

UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET

Aleksej Avramović

**NISKODIMENZIONALNI
PROSTORNO-TEKSTURALNI
DESKRIPTORI MULTISPEKTRALNIH
SLIKA**

doktorska disertacija

Beograd, 2016.

UNIVERSITY OF BELGRADE
FACULTY OF ELECTRICAL ENGINEERING

Aleksej Avramović

**LOW-DIMENSIONAL
SPATIAL-TEXTURAL DESCRIPTORS
OF MULTISPECTRAL IMAGES**

Doctoral Dissertation

Belgrade, 2016.

Mentor:

Dr Irini Reljin, redovni profesor, Univerzitet u Beogradu,
Elektrotehnički fakultet

Članovi komisije:

Dr Milan Bjelica, vanredni profesor, Univerzitet u Beogradu,
Elektrotehnički fakultet

Dr Zdenka Babić, redovni profesor, Univerzitet u Banjoj Luci,
Elektrotehnički fakultet

Dr Miodrag Popović, redovni profesor, Univerzitet u Beogradu,
Elektrotehnički fakultet

Dr Vladimir Risojević, docent, Univerzitet u Banjoj Luci,
Elektrotehnički fakultet

Datum odbrane:

Zahvalnica:

Istraživanje koje je sprovedeno tokom izrade ove disertacije je djelimično finansirano kroz naučno-istraživački projekat Ministarstva nauke i tehnologije Vlade Republike Srpske, ugovor broj 19/6-020/961-37/15.

Posebnu zahvalnost dugujem profesoru Branimiru Reljinu i profesorici Irini Reljin, koji su svojim mentorskim radom pomogli prilikom sprovođenja ovog istraživanja. Takođe, zahvalnost dugujem svim članovima komisije, kolegama i prijateljima na zajedničkom naučno-istraživačkom radu i brojnim korisnim diskusijama i prijedlozima.

Posebno hvala supruzi Milici na podršci i razumijevanju tokom pisanja ove disertacije.

NISKODIMENZIONALNI PROSTORNO-TEKSTURALNI DESKRIPTORI MULTISPEKTRALNIH SLIKA

REZIME:

Prepoznavanje vizuelnog i semantičkog sadržaja u slikama primjenom računarskih programa ima sve veći značaj u raznim granama privrede i industrije, medicini, vojnoj industriji, itd. Prepoznavanje sadržaja u slikama se u većini praktičnih aplikacija oslanja na metode obrade koje na osnovu numeričkih vrijednosti na digitalnim slikama određuju njihov sadržaj. U mnogim slučajevima važno je odrediti koliko je sadržaj dvije slike sličan, da li prikazuju isti objekat ili isti događaj. Sa druge strane, svjedoci smo da se razvojem moderne tehnologije nezaustavljivo povećava broj generisanih digitalnih slika. Savremeni klinički centri opremljeni digitalnom radiologijom, dnevno generišu i do nekoliko desetaka hiljada novih snimaka. Manuelno opisivanje sadržaja tako velikog broja slika predstavlja praktičan problem. Takođe, svakodnevno dobijamo veliku količinu podataka snimljenih tehnikama daljinske detekcije, pri čemu specifične aplikacije zahtjevaju brzo prepoznavanje sadržaja takvih snimaka.

Potreba za prepoznavanjem vizuelnog i semantičkog sadržaja u slikama dovela je do razvoja velikog broja pristupa za opisivanje tog sadržaja na način pogodan za korišćenje u računarskim sistemima. Često se slikama pridružuju odgovarajući deskriptori koji treba da "opišu" sadržaj u slikama. Ti deskriptori su vektori numeričkih vrijednosti ili skup ključnih riječi, koji treba da budu određeni tako da se pomoću njih mogu razlikovati slike različitog vizuelnog ili semantičkog sadržaja ili prepoznati slike sličnog sadržaja. Pošto ljudski vizuelni sistem efikasno koristi informacije o teksturi za prepoznavanje objekata, u praktičnim aplikacijama se često koriste deskriptori teksture.

Razvoj tehnologije omogućio je upotrebu jeftinih multispektralnih kamera, pa se postavlja pitanje kako opisati sadržaj slika sa većim brojem spektralnih opsega. Jednostavno proširivanje deskriptora i upotreba dodatnih podataka može poslužiti da se na odgovarajući način opiše sadržaj multispektralnih slika, ali sa značajnim povećanjem potrebnih memorijskih resursa i računarske kompleksnosti.

U ovoj disertaciji su predložene su metode za izdvajanje niskodimenzionalnih deskriptora multispektralnih slika, pogodnih za automatsku klasifikaciju slika. Takođe, razmotreni su pristupi za uključivanje podataka o prostornom rasporedu lokalnih obilježja na slikama u deskriptor, kako bi se povećala tačnost klasifikacije. Na kraju, predložena je nova metoda za izdvajanje niskodimenzionalnih prostorno-teksturalnih deskriptora za multispektralne slike. Ova metoda koristi postojeće deskriptore tek-

sture sivih slika, računate prema prostornom rasporedu lokalnih obilježja teksture, u kombinaciji sa tehnikama redukcije dimenzionalnosti kako bi se izračunali niskodimenzionalni deskriptori za multispektralne slike.

Postavljen je i izvršen veliki broj eksperimenata, pri čemu su analizirani uticaji pojedinačnih koraka i vrijednosti određenih parametara prilikom automatske klasifikacije sivih slika, slika u boji i multispektralnih slika. Poslije svakog seta eksperimenata izvedeni su zaključci koji su korišćeni za dalja razmatranja problema i postavljanje narednih eksperimenata. Na kraju, predstavljen je i način za klasifikaciju aero snimaka pomoću konvolucionih neuronskih mreža, koji u slučaju korektnog obučavanja značajno popravља tačnost klasifikacije. Radi ilustracije mogućnosti primjene u praktičnim aplikacijama, predstavljena je metoda za automatsku detekciju objekata na aero snimcima visoke rezolucije primjenom konvolucionih neuronskih mreža.

Ključne riječi: Tekstura, deskriptori, multispektralne slike, klasifikacija slika, neuronske mreže, detekcija objekata.

Naučna oblast: Elektronika i računarstvo

Uža naučna oblast: Telekomunikacije

UDK: 621.3

LOW-DIMENSIONAL SPATIAL-TEXTURAL DESCRIPTORS OF MULTISPECTRAL IMAGES

ABSTRACT:

Recognition of visual and semantic content on images using computer programs gained an importance in various fields of agriculture, industry, medicine, military industry etc. Most practical applications use certain methods based on the numerical value of the digital images to determine what is a content of those images. In many cases, it is important to determine a level of visual or semantic similarity between two different images, does two images showing the same object or maybe the same event. We are witnessing that development of modern technology cause unstoppable increase of the number of daily generated digital images. Modern clinical centers, equipped with digital radiology, generate up to tens of thousands of new images per day, so their manual annotation presents a practical problem. Moreover, each day brings a large amount of remotely sensed images and many specific applications require fast identification of their visual content.

The need to recognize visual and semantic content in images initiated the development of a large number of methods for description of image contents, in such a manner suitable for use in specific computer systems. Images are associated with appropriate descriptors that should “describe” the visual or semantic content of those images. These descriptors can be vectors with numerical values, which should be calculated so it is possible to use them to distinguish between images with different visual or semantic content or to recognize images with similar content. Since the human visual system effectively relies on texture to identify objects, texture descriptors are often used in practical applications.

Technology development enable the widespread usage of cheap multispectral cameras, which can capture the data beyond visible spectra. Thus, it is necessary to investigate how to represent and describe the content of multispectral images in the way suitable for practical applications based on image classification. Simple extension of descriptors can increase classification accuracy, but with the cost of more memory resources and computational complexity.

In this dissertation, different methods for extraction of low-dimensional descriptors for multispectral images are proposed, which used for automatic image classification. Moreover, the usage of spatial position of local textural features is discussed as well. It was concluded that extension of texture descriptor of grayscale images with additional data providing spatial-based texture features, can increase classification accuracy. Based on obtained results, a novel method for extraction of

low-dimensional spatial-textural descriptor for multispectral images was proposed. This method is based on existing texture descriptors for grayscale images, expanded with spatial data of local features and combines them with dimensionality reduction techniques in order to calculate low-dimensional descriptors for multispectral images.

A large amount of experiments is set up, to analyze contribution of various steps of automatic classification, as well of different values of parameters. Experiments used grayscale, color and multispectral images. After each experimental set, useful conclusions are made and used in next set of experiments. Step by step, a novel descriptor extraction method is developed. In addition, the usage of convolutional neural networks for automatic image classification and object detection on high resolution aerial images, is discussed.

Keywords: Texture, descriptors, multispectral images, image classification, neural networks, object detection.

Scientific field: Electrical engineering and computer science

Specified scientific field: Telecommunications

UDK: 621.3

Sadržaj

Spisak slika	xi
Spisak tabela	xiii
1 Uvod	1
1.1. Značaj teksture za opisivanje sadržaja slike	2
1.2. Klasifikacija slika, prepoznavanje objekata i detekcija događaja	5
1.3. Tema i doprinos disertacije	7
1.4. Organizacija disertacije	8
2 Automatska klasifikacija slika	10
2.1. Reprezentacija digitalnih slika	10
2.2. Detekcija obilježja, izdvajanje vektora obilježja i deskriptora	11
2.3. Deskriptori teksture	13
2.4. Automatska klasifikacija slika	17
2.5. Redukcija dimenzionalnosti deskriptora	20
2.6. Pregled literature	22
2.7. Istraživački trendovi	25
3 Metodologija istraživanja	26
3.1. Korišćene baze slika	26
3.2. Metodologija nadgledane automatske klasifikacije slika	31
3.3. Problem vizuelne sličnosti slika i tačnosti klasifikacije	37
4 Automatska klasifikacija sivih slika prema teksturi	38
4.1. Postavka eksperimenata	39
4.2. Rezultati automatske klasifikacije sivih slika	40
4.3. Poređenje performansi deskriptora	44
4.4. Uticaj parametara klasifikatora	46
4.5. Zaključci korisni za naredne eksperimente	47
5 Niskodimenzionalni multispektralni deskriptori teksture	49
5.1. Postavka eksperimenata	49
5.2. Rezultati automatske klasifikacije slika u boji i multispektralnih slika	50
5.3. Predložena rješenja za kombinovanje deskriptora različitih kanala	53
5.4. Rezultati klasifikacija multispektralnih slika predloženim deskriptorima	57
5.5. Poređenje rezultata klasifikacije i diskusija	61
5.6. Mogućnosti poboljšanja tačnosti klasifikacije	63

5.7. Zaključci važni za klasifikaciju multispektralnih slika	65
6 Prostorni raspored lokalnih obilježja tekstone	67
6.1. Predloženo rješenje	69
6.2. Postavka eksperimenta	70
6.3. Klasifikacija prostorno-teksturalnim deskriptorima	72
6.4. Redukcija dimenzionalnosti prostorno-teksturalnih deskriptora	76
6.5. Komparativna analiza rezultata	78
6.6. Zaključci o klasifikaciji prostorno-teksturalnim deskriptorima	80
7 Niskodimenzionalni prostorno-teksturalni deskriptori	81
7.1. Predloženo rješenje	81
7.2. Analiza performansi predloženog deskriptora	82
7.3. Komparativna analiza rezultata	84
7.4. Zaključci o klasifikaciji niskodimenzionalnim prostorno-teksturalnim deskriptorima	88
8 Automatska klasifikacija slika konvolucionim neuronskim mrežama	89
8.1. Predloženo rješenje za obučavanje konvolucione neuronske mreže	90
8.2. Automatska detekcija objekata na aero snimcima visoke rezolucije	91
8.3. Zaključci o klasifikaciji konvolucionim neuronskim mrežama	94
9 Zaključak	95
Literatura	97

Spisak slika

1.1.	Primjeri jednostavnih tekstura	3
1.2.	Primjeri složenih tekstura	3
1.3.	Teksture na različitim scenama	4
2.1.	Primjeri sivih slika sa istim histogramima boja a različitim teksturama	13
3.1.	Primjeri slika iz baze UIUC	27
3.2.	Primjeri slika iz baze UMD	28
3.3.	Primjeri slika iz baze UC Merced	29
3.4.	Primjeri slika iz baze NirScene	31
3.5.	Korišćenje rasporeda lokalnih obilježja prilikom računanja deskriptora	34
4.1.	Šematski prikaz algoritma za automatsku klasifikaciju sivih slika	39
4.2.	Analiza uticaja normalizacije deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika	41
4.3.	Analiza uticaja skaliranja deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika	43
4.4.	Analiza uticaja redukcije dimenzionalnosti deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika	44
5.1.	Šematski prikaz algoritma za automatsku klasifikaciju multispektralnih slika	50
5.2.	Analiza uticaja različitih prostora boja na tačnost klasifikacije, za različite deskriptore i baze slika u boji i multispektralnih slika	52
5.3.	Šematski prikaz algoritma za računanje deskriptora teksture za multispektralne slike (<i>DescriptorConc</i>)	54
5.4.	Šematski prikaz algoritma za računanje deskriptora teksture za multispektralne slike (<i>DescriptorBand</i>)	55
5.5.	Analiza uticaja dimenzionalnosti multispektralnih deskriptora na tačnost klasifikacije	60
6.1.	Primjeri različitih prostornih raspodjela tekstura na aero snimcima i odgovarajući prostorni rasporedi regiona	68
6.2.	Primjeri različitih prostornih raspodjela tekstura na aero snimcima i odgovarajući prostorni rasporedi regiona	69
6.3.	Šematski prikaz algoritma za računanje deskriptora teksture koji uključuje podatke o prostornom rasporedu lokalnih obilježja	70
6.4.	Načini podjele prostora slike i slučajevi predloženog algoritma za računanje prostorno-teksturalnog deskriptora	71

6.4.	Analiza uticaja načina normalizacije prostorno-teksturalnih deskriptora sivih slika na tačnost klasifikacije različite deskriptore i baze sivih slika	74
6.5.	Uticaj dodatnog histograma centralnog regiona kod prostorno-teksturalnih deskriptora sivih slika na tačnost klasifikacije za UCM bazu sivih slika	75
6.6.	Poređenje tačnosti klasifikacije za prostorno-teksturalne deskriptore sivih slika redukovane dimenzionalnosti	77
7.1.	Šematski prikaz algoritma za računanje deskriptora teksture za multi-spektralne slike (<i>DescriptorReg</i>)	83
7.2.	Poređenje tačnosti klasifikacije različitih niskodimenzionalnih deskriptora multispektralnih slika	85
8.1.	Primjeri aero snimaka visoke rezolucije	92
8.2.	Rezultati predloženog metoda za automatsku detekciju objekata na slikama visoke rezolucije	93

Spisak tabela

4.1. Poređenje tačnosti klasifikacije u slučajevima kada se deskriptori računaju i procesiraju na različite načine, za UIUC bazu	45
4.2. Poređenje tačnosti klasifikacije u slučajevima kada se deskriptori računaju i procesiraju na različite načine, za UMD bazu	45
4.3. Poređenje tačnosti klasifikacije u slučajevima kada koriste različiti kerneli SVM klasifikatora	46
5.1. Poređenje tačnosti klasifikacije multispektralnih slika u slučajevima kada se deskriptori računaju i procesiraju na različite načine. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama prosječni vremenski intervali potrebno za izvršavanje jedne klasifikacije.	62
5.2. Poređenje tačnosti klasifikacije multispektralnih slika SIFT deskriptorom, za različite veličine kodnih knjiga. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama standardne devijacije za 5 različitih trening/test podjela	64
6.1. Poređenje tačnosti klasifikacije sivih slika u slučajevima kada se koriste obični deskriptori sivih slika i prostorno-teksturalni deskriptori sivih slika. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama dimenzionalnosti deskriptora.	79
7.1. Poređenje tačnosti klasifikacije slika u boji u slučajevima kada se koriste predloženi niskodimenzionalni prostorno-teksturalni deskriptori. Brojevi predstavljaju tačnosti klasifikacije, a dimenzionalnosti deskriptora su date u zagradama.	86
7.2. Poređenje tačnosti klasifikacije multispektralnih slika u boji u slučajevima kada se koriste predloženi niskodimenzionalni prostorno-teksturalni deskriptori. Brojevi predstavljaju tačnosti klasifikacije, a dimenzionalnosti deskriptora su date u zagradama	87
8.1. Poređenje tačnosti klasifikacije za UCM bazu aero snimaka, za metode zasnovane na deskriptorima i metode zasnovane na konvolucionim neuronskim mrežama.	91

1 | Uvod

Značajan naučni napredak i uvođenje novih tehnoloških mogućnosti omogućili su generisanje velikog broja digitalnih slika u različitim oblastima teoretskih i praktičnih primjena, kao što su edukacija, ispitivanje na daljinu, privreda, medicina, robotika i automatika, vojna industrija, prostorno planiranje, urbanizam, hidrometeorologija, nadgledanje poljoprivrednih dobara, nadgledanje saobraćaja, nadgledanje proizvodnih procesa, itd. Obrada digitalnih slika se primjenjuje u svim navedenim oblastima pri čemu se omogućava automatizacija proizvodnih procesa, olakšava nadgledanje, planiranje ili upravljanje i povećava proizvodna efikasnost. U navedenim oblastima primjena, često je potrebno izvršiti analizu digitalnih slika ili video zapisa na način da se porede sadržaji dvije digitalne slike ili dva frejma iz videa, prema nekom vizuelnom ili semantičkom kriterijumu. U idealnom slučaju, digitalne slike treba opisati na način tako da njihov opis sadrži sve potrebne informacije o vizuelnom sadržaju i semantici sadržaja u slikama, tj. o scenama, objektima i/ili događajima koji se prikazuju [1]. Pošto se frejmovi u video zapisima mogu tretirati kao posebne slike, bez gubitka opštosti, u ovoj disertaciji će biti razmatran problem analize i klasifikacije digitalnih slika.

Mnogi savremeni proizvodni procesi su automatizovani na način da se upravljanje robotskim pomagalicama oslanja na analizu sadržaja slika snimljenih prilikom proizvodnog procesa. Nadgledanje poljoprivrednih dobara se može izvršiti automatski na daljinu, tako što se analiziraju uzastopni snimci i traže se razlike između njih, kao što su povećanje prisustva određene boje ili promjene u indeksu vegetacije i sl. Na taj način se može pratiti razvoj usjeva i automatski izvršiti detekcija pojave određenih bolesti, pratiti efekti suša i poplava ili vršiti sigurnosni nadzor. Oblasti primjene kao što su prostorno planiranje i urbanizam, hidrometeorologija i nadzor privrednih dobara, koriste veliku količinu digitalnih slika snimljenih na daljinu, koje se obrađuju na pogodan način da bi se jednostavnije uočile posmatrane promjene ili karakteristike. U ovim slučajevima, algoritmi za obradu digitalnih slika uključuju detekciju objekata ili procjenu kretanja objekata, što se često izvodi poređenjem vizuelnog sadržaja između više slika.

Prethodnih godina došlo je do naglog porasta broja generisanih digitalnih slika u oblasti medicine, naročito razvojem digitalne radiologije. Bolničke baze poda-

taka mogu da sadrže i do nekoliko stotina hiljada medicinskih slika, pa je korisno implementirati algoritme za pretraživanje baze koji omogućavaju nalaženje sličnih kliničkih slučajeva, npr. nalaženje svih rendgenskih snimaka koji prikazuju prelom ruke ili svih sličnih histopatoloških nalaza. Drugim riječima, potrebno je pretražiti bazu medicinskih slika na osnovu semantike njihovog sadržaja. Druga važna primjena obrade digitalnih slika u oblasti medicine je pomoć prilikom postavljanja dijagnoze¹, što uključuje sve algoritme digitalne obrade medicinskih slika koji se koriste za lakše uočavanje anomalija u medicinskim slikama različitih modaliteta, pretraživanje baza medicinskih slika i automatizovanje procesa postavljanja dijagnoze [2].

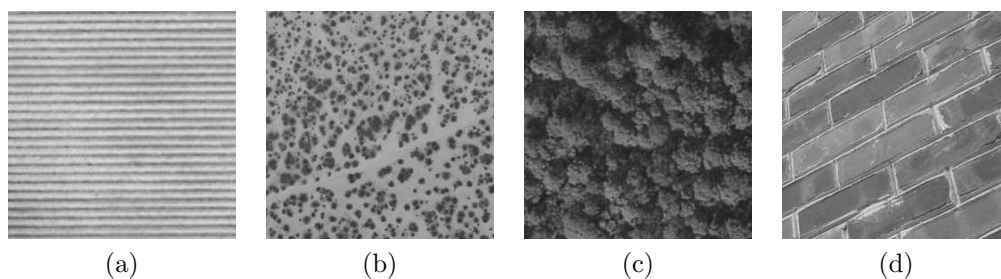
Generalno, u mnogim praktičnim primjenama potrebno je izvršiti analizu digitalnih slika sa ciljem prepoznavanja objekata na slici, prepoznavanja scena, prepoznavanja nekog određenog događaja od interesa, nalaženja razlike između dvije slike ili automatske anotacije slika. Anotacija digitalnih slika predstavlja opisivanje njihovog sadržaja na pogodan način, tako da se taj opis može iskoristiti za praktične primjene. Jedan pristup predstavlja manuelna anotacija na osnovu vizuelnog pregleda korisnika. Tom prilikom se svakoj slici pridružuje spisak ključnih riječi ili odgovarajući tekst, koji treba da pruže što je moguće više korisnih informacija o vizuelnom sadržaju ili semantici sadržaja koji je prikazan u slici. U slučaju da anotaciju vrši mali broj obučениh anotatora, može se postići kvalitetan i jednoznačan opis svake slike iz baze. Međutim, manuelna anotacija nije praktična za obradu velikih baza slika iz razloga što je potrebno pregledati i opisati veliki broj slika, pri čemu je moguće izgubiti jednoznačnost opisa uz veliki utrošak vremena. Pored toga, određena vizuelna obilježja na slikama je teško jednoznačno opisati riječima.

