisne za zadatak klasifikacije tj. ima previše ulaznih varijabli. Zapravo, to ometa zadatak klasifikacije jer bi koeficijenti SVM-a mogli učiti otkrivati nasumične stvari umjesto stvarnih pješaka. Nakon treniranja SMV-a na tisućama takvih HOG skupova značajki, sustav je spreman detektirati pješake [19].

2.4.3. Višezadaćne kaskadne konvolucijske mreže (MTCNN)

Ova metoda razvijena je kao rješenje za detekciju lica i poravnavanje lica. Ova je metoda prvi put predstavljena u radu pod nazivom „Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks” 2016 godine od strane autora Zhang, Zhang i Zhifen. Ova metoda daje najtočnije rezultate od sve tri metode (2.4.1., 2.4.2. i 2.4.3.) spomenute u ovom radu. Radi na slikama lica koje imaju različite orijentacije i može otkriti lica u različitim mjerilima. Može se čak nositi i s okluzijama. Sama po sebi nema nikakav veći nedostatak, ali je relativno sporija u odnosu na 2.4.1. i 2.4.2. [15].

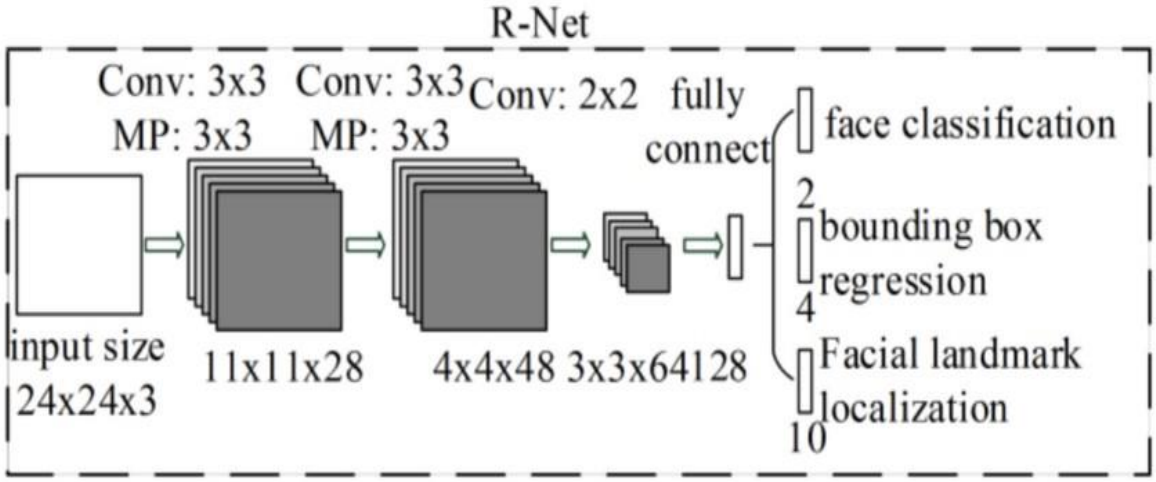
Početni korak kod implementacije MTCNN je uzeti ulaznu sliku i promijeniti joj veličinu u različitim mjerilima kako bi se izgradila slikovna piramida, koja je ulaz sljedeće trostupanjske kaskadne mreže. Nakon toga slijedi provedba triju faza MTCNN-a [20]:

1. **Faza 1: Mreža prijedloga (P-Net) (Slika 2.8.)** - Ova prva faza je potpuno konvolucijska mreža (FCN). Razlika između CNN-a i FCN-a je u tome što potpuno konvolucijska mreža ne koristi gusti sloj kao dio arhitekture. Ova se mreža prijedloga koristi za dobivanje prozora kandidata i njihovih regresijskih vektora graničnog okvira. Regresija graničnog okvira popularna je tehnika za predviđanje lokalizacije okvira kada je cilj detekcija objekta neke unaprijed definirane klase, u ovom slučaju lica. Nakon dobivanja vektora graničnog okvira, radi se određeno usavršavanje kako bi se kombinirala preklapajuća područja. Konačni rezultat ove faze su svi prozori kandidata nakon preciziranja kako bi se smanjio volumen kandidata.



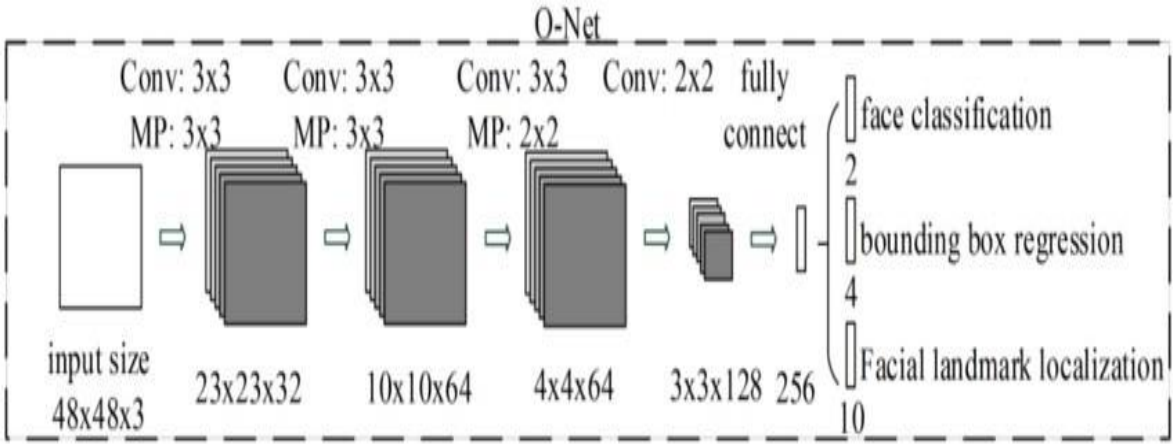
Slika 2.8. Prikaz mreže prijedloga (P-Net) [20]

2. **Faza 2: Mreža pročišćavanja (R-Net) (Slika 2.9.)** - Svi kandidati iz P-Neta ubacuju se u mrežu pročišćavanja. Ovo je CNN, a ne FCN kao ona u prvoj fazi jer postoji gusti sloj u posljednjoj fazi mrežne arhitekture. Mreža pročišćavanja dodatno smanjuje broj kandidata, izvodi kalibraciju s regresijom graničnog okvira i koristi ne-maksimalno potiskivanje (NMS) za spajanje preklapajućih kandidata. Izlazi mreže pročišćavanja su vektor za klasifikacije lica, vektor za granični okvir za lice i vektor za lokalizaciju orijentira lica.



Slika 2.9. Prikaz mreže pročišćavanja (R-Net) [20]

3. **Faza 3: Izlazna mreža (O-Net) (Slika 2.10.)** - Ova je faza slična R-Net-u, ali ova izlazna mreža ima za cilj detaljnije opisati lice i prikazati položaje pet orijentira lica za oči, nos i usta.



Slika 2.10. Prikaz izlazne mreže (O-Net) [20]

Zadatak MTCNN metode je isporučiti tri stvari: klasifikaciju licu, regresiju graničnog okvira i lokalizaciju orijentira lica. Klasifikacija lica označava problem binarne klasifikacije koji koristi gubitak unakrsne entropije. Kod regresije graničnog okvira za svaki prozor kandidata izračunava se pomak između kandidata i najbliže temeljne istine. Za ovaj zadatak koristi se euklidski gubitak. Te za kraj lokalizacija orijentira lica formulirana je kao problem

regresije, u kojem je funkcija gubitka euklidska udaljenost. Postoji pet orijentira lica: lijevo oko, desno oko, nos, lijevi kut usta i desni kut usta [20].

2.5. Poravnanje lica

Poravnanje lica naziva se i normalizacija lica koja pomaže u poboljšanju točnosti prepoznavanja lica. Rezultati ove tehnike normalizacije su ti da se lica nalaze u središtu slike, rotirana tako da je linija koja spaja središte dvaju očiju paralelna s vodoravnom linijom i mijenja veličinu lica u identično mjerilo. Za ovo se može koristiti sljedeća metoda: primjena analitičkog 3D modela lica nakon čega slijedi traženje sličnih konfiguracija fiducijalnih točaka iz vanjskog skupa podataka za zaključivanje iz nenadziranih metoda koje pronalaze transformaciju sličnosti za piksele. Iako je poravnanje naširoko korišteno, trenutno ne postoji potpuno fizički ispravno rješenje u kontekstu neograničene verifikacije lica. 3D modeli su posljednjih godina izgubili naklonost, posebno u okruženjima bez ograničenja jer se većina baza sastoje od 2D slika [21].

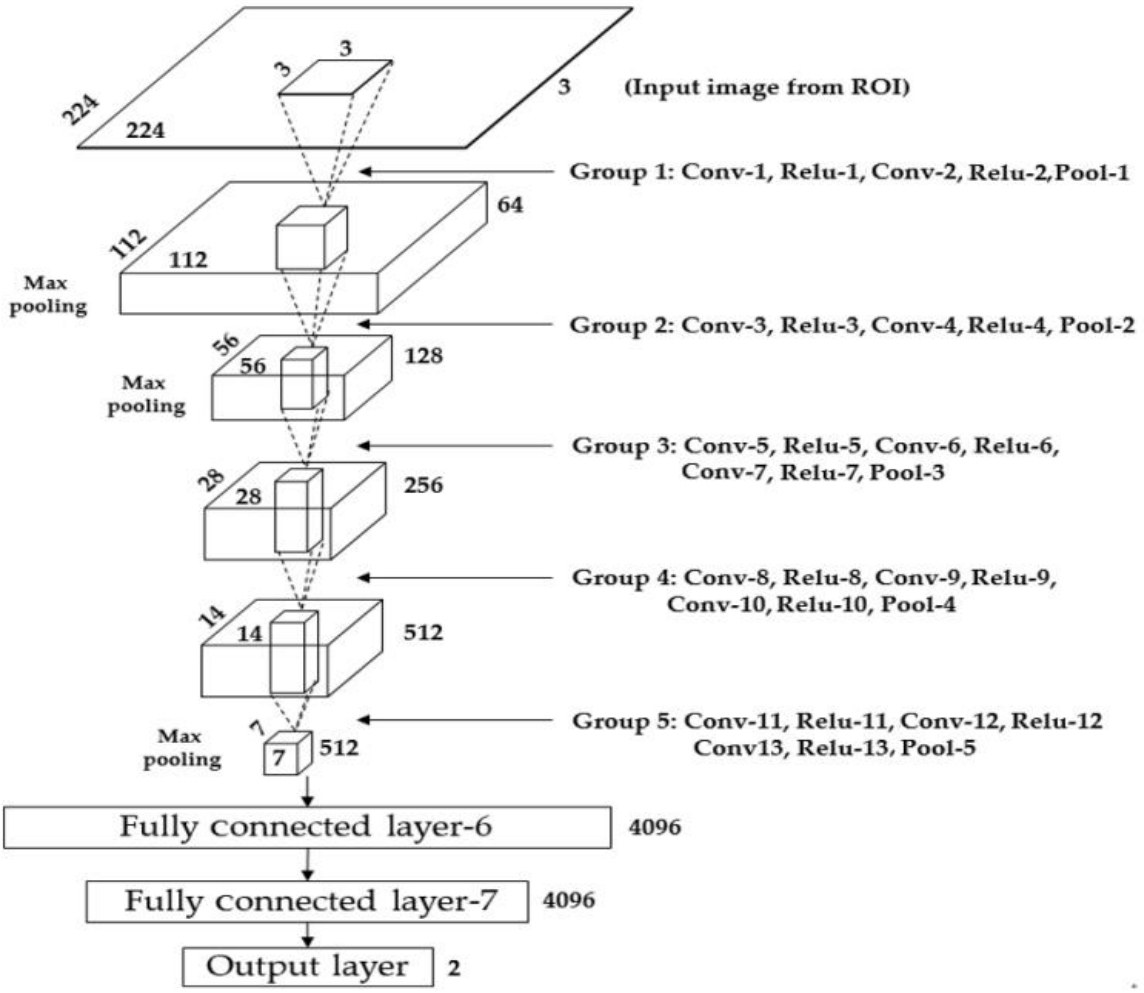
Izdvajanje položaja očiju je najvažniji čimbenik za poravnavanje lica. Nakon dobivanja lokacije očiju otkrivenog lica, može se rotirati slika po 1 stupanj dok oba oka ne budu vodoravna. To će povećati složenost rješenja tako da može poravnati lice na temelju kutova između dva oka pomoću pravila kosinusa. MTCNN također pronalazi neke orijentire lica kao što su lokacije očiju, nosa i usta. Koristeći MTCNN u koracima za prepoznavanje lica, on će automatski izvršiti usklađivanje s detektiranim licem. Google je izvijestio da poravnanje lica poboljšava točnost njegovog FaceNet modela prepoznavanja lica s 98,87% na 99,63%. To je povećanje točnosti od gotovo 1% [15].

2.6. Izdvajanje značajki (metode temeljene na dubokom učenju)

U 2014. i 2015. došlo je do niza istraživanja i publikacija o metodama dubokog učenja za prepoznavanje lica. Izvedbe su brzo dosegle gotovo ljudsku razinu sposobnosti, a zatim su nadmašile izvedbu na ljudskoj razini tijekom trogodišnjeg razdoblja na standardnom skupu podataka za prepoznavanje lica. Ovaj napredak pokreću tri ključna sustava dubokog učenje za prepoznavanje lica: VGGFace, FaceNet i DeepFace [15].

2.6.1. VGGFace

VGGFace odnosi se na niz modela razvijenih za prepoznavanje lica i demonstriranih na referentnim skupovima podataka računalnog vida od strane članova Visual Geometry Group (VGG) sa Sveučilišta u Oxfordu. Sastoji se od 11 slojeva u kojima je osam konvolucijskih slojeva i 3 potpuno povezana sloja. Skup podataka VGGFace2 koji su predložili Cao et al. sadržava 3,31 milijuna slika na kojima se nalaze 9131 jedinstvenih osoba. Varijacije uključuju dob, etničku pripadnost, pozu, profesiju i osvjetljenje. Ovo je najveći skup podataka dostupan za verifikaciju lica. Model VGGFace (Slika 2.11.) može se koristiti za verifikaciju lica također izračunavanjem emeddinga lica za novo lice i usporedbom embeddinga s embeddingom za jedan primjer lica koji je od ranije poznat sustavu. Euklidska udaljenost i kosinusna udaljenost izračunavaju se između dva embeddinga, a lica se podudaraju ili su verificirana ako je udaljenost ispod unaprijed definiranog praga koji je podešen za određene skupove podataka ili primjene [21].



