Универзитет у Београду
Математички факултет

Мастер рад

Проширење функционалности TOC-search апликације препознавањем лица

Студент:
Марко Милошевић

Ментор:
Проф. др. Гордана Павловић Лажетић
Чланови комисије:
Проф. др. Зорица Станимировић
Доц. др. Саша Малков

Београд, 2014.

Садржај

1 Увод **1**
- 1.1 Идентификација особа . . . 1
- 1.2 Препознавање лица . . . 3

2 Препознавање лица **4**
- 2.1 Проблем препознавања лица . . . 4
- 2.2 Алгоритми за решавање потпроблема класификације . . . 4
  - 2.2.1 Haar-like (HL) класификатор . . . 5
  - 2.2.2 LBP Класификатор . . . 6
- 2.3 Основни поступак . . . 7
  - 2.3.1 Припрема фотографије . . . 7
    - 2.3.1.1 Детекција лица . . . 8
    - 2.3.1.2 Детекција очију . . . 9
    - 2.3.1.3 Нормализација лица и обрада фотографије . . . 11
  - 2.3.2 Претварање фотографије у параметре алгоритма . . . 13
    - 2.3.2.1 Eigenfaces алгоритам . . . 13
    - 2.3.2.2 Fisherfaces алгоритам . . . 14
    - 2.3.2.3 Local Binary Patterns Histograms . . . 14

3 Имплементација **16**
- 3.1 Евалуација алгоритама за препознавање лица . . . 16
- 3.2 Рачунарске технологије коришћене у имплементацији софтвера . . . 17
- 3.3 Примена идеје хибридизације на проблем препознавања лица . . . 19
  - 3.3.1 Детекција лица . . . 19
  - 3.3.2 Детекција очију . . . 22
  - 3.3.3 Параметризација фотографија и поређење . . . 23

4 Резултати **27**
- 4.1 Тестирање на стандардним базама лица . . . 27
- 4.2 Утицај смањења тренинг сета на поузданост алгоритма . . . 30

5 Примена софтвера за препознавање лица **34**
- 5.1 Примена на базу података TOC Search . . . 34
- 5.2 Примена у обезбеђивању објеката . . . 35
- 5.3 Примена у обезбеђивању државних граница . . . 36

6 Закључак и правци даљег развоја **38**

Глава 1

Увод

У овом раду представљен је пројекат развијен у сврху унапређења функционалности дигиталне базе тероризма и организованог криминала (TOC-search [10]). Основни задатак развијеног софтвера је проширење скупа информација на основу којих се нека особа (потенцијално терориста) може идентификовати, с обзиром на чињеницу да ниједан текстуални податак у бази није јединствен, односно не обезбеђује једнозначну идентификацију.

Развијено решење пројектовано је на начин који је врло погодан за разноврсне примене. Развијени систем се може инкорпорирати у различите софтвере, обављајући своју основну функцију препознавања лица независно од сврхе коришћења информације коју пружа. Како би се постигао степен независности од сврхе примене, представљено софтверско решење развијено је у форми клијент-сервер сервиса.

1.1 Идентификација особа

Препознавање лица само је један од приступа много ширем концепту проблема идентификације особа. Уобичајен задатак са којим се сусрећу различити системи, јесте утврђивање идентитета особе која се анализира у најразличитије сврхе. Рецимо, гранична полиција мора идентификовати сваку особу која приступа штићеној тачки, како би јој дозволила или забранила пролаз. Други пример, тренутно веома актуелан, јесте идентификација корисника неког уређаја (рачунара, телефона, и слично) у сврху одређивања степена приступа сервисима који су понуђени. Нешто другачији пример истог проблема јесте идентификација кривца или жртве неког злочина, на основу отиска прста/ДНК анализе, записа сигурносних камера, и томе слично. Наравно, ради се о широком спектру могућности, тако да су овде издвојене биометријске методе, којима припада препознавање лица.

Биометријским методама називамо оне алгоритме који подразумевају коришћење биометријских особина човека у сврху идентификације. То могу бити изглед лица, зеница ока, отисак прста, али и глас, начин хода и различите друге карактеристике које раздвајају једну јединку од осталих.

Различите методе дају различит степен поузданости, али и комплексности коришћења. Безбедносно најзаступљенији начин утврђивања идентитета је отисак прста, са добрим разлогом. Ради се о карактеристици која једнозначно одређује особу којој припада, па се резултат може са великом сигурношћу употребити у било које сврхе, чак и као правно признат аргумент. Слично је и са ДНК методом. Лабораторијском анализом се на основу телесних течности, косе или неких других биолошких трагова, може прецизно разликовати једна особа од друге.

У односу на поменуте методе, приступ овде представљен има нешто нижу поузданост, али уз квалитетан скуп улазних информација може задовољити многе потребе. Ипак, од осталих га издваја једна јако битна карактеристика, а то је лакоћа долажења до података, уз сасвим задовољавајуће резултате.

Да бисмо некога препознали уз помоћ ДНК кода, отиска прста или скенирања зенице ока, најпре морамо имати базу таквих информација. Даље, сваки пут када хоћемо да идентификујемо неку особу, неопходно је да та особа сарађује, тако што јој узимамо отисак прста, узорак ДНК или скенирамо зеницу ока и тако добијени податак упоређујемо са подацима у бази.

Са друге стране, фотографије су постале врло уобичајен материјал који нам је доступан и који се са лакоћом проналази, с обзиром на све већу покривеност јавних површина камерама, а и распрострањеност фотографија на друштвеним мрежама и интернету уопште. Природа задатка рачунарске идентификације особа нам омогућује да, углавном, најлакше дођемо до фотографија лица неке особе која нам је од интереса, много једноставније него до неке друге биометријске карактеристике. Рецимо, приликом обезбеђивања државне границе, банке, спортске дворане или другог објекта од важности, може се на одговарајуће место на улазу поставити камера и тако аутоматски прибавити више квалитетних фотографија сваке особе која улази у објекат, без њеног активног учешћа, и без застоја.

Напоменимо да се препознавање лица, као биометријска метода, не може сматрати апсолутно поузданом, али представља веома корисно помоћно средство, које се користи у комбинацији са другим информацијама и базама података које безбедносне службе поседују, у циљу превенције безбедносних претњи и откривања починилаца кривичних дела.

1.2 Препознавање лица

Препознавање лица врло је захтеван задатак, који је традиционално подразумевао одређен степен интелигенције, али који човек са релативном лакоћом испуњава. Међутим, ту логику није једноставно пренети на рачунар, односно имплементирати је тако да се може извршавати аутоматски са великим степеном прецизности.

Ипак, распрострањеност података које можемо употребити је иницирала велики и константни развој различитих алгоритама, који дају задовољавајуће резултате, погодне за коришћење у различите сврхе.

Очекивано, различити алгоритми дају различите резултате. Такође, треба узети у обзир да је један од главних услова за висок степен поузданости резултата добар скуп улазних информација. Квалитетни алгоритми могу да амортизују многе недостатке улазних података, али, за врхунску прецизност, неопходан је квалитетан и довољно обиман скуп фотографија.

У овом раду биће приказани стандардни кораци класичног алгоритма за препознавање лица, са посебним акцентом на специфичности метода који су имплементирани у предложеном решењу.

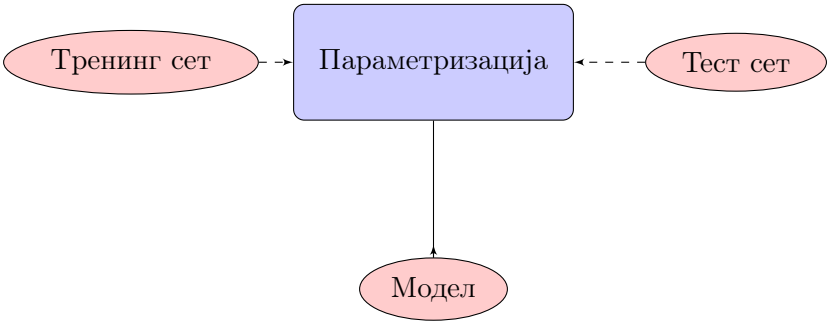
Глава 2

Препознавање лица

2.1 Проблем препознавања лица

Под методом идентификације особа препознавањем лица подразумева се алгоритам који, уз адекватан истрениран модел, за улазну фотографију лица враћа информацију о његовој класи. Другим речима, алгоритам препознавања лица подразумева класификатор фотографија лица, где класе представљају различите особе.

Као и сваки класификатор, алгоритам захтева скуп фотографија са задатим класама - тренинг сет, који се адекватним подалгоритмима претвара у модел класификатора, као што је приказано на Слици 2.1. Након тога, класификатор за сваку нову фотографију из тест сета проналази класу из модела која јој највише одговара, тј. која се од ње најмање разликује по неком унапред задатом критеријуму.



