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

**Sveučilišni studij**

**KLASIFIKACIJA EMOCIJA NA OSNOVI SLIKE LICA**

**Diplomski rad**

**Josip Baketarić**

**Osijek, 2017.**

**Sadržaj**

[1. UVOD 1](#_Toc478939178)

[2. RAČUNALNO RASPOZNAVANJE EMOCIJA 2](#_Toc478939179)

[2.1. Emocije i afekt 2](#_Toc478939180)

[2.2. Modaliteti podataka 3](#_Toc478939181)

[2.2.1. Psihološki modaliteti podataka 3](#_Toc478939182)

[2.2.2. Fiziološki modaliteti podataka 4](#_Toc478939183)

[2.3. Klasifikacija 5](#_Toc478939184)

[2.4. Postupci dohvaćanja značajki 6](#_Toc478939185)

[2.4.1. Histogram 6](#_Toc478939186)

[2.4.2. Lokalni binarni uzorci 7](#_Toc478939187)

[2.4.3. Gaborovi filteri 7](#_Toc478939188)

[2.4.4. Analiza glavnih komponenti 8](#_Toc478939189)

[2.5. Klasifikatori 9](#_Toc478939190)

[2.5.1. Naivni Bayesov klasifikator 9](#_Toc478939191)

[2.5.2. Metoda najbližih susjeda 9](#_Toc478939192)

[2.5.3. Umjetna neuronska mreža 10](#_Toc478939193)

[2.5.4. Stabla odluke 11](#_Toc478939194)

[2.5.5. Šume odluke 11](#_Toc478939195)

[2.5.6. Vektori podrške 12](#_Toc478939196)

[2.6. Viola-Jones algoritam 14](#_Toc478939197)

[2.7. Analiza performansi klasifikatora 15](#_Toc478939198)

[2.8. Prikaz rezultata 16](#_Toc478939199)

[2.9. Komercijalna rješenja za računalno raspoznavanje emocija 16](#_Toc478939200)

[3. RJEŠENJE ZA RASPOZNAVANJE EMOCIJA NA OSNOVI LICA 17](#_Toc478939201)

[3.1. Specifikacije i zahtjevi 17](#_Toc478939202)

[3.2. Korišteni alati i tehnologije 19](#_Toc478939203)

[3.2.1. OpenCV 19](#_Toc478939204)

[3.2.2. EmguCV 19](#_Toc478939205)

[3.2.3. Accord.NET 20](#_Toc478939206)

[3.2.4. AForge.NET 20](#_Toc478939207)

[3.2.5. Weka 20](#_Toc478939208)

[3.2.6. ARFF format 22](#_Toc478939209)

[3.3. Prikaz ključnih elemenata rješenja 22](#_Toc478939210)

[3.4. Prikaz izgleda, rad i uporaba rješenja 26](#_Toc478939211)

[3.5. Analiza ponašanja i performansi 30](#_Toc478939212)

[3.5.1. Analiza rezultata na trening podatcima 31](#_Toc478939213)

[3.5.2. Analiza rezultata na nepoznatim podatcima 32](#_Toc478939214)

[3.6. Osvrt i mogućnost unaprjeđenja 33](#_Toc478939215)

[4. ZAKLJUČAK 35](#_Toc478939216)

[5. LITERATURA 36](#_Toc478939217)

[6. SAŽETAK 39](#_Toc478939218)

[7. ŽIVOTOPIS 40](#_Toc478939219)

[8. PRILOZI 41](#_Toc478939220)

1. UVOD

„*Budućnost leži u projektiranju i prodaji računala za koje ljudi ne shvaćaju da su računala*“

Adam Osborne

Prilikom međuljudske komunikacije, a i komunikacije između ostalih živih bića, govor ima jako važnu ulogu. Jednako važnu, ako ne i važniju ulogu prilikom komunikacije imaju emocije koje vrlo često mogu promijeniti značenje riječi. Iako nema stroge definicije, emocije se često opisuju kao složena stanja ljudskih osjećaja koja rezultiraju u psihičkim promjenama koja utječu na ponašanje i razmišljanje [1]. Nekada je važnije kako je nešto izgovoreno nego što je izgovoreno. U svijetu gdje je sve veća komunikacija između računala i čovjeka, nije dobro zanemariti jedan tako važan aspekt ljudske prirode kao što su emocije. Znanstvena disciplina koja se bavi načinom na koji čovjek komunicira sa računalom se naziva interakcija čovjeka i računala (engl. *Human-Computer Interaction*, HCI). HCI se do sada uglavnom bazirao na kognitivan pristup koji opisuje koliko lako korisnik može koristiti sustav. Slijedeći korak je uvođenje emocija u obzir gdje bi se sustav prilagođavao trenutnoj emociji korisnika. Emocije u velikoj mjeri utječu na čovjekovo odlučivanje. Vrlo je važno za tvrtke, koje posluju preko interneta ili se bave razvojem programske podrške, da zadrže svoje klijente i pridobiju nove jer je konkurencija u skoro svakom području jako velika. Korisnik će se vrlo vjerojatno vratiti na stranicu ili kupiti program koji se prilagođava njemu osobno nego onaj koji to ne radi. Trenutno postoji više različitih tvrtki koje pružaju uslugu računalnog prepoznavanja emocija, ali problem još nije u potpunosti riješen i do sada niti jedno rješenje nije dovoljno dobro za korištenje u ozbiljnije svrhe. Problem prepoznavanja emocija će biti riješen kada (i ako) se uspije razviti rješenje koje će moći prepoznati emocije bolje ili barem kao čovjek.

Zadatak ovog rada je istražiti postojeća rješenja i implementirati vlastito u području računalnog prepoznavanja emocija. Ovaj rad je baziran na prepoznavanju glumljenih emocija sa frontalnih slika lica.

Rad je podijeljen u četiri cjeline od kojih je prva uvod u problem. Druga cjelina detaljnije opisuje emocije, načine na koje se iskazuju te različite računalne postupke. Treće poglavlje sadrži konkretan opis programskog rješenja te dobivene rezultate i njihovu analizu. Prikazani su rezultati testiranja na programu nepoznatoj bazi te na kraju je dan osvrt na implementaciju. Četvrto poglavlje je zaključak u kojemu je dan kratak osvrt na cijeli rad te potencijalne mogućnosti korištenja rješenja.

1. RAČUNALNO RASPOZNAVANJE EMOCIJA

Proučavanje emocija postoji već jako puno godina i veliki broj radova napisan je na tu temu. Jedan od temeljnih radova na spomenutu temu je Darwinov rad iz 1872. godine koji povezuje fizičke aktivnosti, kao što su na primjer podizanje obrva u slučaju iznenađenja ili zbunjenosti, sa stanjem uma [2]. Raspoznavanje emocija složen je proces koji uključuje kontekst, govor tijela, kulturna obilježja, izraze lica itd. [3]. Računalno raspoznavanje emocija bi uvelike pridonijelo velikom broju djelatnosti, od medicine, psihologije, poslova koji, ako nisu odrađeni dobro, mogu rezultirati smrti ili povredom i slično. U kontroli zračnog prometa bi se na temelju emocija kontrolora moglo znati kada je potrebno uzeti odmor. To se može primijeniti na svakog vozača i radnika koji upravlja opremom opasnom po život. Informacije o emocijama u svijetu video igrica i filmova bi pružale jako važnu i nezamjenjivu povratnu informaciju o tome koliko je neka igra/film zanimljiv, strašan, smiješan i slično. Psihijatri bi mogli lakše pratiti promjene raspoloženja njihovih pacijenata i samim time bi lakše napredovali u svojim istraživanjima. Programi za učenje preko računala bi mogli detektirati kada je korisniku postalo dosadno i promijeniti sadržaj ili prikazati neku zanimljivu činjenicu. Zbog brojnih mogućih primjena, jako puno se istražuje na polju računalnog prepoznavanja emocija te je to jedan od glavnih problema u računalnoj industriji ovoga doba.

* 1. Emocije i afekt

Kada je riječ o emocijama, važno je razlikovati osjećaje, emocije i afekte. Dok su emocije socijalne prirode, osjećaji su osobne, a afekti prepersonalne prirode [X4]. Detaljniji opisi spomenutih pojmova su navedeni u slijedećem tekstu, prema [X5]. Osjećaj je senzacija temeljena na iskustvu te je poseban za svakog čovjeka. Emocija prikazuje osjećaj Za razliku od osjećaja, emocija može biti iskrena ili glumljena te je to glavna razlika. Emocije su slika koja prikazuje ostatku svijeta, a osjećaj se zadržava za sebe. Za razliku od emocije i osjećaja, afekt je više apstraktan pojam. Afekt je način na koji se tijelo priprema za reakciju u određenoj situaciji te pojačava svijest o trenutnom stanju.

Postoji više vrsta podjela emocija, u ovom radu će se razmatrati ona od Roberta Plutchika [X6]. Plutchik kaže da postoji osam osnovnih emocija koje su se razvile radi preživljavanja vrste, a da su ostale izvedene iz osnovnih. To su: ljutnja, strah, tuga, gađenje, iznenađenje, iščekivanje, povjerenje i sreća. On je također izradio kotač emocija u kojemu prikazuje osnovne i izvedene emocije. Svaka osnovna emocija ima svoju suprotnost i prikazana je na kotaču suprotno od nje. Povezuju se određene boje sa emocijama, na primjer ljutnja je prikazana crvenom bojom, a radost žutom što se može vidjeti na slici 2.1.



Slika 2.1. Kotač emocija, [X6]

* 1. Modaliteti podataka

Modaliteti podataka koji se koriste za prepoznavanje emocija se mogu grupirati u dvije cjeline: psihološki i fiziološki, [X7]. Psihološki podatci kao i fiziološki mogu se razlikovati od kulture do kulture te je i to potrebno uzeti u obzir. Psihološki modaliteti se mjere na način da se posebni senzori spajaju na ljudsko tijelo, a fiziološki pomoću promatranja pokreta, izraza lica i tako dalje [X7].

Psihološki modaliteti podataka

Jedna od najvažnijih tehnologija u neuroznanosti je elektroencefalografija (engl. *Electroencephalography*, EEG) [X7]. EEG se bazira na očitavanju aktivnosti mozga te pomoću rezultata iz koje je strane mozga došla informacija može odrediti dali je emocija proizašla iz pozitivnog ili negativnog utjecaja. Ovaj način dobivanja podataka zahtjeva da se na glavu korisnika stave elektrode koje očitavaju moždane valove [X7].

