

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 3801

**AUTOMATSKO RASPOZNAVANJE
EMOCIJA IZ SLIKE LICA**

Antonela Mrkalj

Zagreb, lipanj 2014.

Zagreb, 14. ožujka 2014.

ZAVRŠNI ZADATAK br. 3801

Pristupnik: **Antonela Mrkalj (0036468310)**
Studij: Računarstvo
Modul: Računarska znanost

Zadatak: **Automatsko raspoznavanje emocija iz slike lica**

Opis zadatka:

Automatsko raspoznavanje emocija iz slike lica lica zasniva se na tehnikama računalnog vida a može poslužiti u raznim interaktivnim aplikacijama u kojima je poželjno procijeniti emotivno stanje korisnika. Na Zavodu za telekomunikacije je dostupan softver za analizu digitaliziranih slika lica zasnovan na strojnom učenju. Do sada je ovaj softver eksperimentalno primjenjen na procjenu spola i dobi, a u okviru ovog zadatka cilj je primijeniti ga na raspoznavanje emocija.

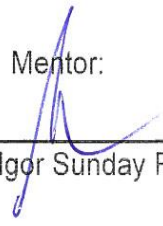
Vaša je zadaća proučiti sustav za analizu lica dostupan na Zavodu za telekomunikacije, pronaći odgovarajuće baze slika lica s anotiranim emocijama, trenirati sustav za raspoznavanje emocija te eksperimentalno utvrditi njegovu točnost.

Svu potrebnu literaturu i uvjete za rad osigurat će Vam Zavod za telekomunikacije.

Zadatak uručen pristupniku: 14. ožujka 2014.

Rok za predaju rada: 13. lipnja 2014.

Mentor:




Prof.dr.sc. Igor Sunday Pandžić

Djelovođa:



Doc.dr.sc. Tomislav Hrkać

Predsjednik odbora za
završni rad modula:



Prof.dr.sc. Siniša Srblić

Mojim rodičeljima, največim učiteljima.

Sadržaj

1. Uvod.....	1
2. Ideja	2
3. Implementacija.....	3
3.1. Baza slika.....	3
3.2. Strojno učenje	4
3.3. Aplikacija za prepoznavanje emocija	4
3.3.1. Grafički aspekti.....	7
4. Rezultati i tablice	8
5. Opisi programa i klasa	11
5.1. Visage SDK.....	11
5.1.1. Face Tracking	11
5.2. OpenCV	11
6. Zaključak	12
7. Literatura.....	13
8. Sažetak	14
9. Abstract.....	15

1. Uvod

Sva živa bića imaju emocije i mogućnost izražavanja istih. Emocija ili osjećaj je naziv za reakciju na neki događaj, no ono što je zanimljivo je kako su međukulturno emocije različite jedino po nazivima.

Ljudi nesvjesno znaju kako će izraziti pojedinu emociju i ona se izražava identično u svim kulturama. Ukoliko se osoba osjeća sretnom svoj osjećaj će izraziti putem osmijeha, ukoliko je tužna putem plača ili jednostavno tužnog pogleda i sl.

Budući da su emocije svojstvene svim ljudima, moguće je izgraditi sustave koji imaju mogućnosti prepoznavati emocije svojih korisnika.

Cilj ovog rada je ostvariti automatsko prepoznavanje emocija iz slike lica pomoću strojnog učenja primijenjenoga do tada za raspoznavanje spola i dobi i utvrditi eksperimentalno njegovu točnost.

U nastavku su opisani ideja i implementacija, korišteni alati, a na kraju su priloženi rezultati i zaključak.

2. Ideja

Budući da se treba utvrditi može li se sustav za procjenu spola i dobi koristiti u slučaju emocija, potrebna je baza slika sa anotiranim emocijama, kod za strojno učenje koji će učiti na temelju slika iz baze te testiranje naučenoga radi rezultata.

