

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Detekcija lica u grupnim scenama

T. Antunović, S. Čolaković, E. Smoljan, F. Stamenković, I. Weber

Voditelj: *Marijo Maračić*

Zagreb, veljača 2015.

SADRŽAJ

| | |
|---|-----------|
| 1. Uvod i problematika | 1 |
| 1.1. Detekcija lica | 1 |
| 1.2. Prepoznavanje lica | 1 |
| 2. Pregled postojećih rješenja | 2 |
| 2.1. Detekcija lica | 2 |
| 2.1.1. Detekcija bazirana na značajkama | 2 |
| 2.1.2. Detekcija bazirana na slici | 3 |
| 2.2. Raspoznavanje lica | 4 |
| 2.3. Primjer primjene: Bostonski maraton | 5 |
| 3. Predložena implementacija | 7 |
| 3.1. Detekcija lica | 7 |
| 3.2. Prepoznavanje lica | 9 |
| 3.2.1. ICA | 9 |
| 3.3. Korištene baze | 11 |
| 3.4. Platforma programske implementacije | 12 |
| 4. Rezultati | 14 |
| 4.1. Detekcija lica | 14 |
| 4.1.1. Detekcija kože | 14 |
| 4.1.2. Klasifikacija područja kože kao lica | 16 |
| 4.1.3. Detekcija, zaključak | 17 |
| 4.2. Prepoznavanje lica | 18 |
| 4.2.1. Smanjivanje dimenzionalnosti vektora značajki PCA algoritmom | 18 |
| 4.2.2. Prepoznavanje lica ICA algoritmom | 19 |
| 5. Zaključak | 21 |

| | |
|----------------------|-----------|
| 6. Literatura | 22 |
| 7. Sažetak | 24 |

1. Uvod i problematika

1.1. Detekcija lica

Detekcija lica je prvi korak u sustavima za raspoznavanje lica u proizvoljnim scenama. Cilj detekcije je lokalizacija i ekstrakcija lica iz pozadine. Ljudsko lice ima visok stupanj varijabilnosti na slikama, što čini detekciju lica teškim problemom u području računalnog vida. Ovaj problem je detaljno istraživao i postoje različiti načini kako dobiti zadovoljavajuće rezultate.

1.2. Prepoznavanje lica

Prepoznavanje lica je zadatak pridodavanja oznake slici lica, pri čemu je oznaka uglavnom identifikator osobe. U tu svrhu potreban je inicijalni skup poznatih lica (slika identificiranih lica) i algoritam prepoznavanja. Primjene sustava prepoznavanja lica u sigurnosti i komercijalnim/korisničkim područjima su brojne, te se ovo područje aktivno istražuje. Rastom računalne snage, padom cijena senzora (kamera) te razvojem sve snažnijih algoritama prepoznavanja kvaliteta raspoznavanja je došla do točke praktične primjenjivosti.

2. Pregled postojećih rješenja

2.1. Detekcija lica

Sustavi za detekciju lica se mogu podijeliti na sustave temeljene na značajkama (engl. *feature-based*) i sustave temeljene na slici (engl. *image-based*) [1]. Oni koji su temeljeni na značajkama mogu vršiti analizu slike niskog nivoa koja se oslanja na detekciju rubova, područja intenziteta ili boje, mogu vršiti analizu značajki ili kreirati modele aktivnih oblika (engl. *active shape models*). Sustavi temeljeni na slici se dijele na tri glavne skupine: neuronske mreže, metode linearnih potprostora te razne statističke pristupe.

2.1.1. Detekcija bazirana na značajkama

Sustav za robusnu detekciju lica u realnom vremenu opisan u [2] može poslužiti kao primjer sustava baziranom na značajkama. Temelji se na prikazu slike koji su nazvali “integralna slika”, jednostavnom klasifikatoru koji koristi AdaBoost algoritam učenja za izbor najbitnijih značajki iz velikog skupa, te na kaskadnom kombiniranju klasifikatora koje omogućuje da pozadinske regije budu brzo odbačene i da što je moguće veći dio računanja koncentrira na regije koje imaju veću vjerojatnost da predstavljaju lice.

Modeli aktivnih oblika

Modeli aktivnih oblika predstavljaju značajke višeg nivoa od prethodno spomenutih modela. Kada se inicijalizira u blizini značajke model aktivnog oblika će kroz interakciju s lokalnim značajkama poput rubova i osvjetljenosti postepeno zauzeti oblik značajke višeg nivoa. Na taj način se mogu koristiti ne samo za detekciju lica nego i za prepoznavanje lica kroz označavanje bitnih regija poput očiju, obrva, usta i nosa [3].

Detekcija temeljena na boji

Postupak detekcije lica temeljene na boji kojeg koriste radovi poput [4] i [5] je sljedeći. Prvo detektirati područja na slici koja odgovaraju koži na temelju boje slikovnih elemenata. Potom pronađena područja klasificirati kao lica ili ne-lica. Klasifikacija je potrebna zato što se segmentacijom po boji izdvajaju i dijelovi slike koji prikazuju ostale dijelove tijela (primjerice ruke). U radu [4] se klasifikaciji pristupa tako da se najprije odrede područja u kandidatima za lica koji odgovaraju očima i ustima te se lica prihvaća ako je ocjena pronađenih kandidata bolja od neke granične vrijednosti.



Slika 2.1: Primjer detekcije lica u radu [4].

Rad [5] klasifikaciji pristupa tako da na temelju postojećeg skupa slika lica stvara sliku prosječnog lica. Klasifikacija se vrši tako da se računa korelacija kandidata sa prosječnim licem. Ovaj pristup detaljnije je opisan u poglavlju 3.1.

2.1.2. Detekcija bazirana na slici

Neuronske mreže

Pri detekciji lica korištenjem neuronskih mreža u [6] korišteno je više mreža koje su obavljale različite zadatke. Prva neuronska mreža je vršila procjenu poze potencijalnih regija koje predstavljaju lica. Nakon nje se vršilo pretprocesiranje s ciljem smanjivanja

varijacija uzrokovanih osvjetljenjem i razlikama vezanim za kamere. Nakon toga za svaku pozu je korišteno nekoliko neuronskih mreža koje su učile različite stvari iz podataka za treniranje i davale različite rezultate. U posljednjem sloju njihovi izlazi su kombinirani koristeći jednostavnu heuristiku s ciljem povećanja točnih detekcija.

Pristup zasnovan na dubokim konvolucijskim neuronskim mrežama [7] na efikasan način izvlači značajke tokom učenja i na FDDB bazi trenutno ostvaruje najbolje rezultate.