Elektrookulografija (engl. *Electrooculography*, EOG) je mjerenje pokreta očiju gdje se stavljaju elektrode oko očiju i mjere se pokreti mišića [X7]. Drugi način prikupljanja ove vrste podataka bi bio pomoću kvalitetne kamere.

Elektromiografija (engl. *Electromyography*, EMG) je metoda pomoću koje se mjere pokreti mišića [X7]. Elektrode se stavljaju na različita mjesta na tijelo i na temelju podataka se mogu prepoznati trzaji, pokreti tijela i slične aktivnosti koje mogu poslužiti za otkrivanje emocije.

Elektrodermalna aktivnost (engl. *Electrodermal Activity*, EDA) je metoda gdje se stavljaju elektrode na vrhove prstiju i provjerava se koliko je koža sposobna voditi struju. Na provodnost struje utječe koliko je koža znojna te se na taj način može jako dobro otkriti razina stresa i tjeskobe [X7]. EDA se obično očitava između dva prsta na istoj ruci.

Volumen krvi (pletizmografija) je dobar indikator stresa i tjeskobe. Senzor se obično stavlja na prst koji mjeri otkucaje srca i zasićenost krvi kisikom [X7].

Disanje je također dobar pokazatelj emocija. Brzina udisaja često može poslužiti kao kvalitetna informacija u prepoznavanju emocija [X7].

Temperatura tijela se također mijenja promjenom emocija, pogotovo ako su to sreća, ljutnja i tuga [X7].

Fiziološki modaliteti podataka

Kao što se emocije očituju psihološki, tako se njihov utjecaj vidi i na „površini“. Govor, izraz lica i pokret su glavni pokazatelji emocija u fiziološkom smislu [X8]. Fiziološke podatke je puno lakše dobiti od psiholoških jer ne trebaju posebni senzori za očitavanje. Glavni načini za prikupljanje fizioloških podataka su kamera i mikrofon.

Lice igra jako važnu ulogu u prikazivanju ljudskih emocija. Izrazi na licu su definirani pokretima mišića ispod kože lica te prikazuju emociju osobe i samim time pripadaju neverbalnoj komunikaciji [X9]. Pokreti mišića lica i njihova manifestacija su organizirani po sustavu akcijskih kodova lica (engl. *Facial Action Coding* System, FACS) [X9]. FACS se sastoji od više akcijskih jedinica (engl. *Action* *Unit*, AU) koji opisuju pojedine pokrete mišića lica. Kombinacija određenih akcijskih jedinica može služiti za otkrivanje emocija. U tablici 2.1 su prikazani osnovi akcijski kodovi, a u tablici 2.2 akcijski kodovi koji grupirani otkrivaju određene emocije [X9]. Intenzitet FASC-a se označava slovima od A do E, gdje A predstavlja minimalnu vrijednost, a B maksimalnu. Uz intenzitete još se označavaju i strane na kojima se pojavljuje FASC. Slovo R označava da se FACS pojavljuje na desnoj strani lica, slovo L označava da se FACS pojavljuje na lijevoj strani lica, a slovo U označava pojavljivanje FACS-a na samo jednoj strani, ali ta strana nije specificirana [X9].

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AU** | **FACS naziv** | **AU** | **FACS naziv** | **AU** | **FACS naziv** |
| 0 | Neutralno lice | 17 | Brada dignuta | 33 | Obraz napuhnut |
| 1 | Dizanje unutarnje obrve | 18 | Nabor od use | 34 | Obraz napuhnut (manje) |
| 2 | Dizanje vanjske obrve | 19 | Jezik se vidi | 35 | Obraz uvučen |
| 4 | Spuštanje obrve | 20 | Usne rastegnute | 36 | Jezi natečen |
| 5 | Dizanje gornje usne | 21 | Vrat zategnut | 37 | Brisanje usana |
| 6 | Dizanje obraza | 22 | Usne oblika O | 38 | Nosnice raširene |
| 7 | Zatezanje usna | 23 | Usne zategnute | 39 | Nosnice skupljene |
| 8 | Usne jedna prema drugoj | 24 | Usne stisnute | 41 | Koža iznad nosa spuštena |
| 9 | Naboran nos | 25 | Usne odvojene | 42 | Unutarnja obrva spuštena |
| 10 | Dizanje gornjeg kapka | 26 | Čeljust spuštena | 43 | Oči zatvorene |
| 11 | Bore od osmjeha | 27 | Usta rastegnuta | 44 | Obrve skupljene |
| 12 | Povlačenje kraja usne | 28 | Sisanje usne | 45 | Treptaj |
| 13 | Oštro povlačenje usne | 29 | Čeljust prema naprijed | 46 | Namigivanje |
| 14 | Rupica | 30 | Čeljust prema strani |  |  |
| 15 | Kut usne pritisnut | 31 | Čeljust stisnuta |  |  |
| 16 | Donja usna pritisnuta | 32 | Griženje usne |  |  |

Tablica 2.1. Osnovni AU kodovi, [X9]

|  |  |
| --- | --- |
| **Emocija** | **AU** |
| Sreća | 6 + 12 |
| Tuga | 1 + 4 +15 |
| Iznenađenje | 1 + 2 + 5B + 26 |
| Strah | 1 + 2 + 4 + 5 + 7 + 20 + 26 |
| Ljutnja | 4 + 5 + 7 + 23 |
| Gađenje | 9 + 15 + 16 |
| Prezir | R12A + R14A |

Tablica 2.2. Grupa AU koje otkrivaju određene emocije, [X9]

Osim izraza lica, jako veliku ulogu u prepoznavanju emocija ima način govora. Govor je osnovna vrsta međuljudske komunikacije. Osim poruke, važno je promatrati kako je poruka izgovorena to jest semantiku. Često način na koji je poruka izrečena može promijeniti njeno značenje. Visina tona, energija, brzina govora, podrhtavanje i slične informacije mogu pružiti jako dobru informaciju o trenutnoj emociji govornika [X8]. Usporenje u brzini govora, smanjen ton i energija te zijevanje odaje da se radi o dosadi dok povećanje u brzini govora i pojačana visina može označavati sreću. Novija istraživanja uključuju kombinaciju akustičnih podataka sa izrečenim riječima te time povećavaju preciznost [X8].

* 1. Klasifikacija

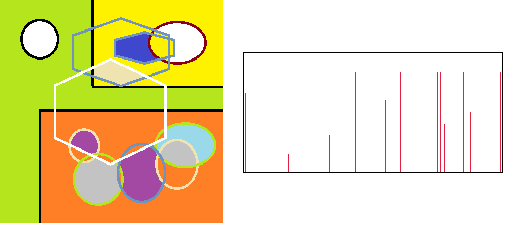
Klasifikacija je znanstvena disciplina koja za zadatak ima razvrstavanje objekata u klase [7, str. 1]. Kako bi se objekt mogao klasificirati potrebno je predstaviti ga sa određenim podatcima koji ga vjerodostojno opisuju. Ti podatci se još nazivaju i značajke (engl. *Features*) [7]. Značajke se dobivaju pomoću različitih operacija nad podatcima. Primjer jedne takve značajke je srednja vrijednost dobivena iz nekog objekta. Iz razloga što jedna značajka ne može vjerodostojno predstavljati neki objekt, barem ne u većini slučajeva, koristi se više značajki. Skup značajki koji opisuju jedan objekt naziva se vektor značajki (engl. *Features* *vector*) [7]. Više vektora značajki čini jedan skup podataka sa kojim se onda može raditi. Takav jedan skup se naziva i trening skup. Algoritmi koji koriste trening skup radi izgradnje modela te pomoću njega mogu odrediti klasu nepoznatog objekta se nazivaju klasifikatori (engl. *Classifiers*) [7]. Preciznost klasifikacije ne ovisi samo o klasifikatoru nego i o kvaliteti i broju značajki.

* 1. Postupci dohvaćanja značajki

Kao što je prethodno spomenuto, značajke imaju jako veliku važnost prilikom klasifikacije i vrlo je bitno da značajke kvalitetno predstavljaju podatke ukoliko se želi postići precizna klasifikacija. Jedna značajka može biti jako dobra za jednu vrstu problema dok za drugo ne baš toliko i iz toga razloga su razvijeni brojni algoritmi koji služe da dobivanje značajki iz podataka.

Histogram

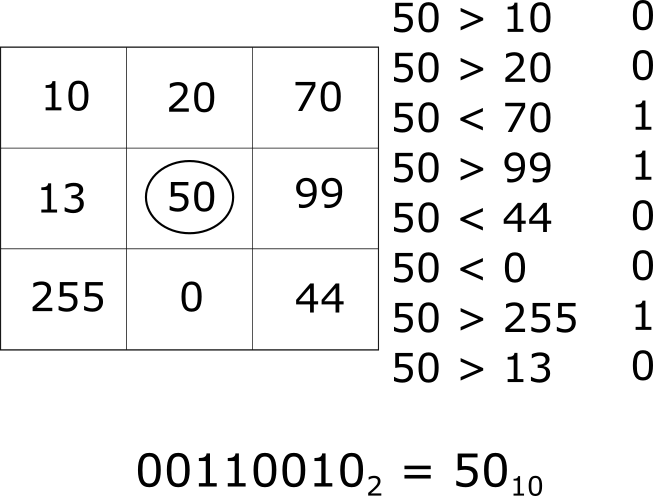
Histogram je metoda grafičkog prikaza podataka koji pokazuje broj pojavljivanja vrijednosti u određenom objektu [X10]. Ako bi se za neku sliku veličine 100x100 koristio histogram, tada bi za svaku moguću vrijednost piksela (0-255) bio dobiven jedan broj koji predstavlja broj pojavljivanja tog piksela u slici. Na taj način bi se dimenzionalnost vektora značajki smanjila sa 10000 na samo 255. Primjer histograma za sliku je prikazan na slici 2.2 gdje se na lijevoj strani vidi slika, a na desnoj broj pojavljivanja piksela.