Automatska anotacija digitalnih slika podrazumijeva primjenu odgovarajućih algoritama za obradu digitalnih slika, tako da se na osnovu njihovog sadržaja (numeričkih vrijednosti) izvrši odgovarajući opis scene, objekta ili događaja koji se prikazuje. Taj opis može biti niz ključnih riječi koje treba da opišu vizuelni sadržaj ili semantiku sadržaja slike ili vektor koji sadrži numeričke vrijednosti. Pogodnom interpretacijom ključnih riječi ili vektora numeričkih vrijednosti mogu se realizovati praktični algoritmi za klasifikaciju slika, detekciju događaja ili objekata na slikama.

1.1. Značaj teksture za opisivanje sadržaja slike

Analizom načina na koji ljudski vizuelni sistem prepoznaje scene i objekte, uočeno je da tekstura slika nosi značajnu količinu informacija potrebnih za prepoznavanje. Reninger i Malik su analizom rezultata identifikacije scene predstavljenih u [3] zaključili da od velike pomoći za tačnost identifikacije može biti jednostavna analiza teksture. Iako ne postoji formalna definicija teksture, može se reći da tekstura opisuje karakteristike površi ili izgleda nekog objekta na osnovu veličine, oblika, gustine, rasporeda i orijentacije osnovnih dijelova površi posmatranog objekta. Prilikom opisivanja teksture, često se koriste pridjevi: glatka ili gruba, meka ili tvrda, hrapava ili fina, mat ili sjajna, orijentisana ili neorijentisana, regularna ili neregularna, itd. Prilikom opisivanja sadržaja slike pomoću teksture, različitim osobinama

¹ *Computer-Aided Diagnosis* (eng.)

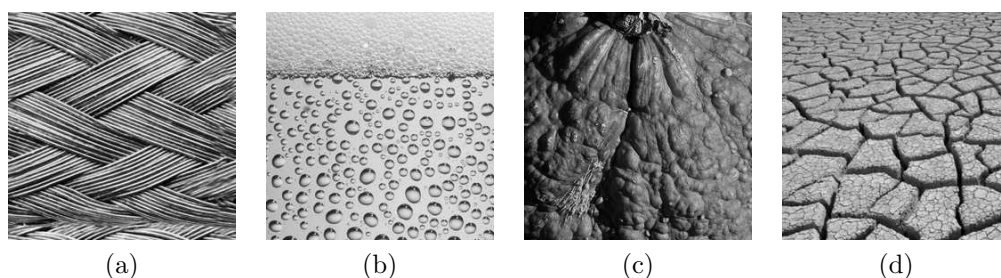


Slika 1.1: Primjeri jednostavnih tekstura

teksture mogu se pridružiti numeričke vrijednosti, koje mogu poslužiti za kvantitativni opis sadržaja slika i određivanje razlike ili sličnosti između pojedinih slika. Posmatrajući objekat ili prikaz na slici sa različitim udaljenosti, navedeni pridjevi mogu biti manje ili više izraženi, pa opisivanje teksture zavisi i od skale na kojoj se ona posmatra. Zbog dobrih deskriptivnih i diskriminativnih svojstava, tekstura ima široku praktičnu primjenu kod automatske klasifikacije slika dobijenih snimanjem iz vazduha, prilikom analize biomedicinskih slika prilikom klasifikacije prirodnih scena, detekcije događaja i prilikom prepoznavanja objekata.

U određenim slučajevima, moguće je precizno opisati teksturu pomoću prethodno navedenih pridjeva. Posmatrajući primjer sa Sl. 1.1a vidimo da su brazde paralelne i horizontalne, pa je usmjerenost dominantna karakteristika date teksture. Tekstura prikazana na Sl. 1.1b nije usmjerena, već se sastoji od granula koje nemaju izraženu grubost i neravnomjerno su raspoređene po cijeloj slici, pa je možemo opisati kao hrapavu i relativno finu. Slično, tekstura prikazana na Sl. 1.1c je hrapava, ali je znatno grublja u odnosu na teksturu prikazanu na Sl. 1.1b. Jasne karakteristike materijala prikazanog na Sl. 1.1d daju osjećaj čvrstine objekta koji se prikazuje i nedvosmisleno ukazuju na semantiku sadržaja slike. U ovom slučaju tekstura se može opisati kao tvrda i djelimično usmjerena.

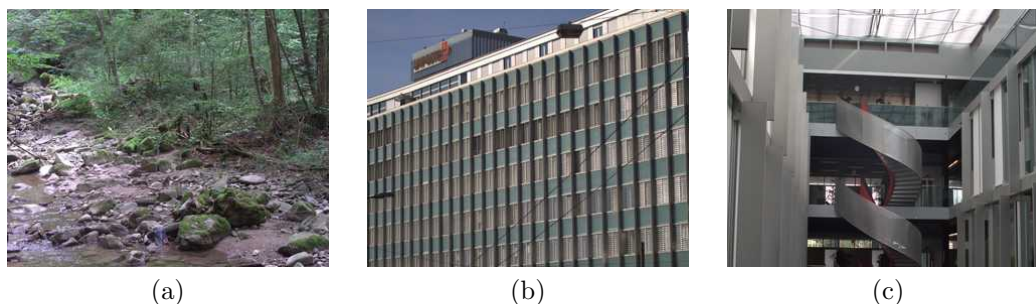
Na prethodnim primjerima cijele slike prikazuju određenu teksturu koja se može opisati sa nekoliko pridjeva. Međutim, u prirodi se češće susrećemo sa objektima i predmetima čija je tekstura složenija, kao što je prikazano na primjerima sa Sl. 1.2. Na Sl. 1.2a prikazana je izbrazdana i isprepletana tekstura, na kojoj vidimo da se različiti dijelovi nepravilnog oblika mogu karakterisati različitim dominantnim orijentacijama. Iako ne postoji jedinstvena orijentacija teksture za cijelu sliku, karakteristike materijala daju osjećaj čvrstine. Tekstura prikazana na Sl. 1.2b karakterisana je mjehurićima, čija je gustina i veličina različita na različitim dijelovima slike. Ka-



Slika 1.2: Primjeri složenih tekstura

rakteristike materijala odgovaraju prikazu gaziranog pića sa pjenom na vrhu. Dalje, susrećemo se sa teksturama koje imaju neravnomjeran raspored svojih lokalnih karakteristika i koje je teže opisati jednostavnim pridjevima, kao što je prikazano na primjerima sa Sl. 1.2c i Sl. 1.2d. Na ovim slikama su redom prikazane neravnomjerno naborane i ispucale površi, pri čemu je veličina lokalnih karakteristika (bora, brazda i ispupčenja) različita na različitim dijelovima slika. Može se zaključiti da je skala lokalnih karakteristika promjenljiva. Sa prethodnih primjera vidimo da se prikazane površi mogu posmatrati iz različitih tačaka gledišta, što takođe može da utiče na raspored ivica i način opisivanja lokalnih karakteristika tekstura.

Posmatrajući primjere sa Sl. 1.1 i 1.2, u većini slučajeva moguće je zaključiti koji objekat je prikazan, tj. moguće je karakteristike objekata i scena poistvjetiti sa karakteristikama tekstura. Teksture prikazane na Sl. 1.2 prikazuju, tkaninu, gazirano piće, dio bundeve i isušenu zemljanu površ, redom. Za razlikovanje tekstura, posebno su značajni oblik, veličina i raspored kontura i ivica na teksturi. Međutim, na jednoj slici često postoje regioni karakterisani različitim teksturama, tj. regioni koji se odnose na različite objekte.



Slika 1.3: Teksture na različitim scenama

Mogućnosti korišćenja teksture prilikom klasifikacije prirodnih slika u boji mogu se razmotriti analizom primjera datih na Sl. 1.3. Na Sl. 1.3a prikazana je prirodna scena koja sadrži regione šumskog zelenila, kamenja i potoka, čije teksture se razlikuju po boji, skali, gustini i orijentaciji ivica. U ovom primjeru, prikazana je složena scena, koju nije moguće opisati jednostavnim pridjevima, već sadrži vizuelni sadržaj složene semantike. Scena na ovoj slici može se pridružiti klasama koje prikazuju šumu i rijeku ili generalno prikaz scene na otvorenom. Scena data na Sl. 1.3b prikazuje stambenu zgradu iznad koje se, na manjem dijelu slike, vidi nebo. Fasada stambene zgrade je jednolika i sadrži prozore koji se ponavljaju na način tako da obrazuju uniformnu teksturu. Ta tekstura je periodična sa ravnomjernim vertikalnim i horizontalnim orijentacijama ivica. Primjer sa Sl. 1.3c je nešto drugačiji jer pokazuje unutrašnjost stambenog objekta koja se karakteriše pravilnim rasporedom ivica koje razgraničavaju površi različite teksture. Na svakom od primjera sa Sl. 1.3 mogu se uočiti regioni sa različitom teksturom, od kojih svaka definiše karakteristike objekta koji se prikazuje. Može se zaključiti da je analiza teksture ključna prilikom opisivanja objekata koji su prikazani na sceni.

Intuitivno je jasno da slike u boji (Sl. 1.3) nose više informacija nego sive slike (Sl. 1.1 i Sl. 1.2), ali se iz navedenih primjera može primjetiti važnost teksture prilikom prepoznavanja scena ili objekata na slikama. Posebno je važno uočiti kakav je

raspored lokalnih karakteristika teksture, kao što su ivice, konture, brazde, izbočine, itd. i način na koji su te karakteristike raspoređene na slikama. Raspored lokalnih karakteristika određuje teksturu, a različite teksture na slici obično odgovaraju različitim objektima ili površima. Pored navedenih lokalnih karakteristika teksture, slike u boji sadrže više vizuelnih informacija. Slične teksture mogu biti različitih boja, pri čemu definišu različite objekte ili dijelove scene. Posmatrajući varijante sivih slika koje prikazuju takve teksture, one mogu biti vizuelno identične, što nije moguće primjetiti ukoliko se ne posmatraju slike u boji. Povećanjem broja korišćenih kanala povećava se količina podataka kojom se mogu opisati teksture. Slično kao što su analizirane teksture slika u prethodnim primjerima (Sl. 1.1, 1.2 i 1.3), mogu se analizirati slike različitih karakteristika: medicinske slike, aero snimci, fotografije, grafikoni, itd.

Zbog izuzetnog značaja prilikom opisivanja sadržaja na slikama, tekstura se može iskoristiti prilikom automatske klasifikacije slika. Primjenom različitih algoritama za obradu digitalnih slika moguće je izračunati numeričke reprezentacije koje opisuju raspored i lokalne karakteristike teksture, kako bi se omogućilo poređenje sličnosti tekstura sa različitih slika ili tekstura sa različitih regiona iste slike. Prethodno opisan način analize lokalnih karakteristika tekstura je često korišćen za razvoj i implementaciju konkretnih algoritama za deskripciju sadržaja slika [4, 5, 6, 7, 8, 9, 10, 11] koji mogu biti prilagođeni konkretnim zadacima automatske klasifikacije slika, prepoznavanja objekata ili detekcije događaja.

1.2. Klasifikacija slika, prepoznavanje objekata i detekcija događaja

Klasifikacija slika, prepoznavanje objekata ili detekcija događaja su slični problemi koja se oslanjaju na grupisanje slika u klase na osnovu njihovog sadržaja. Npr., baza digitalnih medicinskih slika može da sadrži slike velikog broja različitih modaliteta: rentgen, magnetska rezonansa, mamografija, ultrazvuk, radiografija, angiografija, itd. Klasifikacija medicinskih slika prema modalitetu podrazumijeva sprovođenje određene procedure kojom se, na osnovu vizuelnog sadržaja ili semantičke sadržaja slika, vrši grupisanje svih slika iz baze u klase definisane modalitetima. Klase mogu biti definisane i na druge načine, npr. anatomskim regionom, patologijom ili određenom dijagnozom. Prepoznavanje objekata ili događaja podrazumijeva sprovođenje procedure kojom se određuje da li je na slici ili na frejmu videa prikazan neki konkretan objekat ili događaj. Ukoliko je potrebno prepoznati neki objekat ili događaj, koriste se samo dvije klase i koristi se procedura koja slike koje prikazuju objekat ili događaj svrstavaju u jednu klasu, dok se ostale slike svrstavaju u drugu klasu. Moguće je definisati više klasa kako bi se istovremeno izvršilo prepoznavanje više različitih objekata ili događaja. Detekcija objekata ili događaja predstavlja pretraživanje sadržaja slika i traženja određenih karakteristika koje definišu određeni objekat ili događaj pri čemu nije poznato da li se tražene karakteristike, tj. objekat ili događaj, uopšte nalaze u bazi.

Generalno, klasifikacija slika predstavlja njihovo grupisanje u klase na osnovu nekog vizuelnog ili semantičkog kriterijuma. Prepoznavanje objekata/događaja je ta-

kva klasifikacija pri kojoj je barem jedna klasa definisana tako da sadrži sve slike koje prikazuju objekat/događaj koji se prepoznaje, dok se detekcija objekata/događaja odnosi na algoritme za pretraživanje baza slika pri čemu nije poznato da li je traženi objekat/događaj prikazan na bilo kojoj slici iz baze. Obično se u literaturi posebno razmatraju problemi klasifikacije slika, prepoznavanja objekata i detekcije događaja, u zavisnosti od konkretne primjene na koju se razmatrani algoritmi odnose. Svaki od ovih problema podrazumijeva potrebu za posebnim obilježavanjem određene grupe slika iz baze, prema njihovoj vizuelnoj ili semantičkoj sličnosti.

Automatska klasifikacija digitalnih slika predstavlja izvršavanje računarskog programa pri čemu se svakoj od slika iz baze vrši dodjeljivanje odgovarajućih oznaka klase svakoj od slika iz baze. Automatska klasifikacija se izvršava poređenjem obilježja koja se izračunavaju za svaku sliku posebno. Ta obilježja mogu biti numeričke vrijednosti koje opisuju karakteristike tekstura, boje, oblika objekata prikazanih na slikama ili atributi koji opisuju sadržaj prikazan u slikama pomoću ključnih riječi. Dakle, prvo se izvršavaju algoritmi koji vrše obradu digitalnih slika na osnovu njihovog sadržaja i pridružuju im obilježja. Skupovi obilježja mogu biti predstavljeni vektorima obilježja ili deskriptorima. Izdvajanje vektora obilježja predstavlja računanje vektora sa numeričkim vrijednostima koje zavise od sadržaja u slikama ili određenim regionima slika. Deskriptori, pored numeričkih vrijednosti, sadržaj u slikama mogu opisivati i drugim atributima. Možemo usvojiti da vektori obilježja predstavljaju numerički opis vizuelnih obilježja sadržaja slike, dok deskriptori pored numeričkih vrijednosti mogu koristiti i ključne riječi prilikom opisivanja sadržaja u slikama. Često se za reprezentaciju sadržaja u slikama koriste isključivo numeričke vrijednosti u kojem slučaju vektori obilježja i deskriptori predstavljaju sinonime.

U ovoj disertaciji će se koristiti pojam deskriptor koji će označavati vektore obilježja ili njihovu kombinaciju. Npr., deskriptori slika mogu biti jednostavni vektori obilježja koji samo sadrže informaciju o tome koja boja je dominantna na slici, koje boje su prisutne na slici, ili koja je boja i položaj određenog piksela na slici. Takođe, deskriptori mogu predstavljati rezultat složenih operacija na vektorima obilježja tako da daju informaciju prostornom rasporedu složenih lokalnih obilježja u slikama ili da sadrže skup ključnih riječi koje opisuju semantiku sadržaja scene prikazane na slici. U ovoj disertaciji, razmatraju se isključivo deskriptori sa numeričkim obilježjima slika, pa je svaki deskriptor zapravo vektor obilježja i obratno.

Zbog velikih mogućnosti za praktičnu primjenu automatske klasifikacije slika, algoritmi za izdvajanje vektora obilježja i računanje deskriptora treba da daju takve numeričke vrijednosti, iz kojih se može zaključiti da li su dvije slike sličnog vizuelnog sadržaja i/ili semantike sadržaja ili nisu. Želja je računanje takvih deskriptora, čijim poređenjem je moguće izračunati koliko su dvije slike slične ili različite prema nekom određenom kriterijumu. U nekim slučajevima, izdvajanje obilježja i računanje deskriptora se može posmatrati kao procedura redukcije dimenzionalnosti ulaznih podataka, međutim pojedini deskriptori mogu imati mnogo veću dimenzionalnost nego slike čiji vizuelni sadržaj se opisuje. Posmatrano sa aspekta automatske klasifikacije slika, izdvajanje obilježja i računanje deskriptora je transformacija nad digitalnim slikama koja daje izlazni vektor numeričkih vrijednosti.

U daljem radu će biti razmatran problem izdvajanja obilježja sa ciljem da se dobije deskriptor što je moguće niže dimenzionalnosti, tako da uključuje što više

vizuelnih i semantičkih informacija potrebnih za tačniju automatsku klasifikaciju digitalnih slika ili tačnije prepoznavanje i detekciju objekata/događaja. U tu svrhu, biće predstavljeni rezultati eksperimenata nad nekoliko semantički različitih baza slika iz kojih su deskriptori računati na različite načine.

1.3. Tema i doprinos disertacije

Pored toga što tekstura sivih slika ima značajna diskriminativna svojstva, mnogo više informacija se može izdvojiti posmatrajući slike u boji i multispektralne slike. Takođe, od značaja mogu biti informacije o prostornom rasporedu lokalnih obilježja teksture. Uzimajući u obzir informacije o lokalnim obilježjima teksture i njihovom prostornom rasporedu, iz svih spektralnih opsega posebno, može se dobiti deskriptor slike visoke dimenzionalnosti. Pored toga što visoka dimenzionalnost deskriptora može značajno usporiti klasifikaciju, povećava se vjerovatnoća pojave šuma i prikrivanja informacija u samim deskriptorima, što može uzrokovati smanjenje tačnosti klasifikacije.

Tema ove disertacije jeste razmatranje mogućnosti efikasnog izdvajanja deskriptora teksture niske dimenzionalnosti koji sadrže informacije o prostornom rasporedu lokalnih obilježja iz multispektralnih slika. Na osnovu razmatranja efekata metoda za redukciju dimenzionalnosti deskriptora teksture, izdvajanja deskriptora teksture iz multispektralnih slika i korišćenja informacija o prostornom rasporedu lokalnih obilježja teksture, predložen je jedinstven algoritam za izdvajanje kompaktnog deskriptora teksture niske dimenzionalnosti iz multispektralnih slika. U tu svrhu predstavljeni su rezultati automatske klasifikacije baza slika tekstura, aero snimaka i prirodnih scena, u slučajevima kada se koristi različit broj spektralnih opsega. Cilj je da se ispita kakva je tačnost automatske klasifikacije korišćenjem predloženih tehnika za izdvajanje niskodimenzionalnih deskriptora i da se pokaže očuvanje ili poboljšanje tačnosti u poređenju sa slučajevima kada se koriste deskriptori visoke dimenzionalnosti.

Dodatno su razmotrene mogućnosti klasifikacije aero snimaka primjenom konvolucionih neuronskih mreža, koje ukoliko se obuče na odgovarajući način mogu značajno popraviti tačnost klasifikacije. Pored automatske klasifikacije, mogućnosti primjene konvolucionih neuronskih mreža su demonstrirane na problemu automatske detekcije objekata na aero snimcima visoke rezolucije.

Prilikom sprovođenja eksperimenata, korišćene su baze sivih slika (slike koje sadrže jedan kanal), baze slika boji (slike koje sadrže tri kanala iz vidljivog dijela spektra) i baze multispektralnih slika (slike koje sadrže više od tri kanala). Na osnovu istraživanja i rezultata eksperimenata koji će biti izloženi u daljem tekstu, doprinosi disertacije se mogu sažeti na sljedeći način:

- (1) Pregled koncepata za praktičnu realizaciju algoritama za automatsku klasifikaciju slika i prepoznavanja scena, kao i usvajanje metoda za analizu efikasnosti klasifikacije,
- (2) analiza značaja teksture prilikom opisivanja objekata i klasifikacije slika, zatim karakterizacija teksture kroz razmatranje globalnih i lokalnih obilježja teksture prilikom opisivanja vizuelnog sadržaja ili semantike sadržaja slika,

- (3) pregled i analiza efikasnosti predloženih metoda za računanje deskriptora multispektralnih slika, sa aspekta dimenzionalnosti deskriptora i tačnosti klasifikacije,
- (4) analiza uticaja korišćenja informacija o prostornom rasporedu obilježja tekture prilikom klasifikacije slika,
- (5) prijedlog novog algoritma za kombinovanje tehnika za redukciju dimenzionalnosti i grupisanje lokalnih obilježja tekture na osnovu njihovog rasporeda na slici, sa ciljem dobijanja deskriptora što manje dimenzionalnosti a što boljih diskriminativnih performansi i
- (6) razmatranje mogućnosti primjene konvolucionih neuronskih mreža prilikom klasifikacije slika u boji i automatske detekcije objekata na aero snimcima visoke rezolucije.

Tokom istraživanja sprovedenog prilikom izrade ove disertacije, objavljeno je nekoliko naučnih radova čiji su rezultati iskorišćeni u daljem tekstu. U radu [12] analizirani su efekti redukcije dimenzionalnosti lokalnih i globalnih deskriptora tekture prilikom klasifikacije aero snimaka u boji. Zaključeno je da je moguće očuvati tačnost klasifikacije aero snimaka i pored značajnog smanjenja dimenzionalnosti deskriptora i to pogodnim izborom kombinovanja informacija iz različitih spektralnih opsega. Rezultati istraživanja u tom radu su poslužili kao osnova za [13] gdje je izvršen sistematski pregled mogućnosti kombinovanja različitih lokalnih i globalnih deskriptora tekture i tehnika redukcije dimenzionalnosti i to na različitim bazama aero snimaka. Izvršena je analiza tačnosti klasifikacije predloženih niskodimenzionalnih deskriptora, za različite postavke eksperimenata. U [14] razmatrane su različite metode podjele prostora slike prilikom grupisanja lokalnih obilježja tekture. Predloženo je proširenje deskriptora tekture računatih prema lokalnim obilježjima tekture sa dodatnim vektorima obilježja koji sadrže isključivo obilježja iz određenog regiona slike. Testirane su metode koje koriste rotirane podjele prostora slika, pri čemu je pokazana konzistentnost u poboljšanju tačnosti rezultata klasifikacije različitih baza slika, bez obzira na način izdvajanja lokalnih obilježja tekture. Poređenje karakteristika lokalnih obilježja tekture niskog i višeg nivoa prilikom klasifikacije tekstura je izvršeno u [15]. Rad je imao za cilj analizu mogućnosti poboljšanja tačnosti klasifikacije isključivo povećanjem količine korišćenih podataka prilikom računanja deskriptora. Takođe izvršeno je poređenje sa metodama koje koriste lokalna obilježja višeg reda. Mogućnosti upotrebe konvolucionih neuronskih mreža prilikom klasifikacije aero snimaka razmotrene u [16], pri čemu je predložen novi metod za fino podešavanje prethodno obučanih mreža.

