

Sustav za kontrolu pristupa temeljen na prepoznavanju lica

Jagodić, Valentino

Master's thesis / Diplomski rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:225710>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-02-06**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA**

Sveučilišni studij

**SUSTAV ZA KONTROLU PRISTUPA TEMELJEN NA
PREPOZNAVANJU LICA**

Diplomski rad

Valentino Jagodić

Osijek, 2020.

Sadržaj

1. UVOD	1
1.1. Zadatak diplomskog rada	1
2. PREGLED RAZVOJA SUSTAVA ZA PREPOZNAVANJE LICA.....	2
2.1. Povijest razvoja sustava za prepoznavanje lica	3
2.2. Suvremena tehnološka postignuća sustava za prepoznavanje lica	7
3. KORIŠTENE TEHNOLOGIJE.....	9
3.1. Spyder.....	9
3.2. Python.....	9
3.3. OpenCV.....	10
3.4. TensorFlow.....	10
4. PROGRAMSKO RJEŠENJE.....	11
4.1. Pronalazak lica na zadanom mediju	11
4.1.1. Sastavljanje i treniranje neuronske mreže	12
4.1.2. Predobrada podataka	13
4.1.3. Metoda klizajućeg prozora	18
4.2. Točke interesa lica.....	21
4.3. Prepoznavanje osobe na slici.....	22
5. USPOREDBA POSTOJEĆIH RJEŠENJA	25
5.1. Viola – Jones klasifikator temeljen na Haar značajkama	25
5.2. SVM klasifikator istreniran na HOG podatcima	25
5.3. Klasifikator temeljen na CNN modelu	26
5.4. Klasifikator temeljen na MTCNN modelu	26
5.5. Usporedba metoda pronalaženja lica.....	27
5.6. Testiranje algoritma prepoznavanja lica.....	30
6. ZAKLJUČAK	32
LITERATURA.....	33
SAŽETAK.....	36
ABSTRACT	37
ŽIVOTOPIS	38
PRILOZI.....	39

1. UVOD

Razvojem suvremene tehnologije nastale su i različite mogućnosti ugrožavanja tradicionalnog načina osiguravanja ljudi i materijalnih dobara unutar pristupno ograničenih i informacijski osjetljivih područja. Budućnost i moderni svijet donijeli su raznolika unaprjeđenja, što uključuje i različite načine komprimiranja sigurnosti, stoga je nastala potreba za unaprjeđenjem sigurnosnih sustava. Jedan od načina kontrole pristupa korištenje je tehnologije prepoznavanja lica.

U ovom diplomskom radu riječ je o aplikaciji za kontrolu pristupa koji se temelji na prepoznavanju lica. Aplikacija je namijenjena za rad na većini operacijskih sustava koji su poznati danas, što uključuje Microsoft Windows i Linux. Zadaća navedene aplikacije je iz vizualnog materijala (slika ili video) pronaći i zaokružiti lice korisnika te odrediti njihov identitet iz baze podataka lica i naposljetku, razinu njegovog pristupa objektu.

Drugo poglavlje prikazuje razvoj sustava za prepoznavanje lica kroz povijest njegove primjene i aktualne najsuvremenije njegove implementacije. Treće poglavlje diplomskog rada daje pregled i informacije o razvojnim alatima korištenim pri razvoju sustava za kontrolu pristupa, poput programskog jezika Python, koji se nalazi unutar programskog razvojnog okruženja Spyder, programskih paketa TensorFlow i OpenCV, koji su bili neophodni za razvoj ovog programskog sustava. Četvrto poglavlje detaljnije objašnjava programsko rješenje zadanog sustava i različitih problema koje pokriva. Posljednje, peto poglavlje donosi rezultate postojećih rješenja i njihovu usporedbu.

1.1. Zadatak diplomskog rada

Zadatak diplomskog rada je istražiti načine rada postojećih algoritama za prepoznavanje lica te ih usporediti i razviti vlastiti sustav za kontrolu pristupa temeljen na određenim algoritmima prepoznavanja lica. Aplikacija treba odrediti lokaciju lica sa slike te označiti gdje se lice nalazi. Sljedeći korak je prolaženje kroz bazu podataka lica te uspoređivanje pronađenog lica s licima iz baze podataka koristeći različite algoritme koji će donijeti najbolje rezultate.

2. PREGLED RAZVOJA SUSTAVA ZA PREPOZNAVANJE LICA

Današnji svijet tehnologije sa svakim danom doseže svoje nove vrhunce. Također, tehnologija je prisutna u svim životnim aspektima, krenuvši od računarstva, medicine, sustava za zaštitu pa sve do sustava za istraživanje svemirskih prostranstava. Tehnologija se nalazi na svakom koraku kako bi olakšala čovjeku obavljanje svakodnevnih zadataka i osigurala napredak. Međutim, svaki aspekt čovjekovog života podvrgnut je propitivanju i čovjekovim testovima moralnih vertikalna pa je to i slučaj s tehnologijom. Jedan od najvećih tehnoloških napredaka koji se postavlja u moralno sivu zonu razvoj je sustava za prepoznavanja lica, posebice na području sustava za osiguranje posjeda, materijalnih i drugih vrsta dobara. Uz upotrebu u svrhe sigurnosti, sustavi za prepoznavanje lica upotrebljavaju se i za nadzore zračnih luka za prepoznavanje mogućih prijetnja od terorizma i za sprječavanje kriminalnih radnji, česta je upotreba i u vojnim svrhama, geologiji, poljoprivredi, policijskim radnjama i čak današnjim pametnim telefonima te sličnim poljima interesa. [1]

Samo prepoznavanje lica definirano je kao računalni program, odnosno sustav, za automatsko prepoznavanje i provjeru osoba iz određenog vizualnog medija, najčešće upotrebljavanih digitalnih slika ili video izvora. Kako je prethodno navedeno, često se koristi u sigurnosnim sustavima te ga je kao računalni sustav moguće usporediti s otiscima prstiju ili skenerom zjenice oka. [2] Prepoznavanje lica ima prednost nad otiscima prstiju ili skenerom zjenice oka u smislu unikatnosti i prilagodljivosti svog objekta, to jest, lica osobe. Međutim, problem nastaje baš u toj unikatnosti čovjekovog lica, njegovoj velikoj dinamici, gdje se često žrtvuju preciznost i brzina rada ovakvih sustava. [3] Također, neki od problema su velika količina ljudi na jednom mjestu, pogotovo na javnim mjestima, gdje sustavi mogu imati poteškoće s prepoznavanjem individuala uklopljenih u mase te loše kvalitete fotografija ili video materijala dobivenih sa zaštitnih kamera. Pojednostavljen način za objašnjenje kako ovakvi sustavi rade je da se sustav trenira na velikim količinama podataka kako bi mogao razaznati značajke lica i međusobno ih uspoređivao kako bi naposljetku pronašao podatke o osobi na slici.

Kako je riječ o značajkama lica svake individualne osobe, važno je uočiti kako lica nose veliku količinu informacija o osobi, kao što su: dob, spol, unutarnji osjećaji i emotivno stanje osobe, ponekad je moguće zaključiti, pogotovo u modernim vremenima čak i ekonomsko stanje pojedinca. „Čitanje“ izraza lica i promatranje značajki lica ljudima je urođena vještina i osnova na kojoj se razvilo društvo i vrlo važan dio društvenih interakcija. [4] Lice je jedinstveno za svaku osobu i ne postoje dva potpuno identična lica, čak ni u slučaju jednojajčanih blizanaca uvijek je

moguće naći čak i minorne razlike, zbog čega dolazi do stavljanja sustava prepoznavanja lica u moralno sivu zonu.

Problem privatnosti nastaje jer mnoga javna mjesta često imaju nadzorne kamere i poznato je kako je sustav za prepoznavanje lica igrao važnu ulogu u održavanju sigurnosti jer mu nije potrebna suradnja objekta kojeg promatra, u ovom slučaju, čovjeka koji prolazi tim javnim mjestom. Nedostatak dozvole pojedinca čija se slika ili video uzima problem je koji je teško riješiti jer je privatnost čovjekovo pravo, ali je također i pravo na sigurnost od primjera radi, terorizma, koji je moguće spriječiti puno brže ako su postavljene kamere i sustavi za prepoznavanje lica. Međutim, Norton navodi razloge zašto bi sustavi za prepoznavanje lica predstavljali opasnost za privatnost pojedinca, a neki od navedenih razloga su sigurnost, jer ako kamera negdje uslika ili snimi čovjeka, njegova slika ili video može završiti u bazi podataka iz koje je hakeri mogu ukrasti ili zloupotrijebiti. Također, prosječnom čovjeku nije poznato gdje će završiti njegova fotografija i tko će je vidjeti. Naposljetku, svačije lice je osobno vlasništvo svakog čovjeka i ima pravo na njegovu zaštitu, za što se i moguće odlučiti na nekim društvenim mrežama. [5]

2.1. Povijest razvoja sustava za prepoznavanje lica

Sve do nedavno, sustavi za prepoznavanje lica činili su se kao sama fikcija i kao da su se pojavili niotkud, od sigurnosnih sustava na aerodromima, javnim mjestima i slično. Međutim, sustavi za prepoznavanje lica nisu postojali samo kao ideja modernog doba, već njihova povijest seže sve do šezdesetih godina 20. stoljeća.

Mnogi tvrde kako je otac sustava za prepoznavanja lica Woodrow Wilson Bledsoe, američki matematičar, profesor i jedan od znanstvenika koji je najviše pridonio razvoju umjetne inteligencije. [6] Bledsoe je 1960-ih razvio sustav kojim bi se ručno moglo klasificirati fotografije lica poznatiji pod imenom RAND tablet (slika 2.1.) RAND tablet uređaj je koji se mogao upotrebljavati za unos vertikalnih i horizontalnih koordinata na koordinatnoj mreži koristeći posebnu olovku koja je emitirala elektromagnetne valove. Navedeni sustav mogao je biti korišten kako bi se ručno zabilježile koordinate različitih značajki lica, uključujući oči, nos, usta i slično. Navedena mjerenja bila bi potom pohranjena u bazu podataka, što bi značilo kad bi sustav dobio novu fotografiju lica, mogao je izvući fotografiju iz baze podataka koja je najviše nalikovala osobi na toj fotografiji. Naravno, s obzirom na to da se radilo o 1960-ima, sustav za prepoznavanje lica bio je vrlo ograničen tehnološkim napredcima tog doba te je označilo prvi korak u razvoju ovakvih sustava. [7]



Slika 2.1. Woodrow Wilson Bledsoe RAND tablet tehnologija [8]

Idući napredak na sustavu koji je razvio Bledsoe bio je deset godina poslije, u 1970-ima, kad su Goldstein, Harmon i Lesk povećali razinu preciznosti Bledsoevog sustava tako što su dodali 21 točku koje su služile kao označitelji za najvažnije dijelove lica po kojima je lice prepoznatljiv oblik, poput debljine usana i boje kose. [9]

Poslije napretka u 1970-ima, 1988. Sirovich i Kirby počeli su raditi na sustavu za prepoznavanje lica tako što su primijenili linearnu algebru u problemu prepoznavanja lica, koji je nazvan *Eigenfaces*. Njihova ideja bila je pokazati da analiza značajki na skupini fotografija lica može formirati skupinu s osnovnim značajkama lica (slika 2.2.). Također, izračunali su najbolji koordinatni sustav za kompresiju slike, gdje je svaka koordinata zapravo slika, zbog čega se razvio pojam *eigenpicture*. Tri godine poslije, Turk i Pentland proširili su ideju *Eigenface* sustava otkrivanjem kako prepoznati lica na fotografijama, što je dovelo do prvih ideja o automatiziranom sustavu prepoznavanja lica. [9]



Slika 2.2. Prikaz *Eigenface* sustava, originalna slika s lijeve strane i projekcija *Eigenface* na desnoj strani [9]

Učeći na greškama svojih prethodnika, DARPA (*The Defense Advanced Research Projects Agency*) i Nacionalni institut za tehnologiju izbacili su početkom 90-ih godina FERET (*Facial Recognition Technology*), to jest, tehnologiju za razvoj sustava za prepoznavanje lica, gdje su došli na ideju stvaranja velike baze podataka. Baza podataka imala je 3323 frontalne fotografije i pokrenula je ideju o velikim bazama podataka koje bi mogle inspirirati daljnje inovacije i pokrenuti napredak u tehnologiji sustava za prepoznavanje lica. [10]