Slika 2.2. Histogram slike

Lokalni binarni uzorci

Lokalni binarni uzorci (engl. Local Binary Pattern, LBP) je konkurentna metoda koja se koristi prilikom analize lica, tekstura, pokreta [X11]. LBP je baziran na uspoređivanju središnjeg sa susjednim pikselima. Uspoređuju se intenziteti susjednih piksela sa središnjim i ako su manji dodjeljuje im se vrijednost nula, a u slučaju većega intenziteta se dodjeljuje vrijednost jedan. Od dobivenih nula i jedinica se formira binarni broj koji se potom pretvara u dekadski te predstavlja LBP za promatrani piksel [X11]. Proces formiranja jedne LBP značajke prikazan je na slici 2.3.



Slika 2.3. Formiranje LBP značajke

Gaborovi filteri

Gaborov filter je kompleksna sinusoidna funkcija, jednadžba 2-1, koja se primjenjuje na određenim frekvencijama i orijentacijama. Koristi se prilikom detekcije kutova, granica, klasifikacija tekstura i slično [11].

|  |  |  |
| --- | --- | --- |
|  |  | (2-1) |
|  |  | (2-2) |
|  |  | (2-3) |

gdje je:

-*g* – Gaborova funkcija,

-*x*, y - koordinate

-– valna dužina,

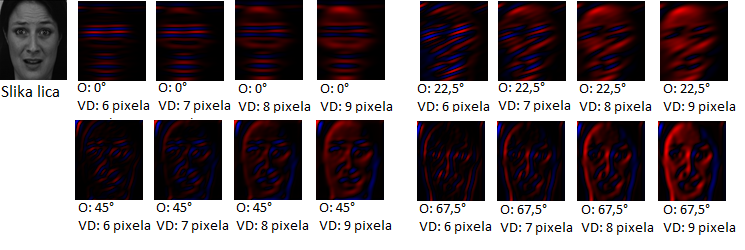
-– orijentacija,

- – Fazni pomak,

- – standardna devijacija,

- - omjer slike.

Na slici 2.4. se nalazi primjer Gaborovih filtera primijenjenih na sliku lica u četiri različite orijentacije i četiri različite valne dužine (sve ukupno 16 filtera).

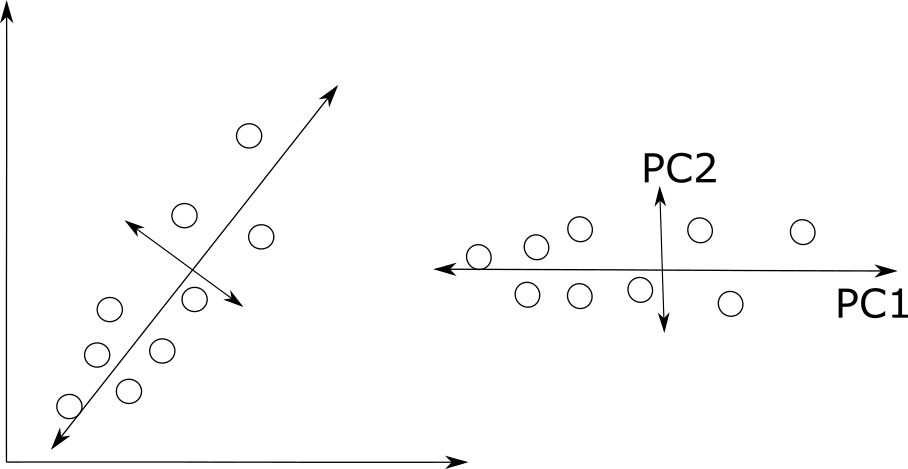


Slika 2.4. Gaborovi filteri primijenjeni na sliku lica

Dvije najvažnije komponente Gaborovih filtera su valna dužina i orijentacija. Orijentacija određuje koje će se linije na slici prepoznavati i vrijednost joj se kreće od 0° do 360°. Što se tiče valne dužine, ona određuje širinu filtera koji se primjenjuje i što je širi, više linija je moguće naći, ali je moguće i pogriješiti, ako je filter preširok. Nakon što su definirani filteri, takozvana banka filtera, oni se primjenjuju na sliku i kao rezultat dobiva se nova slika koja ima visok intenzitet piksela gdje se poklapa s filterom.

Analiza glavnih komponenti

Analiza glavnih komponenti (engl. *Principal component analysis*, PCA) je statistička metoda koja služi za smanjenje dimenzionalnosti podataka, detekciju ne pripadajućih vrijednosti i slično [12]. Jedna od glavnih primjena PCA je smanjivanje dimenzionalnosti. PCA prebacuje podatke u manju dimenziju na način da gleda najrelevantnije podatke, a ostale zanemaruje i time se uz male gubitke informacija postigne velika ušteda u smislu brzine izračuna. PCA određuje za svaku dimenziju koliko ona sadrži kvalitetnih ili korisnih informacija i poreda ih u smislu osnovnih komponenti. Ako bi postojao prostor od 200 dimenzija, PCA bi odredio 200 glavnih komponenti koje bi bile poredane po korisnosti. Prva glavna komponenta bi bila i najkorisnija, pa iza nje druga i tako dalje. Na taj se način može odrediti koliko će se koristiti dimenzija i izbjegavanje nepotrebnih podataka koje uz dodatni napor prilikom računanja mogu unijeti i pogrešku. Na slici 2.5. prikazan je dvodimenzionalni graf s podatcima. Iz grafa se vidi da se podatci prostiru skoro kao po pravcu uz malo odstupanja. PCA prebacuje podatke u dvodimenzionalni prostor u kojemu se umjesto osi, nalaze glavne komponente. Vidi se da su podatci rasprostranjeni po glavnoj komponenti 1, a značajno manje po glavnoj komponenti 2. Ako bi se htjela smanjiti dimenzionalnost podataka bez puno gubitaka, može se promatrati samo projekcija podataka na glavnu komponentu 1 jer ona sadrži većinu informacija.



Slika 2.5. PCA

* 1. Klasifikatori

Kao što je već rečeno, klasifikatori su algoritmi koji služe za prepoznavanje klase ulaznog podatka. Ovisno o količini korištenih značajki bira se odgovarajući klasifikator. Postoji veliki broj klasifikatora koji se koristi i svaki ima neke svoje prednosti i svoje nedostatke. Dalje u tekstu će biti opisano par poznatijih klasifikatora koji se često koriste.

Naivni Bayesov klasifikator

Naivni Bayesov klasifikator je klasifikator koji se temelji na računanju vjerojatnosti [X12]. Ako se koristi veliki broj značajki tada ovaj algoritam postaje spor i težak za korištenje. Da bi se uopće mogao koristiti potrebno je poznavati vjerojatnost da neki objekt pripada određenoj klasi. Ta vjerojatnost se može izračunati iz trening podataka. Pretpostavlja se da su sve značajke međusobno neovisne.

Metoda najbližih susjeda

Metoda najbližih susjeda (engl. *Nearest Neighbour*, NN) je jedna od najjednostavnijih metoda. Računa udaljenost ulaznog vektora značajki od svakog iz trening skupa te se pridodaje klasa koja mu je najbliža [X12]. Udaljenost u n-dimenzionalnom prostoru računa se pomoću formule za Euklidovu udaljenost koja je prikazana jednadžbom (2-4)

|  |  |  |
| --- | --- | --- |
|  |  | (2-4) |

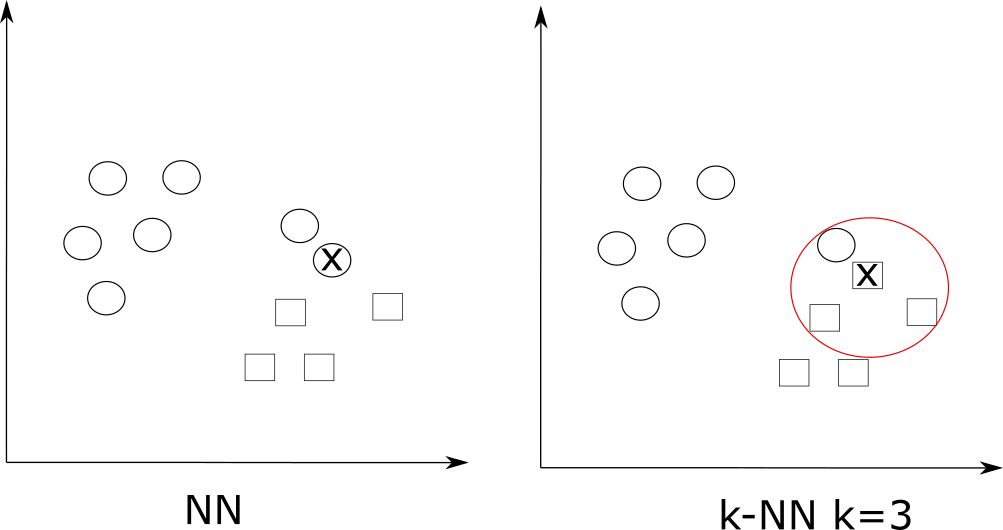
gdje je:

- *d* – udaljenost,

- *n* – broj dimenzija,

- *X* i *P* – točke u prostoru.

K-najbližih susjeda (engl. *k-Nearest Neighbour*, k-NN) je vrsta NN metode koja također računa udaljenosti, ali u obzir uzima k najbližih udaljenosti i dodjeljuje klasu onu koje ima više u blizini te ta klasa ne mora nužno biti i najbliža. Prikaz NN algoritma i k-NN algoritma dan je na slici 2.6. Na slici su prikazane dvije klase označene sa krugom i pravokutnikom. Znak X označava ulaznu klasu kojoj je potrebno pridružiti klasu. U slučaju NN algoritma klasa je pridružena krugu jer je on najbliži, a u drugom slučaju je pridružena pravokutniku iz razloga što ih ima više u njegovoj blizini.

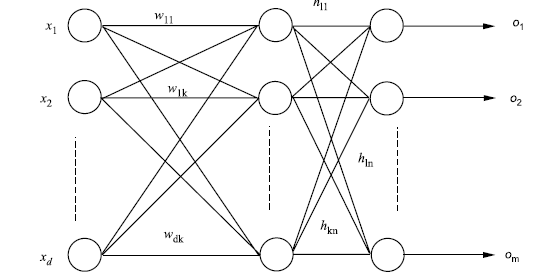


Slika 2.6. NN i k-NN algoritam

Umjetna neuronska mreža

Umjetna neuronska mreža (engl*. Artificial Neural Network*, ANN) je skup umjetnih neurona što se bazira na principu mozga. Neuronska mreža je sastavljena od umjetnih neurona od kojih je svaki funkcija što na izlazu daje logičku jedinicu ako suma ulaza prelazi zadani prag, u slučaju da ne prelazi onda daje logičku nulu [X12].