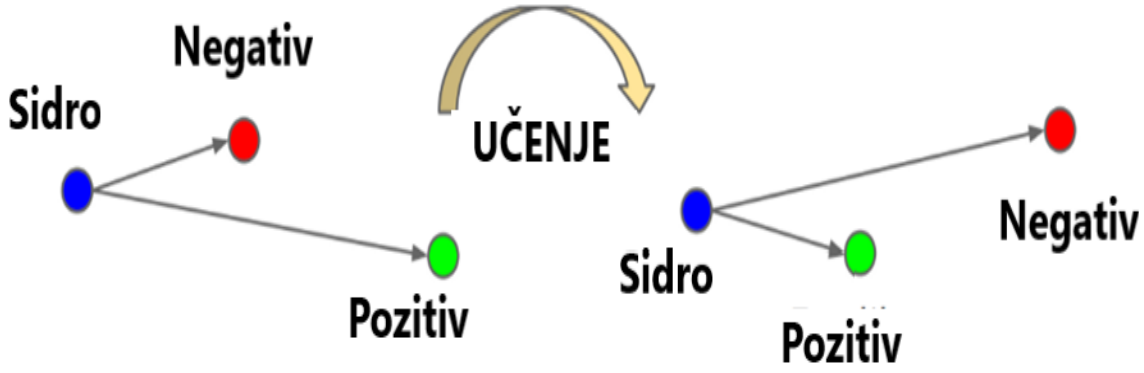
Slika 2.11. Prikaz VGGFace arhitekture [21]

2.6.2. FaceNet

FaceNet je sustav za prepoznavanje lica koji su 2015. godine razvili Googleovi istraživači u svom radu iz 2015. pod nazivom „FaceNet: A Unified Embedding for Face Recognition and Clustering". FaceNet je tada postigao najnaprednije rezultate na nizu referentnih skupova podataka za prepoznavanje lica i predstavio inovaciju nazvanu „triplet gubitak“ koja je omogućila učinkovito kodiranje slika kao vektora značajki koji su omogućili brzi izračun sličnosti i podudaranje putem izračuna udaljenosti. FaceNet sustav može se široko koristiti zahvaljujući višestrukim implementacijama modela otvorenog koda od trećih strana i dostupnosti unaprijed treniranih modela. FaceNet sustav može se koristiti za izdvajanje visokokvalitetnih značajki lica, zvanih embedding lica, koja se mogu koristi se za treniranje sustava za identifikaciju lica [15].

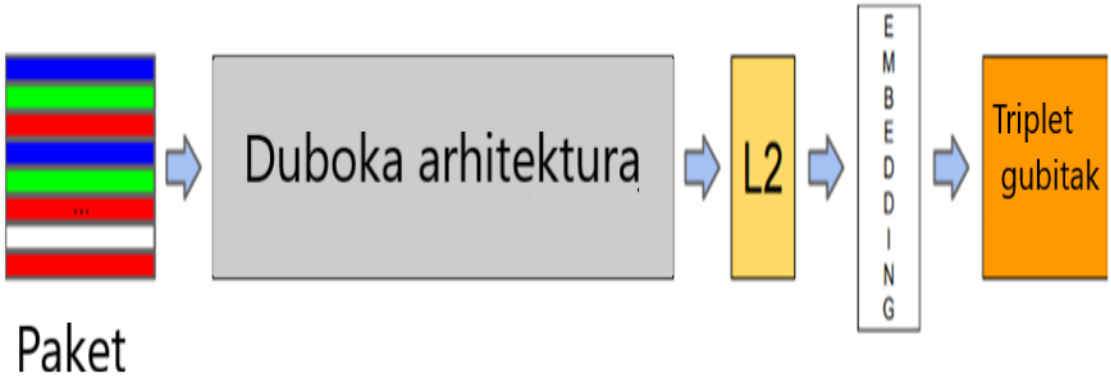
FaceNet uči izravno preslikavanje slika lica u kompaktni euklidski prostor u kojem udaljenosti izravno predstavljaju mjeru sličnosti lica, tj. lica iste osobe imaju male udaljenosti, a lica različitih osoba velike udaljenosti. Upotrebom FaceNet embeddinga kao vektora značajki mogu se implementirati zadatci poput prepoznavanja lica (tko je ova osoba), verifikacije (je li to ista osoba) ili grupiranja (pronađi uobičajene ljude među tim licima). Verifikacija lica također jednostavno uključuje određivanje udaljenosti između dva embeddinga; prepoznavanje postaje problem k-NN klasifikacije, a klasteriranje se može postići korištenjem uobičajenih tehnika kao što su k-srednje vrijednosti. FaceNet koristi duboku konvolucijsku mrežu osposobljenu za izravnu optimizaciju samog embeddinga, umjesto srednjeg sloja uskog grla kao u prethodnim pristupima dubokog učenja. Za treniranje slika pomoću FaceNeta koriste se tripleti grubo poravnatih podudarajućih ili nepodudarajućih zakrpa lica generiranih pomoću nove online metode rudarenja tripleta [21].

Trening triplet gubitka (Slika 2.12.) ima za cilj učenje vektora rezultata verifikacije identiteta usporedbom deskriptora lica u Euklidskom prostoru. Ovo je u duhu slično "metričkom učenju" i, poput mnogih pristupa metričkom učenju, koristi se za učenje projekcije koja je u isto vrijeme distinktivna i kompaktna, postižući istovremeno smanjenje dimenzionalnosti. Triplet gubitak je funkcija gubitka za umjetne neuronske mreže gdje se osnovni (sidro) unos uspoređuje s pozitivnim (istinitim) unosom i negativnim (lažnim) unosom. Udaljenost od osnovnog (sidra) ulaza do pozitivnog (istinitog) unosa je minimizirana, a udaljenost od ulaza osnovne linije (sidra) do negativnog (lažnog) unosa je maksimizirana [21].



Slika 2.12. Prikaz učenja triplet gubitka [21]

U strukturi FaceNet modela (Slika 2.13.), mreža se sastoji od paketnog ulaznog sloja i dubokog CNN-a praćenog L2 normalizacijom, što rezultira embeddingom lica. Nakon toga slijedi tripletni gubitak tijekom treninga [21].



Slika 2.13. Prikaz strukture FaceNet modela [21]

Embedding lica je vektorski prikaz značajki izdvojenih iz lica. Vrlo je korisno pronaći sličnost između dva vektora značajki. Na primjer, jedan vektor koji je blizu (po nekim mjerama od vektora značajki) može biti ista osoba, dok drugi vektor koji je daleko (po nekim mjerama od vektora značajki) može biti druga osoba. FaceNet model će generirati embedding za danu sliku lica. FaceNet model se može koristiti kao dio samog klasifikatora ili se može koristiti za prethodnu obradu lica kako bi se stvorio embedding lica koji se može pohraniti i koristiti kao ulaz u model klasifikatora. Ovaj drugi pristup je poželjan jer je FaceNet model velik i spor za

stvaranje embeddinga lica. Kako bi se usporedile dvije slike, izradi se embedding za obje slike zasebnim provlačenjem kroz model. Tada se može upotrijebiti formulu za pronalaženje udaljenosti koja će rezultirati nižom vrijednošću za slična lica i većom vrijednošću za različita lica [21].

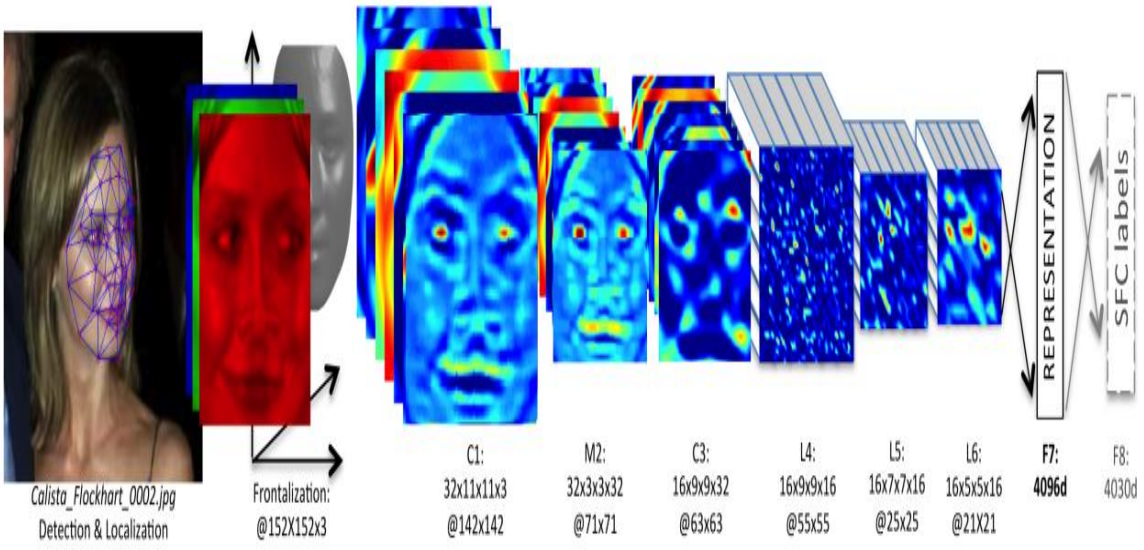
2.6.3. DeepFace

DeepFace je sustav temeljen na dubokim konvolucijskim neuronskim mrežama. On je opisano u dokumentu iz 2014. pod naslovom „DeepFace: Closing the Gap to Human-Level Performance in Face Verification". Bio je to možda prvi veliki korak naprijed korištenju dubokog učenja za prepoznavanje lica, postižući gotovo ljudske performanse na standardnom referentnom skupu podataka [15].

DeepFace koristi duboki CNN istreniran za klasifikaciju lica pomoću skupa podataka od 4 milijuna primjera koji obuhvaćaju 4000 jedinstvenih identiteta. Kreirala ga je istraživačka grupa na Facebooku. Također koristi sijamsku mrežnu arhitekturu, gdje se isti CNN primjenjuje na parove lica kako bi se dobili deskriptori koji se zatim uspoređuju korištenjem euklidske udaljenosti. Cilj treniranja je smanjiti udaljenost između podudarnih parova lica (tj. prikazivanje istog identiteta) i maksimalno povećati udaljenost između nepodudarnih parova, oblik metričkog učenja. DeepFace koristi skup CNN-ova, kao i fazu predobrade u kojoj se slike lica usklađuju s kanonskom pozom pomoću 3D modela. DeepFace je postigao najbolju izvedbu na Labeled Faces in the Wild (LFW) skupu podataka, kao i YouTube Faces DB. DeepFace je zatvorio većinu preostale praznine u najpopularnijem mjerilu za prepoznavanje lica bez ograničenja i sada je na rubu razine ljudske točnosti. Konkretno, s licima, uspjeh naučene mreže u hvatanju izgleda lica na robustan način uvelike ovisi o vrlo brzom koraku 3D poravnanja. Mrežna arhitektura temelji se na pretpostavci da je, kada je poravnanje dovršeno, lokacija svake regije lica je fiksirana na razini piksela. Stoga je moguće učiti iz neobrađenih RGB vrijednosti piksela, bez ikakve potrebe za primjenom nekoliko slojeva konvolucije kao što se radi u mnogim drugim mrežama [21].

Prikaz DeepFace arhitekture nalazi se na slici (Slika 2.14.). Prednji dio arhitekture sastoji se jednog konvolucijskog-pooling-konvolucijskog filtriranja na ispravljenom ulazu,

nakon čega slijede tri lokalno povezana sloja i dva potpuno povezana sloja. Boje ilustriraju mape značajki proizvedene na svakom sloju. Mreža uključuje više od 120 milijuna parametara, gdje više od 95% dolazi iz lokalnih i potpuno povezanih slojeva [21].



Slika 2.14. Prikaz DeepFace arhitekture [21]

2.7. Klasifikacija značajki

Za klasifikaciju značajki mogu se koristiti različiti pristupi koji većinom proizlaze iz statistike. Najčešće korišteni su [15]:

- **Euklidska udaljenost** - to je metoda klasifikacije značajki temeljena na udaljenosti koja izračunava udaljenost između čvorova na licu, a lice koje ima najmanju razliku između tih vrijednosti udaljenosti smatra se podudaranjem. Ova metoda je prikladan za skupove podataka koji imaju manji broj klasa i nižedimenzionalne značajke.
- **Kosinusna sličnost** - razmatra se rješenje koje se dobije nakon izračuna kosinusa kuta. Ovdje bi se usporedile razlike između rezultata. Što je vrijednost bliža 1, veća je vjerojatnost podudaranja. Ova metoda može dati pogrešan rezultat ako značajke testnih podataka nisu potpune.
- **SVM (Metoda potpornih vektora)** - SVM stvara optimalnu hiperravninu za klasifikaciju klasa skupa podataka za trening na temelju različitih značajki lica. Dimenzionalnost hiperravnine je za jedan manja od broja značajki. Različite jezgre

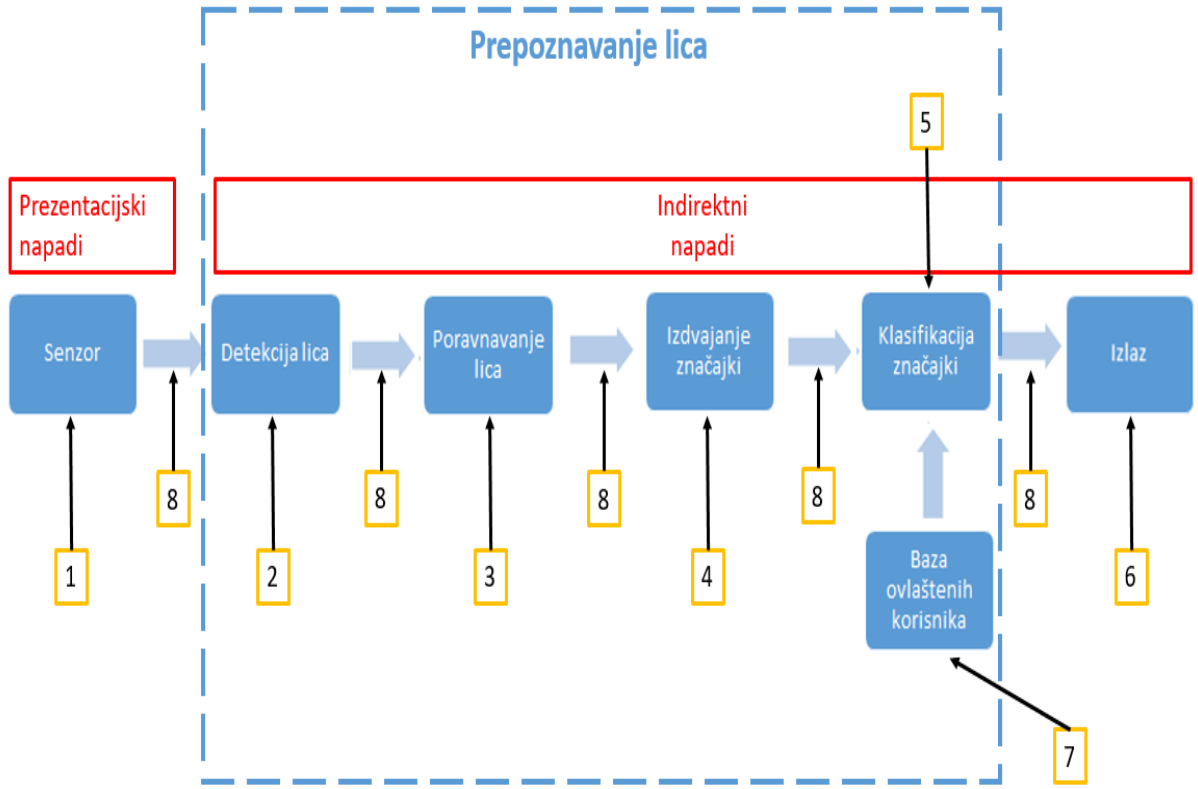
mogu biti primijenjene kako bi se vidjelo koje značajke koristi klasifikator za uklanjanje značajki. To može pomoći u poboljšanju brzine.