Budući da se koristi linearno strojno učenje, dolazi se do zida prilikom ispitivanja emocija. Moguće je odabrati samo dvije emocije za učenje i zbog toga se odlučujem za najpraktičniju osmijeh i neutralan izraz lica. Druge emocije se mogu uzeti u obzir u istoj kombinaciji sa neutralnim izrazom, no rezultati će biti lošiji, jer baze sa anotiranim emocijama su poprilično male, a najzastupljenija emocija je osmijeh.

Praktični dio rada je implementacija aplikacije za prepoznavanje emocije u stvarnom vremenu, koja bi omogućila demonstraciju rezultata strojnog učenja.

3. Implementacija

3.1. Baza slika

Na Zavodu za telekomunikacije postoji baza slika, koja se upotrijebila u svrhu ovog završnog rada.

Problem kod svake baze je nedostatak slika s emocijama, budući da su emocije poprilično slabo zastupljene u područjima računalnog vida, za razliku od drugih karakteristika poput spola ili dobi.

Također, u obzir dolaze samo frontalne slike, prilikom čega se broj slika znatno smanjuje.

Sve slike u bazi moraju biti unaprijed anotirane, što znači da već imaju dodijeljenu oznaku emocije. U našem slučaju, postoje oznake za sreća, ljutnja, tuga, gađenje i strah. Oznake poprimaju vrijednosti od nula do sto.

Sa Zavoda je preuzeta skripta *generatesamples.py* koja je inicijalno bila korištena za stvaranje uzoraka za strojno učenje na temelju spola. Izlaz skripte čine slika u rwi formatu, koja je odrezana oko lica u odnosu na originalnu sliku, te datoteka u kojoj je zapisana vrijednost skupine (jedan ili minus jedan) i položaj i veličina tzv. *bounding box-a*. *Bounding box* je kvadrat koji se pozicionira oko lica.

Skripta je modificirana na način da uzima u obzir vrijednosti za emocije iz anotacijske datoteke. Ukoliko su sve emocije u rangu od nula do pet, slika se označava sa vrijednosti minus jedan, tj. ona spada u klasu neutralno. Ukoliko su sve emocije osim sreće u rangu od nula do pet, a sama sreća od pedeset pa na više, slici se pridjeljuje oznaka jedan, tj. pridružuje se klasi osmijeh.

Kako postoji mogućnost da slike kojima je sreća između pet i pedeset budu označene jednom od te dvije brojke i tako budu pridijeljene nekoj od klasa, dodaje se za kontrolu vrijednost nula, kako se ne bi pomiješale sa drugima.

Budući da se u okviru završnog rada koristi strojno učenje potrebno je, nakon pokretanja skripte i dobivenih .dat i .rwi datoteka, ostvariti podjelu na skup za učenje i skup za testiranje.

Za potrebu slijedećega napisana je jednostavna python skripta koja pretražuje .dat datoteke i uočava pripada li slika klasi osmijeh ili neutralno. Ukoliko ne pripada niti jednoj klasi (oznaka nula) zanemaruje se.

Nakon toga se izgrade dva direktorija sa određenim omjerom slika. U originalu se koristila podjela pola-pola originalnog skupa, no kasnije se koristi dvotrećinski udio slika za skup učenje i jednotrećinski udio za skup testiranje.

3.2. Strojno učenje

Kod za strojno učenje je preuzet od Zavoda za telekomunikacije. Originalno korišten samo na detekciji spola i dobi, prvi put je primijenjen za detekciju emocija.

Koristi se rad linearnog SVM (*support vector machine*), zbog čega se koriste samo dvije klase emocija.

U svrhu ovog završnog rada, strojno učenje se koristilo poput „crne kutije“. Izvršnoj aplikaciji se na ulaz dao skup podataka, prvotno stvoren i nazvan skup za učenje. Nakon završenog učenja se vrši provjera na skupu za testiranje.