Korištenje sustava za prepoznavanje lica u zaštitne svrhe najizraženija je bila za vrijeme 35. *Super Bowl* utakmice koja se dogodila 2002. godine, gdje su zaštitarske službe koristile sustave za prepoznavanje lica kao najveći i najvažniji test ove tehnologije. Međutim, sustav nije bio dovoljno razvijen do tada i bio je poprilično neuspješan, ponajviše zbog prepoznavanja objekata kao lica i manjkavosti u brzini i preciznosti, s čime se zaključilo da još nije spreman za praktičnu upotrebu. Najveća tehnološka ograničenost tog vremena u sustavu za prepoznavanje lica bila je nemogućnost rada s velikim masama koje su bile važne za velike događaje poput *Super Bowl*. [7]

Od 2009. situacija počinje biti bolja za sustave za prepoznavanje lica jer su se počeli primjenjivati u praktičnim svrhama, poput primjera Pinellas predjela u Floridi, Sjevernoj Americi, gdje su stručnjaci napravili forenzičku bazu podataka koja je dopuštala policijskim službenicima pristup arhivama fotografija i kamera s glavnih prometnica. Do 2011. godine, broj kamera i opremljenih policijskih postaja značajno se povećao i omogućio daleko veći broj uhićenih i pokrenutih kriminalnih istraga nego ikada prije. [7]

Posljednja etapa, koja je počela 2010. godine i trenutno traje, obilježena je korištenjem naprednih sustava za prepoznavanje lica na društvenim mrežama, počevši s Facebookom. Facebook je uveo prepoznavanje lica na fotografijama kako bi se prepoznali ljudi na njima koji su „označeni“ na njima (slika 2.3.), što je izazvalo kontroverzu u medijima jer se smatralo kako se ugrožava privatnost njegovih korisnika. [7]



Slika 2.3. Prikaz Facebookovog sustava za prepoznavanje lica na fotografijama [11]

Nakon početka upotrebe sustava za prepoznavanje lica na društvenim mrežama, najvažnija javna upotreba zabilježena je 2011. godine, gdje se u Panami uspješno implementirao sustav za prepoznavanje lica u stvarnom vremenu. Aerodrom u Panami imao je problema s kriminalom te je sustav za prepoznavanje lica uvelike pomogao policijskim službama za suzbijanje kriminala i čak rezultiralo hvatanjem Interpolovih osumnjičenika. Od 2017. godine porastao je trend korištenja sustava za prepoznavanje lica u prodaji, kako bi se stvarao profil kupca i osigurala sigurnost prodajnih mjesta. Međutim, daleko najveći napredak iste godine primijetio se u industriji mobilnih uređaja, gdje je tvrtka Apple po prvi put u mobilne uređaje ugradila sustav za otključavanje mobitela koristeći prepoznavanje lica, što je poslije postao trend i kod drugih proizvođača mobilnih uređaja. Ovakva tehnologija smatra se „*state of the art*“ tehnologijom u ovom području. [7]

2.2. Suvremena tehnološka postignuća sustava za prepoznavanje lica

Dostignuća današnjice na području biometrije bilježe znatan skok, što prvenstveno omogućuje razvoj visokih tehnologija, dostupnost komponenti i znanja široj publici i svima onima koji bi htjeli dublje istražiti svijet sigurnosnih sustava koristeći metode za identifikaciju osoba. Danas je moguće prepoznati neke od najnaprednijih sustava za prepoznavanje lica, od kojih su istaknuti GaussianFace algoritam, DeepFace, Rekognition, LFIS sustav i FaceNet. [12]

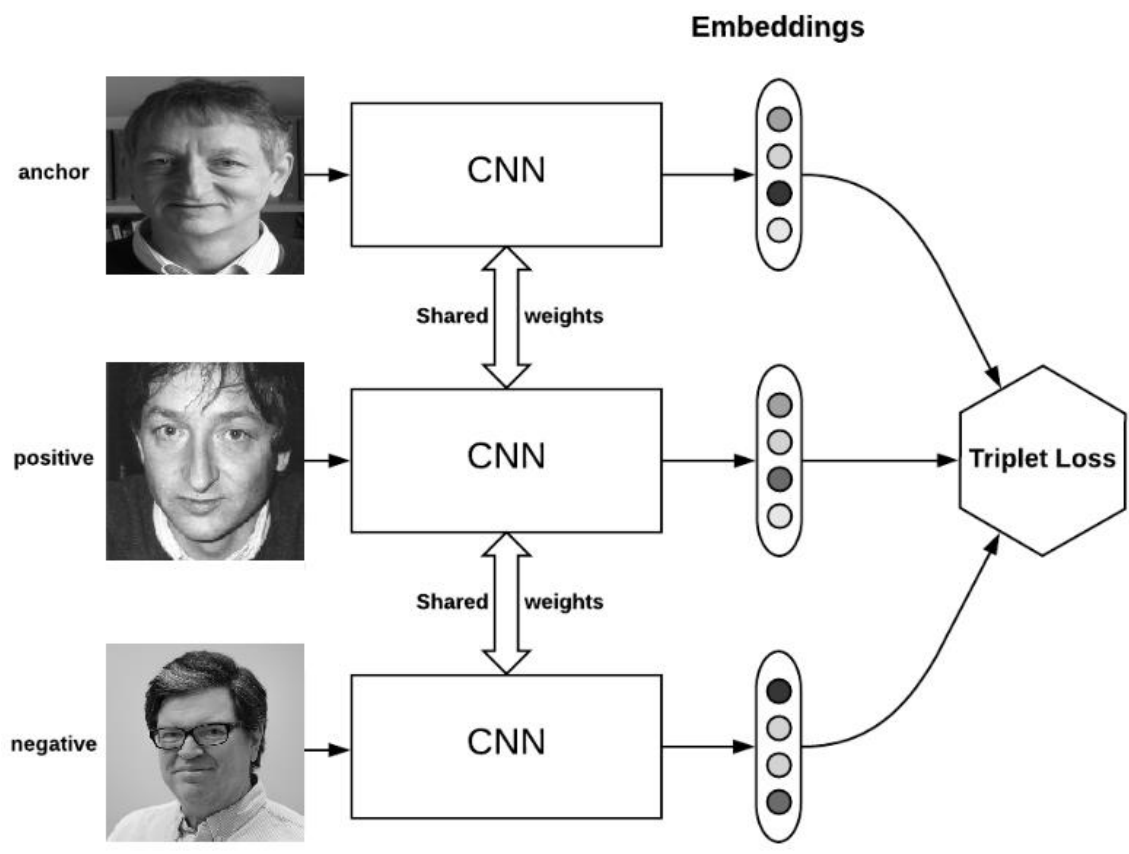
Discriminative Gaussian Process Latent Variable Model (DGPLVM) [13] ili poznatiji pod nazivom GaussianFace je algoritam prepoznavanja lica koji su razvili istraživači na kineskom sveučilištu u Hong Kongu 2014. godine. Ovaj algoritam ostvario je 98.52% točnosti prepoznavanja lica, u odnosu na testnu skupinu ljudi, koja je ostvarila 97.53%. [12]

Iste godine, Facebook je najavio svoj algoritam za prepoznavanje lica, pod nazivom DeepFace. DeepFace je algoritam koji se bazira na neuronskoj mreži od devet dubokih slojeva, ima više od 120 milijuna veza između neurona i trenirana je na 4 milijuna slika Facebook korisnika. [14] Navedeni algoritam može odrediti pripadaju li dvije fotografije lica istoj osobi, s točnošću od 97.25%. [12]

2016. godine Amazon predstavlja svoj algoritam prepoznavanja lica, kojeg nazivaju Amazon Rekognition. Amazonov algoritam je predstavljen kao usluga u oblaku američkim vladajućim tijelima poput američkog imigracijske službe (*U.S. Immigration and Customs Enforcement*) i policije. [15] Algoritam može prepoznati oko 100 ljudi na jednoj slici. U srpnju 2018. godine, Amazonov algoritam pogrešno je identificirao 28 članova američkog Kongresa, zamijenivši ih za ljude s kriminalnom poviješću. [12]

LFIS (*Live Face Identification System*) je Thalesov „*State of the art*“ algoritam za prepoznavanje lica koji se bazira na dubokim neuronskim mrežama. LFIS može obraditi video zapise i identificirati osobe s točnošću od 99.44%. Ovaj algoritam koriste policijske službe. [16]

Uz sve navedeno, važno je spomenuti i Googleov FaceNet, koji ima najveću točnost (99.63%) u odnosu na prethodno navedene sustave za prepoznavanje lica. Ovaj algoritam pruža mogućnost prepoznavanja lica putem učenja mapiranja iz slika. FaceNet koristi duboku konvolucijsku neuronsku mrežu, to jest, CNN. Slike koje FaceNet upotrebljava za trening neuronske mreže su transformiranih veličina, skalirane i vrlo usko izoliraju područje na kojemu se nalazi lice. Ono što razlikuje FaceNet od drugih sustava je korištenje tzv. „*Triplet Loss*“ metode (slika 2.4.) koja koristi tri fotografije nazvane *anchor* fotografija (sidro), *positive* i *negative* fotografija. [17]



Slika 2.4. Prikaz „Triplet Loss“ metode [17]

3. KORIŠTENE TEHNOLOGIJE

Treće poglavlje predstaviti će korištene tehnologije koje su bile potrebne za izradu ovog diplomskog rada. Pozivajući se na prethodno navedene tehnologije i alate iz uvodnog poglavlja, bit će pružen opis svakog pojedinog alata, krenuvši od programskog jezika Python. Python je odabran zbog njegovih raznovrsnih značajki, od kojih je jedna kompatibilnost s većinom operacijskih sustava, uključujući Linux, Windows i mnoge druge ugrađene sustave. Također, korišteni su različiti Python paketi, koji su unutar Pythona vrlo jednostavni za implementaciju. Glavni programski paket koji je bio korišten u procesu izrade ovog diplomskog rada je OpenCV, on omogućuje raznovrsne načine obrade i prepoznavanja značajki na različitim fotografijama. TensorFlow je drugi programski paket koji je korišten zbog lakše implementacije neuronske mreže unutar aplikacije, koja je korištena za optimiziranje algoritma prepoznavanja lica.

3.1. Spyder

Spyder je programsko okruženje napisano u programskom jeziku Python. Navedeno programsko okruženje napravljeno je prvenstveno za korištenje različitim strukama, od inženjera do znanstvenika u mnogim znanstvenim granama. Također, Spyder je *open source* programsko okruženje, namijenjeno za korištenje na više platformi. Razvijen je 2009. godine i konstantno nadograđivan i usavršavan, kako bi ponudio raznolike mogućnosti napredne analize podataka i njihove obrade te mnogih drugih funkcionalnosti koje posjeduje. Spyder posjeduje PyQt5 biblioteku, koja dopušta razvojnim programerima da dodaju nove funkcionalnosti i ugrađene komponente. [18]

3.2. Python

Python je programski jezik visoke razine, radi preko interpretera, a ne preko *debuggera* kao većina programskih jezika. Razvijen je 1991., a njegov tvorac je Guido van Rossum. Prednosti Pythona leže u njegovom dizajnu i upravo zato pomaže programerima olakšati im posao kod pisanja koda kako bi napisali čist i logičan kod, što ga čini pogodnim za kreaciju od najmanjih do najvećih projekata. Također, Python je programski jezik objektnog tipa te radi na principu „skupljanja smeća“ što ga stavlja u jednaku kategoriju gdje su i Java i C# programski jezici. Naposljetku, Python je također *open source*, što ga čini jednostavnim za distribuciju i korištenje u komercijalne i edukativne svrhe. [19]

3.3. OpenCV

OpenCV skraćenica je za *Open Source Computer Vision Library*, odnosno, biblioteka otvorenog tipa namijenjena za računalni vid i strojno učenje. Navedena biblioteka napravljena je kako bi pružila jednu infrastrukturu koja sadrži većinu algoritama za računalni vid i kako bi ubrzala razvoj komercijalnih proizvoda koji koriste računalni vid. OpenCV ima više od 2500 optimiziranih algoritama, koji obuhvaćaju širok spektar klasičnih i modernih algoritama računalnog vida i strojnog učenja. Najčešće upotrebe ovakvih algoritama su za detekciju lica, prepoznavanje objekata, ljudi, radnji i druge. [20]

3.4. TensorFlow

TensorFlow je *open source* biblioteka namijenjena za tijek podataka i diferencijalno programiranje za obavljanje raznolikih vrsta zadataka. Biblioteka TensorFlow nudi mnoge razine apstrakcije tako da je korisniku moguće izabrati onu koja im najviše odgovara za njihove potrebe. TensorFlow je stavljen na tržište 2015. godine, a prethodno tome, korišten je za interne potrebe tvrtke Google. Ova biblioteka prvenstveno je započela kao Googleov DistBelief, algoritam strojnog učenja, koji se bazirao na dubokim neuronskim mrežama. S obzirom da je dobio veliku popularnost unutar različitih tvrtki, Google je okupio grupu znanstveniku, uključujući Jeffa Deana, kako bi pojednostavili i preinačili kod DistBeliefa u puno bržu i robusniju biblioteku, koja je kasnije postala TensorFlow. [21]

4. PROGRAMSKO RJEŠENJE

S obzirom na čovjekovo urođeno znanje i sposobnost prepoznavanja lica drugih ljudi, prepoznavanje lica osobe na fotografiji ili drugim medijima, na prvi pogled je moguće zaključiti kako program koji prepoznaje lice ima jednostavno rješenje. Međutim, rješavanje problema prepoznavanja osobe, iako vrlo jednostavan zadatak za čovjeka, kompleksan je zadatak za računalo. Pri susretanju s vrlo složenim problemima, s obzirom na svoju kompleksnost, često je moguće pomisliti da je rješenje nepostojeće. Najbolji način za rješavanje izuzetno složenih problema je podjela problema na manje cjeline i stvaranje strukturalnog rješenja. Problem prepoznavanja lica osobe na fotografiji ili drugim medijima moguće je podijeliti na dvije veće cjeline: pronalazak lica na zadanom mediju i prepoznavanje pronađenog lica i usporedba s poznatim licima iz baze podataka.