Слика 2.1: Општи поступак рада класификатора

2.2 Алгоритми за решавање потпроблема класификације

С обзиром да су у предложеном софтверском решењу коришћени стандардни, широко популарни класификатори за решавање неколико потпроблема у различитим

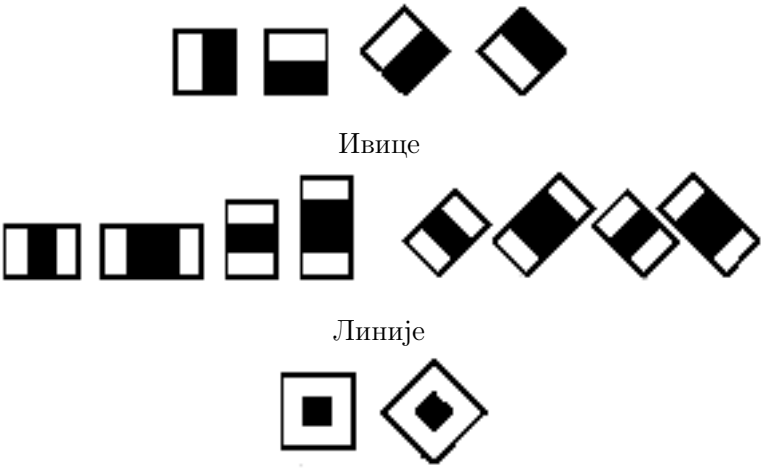
фазама алгоритма, у наредним подсекцијама биће дат њихов кратак приказ.

2.2.1 Haar-like (HL) класификатор

Haar-like (HL) класификатор је општи класификатор за фотографије објеката. Уз добро истениран модел (користећи квалитетан сет фотографија), овај алгоритам са јако великом прецизношћу разликује типове облика на сликама, чак и ако су у питању сложени облици.

Класификатор функционише тако што посматра суседне правоугаоне зоне фотографије, рачуна суме интензитета осветљења за сваку од зона и бележи разлике. Добијене информације класификатор користи за категоризацију делова слика.

На пример, ако тренирамо класификатор да детектује лице, алгоритам ће брзо уочити да је регион очију тамнији од региона образа. Уколико ту карактеристику повежемо са осталим карактеристикама, може се добити изузетно прецизан модел. У решењу које су предложили Viola и Jones [12] детектоване су само хоризонталне и вертикалне карактеристике. С обзиром да се испоставило да доста квалитетних карактеристика није било могуће забележити на тај начин, Lienhart и Maydt су годину дана касније предложили нови концепт који омогућава ротацију оригиналних карактеристика HL класификатора од 45° [6]. Предложена имплементација софтвера за препознавање лица користи карактеристике приказане на Слици 2.2.



Слика 2.2: Карактеристике које HL класификатор користи

Улазни сет података за процес тренирања HL класификатора подразумева коришћење одређеног броја фотографија скалираних на исту величину, које представљају позитивне примере, и одређеног броја насумичних фотографија исте величине, које су негативни примери. Алгоритам користи само интензитете пиксела фотографије, тј. ради са црно-белим (енг. grayscale) фотографијама.

Када имамо сет скалираних и прилагођених фотографија, HL класификатор формира модел у ком бележи карактеристике облика који се понављају у свим позитивним примерима.

Након тога, класификатор се примењује на нову фотографију и тражи карактеристике присутне у фотографијама из тренинг сета. Класификатор је направљен тако да се лако може скалирати, како би пронашао тражени облик произвољне величине. Једна од главних предности HL класификатора је изузетна брзина, јер користи посебну структуру података која брзо рачуна суме квадратних подскупова података (отуда скалирање на квадратни облик) и омогућује детекцију у константном времену.

Облици које овај класификатор може обрађивати могу бити лица или само очи (што користимо у овом систему), али и дрво, кућа, флаша... Интересантна могућа примена у домену безбедности јесте детекција оружја; тренинг сет може садржати велики број фотографија пушака, пиштоља и сличних облика од интереса, помоћу којих бисмо добили модел који успешно упозорава систем на присутност тих објеката у кадру фотографије, нпр. камере.

2.2.2 LBP Класификатор

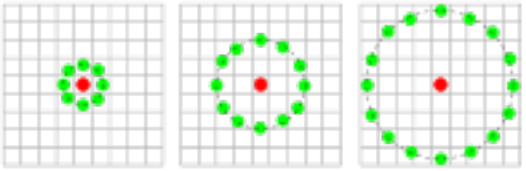
Други метод који је примењен је Local Binary Patterns (LBP) класификатор [8]. То је такође општи алгоритам који има много примена, али у нешто ширем спектру од HL класификатора. У предложеном софтверском решењу, LBP се користи у решавању три потпроблема: детекција лица, детекција очију и само препознавање лица.

За разлику од HL класификатора који посматра целу фотографију, LBP је третира као скуп мањих подскупова.

Вектор LBP карактеристике добија се на следећи начин:

- Улазна фотографија се издели на ћелије унапред одређене величине (у овој имплементацији ћелије су 16×16 пиксела)

- Светлосни интензитет сваког пиксела сваке ћелије упоређује се са свим непосредним суседима. Након тога врши се проширивање околине пиксела, као што је приказано на Слици 2.3. Уколико је интензитет централног пиксела мањи од интензитета пиксела суседа, у вектор ће се уписати бинарна нула, у супротном уписаће се јединица. Резултат се потом претвара у осмобитни број, с обзиром да се обрађује 8 пиксела.
- За сваку ћелију израчунава се хистограм фреквенција свих осмобитних бројева који се јављају.
- Врши се нормализација хистограма и њихово спајање у јединствен вектор ћелије која је анализирана.



Слика 2.3: Суседи пиксела који се анализирају при примени LBP класификатора

Главна предност оваквог приступа је значајно поједностављење и убрзање процеса детекције, јер уместо рада са подацима већих (или великих) димензија, ради се са ограниченим подацима, релативно малих димензија.

2.3 Основни поступак

И поред великог броја различитих приступа у решавању проблема препознавања лица, основни кораци су им у суштини заједнички, иако се могу реализовати на различите начине.

Поступак се, генерално, састоји из више појединачних корака који се баве припремом фотографије за примену алгоритма, и самог алгоритма претварања лица у параметре адекватне за алгоритам. У наредним подсекцијама укратко ће бити описан сваки од појединачних корака и коришћених алгоритама.

2.3.1 Припрема фотографије

Фотографија коју разматрамо може бити произвољног квалитета и садржати произвољне облике. То чак, са рачунарске тачке гледишта, не мора бити фотографија на којој постоји нека особа, или лице. Узмимо као пример безбедносну камеру постављену на улаз неког објекта од значаја, на којој се, очекивано, ниједна особа

не појављује неко време. Систем би у таквој ситуацији требало адекватно да реагује, тј. да игнорише све кадрове послате са камере, све док се на њима не појави особа, заправо лице. То имплицира да је један од првих задатака који софтвер мора обавити детекција, тј. проналажење лица. У општем случају, овај корак се може применити и више пута, јер фотографија коју разматрамо може садржати више лица од интереса.

Након детекције лица (једног или више), примењује се нормализација и поправљање квалитета сваког лица појединачно, које се као такво прослеђује на даљу анализу. Овај корак се може обавити и на почетној фотографији, пре детекције, али принцип остаје непромењен, само је питање редоследа корака.

2.3.1.1 Детекција лица

Овај потпроблем веома је комплексан и има врло широку примену која излази из оквира проблема препознавања лица. На пример, дигиталне камере приликом фотографисања уоквирују лица у кадру, јер је очекивано да корисник жели да му управо те зоне буду добро фокусиране и осветљене. Из тог разлога се параметри као што је отвор бленде, фокус и експозиција прилагођавају.

Улазна информација сваког од алгоритама јесте фотографија која може (али и не мора) садржати лице, а излаз представљају координате зоне фотографије за коју коришћени класификатор процени да одговара форми лица. Слика 2.4 приказује графички уоквирено лице на тест фотографији. Правоугаоник за који детектор сматра да представља лице може се на произвољан начин проследити алгоритму параметризације: овде су коришћене апсолутне координате његове горње-леве и доње-десне тачке. Било би могуће уштедети нешто меморијског простора тако што бисмо бележили само једну тачку и димензију, узевши у обзир да користимо искључиво лица уоквирена у зоне квадратног облика, али због компатибилности са другим системима и могућих надоградњи, као и мале уштеде, задржан је овај робуснији, али и стандарднији формат враћања резултата.

Постоји много решења проблема детекције лица, а свако од њих има предности и мане, и тешко их је поредити на адекватан начин. Због сложености проблема, дешава се да један алгоритам даје јако добре резултате на једном скупу фотографија, али изузетно лоше на другом, и обрнуто.

Управо из тог разлога, у имплементацији овог софтвера користили смо комбинацију два приступа за проблем детекције лица: коришћење класификатора уз помоћ раније објашњених Haar-like карактеристика дигиталне фотографије и Local Binary Patterns класификатора.

Слика 2.4: Пример резултата алгоритма за детекцију лица

Додатно ограничење које је постављено за свако лице јесте минимална величина лица. У сврху што веће поузданости система, овај корак одбацује свако лице мање од 70×70 пиксела.

2.3.1.2 Детекција очију

Када детектор лица издвоји зону која је од интереса, таква фотографија теоретски одмах може бити прослеђена процедури за препознавање лица, тј. може бити претворена у нумеричке параметре које тај алгоритам упоређује.

Међутим, иако би овакав софтвер дао добре резултате на идеалним фотографијама, јасно је да варијација израза лица, као и положај и нагиб главе при фотографисању могу имати огроман негативан утицај на поузданост софтвера.

На пример, уколико се на фотографији налази лице под релативно великим нагибом, за очекивати је да се оно највише подудара са управо таквим претходним примерима из тренинг сета. Црте лица, величина очију или носа постају ирелевантни, ако је облик главе под углом од рецимо 45°.

Слично је и са величином лица. Потребно је да се релативне позиције одговарајућих карактеристика поклапају, како би се њихов облик и величина могли успешно упоредити, и како би имао пресудан утицај на класификацију.