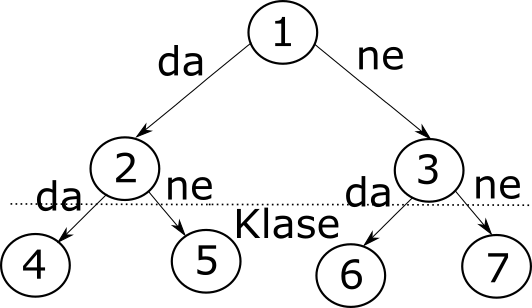
Višeslojni perceptron (engl. *MultyLayer Perceptron*, MLP) je vrsta neuronske mreže koji može raditi sa nelinearnim podatcima. On sadrži jedan ili dva skrivena sloja koji mu to omogućavaju [X12]. Na slici 2.7. se nalazi MLP sa jednim skrivenim slojem.



Slika 2.7. MLP s jednim skrivenim slojem, [X12]

Stabla odluke

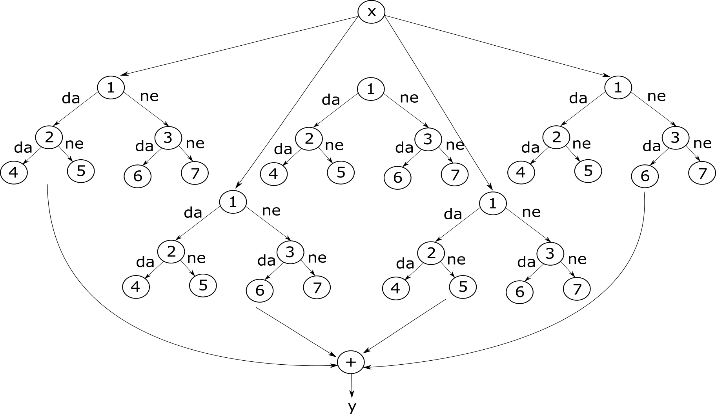
Stabla odluke su klasifikator koji ima strukturu stabla. Sastoji se od čvorova, grana i lista koji predstavlja krajnji čvor. Čvor i grana služe prilikom odluke o klasi, a list na kraju stabla predstavlja klasu [X12]. Stablo koje ima samo po dvije odluke na svakom čvoru se naziv binarno stablo i prikazano je na slici 2.8. Stablo se kreira na način da se najvažnije značajke stavljaju na početak stabla, a ostale manje važne pri dnu [X12]. Ako se koristi veliki broj značajki tada ovakav klasifikator nije primjeren jer veličina stabla se može eksponencijalno povećavati.



Slika 2.8. Binarno stablo

Šume odluke

Šume odluke (engl. Random forest) su klasifikatori koji su sastavljeni od više stabala odluke [X12]. Ovaj algoritam je poznat po preprilagođavanju (engl. Overfitting) što znači da se previše prilagodi trening setu, a na stvarnim podatcima je preciznost jako manja. Svako stablo se kreira sa podskupom trening podataka te se ta metoda naziva odvajanje (engl. Bagging). Postupak klasifikacije se izvodi na način da se svakom stablu da da odredi klasu i na kraju se uzima klasa koja ima najviše glasova. Šuma odluka je prikazana na slici 2.9.



Slika 2.9. Šuma odluka

Vektori podrške

Metoda vektora podrške (engl. *Support Vector Machines*, SVM) je binarni klasifikator koji se koristi pri klasifikaciji i regresiji [7, 119. str. – 138. str]. Prostorno promatrano, SVM pokušava odvojiti značajke suprotnih klasa hiper-ravninom tako da je razmak između najbližih elemenata i hiper-ravnine najveći. Potrebno je dobivene podatke prvo grupirati i potom izračunati optimalnu hiper-ravninu. U slučaju linearno odvojivih podataka, što znači da se može konstruirati hiper-ravnina koja odvaja uspješno dijeli sve podatke na dvije različite klase, hiper-ravnina se opisuje izrazom (2-5).

|  |  |  |
| --- | --- | --- |
|  |  | (2-5) |

gdje je:

-*w* – normala hiper-ravnine,

- - okomita udaljenost hiper-ravnine od ishodišta

- – podatak.

Hiper-ravnine koje odvajaju podatke i između sebe imaju najveću udaljenost su margine i mogu se opisati formulama (2-5) i (2-6), te margine se još nazivaju i vektori podrške.

|  |  |  |
| --- | --- | --- |
|  |  | (2-6) |
|  |  | (2-7) |

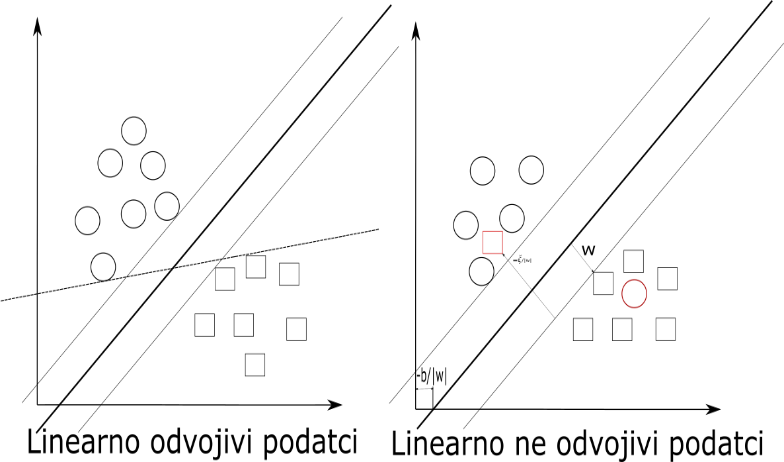
Ako se želi SVM koristiti i na linearno neodvojive podatke, potrebno je dozvoliti određenu grešku prilikom stvaranja hiper-ravnine. Kao i prethodno, u ovom slučaju se margine pokušavaju konstruirati sa što većim razmakom, ali pritom da je što manje pogrešno klasificiranih elemenata. Na slici 2.10. prikazana je hiper-ravnina u slučaju linearno neodvojivih i linearno odvojivih podataka. Odabir hiper-ravnine svodi se na pronalaženje:

|  |  |  |
| --- | --- | --- |
|  | takav da | (2-8) |

gdje je:

-*C* – faktor pogreške,

- - mjera udaljenosti promatranog podatka od pripadajućeg vektora podrške.

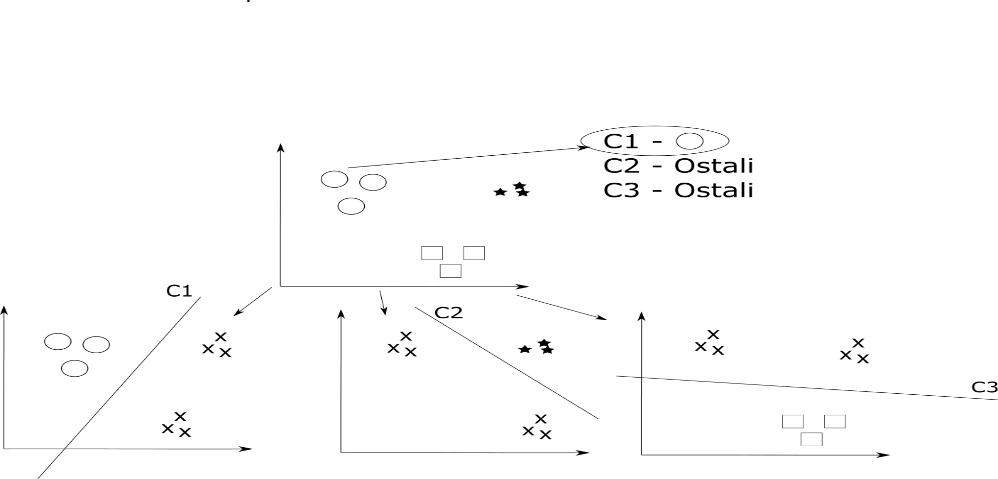


Slika 2.10. SVM sa linearno odvojivim i neodvojivim podatcima

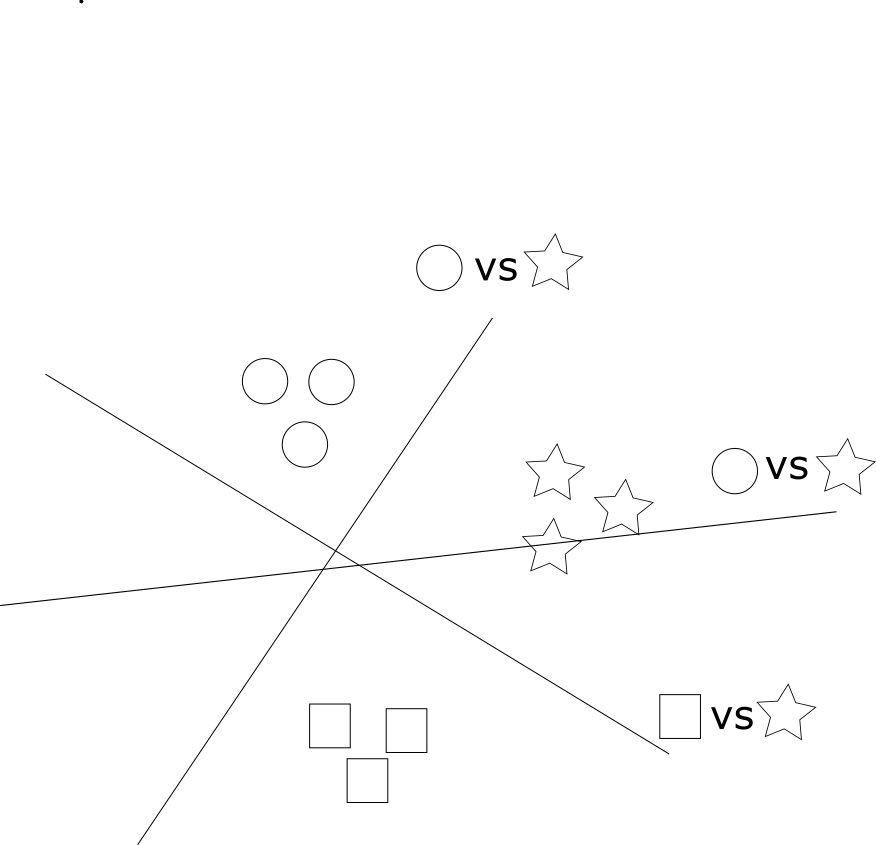
Iako binarni, klasifikator SVM može se koristiti i prilikom više klasne klasifikacije. Dva su moguća načina na koje se SVM može koristiti kao više-klasni klasifikator:

1. Konstruirati n klasifikatora koji svrstava podatke u jednu od klasa ili u preostale n-1 klase. Uz rezultat klasifikacije, klasifikator vraća i mjeru sigurnosti u svoj odabir. Nakon što svi klasifikatori odrade klasifikaciju, u obzir se uzimaju samo oni koji nisu podatku pridružili klasu „ostali“. Rezultat se uzima od klasifikatora sa najvećom mjerom sigurnosti. Ovakav način rada zove se jedan protiv svih (engl. *One versus all*). U slučaju da svi klasifikatori odaberu klasu „ostali“, rezultat se uzima od onog sa najmanjom mjerom sigurnosti. Prikaz rada algoritma nalazi se na slici 2.11.