1.4. Organizacija disertacije

U daljem tekstu, disertacija je organizovana na sljedeći način. U Poglavlju 2 je detaljno opisan problem automatske klasifikacije digitalnih slika i usvojeni su pojmovi potrebni za implementaciju algoritma za automatsku klasifikaciju. Takođe, dat je detaljan pregled relevantne literature u kojoj su opisani najznačajniji naučni doprinosi i tehnike koje su korišćene prilikom istraživanja predstavljenog u ovoj disertaciji. Poglavlje 3 se bavi metodologijom istraživanja, tj. opisuje konkretne

postupke pomoću kojih su vršena testiranja predloženih tehnika i baze slika koje su pri tome korišćene. Poglavlje 4 se bavi izdvajanjem deskriptora teksture i automatskom klasifikacijom sivih slika. Dade su karakteristike deskriptora teksture sivih slika i rezultati eksperimenata na sivim slikama. U Poglavlju 5 su predložene metode za integraciju tehnika redukcije dimenzionalnosti i izdvajanja deskriptora teksture iz multispektralnih slika. Pored analize predloženih rješenja, dati su i diskutovani rezultati eksperimenata. U Poglavlju 6 je razmatran uticaj prostornog rasporeda lokalnih obilježja teksture, dok je u Poglavlju 7 predloženo rješenje koje treba da predstavlja algoritam za izdvajanje niskodimenzionalnih deskriptora teksture uz maksimalno moguće korišćenje informacija o prostornom rasporedu. Poglavlje 8 služi da bi se kompletirao pregled načina za automatsku klasifikaciju slika, gdje je predstavljena mogućnost korišćenja konvolucione neuronske mreže za klasifikaciju. Ukratko je opisano predloženo rješenje i mogućnost primjene prilikom automatske detekcije objekata na aero snimcima visoke rezolucije. Poglavlje 9 predstavlja zaključak.

2 Automatska klasifikacija slika

U uvodnom poglavlju navedeni su osnovni pojmovi potrebni za definisanje problema automatske klasifikacije slika, prepoznavanja objekata i detekcije događaja. Takođe, navedeni su problemi koji nastaju korišćenjem deskriptora teksture visoke dimenzionalnosti, koji se mogu dobiti analizom multispektralnih slika i uz korišćenje informacija o prostornom rasporedu lokalnih obilježja teksture. U ovom poglavlju detaljno se definišu svi pojmovi potrebni za implementaciju algoritama za automatsku klasifikaciju slika, a pored toga dat je pregled naučnih radova u kojima su izloženi doprinosi korišćeni prilikom ovog istraživanja opisanog u ovoj disertaciji.

2.1. Reprezentacija digitalnih slika

Digitalna slika je numerička reprezentacija slike u širem smislu, odnosno način predstavljanja podataka pogodnih za vizuelni prikaz u ravni. Digitalne slike se sastoje od konačnog broja prostornih diskretnih jedinica, piksela¹, čija numerička vrijednost predstavlja nivo svjetline na određenom mjestu slike u širem smislu. Dakle, digitalna slika predstavlja numeričku reprezentaciju stvarne slike. Digitalne slike se predstavljaju u matricnoj formi, pri čemu jedan piksel ima određene koordinate u matrici i odgovarajući nivo svjetline. U daljem tekstu se pod pojmom slika podrazumjeva digitalna slika.

U zavisnosti od broja bita kojima se predstavljaju numeričke vrijednosti piksela, može se predstaviti određeni broj nivoa svjetline. Ukoliko se za predstavljanje koristi jedan bit po pikselu, njegova numerička vrijednost može imati dvije različite vrijednosti, u kojem slučaju se radi o binarnim slikama. U praktičnim aplikacijama, često se koriste sive slike, kod kojih se za predstavljanje numeričke vrijednosti jednog

¹*Pixel, picture element* (eng.) je definisan kao najmanji element na uređaju za prikazivanje kojem se može pripisati adresni prostor. U jednoj digitalnoj slici, piksel predstavlja najmanju komponentu digitalne slike, kojoj je moguće pripisati određenu vrijednost.

piksela koristi nekoliko bita². Slike u boji se predstavljaju kombinacijom podataka iz različitih spektralnih opsega. U skladu sa ljudskom percepcijom boje i vidljivim dijelom spektra, najčešće se koriste crveni, zeleni i plavi kanal, koji se posebno predstavljaju kao pojedinačne sive slike i kažemo da je slika predstavljena u **RGB** prostoru boja³. Kod slika u boji, jedan piksel se predstavlja kao uređena trojka, pri čemu svaka od numeričkih vrijednosti nosi informaciju o intenzitetu odgovarajućeg piksela u kanalu. Ako sivu sliku interpretiramo kao matricu čiji elementi mogu imati jednu od 256 numeričkih vrijednosti nivoa sivila, slika u boji je skup tri matrice istih dimenzija, tj. tenzor. U zavisnosti od praktičnih potreba i konkretnih aplikacija, mogu se koristiti različiti prostori boja [1, 17].

Prema definiciji, multispektralne slike predstavljaju slike koje sadrže više od jednog kanala, tako da su slike u boji jedan slučaj multispektralnih slika. Međutim, obično se koristi terminologija prema kojoj multispektralne slike sadrže podatke iz više od tri kanala. Možemo reći da multispektralne slike predstavljaju prirodno uopštenje slika u boji, gdje se koriste dodatni kanali koji nose podatke iz nevidljivog dijela spektra. Najčešće se dodatni kanali odnose na infracrveni ili ultraljubičasti dio spektra, koji nose korisne informacije nevidljive ljudskom oku, pa multispektralne slike predstavljaju skup više od tri matrice istih dimenzija. Često se koriste za čuvanje podataka prilikom snimanja iz vazduha, kada je od velikog praktičnog značaja analiza zračenja iz nevidljivog dijela spektra. Iako multispektralne slike podrazumjevu slike proširenog spektralnog opsega, za različite praktične aplikacije, dodatni kanali mogu biti definisani prema opsegu iz nevidljivog dijela spektra (npr. NIR⁴ kanal), kao linearne kombinacije podataka iz vidljivog dijela spektra (npr. indeks vegetacije⁵) ili kao rezultat analize nekih od dostupnih kanala (npr. filtriranje nekog od dostupnih kanala).

Hiperspektralne slike koriste mnogo veći broj dodatnih kanala u odnosu na slike u boji i najviše se koriste prilikom snimanja satelitskih slika i drugih snimaka u svemiru i prilikom aero snimanja, gdje se definiše veliki broj kanala iz vidljivog i nevidljivog dijela spektra.

2.2. Detekcija obilježja, izdvajanje vektora obilježja i deskriptora

Već je pokazano kako različite karakteristike teksture zavise od vizuelnog sadržaja prikaza u slikama. Ivice, brazde i konture različitih oblika i veličina, te njihova gustina, veličina i raspored predstavljaju obilježja slika koja je poželjno na pogodan način predstaviti numeričkim vrijednostima. Obilježja i način računanja deskriptora

²U praktičnim aplikacijama često se koriste sive slike sa 8 bita po pikselu.

³*Red, Green, Blue color space* (eng.)

⁴*Near-Infrared* (eng.) Bliski infracrveni opseg je dio spektralnog opsega infracrvenog zračenja talasne dužine od 0.75 do 1.4 μm . Ovaj opseg je definisan tako da obuhvata talasne dužine apsorpcije vode.

⁵Normalizovani indeks vegetacije, *Normalized Difference Vegetation Index* (eng.) predstavlja normalizovanu razliku između vrijednosti piksela iz bliskog infracrvenog opsega i crvenog opsega zračenja.

je poželjno definisati na način da daju što više informacija korisnih za praktičnu aplikaciju za koju su namijenjeni. Obilježja mogu biti jednostavna, kao što je nivo sivila jednog piksela kod sivih slika, boja piksela određene boje kod slika u boji ili pozicija piksela na slici, a mogu predstavljati i numeričke vrijednosti na osnovu vrijednosti boje nekoliko susjednih piksela ili na osnovu vrijednosti boje svih piksela na slici. Ukoliko je potrebno pronaći mjesta na slici koja imaju određene karakteristike u pogledu boje, prisustva i orijentacije ivica, koriste se vrijednosti boje nekoliko susjednih piksela. Vrijednosti boje svih piksela na slici se koriste prilikom primjene integralnih transformacija, kada je potrebno izvršiti reprezentaciju slika u nekom od transformacionih domena, kao što je Furijeov domen [1, 17].

U opštem slučaju, obilježja možemo podijeliti na osnovu nekoliko kriterijuma. Prema načinu interpretacije podataka sa slike, obilježja mogu biti vizuelna ili semantička. Vizuelna obilježja se odnose na prisustvo određene boje, raspored ivica ili karakteristike teksture, dakle, odnose se isključivo na vizuelnu sadržinu slike. Vizuelna obilježja niskog nivoa opisuju elementarni sadržaj slike kao što je boja, tekstura ili oblik. Prilikom analize video frejmova, kao vizuelno obilježje niskog nivoa koristi se i procjena kretanja objekata u uzastopnim frejmovima. Sa druge strane vizuelna obilježja visokog nivoa su usko prilagođena praktičnim aplikacijama, npr. za detekciju lica. Detekcija lica koristi vizuelna obilježja koja se oslanjaju na matematički model poznate morfologije ljudskog lica [18, 19, 20].

Semantička obilježja su obilježja višeg nivoa koja treba da daju informaciju o semantičkom smislu sadržaja prikaza u slici, npr. da li scena u slici prikazuje neki objekat ili određeni događaj. Semantička obilježja se često izvode na osnovu statističke analize vizuelnih obilježja. Npr. vizuelna obilježja se koriste da se prepozna da li se objekti prikazani na sekvenci video frejmova kreću, dok semantička obilježja opisuju da li taj video prikazuje npr. sportski događaj ili o kojem sportu se radi. Nakon detekcije lica u slikama, semantička obilježja se mogu koristiti prilikom prepoznavanja emocija izraženih na licu [21]. Uopšte, semantička obilježja se koriste za opisivanje osobina objekata ili scena koje su prikazane u slikama, tj. da li su materijali od kojih su načinjeni objekti tvrdi, mekani ili savitljivi, i da li je prostor otvoren, zatvoren ili pokriven. Takođe, interesantan je problem pretraživanja baza slika kako bi se pronašle slike koje prikazuju nešto što je tužno ili smiješno. Kod baza medicinskih slika korisno je izvršiti pretragu kako bi se našli primjeri frakture određenog tipa ili sindroma određene bolesti [22, 23].

Analizom sadržaja slike vrši se izdvajanje vektora obilježja, tako što se računaju numeričke vrijednosti koje na pogodan način opisuju prisustvo i karakteristike obilježja na slikama. Tako dobijeni vektori obilježja se mogu koristiti kao deskriptori ili se nad njima mogu izvršiti dodatne matematičke operacije kako bi se dobili vizuelni deskriptori višeg nivoa ili semantički deskriptori. U literaturi se pojmovi vektora obilježja i deskriptora često poistovjećuju, a u ovoj disertaciji je usvojena sljedeća terminologija. Vektor obilježja predstavlja vektor numeričkih vrijednosti koji se dobija analizom sadržaja slika na osnovu jednog usvojenog kriterijuma. Tako vektor obilježja može pokazati koliko ima piksela određene boje, ivica određene orijentacije ili nekih drugih obilježja.

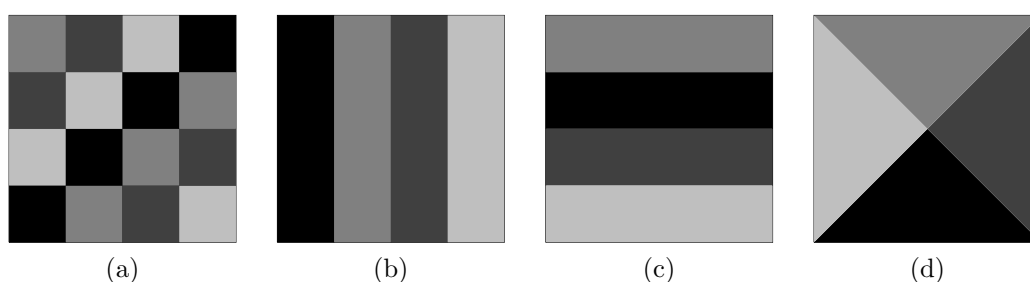
Vektor obilježja se može dobiti analizom slike ili kao rezultat obrade nekoliko vektora obilježja, kao što je spajanje vektora različitih obilježja ili njihova trans-

formacija. Deskriptor može biti skup ključnih riječi koje opisuju sadržaj prikaza u slici ili numerička reprezentacija obilježja u slikama. Kada deskriptor sadrži samo numeričke vrijednosti, pojmovi vektora obilježja i deskriptora se ravnopravno koriste. Prema MPEG-7 standardu za opisivanje multimedijalnog sadržaja, za svaki deskriptor je definisana sintaksa i semantika reprezentacije obilježja [7, 24]. Pošto se u ovoj disertaciji ne koriste ključne riječi, u daljem tekstu će se razmatrati samo obilježja numeričkih vrijednosti i koristiće se isključivo termin deskriptor.

U zavisnosti od načina izdvajanja deskriptori mogu biti lokalni i globalni. Lokalni deskriptori opisuju karakteristike jednog manjeg regiona na slici tj. koriste vrijednosti boje određenog broja povezanih susjednih piksela za računanje numeričkih vrijednosti deskriptora, dok globalni deskriptori opisuju svojstva većih dijelova slike ili cijele slike za računanje numeričkih vrijednosti deskriptora.

2.3. Deskriptori teksture

Tekstura slike je njeno svojstvo koje daje informacije o prostornom rasporedu boja ili nivoa sivila. Posmatrajmo sive slike prikazane na Sl. 2.1. Histogram boja sivih slika pokazuje koliko piksela na slici ima određeni nivo sivila i jednak je za svaku od sivih slika prikazanih na Sl. 2.1. Različit prostorni raspored piksela sa određenim nivoom sivila definiše različite teksture, koje su karakterisane različitim obrazcima i orijentacijama ivica. Konkretno, tekstura prikazana na Sl. 2.1a ima obrazac šahovske table, teksture na Sl. 2.1b i Sl. 2.1c imaju prugaste obrasce različite orijentacije, dok tekstura na Sl. 2.1d ima rotirani blokovski obrazac. Slično kao što je diskutovano u prethodnom poglavlju, u ovom primjeru su različite teksture definisane različitim rasporedom i orijentacijom ivica.



Slika 2.1: Primjeri sivih slika sa istim histogramima boja a različitim teksturama

Pošto tekstura praktično određuje različite segmente ili objekte na slikama, može se koristiti za segmentaciju i uočavanje objekata u slikama. Uzimajući u obzir široke mogućnosti upotrebe teksture, predložen je veliki broj deskriptora koji su se efikasno koristili u različitim praktičnim primjenama. Prema MPEG-7 standardu, definisana su tri deskriptora teksture sivih slika [25]:

- (a) Deskriptor za pretraživanje tekstura (TBD⁶) opisuje perceptivne karakteristike tekstura, kao što su usmjerenost, regularnost i grubost. Ovaj deskriptor

⁶ *Texture browsing descriptor* (eng.)

predstavlja jednostavan vektor koji je definisan sa ukupno 12 bita, 2 za regularnost, 6 za usmjerenost i 4 za grubost. Regularnost predstavlja ocjenu periodičnosti teksture. Ukoliko je tekstura periodična, opisana je sa nivoom regularnosti 3, ukoliko nije periodična, nivo regularnosti je 0. Usmjerenost teksture se predstavlja sa jednom od šest vrijednosti, 0° , 30° , 60° , 90° , 120° ili 150° . Ovaj deskriptor može predstaviti dvije dominantne orijentacije teksture. Grubost teksture je definisana u prostornom i frekvencijskom domenu. Glatka tekstura se opisuje sa nivoom 0, a gruba sa nivoom 3, pri čemu se koriste po dva bita za predstavljanje u prostornom i frekvencijskom domenu.

- (b) Homogeni deskriptor teksture (HTD⁷) je definisan da izvrši kvantitativnu karakterizaciju teksture pogodnu za pretraživanje po sličnosti. Ovaj deskriptor se računa na osnovu rezultata filtriranja slike pomoću banke orijentisanih i skaliranih filtara, čija je baza data pomoću Gaborove funkcije:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + j\Omega x}, \quad (2.1)$$

gdje su σ_x i σ_y propusni opsezi, a Ω centralna frekvencija filtra. Nakon filtriranja, računa se energija i njena devijacija filtrirane slike za svaku od skala i orijentacija.

- (c) Deskriptor histograma ivica (EHD⁸) opisuje prostorni raspored ivica teksture i osmišljen je u cilju efikasnog poređenja tekstura u slučaju da one nisu homogene. Prilikom računanja ovog deskriptora, slika se dijeli na 4×4 nepreklapajuće podslike, pri čemu se jedna podslika dalje dijeli na konačan broj blokova. Analizom vrijednosti nivoa sivila u blokovima se za svaki blok računa jedna od pet definisanih orijentacija ivica (horizontalne, vodoravne, kose pod uglom od 45° i 135° i ostale ivice). Neuniformnom kvantizacijom formira se konačan deskriptor sa 240 bita.

U praktičnim primjenama često su korišćeni i sljedeći deskriptori tekstura:

- (d) GLCM⁹ je statistički metod opisivanja teksture na osnovu prostornih relacija između piksela [26]. Zapravo, radi se o načinu predstavljanja teksture računanjem koliko puta se pojavljuje par piksela sa određenim nivoima sivila sa prethodno definisanim prostornim rasporedom, na osnovu čega se izvodi matrica uzajamnih relacija nivoa sivila. Iz te matrice se mogu računati statistički parametri koji opisuju kontrast, korelaciju, uniformnost i homogenost teksture.
- (e) Autokorelacija i spektar snage slike se mogu koristiti kao deskriptori teksture [27]. Autokorelacija slike detektuje obrazce koji se ponavljaju i može opisati hrapavost ili finoću teksture. Slično, periodičnost ponavljanja obrazaca teksture se može detektovati u frekvencijskom domenu. Analizom spektra snage, računatim primjenom dvodimenzionalne diskretne Furijeove transformacija, može se analizirati hrapavost ili finoća teksture.

⁷Homogeneous texture descriptor (eng.)

⁸Edge histogram descriptor (eng.)

⁹Gray level co-occurrence matrix (eng.)

Neki od prethodno navedenih deskriptora teksture definisani su prema matematičkim modelima, koji se oslanjaju na određene statističke karakteristike tekstura. Te statističke karakteristike su osmišljene tako da kvantitativno predstavljaju jednostavne pridjeve kojima se mogu opisati teksture, kao što su usmjerenost, orijentacija, grubost, finoća, itd. Drugi deskriptori izražavaju spektralne karakteristike teksture korišćenjem određenih integralnih transformacija. Međutim, kada se radi o praktičnim primjenama gdje se koriste slike sa prirodnim teksturama, sljedeći deskriptori su postigli bolje rezultate:

- (f) GIST je globalni deskriptor teksture [8], definisan tako da opiše prostornu strukturu prirodnih scena. Prije računanja deskriptora, slika se filtrira sa skupom skaliranih i rotiranih Gaborovih filtara [4], a zatim se prostor slike dijeli na 4×4 nepreklapajuće podslike. Za svaku od podslika računaju se srednje vrijednosti odziva na filtriranje za svaku od skala i orijentacija. Tako dobijeni statistički parametri se spajaju u jedan vektor numeričkih vrijednosti i koriste se kao deskriptor teksture. Ukoliko se koristi S skala i K orijentacija, GIST deskriptor predstavlja vektor srednjih vrijednosti:

$$\mathbf{gist} = [\mu_{1,1,1} \dots \mu_{1,K,1} \dots \mu_{S,K,1} \dots \mu_{S,K,16}], \quad (2.2)$$

gdje je μ_{ijk} srednja vrijednost Gaborovih koeficijenata na i -toj skali, j -oj orijentaciji u podslici k . GIST deskriptor predstavlja proširenje homogenog deskriptora teksture koji posebno uzima u obzir neuniformnu teksturu često prisutnu u prirodnim scenama.

- (g) CENTRIST¹⁰ je definisan kao vizuelni deskriptor za kategorizaciju scena [10] i predstavlja histogram sive slike nakon primjene Census transformacije. Ovaj deskriptor je definisan tako da naglasi određene strukture, kao što su ivice i konture, a da zanemari površi sa manjim lokalnim varijacijama vrijednosti piksela. Census transformacija svakom pikselu dodjeljuje novu vrijednost koja se računa na osnovu vrijednosti nivoa sivila posmatranog piksela i osam susjednih piksela. Vrijednost Census transformacije se računa računanjem binarnog broja koji se dobija poređenjem vrijednosti centralnog piksela sa susjednim pikselima, a zatim konverzijom tog binarnog broja u decimalni. Ukoliko posmatramo 3×3 piksela, susjednim pikselima se pridružuje “1” ukoliko je njihova vrijednost veća ili jednaka od vrijednosti centralnog piksela, a “0” u suprotnom:

$$\begin{array}{|c|c|c|} \hline 54 & 58 & 100 \\ \hline 78 & 75 & 62 \\ \hline 132 & 128 & 95 \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 1 & & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \Rightarrow (00110111)_2 = (55)_{10}. \quad (2.3)$$

- (h) LBP¹¹ je deskriptor predložen za multirezolucionu klasifikaciju tekstura invarijantno u odnosu na njihovu rotaciju [9]. Prilikom računanja ovog deskriptora, lokalne karakteristike teksture se analiziraju na sličan način kao i prilikom Census transformacije, ali interpretirajući vrijednosti nivoa sivila piksela kao

¹⁰ *Census Transform Histogram* (eng.)