То се постиже дефинисањем кључних тачака лица, у односу на које се врши скалирање и ротирање. Најпогодније зоне су свакако очи, тако да се регион лица прослеђује алгоритмима за детекцију очију, на основу којих се касније врше трансформације.

У општем случају, у процесу препознавања лица овај корак није неопходан, тако да се често не захтева успешна детекција очију. Али, у предложеном софтверском решењу, због повећања поузданости, систем потпуно одбацује лица на којима не може да детектује очи. Разлог томе је одржавање квалитета базе лица, и самих модела класификатора, јер искривљена или непрецизно скалирана лица могу имати значајан, негативан утицај на резултате.

Фаза препознавања очију једна је од најкритичнијих у целом поступку. Идеја је да направимо алгоритам који ће бити што робуснији, као и отпорнији на различите гримасе, близине субјеката камери, ротације главе и томе слично, тако да је тешко идеално пронаћи било коју фиксну зону, па и очи.

Када говоримо о алгоритмима за детекцију лица, њихов задатак је, условно речено, доста лакши. Постоји много карактеристика на које се може обратити пажња; зона целог лица је релативно велика, тако да то амортизује локалне проблеме, као што су отворена/затворена уста, и томе слично. Из тог разлога не изненађују велике варијације у резултатима различитих алгоритама на задатку детекције очију.

Узмимо за пример моделе развијене у Open Source Computer Vision (OpenCV) библиотеци [3]. OpenCV садржи неколико различитих модела за детекцију очију, међутим, како је и у самом извору [3] наглашено, модели се међусобно битно разликују и перфрормансе им значајно варирају од једног до другог тренинг сета података.

OpenCV библиотека укључује модел HL класификатора који са огромном прецизношћу детектује отворене очи (преко 99% на квалитетним базама фотографија), али има прилично лоше резултате за све остале опције. Такође, због повећане прецизности, користе се два модела, један за лево и други за десно око.

За следећи модел класификатора који OpenCV нуди у самом опису стоји да решава проблем затворених очију, што се показује и у тестирању. Ипак, таква флексибилност доноси и смањење поузданости алгоритма у детектовању отворених очију.

Трећи, последњи велики проблем на који наилазимо, јесте проблем детекције очију, у случају када су оне прекривене наочарима. Ради се о прилично незгодној ситуацији за детектор, али је у оквиру библиотеке понуђен и модел HL класификатора

који сасвим коректно амортизује овај проблем, иако, очекивано, не постиже резултате на осталим случајевима као претходно описани модели. Управо из тог разлога, овде представљено решење комбинује чак четири различита модела.

Последњи модел који библиотека нуди јесте LBP класификатор који постиже релативно избалансиране резултате за сваку од горе поменутих ситуација.

2.3.1.3 Нормализација лица и обрада фотографије

Како су фотографије које софтвер добија као улазни податак различите величине и квалитета, неопходно је да лица детектована на фотографији скалирамо и софтверски обрадимо, како би као таква била што адекватнија и унификованија за примену у алгоритму.

У различитим решењима [1] [4] користи се више различитих приступа. Овде дајемо сет поступака који је коришћен у овом конкретном решењу, иако је могуће применити разне процедуре у циљу поправљања квалитета улазних фотографија.

Један од основних проблема са којим се сусрећемо у улазним подацима је различит нагиб лица на фотографијама, као и различита величина лица (нека су фотографисана изблиза у већој резолуцији, нека издалека, итд).

Најпоузданији начин да то исправимо је утврђивање позиције очију на фотографији детектованог лица и нормализација фотографије. Уколико софтвер закључи да очи нису у истој равни, најпре ћемо лице ротирати, како би сви подаци у бази били сачувани на исти начин. Следећи корак је коришћење информације о позицији очију како бисмо скалирали лице, и на тај начин уједначили димензије свих фотографија лица.

Када добијемо фотографију детектованог лица које је ротирано и скалирано на задату величину, приступа се дигиталној обради фотографије, у циљу поправљања њеног квалитета.

Постоје алгоритми који користе информације о нијансама боја, тј. о тоналитету, али овде смо се одлучили за другачији приступ. Наиме, када радимо са колорним фотографијама, сама нијанса боје лица може имати веома добар ефекат у моделу за класификацију (слично је и са другим деловима фотографије, косом, очима итд.), али коришћење те информације може проузроковати велики број негативних појава.

Узмимо као пример фотографију истог лица у јутарњим и поподневним сатима. Познато је да на природној светлости ујутру све боје делују доста хладније (плавичасте), а поподне, нарочито пред залазак сунца, значајно су топлије (црвенкасте).

Долази се до закључка да чак и овако једноставна измена услова може деградирати поузданост алгоритма, иако се ради о особи чији изглед није нимало варирао, чак је и извор светлости остао исти. Услед овога нема сврхе разматрати утицај различитих извора осветљења (сунце, сијалица, неонска сијалица, блиц итд.), а нарочито утицај различитог изгледа самог фотографисаног субјекта (промена тена коже, промена боје косе и томе слично) на квалитет класификатора.

Из тог разлога, предложено софтверско решење бира анулирање свих информација о нијансама боја, и задржавање само податка о интензитету, тј. претварање фотографије у црно-белу, односно у нијансе сиве. Тачно је да у неким ситуацијама то значи губитак драгоцене информације, али, због горе описаних могућих аномалија, мали број решења их користи.

Следећи проблем је отклањање дигиталног шума. Ради се о аномалији која настаје пре свега на фотографијама направљеним у условима слабог осветљења. Како би сензор дигиталног фотоапарата (камере), или филм у аналогној фотографији, прихватио довољно светлости, његова осетљивост се повећава, па долази до насумичних грешака, пиксела са екстремним интензитетима. Такве информације би могле значајно да смање поузданост алгоритма, али се овај проблем релативно лако решава.

Проблем пиксела са екстремним интензитетима превазилази се такозваним „умекшавањем" фотографије. Овај процес детектује велике разлике у интензитетима пиксела, и ако наиђе на, на пример, пиксел са интензитетом 255 окружен пикселима интензитета 60, јасно је да се ради о аномалији. Једноставно уравнотежавање околине пиксела решава проблем, без губљења значајних информација.

Последњи корак у дигиталној обради фотографије коришћен у конкретном систему јесте процедура нелинеарне еквивализације хистограма. Класична еквивализација хистограма решава проблем светлих/тамних фотографија, тако што их своди на неку "нормалну" осветљеност, али овде користимо посебну, нелинеарну варијанту процедуре. Наиме, додатни проблем на који наилазимо су различито осветљене стране лица, као последица лоше постављеног извора светлости. На пример, уколико извор светлости не осветљава уједначено цело лице, већ због своје позиције чини леву страну светлијом од десне или обрнуто, еквивализација целе фотографије не би дала задовољавајуће резултате. Зато најпре обрађујемо одвојено леву и десну страну, па их тек онда спајамо и сређујемо.

Веома је битно да се дигиталном обрадом сачува што више информација које праве разлику између једне особе и осталих, како би класификатор вратио што поузданију информацију.

2.3.2 Претварање фотографије у параметре алгоритма

Претходни кораци обезбеђују нам пут од произвољне фотографије до једне (или више) издвојене зоне лица, која је задатог формата, поправљеног квалитета, и садржи само информације о интензитету осветљења у појединачним пикселима. Такве особине је чине погодном за многе рачунарске примене. Овако обрађене фотографије лица претварају се у нумеричке параметре који су погодни за упоређивање са другим фотографијама (заправо њиховим параметрима), у циљу одређивања најсличније (по неком критеријуму) класе.

И за овај корак, као и за претходне, постоји много приступа, који могу бити изузетно различити. Једна од првих метода којим је проблем препознавања лица решаван је претварање лица у скуп значајних тачака (енг. landmarks) [2]. Значајне тачке лица су, на пример, зенице ока, корен носа, врх носа, вршни и доњи део уха, и тако даље. Након детекције значајних тачака рачунају се односи између значајних тачака (растојања, односи растојања, углови међу дужима које спајају значајне тачке, итд.) и пореде се са одговарајућим односима других класа неким алгоритмом, обично PCA (Главна компонентна анализа, Principal Component Analysis). На тај начин врши се класификација фотографија, тј. препознавање лица. Ипак, проблем одређивања кључних тачака није једноставан (о чему ће бити више речи касније), тако да су у овом решењу коришћени другачији приступи.

2.3.2.1 Eigenfaces алгоритам

Проблем репрезентације фотографије у пуним димензијама $P \times Q$ може се решавати на више начина. Дводимензиона црно-бела фотографија већ је директно представљена као један вектор векторског простора димензија $P \cdot Q$, али нису све димензије једнако битне за класификацију. Да бисмо донели одлуку које димензије узети у обзир, испитујемо варијансу података, али пре свега оних које носе главну различитост међу информацијама.

Principal Component Analysis (Главна компонентна анализа, PCA) [5] је метод који обезбеђује одвајање битних од небитних информација, тј. претварање сета променљивих за које постоји могућност да корелирају у мањи сет података који нису у корелацији. PCA метод проналази смер са највећом варијансом у подацима, који се назива „главним" компонентама.

Eigenfaces алгоритам [11] заснован је на PCA методи и користи сопствене вредности (eigenvalues) и сопствене векторе (eigenvectors) у одређивању параметара. Под Eigenfaces алгоритмом се подразумевају следећи кораци:

- пројектовање свих тренинг података у PCA потпростор
- пројектовање фотографије коју препознајемо у PCA потпростор
- проналажење најближег суседа те нове фотографије са тренинг подацима.