Konstruirati klasifikatora koji svaki pojedinačno određuje klasu podatka. Svaki put kad klasifikator odredi klasu toj klasi se pridodaje jedan glas. Nakon što se svi klasifikatori izvedu i postupak glasovanja završi, vrijeme je za prebrojavanje glasova. Klasa s najvećim brojem glasova pridjeljuje se podatku. U slučaju istog broja glasova, obično se pridjeljuje klasa sa manjim indeksom. Ovakav način rada zove se jedan protiv jednog (engl. *One versus one*). Prikaz rada algoritma nalazi se na slici 2.12.



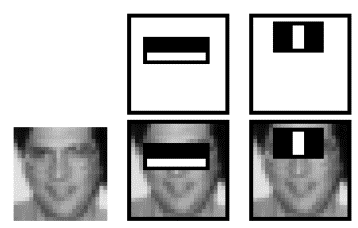
Slika 2.11. SVM način rada algoritma jedan protiv svih



Slika 2.12. SVM način rada algoritma jedan protiv jednog

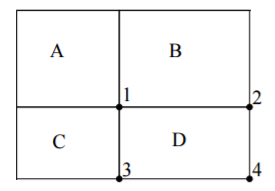
* 1. Viola-Jones algoritam

Jedan od najpoznatijih algoritama za detekciju lica na slikama je Viola-Jones algoritam. Algoritam koristi Haarove značajke koje primjenjuje na sliku i na taj način traži odgovarajući objekt, u ovom slučaju lice, usta, oči i nos [9]. Haarove značajke su slike ili matrice koje na nekim mjestima imaju tamna, a na nekim mjestima svjetla područja i svojim oblikom odgovaraju traženom objektu. Primjer se nalazi na slici 2.2.



Slika 2.13. Primjer Haarove značajke, [9]

Značajke se primjenjuju na cijelu sliku od početka do kraja i pomiču se po jedan piksel nakon primjene. Kada se prođe cijela slika, područje koje značajka zauzima se povećava i sve se opet primjenjuje ponovno. Ovo nije učinkovito raditi jer je jako sporo. Kako bi se ubrzao proces traženja objekata, Viola-Jones koriste integralnu sliku i također se definira minimalna i maksimalna granica od koje i do koje se povećava područje značajka. Računanje zbroja piksela unutar bilo kojeg kvadrata koristeći samo referentna polja se još naziva i integralna slika, što je prikazano na slici 2.3. [9]. Vrijednost integralne slike na lokaciji 1 je zbroj piksela u A kvadratu. Vrijednost na lokaciji 2 je A + B, na lokaciji 3 A + C i na lokaciji 4 A + B +C +D. Zbroj piksela unutar polja D može se izračunati na način da se zbroje vrijednosti na glavnoj dijagonali i oduzmu od zbroja vrijednosti na sporednoj dijagonali (1 + 4 - ( 2 + 3) ).



Slika 2.14. Primjer algoritma integralne slike, [9]

Kako bi se algoritam još ubrzao izbacuju se nepotrebne značajke pomoću algoritma Adaboost. To je algoritam koji izdvaja optimalne značajke iz jako velikog skupa. Nakon izdvajanja značajki, dodjeljuje im se težinska vrijednost [10]. Smatra se da je značajka prihvatljiva, ako može detektirati traženi objekt u barem pola slučajeva.

* 1. Analiza performansi klasifikatora

Opisati (ukrasti iz završnog). Podpoglavlja: metoda izdvajanja, unakrsna provjera, k-struka unakrsna provjera.

* 1. Prikaz rezultata

Opisati confusion matrix, sta je to f-mjera i sve te fore oko tog i dati formule za svaku mjeru pa čak i preciznost.

* 1. Komercijalna rješenja za računalno raspoznavanje emocija

Prepoznavanje emocija nije ništa novo te postoji dosta gotovih rješenja za privatnu, a i komercijalnu upotrebu. Neka od poznatijih rješenja su:

1. Affectiva je jako dobar alat za prepoznavanje emocija u stvarnom vremenu. Besplatan je za osobno korištenje i za tvrtke koje godišnje zarade manje od milijun dolara. Postiže veliku preciznost jer ima ogromnu bazu podataka za treniranje klasifikatora (četiri milijuna lica i konstantno raste). Može se koristiti u razne svrhe, ali glavna ideja je dobivanje povratne informacije od korisnika o nekom proizvodu [13].
2. EmoVu uz prepoznavanje emocija također nudi i prepoznavanje spola, godina, lica, praćenje lica i metrike raspoloženja. Ulazni podatci šalju se pomoću kamere, slike i videa. Pruža se podrška prebacivanja izračuna na grafički procesor kako bi se proces ubrzao [14].
3. Kairos također pruža uslugu prepoznavanja emocija, spola i dobi. Njihovo rješenje može se besplatno isprobati na njihovoj stranici [15].
4. Microsoftov projekt Oksford je popularan alat za prepoznavanje emocija koji radi samo na slikama lica. Prepoznaje sedam emocija i neutralno stanje. Radi na način da se pošalje slika, a kao rezultat dobije se JSON (*JavaScript Object Notation* ) format sa lokacijom lica na slici i postotkom za svaku od pruženih emocija [16].
5. RJEŠENJE ZA RASPOZNAVANJE EMOCIJA NA OSNOVI LICA

Cilj rješenja je pokušati što točnije prepoznati emocije na osnovu slika lica. Prepoznavanje emocija korisnika je korisna informacija u smislu ocjenjivanja neke aplikacije i služi kao dobar povratni odgovor proizvođaču. Postoji dvadesetak rješenja velikih firmi koji omogućuju korištenje njihovih rješenja u poslovne i privatne svrhe. Većina tih rješenja zahtjeva dosta računalne snage jer se koristi veliki broj klasifikatora prilikom analize lica u potrazi za akcijskim jedinicama i za svaku akcijsku jedinicu koju se želi prepoznati potrebno je imati poseban klasifikator. Većinom se sva klasifikacija vrši na njihovim računalima, a rezultat se vraća korisniku. U ovom radu pokušati će se pojednostavniti taj proces uz što manje gubitaka i što brže te će se koristiti samo jedan klasifikator.

* 1. Specifikacije i zahtjevi

Aplikacija mora moći učitati sliku i/ili dohvatiti sliku s kamere. Na učitanoj slici program mora pokušati pronaći lice. U slučaju pronalaska lica na slici, poziva se rutina za obradu slike i prepoznavanje emocije s pripadajuće slike lica. Razvojno okruženje koje se koristi je *Visual Studio*, a programski jezik C#. Aplikacija se dijeli u nekoliko projekata:

* Projekt.Zajedničko– Projekt koji je referenciran u svim ostalim projektima. Koristi se za pomoćne transportne klase i sliče stvari.
* Projekt.Dohvaćanje značajki– Konzolna aplikacija. Sastoji se od tri klase. Klasa program i dvije klase za zapis značajki u datoteku u dva različita formata.

1. Program- Klasa koja služi za učitavanje podataka iz baze (baza sa slikama koja se nalazi lokalno na računalu). Poziva metodu za prepoznavanje lica na slici. Poziva metodu iz servisa koja obrađuje dohvaćeno lice i vraća vektor značajki. Nakon dohvaćenog vektora značajki, poziva metode za spremanje značajki na disk.
2. Zapis značajki– Klasa koja prima listu značajki i zapisuje ih u jednu .txt datoteku.
3. Zapis značajki arff– Klasa koja prima listu značajki i zapisuje ih u .arff formatu koji je kasnije u radu opisan.

* Projekt.Servis– Biblioteka klasa (engl. *Class library*). Glavni dio ovog projekta je klasa za obradu učitane slike, klasa za prepoznavanje lica na slici i klasa koja služi za generiranje Gaborovih filtera koji se primjenjuju na sliku.

1. Obrada slike– Klasa koja prima sliku, a kao rezultat vraća vektor značajki. Primljenoj se slici širina i visina postavlja na fiksnu dimenziju (68x80) i pretvara se u sliku sivih tonova. Slika se potom izrezuje u pet regija: gore lijevo, gore desno, dolje lijevo, dolje sredina i dolje desno. Te se regije koriste prilikom generiranja značajki. Svaka se regija šalje na obradu pomoću Gaborovih filtera koji vraća vektor značajki za tu regiju. Nakon obrade svih regija, generira se jedan vektor značajki koji reprezentira ulaznu sliku.
2. Klasifikator– Ovo je klasa koja je zadužena za pronalaženje lica na slici. Koristi se vanjski *framework* EmguCV i Haarov klasifikator.
3. Klasa filter– Prima ulaznu sliku. Generira šesnaest Gaborovih filtera različitih orijentacija i valnih dužina. Ti filteri se potom primjenjuju na sliku koja onda predstavlja korisne značajke. Kako bi se smanjio broj značajki, koristi se PCA. Poziva se metoda koja prima sliku i pretvara ju u 2D polje to jest matricu. Na matricu se primjenjuje PCA i onda se umjesto cijele slike, koristi samo jedan dio. Na kraju metoda vraća vektor značajki.
4. Pomoćne klase– Služe za pretvaranje slike iz *Bitmap* formata u 2D polje i obrnuto, pretvaranje slike u sliku sivih tonova, rezanje slike, promjenu veličine slike…

* Projekt.Weka– Biblioteka klasa. Služi za treniranje algoritma i klasifikaciju. Sastoji se od dvije klase: klasifikator i klasificiraj. Koristi posebnu Wekinu biblioteku kako bi se mogli koristiti klasifikatori.

