

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Marko Škrjanec

**Avtomatsko razpoznavanje vrste
sadežev z metodami računalniškega
vida**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Matej Kristan

Ljubljana, 2013

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .



Št. naloge: 00081/2013

Datum: 08.04.2013

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **MARKO ŠKRJANEC**

Naslov: **AVTOMATSKO RAZPOZNAVANJE VRSTE SADEŽEV Z METODAMI
RAČUNALNIŠKEGA VIDA
COMPUTER VISION BASED AUTOMATIC FRUIT RECOGNITION**

Vrsta naloge: Diplomsko delo univerzitetnega študija prve stopnje

Tematika naloge:

V diplomski nalogi obdelajte problem avtomatskega razpoznavanja vrst sadežev na podlagi fotografij. Ta problem sodi v splošno domeno računalniškega vida, ki se ukvarja z razpoznavanjem objektov in je zaradi vizualnih lastnosti različnih vrst sadežev še posebno zanimiv. Sadeži iste vrste namreč pogosto izkazujejo veliko variabilnost znotraj vrste, pogosto pa tudi majhno variabilnost preko različnih vrst. V nalogi naredite pregled splošno uveljavljenih opisnikov in razpoznavalnikov za razpoznavanje objektov. Med njimi izberite najprimernejše za nalogo razpoznavanja sadežev. Opisnike in razpoznavalnike primerjajte na veliki zbirki različnih sadežev, podajte kritično analizo ustreznosti glede na performanco in predlagajte najprimernejšo kombinacijo opisnika in razpoznavalnika. Sistem za razpoznavanje ovrednotite v luči razpoznavanja, kakor tudi aplikacije priporočilnega sistema. V kolikor dovolj velika prosto dostopna zbirka slik sadežev ne obstaja, izdelajte lastno.

Mentor:

doc. dr. Matej Kristan

Dekan:

prof. dr. Nikolaj Zimic



IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Marko Škrjanec, z vpisno številko **63100304**, sem avtor diplomskega dela z naslovom:

Automatsko razpoznavanje vrste sadežev z metodami računalniškega vida

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Matjeja Kristana,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 3. septembra 2013

Podpis avtorja:

Zahvaljujem se mentorju doc. dr. Mateju Kristanu za njegovo obsežno pomoč, teoretično podlago in strokovne nasvete pri izdelavi diplomske naloge.

Zahvaljujem se še ostalim profesorjem in asistentom, ki so me učili v teku mojega študija, za vso pridobljeno znanje in navdih za nadaljnji študij računalništva.

Iskreno se zahvaljujem še družini in svojemu dekletu za podporo in pomoč pri izdelavi diplomske naloge.

Kazalo

Povzetek	1
Abstract	3
1 Uvod	5
1.1 Motivacija	5
1.2 Sorodna dela	6
1.3 Cilji in prispevki	10
1.4 Struktura diplomske naloge	11
2 Teoretična podlaga	13
2.1 Preprosti opisniki	14
2.1.1 Histogrami barvnih prostorov	14
2.1.2 Histogram orientacij gradientov	17
2.2 Globalni opisnik oponentnih barv	19
2.3 Vreča vizualnih besed	20
2.4 Klasifikatorji	23
2.4.1 Odločitveno drevo	24
2.4.2 Metoda k-najbližjih sosedov	25
2.4.3 Naključni gozdovi	25
2.4.4 Metoda podpornih vektorjev	26
2.4.5 Ocenjevanje uspešnosti	27
2.5 Arhitektura sistema	28

3	Eksperimentalna analiza	29
3.1	Zbirka slik	30
3.1.1	Zajem zbirke slik	30
3.1.2	Predstavitev zbirke slik	32
3.2	Implementacija sistema	36
3.3	Protokol analize	37
3.4	Rezultati	39
3.4.1	Vrednotenje klasifikacijskih metod	39
3.4.2	Vrednotenje opisnikov	42
3.4.3	Najuspešnejši klasifikacijski model	44
3.4.4	Priporočilni sistem	50
4	Sklepne ugotovitve	53
4.1	Smernice za nadaljnji razvoj	55

Seznam uporabljenih kratic in simbolov

- *BOW* - Bag of visual words (vreča vizualnih besed)
- *CA* - Classification accuracy (klasifikacijska točnost)
- *DWT* - Discrete wavelet transform (diskretna valčna transformacija)
- *HOG* - Histogram of oriented gradients (histogram orientacij gradientov)
- *HSV* - Hue, saturation, lightness color space (barvni prostor HSV)
- *HTTP* - Hypertext Transfer Protocol (protokol HTTP)
- *KNN* - K-nearest neighbors (metoda k-najbližjih sosedov)
- *POI* - Points of interest (značilne točke)
- *RF* - Random forests (klasifikator naključnih gozdov)
- *RGB* - Red, green, blue color space (barvni prostor RGB)
- *RG* - Red-green color space (barvni prostor RG)
- *SIFT* - Scale-invariant feature transform (opisnik SIFT)
- *SVM* - Support vector machine (metoda podpornih vektorjev)

Povzetek

Osrednja tema diplomske naloge je bila analiza primernosti različnih algoritmov računalniškega vida za problem razpoznavanja sadja. Sadje nudi zahtevno domeno za razpoznavanje zaradi svoje raznovrstnosti med sadeži istega razreda, podobnosti sadežev različnih razredov in samega števila različnih sadežev. Za uspešno razpoznavanje sadja je bilo potrebno slike opisati z dobrim atributnim zapisom. Informacije o barvi, teksturi, velikosti in obliki sadežev so bile zajete s pomočjo uveljavljenih opisnikov. Klasifikacija slik na podlagi atributnega zapisa pridobljenega s pomočjo teh opisnikov je potekala z že uveljavljenimi klasifikacijskimi metodami s področja strojnega učenja. Za uspešnost klasifikacijskih metod je bilo potrebno pridobiti veliko in dobro zbirko slik sadja. Ker takšna javno dostopna zbirka slik sadja ne obstoja, jo je bilo potrebno zajeti. Na podlagi analize rezultatov v diplomskem delu je bil zgrajen priporočilni sistem za razpoznavanje sadja, ki je na zahtevni zajeti zbirki slik dosegel kar 85% uspešnost.

Ključne besede: razpoznavanje objektov, barvni opisniki, vreča vizualnih besed, računalniški vid, strojno učenje, razpoznavanje sadja

Abstract

The topic of the thesis was analysing the suitability of different algorithms from computer vision for fruit recognition. Fruit provides a challenging domain for recognition due to its diversity of fruits of the same class, the similarity of fruits of different classes and the sheer number of different fruits. For an efficient fruit recognition it was necessary to describe the images with good descriptions. Color, texture, size and shape information of fruits were obtained through a variety of descriptors. The images were then classified based on the obtained descriptors with well known classification methods from machine learning. For a great performance of classification methods a good fruit image database was needed. Since such a publicly available image database does not exist, an image database had to be obtained. Based on the analysis in the thesis a fruit recognition system was built with an 85% success rate on the obtained image database.

Keywords: object recognition, color descriptors, bag of visual words, computer vision, machine learning, fruit recognition

Poglavje 1

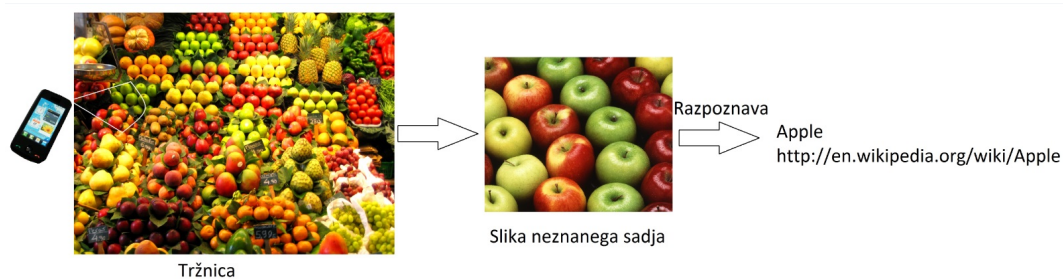
Uvod

1.1 Motivacija

Razpoznavanje objektov in scen [4, 5, 12, 18, 30, 32] je ena izmed tem računalniškega vida, ki se ukvarja z zajemanjem informacij iz slik in predstavitvijo teh informacij za uporabo v računalniških sistemih. Informacije o barvi in teksturi slike se opiše s pomočjo raznih opisnikov. Ti opisi slik so nujni za uspešno razpoznavanje objektov in scen s slik, ki je ključno za interakcijo računalniških sistemov z okoljem in človekom.

Na področju razpoznavanja objektov je v zadnjih letih postalo zanimivo razpoznavanje sadja. Sadje nudi zahtevno domeno za razpoznavanje zaradi svoje raznovrstnosti med sadeži istega razreda, podobnosti sadežev različnih razredov in samega števila različnih sadežev. S pomočjo rezultatov raziskav in analiz razpoznavanja sadja, je mogoče zgraditi priporočilne in druge specializirane sisteme [7, 34, 40]. Priporočilni sistem na podlagi slike objekta predlaga več možnih razredov tega objekta. Takšen preprost priporočilni sistem bi se lahko uporabljal za razpoznavanje neznanih sadežev na tržnici v neki eksotični državi. Uporabniki takšnega sistema bi lahko neznani sadež ali skupino sadežev preprosto slikali in sistem bi jim povedal, katero sadje je na sliki, ter jim podal podrobnejši opis oz. spletno povezavo do njega. Skica takšnega priporočilnega sistema za razpoznavanje sadja je razvidna na

Sliki 1.1. Razpoznavanje sadja je ključno v že razvitih specializiranih sistemih kot sta sistem za detekcijo sadežev na drevesu za robotsko obiranje sadežev [7] in sistem za štetje sadežev na drevesu [34]. Raziskave in analize metod za razpoznavanje sadja so ključne za implementacijo novih specializiranih aplikacij in sistemov za interakcijo z okoljem in človekom.



Slika 1.1: Skica priporočilnega sistema za razpoznavanje sadja.

1.2 Sorodna dela

Razpoznavanje scen in objektov je zadnja leta postala zelo aktualna tema na področju računalniškega vida. O tem pričajo številni članki na to temo, npr., [5,12,18,30,39,40]. Opisniki iz leta v leto postajajo zahtevnejši in seveda boljši. Najsodobnejši opisniki danes upoštevajo barvo, teksturo, obliko in izgled, med drugim so še neodvisni na spremembo pogleda in osvetlitve.

Raziskovalno delo [30] je eno izmed obsežnejših del na temo razpoznavanja scen in objektov. Podrobno opisanih in ovrednotenih je preko 10 opisnikov. Predstavljeni opisniki so vse od preprostih histogramov različnih barvnih prostorov do naprednih opisnikov, kot je recimo vreča vizualnih besed (angl. bag of visual words, BOW). Testiranje je izvedeno na več velikih zbirkah slik z veliko razredi za dokazilo uspešnosti in daje obsežen pregled nad problemom razpoznavanja objektov. Iz rezultatov tega dela so razvidne pomanjkljivosti posameznih opisnikov. Najslabše uspešnosti so dosegli opisniki, ki ne upoštevajo barvne informacije, ki je ključnega pomena pri

tem problemu. Kljub upoštevanju barvne informacije, histogrami barvnih prostorov niso pretirano uspešni zaradi neupoštevanja drugih informacij. Za opisovanje slik v tem članku predlagajo uporabo opisnika BOW, ki zajema informacije o teksturi in barvi.

V članku [39] je ovrednoten opisnik BOW pri uporabi za razpoznavanje scen. Razpoznavanje scen iz slik je podobno kot razpoznavanje objektov pomemben problem na področju računalniškega vida. Zadnje čase se razpoznavanje vrti okoli značilnih točk (angl. points of interest, POI), ki jih je možno zaznati z različnimi detektorji. Točke POI vsebujejo zelo informativno lokalno informacijo o sliki, ki jo je možno opisati z raznimi opisniki. Opisi točk POI se nato z uporabo algoritmov za gručenje združijo v vizualne besede. Opisnik BOW prešteje pojavitve posameznih vizualnih besed v sliki in te informacije se nato uporabijo pri klasifikaciji scen. Opisnik BOW prinaša zelo veliko uspešnost pri problemu razpoznavanja scen.

Avtorji dela [18] pokažejo, da zgodnja združitev informacij o obliki in barvi prispeva k znižanju uspešnosti najsodobnejšega opisnika BOW z prostorskim piramidnim ujemanjem. Detekcija objektov na njihovi zbirki slik, kjer je barva vitalnega pomena, se z njihovim algoritmom izboljša za 14%. Njihova zbirka slik je sestavljena iz risanih junakov bolj in manj znanih risank. Tako kot pri detekciji junakov, je tudi pri razpoznavanju sadja, barva zelo pomembna informacija. V članku je opisan postopek izračuna barvnega histograma orientacij gradientov (angl. histogram of oriented gradients, HOG). Barvni histogrami HOG se uporabljajo za opisovanje prostorskih celic vizualnih besed opisnika BOW. Boljša informativnost barvnega opisnika HOG pa temelji na zmanjšanju nedoločenosti manjših prostorskih celic pri opisovanju z ločenimi informacijami o barvi in obliki. Rezultati so pokazali boljše rezultate v prid barvnemu opisniku HOG, glede na tradicionalni opisnik HOG tudi na ostalih velikih in bolj znanih zbirkah slik, ki se uporabljajo na področju razpoznavanja objektov.

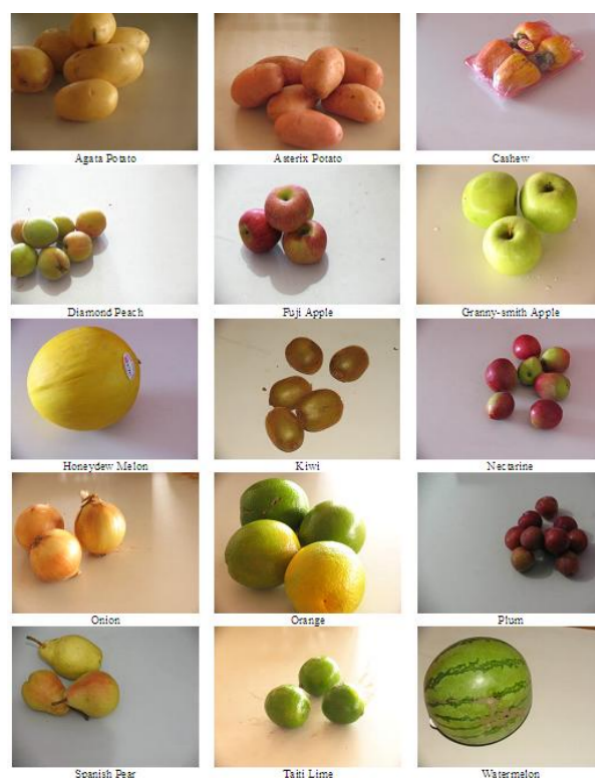
Uporaba opisnika HOG za detekcijo ljudi v slikah je opisana v članku [12]. Detekcija ljudi v slikah je zelo težek problem zaradi spremenljivega izgleda in

širokega razpona poz v katerih se človek lahko pojavi. V članku je predstavljeno, kako lahko z lokalno normaliziranim opisnikom HOG dobimo izredno dobre rezultate detekcije ljudi, kljub šumni okolici in težavni osvetljenosti. Zaradi preprostosti in hitrosti so za klasificiranje slik uporabili metodo podpornih vektorjev (angl. support vector machine, SVM), ki je trenutno ena izmed boljših metod za klasificiranje.