Metode linearnih podprostora

Metode linearnih potprostora su metode poput PCA (engl. *principal component analysis*), ICA (engl. *independent component analysis*), LDA (engl. *linear discriminant analysis*) i FA (engl. *factor analysis*). Pošto unutar predloženog rješenja (poglavlje 3) koristimo PCA i ICA algoritme, ovdje ih nećemo dublje razrađivati.

Statistički pristupi

Kao primjer statističkog pristupa u detekciji lica može poslužiti FloatBoost učenje bazirano na AdaBoost algoritmu [8]. FloatBoost nakon svake iteracije AdaBoost učenja koristi povratni mehanizam za direktnu minimizaciju pogreške. Postiže manju pogrešku učenja i generalizacije koristeći manji broj slabih klasifikatora od AdaBoost algoritma.

2.2. Raspoznavanje lica

Većina algoritama za prepoznavanje lica koriste neki oblik preprocesiranja slike kako bi se dobile značajke pogodne za klasifikaciju. U radu koji nam je dan kao glavna osnova za izradu projekta se koriste metode PCA (principle component analysis) i ICA (independent component analysis), detaljno opisane i uspoređene u [9]. U literaturi ne postoji konceptus o tome koja je metoda bolja u kojim situacijama, ali rad dolazi do zaključka da implementacije ICA metode u prosjeku daju bolje performanse i višu preciznost.

Rad [10] prezentira algoritam raspoznavanja lica u kojem autori nastoje omogućiti precizno raspoznavanje bez obzira na smetnje poput prekrivanja lica, proizvoljnosti položaja i vizualnog šuma. Koriste naučeni koder kako bi dobili opisnik koji se pred kraj obrađuje PCA (principle component analysis) metodom kako bi se smanjile njegove dimenzije i poboljšale performanse algoritma. Detaljnije, nakon detektiranja lica

pronalaže se komponente lica poput nosa, očiju i obraza te se nakon filtriranja šuma dobivaju digitalne reprezentacije njihovih značajki u obliku vektora. Vektori se normaliziraju te se pomoću njih utvrđuje da li se radi o traženom licu. Pretpostavlja se, naravno, da je koder naučen na značajkama lica koja su korištena u treniranju. Algoritam je također u stanju paralelno računanju značajka lica odrediti položaj lica te tu informaciju iskoristiti u konačnoj odluci, čime se poboljšava preciznost. Možda najzanimljiviji dio algoritma je mogućnost kombiniranja više opisnika koristeći poptorni vektor, čime algoritam u konačnosti postiže preciznost od visokih 84.45% na LFW [11] bazi slika.

Rad [12] daje pregled smjera u kojem se trenutno kreću sustavi za raspoznavanje lica i algoritmi koji se pri tome koriste. Opisuje nagli porast interesa za dotično područje u proteklim godinama zbog napretka na području strojnog učenja i računalne grafike te potražnje za sustavima koji koriste tehnologije kojima se bavi računalni put, a neka od primjena su prepoznavanje profilnih slika, automatsko praćenje i promatranje većeg broja ljudi, digitalna rekonstrukcija lica itd. Autor dijeli algoritme za prepoznavanje lica u dvije velike skupine. Prva skupina su algoritmi koji se osnivaju na izgledu, odnosno slikama pojedinaca te se obično koriste u paru sa vektorskim prikazom slika koje obrađuju, što znači da u ovu skupinu pripadaju prije navedeni radovi, a na internetu se nude razne baze podataka slika poput *AT&T Cambridge*¹. Sa druge strane su algoritmi bazirani na modelu, koji nastoje promatrati lica kao 3D modele koji vjerodostojno prikazuju lice sa svim njegovim značajkama te pomoću toga vrše prepoznavanje. Autor navodi neke od algoritama obje skupine te opisuje prednosti i mane svake skupine.

2.3. Primjer primjene: Bostonski maraton

Zanimljiv primjer korištenja računalnog vida u svrhu detekcije i raspoznavanja lica je analiza bombaškog napada na Bostonski maraton 2013. godine.

Detekcija lica na slikama maratona se raspravlja u radu [13], gdje se predlaže algoritam baziran na Viola-Jones algoritmu [14]. Koristeći integralnu sliku i pomični prozor, algoritam prolazi kroz cijelu sliku te provjerava za svaki njen dio da li se na njemu nalazi neko lice. S obzirom da na značajke lica uvelike utječe njegova nakrenutost, algoritam je opremljen detekcijom lica okrenutih prema naprijed, ulijevo ili udesno. Detekcija se vrši nad pomičnim prozorom za svaku od navedenih orijentacija. Algoritam je, osim na slikama maratona, testiran i na FDDB ispitnom skupu te je dao

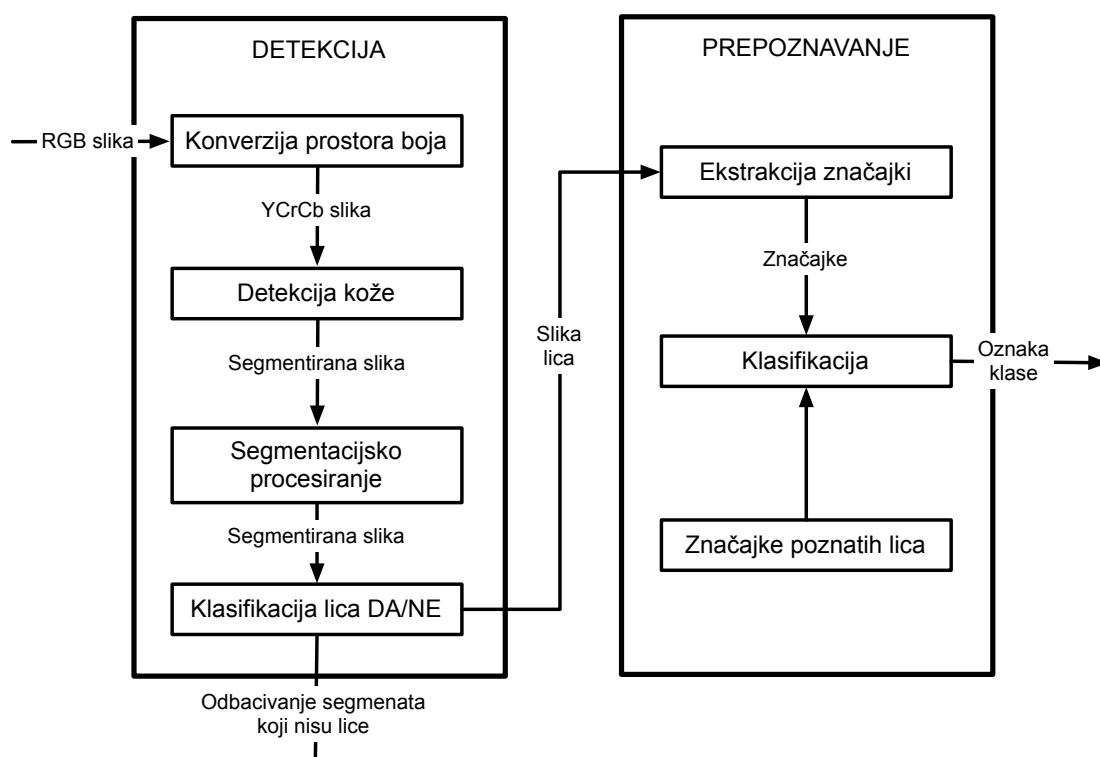
¹<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

bolje rezultate od ostalih rješenja. Nažalost, algoritam se pokazao osjetljivim na utjecaj vizualnog šuma i prekrivanja lica, a problem pozicije lica nije u potpunosti riješen.