Последњи проблем на који наилазимо практичне је природе, јер за већи скуп података величина простора потребног за смештање података и време обраде података експоненцијално расту. Ипак, процедуром која превазилази оквире овог рада, могуће их је свести на разумне величине, користећи особине сопствених вектора и матрице сопствених вектора.

2.3.2.2 Fisherfaces алгоритам

PCA анализа, коришћена у горе описаном алгоритму Eigenfaces, у циљу смањења димензионалности векторског простора одбацује велику количину података, и у пракси врши доста добру селекцију. Подаци које одбацује, тј. које процењује небитним, у великој већини случајева су заиста такви. Али, очекивано, често се међу тим подацима наилази и на потенцијално битне информације, што PCA због свог начина рада није у стању да открије.

Основни проблем PCA алгорима је посматрање сваке класе засебно. У циљу превазилажења овог недостатка, Fisherfaces алгоритам користи другачији приступ смањењу димензије векторског простора под називом Linear Discriminant Analysis (LDA) [7]. Да би пронашао комбинације особина које најбоље раздвајају различите класе, LDA метод максимизује однос унутар-класног и међу-класног одступања, уместо максимизације глобалног одступања информација. Дакле, води се идејом да исте класе треба описати тако да растојање међу њима буде што је могуће мање, а различите класе максимално раздвојити у ниже-димензионој репрезентацији.

2.3.2.3 Local Binary Patterns Histograms

За разлику од претходна два алгоритма који посматрају фотографије као векторе у простору великих димензија, Local Binary Patterns (LBP) алгоритам дели фотографију на ћелије, и појединачно их обрађује, слично као у примени на детекцију лица, што је објашњено у секцији 2.2.2.

За примену LBP алгоритма се, након појединачне анализе, додатно спајају локални хистограми у општи модел, применом процедуре Local Binary Patterns Histograms

(LBPH). Ова процедура решава многе недостатке претходних захваљујући једноставној чињеници да је интересују локалне особине, а не опште. Често се дешава да се информација по којима се једна особа разликује од друге изгуби у генерализацији података, што је превазиђено обрадом података на векторским просторима мањих димензија.

Још једна предност која може бити значајна за различите имплементације софтвера за препознавање лица, укључујући овде описану примену на TOC-search апликацију, јесте једноставно додавање нове фотографије у модел, без промене информација које су раније израчунате, што није случај са прва два алгоритма.

Глава 3

Имплементација

3.1 Евалуација алгоритама за препознавање лица

Један од основних проблема развијања поузданог система за препознавање лица јесте немогућност прецизне евалуације резултујућег софтвера. Ни за један алгоритам коришћен у неком од корака процедуре софтвера не може се рећи да је дефинитивно најбољи и најпогоднији, јер квалитет резултата изузетно зависи од улазних података. На пример, LBPH алгоритам, због чињенице да бележи и анализира локалне особине, одлично амортизује проблем различитог осветљења фотографија, али лоше реагује на значајније разлике у експресијама лица. Са друге стране, холистички приступ Fisherfaces алгоритма релативно лако превазилази проблем гримаса, али разлике у осветљењу могу учинити нумеричке вредности које користимо за поређење целе фотографије драстично различитим. Слично је са фазама детекције лица, детекције очију и тако даље.

Додатни проблем код детекције лица и очију јесте чињеница да се одговарајући класификатори формирају помоћу неког тренинг сета, па исти алгоритам са различитим моделима може дати изузетно различите резултате. Најбољи пример вероватно је детекција очију. Уколико се класификатору зада тренинг сет са квалитетним фотографијама особа које имају отворене очи и формира се модел на том основу, настаће велики проблем са детектовањем лица која имају затворене очи. Ако се овај проблем покуша решити помоћу другог сета који садржи особе са затвореним очима, не може се добити ни приближно квалитетан систем за случајеве првог типа. Уз све то постоји и проблем са особама које носе наочаре на фотографијама.

Могуће решење је формирање модела тако што тренинг сет који се пружа класификатору садржи доста примерака који покривају сва три случаја. Али, имајући

у виду чињеницу да је заиста драстична разлика у изгледу зоне очију од случаја до случаја, очекивано, добијени модел даје незанемарљив број такозваних "False positive-a", тј. често прогласи неку зону зоном очију, иако то није случај.

Имајући у виду велику различитост врста улазних података, предложено софтверско решење у сваком кораку алгоритма комбинује више различитих алгоритама, односно метода, користећи добре стране сваке од њих, у циљу постизања што бољих резултата.

3.2 Рачунарске технологије коришћене у имплементацији софтвера

У циљу независности система од било каквих рестриктивних лиценци, целокупан софтвер развијен је уз коришћење алата отвореног кода.

Циљани оперативни систем за серверску апликацију је Unix, док је коришћени програмски језик C++. Ипак, треба имати у виду да је, због начина функционисања софтвера, могуће користити решење из било ког оперативног система (укључујући и мобилне оперативне системе). Једино је неопходно испоштовати метод комуникације са сервером, а то је коришћење TCP/IP улаза (eng. socket) на произвољном порту. Тако је обезбеђена изолација софтвера од ширег контекста у ком се користи.

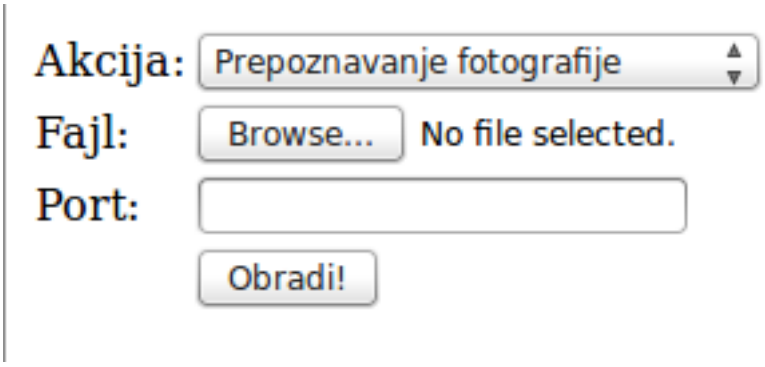
Главне класе које су коришћене у имплементацији софтвера долазе из библиотека OpenCV [3] пакета.

Тип података који садржи објекат саме дигиталне фотографије је "Mat" и представља матрицу података. Прецизније, састављен је из два дела: заглавља (где су записане информације о димензијама матрице, методу чувања података, адреси и томе слично) и показивача на саму матрицу која садржи вредности пиксела. Над овим типом врше се све трансформације (скалирање, ротирање), као и анализа појединачних података, у циљу детекције објеката и параметризације фотографије.

Све класе везане да детекцију лица и очију, као и за само препознавање лица су такође део OpenCV библиотеке, тако да им се без додатних модификација прослеђује објекат типа "Mat".

За потребе комуникације сервиса за препознавање лица са неким другим софтверима (као што је TOC-search база), развијене су класе "ServerSocket" и "ClientSocket". Ове класе омогућују прилагођавање нативних библиотека језика C++ за TCP/IP комуникацију конкретној примени, пружајући помоћне функције за једноставну иницијализацију, повезивање сервера и клијента, као и слање/примање порука.

За потребе тестирања софтвера и демонстрације његових могућности развијен је једноставан веб интерфејс, уз помоћ скрипте писане у програмском језику PHP, приказан на Слици 3.1.



Слика 3.1: Веб интерфејс софтвера за препознавање лица

Команде које сервис прихвата су:

- 'T': Train - тренирање модела. Захтевани параметар је .csv фајл који за сваку фотографију садржи њену адресу праћену нумеричким параметром, класом. Повратна информација говори о успешности тренирања модела; или се враћа буловска вредност "1" (уколико је процедура извршена по плану), или "0", праћена текстуалним описом евентуалне грешке.

- 'R': Recognize - препознавање једне фотографије. Ова команда је праћена адресом фотографије коју систем треба да класификује, а повратна информација се састоји од целобројне позитивне вредности, нумеричке идентификације класе за коју софтвер сматра да је одговарајућа, и децималне вредности, која представља дистанцу пронађеног лица од те класе. Та дистанца се на неки начин може употребити у процени поузданости резултата, као што ће бити објашњено касније. Уколико препознавање није успело из било ког разлога (датотека недоступна, модел недоступан, лице и очи нису детектоване и томе слично), нумеричка вредност класе се поставља на "0", и уместо дистанце се шаље текстуална порука која описује грешку.

- 'A': Append - додавање фотографије у модел. Ова команда захтева као параметар адресу фотографије и њену класу и, као и код акције тренирања, повратна информација само говори о успешности/неуспешности извршења команде.

- 'M': Multiple Recognize - препознавање сета података. Команда која је пре свега коришћена у развоју и тестирању сервиса, мада може имати примену и у реалним системима. Слично као и код тренирања модела, захтевани параметар је .csv фајл који садржи низ фотографија за препознавање, а повратни

податак служи за информисање корисника о статистици примене система, тј. о резултатима који су остварени. Ту се налазе информације као што су број пронађених лица (који су прослеђени на препознавање), број тачно класификованих лица, поузданост, и просечна релативна дистанца свих предикција. Како ове информације изгледају на конкретним базама, биће додатно појашњено у Секцији 4.2.