¹¹ *Local Binary Pattern* (eng.)

vrijednosti dvodimenzionalne kontinualne funkcije. Kako su vrijednosti dvodimenzionalne funkcije poznate isključivo u tačkama gdje se nalaze centri piksela, binarni obrazac lokalne teksture se računa interpolacijom iz vrijednosti dvodimenzionalne funkcije na jednakim rastojanjima od centralnog piksela. Da bi se postigla rotaciona invarijantnost, binarni obrazac se rotira tako da odgovarajuća dekadna vrijednost bude maksimalna. Konačno, LBP deskriptor se računa kao histogram vrijednosti dobijenih nakon svih koraka obrade. U praktičnim primjenama, LBP se najčešće koristi prilikom multirezolucione analize. Računanje lokalnih obilježja na više rezolucija kod LBP deskriptora podrazumjeva računanje binarnih obrazaca za različita rastojanja od centralnog piksela.

- (i) SIFT deskriptor¹² [5, 28] predstavlja vektor obilježja izračunat na osnovu lokalnih karakteristika određenih regiona na slici. Skraćenica SIFT se odnosi na transformaciju koja se koristi prilikom detekcije lokalnih obilježja invarijantnih na skalu i može da se koristi prilikom analize teksture u određenom regionu slike. SIFT detektor predstavlja skup procedura za traženje karakterističnih regiona na slikama, odnosno ključnih tačaka. SIFT deskriptor predstavlja vektor numeričkih vrijednosti dobijen nakon analize detektovanih obilježja. SIFT deskriptor može da se računa na osnovu karakterističnih regiona¹³ na slici koji se traže SIFT detektorom ili tehnikom gustog odabiranja¹⁴, koja analizira teksturu u gusto i pravilno raspoređenim tačkama na slici. Zbog mogućnosti detekcije ključnih tačaka sa karakterističnim obilježjima, pored poređenja slika na osnovu sadržaja, SIFT se može koristiti i za registraciju slika i spajanje slika u mozaike. Prilikom računanja SIFT deskriptora, vrši se agregacija podataka dobijenih SIFT transformacijom. Često se SIFT obilježja koriste u kombinaciji sa BoVW¹⁵ paradigmom [29], koja predstavlja proceduru prebrojavanja sličnih obilježja i konstruisanja odgovarajućeg histograma, međutim postoje i drugi načini za agregaciju lokalnih obilježja koji su pogodniji za određene praktične aplikacije [30, 31, 32, 33].

Aplikacije koje se odnose na klasifikaciju prirodnih i medicinskih slika, kategorizaciju scena i detekciju događaja, u poslednjih nekoliko godina, uglavnom koriste GIST, CENTRIST, LBP i SIFT deskriptore. Osim GIST deskriptora, koji je baziran na integralnoj transformaciji, CENTRIST, LBP i SIFT deskriptori u suštini vrše prebrojavanje lokalnih obilježja detektovanih određenim procedurama. U daljem tekstu disertacije, računanje navedenih deskriptora podrazumjeva primjenu određenih tehnika za obradu sivih slika prilikom kojih se dobija deskriptor sa numeričkim vrijednostima. Posebno će biti razmotreni parametri koji utiču na dimenzionalnost deskriptora, kao i načini kombinovanja podataka iz različitih spektralnih opsega.

U praktičnim slučajevima kada ne postoje veliki memorijski resursi niti računarski kapaciteti (npr. kod prenosivih uređaja sa baterijskim napajanjem), poželjno je koristiti deskriptore što niže dimenzionalnosti. Sa druge strane, multispektralne kamere

¹² *Scale Invariant Feature Transform* (eng.)

¹³ U literaturi se često koristi termin ključne tačke *Keypoints* (eng.) iako se koristi određeni pravougaoni region susjednih piksela.

¹⁴ Tehnika gustog odabiranja *Dense sampling* (eng.)

¹⁵ *Bag of Visual Words* (eng.)

postaju sve jeftinije, manjih dimenzija i pristupačnije. Smanjenjem dimenzionalnosti deskriptora, smanjuje se količina podataka kojima se opisuje sadržaja u slikama, pa se smanjuje tačnost klasifikacije. Kroz naredne eksperimente teži se razvoju algoritma koji bi računao deskriptore niske dimenzionalnosti iz multispektralnih slika, koji postižu dobre rezultate klasifikacije.

2.4. Automatska klasifikacija slika

Da bi se mogla izvršiti automatska klasifikacija na osnovu deskriptora potrebno je ustanoviti nedvosmislene kriterijume prema kojima se slika pridružuje određenoj klasi. Klasifikacija slika na osnovu vizuelne ili semantičke sličnosti može se oslanjati na poznavanje karakteristika slika iz svih klasa prema kojima se vrši klasifikacija, a moguće je izvršiti grupisanje slika sa najbližijim deskriptorima bez prethodnog poznavanja karakteristika klasa. U daljem tekstu smatra se da su klase poznate. Slike koje su već pridružene odgovarajućim klasama ćemo nazvati trening slikama, a slike koje treba klasifikovati testnim slikama. Posmatrajući sliku koju je potrebno klasifikovati, utvrđuje se kojim slikama iz prethodno poznatih klasa je posmatrana slika najbližija prema nekom kriterijumu. Dakle, za klasifikaciju slika potrebno je poznavati karakteristike klasa i utvrditi način na koji se karakteristike neke posmatrane slike porede sa poznatim karakteristikama klasa. To znači da je u praktičnim aplikacijama potrebno imati skup trening slika, na osnovu kojih se obučava klasifikator. Analizom skupa trening slika dobijaju se karakteristike svake od klasa koje služe za obučavanje klasifikatora, odnosno utvrđuje se način na koji se karakteristike neke posmatrane slike porede sa karakteristikama klasa.

Za obučavanje klasifikatora koriste se deskriptori svih slika iz skupa trening slika. Pored tih deskriptora, potrebno je znati kojoj klasi pripada svaka od trening slika, tj. koji deskriptori pripadaju kojoj klasi. Pošto se deskriptori mogu poistovjetiti sa vektorima obilježja, svaki deskriptor predstavlja tačku u višedimenzionalnom prostoru obilježja. Obučavanje klasifikatora predstavlja računanje parametara matematičkog modela, tj. granica, koje će prostor obilježja podijeliti na regione koji pripadaju različitim klasama. Drugim riječima, klasifikator predstavlja funkciju koja prostor obilježja preslikava u skup klasa. U praktičnim aplikacijama najčešće se koriste sljedeći klasifikatori [34, 35]:

- (a) k najbližih susjeda (kNN)¹⁶ je neparametarski algoritam koji se može iskoristiti za automatsku klasifikaciju slika. Tom prilikom računa se udaljenost deskriptora slike od svih uzoraka iz trening skupa, a klasifikacija se vrši na osnovu k najbližih susjeda. Slika se klasifikuje u onu klasu u kojoj se nalazi većina od k najbližih susjeda [35].
- (b) Naivni Bajsov kriterijum se odnosi na skup algoritama za konstruisanje klasifikatora koji se zasnivaju na pretpostavci da su vektori obilježja međusobno nezavisni. Klasifikatori iz ove kategorije mogu se konstruisati prema različitim pretpostavkama o funkciji gustine raspodjele vjerovatnoće podataka iz iz trening skupa. Ovaj način klasifikacije je pogodan u slučajevima kada postoji

¹⁶*k nearest neighbors* (eng.)

opravdanost pretpostavke o nezavisnosti vektora obilježja, u kom slučaju se brzo postiže konvergencija uz dobre rezultate klasifikacije [34]. Međutim, ovaj klasifikator može postići dobre rezultate klasifikacije i kada nije opravdana pretpostavka o nezavisnosti vektora obilježja, npr. prilikom klasifikacije teksta [36, 37].

- (c) Klasifikacija logističkom regresijom se vrši na osnovu linearne kombinacije vektora obilježja iz skupa trening podataka [35]. Obučavanje ovog klasifikatora podrazumjeva računanje težinskih faktora linearne kombinacije elemenata posmatranog vektora obilježja. Ova linearna kombinacija se poredi sa unaprijed određenim pragom, tj. određuje se u kojem dijelu prostora obilježja se nalazi i kojoj klasi pripada.
- (d) Stabla odlučivanja predstavljaju statistički model koji koristi podatke iz trening skupa za klasifikaciju, tako što uči jednostavna pravila odlučivanja. Stabla mogu da se nauče iz numeričkih i nenumeričkih obilježja, kao i da se vizuelno predstave [38].
- (e) SVM¹⁷ je klasifikator koji na osnovu trening podataka generiše hiperravan u prostoru obilježja tako da ona bude maksimalno udaljena od deskriptora različitih klasa [39, 40]. U slučaju kada su deskriptori iz različitih klasa linearno razdvojivi, SVM algoritam će dati optimalan model klasifikacije. Međutim, deskriptori slika iz različitih klasa često nisu linearno razdvojivi, pa je potrebno koristiti dodatnu transformaciju prostora obilježja kako bi se deskriptori slika iz različitih klasa preslikali u prostor u kojem je moguće naći granicu odlučivanja u vidu hiperravni. Praktična realizacija SVM klasifikatora koristi kernel funkcije koje preslikavaju deskriptore u transformisani prostor obilježja, bez potrebnog definisanja transformacije [35]. Zbog dobrih karakteristika SVM se često koristi u praktičnim aplikacijama u oblasti obrade slike, mašinskog učenja i kompjuterskog vida.
- (f) Neuronske mreže predstavljaju matematički model inspirisan biološkim neuronskim mrežama, koji služi za aproksimaciju funkcija sa velikim brojem ulaznih vrijednosti. Prilikom automatske klasifikacije slika, u skorije vrijeme se uglavnom koriste konvolucione neuronske mreže¹⁸, koje daju dobre rezultate klasifikacije ukoliko su trenirane na odgovarajući način. Konvolucione neuronske mreže (CNN) objedinjuju detekciju obilježja i klasifikaciju kroz implementaciju nekoliko uzastopnih konvolucionih i potpuno povezanih slojeva. Prednosti CNN-a jesu inherentno korišćenje prostornog rasporeda lokalnih obilježja i mogućnost korišćenja obilježja višeg reda, što obezbjeđuje bolje rezultate klasifikacije u poređenju sa drugim klasifikatorima. Međutim, CNN-e ne mogu da se koriste u slučaju da je dostupna nedovoljna količina trening podataka i zahtjevaju puno računarskih resursa za učenje. Njihova ogromna prednost jeste mogućnost pretreniranja na jednoj bazi slika i korišćenje na drugim raznorodnim bazama slika [41].

Klasifikacija naivnim Bajsovim klasifikatorom i logističkom regresijom se češće koriste prilikom pretraživanja i klasifikacije tekstualnih dokumenata. kNN i SVM

¹⁷ *Support Vector Machine* (eng.), mašina sa vektorima nosačima.

¹⁸ *Convolutional neural networks* (eng.)

klasifikatori se često koriste prilikom automatske klasifikacije slika, pri čemu je bitno naglasiti neke njihove prednosti i nedostatke. kNN klasifikator praktično nema fazu obučavanja i brzo se izvršava ukoliko je trening skup mali. Greška klasifikacije se smanjuje sa povećanjem trening skupa, ali to zahtjeva proporcionalno veće memorijske resurse i duže izvršavanje. SVM klasifikator može naučiti kompleksne odnose između klasa i postići dobre rezultate klasifikacije čak i kada je trening skup mali. Takođe, u slučaju kada deskriptori iz različitih klasa nisu linearno separabilni u prostoru obilježja, uz odgovarajući odabir kernela SVM postiže dobre rezultate klasifikacije. Automatska klasifikacija slika pomoću konvolucionih neuronskih mreža je pristup koji je nedavno privukao pažnju jer postiže najbolje rezultate u slučaju kada postoji dovoljno veliki broj trening podataka, međutim ako je dostupan samo mali broj trening slika, konvolucione neuronske mreže je teško obučiti kako bi se dobili dobri rezultati klasifikacije.

U većini eksperimenata opisanih u ovoj disertaciji koristi se SVM klasifikator jer je moguće paralelno ispitati efekat redukcije dimenzionalnosti deskriptora na brzinu i tačnost klasifikacije. Takođe, ovaj klasifikator se može efikasno koristiti u slučaju kada postoji malo trening slika. Sa druge strane, konvolucione neuronske mreže nije moguće obučiti sa malo trening slika, pa je pokazano kako ih je moguće iskoristiti za klasifikaciju ukoliko su prethodno obučene na drugoj bazi slika.

Prema načinu na koji se vrši, automatska klasifikacija slika može biti:

- (a) **Nenadgledana klasifikacija**¹⁹, predstavlja pristup automatskoj klasifikaciji gdje korisnik prethodno ne definiše klase, već samo njihov broj. Kada je poznat broj klasa, algoritam za klasifikaciju vrši grupisanje deskriptora u prostoru obilježja, nakon čega korisnik pripisuje značenje svakoj od grupa ili pojedine grupe spaja u jednu. Ovakav pristup klasifikaciji se najčešće koristi prilikom klasifikacije na osnovu piksela kod snimaka iz vazduha satelitskih snimaka niske rezolucije, kada je svakom pikselu potrebno odrediti klasu.
- (b) **Nadgledana klasifikacija**²⁰ se oslanja na prethodno definisane klase od strane korisnika. U ovom slučaju korisnik definiše koje slike pripadaju kojoj klasi, a klasifikator se obučava na osnovu deskriptora svih trening slika. Ovakav pristup klasifikaciji se koristi prilikom prepoznavanja scena, detekcije objekata i događaja, prilikom blokovske klasifikacije aero snimaka visoke rezolucije, itd.
- (c) **Objektno orijentisana klasifikacija**²¹ je takav pristup klasifikaciji koji prepoznaje objekte određenih oblika i veličina na slikama i može biti nadgledana ili nenadgledana. Prilikom klasifikacije koristi se segmentacija koja razdvaja regione na slikama koji čine smislene cjeline proizvoljnog oblika. Takvi objekti se klasifikuju na osnovu sadržaja, najčešće prilikom analize aero snimaka ili prepoznavanja objekata.

Pošto se u ovoj disertaciji razmatra isključivo nadgledana automatska klasifikacija slika, ukratko će biti opisani koraci neophodni za implementaciju. Algoritam za nadgledanu automatsku klasifikaciju slika treba da uključi sljedeće korake: **(1)** Formiranje trening skupa slika, **(2)** računanje deskriptora za svaku od slika iz trening

¹⁹ *Unsupervised classification* (eng.)

²⁰ *Supervised classification* (eng.)

²¹ *Object-oriented image analysis classification* (eng.)

skupa, (3) obučavanje klasifikatora, (4) računanje deskriptora za slike koje se klasifikuju i (5) odlučivanje kojoj klasi pripadaju slike.

Formiranje skupa trening slika

Trening slike treba da budu odabrane tako da sadrže dovoljno vizuelnih ili semantičkih informacija o klasama koje definišu. Poželjno je da njihov broj ne bude veliki kako obučavanje klasifikatora ne bi predugo trajalo, ali previše mali skup trening slika ne garantuje korektno modelovanje podataka koje karakterišu sve klase. Moguće je odabrati isti ili različit broj trening slika za svaku od klasa i svakoj slici treba da bude pridružena informacija o tome kojoj klasi pripada ta slika.

Računanje deskriptora za slike iz trening skupa

Kada je formiran trening skup slika, potrebno je izračunati deskriptore za svaku od tih slika. Na osnovu praktične aplikacije na koju se odnosi automatska klasifikacija, poželjno je razmotriti odgovarajuća obilježja i odabrati odgovarajući algoritam za računanje deskriptora. Različite vrste deskriptora su predložene kako bi se koristile za različite praktične aplikacije. Npr., deskriptori za prepoznavanje lica su izvedeni na osnovu karakteristične morfologije ljudskog lica i nisu pogodni za prepoznavanje objekata dobijenih tehnikama daljinske detekcije.

Obučavanje klasifikatora

Obučavanje klasifikatora predstavlja utvrđivanje pravila na osnovu kojeg se za neku sliku određuje kojoj klasi pripada. Na primjer, u slučaju da se koristi SVM klasifikator, računa se hiperravan koja obezbjeđuje najveću udaljenost između deskriptora trening slika u transformisanom prostoru obilježja.

Računanje deskriptora slika koje se klasifikuju

Za svaku sliku koja treba da se klasifikuje, računa se isti deskriptor koji se koristio za slike iz skupa trening slika.

Odlučivanje kojoj klasi pripadaju slike

Konačno, na osnovu deskriptora slika i ranije obučenog klasifikatora, svakoj slici iz testnog skupa pridružuje se informacija o klasi u koju se svrstava. Klasifikacija neke slike se izvršava tako što se određuje položaj deskriptora posmatrane slike u prostoru obilježja u odnosu na granicu odlučivanja.

2.5. Redukcija dimenzionalnosti deskriptora

Kada se razmatra redukcija dimenzionalnosti deskriptora prilikom automatske klasifikacije slika, taj proces možemo posmatrati sa dva različita aspekta. Svaka slika predstavlja skup podataka koji nose određene vizuelne i semantičke informacije. Prilikom poređenja semantičkog sadržaja slika potrebno je procesirati veliki broj slika velikih dimenzija i na pogodan način opisati njihov sadržaj. Zbog toga nije praktično koristiti podatke o svakom pikselu sa slike, već je poželjno na neki drugi način predstaviti podatke o sadržaju u slikama. U zavisnosti od načina računanja, deskriptori mogu zahtjevati manje ili više memorijskih resursa od slike čiji sadržaj

treba da opisuju. U prvom slučaju, izdvajanje deskriptora predstavlja jedan aspekt redukcije dimenzionalnosti slike, jer se obradom slike dobija deskriptor koji je manjih dimenzija, a koji treba da nosi slične informacije kao i slika. Deskriptori manjih dimenzija su često dovoljni za predstavljanje jednostavnijih tekstura prilikom automatske klasifikacije ili pretraživanja baza slika na osnovu sadržaja [4, 6, 7, 24]. Međutim, sa razvojem tehnologije i povećanjem memorijskih i računarskih resursa, praktične aplikacije mogu da koriste deskriptore veće dimenzionalnosti. Pojedini deskriptori imaju mnogo veće dimenzije od same slike čiji sadržaj opisuju, pri čemu obezbeđuju separabilnost između slika iz različitih klasa [42, 43, 31, 33].

Redukcija dimenzionalnosti deskriptora predstavlja proceduru obrade numeričkih vrijednosti deskriptora kojom se smanjuje dimenzionalnost samog deskriptora prema nekom kriterijumu. Ova procedura može imati nekoliko uticaja na automatsku klasifikaciju slika. Korišćenje vektora sa manjim dimenzijama zahtjeva manje memorijskih resursa prilikom obučavanja klasifikatora i izvršavanja klasifikacije, a procesiranje manje količine podataka zahtjeva izvršavanje manjeg broja matematičkih operacija i omogućava brže izvršavanje automatske klasifikacije. Dalje, u zavisnosti od načina izdvajanja deskriptora, određeni podaci mogu biti korelisani, tj. mogu nositi manju količinu informacija nego neki drugi podaci. Tehnike redukcije dimenzionalnosti mogu da uklone korelisanost i da sačuvaju samo one podatke koji nose najveću količinu informacija. Zbog toga se može očekivati da se u nekim slučajevima redukcijom dimenzionalnosti deskriptora poboljšaju rezultati automatske klasifikacije.

U Sekciji 2.3 opisano je nekoliko deskriptora teksture koji su definisani za sive slike. Dimenzionalnost ovih deskriptora zavisi od različitih parametara koji se mogu postaviti u skladu sa praktičnim potrebama. U slučaju njihove primjene na slike u boji ili multispektralne slike, svaki od kanala se može procesirati kao jedna siva slika. Jedan jednostavan način izdvajanja deskriptora slika u boji ili multispektralnih slika jeste spajanje vektora obilježja izračunatih za svaki od kanala posebno. Tada je dimenzionalnost deskriptora tri ili više puta veća od dimenzionalnosti sive slike. U ovom slučaju, povećanjem broja kanala slike, proporcionalno se povećava dimenzionalnost deskriptora, pa je korisno razmotriti mogućnosti smanjenja dimenzionalnosti.

Tehnike redukcije dimenzionalnosti oslanjaju se na statističku analizu podataka iz deskriptora, da bi se izračunali novi deskriptori niže dimenzionalnosti. U praktičnim primjenama, redukcija dimenzionalnosti deskriptora najčešće se izvodi sljedećim tehnikama:

- (a) Analiza glavnih komponenta, PCA²², predstavlja linearnu ortogonalnu transformaciju koja skup korelisanih uzoraka pretvara u skup nekorelisanih uzoraka [44]. U pogledu automatske klasifikacije slika, jedan deskriptor u prostoru obilježja predstavlja jedan uzorak ulaznih podataka. Iz skupa ulaznih podataka računa se ortogonalan linearna transformacija, koja praktično predstavlja rotaciju koordinatnog sistema prostora obilježja. Kada se na ulazne podatke primjeni PCA transformacija, dobijaju se izlazni podaci poznatih varijansi. Redukcija dimenzionalnosti zasnovana na PCA transformaciji koristi pretpo-

²² *Principal Component Analysis* (eng.)

stavku da podaci veće varijanse nose veću količinu informacija, pa se smanjenje dimenzionalnosti vrši odbacivanjem komponenata sa malom varijansom.

- (b) Linearna diskriminantna analiza, LDA²³ predstavlja statistički metod koji koristi linearnu kombinaciju ulaznih podataka (u ovom slučaju deskriptora) tako da poveća separabilnost između podataka koji pripadaju različitim klasama [45]. U klasifikaciji slika, LDA treba da izvrši takvu transformaciju deskriptora kako bi se smanjila udaljenost između deskriptora iz iste klase, a povećala udaljenost između deskriptora različitih klasa. LDA je transformacija koja čuva diskriminantnost deskriptora, tako što traži projekciju na takvu osu u prostoru obilježja, koja povećava separabilnost između pojedinih klasa.