- **KNN (K-najbliži susjed) -** KNN se temelji na broju susjeda, tj. vrijednosti k. U KNN-u, ako je k=3, tada provjeravamo jesu li podaci blizu koje 3 podatkovne točke. Nakon toga se odlučuje kojoj klasi pripada većina najbližih podatkovnih točaka. Sada se predviđa da će testni podaci biti u ovoj klasi. KNN ima problem „prokletstva dimenzionalnosti" koji se može riješiti primjenom PCA prije upotrebe KNN klasifikatora.
- **ANN (Umjetna neuronska mreža) -** ANN koristi vrlo detaljan algoritam za prepoznavanje lica. Klasificira lokalnu teksturu koristeći višeslojni perceptron za poravnanje lica. Koristi analizu neovisnih komponenti temeljenu na geometrijskim značajkama za izdvajanje značajki i višestruku umjetnu neuronsku mrežu za podudaranje značajki.

2.8. Prepoznavanje lica: vrste napada lažiranjem i načini njihovog sprečavanja

Tehnologija prepoznavanja lica bit će jedan od čimbenika koji određuju budući krajolik umjetne inteligencije. Koliko god obećavajuće bilo prepoznavanje lica, ono ima nedostataka. Korisničke fotografije lako se mogu pronaći putem društvenih mreža i koristiti za lažiranje sustava za prepoznavanje lica. Napadi se mogu izvršiti pomoću papirnatih fotografija, snimki zaslona ili 3D rekonstrukcije lica. Zato je važno da tvrtke i vlade imaju postavljene sustave za zaštitu od lažiranja lica kako bi zaštitile osjetljive podatke, smanjile krađu i umanjile prijevare. Ovi sustavi poboljšavaju postojeća rješenja za prepoznavanje lica provjerom je li osoba stvarno prisutna ili se fotografija koristi za lažiranje sustava [22].

Suprotno glavnom cilju istraživanja prepoznavanja lica tj. poboljšanju performansi pri zadacima verifikacije i identifikacije, sigurnosne ranjivosti sustava za prepoznavanje lica bile su mnogo manje proučavane u prošlosti. Određena pozornost otkrivanju različitih vrsta napada posvećena je tek u posljednjih nekoliko godina, a sastoji se od otkrivanja dolazi li biometrijska značajka od žive osobe ili je lažna. Iz donje slike (Slika 2.15.) je vidljivo da postoji osam modula i točaka koje mogu biti meta napada, a napadi se dijele na dvije vrste: prezentacijski i neizravni napadi [23].



Slika 2.15. Prikaz osam modula sustava prepoznavanja lica koji mogu biti napadnuti

Prezentacijski napadi izvode se na razini senzora (1), bez potrebe za pristupom unutrašnjosti sustava. Prezentacijski napadi povezani su s čisto biometrijskim ranjivostima. U tim napadima uljezi koriste neku vrstu artefakta, obično umjetnog (npr. fotografiju lica, masku, sintetički otisak prsta ili ispisanu sliku šarenice) ili pokušavaju oponašati izgled pravih korisnika (npr. hod, potpis) za lažni pristup biometrijskom sustavu. Budući da "biometrijske osobine nisu tajne", napadači su svjesni ove stvarnosti da je izložena velika količina biometrijskih podataka koji prikazuju lice, oči, glas i ponašanje ljudi, pa iskorištavaju te izvore informacija kako bi pokušali zaobići sustave za prepoznavanje lica koristeći sljedeće primjere [23]:

- Napadači koriste fotografiju korisnika kako bi se lažno predstavljali.
- Napadači koriste videozapis korisnika kako bi se lažno predstavljali.
- Napadači mogu izraditi i koristiti 3D model napadnutog lica kao što je hiperrealistična maska

Neizravni napadi (2-8) izvode se na bazu podataka, komunikacijske kanale, itd. U ovoj vrsti napada, napadač treba pristup unutrašnjosti sustava. Neizravni napadi mogu se spriječiti tehnikama koje se odnose na "klasičnu" kibernetičku sigurnost, a ne na biometriju, pa one neće biti dalje opisane. Za razliku od neizravnih napada za sprečavanje prezentacijskih napada koriste se tehnike protiv lažiranja [23].

Bez implementacije mjera otkrivanja prezentacijskih napada, većina najsuvremenijih biometrijskih sustava lica ranjiva je na jednostavne napade. Korištenje fotografija i videa najčešći su tip napada zbog visoke ekspozicije lica i niske cijene digitalnih fotoaparata visoke rezolucije. Najpoznatije metode prezentacijskih napada su [23]:

- **Foto napadi -** ovi napadi sastoje se od prikazivanja printane fotografije ili fotografije na mobitelu napadnutog identiteta senzoru sustava za prepoznavanje lica.
- **Video napadi** – ovi napadi sastoje se od reproduciranja videa legitimnog korisnika i prikazivanja tog videa senzoru/kameri na bilo kojem uređaju koji reproducira video.
- **Napadi 3D maskom -** ovi napadi sastoje se od toga da napadač gradi 3D rekonstrukciju lica i prikazuje je senzoru/kameri.
- **Ostali napadi** – ovi napadi sastoje se od toga da napadač koristi šminku ili se podvrgne plastičnoj operacija kako bi lažirao legitimnog korisnika.

Danas 2D napadi su popularniji od 3D zbog tehnoloških ograničenja. Kada je riječ o razvoju rješenja za ovaj problem važno se usredotočiti se na tehnike koje: sprječavaju statična i dinamička 2D lažiranja, za koje nije potrebna interakcija korisnika i koje koristite slike, a ne videozapise. Pouzdano rješenje mora postići maksimalnu točnost, zahtijevati malo vremena, dati prednost korisničkom iskustvu te se treba integrirati s postojećim softverom za prepoznavanje lica [22]. U nastavku će biti navedene najkorištenije tehnike za sprečavanje lažiranja kod prepoznavanja lica .

Prepoznavanje treptaja jedan je od testova otkrivanja živosti koji je nevjerojatno precizan. Prirodno treptanje je jednostavan način da se utvrdi je li lice živo ili ne. Prosječan čovjek trepće 15-20 puta u minuti. Oči ostaju zatvorene oko 250 milisekundi tijekom treptaja. Moderne kamere snimaju video s daleko manjim intervalima između sličica (42 milisekundi pri 24 sličica u sekundi). Mogu se upotrijebiti videozapisi kako bi se pronašle sličice sa zatvorenim

očima i zatim se ih može prebrojati za dobivanje očekivanih brojeva. Implementacija detekcije treptaja oka može koristiti analizu orijentira lica i izračunati površinu očiju. Također može se primijeniti duboko učenje za ovaj zadatak [22].

Metode temeljene na dubokom učenju i konvolucijskim neuronskim mrežama (CNN) dodatna su rješenja koja mogu pomoći u borbi protiv lažiranja. Mnogi pristupi sprečavanju lažiranja pristupaju kao problemu binarne klasifikacije. Tim pristupom istrenira se CNN da prepozna koje su prave fotografije, a koje su lažne. Ta mreža nakon što se istrenira uspije diferencirati između pravih i lažnih lica, ali to radi unutar određenih ograničenja. Mreža radi točno samo s određenim skupovima podataka u određenim uvjetima, uključujući stvari poput kvalitete kamere, okoliša, svjetla itd. Promjenom bilo kojeg uvjeta, visoka je vjerojatnost da, neuronska mreža neće dati točne rezultate. Uzimajući sve navedeno u obzir, ova metoda je održiva samo u slučajevima uske upotrebe [22].

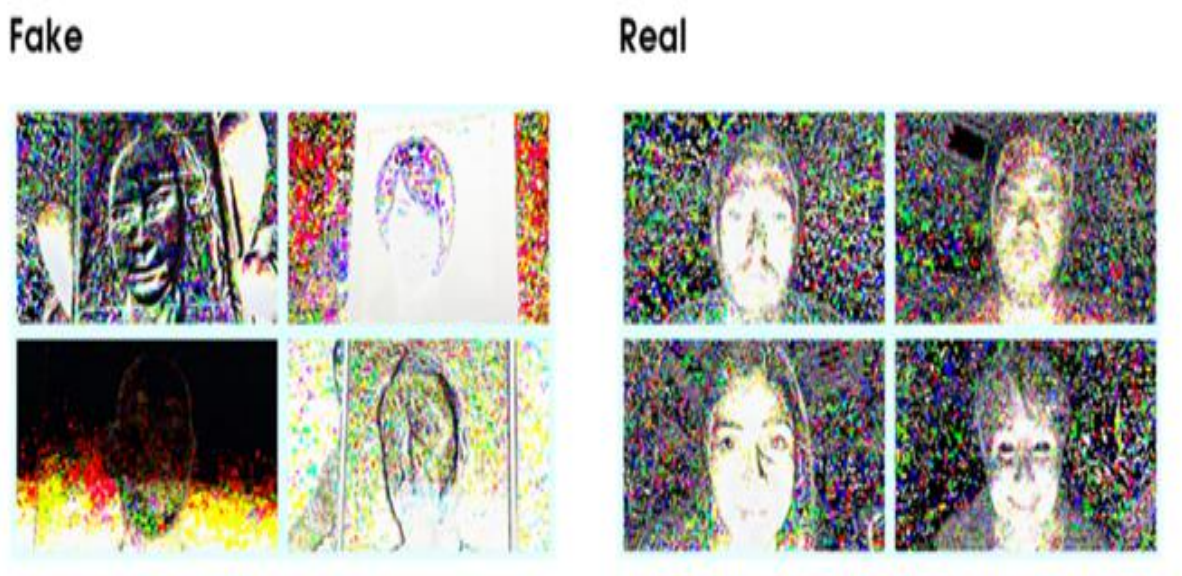
Metode izazova i odgovora su još jedna održiva tehnika za sprečavanje lažiranja. Ova tehnika koristi posebne radnje koja se nazivaju izazovi. Sustav radi na provjeri je li se izazov dogodio tijekom video sekvence. Sustav izazov-odgovor oslanja se na nizu izazova za potvrdu identiteta pojedinca. Ovi izazovi mogu uključivati [22]:

- Osmijehe;
- Izraze određenih emocija kao što je ljutnja ili tuga na licu;
- Određene pokrete glavom kao što je pogled ulijevo.

Ova metoda usprkos svojoj učinkovitosti zahtijeva dodatni unos i radnje te može značajno utjecati na korisničko iskustvo.

3D kamere su najpouzdanije sredstvo zaštite od lažiranja. Precizne informacije o dubini piksela mogu pružiti visoku točnost protiv prezentacijskih napada jer se mogu uočiti razlike između lica i ravnog oblika. 3D napadi mogu prouzročiti poteškoće, ali kamere su još uvijek jedna od najpouzdanijih dostupnih tehnika za sprečavanje lažiranja lica. Unatoč širokoj dostupnosti kamera postoje korisnici koji ih nemaju na svojim računalima. Zbog toga tehnike sprečavanja lažiranja koje rade s običnim RGB slikama su korištenije od 3D kamera [22].

Aktivni bljesak je zanimljiva metoda za sprečavanje lažiranja koja mnogo obećava, za razliku od nekih drugih rješenja ona ne pati od "problema crne kutije". Ovo rješenje omogućava otkrivanje lažiranja pomoću refleksije svjetla na licu. Ideja uključuje korištenje promjenjivog svjetlosnog okruženja koje pruža dodatno svjetlo koje dolazi sa zaslona uređaja. Bijelo svjetlo proizvodi odgovarajuću refleksiju na licu (Slika 2.16). Ova tehnika može se koristiti za diferenciranje pravih lica od lažnih. Uzimaju se slike prije i nakon bljeska te se trenira mreža pomoću tih podataka. Aktivna bljeskalica pomaže u razdvajanju značajki lica i njihovom klasificiranju. Moguće je izgraditi model neovisan o kutu lica (u razumnim granicama). Međutim, ako se izračunava udaljenost u pikselima, poravnanje lica postaje neophodno. Tehnologija bi se mogla učiniti sofisticiranijom na temelju posebnih slučajeva uporabe koje je potrebno riješiti [22].



Slika 2.16. Slike refleksije bijelog svijetla od lažnih i pravih lica primjenom aktivnog bljeska [22]

Svaka od ranije navedenih metoda za sprečavanje lažiranja primjenjiva je na svoj način. Kao i sve ostale tehnologije tako i ove metode imaju svoje prednosti i nedostatke. Neki od njih se mogu poboljšati dodavanjem dodatne složenosti, druge su prikladne samo kao samostalno rješenje, dok bi određene metode bile najbolje u kombinaciji više njih istovremeno.

3. IZRADA I OPIS FUNKCIONIRANJA SUSTAVA

3.1. Odabir alata potrebnih za izradu sustava

Kako bi se sam sustav realizirao potrebno je izabrati i koristiti prikladne alate pomoću kojih se mogu ostvariti razne funkcionalnosti implementirane unutar sustava. Kao programski jezik koji će biti korišten za izradu sustava odabran je Python. On je odabran zbog aktivne zajednice, jednostavnosti i velikog broja podržanih modula i biblioteka. Za integrirano razvojno okruženje izabran je PyCharm zbog raznih alata korisnih za Python koji poboljšavaju produktivnost. Za zadatke vezane uz računalni vid prvenstveno se koristi biblioteka OpenCV. Prilikom treniranja konvolucijske neuronske mreže korišten je Keras. Kod izrade samog rješenja sustava korištene su i mnoge druge Python biblioteke koje će biti spomenute u nastavku kod opisa koraka izrade sustava na konkretnim mjestima gdje su primijenjene.

3.1.1. PyCharm

PyCharm je integrirano razvojno okruženje (IDE) koje se primjenjuje kod računalnog programiranja, s naglaskom na Python programski jezik. Razvijen je od strane češke tvrtke JetBrains. PyCharm je višeplatformski IDE, s dostupnim verzijama za Linux, Windows i MacOS. Postoji besplatno i profesionalno izdanje. Ono što ih razlikuje je to da profesionalno izdanje ima neke dodatne naprednije mogućnosti. PyCharm pruža API koji omogućuje korisnicima da integriraju vlastite značajke kroz dodatke (*engl. plugins*) u PyCharm. Trenutno postoji više od 1000 dodataka koji su kompatibilni s PyCharmom. Neke od poznatijih značajki su mu [24]:

- Analiza i pomoć kod pisanja koda s automatskim dovršavanjem koda;
- Brza i lagana navigacija između projekata i unutar koda;
- Integracija kontrole verzija;
- Podrška za web kroz Flask, web2py i Django
- Automatsko naglašavanje pogrešaka;
- Python program za ispravljanje pogrešaka (engl. *debugger*);
- Podrška za znanstvene biblioteke kao što su SciPy, Matplotlib i NumPy.