Problem razpoznavanja sadja je bil v zadnjih nekaj letih naslovljen v številnih člankih, npr., [4, 7, 32, 34, 40]. V teh člankih so uporabljeni marsikateri že uveljavljeni algoritmi za razpoznavanje objektov in scen.

V članku [4] je predstavljena preprosta rešitev za razpoznavanje sadja večih razredov. Iz slik se najprej odstrani šum okoli sadežev s pomočjo intenzitetne spremembe in morfoloških operacij. Slike so nato pretvorjene v barvni prostor HSV. Iz kanala V , ki predstavlja osvetlitev, se izluščijo informacije o teksturi s pomočjo diskretne valčne transformacije (angl. discrete wavelet transform, DWT) in sopojavljanja značilik kot so kontrast, energija, lokalna homogenost in gručna osvetlitev. Iz kanalov H in S se nato s pomočjo preprostih informacij o povprečju, standardni deviaciji, asimetriji ter sploščenosti barvnega histograma izluščijo barvne informacije. Sliki se nato s pomočjo opisnika, ki vsebuje vse te informacije, poišče razred najbližje slike testne množice po metodi minimalne razdalje. Testiranje in učenje so avtorji članka naredili nad množico 2600 slik, ki so bile razvrščene v 15 razredov. Avtorji so učinkovitost svoje implementacije dokazali z veliko zbirko slik, vendar so slike omejene na en do par sadežev slikanih na beli podlagi. Primeri slik te zbirke so prikazani na Sliki 1.2.

W. C. Seng in S. H. Mirisae [32] sta ena izmed prvih, ki sta opozorila na problematičnost opisnikov, ki slonijo samo na informacijah o barvi ali teksturi. Za izboljšanje natančnosti napovedi predlagata uporabo združenih informacij o barvi, obliki in velikosti sadežev. Informacije se izračunajo na podlagi obrezane slike sadeža s strani uporabnika. Na podlagi te slike se izračunajo povprečja barvnih kanalov slike. Oblika sadeža se izračuna s pomočjo segmentacije slike. Na podlagi izbranega območja se izračunajo ge-



Slika 1.2: Primeri zbirke slik sadja opisane v članku [4].

ometrijske lastnosti. V svojem članku opišeta, kako sta z uporabo vseh teh informacij in metode k-najbližjih sosedov (angl. k-nearest neighbors, KNN) dosegla do celo 10% zvišanje natančnosti napovedi.

V delu [34] je opisana detekcija in štetje sadežev na drevesu. Nad barvno transformiranimi slikami se, s pomočjo učenega klasifikatorja, poiščejo točke POI. Z drsečim oknom se nato algoritem sprehodi čez vse te točke in iz posameznega okna izlušči značilne točke in njihove opise. Končni opis regije okoli točk POI se zgradi s pomočjo vnaprej učenega modela BOW. Ta opis se nato vstavi v klasifikator SVM, ki pove, ali je v oknu prisoten sadež. Sistem je zgrajen na zbirki slik feferonov ter njegovih grmov in mišljen za štetje sadežev feferona na pripadajočem grmu.

V članku [40] je opisan algoritem za lokalizacijo jabolk na sliki za uporabo pri avtomatskem obiranju sadja. Raziskava je pokazala, da se lahko s pomočjo prisotnosti rdeče barve v sliki razlikujejo tako rdeča kot tudi zelena jabolka od preostalega drevesa. Teksturni detektor roba je v kombinaciji z meritvami rdeče barve v območjih prepoznanih, kot približnih krogih, dosegel odlične rezultate pri razpoznavanju jabolk na drevesu. Jabolka na drevesu je uspel sistem razpoznati iz neposredne bližine in tudi iz daljših razdalj. Zbirka slik v tem članku je zajemala le slike dreves z jabolki.

Za razvoj algoritmov računalniškega vida, je zelo pomembna obsežna zbirka slik. Problem pri razvoju algoritmov za razpoznavanje sadja je nedostopnost obsežne zbirke slik. Ustrezne javno dostopne zbirke slik sadja ne obstajajo in jih je potrebno za raziskovalno delo in uporabo najprej pridobiti.

Iz zgornjega pregleda je razvidno, da je bilo na področju računalniškega vida implementiranih in preizkušenih že kar nekaj rešitev. Problemi, ki jih te implementacije rešujejo, niso čisto enako zastavljeni, vendar pa jih je mogoče rešiti s podobnimi rešitvami. Iz letnic teh člankov [4, 18, 25, 30, 32, 34], ki se gibljejo med 2009 in 2013, je razvidno, da je delo na tem področju hitro razvijajoče in zelo aktualno.

1.3 Cilji in prispevki

Glavni cilj in osrednji prispevek diplomske naloge je bila podrobna analiza problema razpoznavanja sadja z algoritmi računalniškega vida. Za potrebe diplomske naloge so bili zbrani in implementirani razni opisniki slik, klasifikacijske in druge metode. Opisniki in klasifikacijske metode so bile izbrane na podlagi popularnosti in primernosti za domeno razpoznavanja sadja. Vsi ti opisniki in metode so bile ocenjene in analizirane glede na rezultate na zbirki slik. Drugi cilj diplomskega dela je bila izdelava ustrezne zbirke slik sadja. Izdelava zbirke slik je bila nujna, ker takšna, javno dostopna zbirka slik sadja, še ne obstaja. Pridobitev in javno dostopna objava zbirke slik je bil drugi prispevek diplomskega dela. Zbirka slik sadja vsebuje približno

1000 slik sadežev. Sadeži pa so razdeljeni v 30 različnih razredov. Izdelava ustrezne zbirke slik sadja in analiza problema in metod za razpoznavanje sadja sta torej bila glavna strokovna prispevka diplomske naloge.

1.4 Struktura diplomske naloge

Preostanek diplomske naloge je razdeljen na tri poglavja. V Poglavju 2 je opisana teoretična podlaga diplomske naloge. Podrobno so opisani vsi opisniki, klasifikacijske in ostale metode, ki so bile uporabljene pri eksperimentalnem delu. Opisani so opisniki od preprostih histogramov različnih barvnih prostorov, do opisnika HOG in BOW. Opisnik BOW je trenutno eden izmed naj sodobnejših in najučinkovitejših opisnikov na tem področju. Poleg opisnikov so opisane še uporabljene metode za klasifikacijo in metode merjenja uspešnosti le teh. Sama arhitektura implementiranega sistema je predstavljena na koncu tega poglavja.

Poglavje 3 opisuje celoten eksperimentalni del in rezultate analize. V tem poglavju je opisan postopek zajema zbirke slik, ki je bila nujna za nadaljnje delo. Zbirka slik sadja je na kratko tudi predstavljena. V drugem delu Poglavja 3 je zajet sam preizkus posameznih opisnikov in metod za klasifikacijo pri razpoznavanju sadja. Poglavje se zaključuje z analizo rezultatov opisnikov in klasifikacijskih metod.

Povzetek izsledkov diplomskega dela sledi v Poglavju 4. V tem poglavju tudi nakažemo smernice za nadaljnje delo.

Poglavje 2

Teoretična podlaga

Teoretična podlaga, ki je bila opisana v tem poglavju, je nujna za razumevanje diplomskega dela. Posamezni opisniki in klasifikatorji so bili izbrani, ker so se odlično izkazali že v opisanih člankih v Poglavju 1.2.

Histogrami barvnih prostorov so preprosti opisniki, ki se uporabljajo za opisovanje celotnih slik in posameznih regij okoli značilnih točk. Preizkušenih barvnih prostorov je bilo precej. Vsak barvni prostor ima svoje značilnosti in je primeren za opisovanje različnih slik oz. za druge probleme. Opisan je tudi histogram orientacij gradientov (angl. histogram of oriented gradients, HOG), ki ne uporablja informacij o barvi, ki je sicer ključnega pomena pri razpoznavanju sadja. Vreča vizualnih besed (angl. bag of visual words, BOW) odpravlja težave barvnih histogramov in opisnika HOG z uporabo vseh bistvenih informacij s slik.

Strojno učenje pozna že precej klasifikacijskih metod, mnoge od teh se tudi uporabljajo pri prepoznavanju predmetov. Klasifikatorji, kot so metoda k-najbližjih sosedov (angl. k-nearest neighbors, KNN), klasifikator naključnih gozdov (angl. random forests, RF) in klasifikator minimalne razdalje se večkrat uporabljajo, kot je razvidno iz sorodnega dela, npr., [4, 32]. Trenutno najbolj uporabljena je metoda podpornih vektorjev (angl. support vector machine, SVM), ki je uporabljena prav zaradi robustnosti in možnosti izboljšav z različnimi jedri, npr., [12, 30, 34, 39].

2.1 Preprosti opisniki

I. Kononenko je v svojem delu [19] zapisal, da je atributni zapis oz. opis slike eden ključnih elementov pri problemu klasificiranja slik. Za uspešno razpoznavanje sadja je bilo potrebno pridobiti kvalitetne opise slik s pomočjo uveljavljenih opisnikov.

2.1.1 Histogrami barvnih prostorov

Histogram intenzitet h slike I podaja informacijo o frekvencah pojavitve posameznih intenzitetnih stopenj v tej sliki. V matematičnem zapisu je histogram enodimenzionalen vektor dolžine števila stopenj. Histogram nam daje neko globalno informacijo o sliki [27]. V naši implementaciji je histogram normaliziran še s številom vseh slikovnih pik. Posamezna celica v histogramu nam torej pove verjetnost pojavitve intenzitetne stopnje v naši sliki.

Implementacija histograma vzeta po [27] z dodanim korakom normalizacije je opisana v Algoritmu 2.1. Rezultat histograma slike [35] je prikazan na Sliki 2.1.

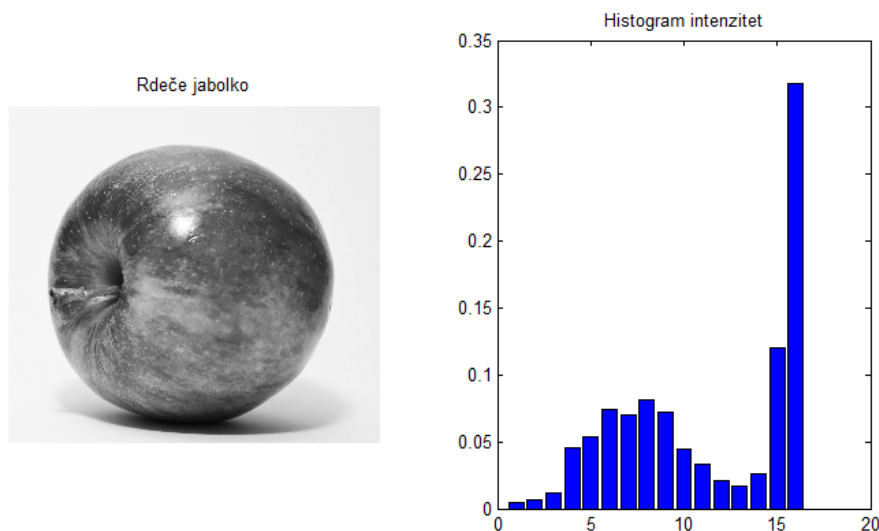
Vhod: Slika I

Izhod: Histogram h

- 1: Napolni vektor h dolžine števila stopenj z nulami.
- 2: Za vsak slikovni pik (x,y) sivinske slike I , povečaj $h(I(x,y))$ za 1.
- 3: Normaliziraj histogram h z vsoto le tega.

Algoritem 2.1: Izračun histograma intenzitet.

Barvni histogram predstavlja porazdeljenost barv v sliki. Barvni histogram se izračuna na slikah podanih v različnih barvnih prostorih [1]. V tem poglavju so predstavljeni histogrami barvnih prostorov RGB, HSV in RG. Poleg teh treh barvnih prostorov sta še predstavljeni histogram barvnega prostora Lab in histogram barvnega prostora oponentnih barv.



Slika 2.1: Slika in njej pripadajoči histogram intenzitet. [35]

V naši implementaciji se barvni histogram zgradi na podlagi enodimenzionalnih histogramov kanalov barvnega prostora. Za vsak kanal barvnega prostora posebej se zgradi enodimenzionalen histogram po Algoritmu 2.1, ki se nato zlepijo eden za drugega. Barvni histogram je torej vektor histogramov posameznih kanalov barvnega prostora in prikazuje porazdelitev intenzitet za vsak barvni kanal posebej.

Histogram barvnega prostora RGB je kombinacija histogramov rdečega R , zelenega G in modrega kanala B . Dandanes so slike največkrat podane v barvnem prostoru RGB, zato je grajenje tega histograma mogoče brez transformacije slike.

Histogram barvnega prostora HSV je kombinacija histogramov barvnega odtenka H , intenzivnosti S in vrednosti V . Barvni prostor HSV je zanimiv zaradi bolj intuitivne predstavitve, ki je posledica preureditve barvnega prostora RGB. Prostor HSV je zanimiv še zaradi neodvisnosti barvnega kanala H na spremembo svetlobe, kot je to opisano v članku [30]. Velik problem tega barvnega prostora je nestabilnost barvnega odtenka blizu sive osi.

Histogram barvnega prostora RG. Normalizacija barvnega prostora RGB za izračun kanalov r , g in b barvnega prostora RG je prikazana v enačbi (2.1). V normaliziranem barvnem prostoru RG sta dovolj le rdeča r in zelena komponenta g za zajetje celotne informacije. Modra komponenta b je odveč po formuli $r+g+b=1$. Histogram prostora RG je torej kombinacija histogramov kanala r in g tega prostora. Zaradi normalizacije sta kanala r in g neodvisna glede na spremembo svetlobe, prisotnost senc in senčenja [30].

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} \frac{R}{R+G+B} \\ \frac{G}{R+G+B} \\ \frac{B}{R+G+B} \end{pmatrix} \quad (2.1)$$

Histogram barvnega prostora Lab. Barvni prostor Lab je bil zasnovan za aproksimiranje človeškega zaznavanja barv. Barvni prostor Lab je homogen, ker so si človeku podobne barve, ki so predstavljene v tem barvnem prostoru, bližje po evklidski razdalji. Kanal L tega prostora se zelo dobro ujema s človeško predstavo svetlosti, kanala a in b pa vsebujeta informacijo o barvi [3, 29]. Histogram barvnega prostora Lab je kombinacija histogramov njegovih kanalov.