2.6. Pregled literature

Nakon što je utvrđeno da tekstura ima veliki značaj prilikom vizuelnog ili semantičkog opisivanja sadržaja slike, deskriptori teksture su često korišćeni prilikom automatske klasifikacije slika, kategorizacije scena i prepoznavanja događaja. Počevši od jednostavnih algoritama koji su izdvajali obilježja teksture i koristili jednostavne klasifikatore, tokom poslednjih godina, razvoj tehnologije i povećanje memorijskih resursa su omogućili razvoj i implementaciju složenijih algoritama za klasifikaciju. Generalno, numerički način interpretacije tekstura u prirodnim i sintetičkim slikama predstavlja interesantan problem od početka razvoja digitalne obrade slike, kompjuterskog vida i mašinskog učenja.

S obzirom da ne postoji stroga definicija teksture, često nije moguće primijeniti ista pravila za kvantitativno opisivanje teksture na baze slika različitih karakteristika. Drugi problem jeste način opisivanja teksture kod slika u boji i multispektralnih slika. Zbog toga, ne postoji strogo definisan način opisivanja i interpretacije teksture. Deskriptori teksture se najčešće računaju prema statističkim karakteristikama vrijednosti piksela [4, 25, 7, 8, 9, 46, 24] koje omogućavaju numeričku reprezentaciju tekstura na osnovu vrijednosti nivoa sivila ili boje na slikama, ali postoje i drugačiji pristupi interpretaciji teksture. Tekstura može biti opisana i tekstonima [47, 48], tj. lokalnim strukturalnim karakteristikama, što predstavlja model zasnovan na ljudskoj percepciji teksture.

Problem opisivanja teksture slika u boji se može posmatrati na sljedeća dva načina. Prva mogućnost je korišćenje teksture sivih slika, tj. posmatranje svakog od dostupnih spektralnih kanala posebno, a zatim, kombinovanje tako dobijenih informacija [49, 27, 29, 50, 11, 13, 51]. Druga mogućnost jeste razmatranje kolor teksture i definisanje odgovarajućih deskriptora u slučaju kada se podaci iz različitih spektralnih opsega integrišu u definiciju teksture [52, 48]. Poslednjih godina, u oblasti automatske klasifikacije slika, istraživanje je uglavnom bilo bazirano na prvom pristupu, tj. na interpretaciji teksture sivih slika i proširenju na slike u boji i multispektralne slike.

Pored globalnih deskriptora teksture [8, 6, 25, 24, 27], posebnu pažnju su privukli lokalni deskriptori, prvenstveno zbog invarijantnosti na određene parametre,

²³*Linear Discriminant Analysis* (eng.)

mogućnosti dalje obrade i prilagođavanja konkretnim problemima i zbog dobrih diskriminativnih karakteristika. Pionirski radovi prilikom definisanja transformacije invarijantne na skalu (SIFT) su [5, 28] u kojima je opisan način detekcije lokalnih karakterističnih tačaka pogodnih za uparivanje slika prema vizuelnoj sličnosti. SIFT deskriptor je kasnije često korišćen u istraživanju, prilikom automatske klasifikacije multispektralnih slika [29, 53], detekcije karakterističnih tačaka kod multispektralnih slika [54], reprezentacije teksture slika u boji [55, 51], prepoznavanja sličnih slika na osnovu prostornog rasporeda lokalnih obilježja [56, 49, 57], itd.

U [9] je predstavljen multirezolucioni pristup prilikom računanja lokalnih deskriptora teksture sivih slika koristeći rotaciono-invarijantna lokalna obilježja (LBP). Ovaj deskriptor je kasnije korišćen u različitim varijacijama za razne probleme automatske klasifikacije slika [58, 59, 60]. Nakon što su obilježja teksture niskog nivoa iskorišćena prilikom klasifikacije aero snimaka, medicinskih slika i prilikom kategorizacije scena, bolja tačnost klasifikacije se postizala boljim modelovanjem obilježja niskog nivoa, tj. prilagođenjem načina predstavljanja obilježja konkretnim zadacima. To je otvorilo diskusiju o karakteristikama teksture koje je ne opisuju statistički već koje najbolje opisuju objekat definisan tom teksturom. U [10] je jedan od motiva za definisanje CENTRIST deskriptora bio način prepoznavanja različitih površi na različitim objektima, dok je u [51] posebna pažnja posvećena vizuelnim karakteristikama teksture koje zapravo karakterišu neki objekat u prirodi. Posebno interesantno pitanje jeste način proširivanja deskriptora definisanih za sive slike i njihova primjena na slike u boji i multispektralne slike. Korišćenjem podataka iz više spektralnih opsega poboljšani su rezultati klasifikacije. U [29], jednostavnim spajanjem deskriptora sivih slika dobijenih iz svakog od dostupnih kanala, izvršena je klasifikacija prirodnih scena. Izvršeno je poređenje performansi lokalnih i globalnih deskriptora teksture, analiza uticaja dekorelacije podataka iz različitih kanala, kao i redukcije dimenzionalnosti, na tačnost klasifikacije. S obzirom na prezentovane rezultate i ranije tehnike klasifikacije aero- i satelitskih snimaka prema pikselima, analiza podataka iz različitih kanala prilikom automatske klasifikacije privukla je veliku pažnju. Uopštenje problema računanja CENTRIST deskriptora na multispektralne slike (mCENTRIST) predloženo je u [11], pri čemu se posebna pažnja obratila na tehnike pretprocesiranja slika i postprocesiranja deskriptora.

Paralelno istraživanju vezanom za reprezentaciju multispektralnih podataka, razmatrane su mogućnosti korišćenja informacija o prostornom rasporedu obilježja na slikama. Još u pionirskim radovima o automatskoj klasifikaciji slika na osnovu tekstura [8], iskorišćen je prostorni raspored globalnih karakteristika scene da bi se izvršilo modelovanje "suštine" koju ta scena prikazuje. Pošto detekcija lokalnih obilježja daje više mogućnosti korišćenja informacija o njihovom prostornom rasporedu, objavljeno je više radova koji se bavi ovom tematikom. U [56] predstavljen je način kako se iskoriste podaci o prostornom rasporedu lokalnih SIFT obilježja slike, tako što se izvršena piramidalna dekompozicija prostora slike na više nivoa. Autori su pokazali da se konačan deskriptor dobija spajanjem svih prostornih histograma pomnoženih odgovarajućim težinskim faktorom. Na ovaj način se značajno povećava dimenzionalnost deskriptora. CENTRIST [10, 11] koristi prostorni raspored lokalnih obilježja dobijenih Census transformacijom da bi se izračunali deskriptori 31 regiona na različitim nivoima piramidalne dekompozicije. Konačni deskriptori CENTRIST

i mCENTRIST dobijaju se spajanjem deskriptora pojedinih regiona.

Dalji koraci su napravljeni u [61, 57] gdje su korišćene orijentisane podjele prostora slike radi bolje prostorne reprezentacije prikaza na slikama. Konkretno u [61] je predložen način grupisanja lokalnih obilježja prema orijentisanim podjelama prostora slike, kako bi se dobio rotaciono-invarijantan deskriptor, dok je u [57] predložen način učenja optimalne orijentisane i neorijentisane podjele prostora slika.

Iako su prethodno opisani deskriptori pokazali zadovoljavajuće diskriminativne karakteristike, njihove numeričke vrijednosti se računaju isključivo prema statističkim parametrima prvog reda. Popravljanje tačnosti klasifikacije može se postići upotrebom statističkih parametara drugog reda i odgovarajućom reprezentacijom lokalnih obilježja. U [42, 43] je predstavljen kompaktan način reprezentacije slika koji koristi statistiku drugog reda, pored centroida klastera u prostoru obilježja koristi i udaljenosti pojedinih deskriptora od centorida. Evaluacija takvog deskriptora na problemu klasifikacije aero snimaka izvršena u [33].

Poboljšanja u performansama klasifikatora se mogu postići i dodatnim procesiranjem podataka ili njihovom drugačijom interpretacijom tokom obučavanja klasifikatora. U [62] je predloženo dodatno procesiranje lokalnih obilježja prije računanja konačnog histograma lokalnih obilježja, kako bi se izbjegla potreba za nelinearnim klasifikatorima. Autori su u [63] predstavili fleksibilan način generisanja histograma lokalnih obilježja koji omogućava raspodjelu doprinosa jednog lokalnog obilježja na više kodnih riječi. Sličan metod je iskorišćen u [64] za klasifikaciju histopatoloških medicinskih slika.

Deskriptori i algoritmi za računanje deskriptora navedeni u prethodnim radovima su primjenjeni na probleme pretraživanja baza slika ili nadgledane klasifikacije slika. Sa druge strane, nenadgledana klasifikacija je pokazala prilikom klasifikacije baza aero snimaka. Tako je nenadgledano učenje obilježja iskorišćeno u [65]. Tehnika gustog odabiranja je iskorišćena za nenadgledanu detekciju obilježja, koja su poslužila modelovanju skupa osnovnih funkcija. Zatim su obilježja niskog nivoa kodovana prema skupu osnovnih funkcija da bi se dobio kompaktan deskriptor slike. Kvaternionaska predstava deskriptora je iskorišćena za nenadgledano obučavanje prilikom klasifikacije aero snimaka u [66].

Najnoviji pristup iskorišćen za automatsku klasifikaciju baza slika koristi konvolucione neuronske mreže [67], opisuju vizuelni sadržaj slika učenjem koeficijentata uzastopnih konvolucionih filtara. Pošto su pokazale veliki potencijal, razmotreni su različiti načini njihovog korišćenja u obradi slike. Autori u [41] su testirali različite arhitekture neuronskih mreža i načine njihovog obučavanja za klasifikaciju aero snimaka. Takođe, u [68] je data analiza mogućnosti primjene konvolucionih neuronskih mreža za klasifikaciju aero snimaka i snimaka dobijenih tehnikama daljinske detekcije za slike u boji i multispektralne slike. Detaljna analiza mogućnosti primjene i prednosti korišćenja konvolucionih neuronskih mreža prilikom klasifikacije aero snimaka i snimaka dobijenih tehnikama daljinske detekcije je objavljena u [69]. Autori u [69] su došli do zaključka da su konvolucione neuronske mreže superiorne prilikom klasifikacije aero snimaka i detekcije objekata, ali da za pojedine probleme klasifikacije snimaka dobijenih tehnikama daljinske detekcije deskriptori niskog nivoa mogu imati bolje performanse.

Autori u [70] su koristili posebnu strategiju “višestrukih pogleda” prilikom obu-

čavanja neuronskih mreža za klasifikaciju aero snimaka. Za učenje mreže korišćeno je nekoliko dijelova sa svake od trening slika, čime su bolje obučene klase koje sadrže izolovane objekte. Diskusija oko primjene različitih arhitektura konvolucionih neuronskih mreža za klasifikaciju slika visoke rezolucije dobijenih tehnikama daljinske detekcije je objavljena u [71].

2.7. Istraživački trendovi

Na osnovu pregleda literature mogu se uočiti dva aktuelna istraživačka trenda u oblasti automatske klasifikacije slika. Prvi se odnosi na iznalaženje što bolje reprezentacije deskriptora obilježja niskog reda, kako bi se bolje prilagodili konkretnim problemima. Drugi se odnosi na razmatranje načina kako što bolje iskoristiti ogroman potencijal koji posjeduju konvolucione neuronske mreže. Pristup koji koristi deskriptore je superioran u slučajevima kada postoji mala količina trening podataka i prilikom nenadgledane automatske klasifikacije. Takođe, deskriptori mogu biti niskih dimenzionalnosti što znači brzo obučavanje klasifikatora i izvršavanje algoritma. Konvolucione neuronske mreže pokazuju superiorne rezultate u slučajevima kada postoji dovoljno velika količina trening podataka, ali treba voditi računa da je njihovo obučavanje računski i vremenski zahtjevno. Pogodnost kod neuronskih mreža jeste mogućnost pretreniranja na bazama slika drugačijih modaliteta [68, 15]. Ovakav pristup ipak ne daje zadovoljavajuće rezultate ako se neuronske mreže koriste za klasifikaciju multispektralnih slika [69], pa način obučavanja neuronskih mreža za multispektralne i hiperspektralne baze slika i dalje predstavlja otvoreno pitanje.

Konvolucione neuronske mreže su preuzele primat kada postoji velika količina trening podataka, ali odgovarajućih baza multispektralnih i hiperspektralnih slika za treniranje CNN-a još uvijek nema ili je njihovo formiranje veoma skupo. Npr. satelitski snimci ili medicinske U ovoj disertaciji, razmatrani su načini kako kreirati deskriptore niskih dimenzionalnosti pogodne za što tačniju klasifikaciju multispektralnih slika.

3 | Metodologija istraživanja

Kao što je već objašnjeno, automatska klasifikacija slika, kategorizacija scena i detekcija događaja, u suštini predstavljaju slične probleme, pa se u daljem tekstu razmatra isključivo problem automatske klasifikacije slika. U ovom poglavlju je dat pregled koraka potrebnih za implementaciju algoritama za nadgledanu automatsku klasifikaciju slika, kao i način za verifikaciju rezultata klasifikacije. Takođe, dat je pregled i opis baza slika korišćenih prilikom istraživanja predstavljenog u ovoj disertaciji.

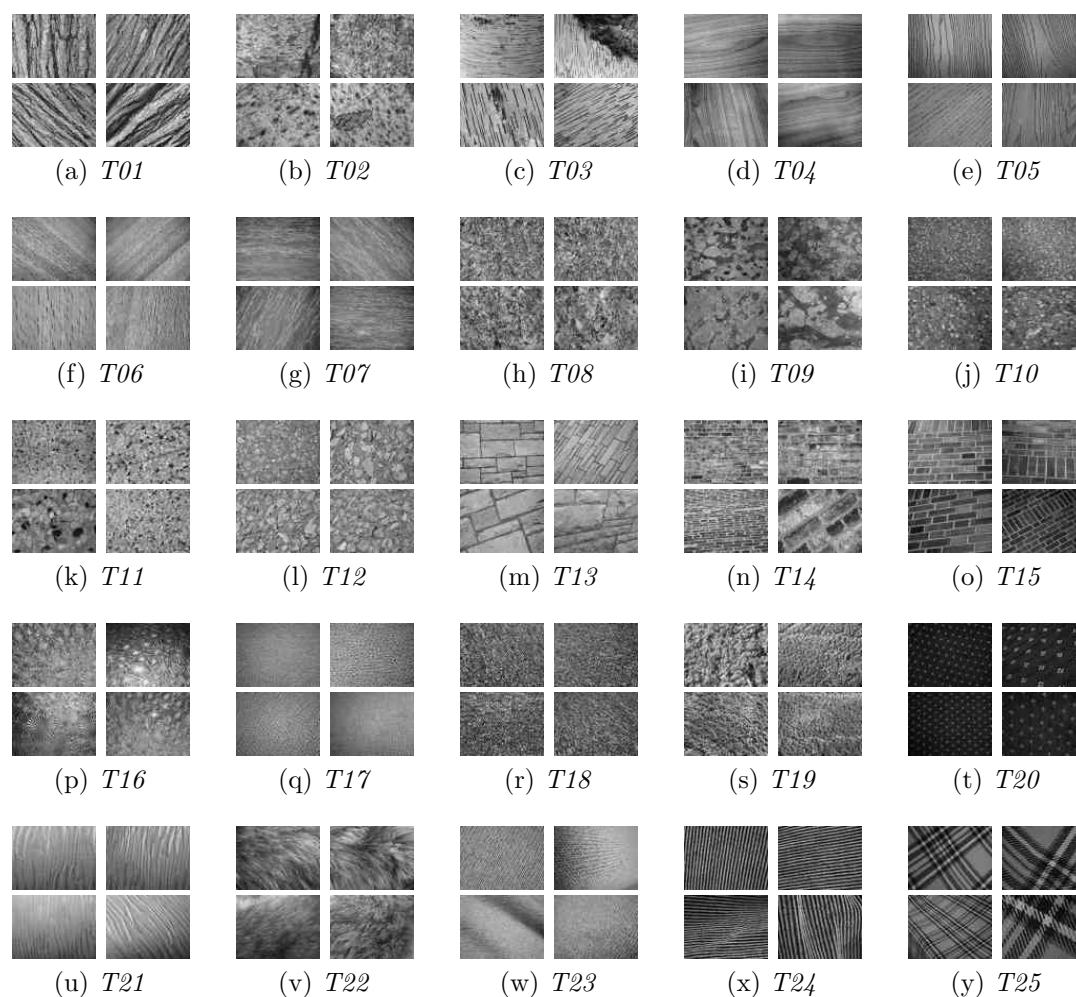
3.1. Korišćene baze slika

Prilikom istraživanja predstavljenog u ovoj disertaciji, korišćene su različite baze slika. Baze sivih slika su korišćene da bi se jasnije izvršila analiza uticaja dimenzionalnosti deskriptora, dok su baze slika u boji i baze multispektralnih slika korišćene za analizu uticaja različitih načina kombinovanja podataka iz različitih spektralnih opsega.

UIUC baza tekstura

UIUC je baza od 1000 nekalibrisanih i neregistrovanih sivih slika koje prikazuju teksture¹. Slike su ručno podjeljene u 25 klasa po 40 slika, veličine 640×480 piksela, kao što je opisano u [46]. Različite klase tekstura su definisane tako da prikazuju ravne površi homogene teksture prikazane iz različitih uglova, kao što su usmjerene teksture koje prikazuju kore od drveta, ili ravne površi obrađenog drveta. Takođe, mogu se uočiti ravne površi drugačijih vrsta materijala, kao što je voda, mermer ili cigla. Različite klase tekstura koje prikazuju šljunak i tkanine su definisane tako da prikazuju površi objekata posmatranih iz različitih uglova. Takođe, postoje klase sa slikama koje prikazuju površi objekata prikazanih ortogonalno i pod različitim uglovima, kao što su slike iz klasa tepih i cigla. Vidimo da se u UIUC bazi nalaze

¹UIUC baza slika je javno dostupna na adresi http://www-cvr.ai.uiuc.edu/ponce_grp/data/index.html



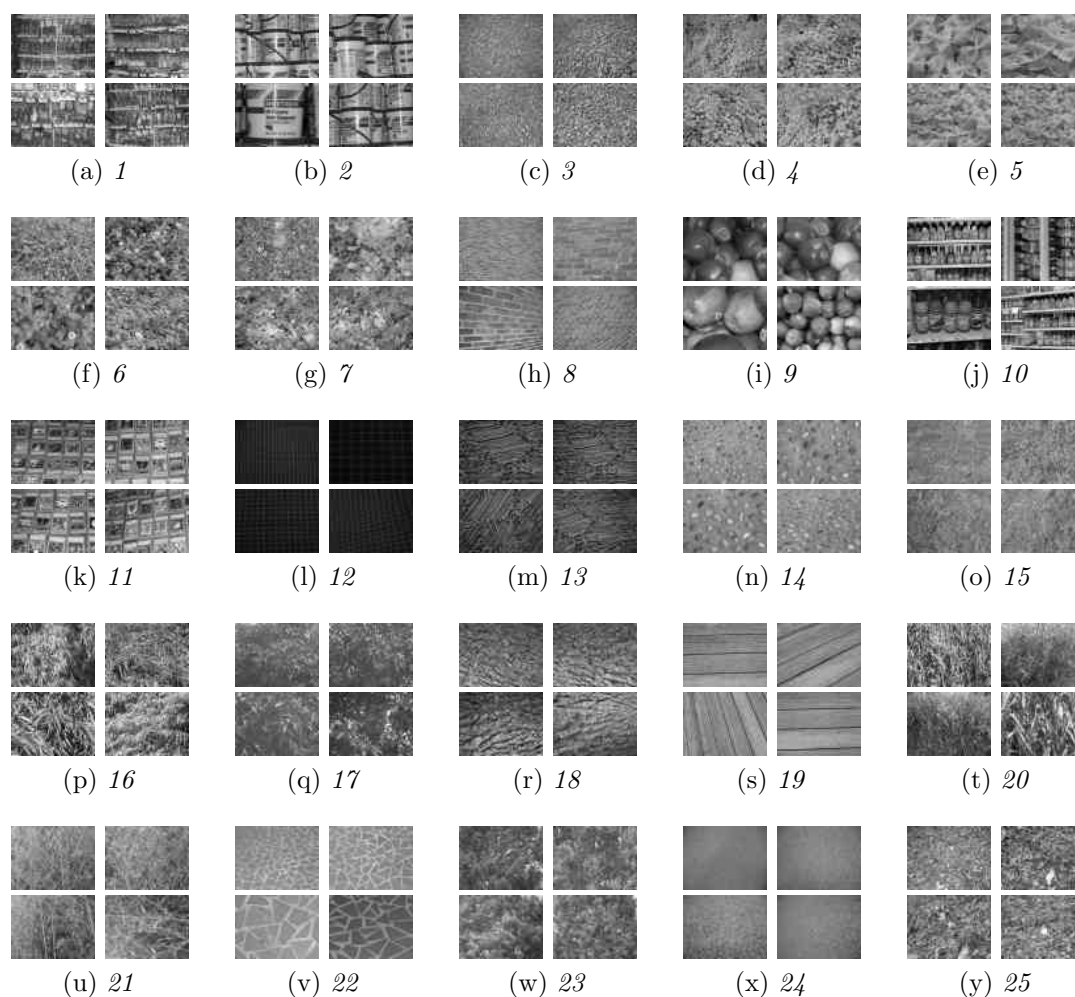
Slika 3.1: Primjeri slika iz baze UIUC

slike tekstura sa velikim stepenom vizuelne sličnosti između slika iz različitih klasa, sa različitim karakteristikama prikazanih površi i različitim uglovima prikazivanja. Klase tekstura su odabrane tako da mogu poslužiti za učenje razlike između objekata sa sličnim vizuelnim karakteristikama, prikazanih pod različitim uglovima.