3.1.2. Python

Python je općenamjenski, interpretirani, objektno orijentiran, interaktivan programski jezik visoke razine. Kreirao ga je Guido van Rossum u periodu od 1985. do 1990. Nekoliko glavnih prednosti Python su [25]:

- **Python je interpretativan** - interpreter obrađuje Python kod u realnom vremenu tijekom izvođenja tj. nije potrebno kompajlirati program prije nego što ga se izvrši. To svojstvo slično je kao kod PHP-a i PERL-a.
- **Python je interaktivan** – kroz Python konzolu može se izravno komunicirati s interpreterom te je tako moguće pisati željene programe.
- **Python je objektno orijentiran** - Python podržava objektno orijentiranu tehniku ili stil programiranja kod koje je kod enkapsulira unutar objekata.
- **Python je prilagođen početnicima** - Python je izvrstan jezik za programere koji tek počinju i podržava razvoj širokog raspona primjena od jednostavne obrade teksta do strojnog učenja i videoigara.

Python je dizajniran da bude vrlo čitljiv. Često koristi engleske ključne riječi tamo gdje drugi jezici koriste interpunkcijske znakove i ima manje sintaktičkih konstrukcija nego drugi programski jezici. Neke karakteristike programiranja u Pythonu su [25] :

- Podržava funkcionalne i strukturirane metode programiranja, kao i objektno orijentirano programiranje.
- Može se lako integrirati s C, C++ i Javom.
- Pruža vrlo visoku razinu dinamičkih tipova podataka i podržava dinamičku provjeru tipa podatka.
- Može se koristiti kao skriptni jezik ili se može kompajlirat u bytecode za izradu velikih aplikacija.
- Podržava automatsko prikupljanje smeća.

3.1.3. OpenCV

OpenCV je biblioteka računalnog vida otvorenog koda. Biblioteka je napisana u C i C++ i primjenjiva je za Linux, Windows i MacOS. Aktivno se razvijaju sučelja za Python, Ruby, Matlab i druge programske jezike. OpenCV je dizajniran za računalnu učinkovitost i s jakim fokusom na aplikacije u stvarnom vremenu. OpenCV je napisan u optimiziranom jeziku C i može iskoristiti višejezgrene procesore. Za daljnju automatsku optimizaciju na Intelovim arhitekturama, mogu se kupiti Intelove biblioteke Integrated Performance Primitives (IPP), koje se sastoje od optimiziranih rutina niske razine u mnogim različitim algoritamskim područjima. OpenCV automatski koristi odgovarajuću IPP biblioteku tijekom izvođenja ako je ta biblioteka instalirana [26].

Jedan od ciljeva OpenCV-a je pružiti infrastrukturu računalnog vida jednostavnu za korištenje koja pomaže ljudima da brzo izgrade prilično sofisticirane aplikacije za računalni vid. Biblioteka OpenCV sadrži više od 500 funkcija koje obuhvaćaju mnoga područja računalnog vida, uključujući inspekciju tvorničkih proizvoda, korisničko sučelje, kalibraciju kamere, medicinsko snimanje, stereo vid, sigurnost i robotiku. Budući da računalni vid i strojno učenje često idu ruku pod ruku, OpenCV također sadrži potpunu biblioteku strojnog učenja opće namjene (MLL). Ova podbiblioteka usmjerena je na statističko prepoznavanje uzoraka i klasteriranje. MLL je vrlo koristan za zadatke vida koji su srž misije OpenCV-a, ali je dovoljno općenit da se koristi za bilo koji problem strojnog učenja [26].

OpenCV je izrastao iz inicijative Intel Researcha za unaprjeđenje CPU-a intenzivnih aplikacija. S tim ciljem, Intel je pokrenuo mnoge projekte uključujući praćenje zraka svjetlosti u stvarnom vremenu. Jedan od autora koji je u to vrijeme radio za Intel posjećivao je sveučilišta i primijetio da su neke vrhunske sveučilišne grupe, kao što je MIT Media Lab, imale dobro razvijene i interno otvorene infrastrukture računalnog vida — kod koji se prenosio od studenta do studenta i koji je davao svakom novom studentu početnu prednost u razvoju vlastite aplikacije računalnog vida. Umjesto ponovnog osmišljavanja osnovnih funkcija od nule, novi učenik mogao bi započeti nadogradnjom onoga što je ranije napravljeno. Intel je vidio priliku za sebe na načina da bi se kroz omogućavanje aplikacija računalnog vida široj populaciji povećala potreba za brzim procesorima. Pokretanje nadogradnji na brže procesore generiralo bi više prihoda za Intel nego prodaja dodatnog softvera. To je vjerojatno razlog zašto je ovaj otvoreni i besplatni kod nastao od strane hardverske, a ne softverske tvrtke [26].

3.1.4. Keras

Keras je biblioteka neuronske mreže visoke razine otvorenog koda, koja je napisana u Pythonu i sposobna je raditi na Theano, TensorFlow ili CNTK. Razvijen je od strane Googleova inženjera Francoisa Cholleta. Napravljen je da bude jednostavan za korištenje, proširiv i modularan za omogućavanje bržeg eksperimentiranja s dubokim neuronskim mrežama. Konvolucijske neuronske mreže i rekurentne mreže nisu podržane pojedinačno, već samo njihova kombinacija. Ne može se nositi s proračunima niske razine, pa se koristi backend bibliotekom za rješavanje tog problema. Backend biblioteka djeluje kao omotač API-ja visoke razine za API niske razine, što mu omogućuje rad na TensorFlow, CNTK ili Theano [27].

Jedina mana Kerasa je što ima svoje unaprijed konfigurirane slojeve, a ne dopušta stvaranje vlastitih apstraktnih slojeva jer ne može koristiti API niske razine. Podržava samo API visoke razine koji radi na vrhu pozadinskog motora (TensorFlow, Theano i CNTK).

Keras ima sljedeće pogodnosti [27]:

- Vrlo je lako razumjeti i uključiti bržu implementaciju mrežnih modela.
- Ima veliku podršku zajednice na tržištu jer ga većina AI tvrtki želi koristiti.
- Podržava multi backend, što znači da se može koristiti bilo TensorFlow, CNTK ili Theano kao backend.
- Budući da ima jednostavnu implementaciju, također ima podršku za više platformi. Uređaji na kojima se Keras može primijeniti su: iOS s CoreML-om, Android s TensorFlow Androidom, web preglednik s podrškom za .js, Cloud engine i Raspberry Pi.
- Podržava paralelizam podataka, što znači da se Keras može trenirati na više GPU-ova odjednom za ubrzavanje vremena treniranja i obradu ogromne količine podataka.

3.2. Konceptualna razrada sustava

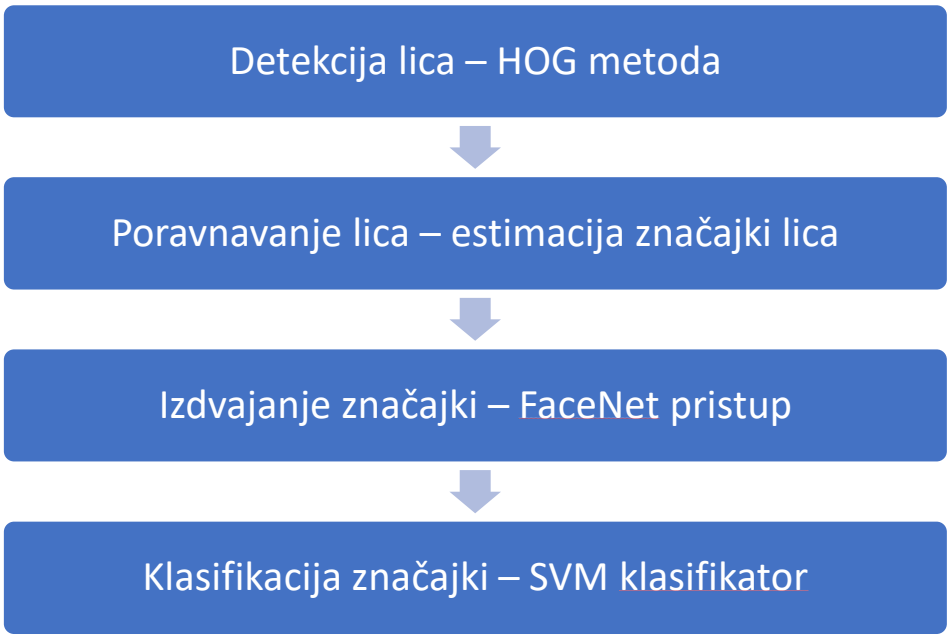
Budući da je postojeći sustav ručne evidencije pohađanja spor, zamoran i podložan greškama, ovaj rad namjerava stvoriti novi sustav evidencije pohađanja koji koristi prepoznavanje lica te ima sposobnost razlikovanja između lažnog i pravog lica u stvarnom vremenu. Kako bi se ovaj sustav lakše shvatio u nastavku će biti razdijeljen u tri dijela:

1. **Opis funkcioniranja rješenja za prepoznavanje lica**
2. **Opis funkcioniranja rješenja za razlikovanje pravih i lažnih lica na temelju treptanja**
3. **Povezivanje rješenja iz prve i druge točke u sustav za evidenciju pohađanja**

Svaki od ovih koraka i potkoraka bit će opširnije opisan te će biti navedene korištene tehnike i algoritmi uz dane slikovne primjere.

3.3. Opis funkcioniranja rješenja za prepoznavanje lica

Povodom nedavnih postignuća u razvoju dubokih konvolucijskih neuronskih mreža (DCNN) za zadatke detekcije i prepoznavanja lica, u ovom radu će se za prepoznavanje lica koristiti rješenje bazirano na dubokom učenju. Kao i ostala rješenja za prepoznavanje lica (2.3.) tako se i rješenje unutar ovog rada sastoji od 4 koraka (Slika 3.1.) koji će zasebno biti opisani u nastavku.



Slika 3.1. Prikaz koraka rješenja za prepoznavanje lica

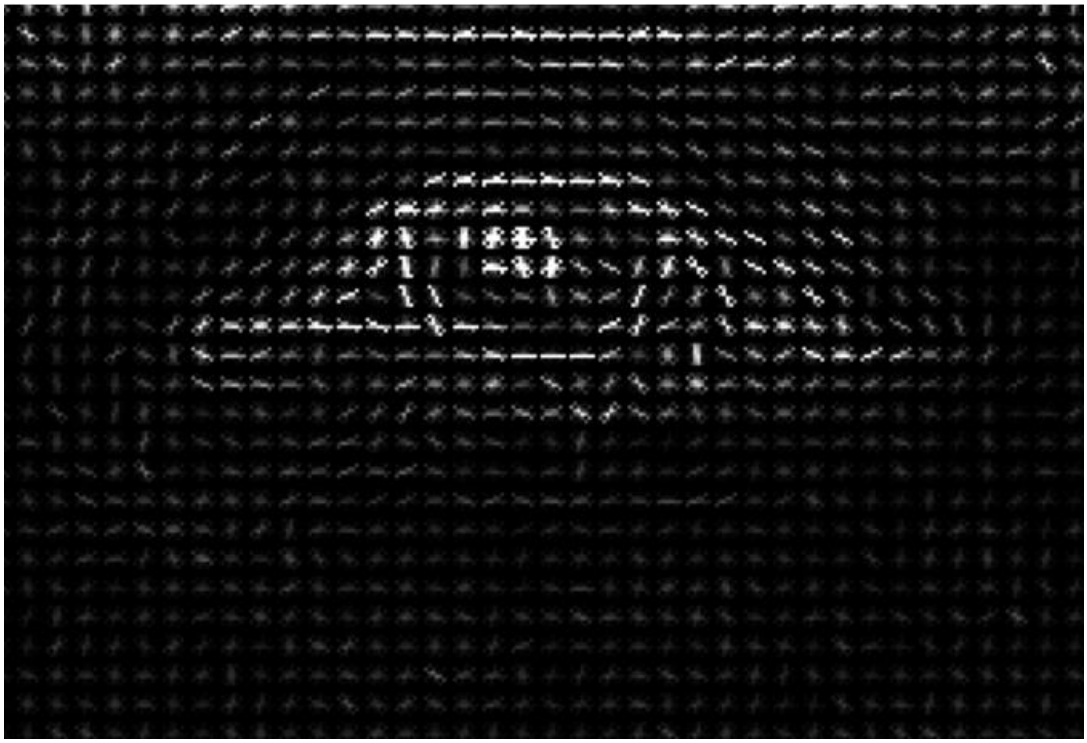
3.3.1. Prvi korak: Detekcija lica

Skoro svi ljudi su se sreli s detekcijom lica jer je ona jedna od značajki implementiranih kod kamere mobitela. Kod kamere se uglavnom koristi za provjeru ako se lice nalazi u fokusu prije nego se okine slika. No u ovom radu detekcija lica će služiti kao prvi korak unutar rješenja za prepoznavanje lica. Iako je najraširenija metoda za detekciju lica Viola Jones algoritam (2.4.1.) u ovom radu će se koristiti naprednija i preciznija HOG metoda (2.4.2.).

Kod Pythona i ostalih programskih jezika na početku se slika mora pretvoriti u interni format na kojem se onda može raditi, a taj format kod Pythona je većinom NumPy polje (engl. array). NumPy polje je n-dimenzionalno polje (nd-polje). RGB slika sadrži podatke u 3 dimenzije (širina, visina, kanal). To prikazano na primjeru za uobičajenu rezoluciju 1024x768 rezultira s ukupno 2 359 296 piksela (1024 *768 * 3). Svaki piksel po kanalu ima vrijednost od 8 bita (1 bajt) u rasponu od 0–255. To znači da svaki RGB piksel ima 3 bajta (24 bita) podataka (1 bajt za svaki kanal: R, G i B). Za lakše manipuliranje potrebno je spomenute podatke o pikselima pretvoriti u nd-polje. To se ostvaruje kombinacijom Python alata poput: Python Image Library (PIL), NumPy i OpenCV koji su korišteni u ovom radu. Kako bi se još olakšalo manipuliranje slika će se prebaciti u crno-bijelu jer za detekciju lica nije potrebna informacija o boji (Slika 3.2.).