Histogram barvnega prostora oponentnih barv. Barvni prostor oponentnih barv opisan v delu [30] se izračuna na podlagi enačbe (2.2). Informacija o intenziteti je predstavljena s kanalom O_3 , informacija o barvah pa z kanaloma O_1 in O_2 . Slednja kanala sta neodvisna glede na spremembo svetlobe, zaradi izničevanja osvetlitve kanalov pri izračunu teh dveh kanalov. Histogram barvnega prostora oponentnih barv je kombinacija histogramov kanalov O_1 , O_2 in O_3 tega prostora.

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix} \quad (2.2)$$

2.1.2 Histogram orientacij gradientov

Barvna informacija je sicer ključnega pomena pri razpoznavanju sadja, vendar intuitivno gledano ni dovolj za doseg velike uspešnosti pri razpoznavanju sadja. Za opis slik so potrebne še informacije o teksturi, obliki in izgledu. N. Dalal in B. Triggs sta v svojem članku [12] zapisala, da je obliko in izgled objekta v sliki mogoče zelo dobro opisati z distribucijo orientacij gradientov. To je tudi glavna ideja histograma orientacij gradientov (angl. histogram of oriented gradients, HOG). Klasični opisnik HOG se gradi na sivinski sliki in zato ne vsebuje informacije o barvi.

Gradient je vektor, ki v neki točki slike kaže smeri največje intenzitetsne spremembe. Orientacija gradienta kaže dejansko smer roba, medtem ko magnituda gradienta ponazarja moč tega roba.

Opisnik HOG se zgradi po naslednjih korakih. Na začetku se vhodna sivinska slika razdeli na manjše celice. Za vsako celico se zgradi enodimenzionalen histogram orientacij gradientov, histogram vsebuje toliko celic kot je število opazovanih orientacij. Preko celotne slike se v naslednjem koraku izračunajo gradienti slike. Za vsako celico se nato preštejejo glasovi orientacij, ki se utežijo s pripadajočo magnitudo. Za boljšo neodvisnost od osvetlitve se histogrami celic normalizirajo. Normalizacija histogramov poteka glede na energijo posameznih histogramov celic preko večjih regij. Na koncu se posamezni histogrami celic zaporedno zlepijo v opisnik HOG. Algoritem za izračun opisnika HOG, ki je še bolj podrobno opisan v članku [12], je prikazan v Algoritmu 2.2. Rezultat prvih treh korakov algoritma je razviden na Sliki 2.2.

Vhod: Sivinska slika

Izhod: Opisnik HOG

- 1: Razdeli vhodno sivinsko sliko na manjše celice (npr. 8 x 8)
- 2: Za vse celice pripravi prazen vektor dolžine števila orientacij (npr. 8)
- 3: Izračunaj gradiente slike
- 4: Za vsako celico preštej glasove orientacij in jih uteži s pripadajočo magnitudo
- 5: Normaliziraj histograme celic iz prejšnjega koraka glede na energijo teh histogramov preko večjih regij
- 6: Zaporedno zlepi posamezne histograme celic v opisnik HOG

Algoritem 2.2: Gradnja opisnika HOG

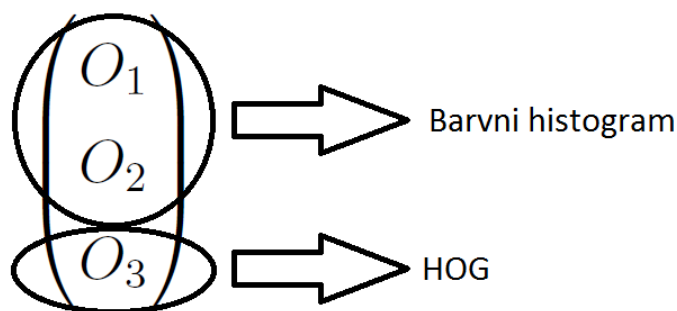


Slika 2.2: Slika po tretjem koraku Algoritma 2.2. [8]

2.2 Globalni opisnik oponentnih barv

Do sedaj opisani preprosti opisniki v Poglavlju 2.1, so zajeli le informacijo o barvi ali obliki ter izgledu. Globalni opisnik oponentnih barv pa zajema vso to globalno informacijo o barvi, obliki in izgledu. Barvni prostor oponentnih barv, kot je opisano v Poglavlju 2.1.1, vsebuje poleg informacije o barvi še informacijo o intenziteti slike v kanalu O_3 . Barvna kanala O_1 in O_2 sta neodvisna na spremembo svetlobe. S pomočjo barvnega histograma slednjih kanalov, lahko iz slike pridobimo od osvetlitve neodvisno informacijo o barvah.

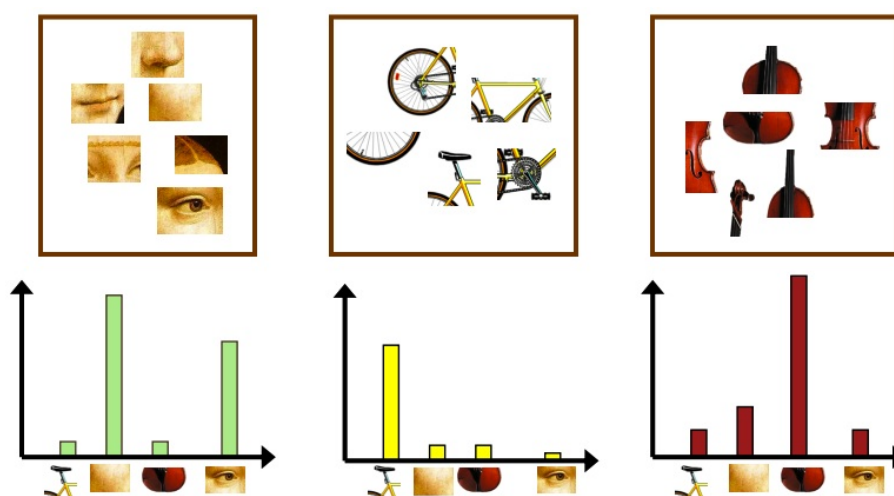
Iz kanala O_3 lahko zajamemo informacijo o obliki in izgledu sadežev s pomočjo opisnika HOG, ki je opisan v Poglavlju 2.1.2. Zajeti opisnik HOG in barvni histogram dveh kanalov nato zlepimo zaporedno v globalni opisnik oponentnih barv. Intuitivno smo tako s pomočjo barvnega histograma kanalov O_1 ter O_2 in opisnika HOG iz slike izluščili informacijo o barvi in obliki sadeža. Teoretično bi ta združena informacija mogla prinesiti večjo uspešnost opisnika pri klasifikaciji.



Slika 2.3: Globalni opisnik oponentnih barv.

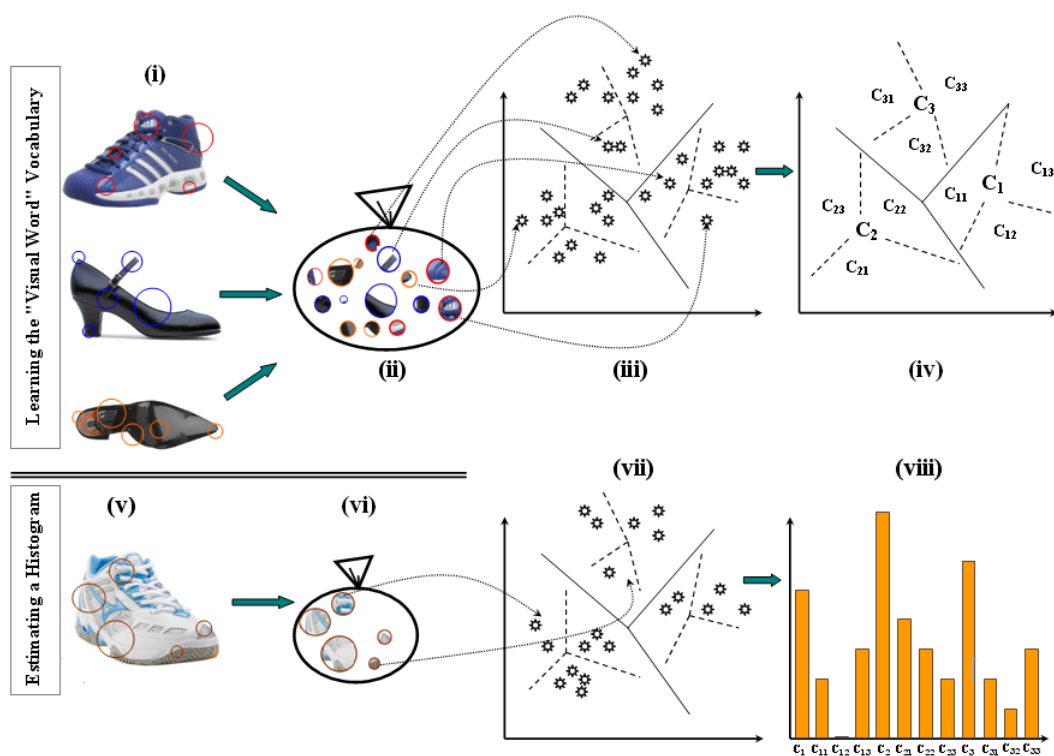
2.3 Vreča vizualnih besed

Do sedaj smo slike opisovali kot celoto, kjer smo posamezne histograme izračunali na podlagi barvnega kanala slike. Pri vreči vizualnih besed (angl. bag of visual words, BOW) je koncept opisovanja slik precej drugačen. Podobno kot besede v tekstovnih dokumentih imajo tudi slike značilne točke, katerih okolica vsebuje veliko uporabnih informacij. Značilne točke z informativno okolico imenujemo vizualne besede. Prav zaradi teh značilnih točk lahko slike zelo uspešno klasificiramo s histogramom frekvenc pojavitve posameznih vizualnih besed [39]. Opisnik BOW je označen kot trenutno eden izmed najsodobnejših opisnikov za razpoznavanje objektov in scen. Še en razlog za uspešnost opisnika BOW pa je uporaba informacije o teksturi in barvi slike. Ta združena informacija prispeva občutno zvišanje uspešnosti razpoznavanja. Število vizualnih besed opisnika BOW se navadno giblje okoli več 1000. Poenostavljen primer histograma BOW je prikazan na Sliki 2.4.



Slika 2.4: Slika prikazuje poenostavljene histograme BOW. [13]

Opisnik BOW opiše slike s histogramom pojavitve posameznih vizualnih besed v sliki. Algoritem je razdeljen v dve fazi, kar je razvidno na Sliki 2.5. V prvi fazi se izračuna slovar terminov, ki je nujen za opisovanje slik. Za izgradnjo slovarja terminov so potrebne slike iz katerih pridobimo značilne točke in opise le teh. V naši analizi se je 30 slik izkazalo optimalno za izgradnjo slovarja. Opise teh značilnih točk nato razporedimo po visokodimenzionalnem prostoru in jih s pomočjo gručenja podobnih točk strnemo v vizualne besede. Ta visokodimenzionalni prostor z vizualnimi besedami imenujemo slovar terminov. V fazi grajenja histograma vizualnih besed najprej iz slike izluščimo značilne točke in opise le teh. Za posamezne točke poiščemo najbližje vizualne besede in preštejemo pojavitve teh besed v sliki. Pojavitve posameznih vizualnih besed predstavlja opisnik BOW [23, 37]. Algoritem za izračun opisnika BOW je prikazan v Algoritmu 2.3.



Slika 2.5: Slika prikazuje obe fazi opisnika BOW. [36]

Vhod: Učna množica slik, testna slika

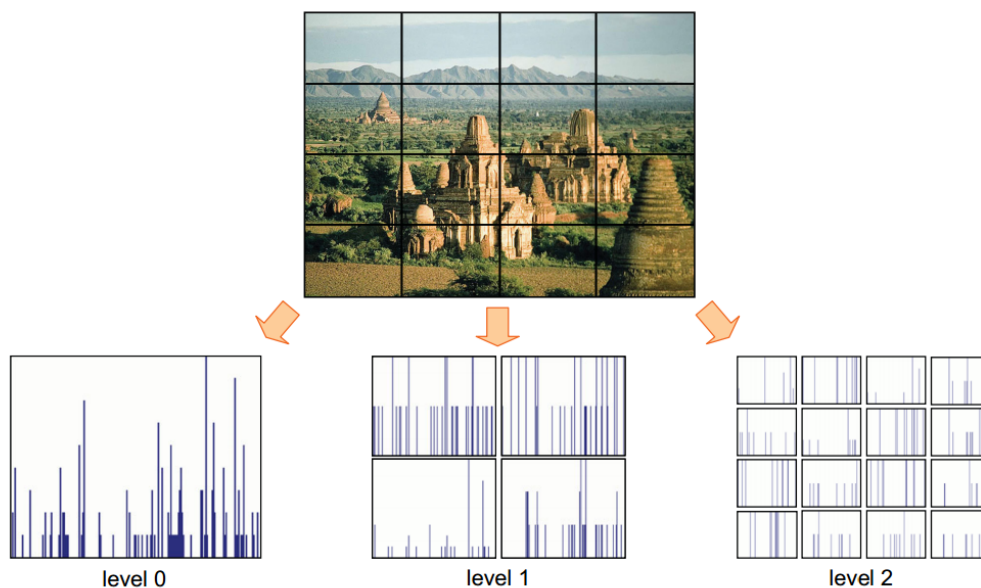
Izhod: Opisnik BOW

- 1: Detekcija in reprezentacija vizualnih besed iz podmnožice učnih slik.
- 2: Grajenje slovarja terminov z gručenjem podobnih vizualnih besed.
- 3: Prepoznavanje vizualnih besed iz slik s pomočjo slovarja terminov.
- 4: Izgradnja histograma frekvenc vizualnih besed.

Algoritem 2.3: Gradnja opisnika BOW.

Opisnik BOW se uspešno uporablja za različne probleme prav zaradi svoje robustnosti. Na voljo je veliko načinov za opisovanje in izbor značilnih točk. Najpogosteje se uporabljajo različne implementacije opisnika SIFT [24] ter opisnika HOG [18]. Najstarejša implementacija opisnika SIFT, ki jo je predlagal D. G. Lowe [24], podobno kot opisnik HOG šteje orientacije gradientov v posameznih celicah slike.

Zadnje čase se največ uporablja izboljšana implementacija opisnika BOW z prostorskim piramidnim ujemanjem (angl. Spatial Pyramid Matching) [21, 22]. Ideja tega opisnika je omiliti probleme pri opisovanju slik različnih skal s klasičnim opisnikom BOW. Pri tem opisniku se slika razdeli na več celic preko različnih skal, kot je razvidno na Sliki 2.6. Histogram na nivoju nič, je enak klasičnemu histogramu BOW. Na nivoju ena se slika razdeli na štiri celice in vse celice se opišejo z histogrami BOW, medtem ko je na nivoju dva teh histogramov že 16. Te, tako imenovane prostorske piramide, se med sabo primerja in uteži s piramidnim jedrom. To jedro v končnem opisu normalizira in bolj uteži histograme na višjih nivojih, ker ti bolj opisujejo lokalne informacije o sliki. V našem implementiranem sistemu je bila uporabljena ta izboljšana različica opisnika BOW.