4.1. Pronalazak lica na zadanom mediju

Što zapravo predstavlja problem pronalaska lica na fotografiji? Kako ljudi prepoznaju lice?, navedena pitanja ključna su pri osmišljavanju načina rješavanja problema pronalaska lica. Čovjek je u stanju prepoznati lice na fotografiji prvim pogledom na nju i bit će u potpunosti siguran da je riječ o licu, neovisno o tome što se na licu nalazi (dodatci poput naočala, nakita, deformacija i slično). Čovjek je pritom u mogućnosti prepoznati lice, ali nije svjestan na koji način ga prepoznaje, odnosno, kako radi njegov unutarnji algoritam prepoznavanja lica. Međutim, kada bi se pretpostavio način na koji radi čovjekov urođeni algoritam prepoznavanja lica, došlo bi se do zaključka da on prepoznaje određene značajke na licu osobe, poput očiju, nosa, usta, obrva, trepavica, ušiju, zubi, boje kože, boje očiju, pjega, bora, oblika lubanje i tako dalje. Navedene značajke samo su dio niza različitih značajki koje ljudsko lice posjeduje. S obzirom na veliku količinu značajki i ljudsku nesvjesnost vlastitog algoritma prepoznavanja lica, nije jednostavno odrediti koje su značajke neophodne kako bi se nešto nazvalo licem, a koje značajke nisu. Zbog prethodno navedenih elemenata, bilo je potrebno pronaći odgovarajući način koji može raditi s velikim brojem varijabli i odrediti koje varijable su najvažnije za određeni rezultat. Za postavljeni problem neuronske mreže pokazale su se kao najbolje rješenje.

4.1.1. Sastavljanje i treniranje neuronske mreže

Kako je navedeno u prethodnom poglavlju, u prepoznavanju lica potrebno je ubrojiti sve elemente koji sačinjavaju lice i čine ga prepoznatljivim ljudskom oku. S obzirom na zadane parametre, neuronske mreže korištene su kako bi se sastavio sustav, koji bi prvenstveno prepoznavao nalazi li se lice na slici. Kako bi neuronska mreža uspješno obavila svoj zadatak, potrebno joj je predati podatke s kojima će raditi. Neuronska mreža zapravo je klasifikator, koja kao rezultat daje pripada li zadani podatak određenoj skupini (može postojati N skupina). U slučaju sustava za prepoznavanje lica, postoji jedna tražena skupina, to jest, ljudsko lice. Međutim, kako bi se prepoznalo lice, potrebno je moći odrediti što je lice, a što nije; što znači, kako bi se neuronska mreža trenirala, potrebne su dvije skupine (klase). Prva klasa je ljudsko lice (pozitivna klasa, slika 4.1.), dok je druga klasa sve ono što nije lice (negativna klasa, slika 4.2.), što je više različitih objekata u negativnoj klasi, to će mreža bolje moći razaznati radi li se o licu ili ne. [22]



Slika 4.1. Primjer podataka pozitivne klase



Slika 4.2. Primjer podataka negativne klase

4.1.2. Predobrada podataka

U većini slučajeva u prikupljenim podacima nalazi se mnoštvo beskorisnih informacija. Ako su koriste takvi neobrađeni podatci, treniranje neuronske mreže bit će znatno duže i zahtijevat će više takvih podataka kako bi neuronska mreža davala precizna rješenja. Iz tog razloga, potrebno je filtrirati beskorisne informacije. Kako se u slučaju prepoznavanja lica radi s fotografijom, a fotografija je trodimenzionalni podatak, koji se sastoji od RGB dimenzija, prvi filter bit će pretvoriti tri dimenzije u jednu (slika 4.3.). Postoji više načina na koji je ovaj korak moguće izvesti, a najjednostavniji način je izračunati srednju vrijednost sve tri komponente. Navedeni korak smanjuje dimenzionalnost podataka tri puta.

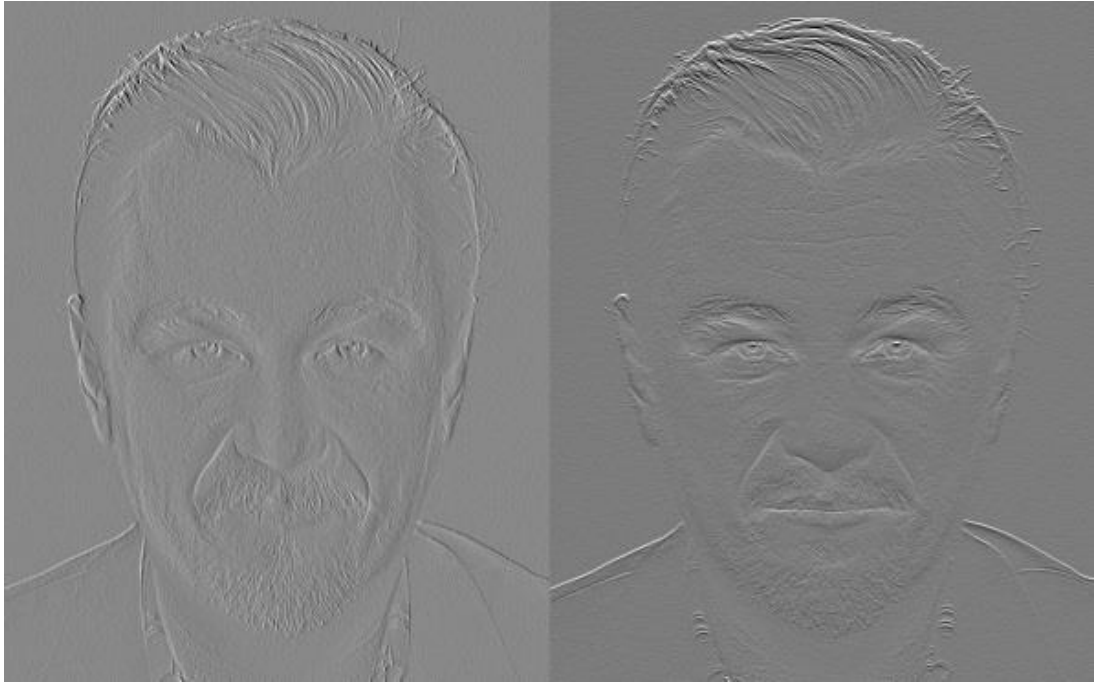


Slika 4.3. Prikaz crno-bijelog filtera

Međutim, iako su dimenzije slike smanjene tri puta, trenutna slika još uvijek ima prevelik broj podataka. Čovjek i u jednoj, i u drugoj slici bez poteškoća može odrediti gdje se nalaze oči, nos, usta i tako dalje. S obzirom na to da se slika sastoji od mnoštva piksela kojima je vrijednost od 0 do 255, računalu je teško odrediti gdje se koja važna značajka nalazi samo iz vrijednosti piksela. Kako čovjek ima mogućnost uočiti siluetu lica osobe, potrebno je napraviti sličnu pojavu za računalu. Siluetu je moguće definirati tako da se detektiraju rubovi lica. Jedan od najčešće korištenih algoritama za detektiranje rubova je Sobelov operator. Sobelov operator radi na način da uzima gradijent svjetla i tako detektira rub. Operator koristi 3 x 3 matricu (formule 4-1 i 4-2) [23], koja zajedno s originalnom slikom daje rezultat gradijenta svjetla, što su zapravo rubovi (slika 4.4.). Nadalje, operator se mora koristiti u x i y smjeru.

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * A \quad (4-1)$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (4-2)$$

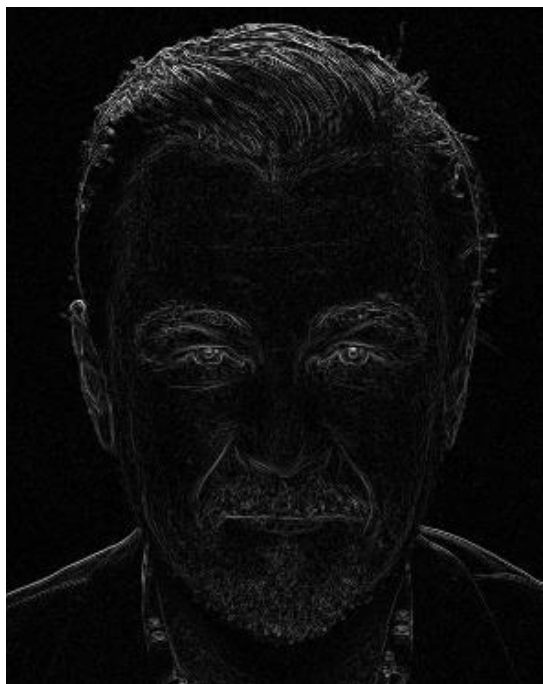


Slika 4.4. Rezultat Sobelovog operatora u x i y smjeru

Idući korak je, iz rezultata Sobelovog operatora dobiti jačinu i smjer gradijenta koristeći formule: 4-3 i 4-4 [24], primjer rezultata jačine gradijenta prikazan ja ne slici 4.5.:

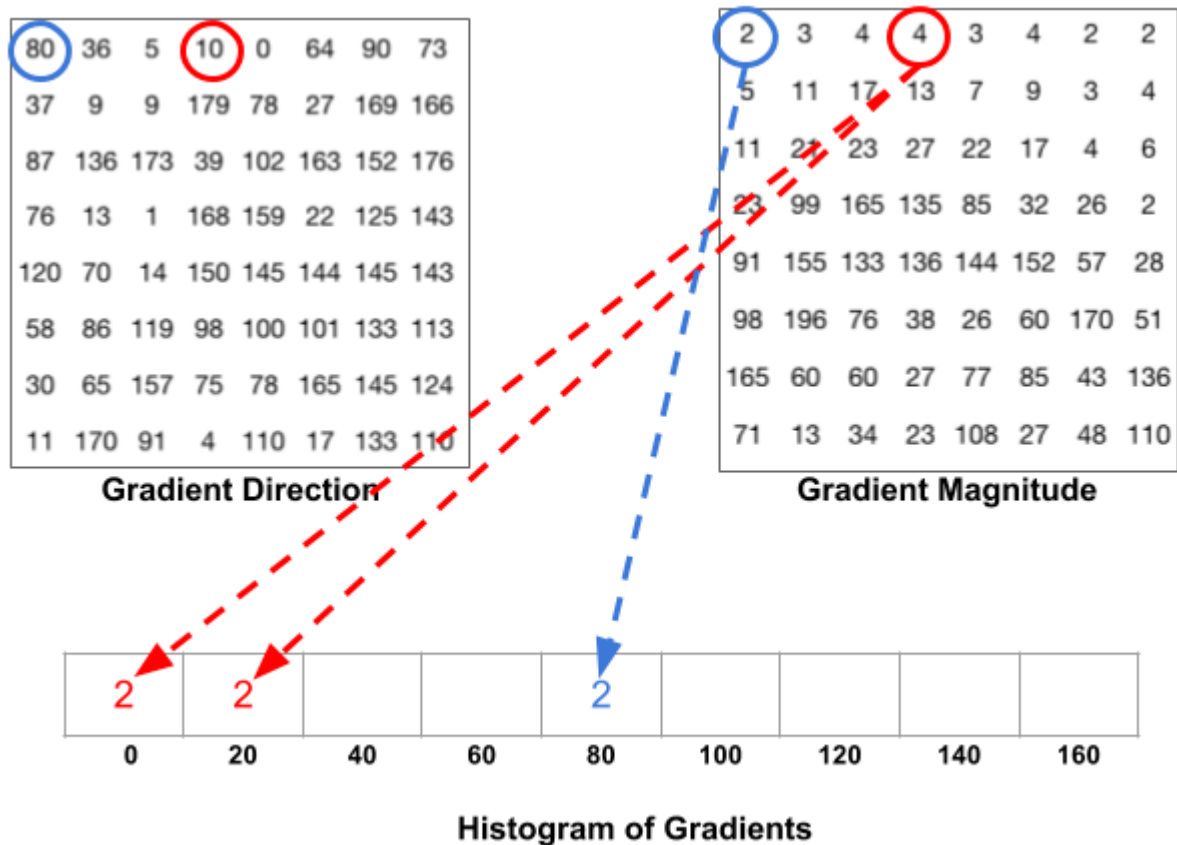
$$g = \sqrt{g_x^2 + g_y^2} \quad (4-3)$$