3.3 Примена идеје хибридизације на проблем препознавања лица

Хибридизација различитих система и решења један је од начина превазилажења проблема који се јављају приликом примене сваког система/решења појединачно. Показало се да принцип хибридизације може бити успешно примењен и при решавању проблема препознавања лица.

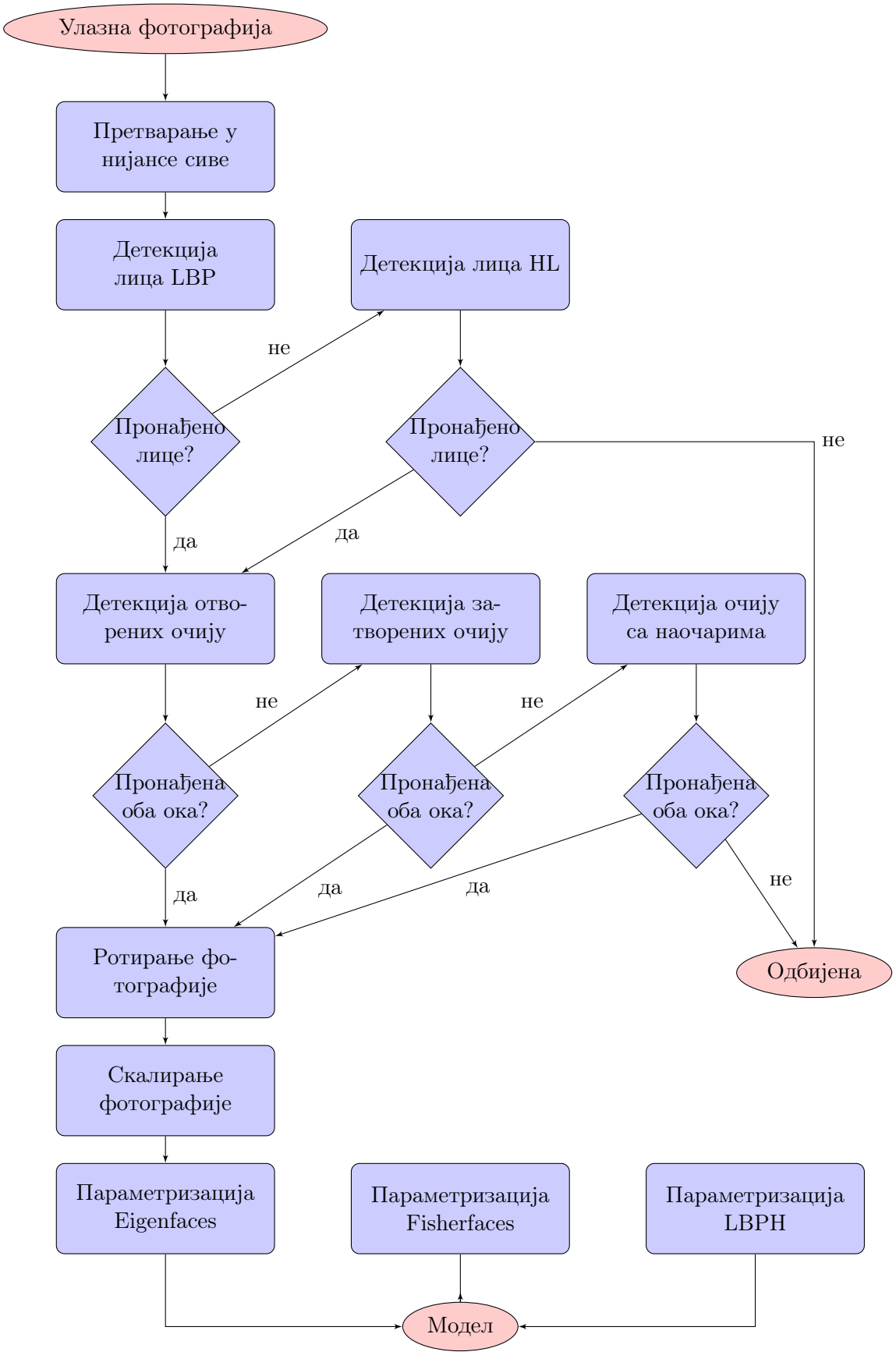
Три кључна корака процедуре препознавања лица, која могу произвести битно различите резултате, јесу детекција лица, детекција очију и параметризација фотографије. Остали кораци су мање осетљиви на различитост улазних података и као такви остављени су да независно, са фиксираним параметрима, обављају задатке. Међутим, могуће је и у овим независним корацима направити неке помаке у смислу побољшања. Рецимо, скалирање и ротирање фотографије би теоретски могло да да боље резултате за неку другу одабрану резолуцију, али ту нема битних разлика. Слично је и са дигиталном обрадом: кораци се примењују сукцесивно, а, иако сигурно комплетно решење може дати боље резултате убацивањем/избацивањем неког од њих, ту је мали простор за напредак, па се одабрани скуп процедура извршава независно од улазног сета фотографија.

Схематски приказ поступка примењеног у софтверском решењу описаном у овом раду дат је на Слици 3.2.

У наставку обрађујемо кораке појединачно, и описујемо начин хибридизације решења за сваки од њих.

3.3.1 Детекција лица

HL и LBP класификатори, који су описани у 2.2.1 и 2.2.2, тренирају се помоћу скупа улазних фотографија издељеног на позитивне и негативне примерке и задатих параметара захтеване прецизности. Предложено софтверско решење користи



Слика 3.2: Алгоритам конкретног софтвера за препознавање лица

OpenCV [3] библиотеку. Софтверски систем за тренирање ових класификатора користи неколико параметара при тренирању модела како би контролисао захтевану прецизност и баланс између false-positive и false-negative резултата. Најважнији од њих је minHitRate који контролише минимални удео погодака за сваку рунду формирања модела и maxFalseAlarmRate који ограничава број примерака који су погрешно класификовани као негативни.

Већи број модела класификације могуће је комбиновати на више начина. Који ће приступ бити одабран, зависи пре свега од захтева апликације и типичног сета улазних података.

У општем случају, грешке класификатора са две класе (позитивна и негативна) деле се у две групе: примерак погрешно класификован као позитиван (false-positive) и погрешно класификован као негативан (false-negative).

Неке примене система као примарни циљ имају детекцију сваког лица, те никако не желе појаву false-negative примерака, тј. лица на које софтвер није указао, а толерантнији су на false-positive примерке, у којима софтвер помисли да се ради о лицу, а то није случај. Рецимо, нека сигурносна камера која покреће снимање када уочи лице у кадру, а иначе га одбацује, никако не сме пропустити његово појављивање. Са друге стране, вишак снимљеног материјала, који се јавља услед погрешне процене алгоритма да се ради о лицу тамо где га нема, не изазива велику штету, осим што остаје забележен додатни неискористив материјал. У неким другим применама, пожељна је обрада искључиво квалитетних фотографија лица, и није проблем уколико софтвер неко од њих пропусти, али није пожељан улазни податак за који се не може са сигурношћу тврдити да представља лице.

Софтвер за препознавање лица тражи неки балансирани детектор лица, са што је могуће већим укупним процентом погодака, независно од тога да ли су грешке једног или другог типа.

С обзиром да у конкретном случају имамо два модела класификације (један LBP и један HL) који дају неки буловски излаз (постоји/не постоји лице), логично комбиновање је повезивање буловских оператора.

Уколико их повежемо "ИЛИ" оператором, имамо ситуацију да ће нека зона бити проглашена лицем, уколико било који од њих мисли да се ради о позитивном примерку.

Такав приступ као резултат има смањење броја false-negative случајева, тј. неко лице треба да буде пропуштено од стране оба алгоритма да би га "ИЛИ" оператор одбацио. То је свакако позитиван резултат, али не треба заборавити да ће такав

приступ неминовно повећати број false-positive случајева. Рецимо, у ситуацији у којој HL класификатор правилно процени да се не ради о лицу, а LBP погреши, "ИЛИ" оператор ће вратити погрешан резултат, прогласиће зону лицем, јер је један од алгоритама погрешио, иако је други био у праву.

Други приступ био би повезивање излаза из појединачних модела "И" оператором ради формирања "крајње одлуке". Позитиван ефекат би свакако била елиминација одређеног броја false-positive случајева. Да би софтвер проценио да се ради о лицу иако то није случај оба подалгоритма би морала да погреше, што је свакако мање вероватно.

Ипак, слично као код "ИЛИ" оператора, ово је мач са две оштрице јер би позитиван ефекат смањења false-positive случајева био праћен негативним ефектом раста броја false-negative случајева.

У предложеном софтверском решењу коришћен је баланс између два поменута приступа, односно, модификација модела класификатора, како бисмо ублажили/елиминисали негативне ефекте буловских оператора примењених на њиховим излазним вредностима.

Ту постоје две прилично равноправне опције: повећати захтевану поузданост погодака подалгоритама и онда их повезати "ИЛИ" оператором, или смањити захтевану поузданост појединачних модела, али их повезати "И" оператором.

Имајући то у виду, за потребе овог софтвера истренирани су по један модел HL и LBP класификатора са повећаном захтеваном поузданошћу (смањеним бројем false-positive случајева), а затим њихово комбиновање "ИЛИ" оператором које обезбеђује да, уколико је лице пронађено у макар једном од подалгоритама, оно буде сматрано лицем у хибридизованом решењу.

3.3.2 Детекција очију

Од претходног корака (корака детекције лица) не захтева се превише строга поузданост, јер резултат који се добија није коначан.