Slika 2.6: Opisnik BOW s prostorskim piramidnim ujemanjem. [21]

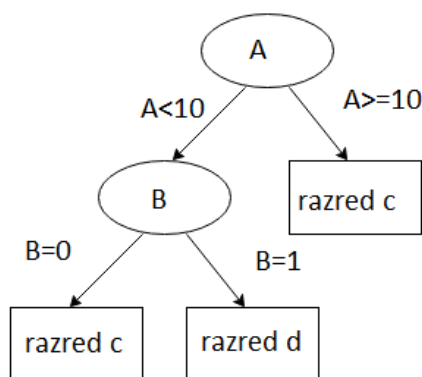
2.4 Klasifikatorji

Napovedovanje pripadnosti razredov oz. klasifikacija slik je poslednji korak pri razpoznavanju objektov. Klasifikatorji v fazi učenja zgradijo model, ki se nato v fazi testiranja uporabi za določanje razreda testne slike. Uspešnost samega napovedovanja je odvisno od opisnikov in samih klasifikacijskih metod. V tem poglavju so opisane le najpopularnejše klasifikacijske metode za razpoznavanje objektov in scen.

2.4.1 Odločitveno drevo

Odločitveno drevo [2] je eno izmed osnovnih metod klasifikacije v strojnem učenju. Odločitveno drevo je sestavljeno iz medsebojno povezanih vozlišč, ki predstavljajo attribute. Posamezne veje vozlišč pa predstavljajo podmnožice vrednosti atributov, ki jih ti lahko zavzamejo.

V fazi učenja se zgradi odločitveno drevo z vsemi vozlišči in pravili. Atributi v vozliščih se izbirajo glede na oceno informativnosti in ustreznih podmnožic njihovih vrednosti. Za izračun informativnosti atributa se uporabljajo mere kot so informacijski prispevek, J-ocena, ReliefF ipd. Bistvo teh mer je ocena primernosti atributa za uporabo v vozlišču. Sama klasifikacija testnih primerov pa poteka s sledenjem vej do končnega vozlišča, ki vsebuje napovedani razred [19]. Odločitvena drevesa so zanimiva zaradi svoje preprostosti in možnosti predstavitve v človeku razumljivemu zapisu. Primer preprostega odločitvenega drevesa je prikazan na Sliki 2.7.



Slika 2.7: Preprosto odločitveno drevo.

2.4.2 Metoda k-najbližjih sosedov

Metoda k-najbližjih sosedov (angl. K-nearest neighbors, KNN) [11,19] je zelo enostavna za razumevanje in preprosta za implementiranje. Model metode KNN zgrajen v fazi učenja predstavlja kar matrika nespremenjenih učnih podatkov. Napoved metode novemu primeru pa je večinski razred izmed k najbližjih učnih primerov. Uspešnost klasifikatorja KNN je v veliki meri odvisna od uporabljenega algoritma za izračun razdalje oz. podobnosti med primeri. Razdalja med dvema primeroma se navadno izračuna kot evklidska razdalja (2.3) med vrednostmi atributov. Časovna zahtevnost klasificiranja novega primera je velika v primerjavi z drugimi klasifikacijskimi metodami. Kljub veliki časovni zahtevnosti pa je hitrost grajenja modela in klasificiranja manjša kot pri drugih metodah, kadar je število učnih primerov majhno.

$$D(u, v) = \sqrt{\sum_{i=1}^a d(u_i, v_i)^2} \quad (2.3)$$

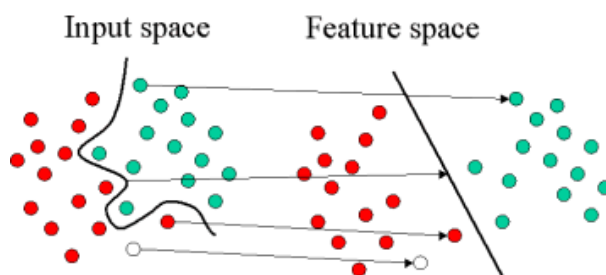
2.4.3 Naključni gozdovi

Metoda naključnih gozdov (angl. random forests, RF) [10,19] je izpeljanka odločitvenih dreves. Pri metodi RF se v fazi učenja zgradi več 100 različnih odločitvenih dreves. Vsako drevo je zgrajeno z naključno izbranimi atributi v vozliščih in ne z najbolj informativnimi, kot pri klasičnih odločitvenih drevesih. V fazi testiranja vsa drevesa glasujejo za razred testnega primera, izbere pa se razred največkrat glasovanega. Ta metoda deluje izredno dobro in je primerljiva z najboljšimi klasifikacijskimi metodami. Zaradi ogromnega števila odločitvenih dreves je razlaga in predstavitev tega modela nemogoča, vendar pa predstavitev modela v našem primeru ni bistvena.

2.4.4 Metoda podpornih vektorjev

Metoda podpornih vektorjev (angl. support vector machine, SVM) je ena izmed v praksi najuspešnejših klasifikacijskih metod [19] in je primerna za probleme z veliko primeri in atributi. Metoda SVM transformira osnovni atributni prostor v kompleksnejši atributni prostor, ki je bolj primeren za razdvojitev razredov s pomočjo hiperravnine. Optimalna hiperravnina je tista, ki maksimizira razdaljo med razredoma in je hkrati enako oddaljena od najbližjih primerov obeh razredov. Najbližje primere obeh razredov imenujemo podporni vektorji. Za transformiranje prostora je na voljo več transformacij, ki se razlikujejo po kompleksnosti. Transformacije prostorov se vršijo z jedri za transformacijo prostora. Po kompleksni transformaciji prostora s pomočjo kompleksnega jedra za transformacijo je razrede mogoče ločiti z linearno funkcijo. Preprosta transformacija atributnega prostora je prikazana na Sliki 2.8.

Metoda SVM rešuje dvorazredne probleme. V primeru problema z večimi razredi, je potrebno vsak razred posebej ločiti od ostalih razredov s pomočjo transformacij prostorov in postavljanjem hiperravnin. Za klasificiranje novega primera je potrebno primer postaviti v transformirani atributni prostor in mu nato na podlagi podpornih vektorjev določiti pripadnost razredu [19].



Slika 2.8: Slika povzeta iz [33] prikazuje transformacijo prostora SVM.

Za boljše rezultate metode SVM je potrebno pred uporabo podatke usrediščiti in normalizirati (2.4). Posamezne vrednosti x_{ik} matrike X je potrebno usrediščiti s povprečjem atributa μ_k in normalizirati s standardno deviacijo

atributa σ_k . Tako pripravljene podatke primerov x'_{ik} so veliko bolj primerni za klasifikacijo s pomočjo metode SVM.

$$x'_{ik} = \left(\frac{x_{ik} - \mu_k}{\sigma_k} \right) \quad (2.4)$$

2.4.5 Ocenjevanje uspešnosti

Poznamo več načinov za izračun uspešnosti modelov, za naš problem je zanimiva klasifikacijska točnost.

Klasifikacijska točnost se uporablja v problemih, kjer je posamezen primer enolično določen z razredom [19]. V našem problemu lahko slika pripada samo enemu razredu, zato je to primarna mera za ocenjevanje uspešnosti. Klasifikacijska točnost (angl. classification accuracy, CA) je količnik števila pravilno napovedanih primerov N_p in števila vseh primerov N . Enačba (2.5) prikazuje izračun klasifikacijske točnosti.

$$CA = \frac{N_p}{N} \quad (2.5)$$

Uspešnost priporočilnega sistema, ki nam kot rezultat poda več napovedi, se je izračunala kot klasifikacijska točnost vseh napovedi. Uspešnost priporočilnega sistema nam torej pove delež prisotnih pravih napovedi med vsemi napovedmi.

Spodnjo mejo uspešnosti predstavlja uspešnost klasifikacije s pomočjo večinskega razreda. V primeru, da je uspešnost metode manjša od uspešnosti ugibanja za večinski razred, moramo to metodo zavreči zaradi preslabe uspešnosti.

2.5 Arhitektura sistema

Arhitektura sistema za evaluacijo različnih kombinacij opisnikov in klasifikacijskih metod je bila zamišljena kot cevovod. Cevovod je zelo enostaven za uporabo ter razumevanje in omogoča preprosto prehajanje med fazami. Slike, podatki in modeli se lahko ne glede na izbran algoritem prenesejo naprej za uporabo v naslednji fazi.

Cevovod sistema je razdeljen v sedem faz. V prvi fazi se na podlagi izbrane zbirke slik preberejo slike. Slike je možno tudi ročno obrezati ali pa nastaviti avtomatsko obrezovanje. V drugi fazi se slike opišejo z izbranim opisnikom. Ti opisi se nato shranijo za kasnejšo uporabo. Shranjevanje opisov je nujno za skrajšanje časa pri vnovični uporabi istega opisnika. V naslednji fazi se slike razdelijo na učno in testno množico glede na izbrano razmerje. Neodvisnost učne in testne množice je zelo pomembna, zato se razdelitev naredi še pred fazo učenja modela. V fazi učenja se na podlagi testne množice zgradi izbrani klasifikacijski model. Ta klasifikacijski model se nato v naslednji fazi uporabi za klasifikacijo razredov testne množice ali pa za napovedovanje priporočilnega sistema. Na koncu se še prikažejo podrobni rezultati klasifikacije ali napovedovanja in uspešnost modela. Cevovod sistema je prikazan v Algoritmu 2.4.

- 1: Branje slik z možnostjo obrezovanja izbrane zbirke slik.
- 2: Opisovanje slik z izbranim opisnikom.
- 3: Shranjevanje opisnikov za kasnejšo uporabo.
- 4: Deljenje slik na učno in testno množico glede na izbrano razmerje.
- 5: Grajenje izbranega klasifikacijskega modela iz učne množice.
- 6: Uporaba klasifikacijskega modela za napoved razredov testne množice.
- 7: Prikaz rezultatov klasifikacije in izračun uspešnosti modela.

Algoritem 2.4: Cevovod implementiranega sistema.

Poglavje 3

Eksperimentalna analiza

V tem poglavju je opisano začetno zajemanje zbirke slik in vse do poteka analize in predstavitve rezultatov.

Pregled literature je pokazal, da uporabna javno dostopna zbirka slik sadja še ne obstaja. Zaradi tega je bilo potrebno zbirko slik zajeti. Zajemanje dobre zbirke slik se je izkazalo za precejšen problem. Samo pisanje skripte za zajem zbirke sicer ni bilo težavno, vendar pa se je tekom samega zajema zbirke pojavilo veliko problemov. Po odpravi teh problemov je bilo možno zajeti veliko in kvalitetno zbirko slik. Ta zbirka je tudi predstavljena in ovrednotena v naslednjih poglavjih.

Zbirka je bila nato opisana z različnimi opisniki. Zgrajeni so bili modeli, ki so testnim slikam določili razrede. Na podlagi uspešnosti in sklepanja iz slik so v tem poglavju ovrednoteni posamezni opisniki in metode. Predstavljena je implementacija priporočilnega sistema, ki na podlagi slike predlaga pet najverjetnejših razredov sadja. Kvalitativna analiza razredov sadja in uspešnosti algoritmov je predstavljena na koncu tega poglavja.

Vsa programska koda, razen skripte za zajem zbirke slik, je bila napisana v programskem paketu MATLAB [28]. Programsko okolje MATLAB se uporablja za analizo in vizualizacijo podatkov. Metode je možno implementirati v programskem jeziku C++ in jih prevesti za uporabo v okolju MATLAB. Prav zaradi tega so operacije in izračuni v okolju MATLAB zelo hitri. Hitrost

pa je eden izmed glavnih vzrokov za razširjenost uporabe tega programskega okolja.

3.1 Zbirka slik

Pregled literature je pokazal, da zahtevna javnodostopna zbirka slik sadja ne obstaja. Eden izmed ciljev te diplomske naloge je bil zajeti uporabno zbirko slik sadja in jo javno objaviti. Za zajem zbirke sadja je bila napisana skripta in je opisana v naslednjem poglavju.

3.1.1 Zajem zbirke slik

Skripta za zajem zbirke slik je bila napisana v programskem jeziku Python [14]. Uporabljen razred *FancyURLopener* iz knjižnice *urllib* [15] omogoča pošiljanje zahtevkov HTTP (angl. hypertext transfer protocol, HTTP) in prejemanje odgovorov nanje. Za iskanje slik so bili zahtevki HTTP naslovljeni na vmesnik za iskanje slik Google [16]. Ta je v odgovoru HTTP vrnil spletne naslove posameznih slik. S pomočjo teh naslovov je bilo možno dostopati do slik in jih shraniti v zbirko.

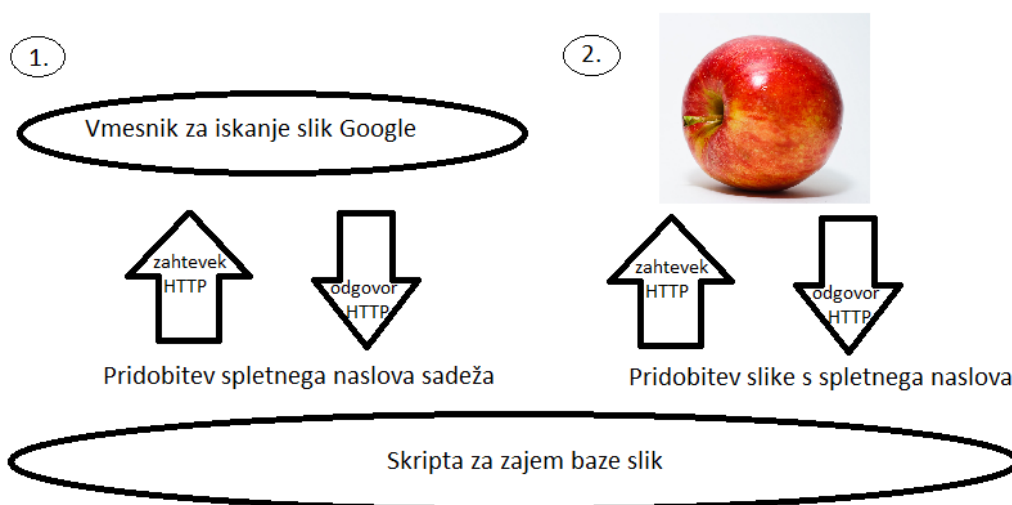
Problema pri zajemu zbirke sta bila dva. Eden izmed problemov je bilo blokiranje s strani vmesnika Google, ki preprečuje preveliko število zahtevkov v nekem časovnem intervalu. Ta problem je bil rešen zelo preprosto s sekundnimi zakasnitvami med poslanimi zahtevki. Pri drugem problemu je pri zahtevkih HTTP prišlo do napak, ki so povzročile nepričakovano zaustavitev izvajajočega programa. Do teh napak je prišlo zaradi izgubljenih zahtevkov HTTP in nedosegljivosti posameznih spletnih strani med poizvedovanjem. Za rešitev tega problema ni bilo dovolj samo obravnavanje napak. Samo z obravnavanjem bi v tem primeru izgubili potencialne slike za v našo zbirko slik. Problem je bil naslovljen s ponovnim pošiljanjem zahtevkov HTTP. V veliki večini je bila to rešitev problema in tako nismo izgubili slik sadja.