Slika 3.2. Prikaz prebacivanja originalne slike u boji u crno-bijelu

Nakon što je slika crno-bijela za svaki piksel posebno se određuje koliko je on taman u odnosu na one koji ga okružuju te se u tom smjeru provlači strelica (Slika 3.3.) tj. gradijent (2.4.2.). To rezultira da su svi pikseli na slici zamijenjeni gradijentom tj. dobije se matrica gradijenta koja olakšava sam postupak. Ova matrica je uglavnom neovisna o varijacijama svjetline u izvornoj slici. Međutim, ona je još uvijek zahtjevna za manipulaciju. Zbog toga se formiraju podmatrice veličine 16x16. Nakon toga se određuje dominantni smjer za svaku podmatricu. To na kraju rezultira slikom (Slika 3.4.) koja prikazuje osnovnu strukturu lica na način koji je pogodan za manipulaciju. U ovom odlomku su ukratko opisani koraci HOG metode nakon kojih se brzo i na laki način pronalazi svako lice na slici što će biti korišteno u daljnjim koracima ovog rada.

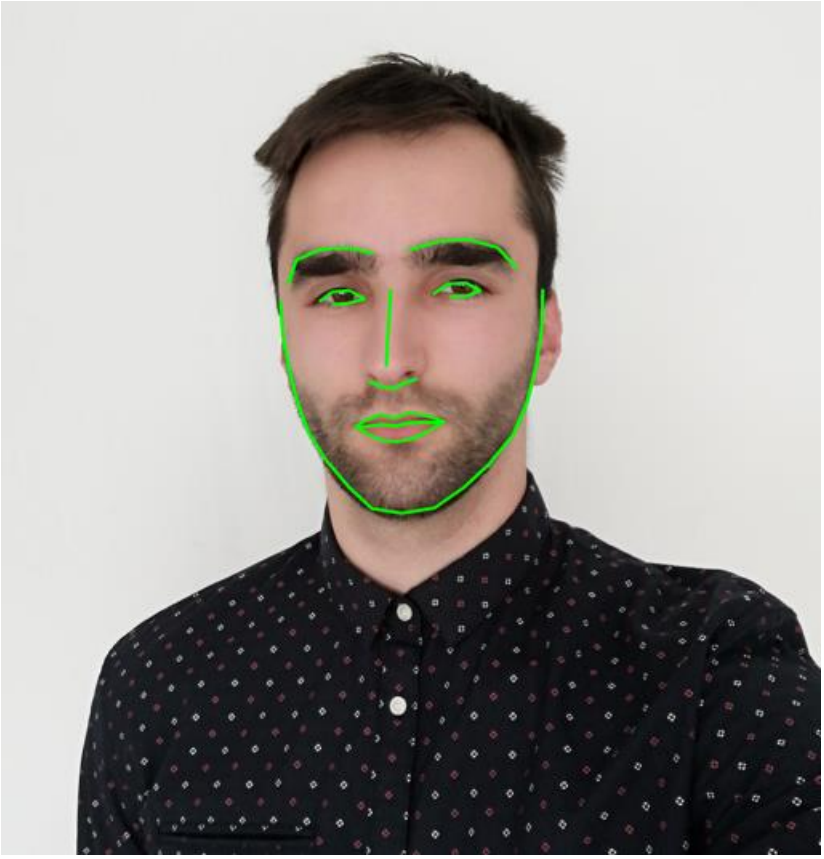


Slika 3.3. Uvećani prikaz slike na kojoj su pikseli zamijenjeni strelicama tj. gradijentima

Slika 3.4. Prikaz originalne slike i rezultata dobivenog provedbom HOG metode

3.3.2. Drugi korak: Poravnanje lica

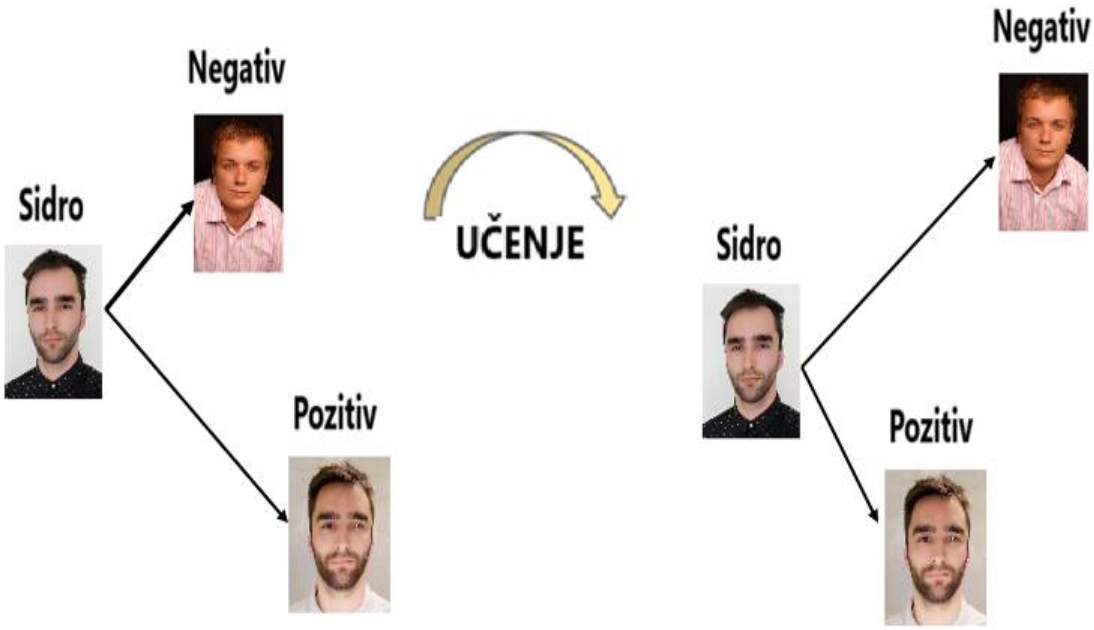
Nakon što su u prvom koraku detektirana i izdvojena lica sada treba riješiti problem kad lica na slici gledaju u nekom smjeru koji je različit od smjera direktnog gledanja u kameru. To je problem jer tada računala smatraju da se radi o potpuno drugoj osobi. Postoji nekoliko pristupa za rješavanje ovog problema, a unutar ovog rada će se primijeniti pristup estimaciji značajki lica. Sam pristup se temelji na 68 značajki lica (Slika 3.5.) koje posjeduju lice svake osobe, a sam algoritam za automatsku detekciju tih značajki opisan je u sljedećem radu [28]. Nakon što su poznate značajke lica potrebno je pozicionirati sliku primjenom afinih transformacija [29] (transformacije kod kojih su očuvane kolinearnost i omjeri udaljenosti) kako bi se orijentiri poput usta i očiju centrirali što je više moguće bez da pritom dođe do iskrivljenja slike.



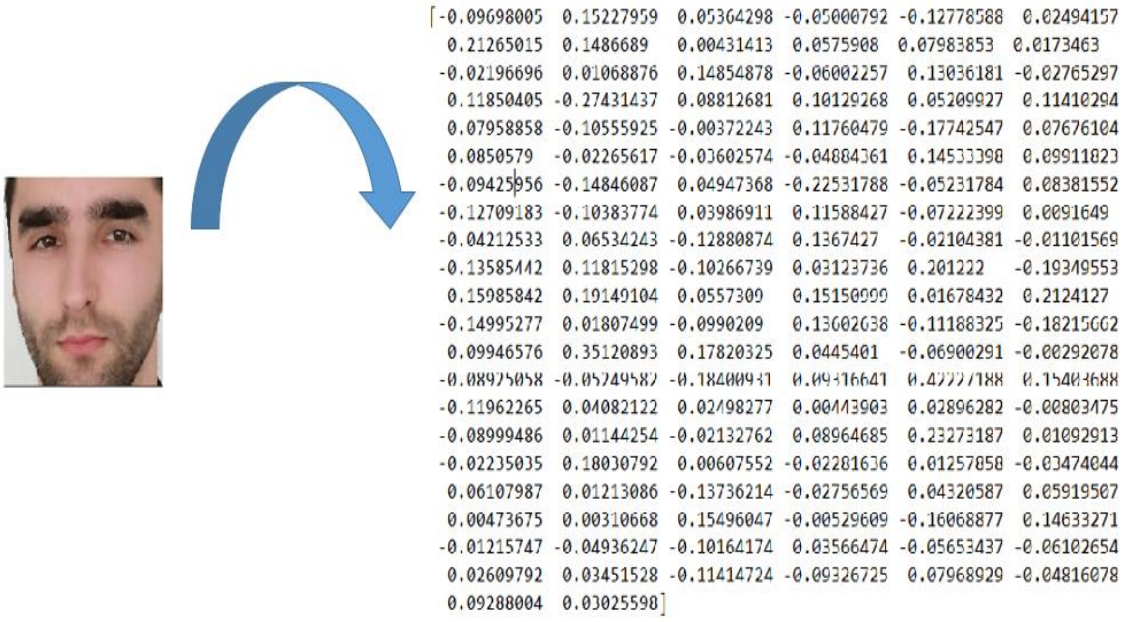
Slika 3.5. Prikaz 68 značajki lica na originalnoj slici

3.3.3. Treći korak: Izdvajanje značajki

Ovaj korak će omogućiti razlikovanje osoba. Kod izdvajanja značajki najbolji način je da se računalu dopusti da samo odabere one najbitnije. U tu svrhu se koristi FaceNet metoda (2.6.2.). Ova metoda koristi duboku konvolucijsku neuronsku mrežu za učenje preslikavanja slika lica u euklidski prostor gdje se udaljenosti podudaraju s mjerama sličnosti lica. To rezultira generiranjem embeddinga lica koji se sastoji od 128 bajta. Trening mreže bazira se na gubitku tripleta (Slika 3.6.), gdje triplet označava: sliku lica ciljane osobe, testnu sliku lica ciljane osobe i sliku lica druge osobe. Ovaj korak zahtijeva veliki skup podataka i puno računalne snage, ali mora se izvršiti samo jednom. OpenFace biblioteka [30] s unaprijed treniranom FaceNet mrežom korištena je za generiranje embeddinga lica. Na slici (Slika 3.7.) je dan prikaz embeddinga lica dobivenog nakon provlačenja slike kroz unaprijed treniranu mrežu.



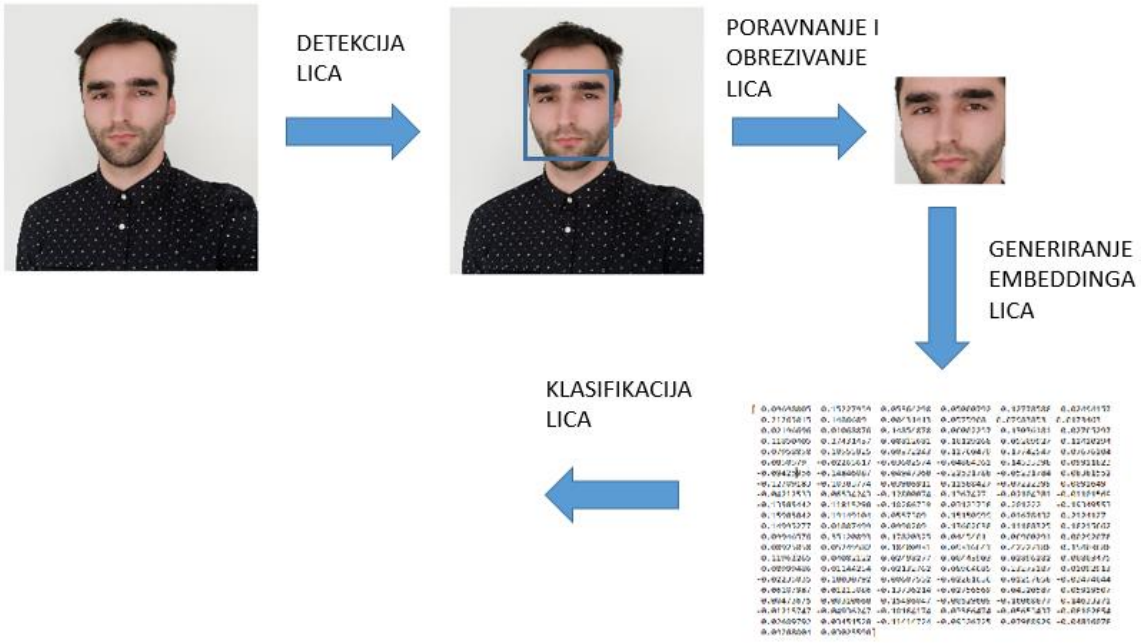
Slika 3.6. Prikaz učenja preko triplet gubitka



Slika 3.7. Prikaz 128d embeddinga lica

3.3.4. Četvrti korak: Klasifikacija značajki

Posljednji korak rješenja za prepoznavanje lica je vrlo jednostavan. Mjere lica koje se analizira se uspoređuju s mjerama lica koja se nalaze u bazi. Za tu svrhu zbog toga što je sam sustav baziran na malom skupu podataka odabran je SVM klasifikator. U nastavku su prikazani svi koraci izrađenog rješenja prepoznavanja lica preko dijagrama toka (Slika 3.8.)



Slika 3.8. Dijagram toka izrađenog rješenja prepoznavanja lica

3.4. Opis funkcioniranja rješenja za razlikovanje pravih i lažnih lica na temelju treptanja

Treptanje je bitna funkcija oka koja pomaže u otklanjanju čestica prašine i širenju suza te time sprječava da se oko osuši. Iako brzina treptanja može varirati s faktorima kao što su umor, ozljeda oka, stres, starost, količina sna, uzimanje određenih lijekova i bolest, stopa spontanog treptanja prosječnog čovjeka iznosi od 15 do 20 treptaja očima u minuti. Odnosno, osoba trepne otprilike jednom svake 3 do 4 sekunde, a prosječno treptanje traje oko 100 milisekundi [31].