$$\theta = \arctan\left(\frac{g_y}{g_x}\right) \quad (4-4)$$



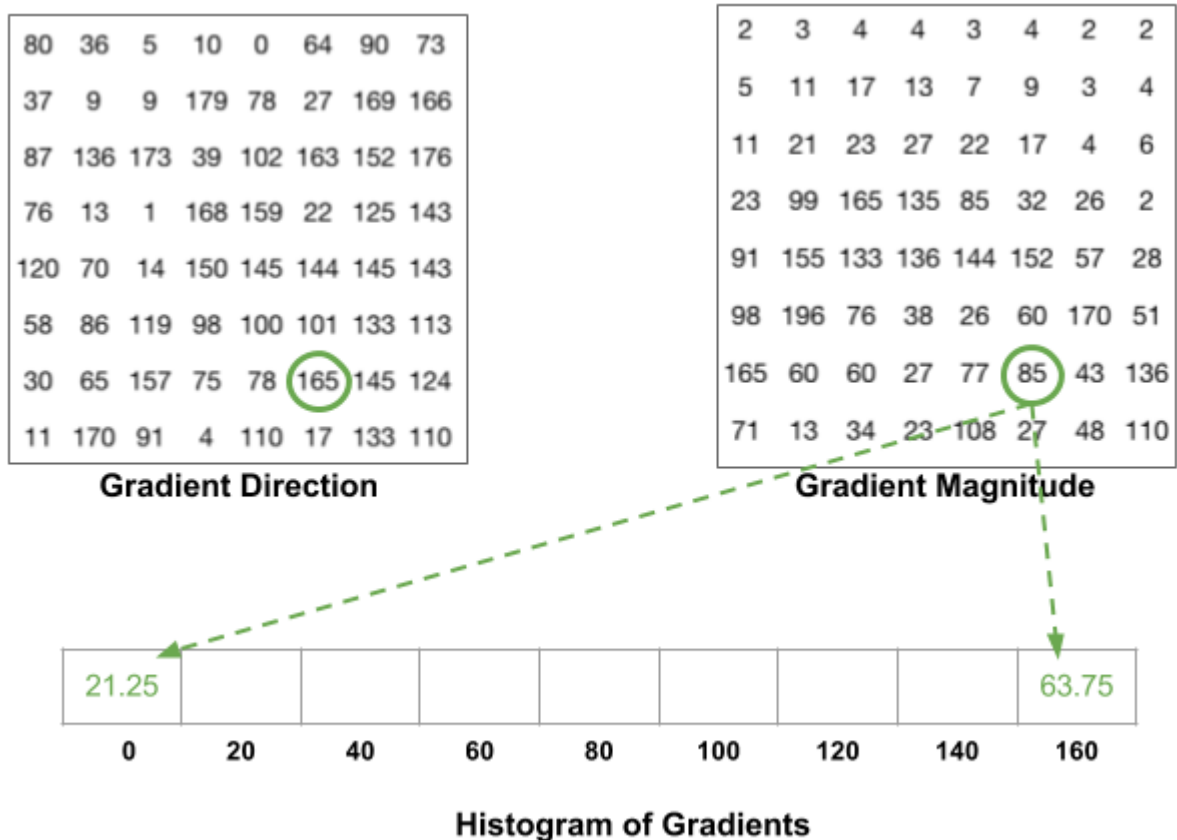
Slika 4.5. Rezultat jačine gradijenta

Iz jačine gradijenta moguće je izračunati histogram orijentiranih gradijenata, što je i posljednji korak predobrade podataka. Histogram orijentiranih gradijenata je histogram koji prikazuje gradijent svjetla pod određenim kutovima. Prvi korak pri računanju histograma orijentiranih gradijenata je podijeliti sliku na 16 x 16 dijelove. Ovdje nema pravila, 16 x 16 odabran je zbog rezolucije trenutne slike, veća slika će zahtijevati veću rezoluciju dijelova na koji se slika dijeli, dok će manja slika zahtijevati manje dijelova na koje se dijeli. U prethodnom koraku izračunata je jačina gradijenta i kut gradijenta za svaki od piksela.



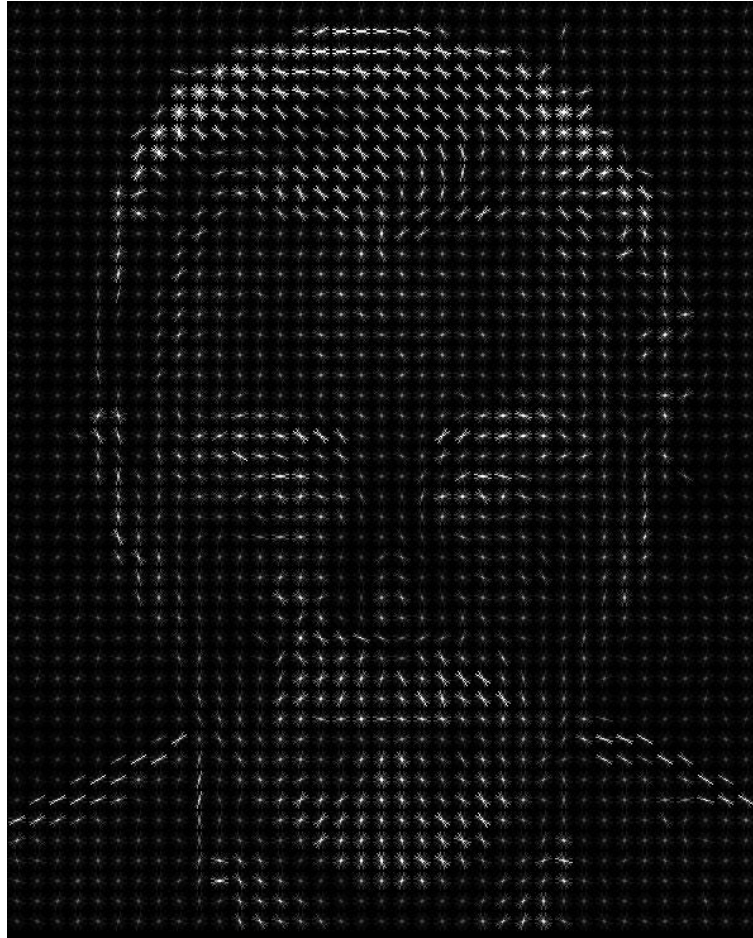
Slika 4.6. Prikaz računanja histograma orijentiranih gradijenata na ćeliji veličine 8 x 8 [24]

Slika 4.6. ilustrira proces računanja histograma orijentiranih gradijenta za ćeliju veličine 8 x 8. Pritom se koristi niz od devet brojeva gdje navedeni brojevi označavaju kutove između 0 i 160. Kao što je uočljivo na slici, gledajući broj označen plavim krugom, on ima kut od 80 stupnjeva i jačinu vrijednosti 2, što znači da se broj 2 postavlja u ćeliju ispod koje se nalazi broj 80. Promatrajući broj označen crvenim krugom, uočljivo je kako on ima kut od 10 stupnjeva i gradijent jačine u vrijednosti 4, s obzirom na to da se kut 10 nalazi točno na pola između 0 i 20, dodaje se pola gradijenta jačine u ćeliju označenu brojem 0 i pola jačine u ćeliju označenu brojem 20. [24]



Slika 4.7. Prikaz računanja histograma orijentiranih gradijenata, poseban slučaj [24]

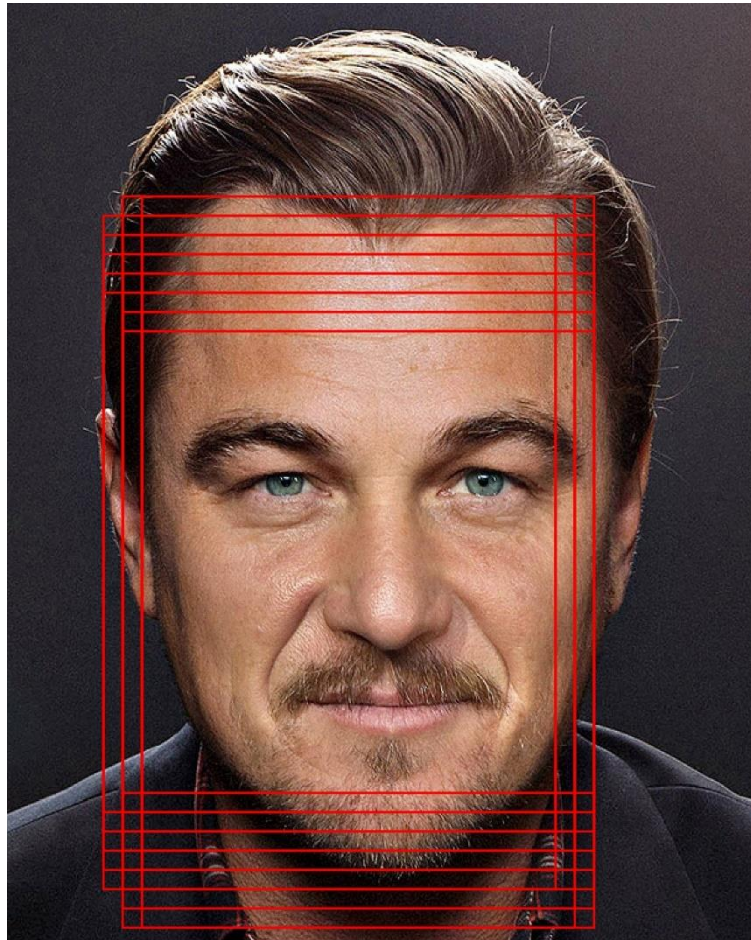
Slika 4.7. prikazuje slučaj gdje se može dogoditi da je kut veći od 160 stupnjeva, kao što je moguće vidjeti broj koji je označen zelenim krugom koji ima vrijednost 165. Poznato je kako 180 i 0 su zapravo isti kutovi pa se može računati kao da je broj između 160 i 180 te se iz toga dobije da je $\frac{1}{4}$ gradijenta jačine ide kod kuta 180, što je zapravo isto što i ćelija označena brojem 0, dok $\frac{3}{4}$ idu u ćeliju označenu sa 160. Nakon što se ovaj korak ponovi za sve piksele unutar ćelije i odradi se za svih N ćelija, dobit će se histogram orijentiranih gradijenata. Nakon toga, potrebna je normalizacija koja se odradi tako da se svi elementi unutar histograma podijele s najvećim elementom i s time se dobije da na histogram orijentiranih gradijenata (slika 4.8.) ne utječe količina svjetla i veličina ćelije. [24]



Slika 4.8. Rezultat histograma orijentiranih gradijenata

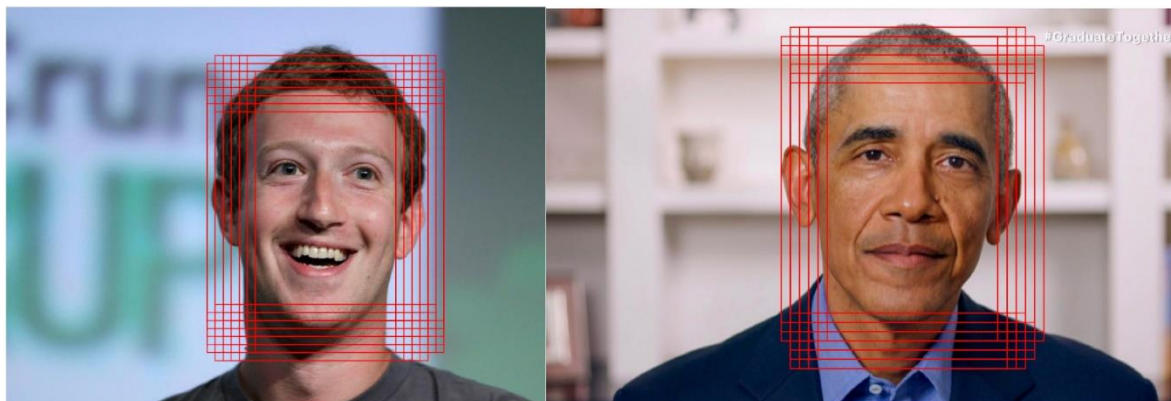
4.1.3. Metoda klizajućeg prozora

Nakon obrade podataka, što je podrazumijevalo računanje histograma orijentiranih gradijenata, sljedeći korak je predati mreži obrađene podatke kako bi mreža bila istrenirana i mogla klasificirati radi li se o ljudskom licu ili ne. Nakon treniranja neuronske mreže, dobivena je preciznost u iznosu od 98% [22]. Ovdje ne završava dizajn sustava za prepoznavanje lica jer je mreža istrenirana kako bi mogla prepoznati lice, ali samo u slučaju gdje je cijela fotografija lice. U slučaju da se na fotografiji nalazi osoba u nekoj okolini i ta osoba zauzima samo neki određeni dio fotografije, mreža će izbaciti kao rezultat za tu fotografiju da ta fotografija nije lice. Navedeni problem se rješava pomoću metode klizajućeg prozora. Metoda klizajućeg prozora radi na principu da postoji određeni „prozor“ koji je zapravo pravokutnik iste veličine kao i fotografije na kojoj je treniran model mreže. Navedeni prozor se pomiče na desnu stranu i prema dolje dok ne prođe cijelu fotografiju i pritom ispituje za svaki od dijelova slike nalazi li se na tom dijelu lice. Rezultat metode klizajućeg prozora daje više pravokutnika centriranih oko lica osobe (slika 4.9.). [22]