UMD baza tekstura

UMD je baza sivih slika tekstura visoke rezolucije, predstavljena u [72], gdje je analizirana multifraktalna reprezentacija slika invarijantna na tačku gledišta². U bazi postoji ukupno 1000 slika tekstura veličine 1280×960 piksela, ručno podjeljenih u 25 klasa po 40 slika. Unutar svake klase, postoje značajne razlike u tačkama gledišta i skalama za koje su teksture predstavljene. Primjeri tekstura iz ove baze su prikazani na Sl. 3.2. Pored klasa tekstura koje prikazuju ravne površi iz različitih uglova, kao što su ciglasti i mermerni zidovi, zemljane površi, obrađeno i neobrađeno drvo, postoje i klase sa trodimenzionalnim teksturama lišća i niskog rastinja. Takođe, postoje i klase koje prikazuju različite posložene predmete, tako da

²UMD baza tekstura je javno dostupna na adresi <http://www.cfar.umd.edu/~fer/website-texture/texture.htm>.



Slika 3.2: Primjeri slika iz baze UMD

se obrazuje relativno uniformna tekstura, kao što su flaše, kante i drugi predmeti na rafovima ili jabuke ili špagete. UMD baza slika ima manji stepen vizuelne sličnosti između slika iz različitih klasa, ali veliki opseg skala i različite uglove gledišta unutar jedne klase. Stoga je pogodna za razmatranje različitih vrsta invarijantnosti deskriptora.

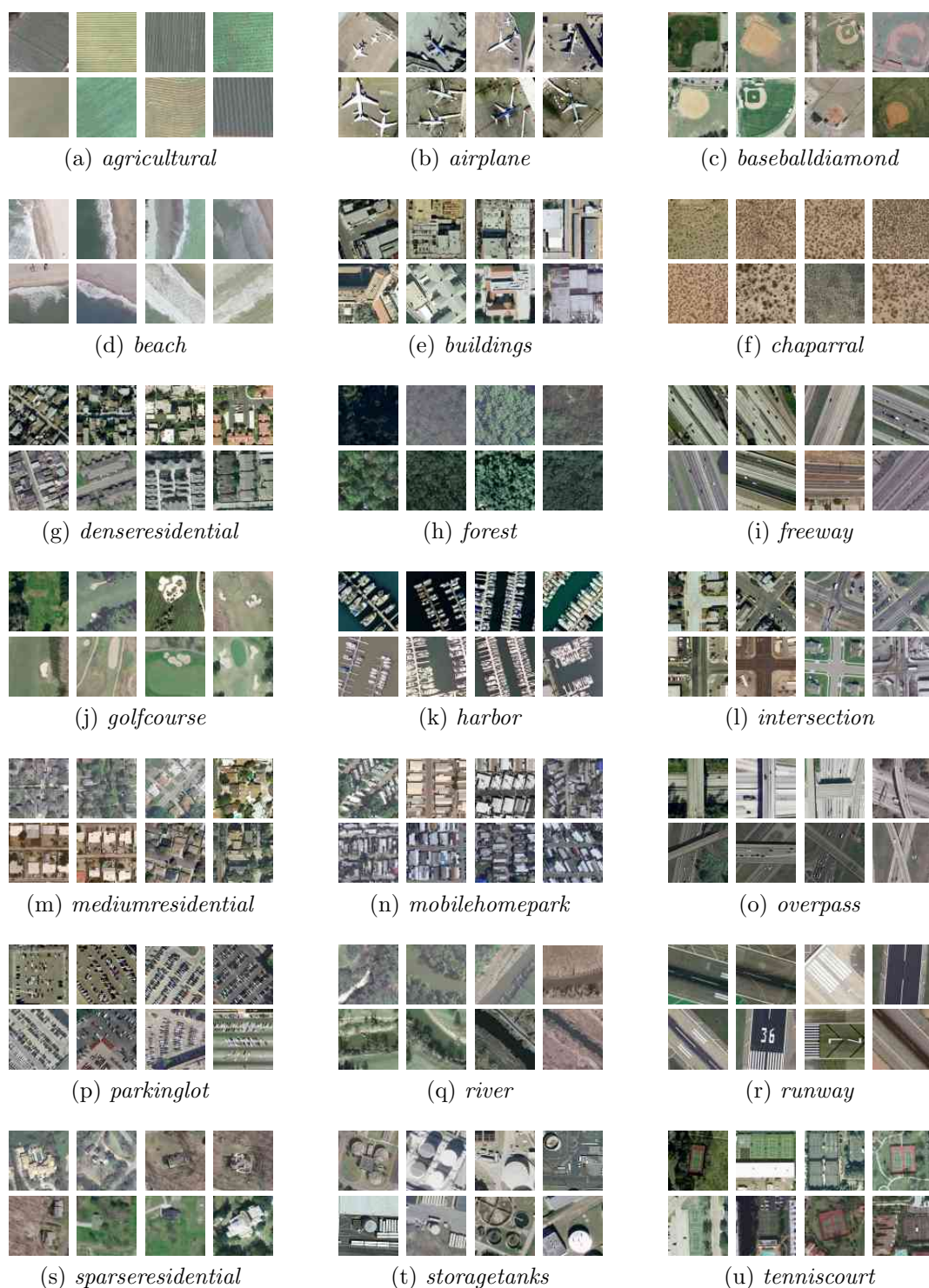
UC Merced baza aero snimaka

UCM je skup od 2100 obilježenih slika u boji preuzetih iz ortometrijskih aero snimaka visoke rezolucije, javno dostupnih na stranicama USGS nacionalne mape³ pogodnih za automatsku klasifikaciju. Sve slike su dimenzija 256×256 piksela, ručno podjeljenih u 21 različitim semantičkih klasa po 100 slika, čiji su primjeri dati na Sl. 3.3. Ova baza slika⁴ je prvobitno iskorišćena u radu [49] da bi se ispitao uticaj prostornog rasporeda lokalnih obilježja teksture prilikom klasifikacije aero snimaka. Slike iz klase *agricultural* prikazuju obradive površine i imaju jasno izbrazdanu teksturu svojstvenu za obrađeno zemljište. Međutim, različiti usjevi su

³United States Geological Survey (eng.) <http://www.usgs.gov/>

⁴UCM baza izdvojenih i ručno klasifikovanih slika je javno dostupna na adresi <http://vision.ucmerced.edu/datasets/landuse.html>

3.1. Korišćene baze slika



Slika 3.3: Primjeri slika iz baze UC Merced

različitih boja. Klasa *airplane* sadrži slike koje prikazuju jedan ili više putničkih aviona, dok klasa *baseballdiamond* sadrži slike koje prikazuju igralište za bejzbol. Ove dvije klase imaju slike koje prikazuju usamljene objekte sa različitim okolinama, pa njihove teksture nisu homogene kao u klasi *agricultural*. Slike iz klase *beach* prikazuju plažu i djelimično morske talase, pa imaju međusobno sličnu teksturu sa

različitom orijentacijom. Slike iz klasa *buildings*, *denseresidential* i *mediumresidential* prikazuju vizuelno slične ali semantički različite sadržaje. Sve te slike prikazuju stambena naselja, manje ili više gusto naseljena, s tim što klasi *buildings* pripadaju slike koje prikazuju velike stambene objekte u poređenju sa slikama iz klasa *denseresidential* i *mediumresidential*. Činjenica da je manja ili veća gustina naseljenosti kriterijum koji razdvaja klase *denseresidential* i *mediumresidential* ne pravi veliku vizuelnu razliku između slika iz ove dvije klase, ali treba naglasiti da takođe ne postoji jasna semantička razlika, osim proizvoljne ocjene gdje je manja a gdje veća naseljenost. Stoga, međusobno razlikovanje ove dvije klase predstavlja teži zadatak prilikom automatske klasifikacije. Takođe, može se uočiti velika vizuelna sličnosti između slika iz klasa *freeway* i *runway*, kao i između slika iz klasa *intersection* i *overpass*. Slično kao slike iz klasa *airplane* i *baseballdiamond*, slike iz klasa *sparseresidential*, *storagetanks* i *tenniscourt* prikazuju usamljene objekte sa velikim stepenom vizuelne različitosti u odnosu na ostale slike, ali sa okolinom koja nema posebno korisnih informacija o samim objektima. Dalje, može se uočiti da slike iz klase *mobilehomepark* imaju određenih vizuelnih sličnosti sa slikama iz klasa koje prikazuju naseljena mjesta. Takođe, slike iz klase *parkinglot* imaju sličnu teksturu, ali sa manjim izolovanim objektima. Slike iz klasa *river*, *forest* i *golfcourse* imaju relativno homogenu teksturu. Možemo primjetiti nekoliko klasa sa vizuelno sličnim, ali semantičkim različitim sadržajem. Na osnovu sadržaja slika, zaključuje se da ova baza slika može predstavljati problem automatske klasifikacije aero snimaka, detekcije usjeva, detekcije naseljenih mjesta i puteva, kao i detekcije usamljenih objekata na aero snimcima.

Multispektralna baza prirodnih scena

NirScene je baza multispektralnih slika koje prikazuju prirodne scene [29]. Baza sadrži ukupno 477 slika visoke rezolucije, različitih dimenzija, podijeljenih u 9 klasa⁵. Svaka slika sadrži tri kanala boja iz vidljivog spektralnog opsega i dodatni kanal sa podacima iz infra-crvenog spektralnog opsega. Primjeri slika iz svake klase ove baze dati su na Sl. 3.4, pri čemu su prikazane slike koje sadrže podatke iz vidljivog dijela spektra. Ova baza se sastoji od slika koje pojedinačno imaju veliki stepen međusobne vizuelne i semantičke sličnosti. Slike iz klase *country* prikazuju prirodne predjele, dok slike iz klase *field* daju prikaze njiva i poljana, vidimo da postoje pojedinačni primjeri slika iz različitih klasa, koje su veoma slične. Takođe, možemo da uočimo veliki stepen semantičke sličnosti između pojedinih slika iz klase *oldbuilding* i pojedinih slika iz klase *urban*. Ostale klase imaju nešto veći stepen semantičke različitosti, iako ponekad imaju vizuelnih sličnosti. Npr. slike iz klase *forest* prikazuju šumske predjele, dok slike iz klase *water* prikazuju pretežno vodene površine. Međutim, u nekoliko primjera, scene šume koje su prikazane na slikama iz klase *forest* prikazuju i riječne tokove, a takođe, u nekoliko primjera slika iz klase *water* prikazani su i djelovi sa zelenim rastinjem. Pored navedenih, definisane su klase *indoor*, *mountain* i *street*. Slike iz klase *indoor* prikazuju scene iz unutrašnjosti različitih prostorija, slike iz klase *mountain* prikazuju planinske vrhove dok slike iz klase *street* prikazuju različite tipove ulica. Analizom sadržaja slika može se za-

⁵NirScene baza slika je javno dostupna na adresi http://ivrl.epfl.ch/supplementary_material/cvpr11/.



Slika 3.4: Primjeri slika iz baze NirScene

ključiti veliki stepen semantičke sličnosti slika iz različitih klasa, ova baza predstavlja izazovan zadatak klasifikacije. Takođe, definisanje klase *indoor* omogućava testiranje algoritama za *indoor/outdoor* prepoznavanje, a postojanje dodatnog kanala sa podacima iz infracrvenog dijela spektra omogućava testiranje algoritama za obradu multispektralnih slika.

3.2. Metodologija nadgledane automatske klasifikacije slika

U Poglavlju 2 opisani su osnovni koraci neophodni prilikom nadgledane automatske klasifikacije slika, dok će u narednim sekcijama biti izloženi detaljni o koracima koji se koriste prilikom implementacije algoritma za nadgledanu automatsku klasifikaciju i načina na koji se vrši verifikacija rezultata. Eksperimenti su osmišljeni tako da je moguće porediti uticaj različitih koraka algoritma klasifikacije na tačnost klasifikacije. Sve korišćene baze slika prethodno su ručno obilježene, tako da se za svaku sliku nedvosmisleno zna kojoj klasi pripada.

Izdvajanje trening slika: Pošto nadgledana automatska klasifikacija podrazumjeva poznavanje ukupnog broja klasa i postojanje trening skupa slika, posebnu

pažnju treba obratiti na način izbora trening slika. Da bi se izvršilo poređenje rezultata klasifikacije kada se koriste algoritmi sa različitim parametrima, u svakom eksperimentu se koriste iste podjele na trening i test skupove slika. U pogledu podjele na trening i test slike, mogu se razmotriti tri scenarija:

- (1) Uopšte, različite klase mogu da sadrže različit broj slika. U prvom scenariju se za trening slike iz svake klase bira slučajni uzorak koji sadrži određeni procenat slika iz svake klase. Često se koristi podjela koja 50 % slika iz svake klase koristi za trening, a ostalih 50 % ta testiranje.
- (2) Da bi se obezbjedila dovoljna raznolikost trening skupa, tako da se sa većom sigurnošću može tvrditi da se prilikom obučavanja klasifikatora koriste svi relevantni podaci koji karakterišu svaku od klasa, koristi se što je moguće više slika za trening. U ovom scenariju, uzima se veliki broj trening slika odakle se može zaključiti koliko korisnih informacija o svakoj klasi deskriptor može da izdvoji. Takođe, veliki broj trening slika treba koristiti kada se radi o obradi i klasifikaciji veoma velikih baza slika, gdje je od velikog praktičnog značaja obezbjeđivanje što više raznovrsnih informacija o svakoj klasi.
- (3) U različitim praktičnim primjenama na raspolaganju je mali broj trening slika u odnosu na ukupan broj slika koje je potrebno klasifikovati. Stoga je korisno razmotriti slučajeve kada se za obučavanje klasifikatora koristi mali broj slika iz svake od klasa.

Računanje deskriptora: Računanje deskriptora slika obavezno uključuje izdvajanje obilježja i formiranje vektora obilježja i deskriptora, a moguće je uraditi postprocesiranje deskriptora ili kombinovanje nekoliko različitih deskriptora u jedan. Takođe, moguće je izvršiti pretprocesiranje slika, prije detekcije obilježja, kao što je filtriranje ili konverzija u određeni prostor boja. U eksperimentima će posebno biti analiziran uticaj različitih deskriptora, različitih pristupa normalizaciji pojedinačnih deskriptora, način kombinovanja deskriptora iz različitih spektralnih opsega, zatim različiti pristupi redukciji dimenzionalnosti deskriptora i različiti načini korišćenja informacija o prostornom rasporedu lokalnih obilježja.

- (a) **Pretprocesiranje** Slike su najčešće predstavljene u RGB kolor prostoru, ali određena obilježja i lokalne karakteristike moguće je jednostavnije detektovati u nekom drugom kolor prostoru. Takođe, određeni kolor prostori su definisani tako da ne postoji korelisanost podataka iz različitih kanala, što može obezbijediti veće mogućnosti redukcije dimenzionalnosti. Pored standardnog RGB kolor prostora, u literaturi se često koriste HSV prostor⁶, LAB kolor prostor. HSV kolor prostor se koristi u alatima za editovanje digitalnih slika i računarskoj grafici, dok LAB kolor prostor koji služi za efikasnu razmjenu podataka između različitih uređaja. U praktičnim aplikacijama pretraživanja baza slika na osnovu sadržaja i automatske klasifikacije slika, korisno je koristiti oponentni kolor prostor oRGB [73, 74, 11], koji je definisan da bude invarijantan na intenzitet svjetlosti na slici. Transformacija iz RGB u oRGB kolor prostor

⁶Hue Saturation Value color space (eng.)

se može obaviti sljedećom jednačinom:

$$\begin{bmatrix} O_1 \\ O_2 \\ O_3 \end{bmatrix} = \begin{bmatrix} 0.7071 & -0.7071 & 0 \\ 0.4082 & 0.4082 & -0.8165 \\ 0.5774 & 0.5774 & 0.5774 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.1)$$

Reprezentacija slika u oponentnom kolor prostoru ne garantuje međusobnu nekorelisanost podataka iz različitih kanala. Pored oponentnog kolor prostora, može se koristiti hiperoponentni kolor prostor [11], koji se definiše tako da obezbijedi nekorelisanost podataka iz različitih kanala. Transformacija za prelazak u hiperoponentni kolor prosto se dobija računanjem PCA transformacije prema podacima iz RGB kanala ili RGB+NIR kanala [29, 11, 13]. Ovom transformacijom se vrši dekorelacija podataka sa slika iz baze koja se klasifikuje, pa je transformacija vezana isključivo za određenu bazu slika. PCA transformacija prema RGB prostoru i PCA transformacija prema RGB+NIR prostoru računaju se iz određenog broja slučajno izabranih slika iz baze. Hiperoponentni kolor prostor izračunat prema podacima iz dostupnih baza slika na osnovu PCA transformacije će u daljem tekstu disertacije biti označeni sa PCA-RGB i PCA-RGB+NIR prema potrebi.

- (b) **Deskriptori** Izdvajanje vektora obilježja se radi prema nekom od algoritama opisanim u prethodnom poglavlju. Prilikom selekcije algoritma za izdvajanje obilježja potrebno je voditi računa o tome da su pojedini algoritmi projektovani da detektuju određenu vrstu obilježja. U narednim eksperimentima su korišćeni GIST, SIFT, LBP i CENTRIST deskriptori.
- (c) **Normalizacija** Ukoliko posmatramo jedan N -dimenzionalni deskriptor sive slike:

$$\mathbf{x} = [x_1, x_2, x_3, \dots, x_{N-1}, x_N], \quad (3.2)$$

prije kombinovanja sa ostalim, moguće je izvršiti normalizaciju njegovih vrijednosti. Normalizacija deskriptora (3.2) podrazumijeva skaliranje njegovih pojedinačnih elemenata x_i , $i = 1 \dots N$, tako da se obezbijedi jedinična norma deskriptora. Normalizacija se može izvesti prema različitim metričkim kriterijumima, ali uvijek je prvi korak centriranje deskriptora. Centriranje predstavlja oduzimanje srednje vrijednosti deskriptora od svakog njegovom elementa $\mathbf{x}_c \leftarrow \mathbf{x} - E(\mathbf{x})$. Ako se koristi L1 metrika, normalizovani deskriptor se dobija kao:

$$\mathbf{x}_{\text{norm}}^{\text{L1}} = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|_1}, \quad (3.3)$$

ako se koristi L2 metrika prilikom normalizacije centriranog deskriptora, dobija se:

$$\mathbf{x}_{\text{norm}}^{\text{L2}} = \frac{\mathbf{x}_c}{\sqrt{\|\mathbf{x}_c\|_2^2}}. \quad (3.4)$$

Takođe, koriste se i varijante normalizacije koje uključuju korijenovanje vrijednosti pojedinačnih odbiraka deskriptora. U [75, 76] je razmatran pristup koji uključuje normalizaciju prema L1 metrici i potom računanje kvadratnog korijena svakog odbirka.

$$\mathbf{x}_{\text{norm}}^{\text{sL1}} = \sqrt{\frac{\mathbf{x}_c}{\|\mathbf{x}_c\|_1}}. \quad (3.5)$$

U [76] predlaže se normalizacija deskriptora koja uključuje L2 normalizaciju necentriranog deskriptora, zatim L1 normalizaciju i konačno korijenovanje svakog elementa vektora posebno. Ovako normalizovan deskriptor će biti obilježen sa oznakom $\mathbf{x}_{\text{norm}}^{\text{rL1}}$.

- (d) **Redukcija dimenzionalnosti** Bez obzira na to koja tehnika se koristi za redukciju dimenzionalnosti, za računanje odgovarajuće transformacije koriste se deskriptori slika iz trening skupa. Ukoliko je na raspolaganju ukupno M deskriptora dimenzionalnosti N , ti deskriptori se predstavljaju matricom:

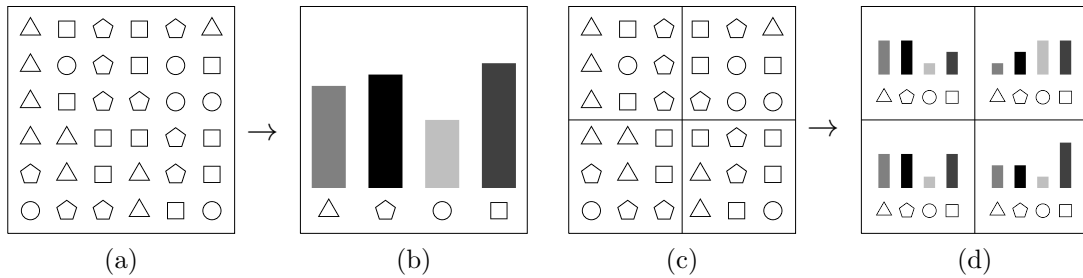
$$\tilde{\mathbf{X}}_{M \times N} = \begin{matrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{matrix} \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{M,1} & x_{M,2} & x_{M,3} & \dots & x_{M,N} \end{bmatrix} \quad (3.6)$$

i računa se transformacija $\mathbf{T}_{N \times K}$, koja redukuje dimenzionalnost svakog deskriptora sa N na K :

$$\tilde{\mathbf{X}}_{M \times K} = \tilde{\mathbf{X}}_{M \times N} \cdot \mathbf{T}_{N \times K}, \quad (3.7)$$

pri čemu se čuva maksimalna količina informacija u $\tilde{\mathbf{X}}_{M \times K}$. Transformacija $\mathbf{T}_{N \times K}$ se može računati prema različitim kriterijumima, kao što je opisano u prethodnom poglavlju.

- (e) **Kombinovanje deskriptora** U slučaju slika u boji i multispektralnih slika, svaki od dostupnih kanala je tretiran kao jedna siva slika iz koje su računati deskriptori. U toku ovog istraživanja, posebna pažnja je posvećena načinu kombinovanja deskriptora dobijenih iz različitih kanala. Način kombinovanja deskriptora iz različitih spektralnih opsega je povezan sa tehnikom redukcije dimenzionalnosti. Detalji oko implementacije algoritama za računanje nisko-dimenzionalnih deskriptora iz multispektralnih slika su dati u Poglavlju 5.
- (f) **Prostorni raspored lokalnih obilježja** Ukoliko se koriste lokalni deskriptori, moguće je implementirati algoritme za računanje deskriptora u zavisnosti od rasporeda lokalnih obilježja. Prema [8, 56] uzimanjem u obzir pozicije obilježja na slikama, moguće je bolje opisati suštinu prikazane scene. Primjer računanja deskriptora prema rasporedu lokalnih obilježja dat je na Sl. 3.5. Različita obilježja su simbolički prikazana različitim geometrijskim oblicima, pri čemu njihovo mjesto na slici odgovara položaju na kojem su detektovani, što je prikazano na Sl. 3.5a. Ukoliko se ne uzima prostorni raspored, histogram lokalnih obilježja je prikazan na Sl. 3.5b. Poređenjem slika koje su



Slika 3.5: Korišćenje rasporeda lokalnih obilježja prilikom računanja deskriptora

opisane takvim deskriptorima poredi se koliko lokalnih obilježja sa određenim karakteristikama ima na kojoj slici, međutim, ne uzima se u obzir gdje se one nalaze. Ukoliko se prostor slike podijeli na četiri dijela, kao što je prikazano na Sl. 3.5c, jednoj slici se pridružuje četiri histograma lokalnih obilježja, koji sadrže informacije o tome koliko lokalnih obilježja sa određenim karakteristikama ima na određenim dijelovima slike. Poređenjem slika pomoću ovakvih deskriptora uzima se u obzir i lokalni raspored karakteristika, ali se povećava dimenzionalnost deskriptora.