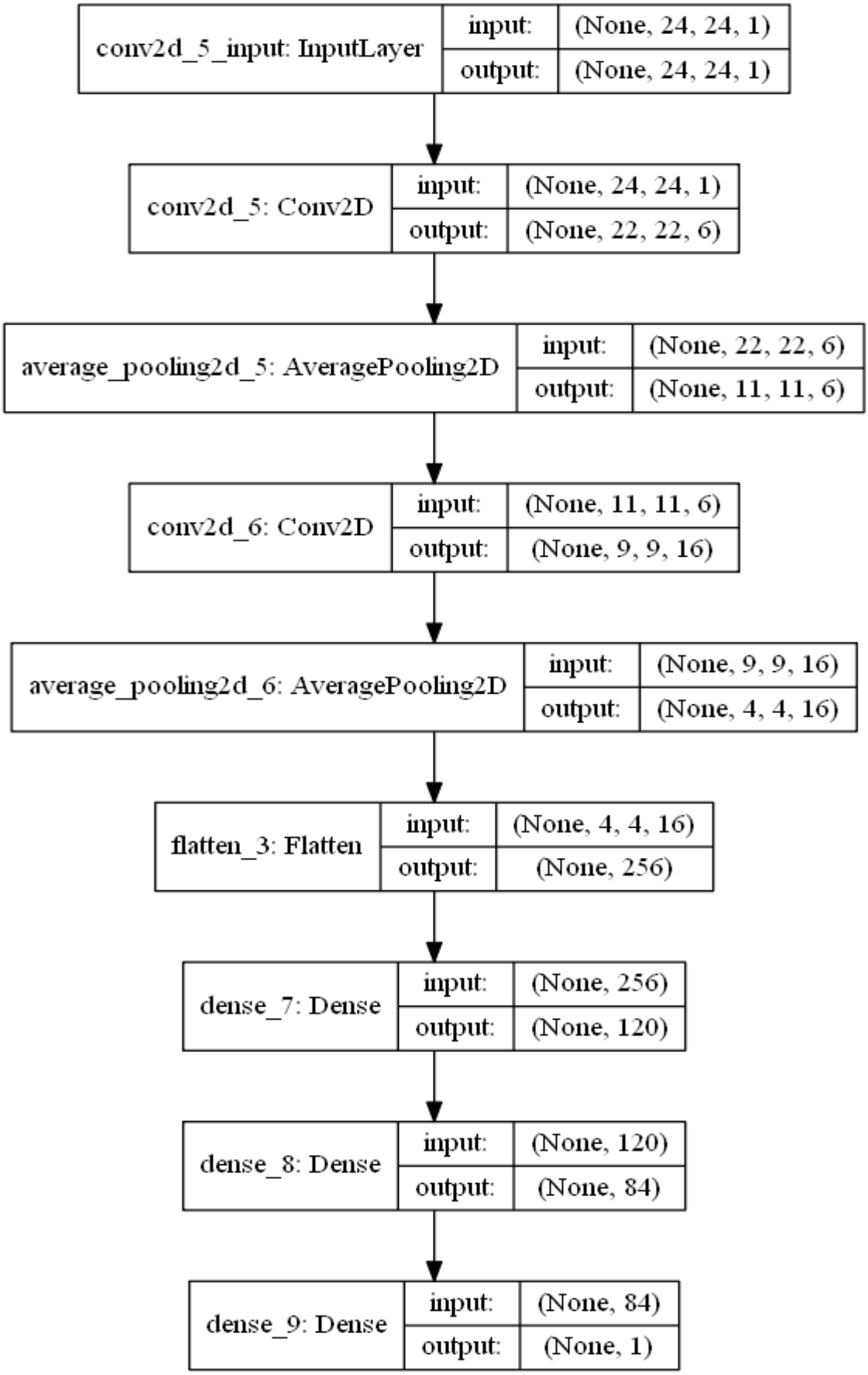
Današnja prosječna kamera može lako snimati video lica s ne manje od 24 fps (sličica u sekundi), tj. interval sličica nije veći od 42 milisekunde. Stoga je prosječnoj kameri lako

zabilježiti dvije ili više sličica za svaki treptaj kada lice gleda u kameru. Na temelju svega navedenog u ovom radu za sprečavanje lažiranja tj. za razlikovanje pravih i lažnih lica odabrano je praćenje treptanja očiju. Prednosti koje karakteriziraju ovaj pristup sprečavanja lažiranja su:

- Treptanje je lako uočljivo ljudsko ponašanje koje služi kao razlikovno obilježe živog lica od fotografije lice.
- Nenapadan pristup za koji nije potrebna svjesna suradnja korisnika što pruža bolje korisničko iskustvo.
- Pristup ne zahtijeva nikakav dodatan hardver i lako se integrira u postojeća rješenja.

Cilj je s ovim rješenjem prepoznati treptanje tj. obrazac otvoreno-zatvoreno-otvoreno oko. Kako bi se ovo rješenje moglo ostvariti potrebno je prvo istrenirati konvolucijsku neuronsku mrežu da prepoznaje između otvorenog i zatvorenog oka. Za treniranje je korišten skup podataka Closed Eyes In The Wild (CEW) [32]. Konkretno skup podataka sadrži, slike veličine 24x24, 2423 ljudskih subjekata, među kojima su 1192 subjekta s oba zatvorena oka, a 1231 subjekt s otvorenim očima. Treniranje modela konvolucijske neuronske mreže vršeno je putem Kerasa, a arhitektura trenirane konvolucijske neuronske mreže dana je na slici (Slika 3.9.). Evaluacijom istrenirani model je postigao točnost od 93%. Postupak treniranja modela mora se izvršiti samo jednom.

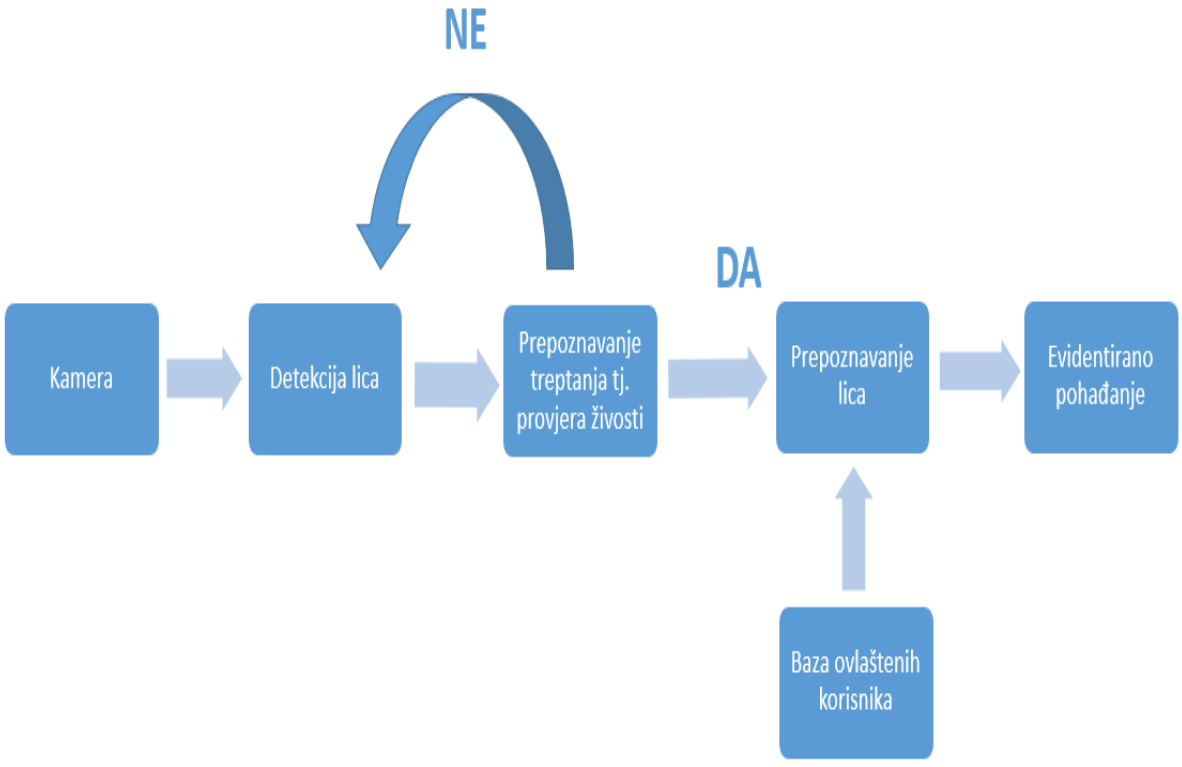
Prvi korak rada rješenja je detekcija lica na crno-bijeloj slici. Drugi korak je iz detektiranog i obrezanog lica detektirati oči (bilo lijevo ili desno oko ili oba oka). Za detekciju očiju se koristi Viola Jones metoda (2.4.1.). Nakon toga se detektirana slika oka smanjuje na veličinu 24x24 kako bi se mogla provući kroz prethodno trenirani model. Svaki put kada je otkriveno oko, predviđa se njegov status pomoću prethodno treniranog modela i prati se status očiju za svaku osobu. Ako je predviđanje < 0,15, tada se pretpostavlja da je oko u zatvorenom stanju. Ako je predviđanje > 0,85, tada se pretpostavlja da je oko otvoreno. Ako status osobe predstavlja u nekom trenutku obrazac otvoreno-zatvoreno-otvoreno onda on ukazuje na treptanje oka (bilo za lijevo ili desno oko) ili treptanje očiju (lijevo i desno oko istovremeno). Prepoznavanjem treptanja može se zaključiti da je lice pravo tj. da je osoba živa.



Slika 3.9. Prikaz arhitekture modela treniranog za razlikovanje otvorenog i zatvorenog oka

3.5. Povezivanje rješenja iz poglavlja (3.3.) i poglavlja (3.4.) u sustav za evidenciju pohađanja

Nakon razumijevanja izrađenih rješenja iz poglavlja (3.3.) i poglavlja (3.4.) potrebno ih je povezati u konačni sustav za evidenciju pohađanja. Sustav za evidenciju pohađanja morao bi evidentirati pohađanje osobe samo ako se ona nalazi u sustavu i ako je stvarna tj. ne radi se o napadu lažiranja. Kako bi se postigao iznad opisan sustav na računalno efikasan način rješenje za razlikovanje pravih i lažnih lica (provjera živosti) implementirano je poslije detekcije lica tj. neposredno prije ostatka rješenja za prepoznavanje lica (Slika 3.10.). Tako je izbjegnuto ponavljanje radnji koje su potrebne za oba rješenja (3.3. i 3.4.). Ako treptanje nije prepoznato tj. lice nije živo ne ide se dalje s procesom prepoznavanja lica nego se sustav vraća na detekciju lica i očiju na novoj sličici (engl. *frame*) i ostaje unutar te petlje sve dok se ne prepozna uzorak otvoreno-zatvoreno-otvoreno oko tj. dok se ne prepozna treptanje. Nakon što je treptanje ostvareno sustav kreće na prepoznavanje lica i uspoređuje ga s ostalim licima u bazi na temelju embeddinga kao što je opisano u poglavlju (3.3.). Postoji li lice te osobe odranije u bazi sustav ga identificira te je njegovo pohađanje evidentirano.



Slika 3.10. Dijagram toka izrađenog sustava za evidenciju pohađanja preko prepoznavanja lica s provjerom živosti u realnom vremenu

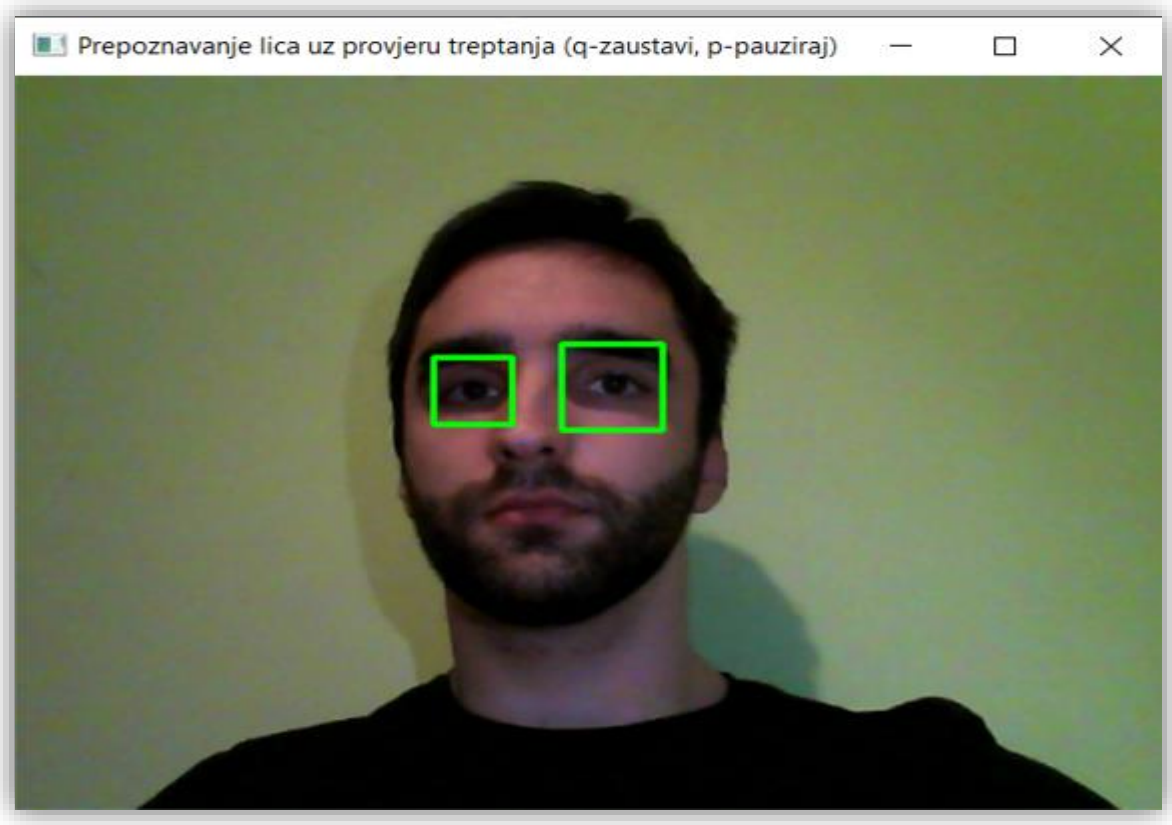
4. EKSPERIMENTALNI REZULTATI

Unutar ovog poglavlja će biti prikazani dobiveni rezultati. Treba napomenuti da je sustav testiran s više osoba, u raznim uvjetima osvjetljenja, raznim pozama lica te raznim udaljenostima. Uzimajući to u obzir u nastavku će biti prikazan samo mali broj izdvojenih slučajeva koji pokazuju suštinu izrađenog sustava.

Cilj je bio napraviti sustav za prepoznavanje lica koji treba što manje podataka za rad. Glavni razlog za ovo ograničenje je činjenica da je za korisnika ovog rješenja jednostavnije i brže trenirati model s jednom ili nekoliko slika za svaku osobu umjesto da mora napraviti veliki skup podataka s mnogo slika za istu osobu. Prije pokretanja sustava potrebno je u jednoj mapi ili bazi podataka napraviti podmapu koja nosi ime osobe te u nju staviti slike te osobe. To se radi kako bi sustav znao tko su ovlašteni korisnici te da bi im kasnije mogao evidentirati pohađanje.