Konačni rezultat učenja jest klasifikator. On će se koristiti kasnije u programskom ostvarenju aplikacije za prepoznavanje emocija.

3.3. Aplikacija za prepoznavanje emocija

Kako ne bi ostali samo na brojčanim vrijednostima, u sklopu ovog rada se radi i aplikacija koja će demonstrirati uživo prepoznavanje emocija.

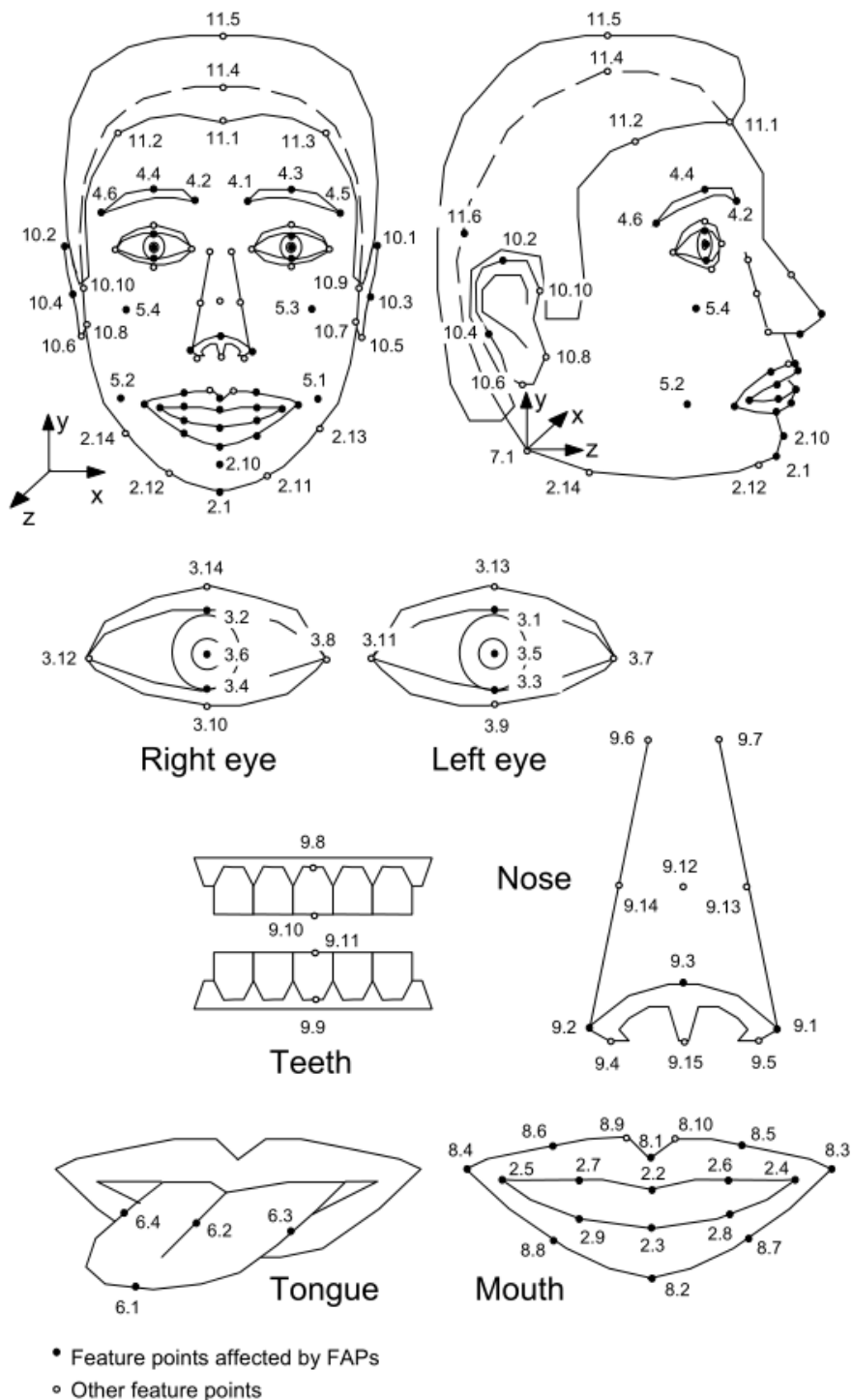
Za potrebe aplikacije koriste se Visual Studio 2010 i Visage|SDK 7.2 beta.