1. Klasifikator- Klasa koja služi za treniranje i ocjenu klasifikatora. Korišteni klasifikator je SVM.
2. Klasificiraj- Klasa koja omogućuje klasifikaciju novih podataka. Kreira novi vektor, koji je prigodan za korištenje od strane klasifikatora, na temelju dobivenog vektora značajki. Kao rezultat vraća predviđenu klasu.

* Projekt.Form- *WindowsForms* aplikacija. Omogućuje korisniku da koristi kameru kako bi predvidio emociju i omogućuje učitavanje slike sa diska. Nakon dohvaćanja slike poziva metodu za obradu slike i metodu za klasifikaciju te na temelju dobivenih značajki ispisuje korisniku rezultat u obliku predviđene emocije.

Dijagram toka aplikacije prikazan je na slici 3.1.

|  |
| --- |
|  |
| Sl. 3.1. Dijagram toka |

* 1. Korišteni alati i tehnologije

Prilikom izrade aplikacije korišteno je nekoliko vanjskih *biblioteka* kao pomoć pri izradi. To su: OpenCV, EmguCV, Accord.NET, Weka i AForge.NET. Svaki od korištenih *frameworka* je slobodan za korištenje.

OpenCV

OpenCV (engl. *Open Source Computer Vision Library*) je *framework* otvorenog koda koja se koristi za računalni vid i strojno učenje [17]. Može se besplatno koristiti u nekomercijalne, ali i komercijalne svrhe. Sadrži više od 2500 optimiziranih algoritama koji se mogu koristiti u svrhu detekcije i prepoznavanja lica, praćenje pokreta, 3D obrada slike i videa, spajanje slika u svrhu proizvodnje slike visoke rezolucije i još mnogo toga [18]. OpenCV koriste svi od osobnih korisnika, manjih tvrtki pa sve do velikih kompanija poput Google-a, Yahoo-a, Intela-a, IBM-a, Microsoft-a… Pisan je u programskom jeziku C++ i postoji mnogo *frameworka* pomoću kojih se može koristiti i u drugim programskim jezicima.

EmguCV

EmguCV je *framework* koji služi kako bi se metode iz OpenCV-a mogle koristiti u .NET kompatibilnim programskim jezicima kao što su C#, Python, VB… Pisan je isključivo u C#. Slobodan je za korištenje ako je rađena aplikacija otvorenog koda [18].

Accord.NET

Accord.NET je *framework* otvorenog koda koji se koristi za strojno učenje, obradu slike i zvuka, obradu signala, statističke primjene, prepoznavanje uzoraka… [19]… Pisan je u programskom jeziku C#. Podijeljen je u biblioteke:

* Znanstveno računanje- matematičke funkcije, statistika, strojno učenje neuronske mreže…
* Obrada slike i signala- razni filteri za slike, spajanje više slika u jednu, kreiranje integralnih slika, filteri za zvuk, detekcija lica u stvarnom vremenu…
* Biblioteke potpore- histogrami, grafovi, komponente za prikaz specifičnih slika i zvuko

Kao i prethodne biblioteke, i ova je slobodna za korištenje.

AForge.NET

AForge.NET je *framework* otvorenog koda razvijen u C#-u. Dizajniran je za rad na područjima računalnog vida, umjetne inteligencije – obrada slike, neuronskih mreža, genetskih algoritama, strojnog učenja, robotike i slično [20]. Konstantno se održava i omogućuju mu se nove funkcionalnosti. Postoji mnogo primjera aplikacija koje su rađene u ovom *frameworku*.

Weka

Weka je alat za rudarenje podatcima koji omogućuje korištenje i analizu performansi postupkom strojnog učenja [21]. Weka je razvijena u programskom jeziku Java. Sastoji se nekoliko načina rada: istraživačkog, pokusnog. toka podataka i naredbenog retka. Prikaz početnog korisničkog sučelja prikazan je na slici 3.2.

|  |
| --- |
|  |
| Sl. 3.2. Wekino početno sučelje |

Način rada u weki:

1. Učitavanje podataka u .arff formatu.
2. Pred obrada podataka.
3. Odabir željene akcija.
4. Klasifikacija- biranje klasifikatora i podešavanje opcija za treniranje.
5. Prepoznavanje nakupina podataka- odabir algoritma i specificiranje načina rada.
6. Učenje asocijacijskih pravila- odabir algoritma i podešavanje te prikaz rezultata.
7. Odabir atributa- algoritmi za izdvajanje relevantnih od irelevantnih atributa u učitanom skupu podataka.
8. Vizualizacija- Vizualni prikaz ovisnosti klasa o atributima
9. Prikaz rezultata. Rezultati se prikazuju pomoću matrice pogrešaka. Matrica pogrešaka je matrica koja na glavnoj dijagonali ima brojeve koji kazuju koliko je točno klasificiranih podataka za svaku od klasa, a ostali koliko je kojih klasa zamijenjeno za tu. Primjer matrice pogrešaka prikazan je na slici 3.3.

Weka omogućuje i uključivanje njezinih funkcionalnosti u neki od programskih jezika. Izvorno je to Java, ali se može prevesti da radi i u drugim jezicima poput C#-a.

|  |
| --- |
|  |
| Sl. 3.3. Matrica pogrešaka |

ARFF format

ARFF je oblik zapisa podataka koji se koristi u Weki. Sastoji se od dva djela. Zaglavlje u kojem se definiraju imena i vrsta atributa te dio gdje su podatci. Zaglavlje počinje s ključnom riječi @RELATION nakon čega se specificira ime relacije i onda slijedi nabrajanje atributa kao što je prikazano na slici 3.4.

|  |
| --- |
|  |
| Sl. 3.4. ARFF zaglavlje |

Podatci se odvajaju od zaglavlja ključnom riječi @DATA. Atributi su odvojeni zarezom, a podatci su svaki u svom redu, slika 3.5.

|  |
| --- |
|  |
| Sl. 3.5. ARFF podatci |

Komentari se pišu tako da se na početak reda u kojemu je komentar stavi znak „%“.

* 1. Prikaz ključnih elemenata rješenja

Aplikacija se sastoji iz dva dijela. Prvi dio je za stvaranje podataka koji se koriste pri treniranju klasifikatora. Drugi dio služi za prepoznavanje emocije preko istreniranog klasifikatora na temelju dobivenih podataka. Program za stvaranje podataka radi na način da iz direktorija učitava jednu po jednu sliku. Na učitanoj slici traži se lice te ako iz nekog razloga ne uspije pronaći, program preskače tu sliku i nastavlja dalje. Dio koda koji je zaslužan za traženje lica na slici je prikazan na slici 3.6.

|  |
| --- |
|  |
| Sl. 3.6. Pronalaženje lica na slici   |  | | --- | |  |   Slika **Error! No text of specified style in document.**.2 Izlistanje koda za "hello world" |

Pomoću imena slike određuje se pripadna emocija što se zapisuje uz pripadni vektor značajki i služi prilikom testiranja. Poziva se metoda iz projekta servisa koja prima sliku lica, a vraća pripadajući vektor značajki. Zadnje što je potrebno napraviti je zapisati značajke u neku datoteku. Ako je prvi prolazak kroz petlju, onda se prvo dodaje pripadajuće zaglavlje, ovisno o formatu zapisa. Značajke se zapisuju u dvije datoteke. Jedna je formata .tsv (engl. *Tab seperated value*), a druga je .arff. Drugi dio aplikacije, koji služi za stvarnu klasifikaciju i korištenje dobivenih podataka radi na način da se prilikom pokretanja trenira klasifikator koji će se potom koristiti. Ovaj proces traje od prilike sedamdesetak sekundi i pokreće se samo jednom. Nakon treniranja, moguće je koristiti kameru za dohvaćanje slika ili učitati sliku sa nekog drugog mjesta. Program koristi usluge biblioteke servisa za dohvaćanje značajke učitane slike koje kasnije prosljeđuje klasifikatoru. Program servis je dio program koji obavlja većinu posla. Taj dio posla je odvojen u poseban program tako da se mogu koristiti njegove usluge sa različitih mjesta, a prilikom promjene implementacije potrebno je na samo jednom mjestu napraviti izmjenu i sve će raditi. Program dijeli sliku na pet regija lica i na njima primjenjuje Gaborove filtere, slika 3.7. Na svaku novu sliku primjenjuje se PCA radi smanjivanje dimenzionalnosti podataka i vraća se vektor značajki u osnovni program.

|  |
| --- |
|  |
| Sl. 3.7. Primjenjivanje Gaborovih filtera |

Zadnji dio aplikacije služi za klasifikaciju. Prostoje dvije klase od kojih jedna služi za treniranje klasifikatora, a druga je napravljena kao jedinstveni objekt (engl. *Singleton*) i služi za klasifikaciju. Jedinstveni objekt je način pisanja koda tako da se iz jedne klase može instancirati samo jedan objekt i on se onda koristi kroz cijeli program. Takav je način pisanja koda dobar za probleme tipa generatora različitih brojeva jer će se uvijek koristi isti objekt za generiranje koji će davati svaki put druge brojeve, a u slučaju da postoji više objekata tog tipa moglo bi doći do slučaja kada bi oba dali isti broj. Način funkcioniranja nalazi se na slici 3.8.

|  |
| --- |
|  |
| Sl. 3.8. Način rada jedinstvenog objekta |

Dio programa zaslužan za treniranje i testiranje klasifikatora nalazi se na slici 3.9. Koristi se SVM klasifikator iz Weke. Treniranje se odvija na način da se podatci podjele na k dijelova. U svakom prolasku se samo jedan dio koristi za testiranje, a svi ostali za trening. Takav se algoritam naziva K-unakrsna provjera (engl. *K-fld cross validation*) čiji je rad prikazan na slici 3.10.