Raspoznavanje lica je također provedeno nad slikama Bostonskog maratona, a time se bavi rad [15], koji testira dva komercijalna sustava za raspoznavanje lica, PittPatt 5.2.2 i NeoFace 3.1 nad slikama nekooperativnih osumnjičenika, zbog čega su same slike slabije kvalitete. Valja napomenuti da je istraga vezana uz maraton mogla biti puno kraća uz valjane algoritme raspoznavanja, budući da su se slike lica braće Tsarnaev, glavnih osumnjičenika, nalazile u službenoj bazi podataka. Iako ne nudi uvid u rad samih sustava, rad prikazuje varijacije u dobivenim rezultatima ovisno o načinu izbora skupa sa kojim se lica uspoređuju. Sustavi su prvo testirani nad cijelom bazom podataka koja se sastoji od milijun profila te su za svaku danu sliku dobivene tri osobe koje najviše odgovaraju traženom licu, pri čemu je samo NeoFace uspio ispravno prepoznati jednu sliku mlađeg brata. Nakon toga su za svakog brata uneseni dodatni podaci poput spola, rase i dobi što je smanjilo broj mogućih osoba za otprilike šest puta u oba slučaja, a povećanje kvalitete je bilo proporcionalno tom smanjenju. Na kraju se isprobala varijanta sa spajanjem rezultata za svaku osobu koju se pokušalo prepoznati u bazi, čime se dobivaju bolji rezultati ako su pojedinačne slike davale slične rezultate, ali gori ukoliko nisu. Sve u svemu, rad dobro pokazuje utjecaj pristupa testnom skupu.

3. Predložena implementacija

U radu koji će poslužiti kao osnova za izradu ovog projekta [5] za detekciju lica se vrši jednostavna analiza niskog nivoa temeljena na segmentaciji boja, slike i višeslojnom filtriranju tako dobivenih regija koristeći različite vrijednosti pragova sličnosti s prosječnim licem. Klasifikacija lica temelji se na PCA i LCA algoritmima opisanim u nastavku. Dijagram sustava prikazan je na slici 3.1

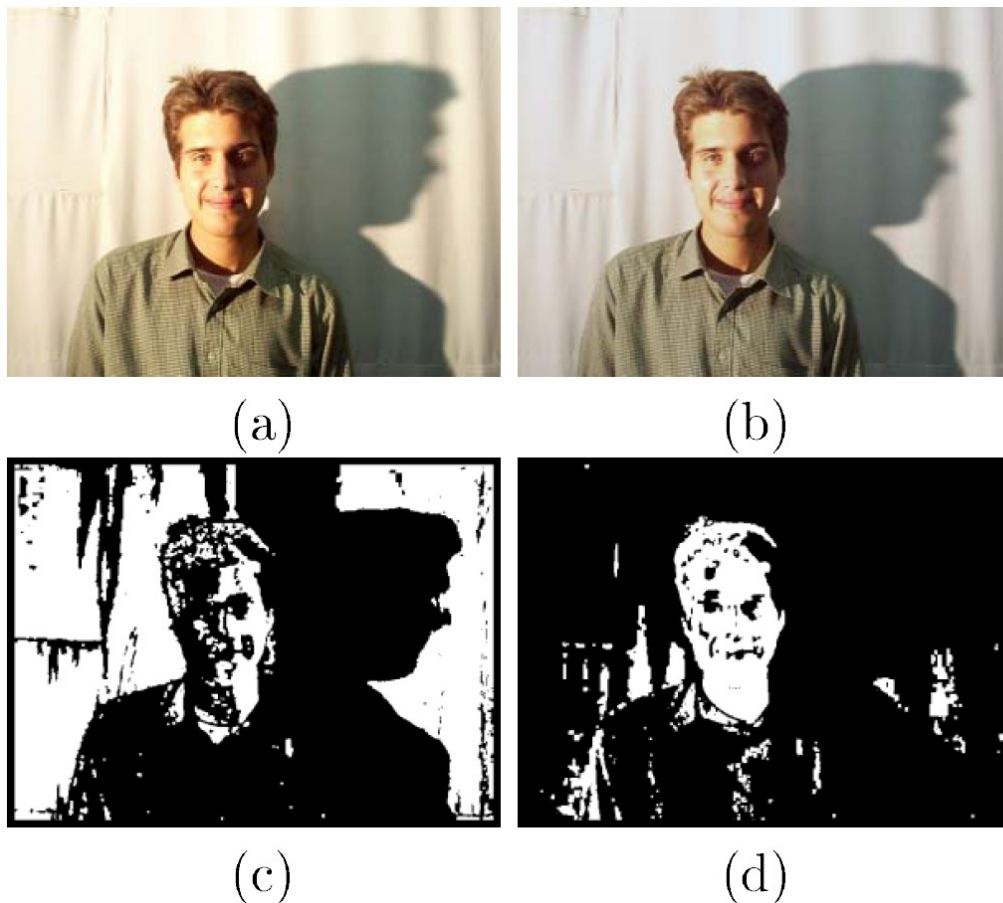


Slika 3.1: Dijagram sustava, protok podataka.

3.1. Detekcija lica

Razmotrimo implementaciju detekcije lica temeljene na boji opisanu u [4] i [5]. Prvi dio postupka obuhvaća pronalazak slikovnih elemenata boje kože. Kako bi se bavio

efikasno, sliku je potrebno iz RGB prostora konvertirati u YCrCb ili YIQ prostor i onda izgraditi binarnu sliku (masku) u kojoj je svaki piksel označen ako komponente piksela zadovoljavaju uvjet pripadanja koži. Sam uvjet pripadanja piksela području kože varira kroz radove: u [16] koji za obradu koristi YCrCb prikaz uvjet glasi $90 < Y < 180$, $90 < Cr < 130$, $80 < Cb < 150$, dok je u radu [4] utvrđena i opisana zavisnost između Cr, Cb i Y komponenti te se prvo izvršava nelinearna transformacija Cr i Cb komponenti i nakon toga ispituje uvjet pripadanja. Ove opisane metode su empirijske i moguće je da svaki istraživački tim definira svoje u sklopu svog rada. U radu [4] se prije samog stvaranja binarne slike početna slika još provlači kroz fazu pretprocesiranja u kojoj se gleda umanjiti utjecaj izvora svjetlosti na boje u slici. Dobivena binarna maska se još dodatno može transformirati operacijama otvaranja, filtriranja, dilatacije, erozije i zatvaranja kako bi se postigle kompaktnije maske koje predstavljaju moguća područja lica. Dobivena područja se iz slike izvlače postupcima segmentacije.