Visage|SDK integrira veliki broj računalnih tehnologija za animaciju likova i računalni vid u jednostavne razvojne alate koji se lako koriste.

U sklopu ovog rada koristi se jedan od tih alata, Face Tracker (praćenje lica).

Budući da se izvorni kod Face Tracker-a može u potpunosti upotrijebiti, no samo unutar Visual Studio 2010 verzije, stvara se novi projekt identičan originalnom.

Face Tracker je alat koji omogućava dohvaćanje točaka lica (eng. *features points*) u stvarnom vremenu. Na slici 1. se vide oznake svake točke lica. U ovom radu su bitne samo dvije točke, točka 3.5 (lijevo oko, zjenica) i 3.6 (desno oko, zjenica).



Slika 1. Točke lica koje dohvaća Face Tracker

Te dvije točke su poslužile za stvaranje „*bounding box-a*“. Budući da se za ulaz klasifikatora mora koristiti slika koja je nastala kao posljedica rezanja okoline oko lica, potrebno je dobiti sve parametre koji će to omogućiti.

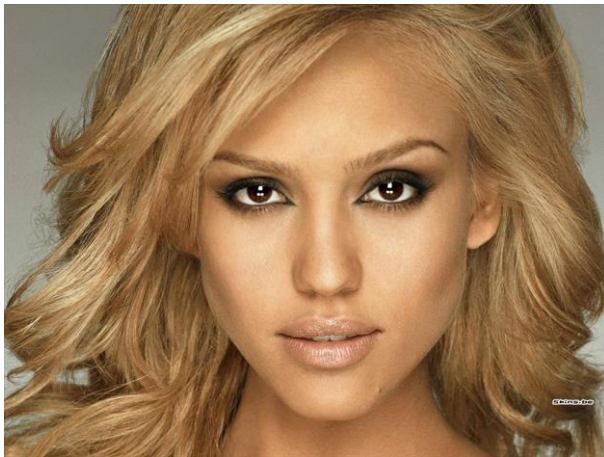
Ideja je da se izračuna točka između očiju, po x i y osi, koja će biti centar kutije. Tipično se za širinu kutije uzima udaljenost između očiju pomnožena sa 1,3 postavljena lijevo i desno od centra, a za duljinu se uzima isto samo se množi sa većim parametrom ukoliko je potrebno (budući da su lica češće dulja nego šira).

Ipak, ovdje to nije slučaj, jer je u kodu za strojno učenje implementirano da kutija mora biti iste duljine i širine.

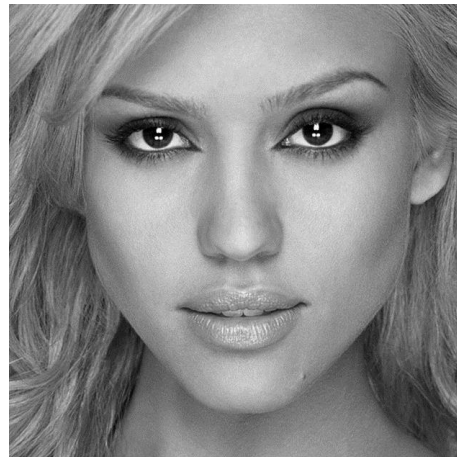
Ugrađena klasa FaceData sadrži podatke poput rezolucije slike i točaka lica. Slika s kojom se radi je tog tipa i zove se trackingData. Točnije, pokazivač na objekt koji sadrži sliku (koja je spremljena kao polje) se zove trackingData i preko njega dohvaćamo točke 3.5 i 3.6, no ono što on vraća nisu slikovne jedinice (eng. *pixels*) već pozicije u jediničnom koordinatnom sustavu. Zato koristimo rezoluciju kako bi pomnožili širinu sa koordinatom x osi i visinu sa koordinatom y osi i dobili točku koja je slikovna jedinica.