Прецизније, иако детектор прогласи неку зону лицем то само значи да ће тај регион бити прослеђен детектору очију, али не мора да значи да ће бити заиста коришћен кроз све остале кораке софтвера.

Са друге стране, када се овај корак заврши, уколико је алгоритам вратио позитиван резултат, таква фотографија се даље прослеђује на обраду која нема више могућност одбацивања примерка, као што је то случај са детекторима лица и очију. Ако

је детектор очију погрешио и пронашао очи тамо где их нема, таква фотографија ће бити скалирана, ротирана (вођена погрешним информацијама), обрађена и прослеђена на параметризацију ради поређења са другим фотографијама. Очекивано, уколико се ради о фотографији која поуздано не садржи лице, тај примерак може озбиљно урушити квалитет модела и његову поузданост.

Као што је интуитивно јасно, детекција очију није нимало лак задатак. Основна два проблема потичу из чињенице да се ради о зони мале резолуције (са мало детаља на које можемо да се ослонимо) и да та, ионако мала, зона може драстично варирати.

Три основна (раније помињана) случаја, на која детектор може наићи су отворене очи, затворене очи и очи прекривене наочарима. Имајући у виду све три могуће комбинације, имплементација једног модела класификатора ипак производи недовољно добре резултате. Са друге стране, уколико се направи један модел који прихвата све три варијанте он ће бити "превише толерантан", често ће различите облике проглашавати очима, иако они то нису, тј. имаће релативно висок број false-positive резултата. Ако се покуша са отклањањем овог проблема повећавањем захтеваног степена поузданости позитивних резултата, јавља се већи број false-negative случајева, па се алгоритам деградира у том смеру.

Приступ који смо користили у предложеној имплементацији релативно је сличан принципу комбиновања модела код детекције лица. Уз помоћ три различита тренинг сета формирали смо три различита модела, тако да користимо модел који одлично проналази отворене очи, али јако лоше затворене и очи прекривене наочарима, затим други који проналази затворене очи и тако даље.

Такође, модели су додатно тренирани за лево и десно око посебно, јер је ипак значајна разлика у изгледу из угла класификатора, иако се из угла човека то сматра само сликом у огледалу.

Начин комбиновања добијених модела идентичан је као код детекције очију. Пролазимо кроз све моделе и повезујемо њихове резултате буловским "ИЛИ" оператором.

3.3.3 Параметризација фотографија и поређење

Пуно је различитих приступа поређењу две фотографије, нарочито када имамо у виду да се на њима налазе лица. Једна од првих метода које су примењене за решавање овог проблема јесте поређење неких величина изведених из значајних тачака лица, такозваних landmark-ова [2].

То је систем који је сасвим коректно функционисао, и давао задовољавајуће резултате, али само на одређеном "типу" лица. Наиме, уколико би те значајне тачке биле

савршено прецизно постављене и уколико би особа била фотографисана са "нормалним" изразом лица, без деформација, вероватно би овај алгоритам и даље био најбољи.

Међутим, ситуација у пракси најчешће није таква. Честа је појава да се значајне тачке или детектују на (мање или више) погрешним местима, или се не детектују уопште. То је могуће превазићи тако што сваку фотографију која пролази кроз систем обучена особа додатно обрађује, потврђујући или модификујући локације тачака које је алгоритам за њихово аутоматско постављање одредио. Ипак, ради се о заморном задатку, нарочито на великим базама, па ово није тако лак начин елиминисања проблема. Додатно, на око врло мале грешке у позицијама тачака праве потенцијално велику разлику у вредностима које алгоритам пореди, па самим тим и у резултатима.

Чак и када имамо фотографије са идеално постављеним значајним тачкама и даље можемо добити лоше крајње резултате овог система. Разлог је једноставан, особа може бити фотографисана са различитим изразима лица и тако учинити услове драстично различитим. На пример, уколико се пореди размак између крајњих тачака усана са леве и десне стране, особе фотографисане са осмехом на лицу ће личити једне на друге, што свакако није добар сценарио.

Наравно, овакав алгоритам има и низ предности, што га чини квалитетнијим на идеалним сетовима података од неких других, али наша идеја је да покријемо што већи спектар могућности.

Управо из тог разлога, из скупа постојећих алгоритама за поређење фотографија овде су одабрана ова три. Сваки од њих има своје предности и мане, али заједно могу квалитетно покрити широк спектар могућих случајева.

Може се рећи да прва два алгоритма, Eigenfaces 2.3.2.1, и Fisherfaces 2.3.2.2 алгоритам имају сличан, холистички приступ, самим тим и сличан спектар различитих варијанти тренинг и тест сета фотографија на којима квалитетно функционишу, а да је Local Binary Patterns Histograms 2.3.2.3 алгоритам који боље третира другачији сет података. Ипак, они имају и својих разлика, па су оба искоришћена. Слично као код детекције лица и детекције очију, излазне вредности ова три алгоритма можемо комбиновати на много различитих начина.

Природни начин који се намеће јесте узимање све три вредности и тражење њихове аритметичке средине. Дакле, просто сабирање и дељење са бројем вредности (у овом случају 3) би требало да обезбеди коректно прецизне резултате.

Ипак, више је проблема у овом приступу. Први на који наилазимо је одабир класа за које тражимо дистанце. Теоретски, најбоље би било пронаћи све дистанце за све класе кроз све алгоритме како бисмо имали пуну слику, али то је (непотребно) временски захтевна операција. Иако постоји вероватноћа да у крајњем скору "победи" класа која није била први одабир ниједног од алгоритама, ипак је процењено да је то довољно редак случај, па су једине класе које улазе у конкуренцију оне које алгоритми врате као први избор.

Дакле, први корак јесте одабир класа које ћемо детаљно обрадити. У најгорем случају то могу бити три класе, јер је могуће да сва три алгоритма врате различиту предикцију. Ипак, реална ситуација је таква да се обично то сведе на једну или две класе, али треба бити спреман и на најкомплекснији случај.

Након одабира подскупа класа које сучељавамо остаје да проценимо која од њих је највероватнија. За сваку од класа проналазимо све три дистанце (процене сваког алгоритма) и на основу њих одређујемо јединствену вредност коју даље можемо поредити.

Проблем на који наилазимо јесу такозвани outliner-и, тј. вредности које су далеко од "очекиваних", и које имају изузетно снажан утицај на рачунање обичне аритметичке средине. Рецимо, на фотографији са екстремном фацијалном експресијом LBPH алгоритам може вратити неку веома велику дистанцу и потпуно умањити чињеницу да остала два алгоритма могу донети добру одлуку.

Управо из тог разлога, за рачунање обједињене дистанце не користе се све три вредности, већ се бирају две најповољније, тј. најмање. Дакле, одбацујемо највећу од њих, и онда формирамо резултат помоћу преостале две.

Због чињенице да слични типови алгоритма враћају у општем случају сличне вредности, LBPH алгоритам може бити надгласан управо том неравноправношћу, од стране два алгоритма холистичког приступа. Управо са циљем превазилажења овог проблема, дистанца коју враћа LBPH алгоритам рачуна се двоструко, како би јој се повећао значај. На пример, ако приликом селекције два алгоритма који су вратили мању дистанцу остану резултати Eigenfaces и LBPH алгоритма, они неће бити узети са истим коефицијентом, већ ће LBPH дистанца бити удвостручена, и онда збир дистанци бити подељен са 3, уместо са 2. Ако је LBPH дистанца одбачена као outliner онда све остаје исто, преостале две вредности се уобичајено упросечују дељењем њиховог збира са 2.

Последњи проблем који је имао озбиљног негативног утицаја на квалитет резултата је чињеница да дистанце које алгоритми враћају нису равноправне.

Наиме, просечна вредност свих дистанци које враћа сваки од алгоритама међусобно се битно разликује. Управо из тог разлога је на већем тренинг сету фотографија најпре израчунат просечан резултат најмањих дистанци (вредности које су алгоритми вратили за ону класу за коју мисле да јој тест фотографија припада), а затим је нормализован. Дакле, пре коришћења било које од вредности које алгоритми враћају оне су помножене адекватним коефицијентима како би биле међусобно равноправне.

Имајући у виду да је проблем који софтвер третира, проблем препознавања лица, изузетно актуелна тема у свету алгоритама, процедура рачунања обједињеног скора класа направљена је тако да омогући евентуално касније убацивање додатних алгоритама, или искључивање неког од постојећих, тј. није фиксно везана за три дистанце.

Глава 4

Резултати

4.1 Тестирање на стандардним базама лица

Као што је раније наглашено, јако је тешко међусобно поредити алгоритме за препознавање лица, због различитости резултата на различитим базама.

У сврху анализе поузданости софтвера представљеног овим радом, коришћен је сет од 4 различите базе података др Либора Спацека [9], које су различите тежине и формиране су у различитим условима. Свака од њих садржи по 20 фотографија за сваку од особа (сваку класу).

Подела сваког од разматраних сетова података извршена је поштујући опште прихваћену праксу, тако што је две трећине информација искоришћено за тренинг сет, и једна трећина за тест сет. У конкретном случају, 13 фотографија сваке особе употребљено је за тренирање модела, а 7 фотографија за тестирање.

Прва база (faces94) спада у ред лаких сетова података. Садржи фотографије 153 особе у резолуцији 180×200 пиксела, фотографисане у идентичним условима. Позадина је једнобојна, нема варијација у величини и позицији лица на фотографијама уз присутност јако малих ротација. Неколико примерака фотографија из ове базе приказано је на Слици 4.1.