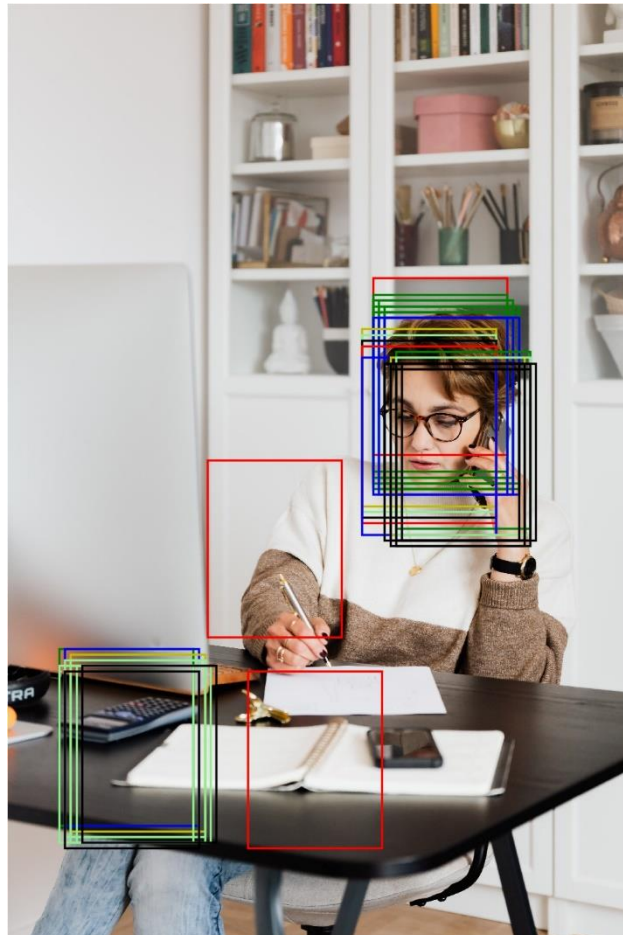
Slika 4.9. Rezultat metode klizajućeg prozora

Koristeći istrenirani model zajedno s metodom klizajućeg prozora na nekoliko različitih fotografija dobiveno je da je model vrlo precizan ukoliko se radi o portretu fotografije bez okoline ispunjene raznolikim objektima (slika 4.10.).



Slika 4.10. Rezultat na fotografijama gdje su osobe u prvom planu

U slučaju da se za testni primjer uzme fotografija koja ima više različitih objekata i gdje osoba nije u prvom planu, s obzirom na jednostavnost neuronske mreže i na relativno malu količinu podataka, metoda klizajućeg prozora, zajedno s istreniranim modelom neuronske mreže ne daje vrlo precizne rezultate, kao što je moguće uočiti na slici 4.11.:



Slika 4.11. Primjer točnosti neuronske mreže na fotografiji gdje se nalazi više različitih objekata i lice nije u prvom planu

Nakon testiranja rješenja na različitim fotografijama, moguće je primijetiti kako rješenje radi relativno dobro, ali donosi i određene svoje probleme. Jedan od problema je i određivanje skale slike; s obzirom na to da je mreža trenirana na fiksnoj rezoluciji podataka fotografije kako bi ovakav model imao apsolutnu točnost, lice na testnoj slici mora biti otprilike jednake rezolucije kao i veličina prozora koji je jednake rezolucije kao i testni podatci. Navedena pojava dovodi do problema određivanja skale slike ili prozora. Drugi problem je jednostavan model i relativno mala količina podataka što dovodi neuronsku mrežu da ima „*false positive*“ rezultate na zadanim fotografijama. Rješenje prvog problema bilo bi treniranje mreže s podacima različitih skala, a rješenje drugog problema bilo bi korištenje sofisticiranije neuronske mreže s velikom količinom

različitih podataka. Oba rješenja dovela bi do znatnog povećanja vremena treniranja mreže, zbog čega će se u ovom radu koristiti neki od većih istreniranih modela neuronskih mreža (navedenih u poglavlju usporedbe postojećih rješenja), ovisno o brzini i preciznosti navedenih modela.

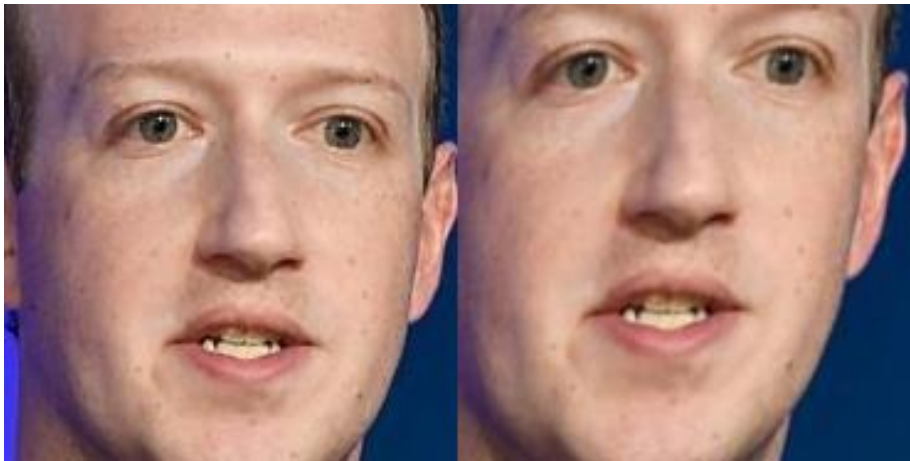
4.2. Točke interesa lica

Nakon što je pronađeno i označeno lice na fotografiji koristeći neuronske mreže potrebno je to isto lice usporediti s postojećim licem iz baze podataka kako bi se odredilo nalazi li se na slici osoba koju algoritam „poznaje“. Ovaj korak bio bi nepotreban kada bi postojali testni podatci gdje je lice uvijek pozicionirano izravno pred kamerom. S obzirom na to da je takav slučaj vrlo rijedak u stvarnom svijetu, potrebno je imati određeni način kako bi lice uvijek bilo pozicionirano prema kameri. Kako bi se dobio traženi rezultat gdje je lice uvijek okrenuto prema kameri, koristi se metoda 68 točaka interesa lica, koju su 2014. osmislili Vahid Kazemi i Josephine Sullivan. [25] Navedena metoda istrenirana je na velikoj količini podataka (68 točaka na slikama lica). Nakon treniranja mreže na podatcima ovog tipa, koristi se neuronska mreža kako bi se istrenirao model 68 točaka interesa lica, čiji je rezultat vidljiv na slici 4.12.:



Slika 4.12. Prikaz 68 točaka interesa na fotografiji

Sljedeći korak nakon pronalaženja točaka interesa na licu, pronađene točke se uspoređuju sa zadanim *templateom* točaka interesa lica. *Template* se nalazi u vlastitom koordinatnom sustavu koji je uvijek okrenut prema kameri te se koriste različiti algoritmi kako bi se lice dovelo u poziciju *templatea*, gdje je okrenuto prema kameri. Postoje kompleksni algoritmi poput 3D poravnanja [26] gdje se navedene točke koriste u 3D prostoru što je potencijalno jedan od boljih modela, ali zbog jednostavnosti programskog rješenja ovog rada korištene su najjednostavnije matematičke operacije transformiranja slike *shear* i *warp*, koje su automatski izvedene u metodi pronalaska točaka interesa lica. Prikaz fotografije s transformiranim licem pokazuje slika 4.13.:



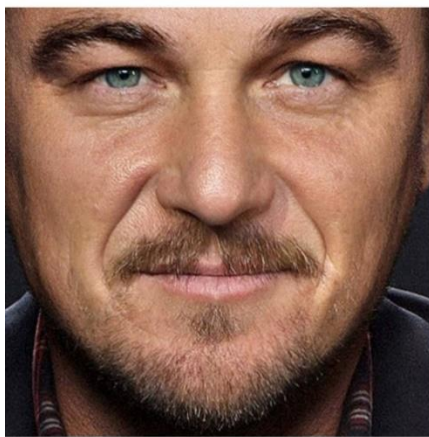
Slika 4.13. Prikaz poravnanja fotografije koristeći *shear* i *warp*

4.3. Prepoznavanje osobe na slici

U prethodnim koracima objašnjeno je i pokazano kako razviti sustav za pronalaženje lica na fotografiji veličine $X \times Y$, također, objašnjeno je i kako za uspješno prepoznavanje lica potrebno je imati 68 točaka interesa na licu kako bi lice uvijek bilo okrenuto prema kameri. Kad su čovjeku dane 2 identične slike lica, osoba će bez problema i bez dvoumljenja moći reći da je jednaka osoba na obje slike. U navedenoj situaciji potrebno je uzeti u obzir da se ljudi svakodnevno susreću s različitim licima u svojoj okolini, što bi možda dovelo do ideje da se dvije slike lica izravno uspoređuju jedna s drugom. Ovakva ideja nije loša, ali, primjera radi, kad bi postojala RGB fotografija lica veličine 1024×1024 i ako se pretpostavi da računalo obradi milijun brojeva u jednoj sekundi s bazom od 1000 osoba dobit će se: $1024 \cdot 1024 \cdot 3 = 3145728 \div 1000000 \frac{\text{operacija}}{\text{sekunda}} \cdot 1000 = 3145,73s = 52,43 \text{ minute}$, s čime se dobije da za pronalazak jedne osobe od 1000 ljudi je potrebno ≈ 52 minute. Kad se ovaj primjer prebaci u kontekst društvenih mreža, poput Facebooka, koji ima više od 2.7 milijardi korisnika, takav algoritam se ne bi mogao

primijeniti u stvarne svrhe, što dovodi do problema; kako fotografiju lica osobe precizno prikazati s manjom rezolucijom podataka. [27]

Kako bi se smanjila dimenzionalnost podataka potrebno je sliku opisati s puno manje dimenzija. U ovom slučaju postoje različiti načini, od kojih je najpoznatiji Encoding Base 64, koji pretvara sliku u niz znakova kako bi je bilo lakše koristiti na web stranicama. Ovakav način nije moguće koristiti u prepoznavanju lica jer i mala varijacija na slici će rezultirati u potpunosti drugačijim nizom znakova. Kako lica imaju određene značajke na sebi, poput boje očiju, trepavica, obrva, boje kose, veličine ušiju, nosa, usta i slično, potrebno je napraviti način koji će opisati što se nalazi na slici. Navedeni postupak zove se enkodiranje lica, koji se radi pomoću Resnet neuronske mreže koja daje vektor od 128 značajki koji opisuju zadano lice (slika 4.14.). [25]