Slika 3.2: Određivanje područja na slici koja pripadaju koži uz normalni i smanjeni utjecaj svjetlosti.

Potom je potrebno područja slike dobivena segmentacijom klasificirati kao lica od-

nosno ne-lica. U skladu sa [5] klasifikacija se obavlja na sljedeći način. Na samom početku postoji skup lica koja čine skup za učenje, od tih se lica stvara slika prosječnog lica i za svakog kandidata lica se računa korelacija kandidata sa prosječnim licem. Ako je ta korelacija niska kandidat se odbacuje, inače ako je područje kandidata dovoljno veliko on se prihvaća kao područje lica. Test korelacije je sličan načinu na koji se maximal rejection classifier (MRC) koristi za detekciju lica opisanog u radu [17]. Prihvatanje kandidata na temelju veličine područja se obavlja koristeći prilagođavajuću granicu prihvatanja kako bi se omogućilo prihvatanje malih lica na slici, a istodobno odbacivalo manja područja za koje je prolazak testa korelacije moguć (dijelovi tijela poput ruku).

3.2. Prepoznavanje lica

Nakon što su lica detektirana, izrezana i skalirana lica se prepoznaju korištenjem kombinacije PCA i ICA algoritama. Iako se PCA algoritam može koristiti samostalno za prepoznavanje lica u ovom ćemo ga radu koristiti kao korak predprocesiranja slike. Uz pomoć PCA smanjujemo dimenzionalnost vektora značajki. Ovaj pristup pridonosi kvaliteti rješenja iz više razloga, značajke sa malom varijansom su vjerojatno posljedica šuma i nepoželjne su te se smanjuje složenost modela [9]. Nakon toga, lica se prepoznaju ICA metodom. Postoje dvije arhitekture te više algoritama ICA metode koji se mogu primjeniti u ovome slučaju.

3.2.1. ICA

ICA (eng. Independent Component Analysis) metoda slična je PCA (eng. Principal Component Analysis) te se može smatrati njenom generalizacijom. Razlika između njih leži u tome što PCA daje nekorelirane bazične vektore, a ICA statistički neovisne vektore. ICA rješava BSS (eng. blind source separation) problem, pokušava prikazati signal kao linearnu kombinaciju nezavisnih signala.

Ako je \mathbf{s} vektor nepoznatih nelinearnih signala, \mathbf{x} opaženi signal i \mathbf{A} matrica transformacije, pokušavamo riješiti sljedeću jednakost:

$$\mathbf{x} = \mathbf{A}\mathbf{s}$$

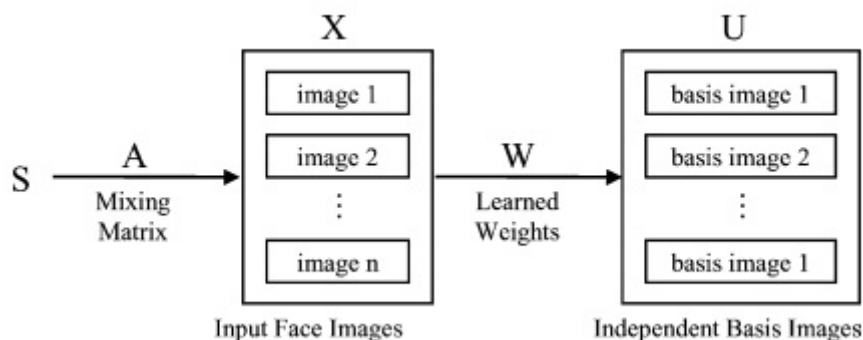
Pretpostavka je da \mathbf{A} nema inverz. ICA algoritmi pokušavaju naći \mathbf{A} ili \mathbf{W} iz sljedeće jednačbe:

$$\mathbf{u} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s}$$

Moguće je da ne postoji matrica \mathbf{W} koja u potpunosti zadovoljava ovu jednadžbu. \mathbf{W} se aproksimira iterativnim postupcima koji maksimiziraju neovisnost bazičnih vektora. Ona se ne maksimizira izravno već se odabiru funkcije koje imaju maksimum kada su vektori nezavisni. ICA algoritam InfoMax maksimizira entropiju postupkom gradijentnog spusta, dok FastICA maksimizira gdje je G nekvadratna funkcija, v slučajna varijabla, a c neka pozitivna konstanta. Svi ICA algoritmi konvergiraju u isto rješenje te ne postoje velike razlike u njihovim performansama [9].

Arhitektura I

Postoje dva načina za primjenu ICA-e na problemima prepoznavanja lica. Kod prvog načina ulazne slike lica \mathbf{X} promatraju se kao linearna kombinacija nezavisnih slika \mathbf{S} . Primjenom InfoMax algoritma dobivamo \mathbf{U} , aproksimaciju nezavisnih bazičnih slika. Slika 3.3 prikazuje opisanu arhitekturu. Ova arhitektura daje lokalizirane značajke te je bolja za prepoznavanje pokreta lica [9].