|  |
| --- |
|  |
| Sl. 3.9. Dio koda za treniranje klasifikatora |

|  |
| --- |
|  |
| Sl. 3.10. K-unakrsna provjera |

* 1. Prikaz izgleda, rad i uporaba rješenja

Da bi se aplikacija mogla koristiti prvo je potrebno stvoriti set vektora značajki iz odgovarajuće baze podataka. Kako bi program koji služi stvaranje baze vektora značajki radio, potrebno je bazu podataka imena „KDEF – DATABASE“ imati na radnoj površini. Svi podatci nastali spremaju se u „Data“ direktorij i biti će potrebni za rad programa. Izgled programa koji stvara bazu vektor značajki prikazan je na slici 3.11.

|  |
| --- |
|  |
| Sl. 3.11. Izgled programa za stvaranje baze vektora značajki |

Izgleda tek pokrenutog programa nalazi se na slici 3.12.

|  |
| --- |
|  |
| Sl. 3.12. Izgled tek pokrenutog programa |

Sa slike 3.12. vide se mogućnosti programa. Moguće je učitati sliku s diska i na njoj izvršiti klasifikaciju pomoću gumba „Učitaj sliku“. Prilikom prve klasifikacije potrebno je pričekati sedamdesetak sekundi da se klasifikator istrenira. Taj proces obavlja se samo jednom. Također je moguće odabrati koji se klasifikator želi koristiti prilikom prepoznavanja lica i vrijeme okidanja za automatsko prepoznavanje. Zadnja opcija testiranja je na drugoj bazi podataka (predefinirana baza koju je potrebno imati na radnoj površini). Na slici 3.13. prikazan je izgled programa prilikom prve klasifikacije gdje se prikaže dodatan prozor koji izvještava korisnika da treba pričekati dok se klasifikator ne istrenira.

|  |
| --- |
|  |
| Sl. 3.13. Izgled programa prilikom prve klasifikacije |

Nakon što se treniranje klasifikatora završi, pojavljuje se rezultat u obliku teksta koji govori koja je emocija prepoznata na slici, slika 3.14.

|  |
| --- |
|  |
| Sl. 3.14. Rezultat klasifikacije |

Gumb „Info“ otvara novi prozor koji daje informacije o klasifikatoru. Prikazana je sveukupna preciznost klasifikacije, matrica pogrešaka i preciznost klasifikacije po pojedinoj emociji, slika 3.15.

|  |
| --- |
|  |
| Sl. 3.15. Informacije o klasifikatoru |

Jedna od opcija je klasifikacija slike sa kamere uređaja. Pomoću padajućeg izbornika odabire se kamera koja se želi koristiti i pritiskom na gumb „Pokreni“ kamera se pokreće i prikazuje sliku u sredini programa. Klasifikacija se može vršiti „ručno“ ili „automatski“ Ručna klasifikacija odnosi se na to da korisnik pritisne gumb „Prepoznaj“ i u tome trenutku se uzima trenutna slika i šalje na klasifikaciju. Automatska klasifikacija odnosi se na to da će program svaki interval vremena uzimati trenutnu sliku i slati je na klasifikaciju te vraćati rezultat. Primjer korištenja kamere prilikom dohvaćanja slike za klasifikaciju prikazan je na slici 3.16.

|  |
| --- |
|  |
| Sl. 3.16. Klasifikacije pomoću slike sa kamere |

Slika 3.17. prikazuje rezultate testiranja klasifikatora na drugoj bazi podataka. Prikazana je matrica pogrešaka i ukupna preciznost klasifikacije.

|  |
| --- |
|  |
| Sl. 3.17. Testiranje klasifikacije na drugoj bazi podataka |

* 1. Analiza ponašanja i performansi

U ovom poglavlju prikazani su rezultati uspješnosti klasifikacije u poznatim okolnostima i sa nepoznatim podatcima. Jedan od parametara koji se uzima u obzir je vrijeme potrebno za treniranje i testiranje klasifikatora. Uz samo jedan klasifikator i 1856 značajki, ne uključujući oznaku klase, vrijeme potrebno za treniranje i testiranje klasifikatora iznosi oko 70 sekundi. Program se izvodio na računalu čije su osnovne specifikacije dane u tablici 3.1. Testiranje klasifikatora vršilo se k-unakrsnom provjerom gdje je broj ponavljanja bio postavljen na deset.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Procesor | Intel(R) Core(TM) i5-5200U CPU @ 2,20GHz, 2195 MHz | | Broj stvarnih jezgri | 2 | | Broj logičkih jezgri | 4 | | Radna memorija | 6 GB DDR3 L | | Tvrdi disk | 1 Disk - WDC WD10JPVX-22JC3T0 (931 GB) | | Matična ploča | Acer Aspire E5-573G | |
| Tablica 3.1. Osnovne specifikacije računala na kojoj je provedena analiza performansi |

Prednosti ovog načina pristupa problemu su u tome da se koristi jedan klasifikator. Kada bi se koristio po jedan klasifikator za svaku akcijsku jedinicu, vrijeme potrebno za testiranje i treniranje svih tih klasifikatora bi bilo značajno veće. Uz manji broj klasifikatora i samo programsko rješenje je puno jednostavnije i stoga lakše za održavati i nadograđivati, ali dolazi s određenim žrtvovanjem preciznosti klasifikacije.

Analiza rezultata na trening podatcima

Što se tiče prosječne preciznosti klasifikacije na trening podatcima, ona iznosi 82,06%. Prilikom treninga korišteno je 797 različitih vektora značajki koji predstavljaju različite izraze lica. Neke emocije su jako dobro prepoznate, a neke su malo slabije što je mana korištenja jednog klasifikatora jer su za neke emocije slični izrazi lica i dosta ih je teže prepoznati nego druge. Uspješnost klasifikacije po pojedinoj emociji dana je u tablici 3.2.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Emocija | Uspješnost klasifikacije [%] | | Strah | 71,68 | | Srdžba | 75,86 | | Gađenje | 88,79 | | Radost | 96,49 | | Neutralno | 75,38 | | Tuga | 77,36 | | Iznenađenje | 90,09 | |
| Tablica 3.2. Uspješnost klasifikacije po pojedinoj emociji |

Iz tablice 3.2. se vidi da su emocije: strah, srdžba, tuga i neutralno ispod prosjeka po uspješnosti, a ostale odstupaju od njih za minimalno deset posto. To je možda iz razloga jer izrazi lica za te emocije odstupaju od drugih, a i ljudi ih lakše prepoznaju. . Cijela matrica pogrešaka s detaljima klasifikacije prikazana je u tablici 3.3.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Strah | Srdžba | Gađenje | Radost | Neutralno | Tuga | Iznenađenje | | Strah | 81 | 5 | 2 | 2 | 7 | 8 | 9 | | Srdžba | 7 | 88 | 8 | 1 | 6 | 4 | 0 | | Gađenje | 5 | 9 | 95 | 1 | 1 | 3 | 0 | | Radost | 2 | 0 | 0 | 110 | 3 | 0 | 0 | | Neutralno | 3 | 4 | 0 | 0 | 98 | 9 | 1 | | Tuga | 9 | 10 | 2 | 0 | 10 | 82 | 1 | | Iznenađenje | 6 | 0 | 0 | 0 | 5 | 0 | 100 | |
| Tablica 3.3. Matrica pogrešaka |

Iz matrice pogrešaka vidi se da je srdžba najviše miješana sa tugom i strahom, a neutralno s tugom. Tuga je najviše zamijenjena sa neutralnim, a potom sa strahom, dok je iznenađenje daleko najviše mijenjano sa strahom. Najbolje klasificirana emocija je radost koja skoro pa i nije pogrešno klasificirana.

Analiza rezultata na nepoznatim podatcima

Kao testni podatci korištene su slike iz Cohn-Kanade baze podataka [22], [23]. Primjer slika nalazi se na slici 3.18. Baza sadrži frontalne slike lica sa emocijama. Za svaku emociju postoji više slika, to jest od neke neutralne emocije pa postupno kako dolazi do zadane emocije.

|  |
| --- |
|  |
| Sl. 3.18. primjer slika iz Cohn-Kanade baze podataka (©Jeffrey Cohn) |

Prilikom testiranja korišteno je 299 slika lica. Prosječna točnost klasifikacije iznosi 66,22%, a točnost klasifikacije po emociji dana je u tablici 3.4.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Emocija | Uspješnost klasifikacije [%] | | Strah | 26,47 | | Srdžba | 48,78 | | Gađenje | 90,00 | | Radost | 91,67 | | Neutralno | 0,00 | | Tuga | 55,17 | | Iznenađenje | 100,00 | |
| Tablica 3.4. Uspješnost klasifikacije po pojedinoj emociji |

Najlošije prepoznata emocija je strah, a potom slijede srdžba i tuga što se podudara s rezultatima testiranja koji su provedeni na trening podatcima. Ostale tri emocije su jako dobro prepoznate, a iznenađenje je prepoznato sa čak 100%. Neutralnih izraza nije bilo u bazi (u završnom okviru skupa slika od kojih svaki okvir predstavlja postupno dolaženje do određene emocije). Ako bi se izuzeo strah, rezultati su dosta dobri. Matrica pogrešaka prikazana je u tablici 3.5 i prikazuje detaljniji prikaz klasifikacije u smislu koja emocija je koliko puta klasificirana ispravno, a koliko puta kao neka druga (i koja druga).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Strah | Srdžba | Gađenje | Radost | Neutralno | Tuga | Iznenađenje | | Strah | 18 | 1 | 0 | 3 | 1 | 2 | 0 | | Srdžba | 12 | 20 | 2 | 2 | 2 | 7 | 0 | | Gađenje | 19 | 15 | 18 | 1 | 2 | 3 | 0 | | Radost | 1 | 0 | 0 | 66 | 0 | 1 | 0 | | Neutralno | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | Tuga | 5 | 4 | 0 | 0 | 2 | 16 | 0 | | Iznenađenje | 13 | 1 | 0 | 0 | 2 | 0 | 60 | |
| Tablica 3.5. Matrica pogrešaka |