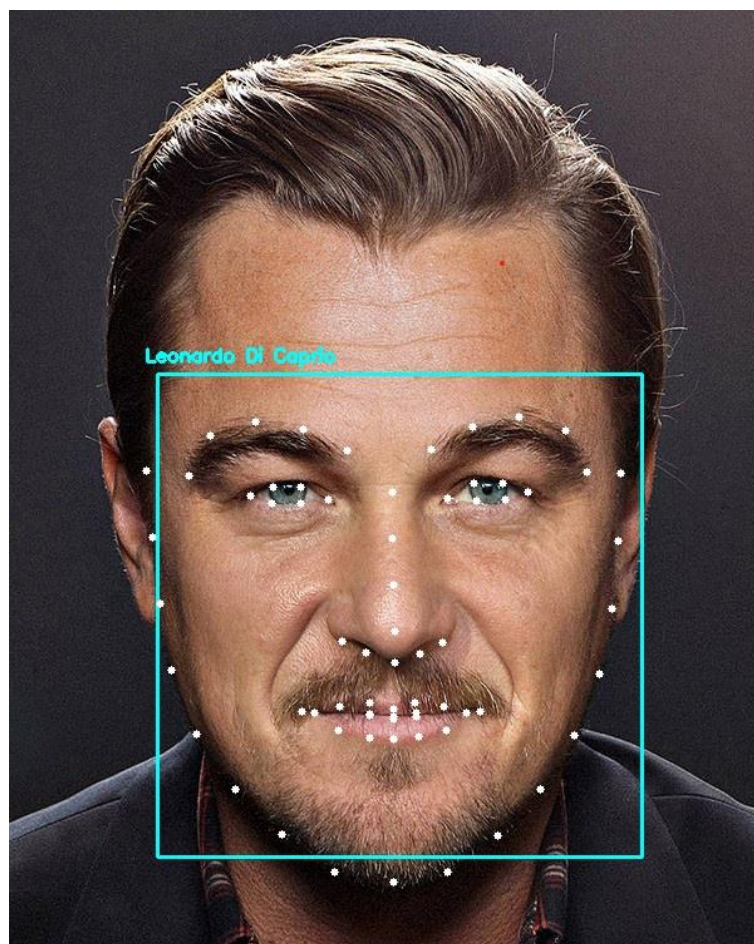


-0.076434865593910	-0.152869731187820	-0.229384596781731	-0.305739462375641
0.075654089450836	0.151308178901672	0.226962268352509	0.302616357803345
0.124168574810028	0.248337149620056	0.372505724430084	0.496674259924012
-0.034278389066458	-0.068556778132915	-0.102835167199373	-0.137113556265831
-0.121938303112984	-0.243876606225967	-0.365814909338951	-0.487753212451935
-0.023327343165874	-0.046654686331749	-0.069982029497623	-0.093309372663498
0.077418811619282	0.154837623238564	0.232256434857845	0.309675246477127
-0.087471380829811	-0.174942761659622	-0.262414142489433	-0.349885523319244
0.140589997172356	0.281179994344711	0.421769991517067	0.56235998689423
0.015179047524929	0.032358095040858	0.048537142574787	0.064715190899716
0.258149027824402	0.516298055648804	0.774447083473206	1.032596111297607
-0.032327276418924	-0.064745552837849	-0.097118329256773	-0.129491105657697
-0.287813514471054	-0.575627028942108	-0.863440543413162	-1.151254057884216
-0.027534145861864	-0.055068291723728	-0.082602437585592	-0.118136583447456
-0.00858589823735	-0.017171779647470	-0.025757669471204	-0.03434359294939
0.063795156776905	0.127590313553810	0.191385470330715	0.255180621707620
-0.05446840076386	-0.108936801552773	-0.163405202329159	-0.217873683105545
-0.141384214162827	-0.282768428325653	-0.424152642488480	-0.565536856651306
-0.079739645123482	-0.159479290246964	-0.239218935378445	-0.318958580493927
-0.050465591251850	-0.100931182503700	-0.151396773755550	-0.201062365807401
-0.060583144426346	-0.121166288852692	-0.181749433279037	-0.242332577705383
-0.004128032363951	-0.008256064727902	-0.0123840897091854	-0.016512129455805
-0.008044233785756	-0.001688467571512	-0.002532701357268	-0.003376935143024
0.020590972155333	0.041181944310665	0.061772916465998	0.082363888621330
-0.156765401363373	-0.313530802726746	-0.470296204090118	-0.627061605453491
0.295273125171661	0.590546250834323	0.885819375514984	1.181092590686646
-0.026537643745542	-0.053075287491083	-0.079612931236625	-0.106150574982166
0.009965099394321	0.019930198788643	0.029895298182964	0.039860397577286
0.075807109475136	0.151614218950272	0.227421328425407	0.303228437900543
-0.182376101613045	-0.364752203226089	-0.547128304839134	-0.72958406452179
-0.00694582897192	-0.013891657814384	-0.020837486721575	-0.027783315628767
0.199150711297989	0.398301422595978	0.597452133893967	0.796602845191956

Slika 4.14. 128 značajki enkodirane slike lica

Ako bi se ponovio račun iz prethodnog odlomka i ako se pretpostave jednaki parametri, dobit će se: $128 \div 1000000 \frac{\text{operacija}}{\text{sekunda}} \cdot 1000 = 0.128s$, što je puno brže nego kada bi se išlo s prethodnom pretpostavkom. Također, ovakav algoritam je jako dobar jer neovisno o rezoluciji fotografije, algoritam će uvijek davati vektor od 128 značajki. Kako bi se značajke izračunale, korišten je gotov istrenirani model, s obzirom na činjenicu da je na grafičkoj kartici Nvidia Tesla trebalo 24 sata treniranja kako bi se dobila dobra preciznost modela. [25] Izračunavanjem značajki napravljen je i zadnji korak algoritma za prepoznavanje lica. Način na koji se dobije identitet osobe na slici je uspoređivanje baze podataka s enkodiranim licima i enkodirano traženo lice (slika 4.15.). Algoritam uspoređivanja je Euklidska udaljenost, koja se računa po formuli 4-5:

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (4-5)$$



Slika 4.15. Rezultat raspoznavanja lica

Kako bi se maksimizirala preciznost i točnost algoritma, jedna osoba će biti definirana s više enkodiranih fotografija, najčešće se uzima deset fotografija po jednoj osobi za dobru točnost. Nakon što se izabere deset različitih fotografija za N osoba, ti podatci se mogu ubaciti u najjednostavniji SVM klasifikator te istrenirati njega da raspoznaje zadane osobe. [25] Nakon što je završen proces treniranja klasifikatora, proces prepoznavanja lica je tim korakom završen. Prateći prethodne korake i objašnjenja, ako bi se svi modeli išli trenirati, to jest, da se ne uzimaju gotovi modeli. Prethodni koraci su svojevrsan „recept“, prema kojemu se može napraviti sustav za raspoznavanje bilo kakvog objekta, koji ne mora biti ljudsko lice. Posljednji korak je napraviti *whitelist* i/ili *blacklist* te usporediti dobivene identifikacijske podatke s postojećim listama te odrediti odgovarajuće sigurnosne mjere ovisno o dobivenom rezultatu.

5. USPOREDBA POSTOJEĆIH RJEŠENJA

Kako se ova znanstvena grana počela razvijati 60-ih godina 20.stoljeća, razumno je pretpostaviti kako postoje već određene metode pronalaženja lica na fotografijama, stoga su uspoređene najpoznatije metode. Odabrane su: Viola - Jones, HOG+SVM, CNN i MTCNN, koje su najpoznatije četiri metode pronalaženja lica, koje su zatim testirane i uspoređene.

5.1. Viola – Jones klasifikator temeljen na Haar značajkama

Pri pronalaženju lica, najvažniji čimbenici su otkrivanje najvažnijih značajki na ljudskom licu poput očiju, nosa, obrva, usta i slično. Ono što je u ovom slučaju potrebno je otkrivanje ovih značajki u stvarnom vremenu na fotografiji, za što služe Haar značajke, koje su nazvane prema mađarskom matematičaru Alfredu Haaru. Haar je predložio korištenje navedenih značajki 1909. godine. Značajke na slici 5.1. predstavljene su kvadratima i pravokutnicima s tamnom i svijetlom stranom, što je način na koji računalo određuje koja je to značajka. [28]



Slika 5.1. Prikaz Haarovih značajki [28]

5.2. SVM klasifikator istreniran na HOG podacima

HOG algoritam (*Histogram of Oriented Gradients*) deskriptor je značajki koji se koristi u računalnom vidu i obradi slike u svrhu otkrivanja objekata. 1986. godine Robert K. McConell je prvi put opisao koncept HOG algoritma, pritom ne koristeći naziv HOG pri patentiranju. Nakon toga, 1994. godine, Mitsubishi Electric Research Laboratories jedni su od prvih korisnika koncepta HOG algoritma. Međutim, HOG algoritam se proširio 2005. godine, kad su istraživači s

Francuskog nacionalnog instituta, Navneet Dalal i Bill Triggs, postavili svoj rad koji se bazirao na HOG deskriptorima na Konferenciji računalnog vida i prepoznavanja uzoraka. [29]

5.3. Klasifikator temeljen na CNN modelu

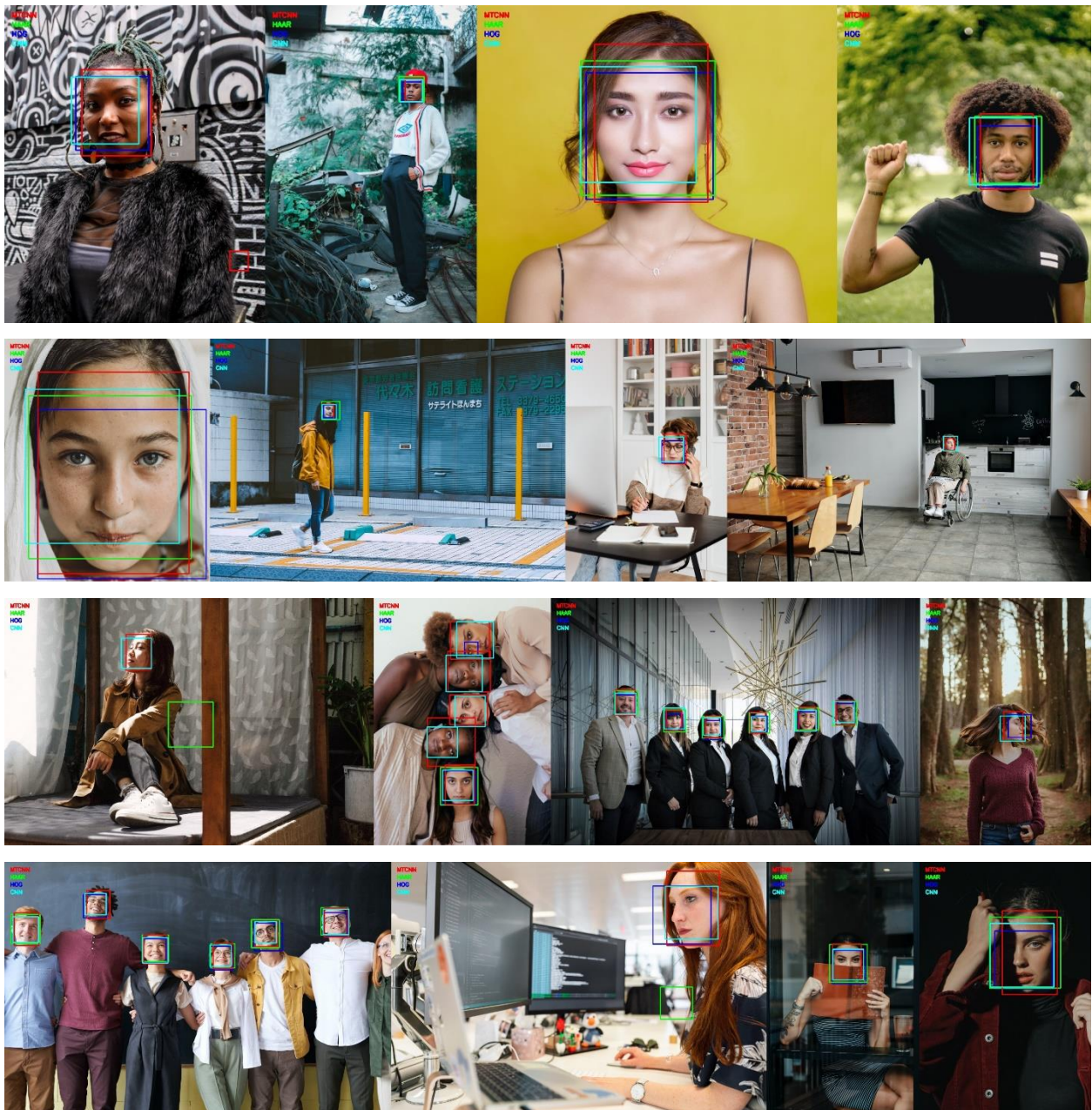
CNN (*Convolutional Neural Network*) klasifikator temeljen je na modelu duboke konvolucijske neuronske mreže, zbog čega je i dobio svoje ime. Mreža na kojoj je treniran model sastoji se od ulaznog i izlaznog sloja. S obzirom na to da mreža radi sa slikama, postoji i konvolucijski sloj, iza kojeg idu potpuno povezani slojevi, *flatten* i *softmax* sloj. [30]