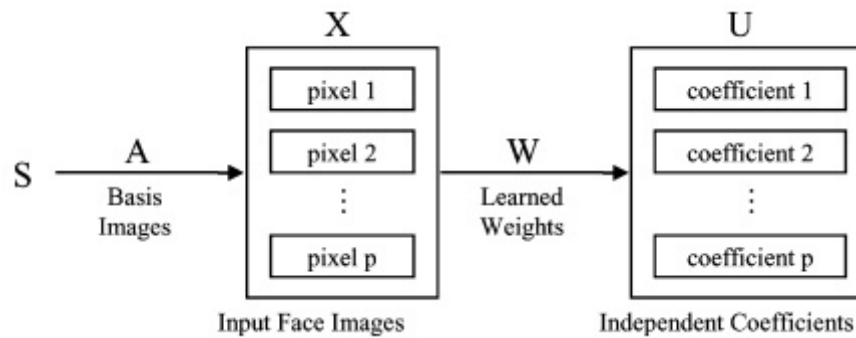


Slika 3.3: Arhitektura I

Arhitektura II

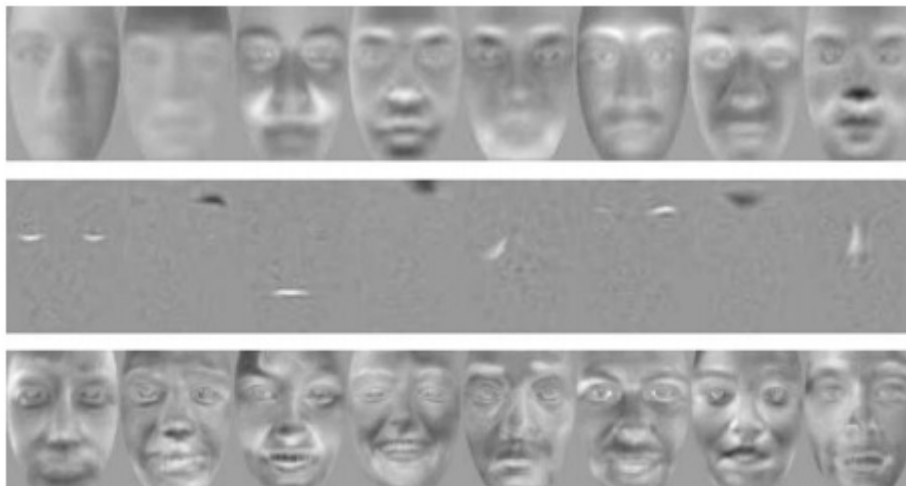
Za razliku od prve, druga arhitektura daje globalne značajke te postiže bolje rezultate kod prepoznavanja lica. Cilj ovog načina ICA jest pronalazak nezavisnih koeficijenata za ulazne podatke. Matrica \mathbf{X} predstavlja slikovne elemente (eng. pixels) slike, \mathbf{S} nezavisne koeficijente ulazne slike, dok je \mathbf{A} matrica bazičnih slika.

Na slici 3.5 možemo vidjeti kakve se značajke dobiju prikazanim arhitekturama. Gornji red prikazuje PCA značajke, srednji ICA značajke uz arhitekturu I, a donji red



Slika 3.4: Arhitektura II

ICA uz arhitekturu II. Jasno je vidljivo da su značajke ICA II arhitekture globalne. U radu [9] testirani su ICA i PCA algoritmi različitih arhitektura. Vidljivo je da ICA arhitektura II daleko najuspješnija u prepoznavanju lica, pa ćemo je koristiti i u ovom radu. ICA algoritam FastICA postiže malo bolje rezultate od InfoMax-a te njega biramo za ovaj rad.



Slika 3.5: Značajke dobivene PCA, ICA arhitektura I i ICA arhitektura II

3.3. Korištene baze

Zadatak projekta je detekcija i prepoznavanje lica u proizvoljnim scenama. Implementaciju sustava potrebno je testirati na standardiziranoj bazi podataka kako bi se rezultati mogli usporediti sa drugim implementacijama.

U trenutku pisanja nismo pronašli bazu koja bi se mogla koristiti za oba koraka algoritma (detekciju i prepoznavanje). Takva baza morala bi sadržavati slike proizvoljnih

scena u kojima su lica označena lokacijski (položaj lica na slici) i labelirana (identifikator osobe svakog označenog lica). Stoga predlažemo evaluaciju cjelokupnog algoritma u dva koraka. Prvo evaluiramo detekciju lica na prikladnoj bazi, a potom na drugoj bazi evaluiramo algoritam prepoznavanja. Iako ovaj pristup nije idealan, omogućava evaluaciju na standardnim bazama i usporedbu sa drugim sustavima.

Detekcija lica

Postoje brojne baze u kojima je lice prikazano frontalno ili pod nekoliko različitih kutova ¹. Većina tih baza sadrži vrlo strukturirane fotografije. Lice najčešće zauzima veliki dio slike, centralno je pozicionirano i prikazano u cjelokupnosti. U tom smislu one nisu prikladne za evaluaciju sustava za detekciju lica u proizvoljnim scenama.

Baza podataka FDDB ² jedna je od rijetkih koje sadrže slike uistinu proizvoljnih scena, s označenim područjima koja prikazuju lica. Sadrži 5171 lica označenih unutar 2845 fotografije. Često se koristi te postoje dobro dokumentirani rezultati za velik broj sustava za detekciju, opisano u [18]. Stoga smatramo da je baza prikladna za evaluaciju našeg sustava detekcije.

Prepoznavanje

Mnoge baze prikladne su za evaluaciju sustava prepoznavanja lica. Za takve baze bitno je da sadrže slike isključivo lica, kako bi se smanjila količina šuma u algoritmu prepoznavanja. Naravno, svaka slika mora biti označena jedinstvenim identifikatorom prikazane osobe. Kako bi sustav za detekciju bio što primjenjiviji poželjno je da su lica prikazana iz različitih kutova u različitim uvjetima osvjetljenja.

Od mnogih raspoloživih baza predlažemo evaluaciju na *Yale Extended Face Database B*. Radi se o bazi koja sadrži 16128 fotografija lica 28 osoba u 9 različitih poza i 64 uvjeta osvjetljenja. Prvi puta je predložena u radu [19].

3.4. Platforma programske implementacije

Za implementaciju sustava predlažemo korištenje programskog jezika Python u kombinaciji sa sljedećim potpornim knjižnicama.

¹http://robotics.csie.ncku.edu.tw/Databases/FaceDetect_PoseEstimate.htm

<http://www.face-rec.org/databases/>

²<http://vis-www.cs.umass.edu/fddb/>

NumPy

NumPy³ je knjižnica za numerički izračun. Sučelje (API) knjižnice dostupno je u jeziku Python, dok je implementacija pisana u jeziku C++ zbog brzine izvođenja. U našem sustavu NumPy će se koristiti za memorijske strukture pohrane slika te sve numeričke obrade.

OpenCV

OpenCV⁴ (engl. *Open Computer Vision*) knjižnica je za obradu digitalnih slika u kontekstu računalnog vida. Sučelje za jezik Python je dostupno. Koristimo ju za predprocesiranje i segmentaciju slike.

scikit-learn

Scikit-learn⁵ je potporna knjižnica za sustave strojnog učenja. Koristimo ju za implementaciju PCA i LCA algoritama.

³<http://www.numpy.org>

⁴<http://opencv.org>

⁵<http://scikit-learn.org>

4. Rezultati

4.1. Detekcija lica

Detekcija lica evaluira se na standardnoj bazi FDDB¹. Evaluacija se provodi 10-fold-cross-validacijom. Baza FDDB podjeljena je u 10 disjunktnih skupova istih veličina. Evaluacija se provodi slijenim izdvajanjem jednog skupa za testiranje, a treniranjem algoritma na ostalih 9 skupova. Treniranje se svodi na pronalazak optimalnih parametara (ugnježđenom 9-fold-cross-validacijom) za najbolju detekciju lica. Dobiveni detektor se koristi za mjerenje rezultata na testnom skupu.