Implementacija skripte za zajem zbirke slik je dokaj preprosta in ne vsebuje zahtevnih odsekov programske kode. Na začetku se iz datoteke preberejo imena razredov sadja. Na podlagi teh imen se nato kreira prazna zbirka slik. Za vsak razred sadja se v zbirki slik kreira mapa z imenom razreda. Za vsak razred sadja se nato pridobi spletne naslove slik sadežev. S pomočjo spletnih naslovov slik posameznih sadežev se dostopa do njih in se jih shrani na pravo mesto v zbirki slik. Ker slike niso naša last je bilo potrebno spletne naslove slik shraniti v eno datoteko. Potek skripte je prikazan v Algoritmu 3.1 in na Sliki 3.1.

- 1: Preberi imena razredov sadja iz datoteke.
- 2: Kreiraj zbirko slik.
- 3: **for** vsak razred sadja **do**
- 4: Pridobi spletne naslove slik razreda.
- 5: **for** vsak spletni naslov **do**
- 6: Pridobi sliko in jo shrani na pravo mesto v zbirki.
- 7: Shrani spletni naslov slike.
- 8: **end for**
- 9: **end for**

Algoritem 3.1: Psevdokoda skripte za zajem zbirke slik.

Zajemanje zbirke slik s to skripto lahko traja tudi po več ur odvisno od števila zajetih slik. V našem primeru je zajemanje zbirke trajalo približno dve uri. Zajetih je bilo 30 razredov sadja in okoli 60 slik na razred.



Slika 3.1: Poenostavljena skica pridobivanja slike sadeža [35].

3.1.2 Predstavitev zbirke slik

Za diplomsko delo je bila zajeta zbirka 1800 slik s pomočjo skripte opisane v Poglavlju 3.1.1. Zbirka vsebuje razrede 30 različnih razredov sadja. Razredi so poimenovani z angleškimi izrazi, zaradi večje množice zadetkov pri iskanju z vmesnikom Google. Za vsak razred je bilo sprva zajetih 60 različnih slik. Zajete slike so bile filtrirane ročno, zaradi obilice neprimernih in napačnih zadetkov. Vmesnik za iskanje slik Google je dober, vendar se zanaša na informacije podane s strani avtorja in konteksta slike in zaradi tega je prihajalo do napačni zadetkov. Nekaj zajetih slik je bilo premajhnih za opisovanje z opisnikoma HOG in BOW. Te slike so bile izločene z nastavljivo mejo pri branju slik za uporabo v implementiranem programu.

Izmed 1800 slik jih je bila približno polovica neprimernih oz. napačnih. Pojavljale so se slike mest z imeni sadežev. Poleg sadeža *kiwi* obstaja tudi žival *kiwi*, zato je bil ta razred preimenovan v *kiwifruit*. Pojavljale so se slike cvetov rastlin in dreves, ki obrodijo te sadeže. Poleg vsega tega so se pojavljale še slike sladice ter šamponov z okusi teh sadežev in risane slike. Moteči so bili tudi napisi, ki se ponekod raztezajo čez celotno sliko. Slike

z vsemi temi motečimi elementi so bile odstranjene. Navkljub potrebi po ročnem filtriranju zbirke, je bil zajem s skripto dosti hitrejši, kot bi bilo ročno iskanje slik. Prisotnost uporabnika ob zajemu zbirke s skripto prav tako ni bilo potrebno. S skripto se tudi deloma izognemo subjektivnemu izbiranju slik pri ročnem iskanju.

Graf prikazan na Sliki 3.2 prikazuje število slik v posameznih razredih po izbrisu napačnih, neprimernih in risanih slik. Zbirka slik vsebuje 971 slik. Slike so razvrščene v 30 angleško poimenovanih razredov. Povprečno število μ slik na razred je približno 32, standardni odklon σ slik na razred je približno 6. Slike so podane v formatu JPEG, ki je eden izmed najrazširjenih in najpogosteje uporabljenih formatov za fotografije. Slike v zbirki so velike od nekaj KB pa vse do visokoresolucijskih slik velikosti nekaj MB.

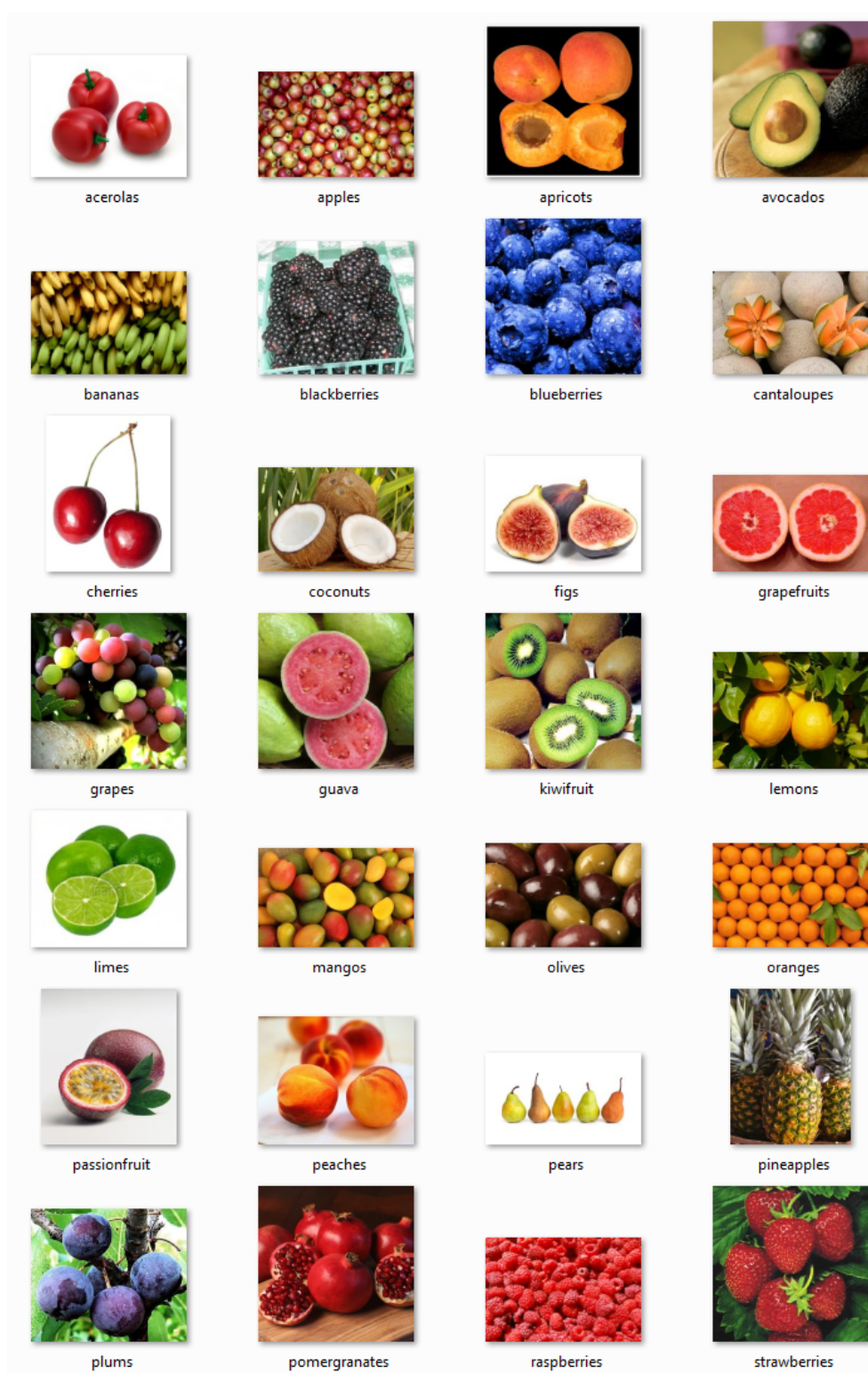


Slika 3.2: Število slik posameznih razredov sadja v zajeti zbirki.

Primeri zajete zbirke slik so prikazani na Sliki 3.3. Zbirka slik je zanimiva zaradi raznolikosti znotraj razredov in podobnosti med nekaterimi razredi. Na nekaterih slikah je samo en sadež. Na nekaterih drugih slikah je zgolj nekaj sadežev, na ostalih pa je zelo veliko sadežev in jih je že nemogoče prešteti. Nekateri sadeži so obrani, spet drugi so slikani na drevesih. Veliko sadežev nima bele podlage, ampak neko drugo ozadje recimo drevo, krožnik, itd. Za vse slike, ki ne vsebujejo sadeža na beli podlagi, rečemo, da vsebujejo šum. Šumne slike je dosti težje dobro opisati in klasificirati. Kljub vsemu šumu in raznolikosti lahko za vse slike, ki so v zbirki, rečemo, da je bistvo slike sadež oz. več sadežev.

Dostopnost zbirke

Zajeta zbirka slik sadja (angl. fruit image data set 30, FIDS30) je dostopna na spletni strani laboratorija VICOS [20]. Slike v zbirki niso naša last, saj so bile pridobljene iz spletnih virov. Spletni naslovi virov slik so navedeni v zbirki slik sadja. Pridobljene slike niso bile uporabljene za komercialne namene.



Slika 3.3: Primeri slik sadja zajete zbirke slik [20].

3.2 Implementacija sistema

V sklopu diplomske naloge je bilo implementiranih in uporabljenih kar nekaj opisnikov, klasifikacijskih in ostalih metod. Integracija že implementiranih metod je bila večkrat težavna zaradi pomanjkljive dokumentacije. Pojavila se je še težava s kompatibilnostjo. Večkrat so metode bile implementirane za točno določene operacijske sisteme oz. verzije le teh. Prenos programa na drug sistem je večkrat povzročilo nekompatibilnost pri prevajanju metod napisanih v jeziku C++ za programsko okolje MATLAB.

V Poglavju 2 je bila opisana teoretična podlaga opisnikov, klasifikacijskih metod, izračun uspešnosti in arhitekture samega sistema. Za uporabo pri diplomskem delu so bile implementirane naslednje metode oz. algoritmi:

- Ogrodje celotnega sistema
- Bralnik zbirke slik z možnostjo obrezovanja in izbrisa posameznih slik
- Delitev zbirke slik na učno in testno množico
- Različni histogrami barvnih prostorov
- Globalni opisnik oponentnih barv s pomočjo opisnika HOG [26]
- Usrediščenje in normalizacija podatkov za klasifikacijo metode SVM
- Večinski klasifikator realiziran s pomočjo klasifikatorja KNN
- Priporočilni sistem na podlagi modela SVM in opisnika BOW
- Izračun klasifikacijske točnosti
- Prikaz rezultatov klasifikacije
- Prikaz pozitivnih in negativnih klasifikacij

V naslednjih odstavkih so navedeni viri opisnikov in klasifikacijskih metod, katere nismo implementirali v sklopu diplomske naloge. Te metode so bile že implementirane v sklopu odprtokodnih knjižnic.

Implementaciji odločitvenih dreves in metode KNN sta bili implementirani v sklopu programskega paketa MATLAB [28].

Uporabljena implementacija opisnika HOG v našem sistemu je bila implementirana s strani piscev članka [25]. Programska koda je na voljo na spletni strani [26].

Uporabljen opisnik BOW s prostorskim piramidnim ujemanjem je bil implementiran v knjižnici VLFEAT. Programska koda za okolje MATLAB je dostopna na spletni strani [38]. Ta odprtokodna knjižnica vsebuje implementacije mnogih algoritmov, ki se danes uporabljajo na področju računalniškega vida.

Programska koda implementacije algoritma RF je javno dostopna za uporabo na spletni strani [17].

Implementacija klasifikatorja SVM je bila razvita v knjižnici LIBSVM. Programska koda je na voljo na spletni strani [9]. Poleg te implementacije metode SVM je bila uporabljena še implementacija v knjižnici VLFEAT [38].

3.3 Protokol analize

Opisnikov in klasifikacijskih metod uporabljenih in implementiranih v našem sistemu je bilo precej. Analiza teh modelov je zato trajala precej časa, že zaradi samega časa opisovanja slik in grajenja modelov. Problem časovne zahtevnosti grajenja opisnikov je bil naslovljen s shranjevanjem opisov. Shranjevanje opisov zbirke je bila dobra praksa za znižanje časa pri vnovični uporabi istega opisnika.

Največ časa je bilo posvečeno opisniku BOW in dobri klasifikaciji le tega. Opisovanje slik z opisnikom BOW in klasifikacija teh slik s pomočjo metode SVM je trajalo približno pol ure. Glede na nastavljene parametre pa je lahko klasifikacija s pomočjo opisnika BOW trajala tudi do ene ure. Ta opisnik je imel največ nastavljivih parametrov od vseh metod in s tem največ možnosti za zvišanje uspešnosti. Za model SVM na opisniku BOW je bil implementiran tudi priporočilni sistem.

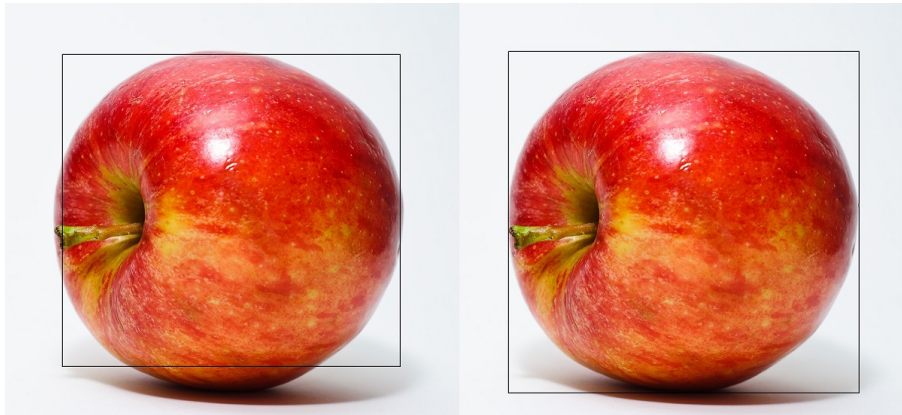
Priporočilni sistem je različica klasifikacijskega modela, pri katerem namesto enega razreda model predlaga več najverjetnejših razredov. V našem primeru priporočilni sistem za sliko predlaga pet najbolj verjetnih sadežev, ki bi se lahko nahajali na sliki. Za ocenjevanje uspešnosti tega modela se je uporabila klasifikacijskih točnost vseh petih napovedi.