5.4. Klasifikator temeljen na MTCNN modelu

MTCNN, ili *Multi-task Cascaded Convolutional Networks* je posebno dizajnirana neuronska mreža koja služi za pronalaženja lica unutar određene fotografije, gdje fotografija može sadržavati jedno ili više lica. U ovakvim neuronskim mrežama koristi se korelacija između detekcije i poravnanja kako bi se poboljšalo njihovo izvođenje. MTCNN koristi stepenastu strukturu s tri razine pažljivo dizajniranih dubokih konvolucijskih neuronskih mreža kako bi odredile lokaciju lica. Ovakve mreže prvo pronalaze važne značajke ljudskog lica, poput očiju, nosa i usta, te nakon toga određuje krajnji rub lica. Najveća razlika između CNN modela i MTCNN modela je to što MTCNN koristi multirezolucijski model, što znači da ova mreža ima više rezolucija iste slike, takozvani „*image pyramid*“ za trening podatke, dok CNN koristi jednu sliku u $X \times Y$ rezoluciji. [31]

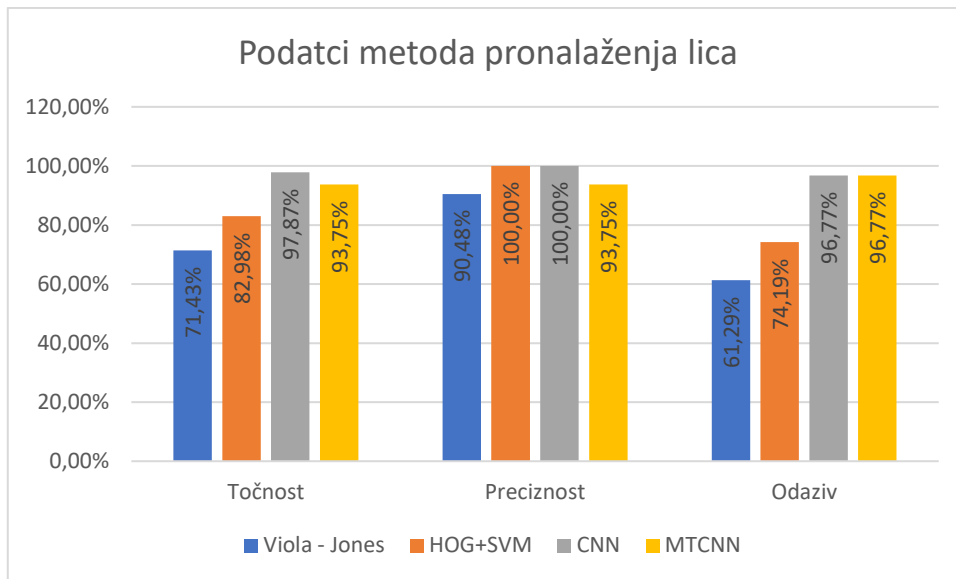
5.5. Usporedba metoda pronalazjenja lica

Kako bi se prethodno navedene metode prepoznavanja lica mogle usporediti, potrebno je odabrati testni skup podataka koji će testirati sve aspekte prethodno navedenih metoda. Odabrani testni skup sastojao se od 32 slike, od kojih je 16 predstavljalo osobe (slika 5.2.), a 16 objekta. Fotografije s osobama grupirane su po određenim kriterijima: lice u prvom planu, osoba i jako velik broj objekata, takozvani „*busy environment*“, fotografije gdje se nalazi više osoba i fotografije gdje je određeni dio lica prekriven.

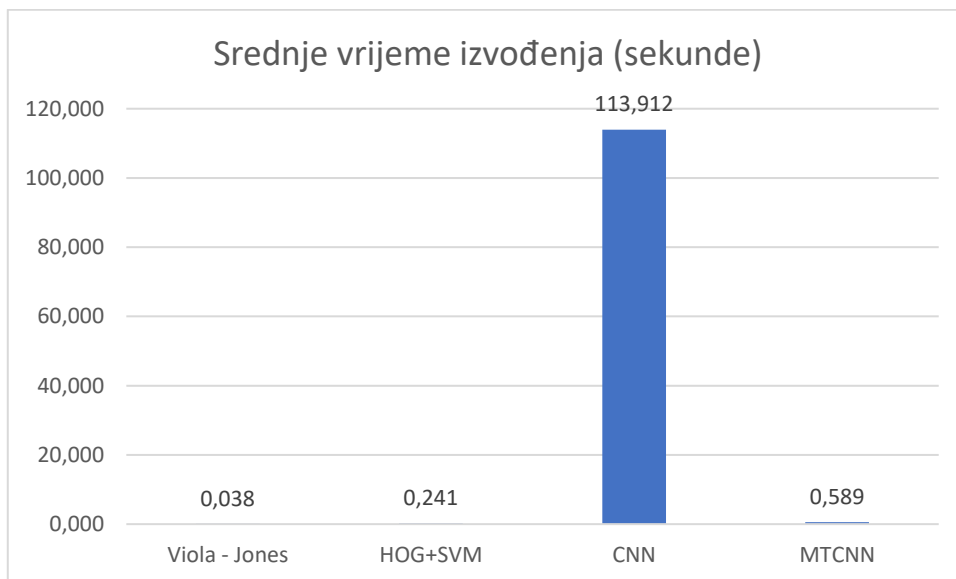


Slika 5.2. Prikaz točnosti na testnom skupu za sve modele

Nakon testiranja metoda pronalaženja lica Viola - Jones, HOG+SVM , CNN i MTCNN na prethodno navedenim testnim podacima, dobiveni su sljedeći rezultati:



Slika 5.3. Podatci metoda pronalaženja lica

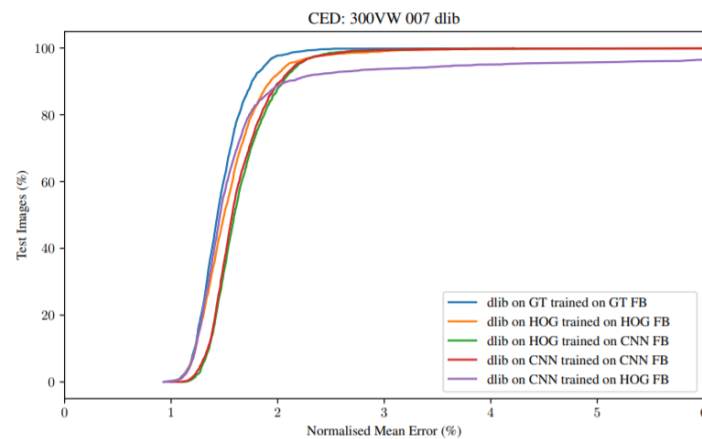


Slika 5.4. Srednje vrijeme izvođenja u sekundama

Promatrajući dobivene podatke sa slike 5.3., moguće je uvidjeti kako najveću točnost i najveću preciznost ima CNN model prepoznavanja lica. Slika 5.4. prikazuje kako CNN model ima daleko najveće vrijeme izvođenja u odnosu na druge modele. Viola - Jones model je najbrži, ali pritom ima najmanju točnost i najmanju preciznost. MTCNN ima jako veliku točnost i relativno nisko vrijeme, što ovaj model čini najboljim za prepoznavanje osoba na fotografijama. Iz navedenih

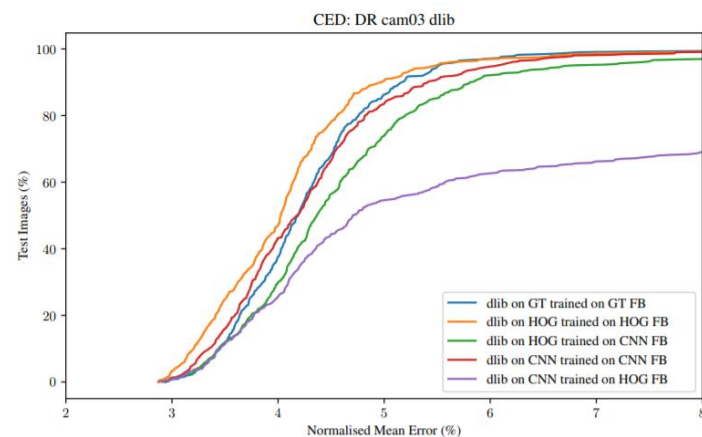
podataka moguće je uočiti kako je Viola - Jones model najbolji za primjenu na video materijalima, s obzirom na to da ima brzinu izvođenja daleko ispod jedne sekunde.

S obzirom na to da je ovo testiranje trajalo oko jedan sat za 32 različite fotografije, zbog ograničenja računala (nemogućnost obrade velike količine podataka u realnom vremenu) i nemogućnosti pristupa drugim vrstama modela, te kako su ovakve testove obavile različite tvrtke, znanstvenici i ljudi koji su razvijali modele za pretraživanje lica, istaknuti su podatci istraživanja Philippa Koppa, koji prikazuju preciznost metoda pronalaženja lica, navedeno u radu *Analysis and Improvement of Facial Landmark Detection* iz 2019. godine. [32]

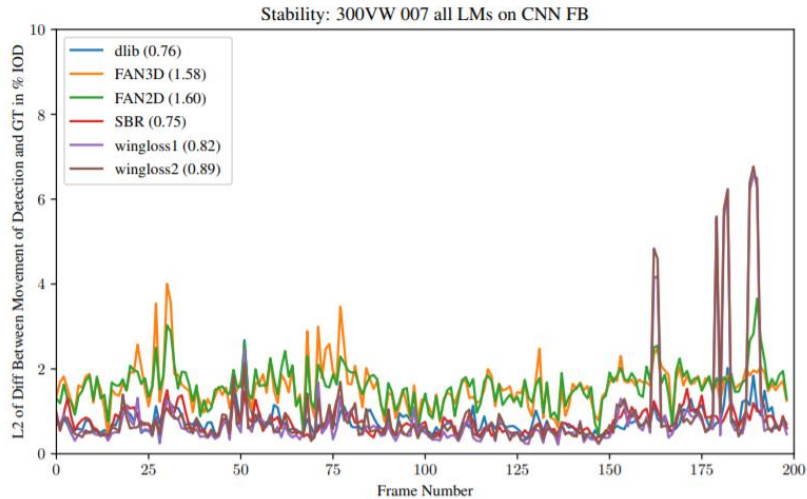


Slika 5.5. Usporedba modela pronalaženja lica treniranih na različitim podacima koristeći sekvencu slika 300VW 007 [32]

Grafovi na slici 5.5. i 5.6. prikazuju točnost dlib metoda pronalaženja lica treniranih na različitim podacima. Prema Kopp, metode ne uspiju prepoznati lice u 35% slučajeva ako je model istreniran na HOG podacima. [32]



Slika 5.6. Usporedba modela pronalaženja lica treniranih na različitim podacima koristeći sekvencu slika DR cam03 [32]

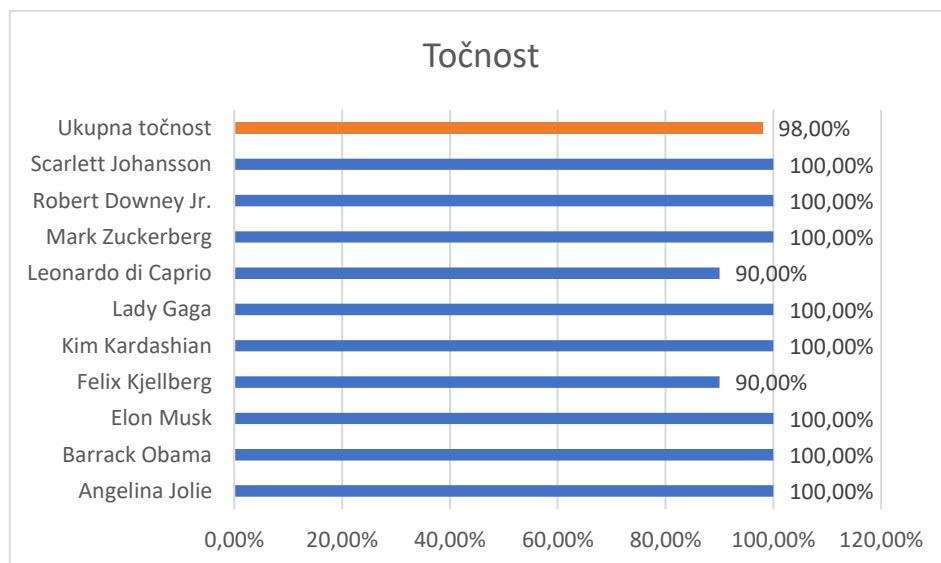


Slika 5.7. Prikaz *drifta* različitih modela pronalaženja točaka interesa [32]

Graf na slici 5.7. prikazuje količinu *drifta* (za isti podatak, ukoliko se algoritam izvrši više puta, davat će slične, ali nikada identične rezultate) različitih modela pronalaženja točaka interesa te se može uočiti kako dlib i SBR model pronalaženja točaka interesa imaju najmanji *drift*, iz kojeg razloga se u ovom diplomskom radu koristi dlib model.