- (g) **Postprocesiranje** Nakon računanja deskriptora slike, bilo da se radi o sivim ili multispektralnim slikama, i da li se koristi normalizacija, redukcija dimenzionalnosti i podaci o rasporedu lokalnih obilježja, moguće je izvršiti dodatnu obradu deskriptora. Postprocesiranje može poslužiti daljoj redukciji dimenzionalnosti deskriptora, normalizaciji, dekorelaciji ili naglašavanju određenih komponentata deskriptora. U narednim eksperimentima će biti razmotren uticaj dodatne normalizacije, sa ciljem što uniformnijeg predstavljanja podataka prilikom obučavanja klasifikatora. U ovom slučaju, počinje se od deskriptora slika iz trening skupa predstavljenih matricom $\tilde{\mathbf{X}}_{M \times N}$ ili $\tilde{\mathbf{X}}_{M \times K}$, ukoliko je izvršena redukcija dimenzionalnosti. Cilj je skaliranje pojedinačnih odbiraka svakog od deskriptora slika iz trening skupa, kako bi njihove vrijednosti bile u opsegu od 0 do 1. Računaju se minimalne vrijednosti po svakoj od dimenzija $i = 1 \dots N$:

$$\begin{aligned} \tilde{\mathbf{X}}_{M \times N} &= \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{M,1} & x_{M,2} & x_{M,3} & \dots & x_{M,N} \end{bmatrix}, \\ &\Downarrow \\ \tilde{\mathbf{X}}_{1 \times N}^{\min} &= \begin{bmatrix} x_{\min}^{(1)} & x_{\min}^{(2)} & x_{\min}^{(3)} & \dots & x_{\min}^{(N)} \end{bmatrix} \end{aligned} \quad (3.8)$$

gdje je $x_{\min}^{(i)} = \min(x_{ji})$, $j = 1 \dots M$, i maksimalne vrijednosti po svakoj od dimenzija:

$$\begin{aligned} \tilde{\mathbf{X}}_{M \times N} &= \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{M,1} & x_{M,2} & x_{M,3} & \dots & x_{M,N} \end{bmatrix}, \\ &\Downarrow \\ \tilde{\mathbf{X}}_{1 \times N}^{\max} &= \begin{bmatrix} x_{\max}^{(1)} & x_{\max}^{(2)} & x_{\max}^{(3)} & \dots & x_{\max}^{(N)} \end{bmatrix} \end{aligned} \quad (3.9)$$

gdje je $x_{\max}^{(i)} = \max(x_{ji})$, $j = 1 \dots M$. Normalizacija vrijednosti pojedinačnih odbiraka deskriptora se izvršava primjenom jednačine:

$$\tilde{\mathbf{X}}_{M \times N}^u = \left(\tilde{\mathbf{X}}_{M \times N} - \begin{bmatrix} \tilde{\mathbf{X}}_{1 \times N}^{\min} \\ \tilde{\mathbf{X}}_{1 \times N}^{\min} \\ \vdots \\ \tilde{\mathbf{X}}_{1 \times N}^{\min} \end{bmatrix}_{M \times 1} \right) \left(\tilde{\mathbf{X}}_{1 \times N}^{\max} - \tilde{\mathbf{X}}_{1 \times N}^{\min} \right)^{-1}. \quad (3.10)$$

Obučavanje klasifikatora: Kao što je definisano u Poglavlju 2, klasifikator definiše način na koji se odlučuje kojoj klasi pripada koja slika na osnovu njenog deskriptora. Prilikom izvođenja eksperimenata predstavljenih u ovoj disertaciji koristi se SVM klasifikator koji može modelovati kompleksne odnose između deskriptora različitih klasa čak i kada ne postoji linearna separabilnost između deskriptora slika iz različitih klasa. Deskriptori testnih slika se prosljeđuju SVM klasifikatoru sa linearnih kernelom, osim ako posebno nije naglašeno da se radi o drugom kernelu. Koristi se SVMlib [40] implementacija SVM klasifikatora.

Računanje deskriptora za slike koje se klasifikuju: Da bi se mogla izvršiti klasifikacija slika iz testnog skupa, potrebno je izračunati njihove deskriptore. Deskriptori testnih slika treba da budu procesirani istim redoslijedom i istim transformacijama kao i deskriptori trening slika. Ukoliko je izvršena promjena kolor prostora, normalizacija i redukcija dimenzionalnosti deskriptora trening slika, iste transformacije treba primijeniti i na deskriptore testnih slika, kako bi sačuvala jedinstvenost reprezentacije podataka.

Klasifikacija i verifikacija rezultata: Nakon što je izvršeno računanje deskriptora slika iz testnog skupa, vrši se klasifikacija. Obučeni klasifikator predstavlja matematički model u prostoru obilježja koji razgraničava dijelove prostora koji pripadaju određenim klasama. Klasifikacija jedne slike predstavlja određivanje kojem dijelu prostora pripada deskriptor te slike u prostoru obilježja. Nakon što se svakoj slici pridruži oznaka klase u koju je klasifikovana, može se provjeriti tačnost klasifikacije, tj. izvršiti verifikacija rezultata. Slika je tačno klasifikovana ukoliko je klasa pridružena slici od strane klasifikatora ista kao i prethodno definisana klasa u bazi slika, u suprotnom klasifikacija je pogrešno izvršena. Pošto obučavanje klasifikatora zavisi od slika iz trening skupa, bolji ili lošiji izbor trening slika može značajno uticati na rezultate klasifikacije. Stoga je poželjno ponoviti proces klasifikacije za različite trening/test podjele slika iz baze.

- (1) Unakrsna validacija jeste tehnika validacije koja služi za procjenu kako se rezultati statističke analize mogu generalizovati. Često se koristi prilikom testiranja modela predikcije, a prilikom automatske klasifikacije slika korisno je upotrebiti tzv. k -tostruka unakrsnu validaciju⁷. Ukupan broj slika se dijeli na k disjunktivnih skupova, od kojih se jedan bira za trening a ostali se koriste za test. Cjelokupan proces se ponavlja dok se svaki od k skupova ne iskoristi kao trening, tako da se dobija ukupno k različitih rezultata klasifikacije. Konačan rezultat klasifikacije se dobija usrednjavanjem svih k rezultata sa različitim trening/test podjelama. U eksperimentima se najčešće koristi 5-tostruka krosvalidacija.
- (2) Trening skup je moguće definisati slučajnim izborom određenog broja slika iz svake od klasa. Ponavljanjem ovog slučajnog izbora više puta dobijaju se različiti preklapajući skupovi trening slika. Za svaku od ovako definisanih trening/test podjela se izvršava klasifikacija i konačni rezultati klasifikacije se dobijaju usrednjavanjem rezultata dobijenih za svaku od slučajnih podjela.

Predstavljanje rezultata klasifikacije: S obzirom da se u svakom eksperimentu vrši klasifikacija za nekoliko različitih trening/test podjela slika iz baze, pored

⁷*k-fold cross validation* (eng.)

srednje vrijednosti tačnosti klasifikacije računa se i standardna devijacija rezultata. Ukoliko je standardna devijacija velika, može se zaključiti da testirani algoritam za klasifikaciju mnogo zavisi od izbora trening slika. U slučaju male devijacije, dobijeni rezultati su pouzdani. Više detalja o tačnosti klasifikacije se može vidjeti iz matrice konfuzija. Matrica konfuzija pokazuje koliko procenata slika iz određene klase je klasifikovano u svaku od definisanih klasa. Analizom matrice konfuzija moguće je ustanoviti koliko je određeni deskriptor uspio prepoznati vizuelne ili semantičke razlike između klasa slika. U slučaju većeg broja klasa, matricu konfuzija je korisno predstaviti grafički. Kada je potrebno izvršiti analizu uticaja određenog parametra na tačnost klasifikacije, obično se koristi tabelarno ili grafičko predstavljanje tačnosti klasifikacije u odnosu na vrijednost parametra.

3.3. Problem vizuelne sličnosti slika i tačnosti klasifikacije

Tačnost klasifikacije slika može zavistiti od vrijednosti velikog broja parametara i od izbora koraka klasifikacije koji se implementiraju. Kroz naredne eksperimente će biti razmatran uticaj nekoliko koraka klasifikacije, kao što je pretprocesiranje slika ili postprocesiranje deskriptora. Takođe, biće razmotren i uticaj smanjene dimenzionalnosti deskriptora na tačnost klasifikacije.

Pošto se u ovoj disertaciji razmatra nadgledana klasifikacija slika, potrebno je poznavati karakteristike slika iz svake klase, prije nego što se obučiti klasifikator. Rezultat klasifikacije jedne slike je isključiv, tj. slika se može pridružiti isključivo jednoj klasi. Rezultat je tačan ukoliko je tako određena klasa ista kao i klasa kojoj je slika prethodno pripisana. Upravo je to drugi razlog zašto ponekad rezultate klasifikacije treba uzeti sa rezervom, jer različiti korisnici bi zasigurno na različite načine izvršili manuelnu klasifikaciju bilo koje baze slika.

Dalje, kada se porede rezultati klasifikacije za dvije različite metode, u oba eksperimenta treba omogućiti što je moguće više jednake uslove, kao što je jednaka podjela na trening i test slike ili jednake vrijednosti parametara. Međutim, treba naglasiti da različiti deskriptori mogu imati najbolje rezultate klasifikacije za potpuno različite vrijednosti parametara klasifikatora. Iako se ovaj problem donekle može riješiti jednakom metodom normalizacije, on predstavlja treći razlog zašto je tačnost klasifikacije relativna za ocjenu kvaliteta klasifikacije.

4 Automatska klasifikacija sivih slika prema teksturi

Performanse algoritama za automatsku klasifikaciju mogu zavisi od izbora velikog broja parametara i izbora odgovarajućih koraka realizacije algoritma za klasifikaciju. U ovom poglavlju će biti izvršena analiza performansi algoritama za klasifikaciju na bazama sivih slika, kako bi se bolje uočili uticaji pojedinih koraka koji se mogu uključiti u algoritam. Pošto tačnost klasifikacije zavisi od parametara klasifikatora, svi eksperimenti će biti izvršeni pod istim uslovima, koristeći linearni SVM klasifikator, da bi se omogućilo korektno poređenje različitih načina za računanje deskriptora. Na osnovu rezultata eksperimenata, biće analiziran uticaj pojedinih metoda za normalizaciju deskriptora i redukciju njihove dimenzionalnosti. Optimalne vrijednosti parametara dobijene prilikom ove analize će biti iskorišćene u daljim eksperimentima.

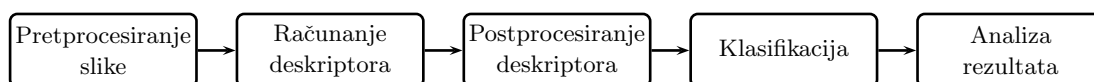
S obzirom da u literaturi postoji veliki broj deskriptora, od kojih su neki posebno dizajnirani za specifične zadatke, za naredne eksperimente su izabrani deskriptori različitih karakteristika, kako bi se dobio što precizniji pregled mogućnosti razmatranih tehnika. U Poglavlju 2 dat je opis nekoliko deskriptora teksture sivih slika. U daljem istraživanju će biti korišćene varijante GIST, CENTRIST, LBP i SIFT deskriptora i to iz sljedećih razloga. GIST deskriptor [8] se koristi kao predstavnik globalnih deskriptora teksture koji je pokazao dobre performanse prilikom klasifikacije scena i aero snimaka [27, 12]. CENTRIST je lokalni deskriptor, jednostavan za računanje i koristi mali broj susjednih piksela za računanje lokalnih obilježja. U narednim eksperimentima će biti korišćen uprošćeni algoritam za računanje ovog deskriptora¹, kako bi se što bolje ispitale njegove suštinske karakteristike. LBP je nešto komplikovaniji u poređenju sa CENTRIST-om, jer koristi interpolaciju i multirezolucioni pristup analizi. Biće korišćena njegova uniformna i rotaciono-invarijantna verzija [9]. SIFT se koristi jer je lokalna obilježja izračunata pomoću SIFT detektora moguće kombinovati na veliki broj načina prilikom različitih zadataka automatske

¹Uprošćeni algoritam za računanje CENTRIST deskriptora koristi samo Census transformaciju i formiranje histograma.

klasifikacije slika [55, 62, 63, 29, 76, 53, 13, 51]. Pri tome različite varijante SIFT deskriptora pokazuju dobre diskriminativne karakteristike. Podešavanjem parametara prilikom agregacije lokalnih SIFT obilježja moguće je kreirati SIFT deskriptor različitih dimenzionalnosti.

4.1. Postavka eksperimenata

Šematski prikaz algoritma za automatsku klasifikaciju prikazan je na Sl. 4.1. U slučaju da se posebno analizira uticaj nekog od koraka, to će biti naglašeno na dijagramu toka sjenčenjem odgovarajućeg koraka, i dato uz rezultate eksperimenata.



Slika 4.1: Šematski prikaz algoritma za automatsku klasifikaciju sivih slika

U narednim eksperimentima posebno je ispitan uticaj izbora različitih metoda za pretprocesiranje, izbor različitih deskriptora i metoda za postprocesiranje deskriptora, zatim izbor različitih parametara klasifikatora, na tačnost klasifikacije.

- (a) **Pretprocesiranje slike:** Neki od najjednostavnijih koraka za pretprocesiranje uključuju promjenu dimenzija slike ili transformaciju u drugi kolor prostor. Pošto se u narednim eksperimentima radi isključivo sa sivim slikama, biće ispitan uticaj količine ulaznih podataka, tj. tačnost klasifikacije u zavisnosti od veličine ulaznih slika. Što su slike manjih dimenzija, potrebno je manje računskih operacija za računanje deskriptora i manje memorije. Konkretno, biće ispitani slučajevi kada se ulazna slika smanji na 50%, 20% i 10% u odnosu na svoje prvobitne dimenzije.
- (b) **Izdvajanje deskriptora:** Kao što je već naglašeno, biće ispitan uticaj izbora GIST, CENTRIST, LBP i SIFT deskriptora.
- (c) **Postprocesiranje deskriptora:** Nakon računanja, deskriptori se mogu dodatno numerički obraditi, u skladu sa potrebama konkretnog problema. U narednim eksperimentima će posebno biti ispitan uticaj normalizacije deskriptora, i to L1 normalizacije (3.3), L2 normalizacije (3.4) i korijenske L1 normalizacije (3.5) i normalizacije predložene u [76]. Takođe, biće ispitan uticaj redukcije dimenzionalnosti deskriptora pomoću PCA transformacije, tj. koliko je moguće smanjiti dimenzionalnost deskriptora, a da se značajno ne naruše performanse klasifikatora. Na kraju, ispituje se uticaj skaliranja deskriptora radi njihove uniformne reprezentacije, kao što je dato jednačinom (3.10).
- (d) **Izbor parametara klasifikatora:** Rezultati klasifikacije zavise od parametara klasifikatora i načina njegovog obučavanja. Stoga, biće razmotren uticaj izbora kernela SVM klasifikatora na tačnost klasifikacije. U slučajevima kada nije moguće postići linearnu separabilnost koristi se regularizacioni parametar C kako bi se dozvolila određena odstupanja i omogućilo izračunavanje odgo-

varajućih hiperravni². Kako bi se omogućilo poređenje rezultata klasifikacije iz jednog skupa eksperimenata, regulacioni parametar se postavlja na jedinstvenu vrijednost za razmatrani eksperiment. Ukoliko je potrebno, vrijednost korišćenog regulacionog parametra je posebno naglašena. Više detalja o načinu obučavanja SVM klasifikatora može se naći u [35].

U rezultatima gdje je naglašen broj trening slika po klasi, iz svake klase je slučajno izabran naznačeni broj slika za trening, dok su sve ostale slike korišćene prilikom testiranja. Cijeli postupak podjele na trening i test slike je ponovljen 5 puta i konačna tačnost klasifikacije se dobija kao srednja vrijednost tačnosti klasifikacije dobijenih za svih 5 pojedinačnih podjela. Za sve eksperimente izvršena je jedinstvena podjela na trening i test skup slika, tako da se može izvršiti što korektnije poređenje performansi različitih varijanti klasifikatora. Drugim riječima, u svakom eksperimentu na određenoj bazi slika gdje se koristi jedan određeni broj trening slika po klasi, koristi se ista podjela na trening/test slike za svaku od 5 pojedinačnih klasifikacija.

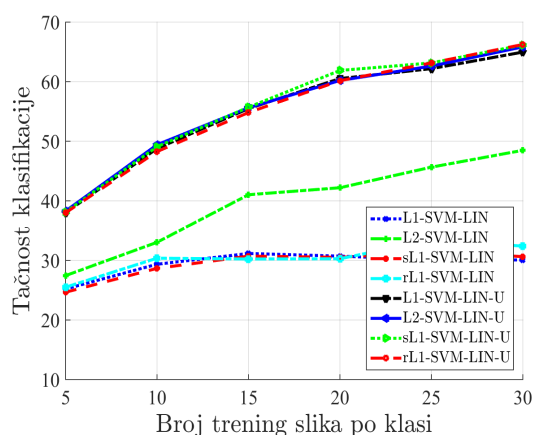
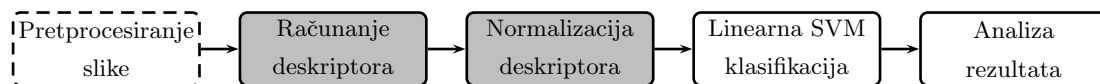
4.2. Rezultati automatske klasifikacije sivih slika

U narednim eksperimentima korišćene su baze sivih slika UIUC i UMD, koje su opisane u Poglavlju 3. Koristi se SVM klasifikator sa linearnim kernelom, a vrijednost regularizacionog parametra je $C = 1$. U posebnim eksperimentima ispituju se uticaji različitih pristupa normalizaciji deskriptora, redukciji dimenzionalnosti slika i redukcije dimenzionalnosti deskriptora.

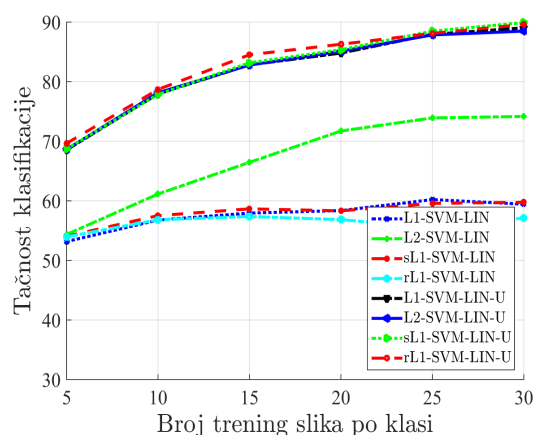
Za sve eksperimente sa GIST deskriptorom, korišćene su 4 orijentacije sa po 8 skala, što je pokazano kao optimalan izbor u [12]. Pošto se filtriranje radi na 16 nepreklapajućih podslika, računa se 16 srednjih vrijednosti filtrirane podslike za 8 različitih skala i 4 različite orijentacije, pa je dimenzionalnost GIST deskriptora za jednu sivu sliku jednaka $16 \times 8 \times 4 = 512$. Prilikom računanja SIFT deskriptora može se postaviti nekoliko različitih parametara. Ukoliko drugačije nije naglašeno, u svim narednim eksperimentima koriste se blokovi veličine 32×32 prilikom računanja lokalnih vektora obilježja. Lokalna obilježja su računata prema principu gustog odabiranja sa razmakom od 2 piksela. Pošto konačni SIFT deskriptor zavisi od broja kodnih riječi koje se računaju prema BoVW paradigmi, u narednim eksperimentima je korišćeno 1000 kodnih riječi, jer je ustanovljeno da veći broj kodnih riječi obezbjeđuje dobre rezultati klasifikacije [13]. Dimenzionalnost SIFT deskriptora zavisi od broja kodnih riječi, a dimenzionalnost SIFT deskriptora za jednu sivu sliku jednaka je 1000. Prilikom računanja CENTRIST deskriptora, računa se histogram Census transformacije originalne sive slike, što predstavlja pojednostavljenu varijantu deskriptora. Pošto se vrijednosti Census transformacije jednake 1 i 256 ne uzimaju u obzir, konačni CENTRIST deskriptor ima dimenzionalnost 254. LBP deskriptor se koristi u uniformnoj i rotaciono-invarijantnoj verziji, pri čemu se lokalni binarni obrazac računa na tri rezolucije. Koriste se poluprečnici 1, 2 i 3 sa po 8, 16 i 24 tačke. Uniformni LBP je deskriptor dimenzionalnosti za dva veće od broja

²Princip meke margine *Soft margin method* (eng.)