Осветљење фотографија такође не варира, једине компликације јесу одређене (врло ограничене) варијације у изразима лица. Такође, све фотографије направљене су у једној сесији, па су и фактори као што је различитост фризура искључени.

Овакви услови чине базу релативно једноставном за препознавање, што је очекивано довело до изузетно високог процента поузданости предложеног софтвера - преко 99%. Ипак, велика већина широко доступних алгоритама би дала на овој бази

Слика 4.1: Неколико фотографија из базе faces94

блиско квалитетне резултате, тако да је неопходно тестирати софтвер на тежим случајевима да би се добила права слика.

Следећа база (faces95) садржи фотографије 72 особе у истом формату (180×200 пиксела). Особе су током сесије фотографисања прилазиле један корак ка камери, што је довело до значајне варијације у величини лица (скалираности). Фотографије су прављене на сваких просечно пола секунде.

Позадина више није једнобојна као у faces94 бази, али није ни превише компликована. Ради се о завеси која стоји на фиксном месту, али сенке које особе праве крећући се стварају одређене проблеме софтверу за препознавање. И даље нема великих ротација лица, али је присутна озбиљна разлика у осветљености фотографија, имајући у виду да се особе крећу и да је додатно намерно мењано осветљење током рада. Слично као код претходне базе, мале су варијације у изразима лица, и све су фотографије направљене у једној сесији. Насумичне фотографије из базе faces95 приказане су на Слици 4.2.

Ипак, ни ова база не сматра се тешком, али се сматра средње захтевном, па се добијени скор софтвера од 99% сматра јако добрим резултатом.

Трећа база (faces96) спада у ред компликованијих тест база. У њој се налазе фотографије 152 особе, које су направљене испред различитих позадина са јако агресивним и значајним варијацијама (сјајни постери). Лица на фотографијама снимљена су у различитим позицијама (крећу се), различитих су величина, са мањим разликама у нагибу главе, итд. Осветљење доста варира, као и код претходне базе, с тим што су овде додате и одређене варијације фацијалних експресија субјеката, у мањој

Слика 4.2: Неколико фотографија из базе faces95

мери. На Слици 4.3 приказано је по неколико фотографија из две класе. Иако је база јако захтевна, предложени софтвер постигао је преко високих 97% прецизности.

Слика 4.3: Неколико фотографија из базе faces96

Последња тестирана база (grimace) такође спада у ред тешких, али по другом основу. За разлику од базе faces96, она нема агресивну позадину, нити великог осцилирања по питању осветљења, али су огромне разлике у експресијама особа. Особе на фотографијама праве широк спектар најразличитијих гримаса (смех, туга, затворене-/отворене очи, значајно нагнута глава, итд). Степен варијација изгледа на фотографијама могуће је видети на Слици 4.4. Захваљујући одличној претходној обради фотографија, свако лице је препознато (поузданост 100%), али треба узети у обзир да ова база садржи фотографије само 18 особа. Ипак, стопроцентна поузданост на овој тешкој бази свакако је значајан резултат.

Слика 4.4: Неколико фотографија из базе grimaces

4.2 Утицај смањења тренинг сета на поузданост алгоритма

Као што је и претпостављено, не можемо очекивати једнаку поузданост класификатора ако је трениран базом која садржи једну, две или тринаест фотографија по особи (класи). Наравно, значајно је теже забележити особине које су карактеристичне за једну класу ако имамо мали број фотографија које користимо. Ипак, то не значи ни да би тренинг сет од троцифреног броја фотографија по особи довео прецизност сваког модела на максималну.

Како би био утврђен утицај промене броја фотографија на поузданост софтвера, коришћене су различите поделе комплетног скупа података на тренинг сет и тест сет.

Број фотографија по класи који ове базе пружају (20) омогућио је квалитетно тестирање софтвера на максимално 13 примерака по класи тренинг сета, али је идеја утврдити колико нагло и у којој мери се мења прецизност класификатора са изменом дистрибуције података између тренинг и тест сета.

Проблем који настаје јесте начин на који алгоритми враћају своју процену поузданости. Наиме, изузетно је тешко осмислити квалитетно скалу од 0 до 1 када се узме у обзир да из алгоритма добијамо вредности за које само знамо да су веће од 0, без горње границе.

Како бисмо на неки начин усагласили скалу, за јединичну вредност дистанце узета је просечна дистанца погодака за најлакшу базу (faces94) са највећим тренинг сетом (13 фотографија по класи).

Комплетна статистика тестирања алгоритма на бази података faces94 приказана је у Табели 4.1.

Табела 4.1: Резултати тестирања на faces94 бази лица

	Број погодака	Број промашаја	Поузданост	Дистанца
1	2494	383	0.866875	2.452650
2	2492	233	0.914495	2.354526
3	2467	106	0.958803	2.214526
4	2355	67	0.972337	2.054526
5	2221	49	0.978414	1.924526
6	2076	44	0.979245	1.794526
7	1930	38	0.980691	1.689526
8	1789	28	0.984590	1.352817
9	1642	25	0.985003	1.238405
10	1493	21	0.986130	1.153852
11	1346	22	0.988252	1.135690
12	1198	13	0.989265	1.083159
13	1053	9	0.991525	1.000000

Статистика примене алгоритма на нешто компликованију базу (faces95) даје сличне резултате, као што приказује Табела 4.2.

Табела 4.2: Резултати тестирања на faces95 бази лица

	Број погодака	Број промашаја	Поузданост	Дистанца
1	1197	150	0.888641	2.758261
2	1136	140	0.890282	2.608936
3	1154	52	0.956882	2.460803
4	1089	45	0.960317	2.308441
5	1033	31	0.970865	1.994333
6	969	23	0.976814	1.790807
7	902	20	0.978308	1.66644
8	835	11	0.978898	1.540081
9	766	15	0.980794	1.385355
10	693	18	0.974684	1.237703
11	626	15	0.976599	1.254123
12	561	10	0.982487	1.240878
13	491	8	0.983968	1.127826

База података која је направила највеће проблеме софтверу услед великих варијација у осветљењу, скалирању и ротирању, као и компликованим позадинама (faces96) је и поред свега сасвим задовољавајуће савладана, као што се види у Табели 4.3.

Последња база коришћена за тестирање (grimaces) је, делимично захваљујући релативно малом броју класа, дала такође изванредне резултате. Из података приказаних у Табели 4.4 такође се може уочити да је кључна разлика тестирања на мањој количини тренинг података у односу на већу релативна дистанца. Наиме,

Табела 4.3: Резултати тестирања на faces96 бази лица

	Број погодака	Број промашаја	Поузданост	Дистанца
1	2249	604	0.788293	3.051672
2	2192	514	0.810051	2.936757
3	2412	139	0.945512	2.811274
4	2288	117	0.951351	2.695642
5	2152	99	0.956019	2.509219
6	2019	84	0.960057	2.384538
7	1881	72	0.963134	2.231274
8	1741	61	0.966148	2.046011
9	1599	52	0.968503	1.900115
10	1455	46	0.969353	1.734274
11	1312	39	0.971132	1.569005
12	1170	30	0.975000	1.412463
13	1028	23	0.978116	1.248340

као последица малог броја класа и поред чињенице да алгоритам са релативно малом поузданошћу предвиђа класе неког следећег лица, ипак јако ретко долази до грешака. Што је већи број класа, очекивано, већа дистанца од "праве" класе доноси већу вероватноћу да ће нека друга класа "збунити" алгоритам, јер би лице могло довољно да личи на њу, због већег избора могућих опција.

Табела 4.4: Резултати тестирања на grimaces бази лица

	Број погодака	Број промашаја	Поузданост	Дистанца
1	334	6	0.982353	7.582706
2	317	5	0.984472	7.017969
3	301	4	0.986885	6.393834
4	283	3	0.989510	5.861686
5	267	2	0.992565	5.319992
6	250	1	0.996016	4.704779
7	231	1	0.995689	3.958964
8	214	1	0.995349	3.407479
9	195	1	0.994898	2.662216
10	178	1	0.994413	2.13079
11	161	0	1.000000	1.658637
12	143	0	1.000000	1.567094
13	125	0	1.000000	1.492736

Ипак, треба напоменути да на ове неуспешно препознате примерке можемо додати и одређен број фотографија на којима није пронађено лице или нису пронађене очи. Наиме, као што је описано у Секцији 2.3, овај софтвер (иако то није случај са неким другим алгоритмима) захтева успешно проналажење и лица и оба ока на фотографији. Уколико нешто од тога не буде детектовано, систем одбија фотографију, и не класификује је уопште. Ти случајеви нису представљени у табелама, а њихов удео

варира између 0.5 и 2 процента у овим базама. У развоју овог софтвера, као приоритет је постављена елиминација погрешних процена, а не максимизовање броја погодака. То је уобичајена пракса за системе који су развијани пре свега у безбедносне сврхе. Ипак, треба имати у виду да за примену софтвера за препознавање лица у апликацијама забавног карактера ово није пожељна особина, јер се обично очекује да алгоритам врати било какав резултат, с обзиром да грешка нема велику цену.

Глава 5

Примена софтвера за препознавање лица

5.1 Примена на базу података TOC Search

Иако је софтвер иницијално писан за примену на бази тероризма организованог криминала TOC-search, касније у току рада установљено је да, на жалост, не постоје услови за висок степен поузданости система, користећи информације тренутно доступне у бази.