Uspešnosti metod

Uspešnosti posameznih opisnikov in klasifikacijskih metod so bili izračunani na podlagi zajete zbirke slik. Grajenje klasifikacijskega modela je potekalo na podmnožici zbirke slik, ki jo imenujemo učna množica. Testiranje je potekalo na drugi podmnožici zbirke slik, ki jo imenujemo testna množica. Testna in učna množica slik nista imeli nobenih skupnih primerov. Uspešnost modela je bila izračunana kot povprečje 100 meritev uspešnosti modela, razen pri opisniku BOW. Grajenje modela na podlagi opisnika BOW je trajalo tudi do ene ure, zato se je uspešnost tega modela izračunala kot povprečje petih meritev. Učna in testna podmnožica sta bili vsakič zgrajeni naključno v razmerju približno 7:3.

Anotacije

Velika večina slik prikazuje sadež obdan z neko šumno okolico. Vpliv okolice lahko preprosto zmanjšamo z obrezovanjem slik. Sadež je večinoma centriran oz. se nahaja na sredini slike. Prav zaradi tega lahko uvedemo avtomatsko obrezovanje slik z neko relativno mejo. Avtomatsko obrezovanje slik se je izkazalo za dobro pri nakaterih opisnikih. Slika 3.4 prikazuje avtomatsko obrezovanje roba slike velikosti $1/8$. Za najboljše pa se je izkazalo ročno obrezovanje slik. Lastnoročno obrezovanje je odlično in doprinese k zvišanju uspešnosti vseh modelov, ker lahko obrezani del točno prilagodimo naši sliki in sadežu na njej.



Slika 3.4: Leva anotacija prikazuje avtomatsko odrezani rob slike [35], desna anotacija pa prikazuje ročno obrezovanje.

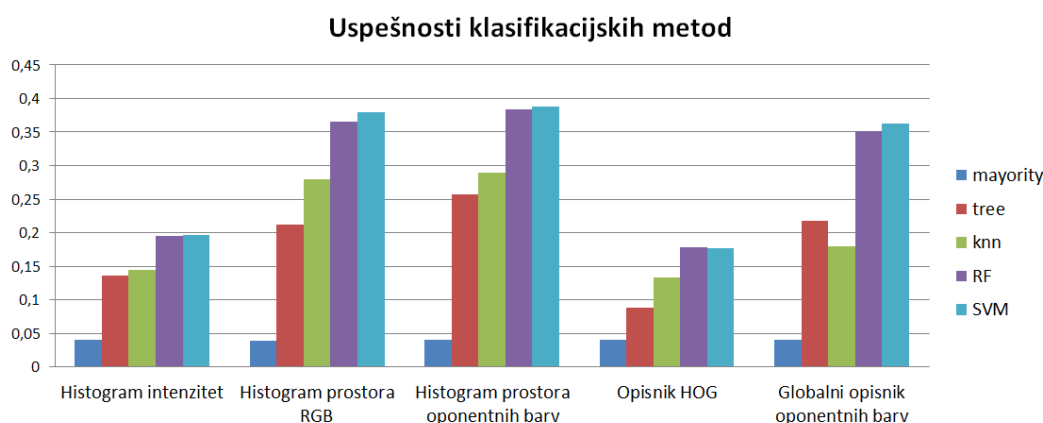
3.4 Rezultati

V tem poglavju so ovrednoteni rezultati opisnikov in klasifikacijskih metod na zajeti zbirki slik. Ovrednotene so anotacije in usrediščenje ter normalizacija podatkov. Na koncu poglavja je še predstavljen priporočilni sistem, ki temelji na opisu slik z opisnikom BOW in klasifikaciji s pomočjo metode SVM

3.4.1 Vrednotenje klasifikacijskih metod

Minimalna klasifikacijska točnost naših metod je morala preseči uspešnost večinskega klasifikatorja, ki je v našem primeru približno 4%. V tem poglavju so ovrednotene metode za klasifikacijo, ki so bile uporabljene v nalogi. Mejo večinskega klasifikatorja so presegle vse uporabljene metode.

Graf na Sliki 3.5 prikazuje klasifikacijske točnosti posameznih metod zgrajenih na različnih opisnikih. Na grafu so prikazani rezultati klasifikacij le s pomočjo petih različnih opisnikov, kljub temu so rezultati zelo jasni. Iz te slike je razvidno, da sta metodi RF in SVM bili daleč pred odločitvenimi drevesi in metodo KNN. Razvrstitev metod po uspešnosti se malo spreminja, vendar sta metodi RF in SVM približno enako uspešni.



Slika 3.5: Klasifikacijske točnosti metod na različnih opisnikih.

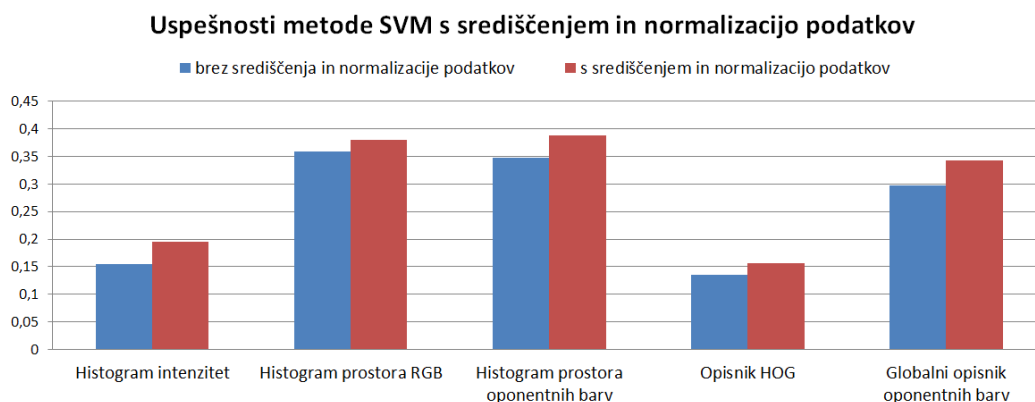
Po pričakovanjih se je uspešnost odločitvenih dreves izkazala za najslabšo pri skoraj vseh opisnikih. Preizkušeni opisniki so opisali slike z različno dolgimi vektorji decimalnih števil. Odločitvena drevesa so ena izmed najpreprostejših klasifikacijskih metod in preprosto niso predvidena za takšne velike in težke probleme. Odločitvena drevesa so dosegle 10-20% nižjo uspešnost, kot metoda SVM na istih opisnikih.

Metoda KNN je dosegla povprečno nekaj odstotkov boljše rezultate kot odločitveno drevo. Podobno kot odločitveno drevo je tudi metoda KNN, ena izmed preprostejših metod za klasifikacijo. Pri uporabljenih opisnikih se vrednosti celic v histogramih raztezajo čez neko območje največkrat od nič do ena. Metoda KNN ta problem naslavlja boljše kot odločitvena drevesa, vendar še vedno ne dosega prav dobrih rezultatov. Metoda KNN postane problematična pri daljših opisnikih kot sta opisnik HOG in globalni opisnik

oponentnih barv, kot je razvidno na Sliki 3.5. Tudi ta klasifikacijska metoda torej ni bila dovolj za doseg največje uspešnosti pri našem problemu.

Metoda naključnih gozdov se je izkazala zelo primerna za razpoznavanje sadja, saj je dala približno enake rezultate kot metoda SVM. Za klasifikatorja RF in SVM tudi na splošno velja, da dajeta podobne uspešnosti. Na opisnikih prikazanih na Sliki 3.5 je metoda RF dosegla klasifikacijske točnosti okoli 40%. Ta uspešnost je že kar dobra glede na to, da zbirka slik vsebuje 30 razredov sadja. V primerjavi z klasifikatorjem večinskega razreda je to desetkrat izboljšana uspešnost.

Metoda SVM se je izmed izbranih klasifikatorjev izkazala kot najbolj primerna za razpoznavanje sadja. Metoda SVM je uporabna prav zaradi odličnih in številnih implementacij, ki so prosto dostopne na spletnih straneh oddelkov, ki so jih razvili. V našem primeru sta bili uporabljeni dve implementaciji, ki sta bili navedeni v Poglavju 3.2. Metoda SVM je na opisnikih na Sliki 3.5 dosegala točnosti nekje do 40%, malo več kot metoda RF. V nadaljevanju je bila uporabljena za klasifikacijo različnih opisnikov.

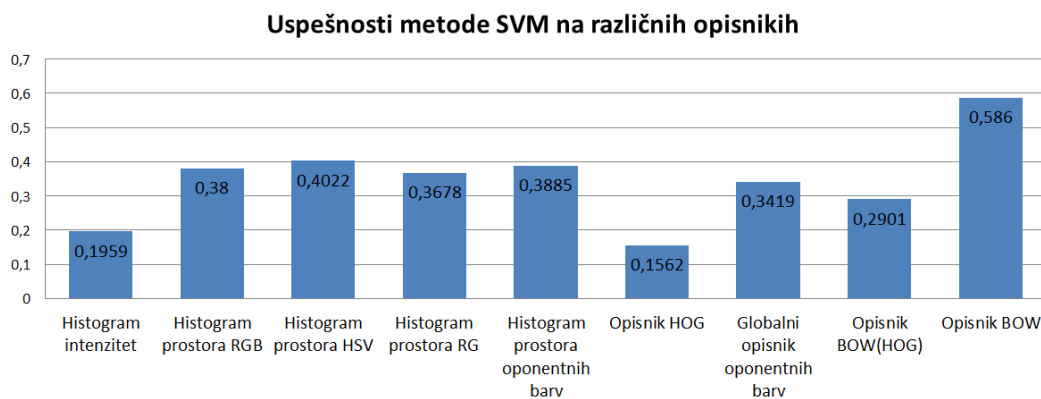


Slika 3.6: Klasifikacijske točnosti metode SVM s in brez središčenja in normalizacije podatkov.

Graf na Sliki 3.6 prikazuje zvišanje klasifikacijske točnosti s središčenjem in normalizacijo podatkov za klasifikacijo z metodo SVM, kot je bolj podrobno opisano v Poglavju 2.4.4. Iz tega grafa je razvidno tudi do 5% zvišanje uspešnosti metode SVM s središčenjem in normalizacijo podatkov.

3.4.2 Vrednotenje opisnikov

Slika 3.7 prikazuje klasifikacijske točnosti metode SVM zgrajene na različnih opisnikih slik. Pred klasifikacijo so bili podatki usrediščeni in normalizirani, kot se je izkazalo za dobro v Poglavju 3.4.1.



Slika 3.7: Klasifikacijske točnosti metode SVM na različnih opisnikih.

Opisnik HOG je dosegel najslabše rezultate med opisniki. Informacija o obliki in izgledu sadežev ni bila dovolj za dobro napoved razreda slike. Slabi rezultati opisnika HOG niti niso tako presenetili, zaradi zelo šumne in težke zbirke slik. Veliko slik ni vsebovalo osamelih sadežev, ampak kopico sadežev, na katerih je bilo težko zajeti nek reprezentativen opis za kvalitetno klasifikacijo s pomočjo opisnika HOG.

Histogram intenzitet se je izkazal kot zelo slab opisnik za opisovanje slik. Ta rezultat niti ni presenetil, saj je bil ta opisnik uporabljen le zaradi začetne izgradnje in priprave sistema. Razlog za zelo slab rezultat je bila neuporaba barvnih in teksturnih informacij o sliki. Najvišja dosežena točnost tega kla-

sifikatorja je bila okoli 24% s pomočjo obrezovanja slik, ki je prikazano na Sliki 3.8.

Opisnik BOW, ki za opis značilni točk uporablja opisnik HOG, je podvojil uspešnost opisnika HOG. Ta opisnik je zajel informacije o teksturi, kljub temu je uspešnost še vedno slabša kot pri preprostih barvnih histogramih. Iz teh rezultatov smo lahko sklepali, da opisnik HOG preprosto ni bil dovolj za opisovanje značilnih točk, kaj šele celotnih slik.

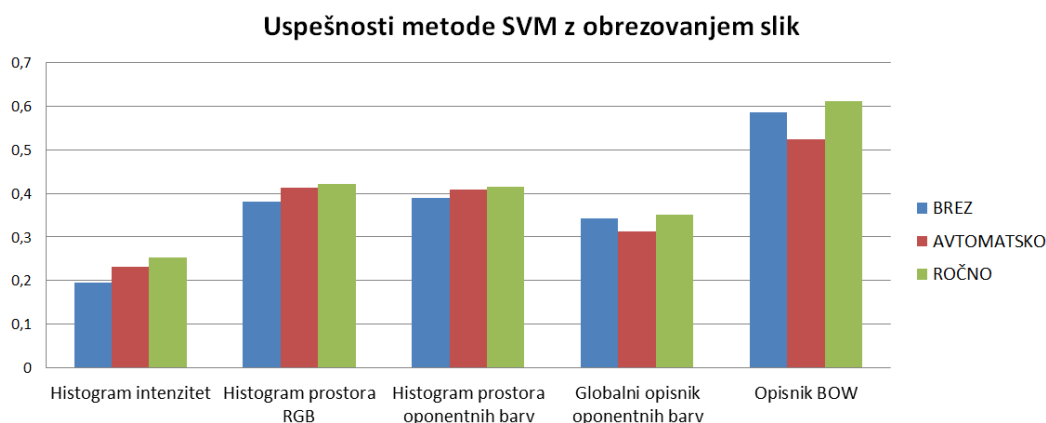
Združena informacija o barvah in obliki je bila predstavljena z globalnim opisnikom oponentnih barv. Kljub obetom je ta opisnik dosegel nižjo točnost kot vsi barvni opisniki. Dodana informacija o obliki in izgledu je celo znižala uspešnost metode SVM na tem opisniku. Za zajem združenih informacij o barvi, teksturi ter obliki smo potrebovali resnično dober in sodoben opisnik in ne le združevanje posameznih opisnikov v nek mogoče dober opisnik.

Barvni histogrami so dali veliko boljše rezultate. Njihova prednost je bila upoštevanje barvne informacije, ki je ključna pri klasifikaciji slik. Klasifikacijske točnosti teh opisnikov so se gibale okoli 36-40%. S pomočjo obrezovanja slik se je ta točnost še povečala za nekaj odstotkov in presegla mejo 40%. Obrezovanje slik je pri barvnih histogramih dalo najboljše rezultate, kar je razvidno na Sliki 3.8. Točnost večinskega klasifikatorja je bila okoli 4%, kar je za barvne histograme pomenilo desetkratno izboljšanje uspešnosti večinskega klasifikatorja.

Opisnik BOW, ki je podrobno opisan v Poglavju 2.3, je dosegel najboljši rezultat na zbirki slik in skoraj 20% višjo uspešnost kot ostali opisniki. Opis teksture na podlagi opisov značilnih točk, ki upoštevajo barvno informacijo, se je torej v našem primeru izkazal za najbolj informativen opis posameznih slik. Metoda SVM je z opisnikom BOW dosegla klasifikacijsko točnost 58,6%. Najuspešnejši model in njegovi problemi so podrobno opisani v Poglavju 3.4.3.