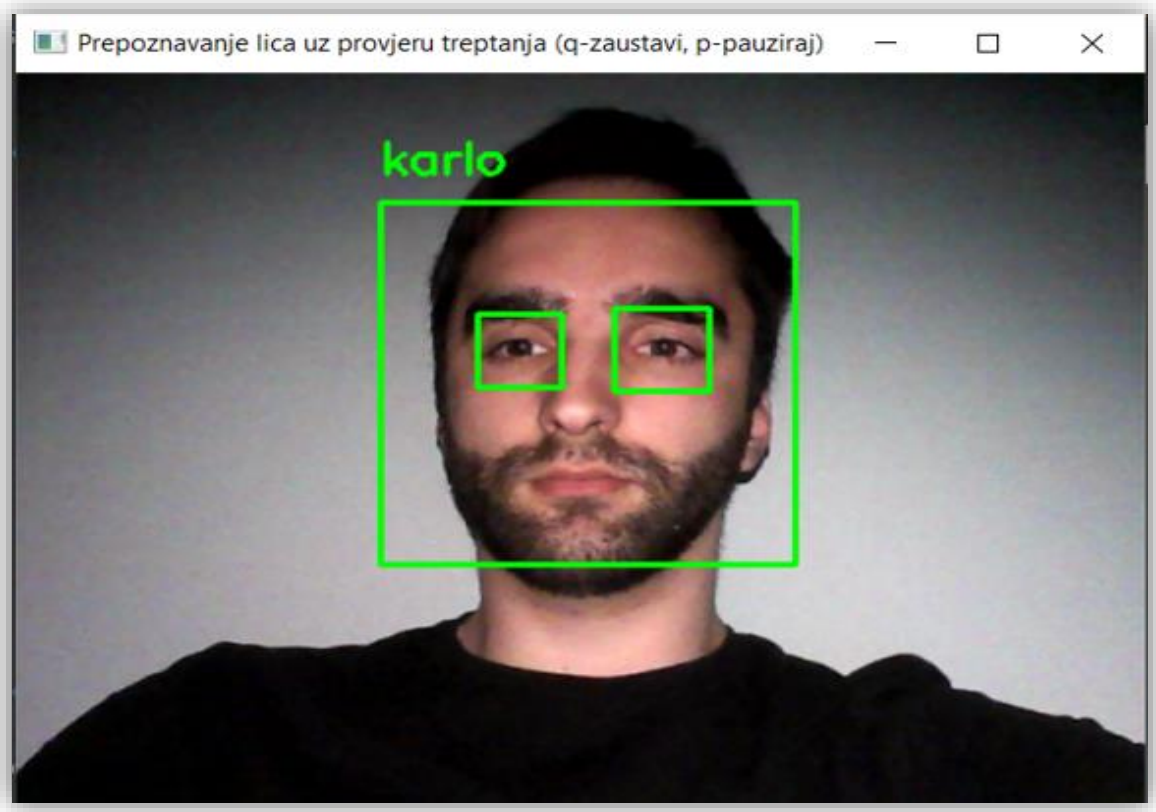
Pokretanjem sustava nudi se odabir da li se evidencija pohađanja želi provesti na videozapisu ili na web kameri. U nastavku će biti prikazan rad sustav na web kameri. Jedina razlika u odnosu na opciju videozapisa je ta da je tu ulazni podatak u sustav snimka u realnom vremenu iz web kamere, a kod opcije videozapis ulazni podatak sustava bi bio ranije snimljen videozapis. Odabirom opcije web kamera pokreće se kamera laptopa te se istovremeno prikazuju podatci iz web kamere na zaslonu laptopa. To se ostvaruje preko while petlje koja se postavi u True, a ona tj. sustav se može pauzirati pritiskom na tipku p ili prekinuti pritiskom na tipku q.

Dok se sve to odvija u pozadini se učitavaju potrebni modeli za detekciju lica i očiju, istrenirani modeli konvolucijskih neuronskih mreži itd. Nakon što se sve to odradi dobije se prikaz na zaslonu kao na slici (Slika 4.1). Iz tog prikaza je vidljivo da je sustav prvo detektirao lice te nakon toga dobro detektirao oči što je vidljivo iz zelenih pravokutnika koji obrubljuju oči.

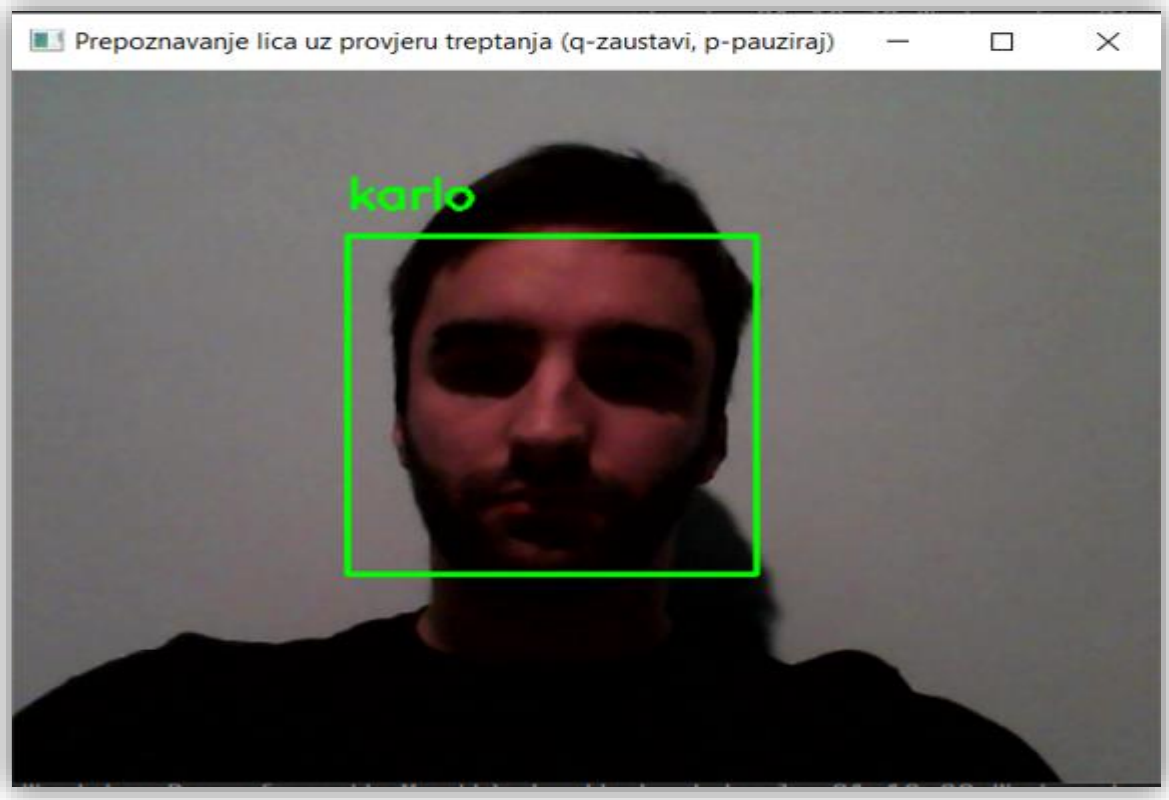


Slika 4.1. Prikaz rezultata detekcije očiju u realnom vremenu

Nakon što korisnik trepne sustav kreće s prepoznavanjem lica te ako je korisnik ovlašteni sustav ga identificira (Slika 4.2.) i (Slika 4.3.), od strane sustava korisnik se oslovi imenom te je obaviješten da mu je pohađanje evidentirano. Iz slika (Slika 4.2.) i (Slika 4.3.) vidljivo je da sustav pravilno identificira ovlaštene korisnike i to od izrazito povoljnih svjetlosnih uvjeta (Slika 4.2.) pa do blizu uvjeta mraka (Slika 4.3.). Što je samo okružje mračnije sustav sve teže prepoznaje oči jer je za detekciju očiju korišten jednostavniji detektor od onog korištenog za detekciju lica, a to sve je vidljivo iz slike (Slika 4.3.).



Slika 4.2. Prikaz rezultata prepoznavanja korisnika u povoljnim svjetlosnim uvjetima

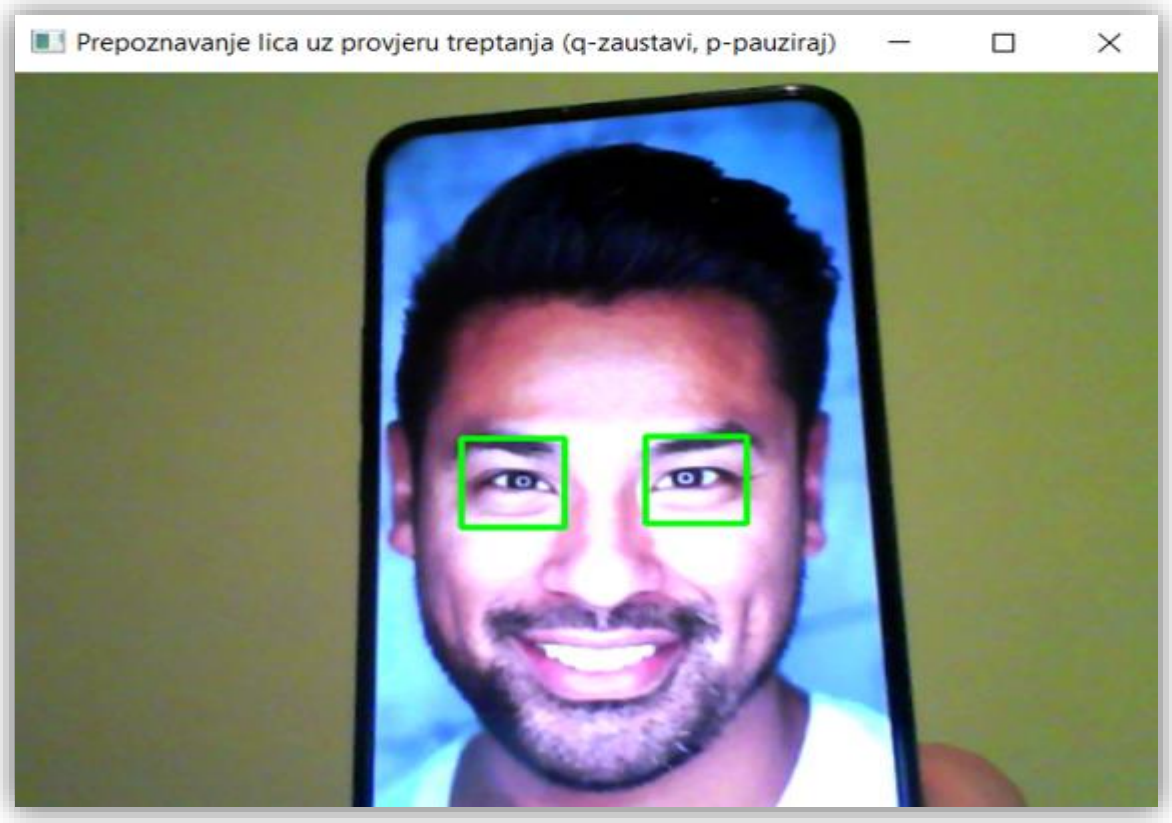


Slika 4.3. Prikaz rezultata prepoznavanja korisnika u nepovoljnim svjetlosnim uvjetima

Kad se određena sesija završi generira se Excel dokument evidencije pohađanja (Slika 4.4.). Kako bi se sustav testirao na razlikovanje između pravih i lažnih lica tj. na provjeru živosti ispred njega je stavljena slika korisnika na mobitelu koji je prethodno unesen u bazu kao ovlašteni, te se tako pokušao zavarati sustav (Slika 4.5.). Usprkos tome iz slike (Slika 4.5.) je vidljivo da je sustav dobro detektirao oči i sukladno tome lice osobe. Na temelju toga vidi se da rješenje za sprečavanje lažiranja unutar sustava dobro detektira kad se ne radi o pravom licu jer osoba u tom slučaju ne trepće.

	A	B	C	D	E
1	Name	Time	Weekday	Day of month	Month
2					
3	karlo	21:51:42	Wednesday	29	June

Slika 4.4. Prikaz generiranog izvješća evidencije pohađanja u Excel datoteci

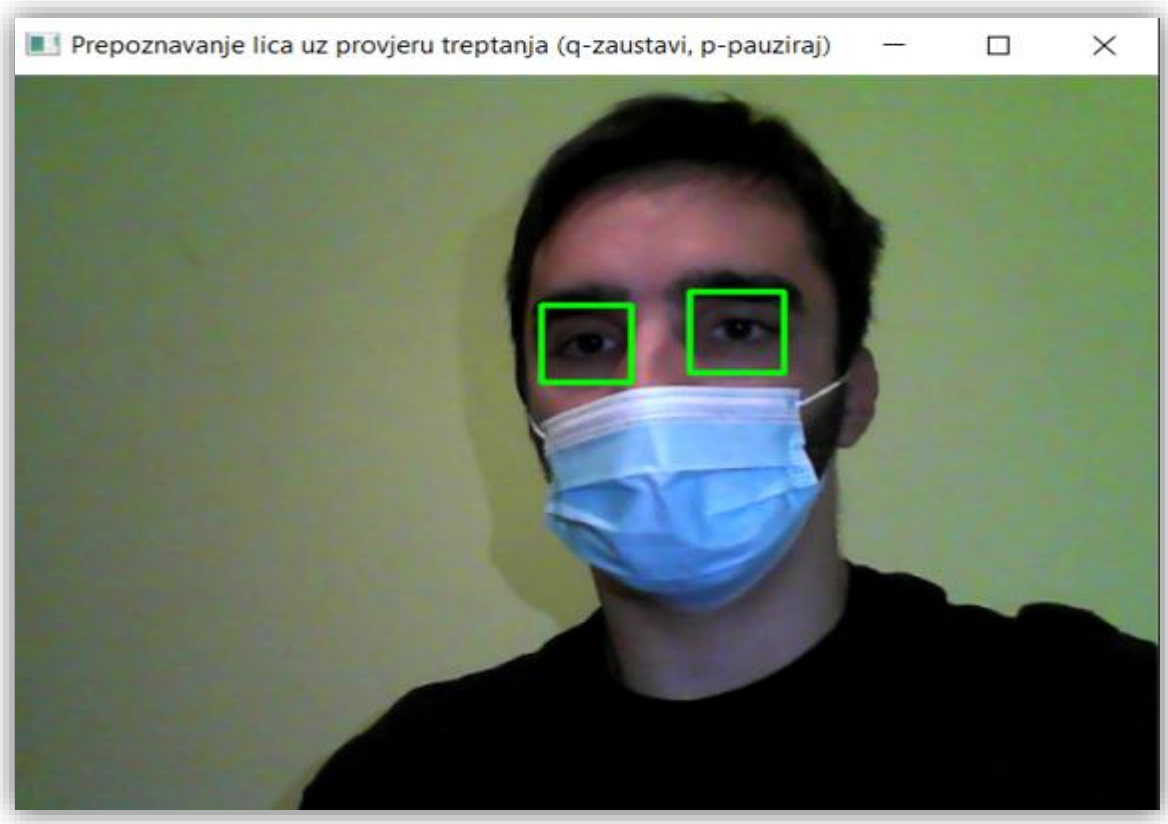


Slika 4.5. Prikaz rezultata dobivenog za pokušaj lažiranja pohađanja korištenjem fotografije ovlaštenog korisnika na mobitelu