Unutar Visage|SDK postoje brojne ugrađene klase i funkcije koje omogućavaju određene funkcionalnosti, no on koristi i OpenCV.

Zahvaljujući njemu omogućeno je tzv. rezanje originalne slike i stvaranje nove. Za kontrolu omogućeno je spremanje modificirane i originalne slike u direktoriju od Face Tracker-a.



Originalna slika
Dimenzije: 1024 x 768



Slika dobivena
„bounding box-om“
Dimenzije: 632 x 632

Slika 2. Prikaz utjecaja *bounding box-a*

Nakon toga slijedi poziv funkcije validacije iz koda za strojno učenje. To znači da se dio koda mora ugraditi u aplikaciju. S obzirom da je potrebno samo procijeniti kojoj klasi slika pripada, nije potrebno treniranje već samo validiranje.

Prilikom poziva funkcije potrebno je prosljediti točne podatke. U nastavku je prikazan pseudokod pozivne funkcije.

```
value_s(klasifikator, 0, pokazivač na zapis originalne slike, visina slike,  
širina slike, widthStep slike, centralna točka (koordinata x), centralna  
točka (koordinata y), veličina stranice kutije);
```

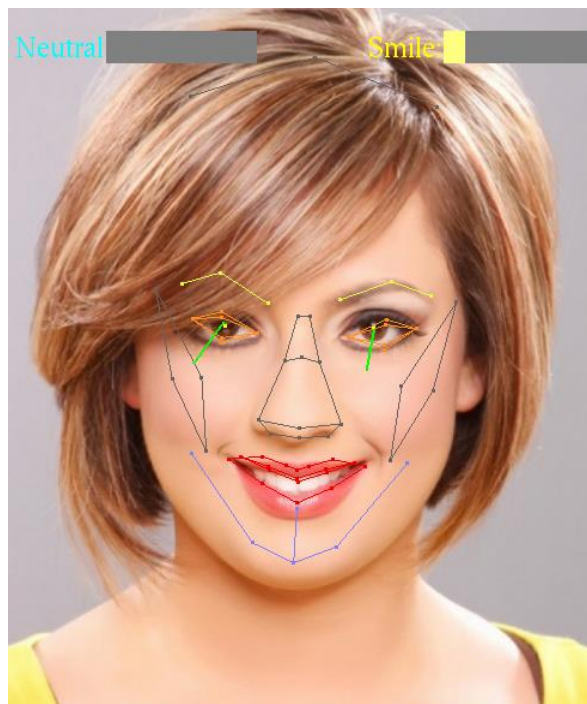
Ono što tada vraća validator je broj čiji iznos može biti od minus beskonačno do plus beskonačno. Ukoliko je negativan broj, procijenjeno je da slika pripada klasi neutralno, inače klasi osmijeh.

Prilikom dohvaćanja slike, koristi se trostruki spremnik (eng. *buffer*), pri čemu se uvijek čita iz jednog spremnika dok se u druga dva učitavaju slike. Kako bi se smanjilo opterećenje da se prilikom svake slike (eng. *frame*) vrši prepoznavanje, ograničeno je na svaku šestu sliku.

3.3.1. Grafički aspekti

Umjesto brojčanog prikaza vanjskog rezultata, koristi se grafički prikaz.

U gornjem lijevom uglu postavljen je mali spremnik za neutralno, a u desnom za osmijeh. Uvijek je samo jedan spremnik aktivan, nikad dva istovremeno. Spremnici su skalirani na veličinu od nula do dvadeset, budući da vrijednosti mogu ići u beskonačno.