4.1.1. Detekcija kože

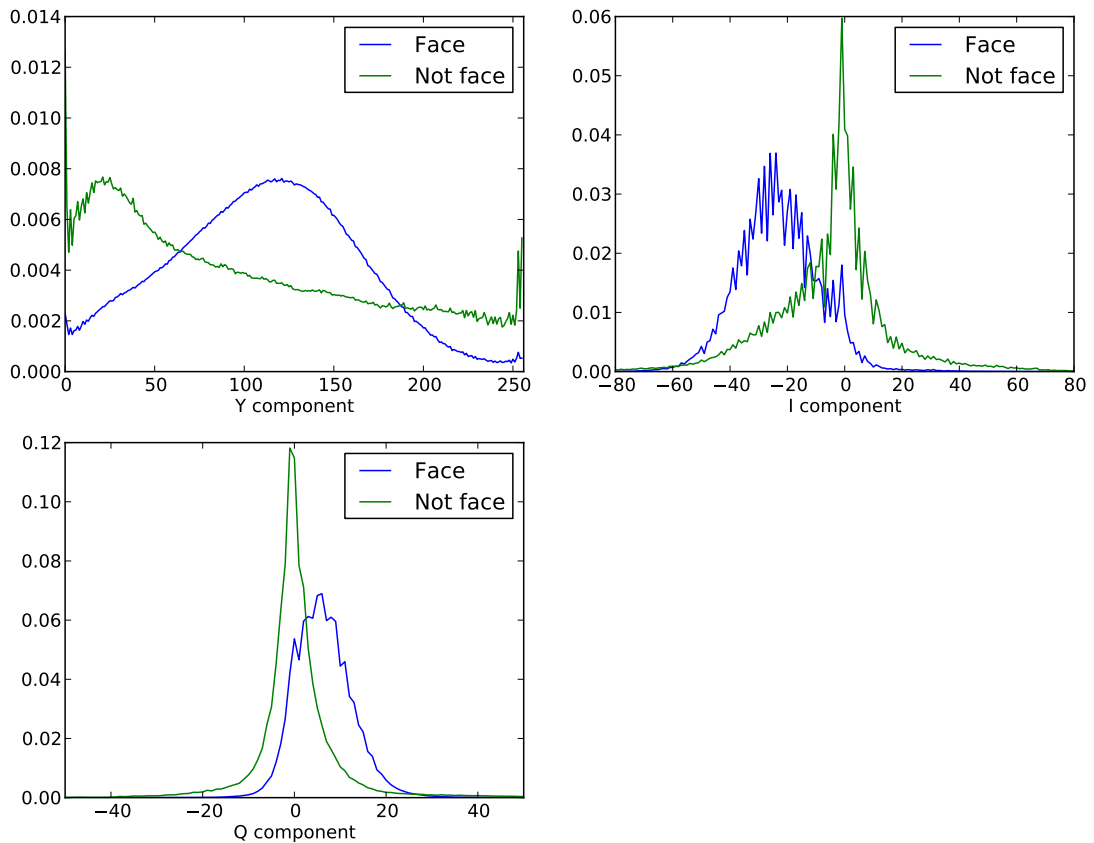
Detekcija kože provodi se u YIQ prostoru boja. Za određivanje tonova koji odgovaraju koži rade se histogrami boje (po komponentama) za sve piksele koji su označeni kao dio lica, te isto tako za piksele koji nisu dio lica.

Skup slika za treniranje se konvertira u YIQ prostor boja. Potom se koriste oznake FDDB baze za određivanje pripadnost piksela klasama (lice, ne-lice). Histogrami dobiveni za čitavu bazu FDDB prikazani su na slici 4.1.

Iz histograma se očituje da je razdvojenost razreda prisutna u svim komponentama, ali uz snažna preklapanja. Pikslel se stoga klasificira kao boja kože (klasa "lice") samo ako je unutar raspona označenog za lice svih triju komponenti. Rasponi klase lice se za svaku komponentu određuju algoritamski na sljedeći način:

1. Računa se razlika između histograma za "lice" i "ne-lice"
2. Pronalazi se maksimum M dobivene razlike
3. Područje klase "lice" se širi na obje strane od maksimuma
4. Širenje se zaustavlja kada je razlika manja od $\Delta * M$

¹<http://vis-www.cs.umass.edu/fddb/>



Slika 4.1: Histogrami boje (po YIQ komponentama) za razrede "lice" i "ne-lice", izračunati iz FDDB baze.

5. Parametar Δ se određuje eksperimentalno

Rezultati opisanog provođenja detekcije piksela kože prikazani su na slici 4.2. Gornji red na slici su originalne fotografije, ispod njih su pripadne maske.

Očituje se par reprezentativnih situacija. Na lijevom originalu koža je detektirana samo djelomično, uz velik komad krivo detektirane pozadine. Na srednjem originalu skoro uopće nije detektirana koža, dok je na desnom detekcija velikim dijelom uspješna. Inspekcija velikog broja dobivenih maski ukazuje da je detekcija kože temeljena na boji srednje uspješnosti. Zbog velike varijance boje kože (uslijed rasne raznovrsnosti ljudi, različitih uvjeta osvjetljenja i bijele točke fotografije), koža se često ne detektira. Istovremeno, pozadina fotografija sasvim legitimno može sadržavati piksele boje slične kožinoj, koji su posljedično označeni u maskama kao koža.

U radu [5], na kojem se temelji ovaj eksperiment, nije opisano koji postupci procesiranja maske se koriste. Stoga smo koristili procesiranje koje je davalo najbolje rezultate. Koristili smo algoritme otvaranja i dilatacije (koristeći knjižnicu OpenCV). U rezultirajućoj maski smo označili spojena područja (koristeći knjižnicu SciPy) te



Slika 4.2: Primjeri detekcije piksela kože na FDDB bazi.

izbacili područja koja veličinom iznimno odstupaju od prosječne veličine područja. Cijeli postupak rezultirao je osjetnim povećanjem kvalitete maski.

4.1.2. Klasifikacija područja kože kao lica

Konačni korak detekcije je klasifikacija segmenata maske kao lica ili ne korištenjem metrike udaljenosti od "prosječnog" lica, po uzoru na rad [5].