Детаљнијом анализом квалитета и броја фотографија у бази, дошли смо до закључка да није могуће адекватно применити предложени софтвер на фотографије смештене у бази. Штавише, вероватно је немогуће направити софтвер за препознавање лица који би био успешан у овом случају.

Чињеница да се у бази налазило око 2000 класа у датом тренутку и да је преко 99% њих имало само по једну фотографију чини тестирање овог софтвера практично немогућим. Додатни проблем прави чињеница да је велика већина тих фотографија веома слабог квалитета, мале резолуције и направљена у тешким условима. Јасно је да ни човек са изузетно развијеним осећајем за препознавање лица не може постићи иоле висок проценат погодака. Разлог из којег човек тако добро препознаје лица јесте што их довољно пута види. За само пар секунди лице се довољно квалитетно може меморисати, јер лице уживо даје доста квалитетнији скуп информација него једна фотографија. Очекивано, уколико располажемо само једном фотографијом јако тешко можемо правилно разликовати неколико хиљада класа.

Као закључак намеће се чињеница да је неопходно значајно подићи квалитет информација које база садржи, да би било могуће спровести било који алгоритам за препознавање лица, па тако и овај овде описан.

Још једна чињеница која значајно компликује примену софтвера на ову базу јесте "смер" коришћења резултата алгоритма. Наиме, природа примене базе TOC-search је таква да особа коју обрађујемо има за циљ да не буде пронађена у бази, па јој је у интересу да на што је могуће више начина омете поступак.

То је доста тежи случај од, рецимо, ситуације у којој особа да би била примљена у зграду или да би прошла гранични прелаз мора да буде препозната, тј. да учини све што је у њеној моћи да помогне алгоритму.

Једна могућа апликација била би прављење подбазе за специфичне потребе безбедности појединих служби, као што су војска, полиција, агенције за физичко-техничко обезбеђивање лица и објеката и томе слично. Таква база би подразумевала мањи скуп особа, што би резултирало мањим бројем могућих класа, и омогућило лакше прикупљање квалитетног сета података.

5.2 Примена у обезбеђивању објеката

Уколико специфична примена софтвера због своје природе може да обезбеди квалитетан улазни сет података, могуће је постићи прецизност врло блиску теоријском максимуму.

Ако узмемо као пример обезбеђење улаза у зграду неке компаније, служба која се бави сигурносном провером може за веома кратко време направити велики број фотографија особа које желе дозвољен пролазак кроз чувану тачку. Служба која се бави провером може осигурати висок квалитет улазних података правилним позиционирањем камере, тако да снима анфас фотографије лица која улазе у зграду, имајући у виду да се фотографије праве у сврху препознавања лица. Такође, на улазу у чувано место, особа која жели да уђе ће се потрудити да максимално квалитетно буде забележена од стране софтвера, неће покушавати да га наведе на лош траг тако што ће крити лице, носити тамне наочаре итд.

Слична ситуација би била на проласку кроз гранични прелаз. Службено лице на граничном прелазу има овлашћења којима може да обезбеди да особа буде квалитетно фотографисана, што би значило да је алгоритму задатак олакшан. Особи која пролази кроз граничну контролу није дозвољено да константно гледа у страну, крије лице или очи, и слично.

5.3 Примена у обезбеђивању државних граница

Позната је чињеница да све службе безбедности имају великих проблема са фалсификованим и туђим документима, што је нарочито изражен и критичан проблем управо при преласку државних граница.

Свакако, додатни метод идентификације помоћу предложеног софтвера за препознавање лица би бар могао донекле да ублажи тај проблем, по више основа. Најпре, ради се о додатном слоју провере информација, а нарочито велику примену би могао да оствари у комбинацији са подацима које база TOC-search поседује, што је и један од основних мотива развоја овог пројекта.

Наиме, један пример преваре коју та база у комбинацији са софтвером за препознавање лица може осујетити јесте откривање лажног документа који је особа добила нелегалним путем. У таквој ситуацији, постоји релативно висока вероватноћа да ће софтвер препознати лице које се налази у бази (због актуелних или ранијих преступа) и сугерисати особи која проверава идентитет на граничном прелазу да се ради о могућој превари.

Посебна погодност коришћења овог софтвера је чињеница да додатно време које се троши на овакав филтер не би било проблематично, и не би озбиљније угрозило проток људи преко границе.

Са друге стране, јавља се технички проблем у коришћењу овог система на граничним прелазима одређеног типа. Чињенично стање је да у тренутним техничким могућностима доста друмских граничних прелаза нема услове да провери сваку особу проласком кроз софтвер, јер то захтева фотографисање лица, што често није случај. Имајући то у виду, софтвер би највећу примену у актуелним техничким условима које границе нуде имао на аеродромима. У многим земљама света већ је развијена пракса да постоји сигурносна камера и да се захтева од лица које жели да пређе границу да стајањем на унапред предвиђено место омогући запис квалитетне анфас фотографије. Са таквим потенцијалима, без додатних проблема би било могуће искористити добијену информацију као улаз у софтвер за препознавање лица.

Иако би било доста компликованије, било би изводљиво да се слична провера обавља и на друмским границама. Ипак, неопходан предуслов јесте да службеник одвоји особе за које има неке сумње на страну, како би пред неком сигурносном камером били обезбеђени подаци уз помоћ којих би се извршила провера.

Ипак, како време буде одмицало и сваки сегмент друштва буде технолошки опремљенији, утицај овог софтвера биће све већи у разним сегментима живота, а не само

у области безбедности, захваљујући новим методама прикупљања информација од интереса - фотографија.

Глава 6

Закључак и правци даљег развоја

Иако је софтвер дао изузетне резултате на захтевним базама, планиран је наставак истраживања у циљу даљег унапређења.

Имајући у виду темпо којим се долази до нових открића у областима Computer Vision-a, као и природу идеје хибридизације, има основа да очекујемо напредак истог система инкорпорирајући неке нове, будуће алгоритме, или унапређења постојећих.

Негативан ефекат хибридизације је чињеница да је просторна и временска сложеност алгоритма сразмерно већа. Ипак, разлика у комплексности у односу на друге методе је линеарна, што у модерном рачунарском свету снажних хардвера не мора бити велика препрека.

Унапређење које би делимично превазишло овај проблем је већа конфигурабилност система. Конкретно, за различите примене требало би обезбедити различито време извршавања; уколико поузданост није најзначајнија, треба обезбедити кориснику могућност да балансира између квалитета и перформанси. Такође, за конкретну примену, може се открити да неки од алгоритама константно дају квалитетне резултате, па опција која би напредном кориснику била корисна јесте мењање интензитета утицаја алгоритама у рачунању обједињене вредности у кораку параметризације лица и рачунања дистанци, описаног у секцији 3.3.3.

Поред интеграције предложеног софтвера са базом TOC-search, планира се и интергација са неким другим професионалним базама података. Иако информације које даје предложени софтвер имају висок степен поузданости, тек комбинацијом са другим методама и унакрсном провером са подацима из професионалних база можемо добити комплетан, поуздан систем безбедности.

И поред чињенице да не може бити стопроцентно прецизан, софтвер за препознавање лица има широке примене у различитим сферама рачунарства, безбедности и забаве.

С обзиром на широк спектар могућих примена, идеја која је коришћена у развоју овог софтвера је прављење општег система, прилагодљивог конкретним потребама и независног од било које примене појединачно.

Кључни принцип који се мора поштовати јесте да је неопходна супервизија резултата од стране корисника, јер се не ради о егзактном решењу. Чак и човек са значајно вишим потенцијалима за препознавање особа може да погреши.

Ограничење за практичну примену овог, али и других софтвера за препознавање лица (у смислу поузданости резултата), су недовољно квалитетне реалне базе фотографија. То је проблем за сваки алгоритам са овом наменом, јер није реално очекивати да у бази од неколико хиљада различитих особа са само пар фотографија по особи софтвер има високу прецизност.

Намеће се закључак да, уколико се софтверу за препознавање лица дају довољно квалитетни подаци (које је чак и једноставније прибавити него улазне информације за препознавање отиска прста и, рецимо, зенице ока), резултати ће бити упоредиви са ова два изузетно поуздана система.

Ипак, предложени софтвер за препознавање лица представља користан алат за асистенцију у одлучивању, и може се користити у различитим областима, међу којима је најзначајнија област безбедности.

Библиографија

[1] S. Anila and N. Devarajan. Preprocessing technique for face recognition applications under varying illumination conditions. *Global Journal of Computer Science and Technology*, 12:No.11–F, 2012.

[2] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Trans. on PAMI*, 15(10):1042–1052, 1993.

[3] Intel Corporation. Open CV, 2012. [Online; приступљено 1. јун 2014].

[4] R. Gross and V. Brajovic. An image preprocessing algorithm for illumination invariant face recognition. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication*, page 10–18, 2003.

[5] I.T. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[6] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. *ICIP02*, 1:900–903, 2002.

[7] A. Martinez and A. Kak. Pca versus lda. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 23:228–233, 2001.

[8] M. Pietikainen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Springer, 2011.

[9] L. Spacek. Face recognition data, 2008. [Online; приступљено 1. јун 2014].

[10] Z. Stanimirovic and D. Trifunovic. Introducing the terrorist organized criminal search database TOC-s. *Terrorism and Political Violence*, 29(4):508–514, 2010.

[11] M.A. Turk and A.P. Pentland. Face recognition using eigenfaces. *Computer Vision and Pattern Recognition*, pages 586–591, 1991.

[12] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference*, 1:511–518, 2001.