Slika 3. Prikaz grafičkih barova

4. Rezultati i tablice

Rezultati nisu bili zadovoljavajući, a glavni problem je baza slika. Ne postoji dovoljno velika baza sa emocijama čime klasifikator nema visoku točnost.

Na početku se koristila nepromijenjena baza, u kojoj su se, na temelju anotiranih emocija, radile podjele. U tablici se vidi učinak takve baze.

Tablica 1. Učinak podjele baze slika na učenje i testiranje

Broj slika Neutral	Broj slika Osmijeh	Priprema uzoraka: podjela na klase	Način podjele slika	Broj slika za učenje	Broj slika za testiranje	Točnost prepoznavanja neutralnih izraza	Točnost prepoznavanja osmijeha
5240	1916	neutral: slike anotirane od 0 do 5 u svim emocijama osmijeh: slike anotirane od 30 do 100 u emociji osmijeh, druge emocije 0 do 5	pola neutralnih pola osmijeh	3578	3578	69	73
3368	838	neutral: slike anotirane od 0 do 5 u svim emocijama osmijeh: slike anotirane od 50 do 100 u emociji osmijeh, druge emocije 0 do 5	pola neutralnih pola osmijeh	2103	2103	90	82
3368	838	isto	prva 1/3 neutralnih i osmijeh za testiranje 2/3 za učenje	2805	1401	97	63
3368	838	isto	pola osmijeh 2/3 neutralne za učenje	2665	1541	89	72
3368	838	isto	prve 2/3 neutralnih i osmijeh za učenje 1/3 za testiranje	2805	1401	90	92

Iako rezultati djeluju poprilično dobri, u aplikaciji za prepoznavanje uživo nisu pokazali svoju točnost.

Ponajviše jer su u obzir uzete i slike koje su profili ili slike lica koja nisu usmjerena prema naprijed već su nagnuta u stranu.

Nakon brisanja takvih slika i ručnog ispravljanja pogrešno anotiranih slika dolazi se do rezultata navedenih u tablici 2.

Tablica 2. Rezultat učenja na pročišćenju bazi

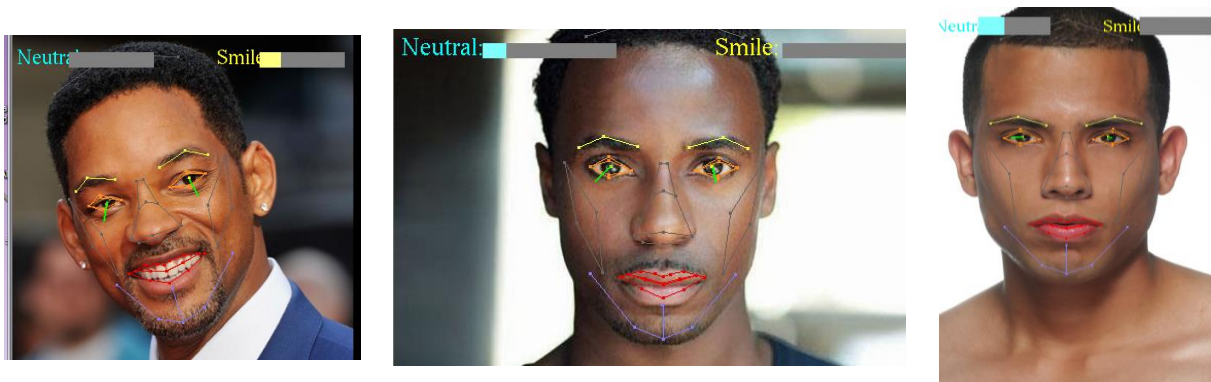
Broj slika Neutral	Broj slika Osmijeh	Priprema uzoraka: podjela na klase	Način podijele slika	Broj slika za učenje	Broj slika za testiranje	Točnost prepoznavanja neutralnih izraza	Točnost prepoznavanja osmijeha
1596	473	neutral: slike anotirane od 0 do 5 u svim emocijama osmijeh: slike anotirane od 50 do 100 u emociji osmijeh, druge emocije 0 do 5	prva 1/3 neutralnih i osmijeh za testiranje 2/3 za učenje	992 (688 neutral+ 304 osmijeh)	604 (435 neutral + 169 osmijeh)	92	62

Korištenjem klasifikatora dobivenog na temelju zadnje tablice postoje greške prilikom klasificiranja slika što je vidljivo na slici 4.