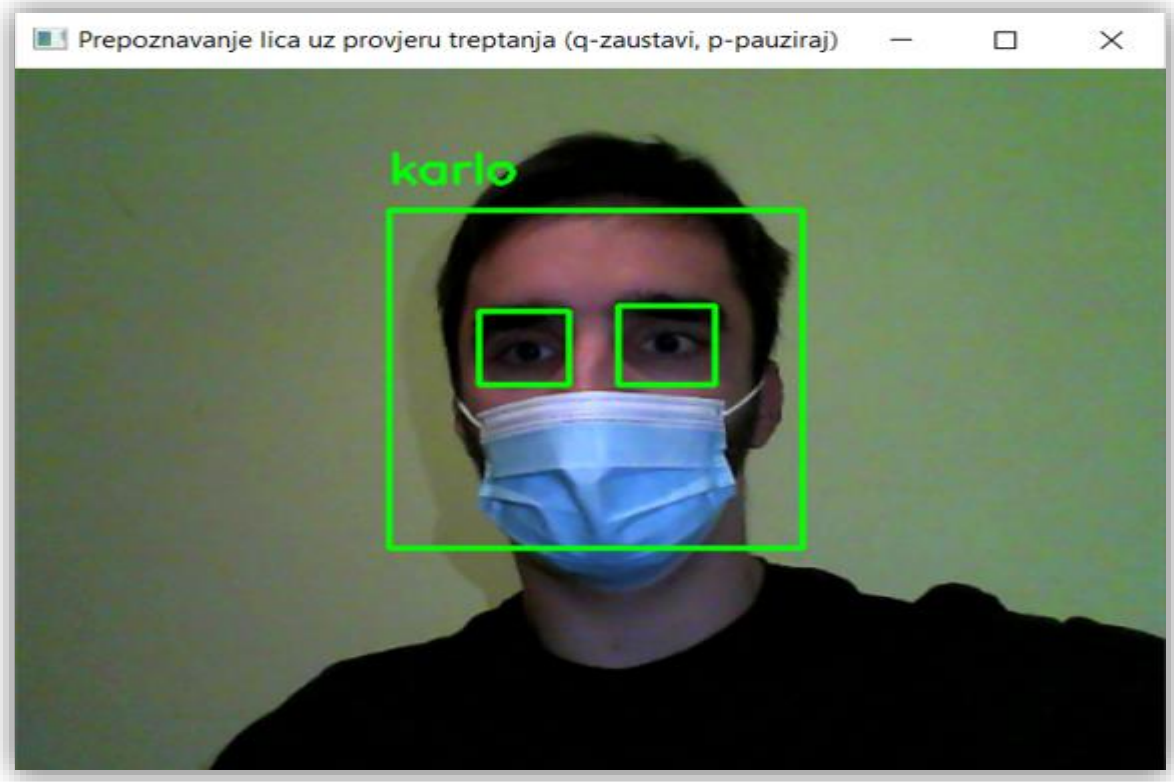
Sljedeći slučaj kojim se pokušao zavarati sustav je sadržavao istovremeno u kadru dva ovlaštena korisnika od kojih je jedan bio stvarna osoba, a za drugog ovlaštenog korisnika bila je prikazana fotografija lica na mobilnom uređaju (Slika 4.6.). Iz slike je vidljivo da je sustav opet uspio dobro diferencirati između pravog i lažnog lica osobe i to istovremeno kad je uz to lice bilo prisutno i stvarno lice drugog ovlaštenog korisnika.

Slika 4.6. Prikaz rezultata dobivenog za pokušaj lažiranja pohađanja korištenjem fotografije ovlaštenog korisnika na mobitelu uz istovremeno prisustvu druge stvarne ovlaštene osobe

Kod testiranja samog sustava došlo se do zanimljivog i neočekivanog svojstva sustava koje je prikazano kroz slike (Slika 4.7.) i (Slika 4.8.). Sustav je uspio prepoznati ovlaštenog korisnika usprkos tome što je nosio masku na licu i time prekrivao većinu karakteristika lica. Na temelju tih karakteristika sustav vrši generiranje embeddinga lica preko kojeg uspoređuje osobu s osobama koje se nalaze u bazi. To je neočekivano svojstvo jer kod izrade rješenja za prepoznavanje lica nigdje nije bio dan skup podataka na kojem su ljudi nosili maske na licu.



Slika 4.7. Prikaz rezultata dobivenog za slučaj nošenja plave maske prije treptanja



Slika 4.8. Prikaz rezultata dobivenog za slučaj nošenja plave maske nakon treptanja

Kao što je na početku ovog poglavlja napomenuto ovdje su predstavljeni probrani slučajevi testiranja sustava preko kojih se htjelo prikazati osnove mogućnosti sustava. Iz njih je vidljivo da sustav može nastaviti ispravno funkcionirati u različitim intenzitetima svjetla, različitim udaljenostima korisnika od kamere te pritom može i spriječiti napade lažiranja fotografijama bilo fotokopiranom fotografijom bili fotografijom prikazanom na mobilnom uređaju. Što u gornjim primjerima slučajeva nije prikazano su napadi lažiranja izvedeni prikazivanjem videozapisa korisnika mobilnim uređajem na kojima korisnik trepće te napadi lažiranja gdje je druga osoba stavila realističnu hiperelastičnu masku druge osobe koja je ovlašteni korisnik. Za te navedene slučajeve sustav ne bi uspio uočiti da se radi o napadu lažiranja tj. identificirao bi u oba slučaja da se radi o pravoj osobi. Ti napadi mogli bi se spriječiti kombiniranjem trenutnog rješenja s ostalim rješenjima navedenim u poglavlju (2.8.). Kod slučajeva kad je lice korisnika u nekom vrlo nepovoljnom položaju, npr. nagnuto jako u lijevo ili desnu stranu, sustav tada ne može uopće detektirati lice, a još manje oči pa se koraci sustava koji slijede nakon njega ne mogu odviti. Uzimajući ovo sve iznad spomenuto u obzir može se donijeti zaključak da je sustav evidencije pohađanja sposoban za korištenje u stvarnom vremena i uz to može ispravno identificirati istovremeno više osoba te spriječiti napade lažiranja fotografijama.

5. ZAKLJUČAK

Prepoznavanje lica je trenutno među najproduktivnijim primjenama obrade slika i ima ključnu ulogu kako u industriji tako i znanstvenom istraživanju. Prepoznavanje ljudskog lica je aktualno pitanje u svrhu identifikacije posebno u kontekstu pohađanja nastave. Sustav evidencije pohađanja pomoću prepoznavanja lica je postupak prepoznavanja učenika korištenjem karakteristika lica koje razlikuju osobe međusobno. Sustav izrađen unutar ovog rada ima za cilj ostvariti digitalizaciju i automatizaciju trenutnog sustava koji je baziran na javljanju nakon prozivanja i vođenju evidencije pohađanja na papiru. Trenutne strategije za evidenciju pohađanja karakterizira to da uzimaju puno vremena i pritom su podložne raznim greškama. Također, trenutne strategije su ranjive na lažiranje pohađanja. Stoga se ovaj rad pozabavio svim tim problemima.

Unutar teorijskih osnova rada detaljno su razrađena i opisana područja biometrije, prepoznavanje lica, konkretni koraci unutar sustava prepoznavanja lica i moguće vrste lažiranja sustava prepoznavanja lica. U praktičnom dijelu ovog rada predstavljen je novi pristup evidencije pohađanja prepoznavanjem lica u stvarnom vremenu s implementiranim sprečavanjem lažiranja pohađanja. Unutar rješenja za prepoznavanje lica korišteni su HOG detektor lica, estimacija značajki lica, generiranje 128 dimenzionalnog vektora raznih mjera lica koje služe za razlikovanje između osoba te SVM klasifikator. Za sprečavanje lažiranja pohađanja trenirana je konvolucijska neuronska mreža koja razlikuje otvoreno i zatvoreno oko te se na temelju njenih predviđanja prati ako osoba trepće. Taj pristup je odabran jer je neinvazivan te ne zahtijeva dodatni napor od strane korisnika. Izvještaj evidencije pohađanja automatski se generira u obliku Excel datoteke na kraju sesije od strane sustava.

Sustav je testiran u raznim uvjetima kao što su varijacije osvjetljenja, razni zakreti glave i varijacije udaljenosti od kamere. Dobiveni rezultati pokazuju da je izrađeni sustav evidencije pohađanja poprilično efikasan i robustan, a da istovremeno štedi vrijeme i smanjuje potrebu za manualnim radom. Sustav bez problema otkriva pokušaje lažiranja fotografijama bilo da su one printane ili prikazane na zaslonu mobitela. Usprkos tome sustav je još uvijek podložan napadima u obliku videozapisa i hipereleastičnih maski lica ovlaštenih osoba. Zanimljivo opažanje iz rezultat je da sustav prepoznaje ovlaštenog korisnika i sa zaštitnom maskom iako unutar rješenja za prepoznavanje lica nije korišten skupa podataka osoba koje nose masku što bi moglo biti korisno za primjenu u doba koronavirusa ili neke druge pandemije.

Sposobnost sustava za izvođenje u stvarnom vremenu čini ga prikladnim za prepoznavanje ljudi na CCTV snimkama u trgovačkim centrima, zdravstvenim ustanovama, zračnim lukama, obrazovnim ustanovama, bankama i drugim visoko sigurnosnim zonama koje zahtijevaju odobrenje pristupa. Unaprjeđenja ovog rada mogu se ostvariti kroz nekoliko aspekata. Jedan od njih bi bio implementacija rješenja koje sprječava napade lažiranja pohađanja videozapisima ili hiperelastičnim maskama. Primjeri takvog rješenja bili bi upotreba aktivnog bljeska, upotreba 3D kamere, treniranje nove konvolucijske neuronske mreže da razlikuje teksturu lica i teksturu zaslona ili njihove kombinacije. Sustava bi se također mogao proširiti s prepoznavanjem emocija kako bi se ostvarila bolja interakcija s korisnikom. Te za kraj sustav bi mogao biti implementiran za rad na mobitelima.

LITERATURA

[1] Curse of Dimensionality, https://towardsdatascience.com/curse-of-dimensionality-a-curse-to-machine-learning-c122ee33bfeb, Pristupljeno: 30.4.2022.

[2] Applications of Facial Recognition Technology, https://www.trio.dev/blog/facial-recognition-applications, Pristupljeno: 30.4.2022.

[3] Biometrics: definition, use cases, latest news, https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/biometrics, Pristupljeno: 2.5.2022.

[4] Biometrics, https://www.techtarget.com/searchsecurity/definition/biometrics, Pristupljeno: 2.5.2022.

[5] Types of Biometrics, https://www.biometricsinstitute.org/what-is-biometrics/types-of-biometrics/, Pristupljeno: 5.5.2022.

[6] Biometric recognition and authentication systems, https://www.ncsc.gov.uk/collection/biometrics/, Pristupljeno: 7.5.2022.

[7] Biometric Devices — Complete Guide on Technology, https://recfaces.com/articles/articles-biometric-devices, Pristupljeno: 11.5.2022.

[8] Iris recognition edges fingerprint, face and palm biometrics for most willing use by UK consumers, https://www.biometricupdate.com/201912/iris-recognition-edges-fingerprint-face-and-palm-biometrics-for-most-willing-use-by-uk-consumers Pristupljeno: 13.5.2022.

[9] Fujitsu: The future of biometrics, https://www.banking-gateway.com/projects/fujitsu-the-future-of-biometrics-dr-joseph-reger/fujitsu-the-future-of-biometrics-dr-joseph-reger2.html, Pristupljeno: 15.5.2022.

[10] What is facial recognition? How facial recognition works, https://us.norton.com/internetsecurity-iot-how-facial-recognition-software-works.html, Pristupljeno: 25.5.2022.

[11] Challenges Faced by Facial Recognition System, https://www.pathpartnertech.com/challenges-faced-by-facial-recognition-system/, Pristupljeno: 27.5.2022.

[12] One-shot learning, https://en.wikipedia.org/wiki/One-shot_learning, Pristupljeno: 27.5.2022.

[13] What is Facial Recognition – Definition and Explanation, https://www.kaspersky.com/resource-center/definitions/what-is-facial-recognition

Pristupljeno: 29.5.2022.

[14] Face Recognition, https://www.eff.org/pages/face-recognition, Pristupljeno: 2.6.2022.

[15] Face Recognition Pipeline Clearly Explained, https://medium.com/backprop-labs/face-recognition-pipeline-clearly-explained-f57fc0082750, Pristupljeno: 3.6.2022.

[16] face detection, https://www.techtarget.com/searchenterpriseai/definition/face-detection, Pristupljeno: 4.6.2022.

[17] Face Detection For Beginners, https://towardsdatascience.com/face-detection-for-beginners-e58e8f21aad9, Pristupljeno: 5.6.2022.

[18] Face Detection using Viola Jones Algorithm, https://www.mygreatlearning.com/blog/viola-jones-algorithm/, Pristupljeno: 7.6.2022.

[19] Histograms Of Oriented Gradients for Human Detection, https://medium.com/@ktiwary2/scattered-thoughts-on-ml-68d30f44da19, Pristupljeno: 9.6.2022.

[20] Multi-task Cascaded Convolutional Networks (MTCNN) for Face Detection and Facial Landmark Alignment, https://medium.com/@iselagradilla94/multi-task-cascaded-convolutional-networks-mtcnn-for-face-detection-and-facial-landmark-alignment-7c21e8007923, Pristupljeno: 15.6.2022.

[21] Face Recognition, https://bhashkarkunal.medium.com/face-recognition-real-time-webcam-face-recognition-system-using-deep-learning-algorithm-and-98cf8254def7, Pristupljeno: 19.6.2022.

[22] Anti-Spoofing Techniques For Face Recognition Solutions, https://towardsdatascience.com/anti-spoofing-techniques-for-face-recognition-solutions-4257c5b1dfc9, Pristupljeno: 22.6.2022.

[23] Facial Recognition: Types of Attacks and Anti-Spoofing Techniques, https://towardsdatascience.com/facial-recognition-types-of-attacks-and-anti-spoofing-techniques-9d732080f91e, Pristupljeno: 23.6.2022.

[24] PyCharm, https://en.wikipedia.org/wiki/PyCharm, Pristupljeno: 24.6.2022.

[25] Python Tutorial, https://www.tutorialspoint.com/python/index.htm, Pristupljeno: 24.6.2022.

[26] What Is OpenCV?, https://www.oreilly.com/library/view/learning-opencv/9780596516130/ch01.html, Pristupljeno: 24.6.2022.

[27] Keras Tutorial, https://www.javatpoint.com/keras, Pristupljeno: 24.6.2022.

[28] V. Kazemi, and J. Sullivan, "One millisecond facealignment with an ensemble of regression trees,"Proceedings of the IEEE Conference on ComputerVision and Pattern Recognition, pp. 1867-1874,2014.

[29] Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning, https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78, Pristupljeno: 27.6.2022.

[30] Openface, https://github.com/cmusatyalab/openface/tree/master/models/openface Pristupljeno: 28.6.2022.

[31] How Fast is the Average Blink?, https://www.somatechnology.com/blog/thursday-thoughts/fast-average-blink/, Pristupljeno: 29.6.2022.

[32] Closed Eyes In The Wild (CEW), http://parnec.nuaa.edu.cn/_upload/tpl/02/db/731/template731/pages/xtan/ClosedEyeDatabases.html, Pristupljeno: 30.6.2022.