Obrezovanje slik se je izkazalo za učinkovito, kot je to razvidno iz grafa na Sliki 3.8. Avtomatsko obrezovanje, ki je v naši implementaciji odrezalo zunanji rob širine $1/8$ celotne slike, je pripomoglo 3-4% k uspešnosti pri barv-

nih opisnikih. Ročno obrezovanje slik, kjer uporabnik sam obreže slike, pa je prineslo do celo 5% zvišanje uspešnosti in to pri vseh opisnikih. Klasifikacijska točnost najuspešnejšega modela je z obrezovanjem slik tako narasla na 61,2%. Obrezovanje in odstranjevanje šuma s slik se je izkazalo kot zelo dobra ideja.

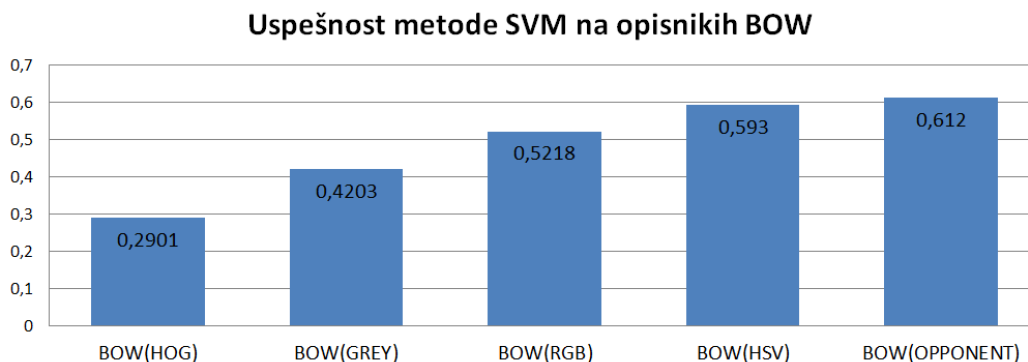


Slika 3.8: Klasifikacijske točnosti metode SVM z obrezovanjem slik.

3.4.3 Najuspešnejši klasifikacijski model

Klasifikacijski model zgrajen s pomočjo metode SVM na opisih opisnika BOW je dosegel najvišjo uspešnost na zajeti zbirki slik sadja. Opisnik BOW je bil zgrajen na ročno obrezanih slikah v barvnem prostoru oponentnih barv. Ta model je dosegel klasifikacijsko točnost 61,2%. Dosežena uspešnost je dobra, zaradi obilice šuma in raznolikih slik prisotnih v zbirki slik sadja. Opisovanje zajete zbirke slik z opisnikom BOW in grajenje modela SVM traja odvisno od nastavitve od pol do ene ure.

Graf na Sliki 3.9 prikazuje uspešnosti opisnika BOW zgrajenega na različnih barvnih prostorih slik. Nastavitev barvnega prostora je bila ključna za pridobitev najvišje uspešnosti modela, saj je barvni prostor slik dajal največje spremembe pri uspešnosti klasifikacije opisnika BOW. Barvna prostora oponentnih barv in HSV sta dala daleč najboljše rezultate.



Slika 3.9: Točnosti metode SVM na opisnikih BOW z ročnim obrezovanjem slik.

Slika 3.10 prikazuje deleže pravilno razvrščenih slik posameznih razredov. Ti deleži se spreminjajo in so lahko različni za isti klasifikacijski model. Ne glede na spremembe pa veljajo splošna pravila, zakaj je posamezna slika napačno klasificirana. Na Sliki 3.11 so prikazani primeri napačnih klasifikacij opisanih v tem poglavju. Podrobnejše informacije napačnih napovedi so prikazane na Sliki 3.12, ki prikazuje matriko zamenjav (angl. confusion matrix) razredov najuspešnejšega klasifikacijskega modela. Vrstice v matriki zamenjav prikazujejo delež napovedi za različne razrede na slikah opazovanega razreda.

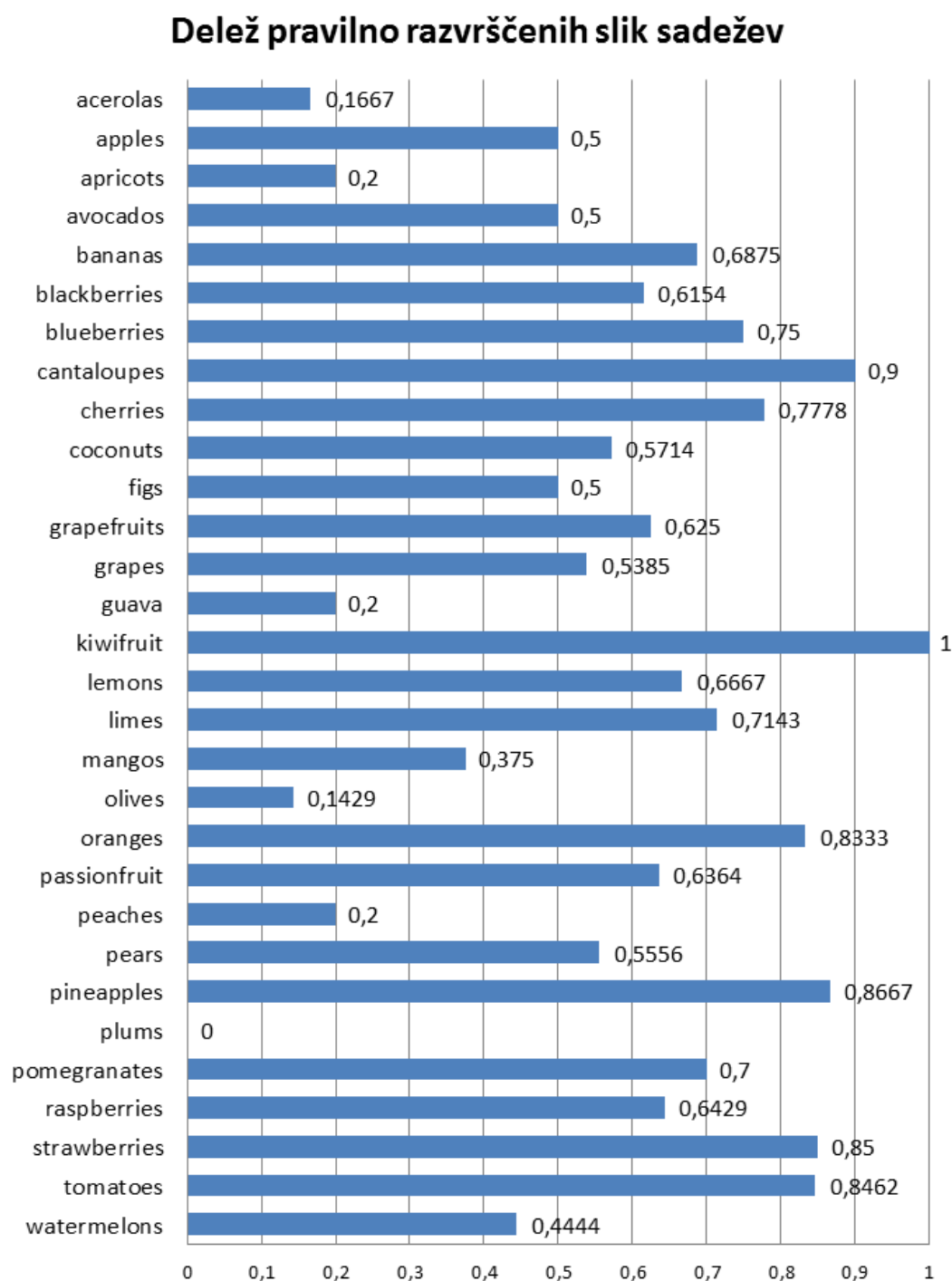
S Slike 3.10 je razvidno, da je precej razredov klasificiranih zelo dobro. Pravilne napovedi teh razredov presežejo delež 75%. Ti razredi so unikatni in jih je preprosto prepoznati s slik. Razredi kivi, melona, ananas in pomaranče so edini s svojimi nabori barv. Težko jih je zamenjati z drugimi razredi, ker so unikatni po barvi ter po teksturi znotraj in zunaj sadeža. Iz Slike 3.12 je razvidno, da se ti razredi skoraj ne zamenjujejo z drugimi.

Drugi razlog za dobro klasifikacijo je velika in raznolika učna množica slik razreda, ki prinese širino k napovedovanju tega razreda. Problem te širine je veliko število napačnih klasifikacij ostalih razredov, kar je razvidno na Sliki 3.12 iz razredov paradižnika in jagod. Uspešnost klasifikacije pa-

radižnika in jagod je dobra, vendar klasifikacijski model klasificira marsikatero sliko kot paradižnik ali jagodo in tako zniža uspešnost celotnega modela. Ta problem se zelo spreminja od modela do modela zaradi razmerja učnih in testnih slik razredov.

Razredi, ki dosegajo nižje uspešnosti kot 75%, a še vedno več kot 50%, sistem večinoma dobro razpozna, vendar so problematični ker so si po dva ali trije razredi zelo podobni med seboj in se zamenjujejo. Razreda limona in limeta sta si po teksturi zelo podobna in sistem jih zaradi tega zamenjuje, kar je zelo razvidno na Sliki 3.12. Še ena zelo vidna skupina so grozdje, robide in borovnice, ki se med seboj zelo pogosto zamenjujejo.

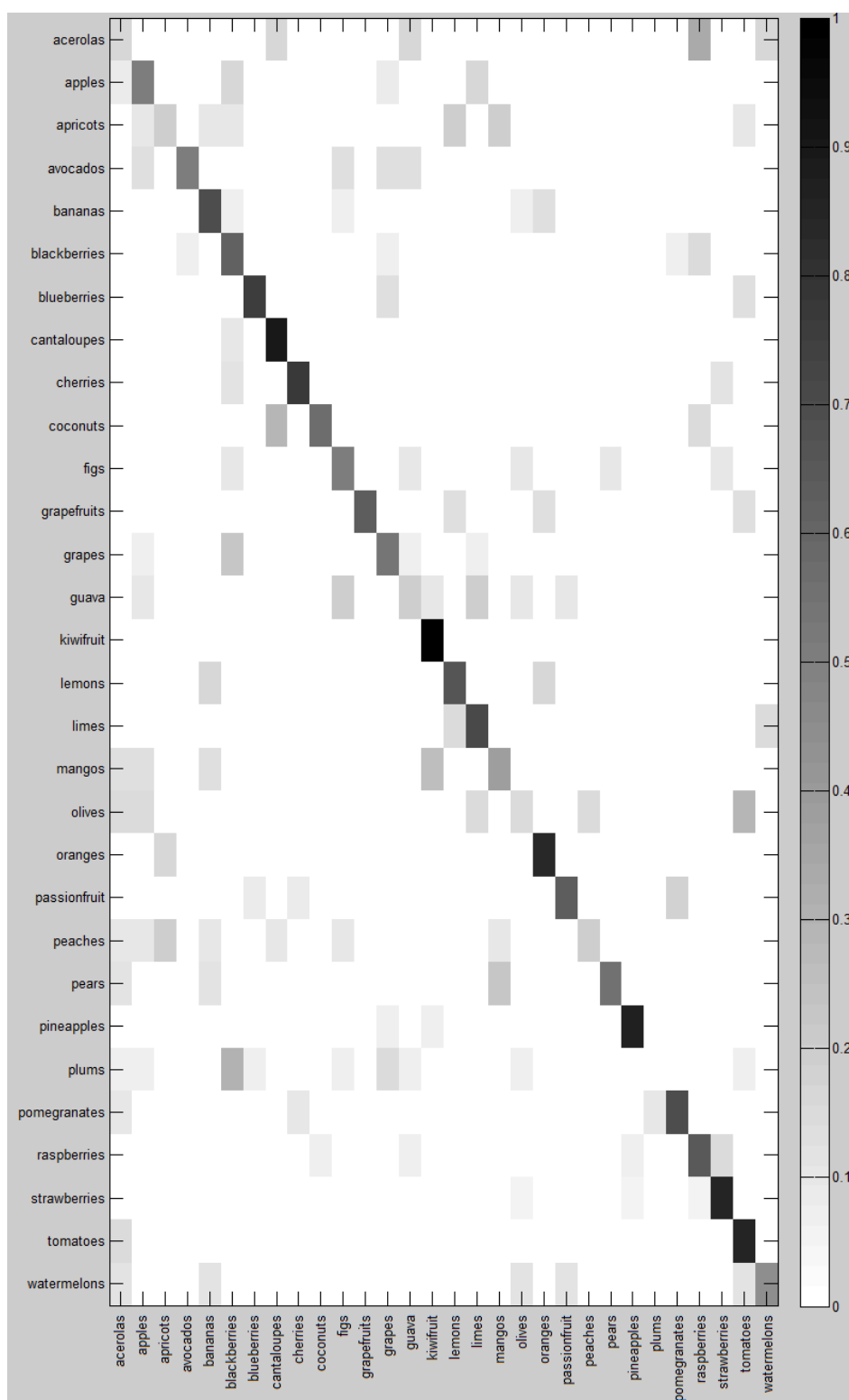
Slike razredov z uspešnostjo pod 40% so večinoma zelo raznolike, težke za prepoznavo in imajo poleg vsega tega še veliko podobnih razredov. Olive in slive so dosegle najslabšo razpoznavo v našem sistemu. Ta dva razreda sta razpoznana kot marsikatero sadje in ne vsebujeta nekih logičnih povezav z napačno razpoznanimi razredi. Pregled napačno klasificiranih slik oliv in sliv je pokazal, da so slike teh dveh razredov zelo raznolike z obilico šumnega ozadja, zato tudi slaba ocena teh slik ni presenetljiva.



Slika 3.10: Delež pravilno razvrščenih slik razredov na podlagi opisnika BOW in metode SVM.



Slika 3.11: Primeri napačno razvrščenih slik zbirke [20] z napovedanimi klasifikacijami.

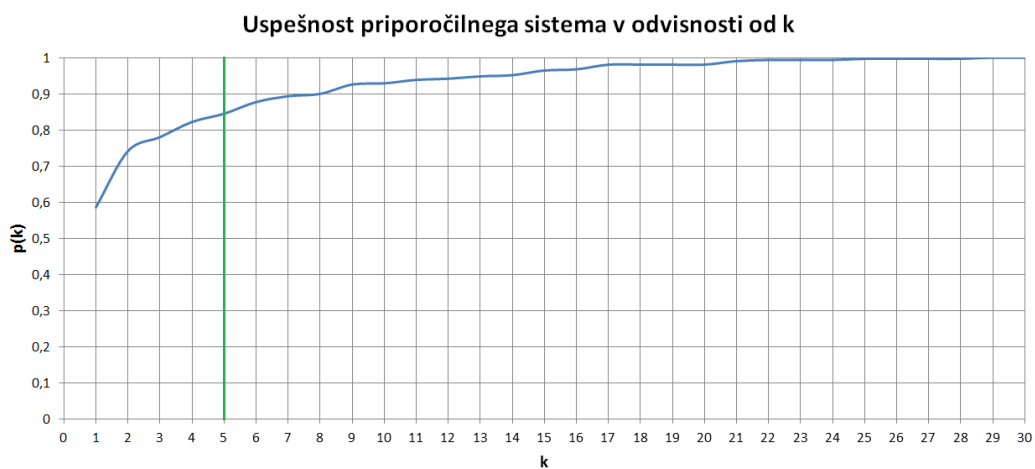


Slika 3.12: Matrika zamenjav razredov najuspešnejšega klasifikacijskega modela.