5.6. Testiranje algoritma prepoznavanja lica

Na kraju, obavljeno je testiranje sustava za prepoznavanje lica na deset različitih poznatih osoba kako bi se utvrdilo kolika je uspješnost napravljenog algoritma. Za testiranje odabrana je jedna fotografija poznate osobe (osoba gleda izravno u kameru) s kojom će se uspoređivati i deset fotografija (osoba ne gleda izravno u kameru) na kojima će se testirati.



Slika 5.8. Točnost sustava na testnim podacima



Slika 5.9. Primjer testnih fotografija i rada algoritma na njima

Slika 5.8. prikazuje grafikon gdje je moguće vidjeti kako sustav ima 98% točnost na testnom skupu. Točnost ovakvog sustava ovisi o baznoj slici, koja se koristi kao glavni opis osobe i sprema u bazu podataka. Kako bi se postigli najbolji rezultati, bazna slika mora biti veće rezolucije i lice bi trebalo biti orijentirano prema kameri te bi trebalo biti u cijelosti vidljivo. Sa slike 5.9. moguće je uočiti kako sustav nije uspio prepoznati dvije fotografije, čiji je zajednički čimbenik bio to da su bile puno manje rezolucije nego njihova bazna fotografija. Nakon povećavanja rezolucije obje fotografije, algoritam je imao stopostotnu točnost.

6. ZAKLJUČAK

Tijekom nekoliko prethodnih godina, razvoj čitavog svijeta dosegnuo je svoj vrhunac u pozitivnom smislu, ali i u negativnom, gdje zbog opće dostupnosti informacija i tehnoloških dobara uvijek postoji opasnost od njihovog zloupotrebljavanja. Modernizacija suvremene svakidašnjice čovjekova života dovela je do potrebe za povećanjem stupnja sigurnosti na svim poljima javnog života, što je dovelo do pojačanja javne sigurnosti integracijom kamera i nadzornih sustava u čovjekovu svakidašnjicu. Ideja kontrole pristupa koristeći prepoznavanje lica nije novost, već ona seže sve do 60-ih godina 20. stoljeća, kad se prvi put spominje. Međutim, zbog tehnološke ograničenosti ideju nije bilo moguće provesti u djelo. Prijelaz preko ograničenja koja su bila prepreka u razvoju sigurnosnog sustava za prepoznavanje lica događa se razvojem društvenih mreža koje su donijele velike baze podataka lica i naglim napretkom računalnih tehnologija, gdje je moderni procesor u stanju obraditi više od milijun operacija u jednoj sekundi.

U ovom diplomskom radu riječ je o istraživanju i izradi sustava za kontrolu pristupa temeljenom na prepoznavanju lica. Prvenstveno, bilo je važno proučiti razvoj sustava za prepoznavanje lica kroz povijest, počevši od ručnog unosa sredinom 20. stoljeća pa do automatiziranih sustava koji se nalaze u mobilnim uređajima. Pri razvoju sustava za prepoznavanje lica korišten je programski jezik Python zbog jednostavnosti rada s neuronskim mrežama.

Izrada ovakvog sustava zahtijevala je kreaciju i treniranje neuronske mreže. Kako bi se neuronska mreža uspješno istrenirala, jedan od važnijih dijelova bio je obraditi podatke koristeći metodu histograma orijentiranih gradijenata. Na dobivenom rezultatu koristeći model 68 točaka interesa, slika se ispravlja putem algoritma *warp* i *shear*. Posljednji korak bio je enkodirati fotografiju u 128 ključnih značajki i usporediti je s bazom lica kako bi se dobio identitet osobe.

Uz sve prethodno navedeno, ovaj diplomski rad obuhvaćao je usporedbu gotovih modela, gdje su uspoređena četiri modela, od kojih je moguće izdvojiti dva, MTCNN kao onaj s najboljom točnošću i kao najadekvatniji za prepoznavanje na fotografiji te Viola - Jones, s najbržim vremenom izvođenja čija svrha najviše dolazi do izražaja u prepoznavanju lica u video zapisima.

Naposljetku, uočeno je kako ovakav sustav nije optimalan za klasični način sigurnosti jer ga je vrlo lako prevariti koristeći fotografiju. U ovom slučaju preporuča se korištenje minimalno dvije kamere kako bi se dobila stereo snimka koja bi mogla prepoznavati lice u 3D prostoru. Ovakav sustav bio bi idealan za frekventna javna mjesta kao obavještajni sustav bilo kakvog nepoželjnog ponašanja koji je korisnik zadao.

LITERATURA

- [1] M. Hrga, Računalni vid, Zbornik radova Veleučilišta u Šibeniku, No. 1-2/2018, pp. 207-216, 2018.
- [2] Računalni vid [online], wikipedia.org, dostupno na: https://hr.wikipedia.org/wiki/Ra%C4%8Dunalni_vid [23.9.2020]
- [3] F. Ahmad, A. Najam, Z. Ahmed, „Image-based Face Detection and Recognition: „State of the Art“, Beijing University of Aeronautics & Astronautics, Beijing, 2013.
- [4] L.A.Thomas, M.D.De Bellis, R.Graham, K.S.LaBar, Development of emotional facial recognition in late childhood and adolescence, Developmental Science, No. 10:5/2007, pp. 547-558, 2007.
- [5] How does facial recognition work? [online], us.norton.com, dostupno na: <https://us.norton.com/internetsecurity-iot-how-facial-recognition-software-works.html>, [23.9.2020.]
- [6] Woody Bledsoe [online], wikipedia.org, dostupno na: https://en.wikipedia.org/wiki/Woody_Bledsoe, [23.9.2020.]
- [7] The History of Face Recognition Software [online], facefirst.com, dostupno na: <https://www.facefirst.com/blog/brief-history-of-face-recognition-software/#>, [23.9.2020.]
- [8] S.Raviv, Secret History of Facial Recognition [online], 2020., wired.com, dostupno na: <https://www.wired.com/story/secret-history-facial-recognition/>, [24.9.2020.]
- [9] M.Turk, A.Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, No.1, Vol.3, pp. 71-86, 1991.
- [10] P.J.Phillips, H.Moon, S.A.Rizvi, P.J.Rauss, The FERET Evaluation Methodology for Face-Recognition Algorithms, Gaithersburg, 1999.
- [11] L.Vaas, Facebook expands use of face recognition [online], 2019., nakedsecurity.sophos.com, dostupno na: <https://nakedsecurity.sophos.com/2019/09/06/facebook-expands-use-of-face-recognition/>, [24.9.2020.]

- [12] Facial Recognition – fascinating and intriguing [online], thalesgroup.com, dostupno na: <https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/facial-recognition>, [27.9.2020.]
- [13] C.Lu, X.Tang, Surpassing Human – Level Face Verification Performance on LFW with GaussianFace, The Chinese University of Hong Kong, Hong Kong, 2014.
- [14] DeepFace [online], wikipedia.org, dostupno na: <https://en.wikipedia.org/wiki/DeepFace>, [27.9.2020.]
- [15] Amazon Rekognition [online], wikipedia.org, dostupno na: https://en.wikipedia.org/wiki/Amazon_Rekognition, [27.9.2020.]
- [16] Video – based facial recognition system [online], thalesgroup.com, dostupno na: <https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/biometric-software/live-face-identification-system>, [27.9.2020.]
- [17] E.Tyantov, Deep Learning Achievements Over the Past Year [online], 2017., blog.statsbot.co, dostupno na: <https://blog.statsbot.co/deep-learning-achievements-4c563e034257> [26.9.2020.]
- [18] Spyder, The Scientific Python Development Environment [online], spyder-ide.org, dostupno na: <https://www.spyder-ide.org/>, [23.9.2020.]
- [19] Python [online], wikipedia.org, dostupno na: [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)), [23.9.2020.]
- [20] OpenCV [online], opencv.org, dostupno na: <https://opencv.org/about/>, [23.9.2020.]
- [21] TensorFlow [online], wikipedia.org, dostupno na: <https://en.wikipedia.org/wiki/TensorFlow>, [23.9.2020.]
- [22] Application: A Face Detection Pipeline [online], jakevdp.github.io, dostupno na: <https://jakevdp.github.io/PythonDataScienceHandbook/05.14-image-features.html> [23.9.2020.]
- [23] Sobel operator [online], wikipedia.org, dostupno na: https://en.wikipedia.org/wiki/Sobel_operator [23.9.2020.]

- [24] S. Mallick, Histogram of Oriented Gradients [online], 2016., learnopencv.com, dostupno na: <https://www.learnopencv.com/histogram-of-oriented-gradients/> [23.9.2020.]
- [25] A. Geirgay, Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning [online], 2016., medium.com, dostupno na: <https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78> [23.9.2020.]
- [26] face-alignment [online], github.com, dostupno na: <https://github.com/1adrianb/face-alignment>, [25.9.2020.]
- [27] J. Clement, Number of monthly active Facebook users worldwide as of 2nd quarter 2020 [online], 2020., statista.com, dostupno na: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>, [24.9.2020.]
- [28] R. Gupta, Breaking Down Facial Recognition: The Viola-Jones Algorithm [online], 2019., towardsdatascience.com, dostupno na: <https://towardsdatascience.com/the-intuition-behind-facial-detection-the-viola-jones-algorithm-29d9106b6999> [23.9.2020.]
- [29] Histogram of oriented gradients [online], wikipedia.org, dostupno na: https://en.wikipedia.org/wiki/Histogram_of_oriented_gradients, [24.9.2020.]
- [30] Convolutional neural network [online], wikipedia.org, dostupno na: https://en.wikipedia.org/wiki/Convolutional_neural_network, [24.9.2020.]
- [31] J. Brownlee, How to Perform Face Detection with Deep Learning [online], 2019., machinelearningmastery.com, dostupno na: <https://machinelearningmastery.com/how-to-perform-face-detection-with-classical-and-deep-learning-methods-in-python-with-keras/> [23.9.2020.]
- [32] P. Kopp, Analysis and Improvement of Facial Landmark Detection, Swiss Federal Institute of Technology Zurich, Zurich, 2019.
- [33] dlib-models [online], github.com, dostupno na: <https://github.com/davisking/dlib-models>, [24.9.2020.]

SAŽETAK

Potreba za većim stupnjem sigurnosti u modernom dobu dovela je do razvoja naprednih sustava poput onog za prepoznavanje lica. Povijest sustava za prepoznavanje lica prati se od 60-ih godina 20. stoljeća, od ručnog unosa do kompleksnih modela koji se nalaze čak i na mobilnim uređajima. Zadatak ovog diplomskog rada bio je istražiti i izraditi sustav za prepoznavanje lica koristeći neuronske mreže. Kako bi neuronska mreža bila uspješno istrenirana, jedan od problema bio je način predobrade podataka. Dimenzionalnost podataka smanjena je korištenjem histograma orijentiranih gradijenata. Sustav je programiran u programskom jeziku Python zbog jednostavnosti rada s neuronskim mrežama. Također, izvršena je usporedba postojećih modela kako bi se došlo do najboljeg modela. Koristeći neuronske mreže istrenirane na prethodno obrađenim podacima, sustav uspješno pronalazi i prepoznaje osobu na fotografiji.

Ključne riječi: prepoznavanje lica, neuronske mreže, duboko učenje, sigurnost, računalni vid

ABSTRACT

System for access control based on face recognition

Need for increased security in modern age lead to development of advanced systems such as one for facial recognition. The history of face recognition systems dates back to the 1960s, from manual input to complex models found even on mobile devices. The task of this thesis was to research and develop a system for facial recognition using neural networks. In order for the neural network to be successfully trained, one of the problems was the way to pre-process data. Data dimensionality was reduced using histograms of oriented gradients. System was developed in programming language Python for its simplicity of working with neural networks. Furthermore, a comparison of existing models was conducted in order to discover what the best model is. Using neural networks trained on previously processed data, the system successfully finds and recognizes the person in the given photo.

Keywords: facial recognition, neural networks, deep learning, security, computer vision

ŽIVOTOPIS

Valentino Jagodić rođen je 11. srpnja 1995. godine u Požegi, s prebivalištem u Pleternici. Završio je Osnovnu školu fra Kaje Adžića u Pleternici, nakon čega upisuje Gimnaziju u Požegi, smjer Prirodoslovno – matematički koju završava 2014. godine. 2015. godine upisuje preddiplomski sveučilišni studij računarstva na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku i završava ga 2018. godine, nakon čega upisuje diplomski sveučilišni studij Robotike i umjetne inteligencije na istom fakultetu.

PRILOZI

- [1] Izvorni kod sustava za kontrolu pristupa temeljen na prepoznavanju lica (verzija za video)
- [2] Izvorni kod sustav za kontrolu pristupa temeljen na prepoznavanju lica (verzija za slike)