Slika 4. Pogrešna detektiranja emocija

Također, na slikama sličnih oblika lica i izraza postoje i točne klasifikacije što se vidi na slici 5.



Slika 5. Točna detektiranja emocija

5. Opisi programa i klasa

Kratak opis korištenih programa i karakterističnih klasa.

5.1. Visage|SDK

Visage|SDK integrira veliki broj računalnih tehnologija za animaciju likova i računalni vid u jednostavne razvojne alate koji se lako koriste.

5.1.1. Face Tracking

Visage|SDK paket koji omogućava pronalazak i praćenje lica i njegovih točaka u videu i preko kamere.

Struktura VisageSDK::FaceData

- *FDP *VisageSDK::FaceData::featurePoints2D*

Dohvat točaka lica. Vraća točke u dvodimenzijском prostoru koji je normaliziran na veličinu slike. Ishodište koordinatnog sustava je postavljeno u gornjem lijevom kutu slike.

- *IplImage *VisageSDK::FaceData::frame*

Pokazivač na sliku iz videa ili kamere. IplImage je klasa za spremanje slika iz OpenCV-a. Pristup veličini slike je moguć preko `frame -> width`, `frame -> height`, dok se zapisu slikovnih jedinica slike pristupa preko `frame -> imageData`.

5.2. OpenCV

OpenCV (Open Source Computer Vision Library) je biblioteka otvorenog koda koja uključuje nekoliko stotina algoritama vezanih uz područje računalnog vida.

6. Zaključak

Prepoznavanje emocija je moguće ostvariti sa navedenim alatima. Ipak, najveći problem je visoka pogreška prilikom prepoznavanja.

Problem je moguće riješiti ukoliko bi se izgradila baza slika emocija koja bi imala pokrivenost sve emocije i to u nekoliko tisuća primjeraka. Slike bi morale biti frontalne sa jasnim emocijama, nikako dvojbene, jer utječe na razlikovanje emocija.

Također bi se moglo omogućiti treniranje sa drugim oblicima strojnog učenja, kako bi se moglo spojiti prepoznavanje više emocija.

7. Literatura

[1] Visage|SDK, <http://www.visagetechologies.com/>

[2] OpenCV documentation, <http://docs.opencv.org/>

8. Sažetak

U sklopu završnog rada provodi se eksperimentalno testiranje postojećeg programa za strojno učenje u vidu automatskog prepoznavanja emocija. Prilikom testiranja potrebno je pronaći bazu sa što više slika koje bi odgovarale traženoj temi te ih iskoristiti za učenje. Na temelju dobivenog klasifikatora ostvaruje se testiranje naučenog.

Rezultat testiranja je procijenjena emocija, koja može biti točna ili pogrešna. Poboljšavanje točnosti ostvaruje se sa većom bazom, koja u ovim trenucima nije moguća.

Ključne riječi: emocije, automatsko prepoznavanje, praćenje lica, strojno učenje

9. Abstract

As a part of this thesis, experimental testing of existing software for machine learning was implemented in terms of automatic emotion recognition.

During period of testing, it was necessary to find a database with enough pictures which will be right for requested topic and use them for learning.

Testing of what is learned is based on required classifier.

Result of testing is estimated emotion, which can be right or wrong.

Accomplishing improvements of estimation is possible with bigger database, which, in this time, is not available.

Key words: emotions, automatic recognition, face tracking, machine learning