3.4.4 Priporočilni sistem

Do sedaj smo obravnavali klasifikacijske modele, ki so kot rezultat klasifikacije podali razred slike. V tem poglavju je na kratko predstavljen priporočilni sistem, ki predlaga več najverjetnejših razredov za sliko. Priporočilni sistem je bil zgrajen na podlagi najuspešnejšega klasifikacijskega modela, ki je opisan v Poglavju 3.4.3.

Uporabnik priporočilnega sistema lahko na podlagi predlogov sistema poišče pravi razred slike recimo na spletu. Z več predlogi, priporočilni sistem doseže višjo uspešnost kot klasifikacijski modeli. Uspešnost priporočilnega sistema se izračuna kar kot klasifikacijskih točnost vseh predlogov, kot je to opisano v Poglavju 2.4.5. Ta sistem lahko doseže zelo visoko uspešnost, vendar pa je le ta pogojena s številom predlaganih razredov. Graf na Sliki 3.13 prikazuje uspešnost priporočilnega sistema v odvisnosti od števila predlogov k .



Slika 3.13: Uspešnost priporočilnega sistema v odvisnosti od števila predlogov k .

Graf na Sliki 3.13 nam daje jasno predstavo o uspešnosti priporočilnega sistema. Iz grafa je razvidna 85% verjetnost za prisotnost pravega razreda slike med prvimi petimi predlogi. Pregled petih različnih sadežev za 85%

zagotovljen pravilni rezultat, še ni tako problematičen. Za večjo uspešnost bi bilo potrebno število zadetkov zelo povečati, s tem pa bi se dejanska uporabnost takšnega sistema močno zmanjšala.

Poglavje 4

Sklepne ugotovitve

Razpoznavanje sadja je zahteven problem, ki je pritegnil pozornost številnih raziskovalcev. O tem priča obilica sorodnih člankov, npr., [4, 7, 32, 34, 40]. Delo na tem zelo aktualnem področju še ni končano in se bo razvijalo naprej.

V diplomski nalogi smo obravnavali problem razpoznavanja sadja. Sadje nudi zahtevno domeno za razpoznavanje zaradi raznovrstnosti sadežev istega razreda, podobnosti med posameznimi razredi sadja in samega števila različnih razredov sadja.

Cevovodna arhitektura našega sistema je zelo poenostavila potek razpoznavanja sadja, izračuna uspešnosti in pregledovanja rezultatov. Posamezne faze programa so se odvijale neodvisno od uporabljenih procedur in algoritmov. Na ta način smo lahko preizkusili poljubne kombinacije raznih metod za pripravo podatkov, opisnikov in klasifikacijskih metod.

V sklopu diplomske naloge je bila zajeta velika zbirka slik sadja, ker takšna uporabna, javnodostopna zbirka slik sadja še ne obstaja. Zbirka je bila zajeta s pomočjo napisane skripte in objavljena na spletni strani laboratorija VICOS [20]. Javnodostopna objava anotirane zbirke slik sadja predstavlja enega izmed prispevkov te diplomske naloge.

Opisane in preizkušene klasifikacijske metode in opisniki so bile izbrane na podlagi dobrih rezultatov v člankih [4, 12, 30, 32, 39]. Metode, opisane v teh člankih, rešujejo problem razpoznavanja sadja in druge podobne probleme.

Te metode so bile preizkušene na zajeti zbirki slik z zelo raznolikimi slikami sadja. Pridobitev kvalitetnega atributnega zapisa je bilo nujno za dobro razpoznavanje sadja.

Poglobitev v lastnosti in postopke posameznih opisnikov je bila nujna za dobro razpoznavanje sadja in razumevanje ter razlago rezultatov. Opisniki so zajeli informacije o barvi, teksturi, obliki in izgledu sadežev. Barvni histogrami so bili uporabni pri opisovanju celotnih slik in so zelo dobro zajeli barvno informacijo. Informacije o obliki in izgledu sadežev, ki jih opiše opisnik HOG, se niso izkazale dovolj pri opisovanju slik in so dajale zelo slabe rezultate. Opisnik BOW se je izkazal kot najbolj primeren za opisovanje pridobljene zbirke slik, kar tudi ni bilo presenetljivo, glede na res dober in informativen opis teksture sadežev.

Na področju strojnega učenja se je uveljavilo že veliko klasifikacijskih metod. V diplomskem delu je bilo opisanih in preizkušenih nekaj teh metod, ki so dajale zelo različne rezultate. Pri razpoznavanju sadja se je najbolj izkazala metoda SVM, ki je danes ena izmed najbolj popularnih in robustnih klasifikacijskih metod. Zelo dobre rezultate je dosegla tudi metoda naključnih gozdov in je bila resen tekmelec metode SVM. Zaradi slabših rezultatov metode RF, je bila na koncu pri grajenju najboljšega modela uporabljena odlična implementacija metode SVM [38].

Klasifikacijska točnost implementiranega sistema je dosegla zgornjo mejo pri 62%. Ta uspešnost je zelo dobra glede na težavnost zbirke slik in samega zastavljenega problema. Poleg klasifikacijskega modela je bil implementiran še priporočilni sistem, ki pa nam na podlagi petih predlogov zagotovi kar 85% uspešnost.

Osrednji prispevek diplomske naloge je bila torej podrobna analiza opisnikov in klasifikacijskih metod za problem razpoznavanja sadja. Na podlagi analize in rezultatov je bil zgrajen priporočilni sistem. Možnosti za izboljšanje priporočilnega sistema je veliko in so predstavljene v Poglavju 4.1.

4.1 Smernice za nadaljnji razvoj

Manevrskega prostora za izboljšanje uspešnosti implementiranega sistema je veliko. Za začetek bi lahko poizkusili z boljšimi klasifikacijskimi metodami. Lahko bi uporabili uteženo glasovanje različnih modelov ali pa še kakšno drugo metodo za klasifikacijo, npr., [6, 31]. Uporaba boljših in večih klasifikacijskih modelov bi lahko izboljšalo uspešnost napovedi samo do neke mere. Za nadaljne izboljšanje bi potrebovali še bolj informativne opisnike.

Hiter razvoj na področju razpoznavanja stalno prinaša nove opisnike in izboljšane implementacije že obstoječih. Najprej bi bilo vredno poizkusiti implementacijo barvnega opisnika HOG opisanega v članku [18]. Poleg tega opisnika bi bilo vredno poizkusiti še kakšne opisnike za razpoznavanje scen in kakšnih drugih podobnih problemov [5, 12, 30, 39, 40].

V diplomskem delu je bila uspešnost napovedovanja izboljšana z obrezovanjem slik, ker smo s tem odstranili nekaj šuma iz slike. Šum bi lahko še bolje odstranili s slike z izrezovanjem sadežev s slik, glede na spremembo intenzitete, kot je to narejeno v članku [4]. Izrezovanje sadežev bi še boljše pripomoglo k zvišanju uspešnosti, kot le obrezovanje slik.

Uspešnost modelov bi lahko izboljšali še z večjo zbirko slik sadja, ki jo lahko brez težav zajamemo s pomočjo implementirane skripte.

Literatura

- [1] Color histogram. Dostopno na: http://en.wikipedia.org/wiki/Color_histogram.
- [2] Decision tree. Dostopno na: http://en.wikipedia.org/wiki/Decision_tree.
- [3] Lab color space. Dostopno na: http://en.wikipedia.org/wiki/Lab_color_space.
- [4] S. Arivazhagan, R. N. Shebiah, S. S. Nidhyanandhan, and L. Ganesan. Fruit recognition using color and texture features. *Journal of Emerging Trends in Computing and Information Sciences*, 1(2):90–94, 10 2010. Dostopno na: http://www.cisjournal.org/archive/vol1no1/vol1no1_12.pdf.
- [5] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 509–522, 4 2002. Dostopno na: http://doczine.com/bigdata/2/1367193725_305b572417/bmp-shape.pdf.
- [6] Christopher M. Bishop. *Neural Networks and Machine Learning*. Springer, 1998.
- [7] D. M. Bulanon and T. Kataoka. Fruit detection system and an end effector for robotic harvesting of fuji apples. *Agricultural Engineering International*, 12(1):203–210, 3 2010. Dostopno

- na: <http://www.cigrjournal.org/index.php/Ejournal/article/viewFile/1285/1319>.
- [8] Pavel Campr. Hog lena old. Dostopno na: http://pavel.campr.cz/public/scikits-image-hog-fix/hog_lena_old.png, 2012.
- [9] C. Chang and C. Lin. Libsvm. Dostopno na: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 4 2013.
- [10] A. Criminisi, J. Shotton, and E. Konukoglu. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Journal Foundations and Trends® in Computer Graphics and Vision*, 7:81–227, 2 2012. Dostopno na: http://research.microsoft.com/pubs/158806/CriminisiForests_FoundTrends_2011.pdf.
- [11] P. Cunningham and S. J. Delany. k-nearest neighbour classifiers. *International Conference on Electrical Engineering and Informatics*, 3 2007. Dostopno na: <http://www.csi.ucd.ie/UserFiles/publications/UCD-CSI-2007-4.pdf>.
- [12] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:886–893, 6 2005. Dostopno na: http://hal.archives-ouvertes.fr/docs/00/54/85/12/PDF/hog_cvpr2005.pdf.
- [13] Li Fei-Fei. Bow histogram.
- [14] Python Software Foundation. Python programming language. Dostopno na: <http://www.python.org/>.
- [15] Python Software Foundation. urllib. Dostopno na: <http://docs.python.org/2/library/urllib.html>.
- [16] Google. Google images. Dostopno na: <http://www.google.com/imghp>.

-
- [17] A. Jaialtilal, A. Liaw, M. Wiener, L. Brieman, and A. Cutler. Random forest matlab. Dostopno na: <https://code.google.com/p/randomforest-matlab/>, 12 2010.
- [18] F. S. Khan, R. M. Anwer, J. Weijer, A. D. Bagdanov, M. Vanrell, and A. M. Lopez. Color attributes for object detection. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3306–3313, 6 2012. Dostopno na: <http://www.micc.unifi.it/publications/2012/KAVBVL12>.
- [19] Igor Kononenko. *Strojno učenje*. Založba FE in FRI, Ljubljana, 2 edition, 2005.
- [20] Marko Škrjanec. Fruit image data set. Dostopno na: <http://www.vicos.si/Downloads/FIDS30>, 9 2013.
- [21] S. Lazebnik, C. Schmid, and J. Ponce. Spatial pyramid matching. Dostopno na: http://www.cs.unc.edu/~lazebnik/publications/pyramid_chapter.pdf.
- [22] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:2169–2178, 2006. Dostopno na: http://www-cvr.ai.uiuc.edu/ponce_grp/publication/paper/cvpr06b.pdf.
- [23] S. Lazebnik, A. Torralba, L. Fei-Fei, D. Lowe, and C. Szurka. Bag of words models. Dostopno na: http://cs.nyu.edu/~fergus/teaching/vision_2012/9_BoW.pdf, 3 2011.
- [24] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 1 2004. Dostopno na: <http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>.
- [25] O. Ludwig, D. Delgado, V. Goncalves, and U. Nunes. Trainable classifier-fusion schemes: an application to pedestrian

- detection. *IEEE Conference On Intelligent Transportation Systems*, 1:1–6, 10 2009. Dostopno na: http://www2.isr.uc.pt/~urbano/PDFs_PUB/Conferencias_Internacionais/2009/%5B113%5D2009ITSC09-129-FI.pdf.
- [26] Oswaldo Ludwig. Hog descriptor for matlab. Dostopno na: <http://www.mathworks.com/matlabcentral/fileexchange/28689-hog-descriptor-for-matlab>, 3 2011.
- [27] Sonka M., Hlavac V., and Boyle R. *Image Processing, Analysis, and Machine Vision*. Thompson Learning, 3 edition, 2008.
- [28] MathWorks. Matlab. Dostopno na: <http://www.mathworks.com/products/matlab/>.
- [29] Rick Rys. What is lab color space? Dostopno na: <http://www.hidefcolor.com/color-management/lab-color-space/>.
- [30] K. E. A. Sandle, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 9 2010. Dostopno na: <http://www.science.uva.nl/research/publications/2010/vandeSandeTPAMI2010/vandeSandeTPAMI2010-EvaluatingColorDescriptors.pdf>.
- [31] Robert E. Schapire. The boosting approach to machine learning an overview. 12 2001. Dostopno na: <https://www.cs.princeton.edu/courses/archive/spring07/cos424/papers/boosting-survey.pdf>.
- [32] W. C. Seng and S. H. Mirisae. A new method for fruits recognition system. *International Conference on Electrical Engineering and Informatics*, 01:130–134, 8 2009. Dostopno na: <http://www.mncc.com.my/tranx/Woo%20CS%20FruitRecognitionjournal.pdf>.
- [33] Stat soft. Svm transformation. Dostopno na: <http://www.statsoft.com/textbook/support-vector-machines/>.

-
- [34] Y. Song, C. A. Glasbey, G. W. Horgan, G. Polder, J. A. Dieleman, and G. W. A. M. Heijden. Automatic fruit recognition and counting from multiple images. *Biosystems Engineering*, 3 2013. Dostopno na: <http://www.bioss.ac.uk/people/chris/fruit.pdf>.
- [35] Abhijit Tembhekar. Red apple. Dostopno na: http://upload.wikimedia.org/wikipedia/commons/1/15/Red_Apple.jpg, 3 2009.
- [36] B. Tomasik, P. Thiha, and D. Turnbull. Bag of visual words approach. Dostopno na: <http://www.sccs.swarthmore.edu/users/09/btomasil/tagging-products.html>.
- [37] B. Tomasik, P. Thiha, and D. Turnbull. Tagging products using image classification. 2009. Dostopno na: <http://www.sccs.swarthmore.edu/users/09/btomasil/tagging-products.pdf>.
- [38] VLFEAT. Bag of words implementation. Dostopno na: <http://www.vlfeat.org/applications/apps.html>.
- [39] J. Yang, Y. Jiang, A. G. Hauptmann, and C. Ngo. Evaluating bag-of-visual-words representations in scene classification. *Proceedings of the international workshop on Workshop on multimedia information retrieval*, pages 197–206, 9 2007. Dostopno na: <http://lastlaugh.inf.cs.cmu.edu/alex/BagOfVisWords.MIR07.pdf>.
- [40] J. Zhao, J. Tow, and J. Katupitiya. On-tree fruit recognition using texture properties and color data. *Intelligent Robots and Systems*, pages 263–268, 8 2005. Dostopno na: http://automation.tkk.fi/attach/AS-84-2161/harj_omena.pdf.