* 1. Osvrt i mogućnost unaprjeđenja

Prilikom prepoznavanja emocija koristili su se Gaborovi filteri u kombinaciji s PCA za smanjivanje dimenzionalnosti, a kao klasifikator SVM. Koristi se jedna baza podataka za treniranje klasifikatora sadosta kvalitetnim slikama koje su napravljene u kontroliranim uvjetima te su sve emocije odglumljene. Program je rađen da prepoznaje emocije na osnovu frontalnih slika lica. Rezultati klasifikacije na trening podatcima su dosta visoki, a na drugim bazama podataka oko 20% manji. Preciznost klasifikacije srozavaju određene emocije koje program jako teško prepoznaje, a to su strah, srdžba i tuga. Strah je daleko najlošije prepoznata emocija. Ako bi se samo preciznost klasifikacije straha dovela na razinu druge najlošije prepoznate emocije (srdžba, 48,78%) tada bi se ukupna preciznost popravila za više od 7% i bila bi 73.53%. Ako bi se zanemarilo i prepoznavanje neutralne emocije, koje ni nema, tada bi se ukupna preciznost popravila za skoro 10% i bila bi 76,05%. Glavni problem i najveća mogućnost unaprjeđivanja preciznosti bi bila traženje rješenja za veću preciznost tri najlošije prepoznate emocije. To bi se moglo raditi promjenom ili unaprjeđenjem kvalitete značajki koji se dobivaju iz slika. Još jedan način unaprjeđenja cjelokupnog rješenja bi bio korištenje više različitih baza slika prilikom konstruiranja klasifikatora koji bi bili u različitim okruženjima i različite kvalitete što bi dalo realniju situaciju u kojoj bi se primjenjivala aplikacija. Ako bi se koristile kamere s korisničkih računala tu kvaliteta može varirati od jako niske pa do jako visoke te bi u tom slučaju pomogle različite baze podataka sa različitim kvalitetama. Ako bi se stvarno htjela poboljšati preciznost te se koristiti jedna ovakva aplikacija u poslovnom svijetu, bilo bi potrebno obratiti pažnju i na druge aspekte ljudske prirode preko kojih se mogu prepoznati emocije. Na primjer način govora često može odati o kojoj se emociji radi ili barem suziti izbor i pomoći pri klasifikaciji. Pokreti tijela bi se trebali pratiti, a na kraju i kontekst u kojoj se osoba nalazi uvelike pomažu prilikom prepoznavanja emocija. Kada bi se svi ti aspekti objedinili i koristili za prepoznavanje emocija, tada bi se takvo rješenje moglo koristiti za neka ozbiljnija rješenja. Na kraju i uz sve to, neka određena pogreška je neizbježna.

1. ZAKLJUČAK

Prepoznavanje emocija osoba je zamršen proces koji je često kompliciran za ljude, a kamoli za računala. Čovjek prepoznaje emocije na određenoj osobi pomoću jako puno značajki. Na to utječe izraz lica, ton, kontekst, prijašnja iskustva s određenom osobom ili kulturom i uz sve to često se zna pogriješiti. Ovaj se rad bazira samo na slikama lica što je dovoljno za neke privatne upotrebe i za prepoznavanje glumljenih emocija koji se jako manifestiraju na licu. S obzirom na to da se koriste samo slike, rezultati su zadovoljavajući. Rezultati variraju za određene emocije. Neke je lakše prepoznati na licu, a neke teže. Sva testiranja i treniranja rađena su samo s glumljenim emocijama te bi bila dosta veća pogreška kada bi se pokušavale prepoznati stvarne emocije u nekoj realnoj situaciji. Ovakvo rješenje bi se moglo koristiti jedino na mjestima gdje bi se emocije jako izražavale na licu, a to je prilikom gledanja nekih videa ili slika koji su ili izrazito šaljivi, tužni ili imaju zastrašujuće scene. Na takvim mjestima bi se moglo koristiti za mjerenje uspješnosti neke aktivnosti (na primjer: nasmijavanje ljudi, pokušaja preplašivanja).

1. LITERATURA

[1] D. G. Myers, "Theories of emotion." Psychology: Seventh Edition, NY: Worth Publishers 500, New York, 2004.

[2] C. Darwin, P. Ekman, i P. Prodger, The expression of the emotions in man and animals, Oxford University Press, USA, 1998.

[3] Y. L. Tian, T. Kanade i J. F. Cohn, Facial expression analysis. In Handbook of face recognition (pp. 247-275). Springer, New York, 2005.

[X4] G. Deleuze i F. Guattari, A Thousand Plateaus: Capitalism and Schizophrenia, Les Éditions de Minuit, Francuska, 1980.

[X5] E. Shouse, Feeling, Emotion, Affect, M/C Journal, br. 6, sv. 8, str. -, Prosinac 2005.

[X6] R. E. Plutchik, The circumplex as a general model of the structure of emotions and personality, American Psychological Association, br. -, sv. -, str. 17-45., Siječanj 1997.

[X7] T. Christy, L. I. Kuncheva i K. W. Williams, Selection of Physiological Input Modalities for Emotion Recognition, Technical Report# CS-TR-002-2012, Bangor University, UK

[X8] Z. Zhiong, M. Pantic, G. I. Roisman i T. S. Huang, A survey of affect recognition methods: Audio, visual, and spontaneous expressions, IEEE transactions on pattern analysis and machine intelligence, br. 1, sv. 31, str. 39. – 58., 2009

[X9] A. J. Fridlund, Human facial expression (1 ed.), Academic Press, San Diego, 1994.

[X10] D. Popescu, R. Dobrescu i M. Nicolae, Texture Classification and Defact Detection by Statistical Features, International Journal of Circuits, System and Signal Processing, Br. 1, Sv. 1, str. 79 – 84, 2007.

[X11] T. Ahonen, A. Hadid i M. Pietikäinen, Face Description with Local Binary Patterns: Application to Face Recognition, Pattern Analysis and Machine Intelligence, IEEE Transaction on, Br. 12, Sv. 28 , str. 2037 – 2041, 2006.

[X12] M. N. Murty i V. S. Devi, Pattern Recognition An Algorithmic Approach, Springer, SAD, 2011.

[4] Johnson-Laird, P. Nicholas, and K. Oatley, The language of emotions: An analysis of a semantic field, Cognition and emotion, Vol. 3, No. 2, pp. 81-123, 1989.

[5] <http://www.humintell.com/2010/06/the-seven-basic-emotions-do-you-know-them/>, pristupljeno 26.11.2016.

[6] D. Lundqvist, A. Flykt, & A. Öhman, The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9, 1998

[7]S. Theodoridis i K. Koutroumbas, Pattern Recognition (Fourth Edition), Elsevier Inc., SAD, 2009.

[8]M. N. Murty i V. S. Devi, Pattern Recognition An Algorithmic Approach, Springer, SAD, 2011.

[9] P. Viola, M. Jones, Rapid object detection using boosted cascade of simple features, In Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference, Vol. 1, pp. I-511-I-518, 2001.

[10] Y. Freund & R. E. Schapire, A desicion-theoretic generalization of on-line learning and an application to boosting. In European conference on computational learning theory, pp. 23-37 Springer Berlin Heidelberg, 1995.

[11] S. Padmavathi, P. Prem i D. Praveenn, Locating Fabric Defects Using Gabor Filters, International Jurnal of Scientific Research Engineering & Technology, ISSN 2278 – 0882, Vol. 2 Issue 8, pp. 472-478, Studeni 2013.

[12] S. Wold, K. Esbensen, and P. Geladi, Principal component analysis, Chemometrics and intelligent laboratory systems, Elsevier Science Publishers B.V, pp. 37-52, Nizozemska, 1987.

[13] <http://www.affectiva.com/>, pristupljeno 26.11.2016.

[14] <http://emovu.com/>, pristupljeno 26.11.2016.

[15] <https://www.kairos.com/>, pristupljeno 26.11.2016.

[16] <https://www.microsoft.com/cognitive-services/>, pristupljeno 26.11.2016.

[17] <http://opencv.org/>, pristupljeno 11.12.2016.

[18] <http://www.emgu.com/>, pristupljeno 11.12.2016.

[19] <http://accord-framework.net/>, pristupljeno 11.12.2016.

[20] <http://www.aforgenet.com/framework/>, pristupljeno 11.12.2016.

[21] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann i I. H. Witten, The WEKA Data Mining Software: An Update, SIGKDD Explorations, Vol. 11, pp. 10-18, 2009

[22] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

[23] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

1. SAŽETAK

U radu je implementirana metoda za prepoznavanje emocija na slikama lica. Korišteni su Gaborovi filteri kako bi se dobile crte na licu i potom PCA kako bi se smanjila dimenzionalnost. Kao klasifikator koristi se SVM. U teorijskom dijelu objašnjeni su pojmovi kao što su emocija i afekt. Opisan je klasifikator i korištene metode. Praktični dio se sastoji od programa za dohvaćanje značajki iz baze slike i od dijela koji koristi te značajke kako bi se istrenirao klasifikator te kasnije koristio prilikom prepoznavanja emocija, što s korisnikove kamere, što sa učitane slike. Ovakvo rješenje je dovoljno dobro za korištenje u osobne svrhe.

**Ključne riječi:** Emocije, Gaborovi filteri, PCA, SVM, izrazi lica

**ABSTRACT**

A method for recognizing emotions in images of faces is implemented in the paper. Gabor filters are used in order to extract facial features and then PCA afterwards is applied in order to reduce dimensionality. SVM is used as a classifier. In the theoretical part, terms such as emotion and affect are explained. The classifier is described, as well as the applied methods. The practical part consists of an application which extracts features from a base of images and from a part which is using those features to train the classifier in order to use it later on when recognizing emotions in an image from a personal camera or from an uploaded one. This sort of solution is good enough for personal use.

**Key words:** Emotions, Gabor filters, PCA, SVM, facial expression

1. ŽIVOTOPIS

Josip Baketarić, rođen u Požegi 10. kolovoza 1993. Osnovnu školu završio u Pleternici u razdoblju od 2000. do 2008. godine. 2008. godine upisuje srednju tehničku školu u Požegi, smjer tehničar za računalstvo koju završava 2012. godine. Nakon toga upisuje Elektrotehnički fakultet u Osijeku, preddiplomski studij, smjer računarstvo kojega završava 2015. godine. 2016. godine upisuje diplomski studij na prijašnje spomenutom fakultetu, sada Fakultet elektrotehnike, računarstva i informacijskih tehnologija, smjer računarstvo, izborni blok programsko inženjerstvo kojega trenutno pohađa.

1. PRILOZI

Na CD-u:

1. Diplomski rad „Klasifikacija emocija na osnovi slike lica.docx“
2. Diplomski rad „Klasifikacija emocija na osnovi slike lica.pdf“
3. Izvorni kod programa
4. Instalacija programa