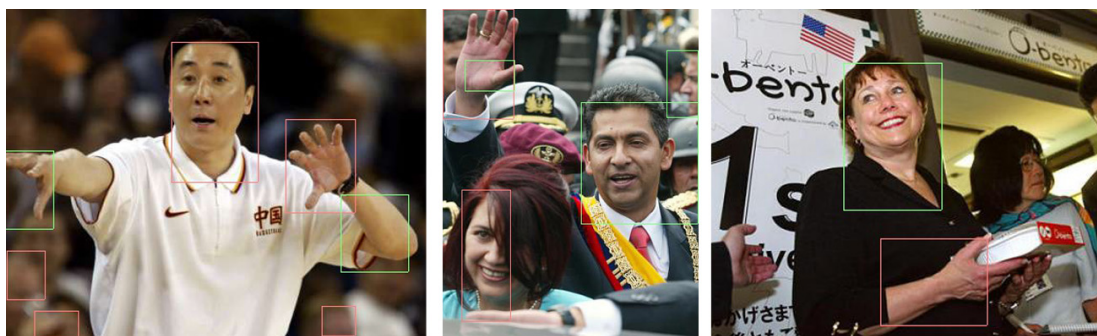
Možemo primjetiti da su prosječna lica nalik zamućenom licu, snimljenom frontalno u neutralnom osvjetljenju. Ovo je u skladu sa očekivanjima, ali otežava usporedbu sa licima koja su snimana iz proizvoljnih kutova, u različitim uvjetima osvjetljenja te su ponekad djelomično prekrivena. Dodatni faktor koji otežava usporedbu jest to da su maske (proizašle iz detekcije boje kože) vrlo rijetko ograničene isključivo na kožu lica. Često se jedan segment maske proteže na vrat, ramena i ruke. Ponekad su lica na fotografiji vrlo blizu jedno drugome te se bez sofisticiranijeg segmentiranja smatraju jednim područjem. Zbog ovih razloga, kao i već spomenutog problema kvalitetne detekcije boje kože, klasifikacija segmentiranih područja kao lica ili ne-lica korištenjem metrike distance od prosječnog dala je loše rezultate. Prosječno lice dobiveno za podskup baze prikazano je na slici 4.3.

Primjer konačne klasifikacije prikazan je na slici 4.4. Na tri prikazane fotografije označeni su zeleni i crveni pravokutnici. Oba tipa pravokutnika predstavljaju područja određena kao potencijalna lica na temelju maske piksela kože. Isječci slike svakog pravokutnika predloženi su klasifikatoru baziranom na prosječnom licu. Zeleni pravokutnici su područja koja je klasifikator označio kao lica, crveni pravokutnici područja



Slika 4.3: Prosječno lice dobiveno iz FDDB baze (skalirano, u originalu veličine 64×64 piksela).

koja je označio kao ne-lica. Vidimo da postoji relativno mala konzistentnost klasifikacije, koju smo opazili na većini slika testnog skupa.



Slika 4.4: Primjeri konačne detekcije lica na FDDB bazi.

Detektor lica se evaluira koristeći 10-fold cross-validaciju (standard predložen u tehničkom izvještaju FDDB baze). Evaluiraju se dvije mjere uspješnosti. Prva mjera je "kontinuirani" rezultat koji se računa kao omjer preklapanja detektiranih lica i označenih. Rezultat takve evaluacije je $14 \pm 0.01\%$ preklapanja. Najuspješniji detektori trenutno postižu oko 45% [18]. Druga mjera evaluacije je "diskretni" rezultat koji mjeri broj detektiranih lica, pri čemu se detekcija smatra uspješnom ako je preklapanje oznake i detekcije veća od 50%. Naš detektor uspješno locira 35% lica, dok trenutno najuspješniji detektori lociraju oko 65%. Pri tome je preciznost našeg detektora (broj točnih prepoznavanja naspram broja svih detekcija) 13%.

4.1.3. Detekcija, zaključak

U radu [5], na kojem se bazira naš postupak, mnogi koraci detekcije nisu kvalitetno opisani. Za evaluaciju se ne koristi neka od standardnih baza, već neobjavljena baza koju su pripremili sami autori. Stoga je nemoguće rekreirati njihov eksperiment i potvrditi impresivne rezultate. Smatramo da bi bilo poželjno da autori svoj sustav evaluiraju na FDDB (ili nekoj drugoj standardnoj) bazi, kako bi se realističnije utvrdila kompetitivnost njihovog algoritma postojećim rješenjima.

4.2. Prepoznavanje lica

Prepoznavanje lica evaluirano je na Yale² bazi podataka. Ona se sastoji od slika 39 osoba, s tim da je za svaka osoba predstavljena sa 64 slike. Slike se razlikuju po kutu izvora svjetlosti u odnosu na os kamere. Slika naziva yaleB03 _P00A-060E+20.pgm prikazuje osobu B03 sa izvorom svjetlosti zarotiranim za kut od -60 stupnjeva po osi x i 20 stupnjeva po osi y u odnosu na kameru. Koristena je baza sa izdvojenim licima (eng. *cropped data base*). Primjeri slika iz baze mogu se vidjeti na slici 4.5.



Slika 4.5: Lice sa različitim rotacijama izvora svjetlosti.

Baza je podijeljena na skup za učenje i skup za ispitivanje. Skup za učenje sastavljen je n slučajno odbranih slika svake osobe, gdje je n promjenjivi parametar. Potreba za uzimanjem više referentnih slika svake osobe javila se zbog raznovrsnosti slika u bazi. Na slikama 4.7 i 4.8 može se vidjeti utjecaj broja n na preciznost klasifikacije.

4.2.1. Smanjivanje dimenzionalnosti vektora značajki PCA algoritmom

Za smanjivanje broja značajki koristi se PCA algoritam. Ovim korakom iz slike izbacujemo PCA značajke male frekvencije za koje se pretpostavlja da pripadaju šumu. Također, smanjivanje dimenzionalnosti potrebno je jer smanjuje vrijeme generiranja ICA značajki, čime se postiže značajno ubrzanje učenja i klasifikacije. Uzima se broj najrelevantnijih značajki takav da suma njihovih frekvencija zajedno čini 99%.

²<http://vision.ucsd.edu/content/yale-face-database>

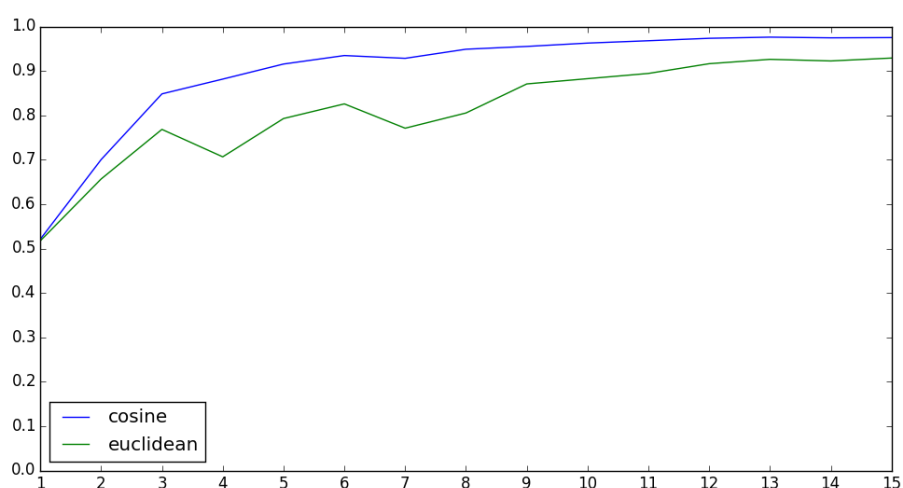
4.2.2. Prepoznavanje lica ICA algoritmom

Prepoznavanje lica provodi se tako da se nad skupom za učenje provede ICA algoritam. Dobiveni vektori značajki nazivaju se baznim slikama, čiji se primjeri mogu vidjeti na slici 4.6. Za sliku lica koja se klasificira izračunava se ICA vektor značajki. Na temelju njega izračunava se mjera sličnosti slike sa svakom od baznih slika. Mjera sličnosti može se temeljiti na kosinusnoj ili euklidskoj udaljenosti. Slika se klasificira u razred bazne slike s najvećom sličnošću. Najveća dobivena preciznost iznosi 97.6% za testni skup sa $n = 15$ slika za svaku osobu.



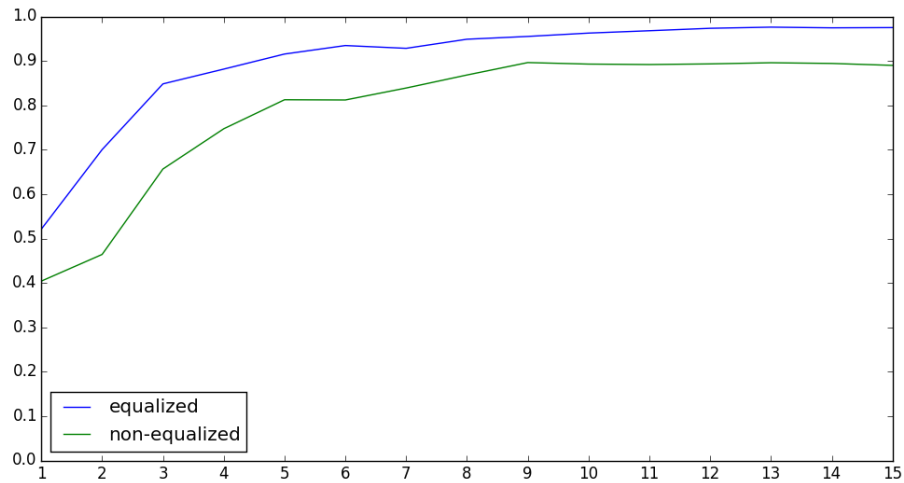
Slika 4.6: Bazne slike lica.

Na slici 4.7 možemo vidjeti preciznost za euklidsku i kosinusnu mjeru sličnosti u odnosu na veličinu skupa za učenje, n .



Slika 4.7: Preciznost klasifikacije za euklidsku i kosinusnu mjeru sličnosti u odnosu na veličinu skupa za učenje.

U radu [5] normalizacija histograma ulazne slike koristi se kao korak predprocesiranja koji poboljšava klasifikaciju. Na slici 4.8 vidi se kako normalizacija histograma utječe na preciznost klasifikacije.



Slika 4.8: Preciznost klasifikacije u odnosu na veličinu skupa za učenje sa i bez normalizacije histograma.

4.2.3. Prepoznavanje, zaključak

5. Zaključak

Zaključak.

6. Literatura

- [1] Erik Hjelmås and Boon Kee Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, September 2001.
- [2] Paul Viola and Michael Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:137–154, 2004.
- [3] Prabhu Utsav and Keshav Seshadri. Facial recognition using active shape models, local patches and support vector machines.
- [4] Andrew Senior, Rein-Lien Hsu, Mohamed Abdel Mottaleb, and Anil K. Jain. Face detection in color images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):696–706, May 2002.
- [5] D. N. Chandrappa and M. Ravishankar. Automatic face recognition in a crowded scene using multi layered clutter filtering and independent component analysis. In Ajith Abraham, Albert Y. Zomaya, Sebastián Ventura, Ronald Yager, Václav Snásel, Azah Kamilah Muda, and Philip Samuel, editors, *ISDA*, pages 552–556. IEEE, 2012.
- [6] Henry A. Rowley, Student Member, Shumeet Baluja, and Takeo Kanade. Neural network-based face detection. *IEEE Transactions On Pattern Analysis and Machine intelligence*, 20:23–38, 1998.
- [7] Cha Zhang and Zhengyou Zhang. Improving multiview face detection with multi-task deep convolutional neural networks. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 1036–1041. IEEE, 2014.
- [8] Stan Z. Li and ZhenQiu Zhang. Floatboost learning and statistical face detection. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 26(9):2004, 2004.

- [9] Bruce A. Draper, Kyungim Baek, Marian Stewart Bartlett, and J. Ross Beveridge. Recognizing faces with pca and ica. *Comput. Vis. Image Underst.*, 91(1-2):115–137, July 2003.
- [10] Zhimin Cao, Qi Yin, Xiaoou Tang, and Jian Sun. Face recognition with learning-based descriptor. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2707–2714, June 2010.
- [11] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments.
- [12] Xiaoguang Lu. Image analysis for face recognition.
- [13] Jeremiah R Barr, Kevin W Bowyer, and Patrick J Flynn. The effectiveness of face detection algorithms in unconstrained crowd scenes. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 1020–1027.
- [14] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. pages 511–518, 2001.
- [15] Joshua C. Klontz and Anil K. Jain. A case study on unconstrained facial recognition using the boston marathon bombings suspects, 2013.
- [16] Md. Abul Bashar Md. Hafizur Rahman, Suman Chowdhury. An automatic face detection and gender classification from color images using support vector machine.
- [17] Michael Elad, Yacov Hel-or, and Renato Keshet. Pattern detection using maximal rejection classifier. In *In Int. Workshop on Visual Form*, pages 28–30, 2000.
- [18] Vidit Jain and Erik Learned-Miller. Fddb: A benchmark for face detection in unconstrained settings. Technical Report UM-CS-2010-009, University of Massachusetts, Amherst, 2010.
- [19] A.S. Georgiades, P.N. Belhumeur, and D.J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6):643–660, 2001.

7. Sažetak

Sažetak.