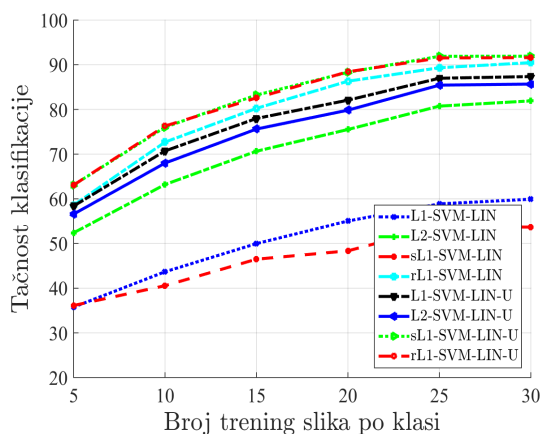
4.2. Rezultati automatske klasifikacije sivih slika



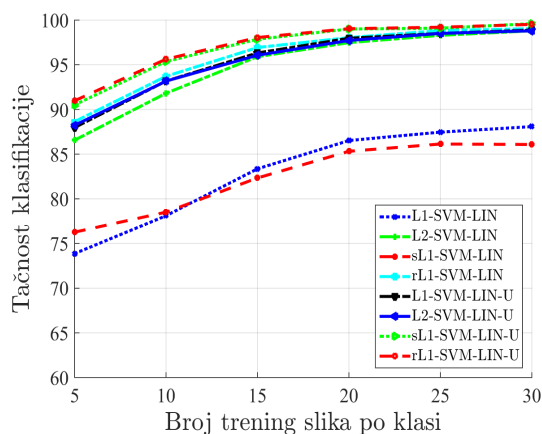
(a) UIUC baza, GIST deskriptor



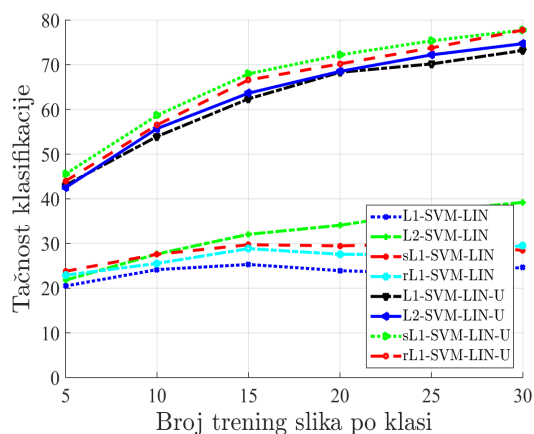
(b) UMD baza, GIST deskriptor



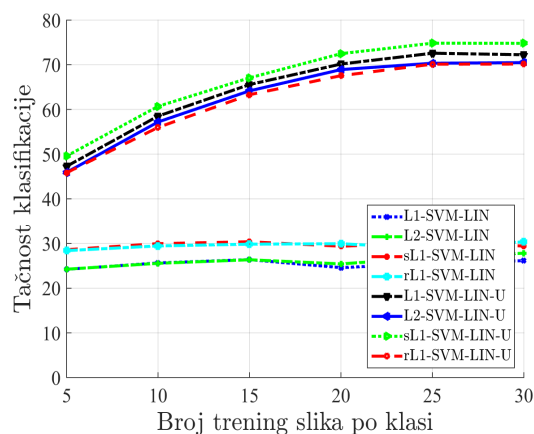
(c) UIUC baza, SIFT deskriptor



(d) UMD baza, SIFT deskriptor



(e) UIUC baza, CENTRIST deskriptor



(f) UIUC baza, LBP_{riu2} deskriptor

Slika 4.2: Analiza uticaja normalizacije deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika

tačaka u kojima se računa lokalni obrazac, pa je ukupna dimenzionalnost korišćenog LBP deskriptora za jednu sivu sliku jednaka $10 + 18 + 26 = 54$.

Uticaj različitih metoda za normalizaciju

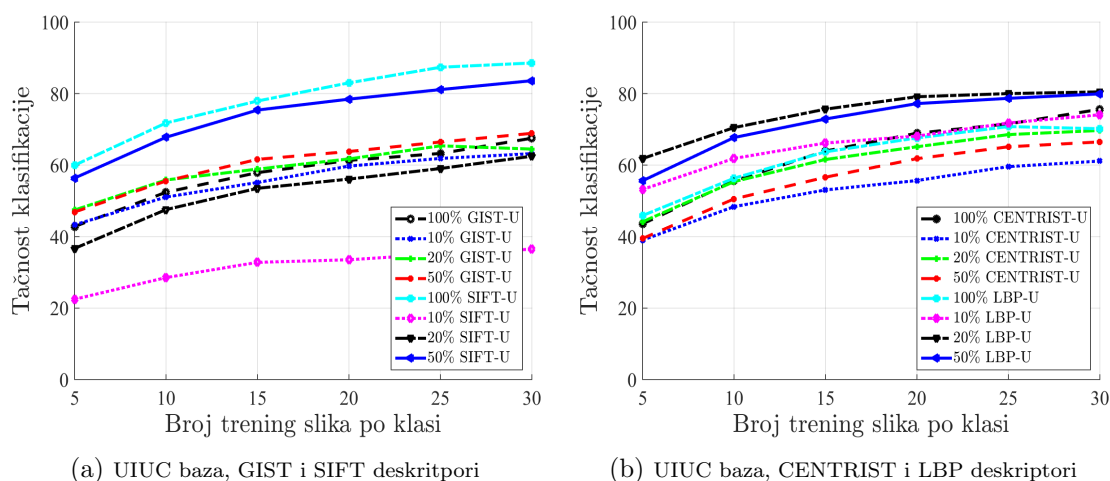
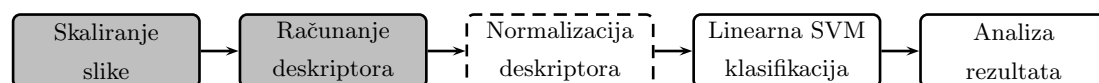
Iako su različite tehnike za normalizaciju deskriptora već razmatrane u literaturi, na ovom mjestu je napravljen detaljan komparativni pregled njihovog uticaja na tačnost klasifikacije. Naime, pojedine pristupi za normalizaciju deskriptora su samo navedeni kao optimalni za određene deskriptore [10, 76, 11] dok je konkretno poređenje njihovog uticaja na tačnost klasifikacije, provedene pod jednakim uslovima, izostala. Oznaka SVM-LIN označava linearni SVM klasifikator dok dodatak -U označava slučajeve kada je korišćeno skaliranje deskriptora prema (3.10). Prefiks L1 označava normalizaciju (3.3), L2 normalizaciju (3.4), sL1 normalizaciju (3.5), dok rL1 označava normalizaciju opisanu u [76].

Pregledom rezultata datih na Sl. 4.2 mogu se primjetiti značajna poboljšanja tačnosti klasifikacije u slučajevima kada se koristi postprocesiranje deskriptora skaliranjem datim sa (3.10), bez obzira na izbor deskriptora ili baze slika koja se klasifikuje. Korišćena transformacija predstavlja vrijednosti elemenata deskriptora tako da je njihova raspodjela uniformna, što omogućava da linearni klasifikator bolje modeluje granice između klasa. Takođe, može se primjetiti da metode za normalizaciju koje koriste korijenovanje (označene sa sL1 i rL1) u većini slučajeva poboljšavaju performanse u poređenju sa L1 i L2 normalizacijom, barem za nekoliko procenata. Tačnost klasifikacije za bazu UMD je veća u poređenju sa bazom UIUC za sve deskriptore, iako se u oba slučaja radi o sivim slikama sa teksturama. Ovakvi rezultati se mogu objasniti većim vizuelnim razlikama između tekstura iz različitih klasa u bazi UMD. Konkretno, SIFT deskriptor ima najbolje performanse, s obzirom da najbolje modeluje lokalne karakteristike tekstura.

Uticaj različitog skaliranja slika

Prije izdvajanja deskriptora, često se vrši promjena dimenzija ili skaliranje slika, kako bi se vršila obrada manje količine podataka i ubrzalo izdvajanje deskriptora. Važno je da korišćeni deskriptori zavise od skale slike u što manjoj mjeri. U ovim eksperimentima analizira se uticaj smanjenja dimenzija slika prije izdvajanja deskriptora, kako bi se ispitao uticaj skale tekstone na performanse klasifikacije. Oznaka 50% pored imena deskriptora znači da je slika skalirana na 50% svoje originalne veličine prije računanja deskriptora. Kako bi se korektno poredio uticaj skaliranja slika na tačnost klasifikacije, u ovom setu eksperimenta nije vršena normalizacija deskriptora, već isključivo skaliranje slika i rezultati klasifikacije slika iz UIUC baze prikazani su na Sl. 4.3.

Posmatrajući performanse dobijene korišćenjem deskriptora SIFT i CENTRIST, očigledno je da se smanjivanjem dimenzija slike gubi određeni dio informacija korisnih za klasifikaciju. Može se primjetiti da smanjenje dimenzija slika za 50 % prije izdvajanja deskriptora smanjuje tačnost klasifikacije do 5% na obje posmatrane baze, tj. održava relativno dobre performanse klasifikacije. GIST i LBP deskriptori imaju nešto bolje performanse ukoliko je izvršeno smanjenje dimenzija slika prije računanja deskriptora. Ovakvi rezultati mogu se objasniti izborom parametara sa kojima je izvršeno računanje deskriptora. GIST je računat na 8 skala po svakoj



Slika 4.3: Analiza uticaja skaliranja deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika

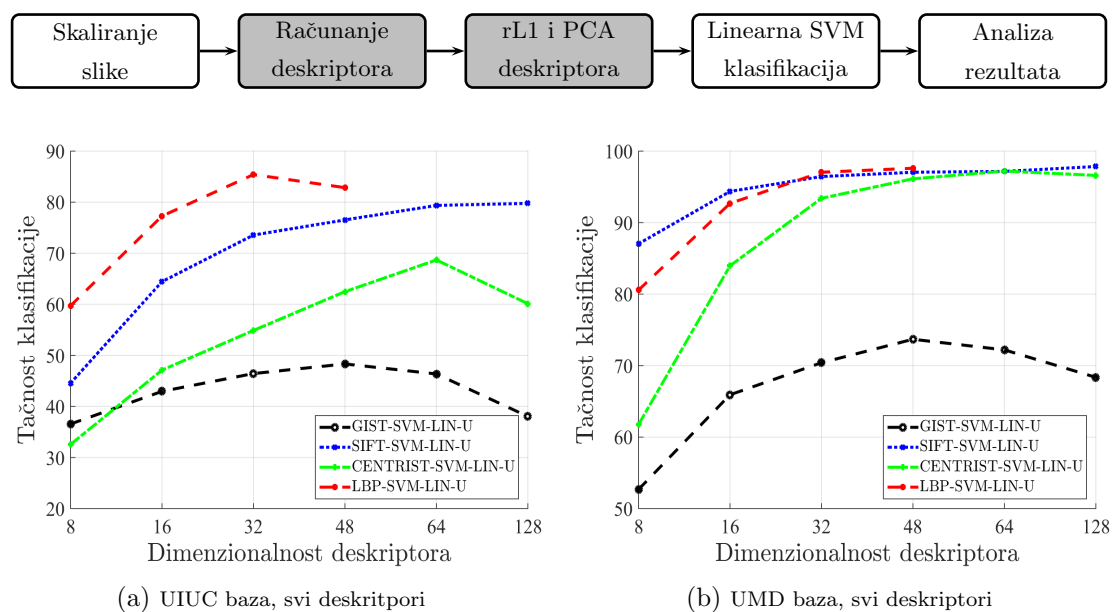
od 4 orijentacije, dok se LBP koristi u uniformnoj i rotaciono-invarijantnoj varijanti na 3 različite rezolucije. Iako ima mogućnost detekcije lokalnih obilježja na više skala, SIFT je računat na samo jednoj skali, dok CENTRIST ne uzima u obzir različite skale. Može se zaključiti da je računanje deskriptora na više skala je korisno, jer je moguće smanjiti dimenzije slika i ukupnu količinu podataka koju treba obraditi prilikom računanja deskriptora i pri tome sačuvati ili poboljšati rezultate klasifikacije.

Uticaj smanjenja dimenzionalnosti deskriptora

Dimenzionalnost deskriptora najviše zavisi od izbora parametara sa kojima se računaju deskriptori. U slučaju kada se analiziraju multispektralne slike, dimenzionalnost deskriptora zavisi od broja dostupnih kanala. U ovom setu eksperimenata analizira se uticaj smanjenja dimenzionalnosti deskriptora na tačnost automatske klasifikacije slika. Prethodno su navedene dimenzionalnosti korišćenih deskriptora, a u ovim eksperimentima je njihova dimenzionalnost značajno smanjena kako bi se ispitala mogućnosti korišćenja niskodimenzionalnih deskriptora prilikom klasifikacije sivih slika. U literaturi je slična analiza izvršena u [12], ali sa manjim brojem deskriptora i na drugoj bazi slika. Stoga je ova analiza korisna za dodatno poređenje performansi deskriptora.

U skladu sa rezultatima prethodnih eksperimenata, izvršeno je pretprocesiranje slika tako što je njihova veličina smanjena za pola prije računanja deskriptora, a nakon računanja deskriptora izvršena je rL1 normalizacija i skaliranje prema jednačini (3.10). Kao tehnika za redukciju dimenzionalnosti koristi se PCA, pošto pored mogućnosti odbacivanja neinformativnih komponenata, vrši i dekokorelaciju deskriptora. Iz svake klase je korišćeno po 20 slika za trening klasifikatora.

Na osnovu rezultata koji su prikazani na Sl. 4.4 može se primjetiti da u većini slučajeva tačnost klasifikacije na sličan način zavisi od dimenzionalnosti deskriptora,



Slika 4.4: Analiza uticaja redukcije dimenzionalnosti deskriptora na tačnost klasifikacije, za različite deskriptore i baze sivih slika

bez obzira o kojem deskriptoru se radi. Kada su dimenzionalnosti deskriptora veoma niske, redom 8 i 16, svi deskriptori imaju slabije performanse. Međutim, u slučaju klasifikacije UMD baze, kada se koriste deskriptori SIFT i LBP smanjene dimenzionalnosti na 8, postiže se tačnost klasifikacije preko 80%, što se može objasniti dovoljno velikom vizuelnom razlikom između slika iz različitih klasa u bazi.

4.3. Poređenje performansi deskriptora

Pored analize uticaja pojedinačnih koraka automatske klasifikacije na tačnost klasifikacije, korisno je izvršiti njihovo direktno poređenje. U tabelama 4.1 i 4.2 dati su rezultati automatske klasifikacije za slučajeve kada su slike i deskriptori procesirani na različite načine, za baze UIUC i UMD, redom. U svakom eksperimentu uključeni su različiti koraci prilikom klasifikacije, ali uvijek je korišćena ista podjela na trening i test skup slika. Iz svake klase je korišćeno 10 slika za trening, a ostale za testiranje. Rezultati u tabelama predstavljaju prosječne tačnosti klasifikacije prilikom za 5 različitih slučajnih podjela na trening i test skup slika, a brojevi u zagradama predstavljaju standardne devijacije. Posebno su date tačnosti klasifikacije ukoliko se koriste deskriptori punih dimenzionalnosti i ukoliko je dimenzionalnost deskriptora smanjena na 16 primjenom PCA transformacije. Koristi se redukovana dimenzionalnost 16 koja je znatno manja od originalne dimenzionalnosti, iako prema rezultatima sa Sl. 4.4 vidimo da se za nešto veće dimenzionalnosti dobijaju nešto bolje tačnosti klasifikacije. U svakom od eksperimenata korišćen je linearni SVM klasifikator.

Posmatrajući rezultate klasifikacije za UIUC bazu, SIFT-1000 i LBP deskriptori punih dimenzionalnosti pokazuju bolje rezultate u poređenju sa GIST i CENTRIST deskriptorima. SIFT deskriptor postiže najbolje rezultate u kombinaciji sa normali-

4.3. Poređenje performansi deskriptora

Tabela 4.1: Poređenje tačnosti klasifikacije u slučajevima kada se deskriptori računaju i procesiraju na različite načine, za UIUC bazu

	GIST	SIFT	CENTRIST	LBP
dimenzionalnost	512	1000	254	56
100 % + LIN	53.76 ± 2.22	61.97 ± 1.43	56.00 ± 1.47	79.89 ± 2.41
50 % + LIN	56.64 ± 2.68	58.48 ± 2.00	52.93 ± 2.13	80.88 ± 1.79
rL1 + LIN-U	48.26 ± 2.57	76.29 ± 1.99	56.53 ± 1.59	55.97 ± 1.52
50 % + LIN-U	55.54 ± 2.18	67.78 ± 2.63	50.53 ± 3.54	67.71 ± 3.05
50 % + rL1 + LIN-U	52.54 ± 2.18	70.82 ± 1.63	52.08 ± 3.32	66.82 ± 2.26
dimenzionalnost	16	16	16	16
100 % + LIN	42.34 ± 2.56	39.57 ± 2.88	50.42 ± 1.67	77.76 ± 2.76
50 % + LIN	47.41 ± 2.38	34.40 ± 1.55	47.81 ± 2.47	78.61 ± 1.54
rL1 + LIN-U	33.57 ± 2.23	59.55 ± 3.47	47.20 ± 3.34	72.29 ± 1.07
50 % + LIN-U	30.11 ± 1.41	32.16 ± 1.63	32.61 ± 3.32	63.76 ± 1.42
50 % + rL1 + LIN-U	37.87 ± 1.41	53.86 ± 2.91	39.28 ± 2.84	66.83 ± 2.19

Tabela 4.2: Poređenje tačnosti klasifikacije u slučajevima kada se deskriptori računaju i procesiraju na različite načine, za UMD bazu

	GIST	SIFT	CENTRIST	LBP
dimenzionalnost	512	1000	254	56
100% + LIN	80.96 ± 0.85	89.62 ± 1.83	87.76 ± 1.42	92.13 ± 1.92
50% + LIN	81.63 ± 0.27	88.74 ± 1.15	84.53 ± 1.95	91.81 ± 2.25
100% rL1 + LIN-U	78.69 ± 1.33	95.63 ± 1.26	90.45 ± 1.71	84.69 ± 1.41
50% + LIN-U	80.48 ± 1.11	91.60 ± 0.57	87.15 ± 1.44	79.49 ± 2.29
50% + rL1 + LIN-U	77.63 ± 1.22	94.11 ± 0.39	87.71 ± 1.05	78.83 ± 1.86
dimenzionalnost	16	16	16	16
100% + LIN	63.78 ± 1.67	80.29 ± 3.21	80.29 ± 3.21	88.29 ± 1.65
50% + LIN	60.93 ± 2.25	74.85 ± 3.25	66.85 ± 2.00	84.96 ± 1.20
100% rL1 + LIN-U	62.51 ± 1.74	91.47 ± 1.48	81.65 ± 1.11	89.78 ± 1.82
50% + LIN-U	60.93 ± 2.25	74.85 ± 3.25	66.77 ± 2.00	84.96 ± 1.20
50% + rL1 + LIN-U	58.98 ± 1.59	89.31 ± 1.48	74.93 ± 0.47	87.07 ± 2.88

zacijom uniformnim skaliranjem, kao što je opisano u [76]. Interesantno je primjetiti da LBP deskriptor postiže veću tačnost klasifikacije u slučaju kada se koristi samo promjena veličine slike prije računanja deskriptora, a ne dodatno procesiranje deskriptora. Pored toga, LBP ima znatno manju dimenzionalnost. Može se zaključiti da SIFT i LBP mogu “opisati” diskriminativna svojstva tekstura iz UIUC baze i postići dobre rezultate klasifikacije sa malim brojem trening slika. GIST i CENTRIST deskriptori imaju nešto lošije performanse na ovoj bazi, koje nije moguće popraviti dodatnim procesiranjem deskriptora.

Ukoliko se uporede rezultati klasifikacije u slučaju kada se koriste deskriptori punih dimenzionalnosti sa slučajevima kada je njihova dimenzionalnost smanjena na 16. Osim što je izvršena redukcija dimenzionalnosti, svi ostali uslovi u eksperi-

mentima su ostali isti. Pored očekivano lošijih rezultata, može se primjetiti da LBP deskriptor zadržava visoku tačnost klasifikacije. U slučaju kada se koristi skaliranje slike, bez postprocesiranja deskriptora, tačnost klasifikacije je smanjena za 2%, dok je u slučaju kada se koriste normalizacija i uniformno skaliranje deskriptora, tačnost klasifikacije povećava za preko 15%. Uopšte, može se primjetiti da se tačnost klasifikacije najviše smanjuje za deskriptore većih originalnih dimenzionalnosti. GIST i CENTRIST deskriptori ne ostvaruju dobre tačnosti klasifikacije za UIUC bazu, prvenstveno zbog toga što su prilagođeni za aplikacije prepoznavanja scena.

Rezultati klasifikacije UMD baze, koji su dobijeni u eksperimentima izvršenim pod istim uslovima kao i za UIUC bazu, dati su u tabeli 4.2. Može se primjetiti da su tačnosti klasifikacije znatno bolje za UMD bazu u poređenju sa UIUC bazom, iz razloga veće vizuelne raznolikosti između slika iz različitih klasa. Za slučajeve pune dimenzionalnosti, može se primjetiti da je kombinacija normalizacije i uniformnog skaliranja deskriptora poboljšava performanse SIFT i CENTRIST deskriptora. Što se tiče GIST i LBP deskriptora, najbolje performanse dobijaju se bez navedenih koraka postprocesiranja. Kada se analiziraju rezultati klasifikacije za slučaj redukovane dimenzionalnosti deskriptora, kombinacija normalizacije i uniformnog skaliranja deskriptora daje najveću tačnost klasifikacije za svaki od deskriptora.

4.4. Uticaj parametara klasifikatora

Iako je tema ove disertacije način kreiranja deskriptora na ovom mjestu će biti razmotren uticaj parametara klasifikatora. U narednim eksperimentima slike su skalirane na pola svoje originalne veličine, nakon čega su računati deskriptori, koji su normalizovani i uniformno skalirani. Klasifikacija je izvršena SVM klasifikatorom, sa različitim kernelima. Upotreba kernela je korisna ukoliko deskriptori iz različitih klasa u prostoru obilježja nisu linearno separabilni, u kojem slučaju transformacija u prostor obilježja više dimenzionalnosti može povećati separabilnost. Bez objašnjavanja detalja, ispituje se uticaj uniformnog skaliranja deskriptora u slučajevima kada se koriste različiti kerneli. Konkretno, porede se rezultati klasifikacije kada se koristi SVM klasifikator sa linearnim (LIN), polinomskim (POL),

Tabela 4.3: Poređenje tačnosti klasifikacije u slučajevima kada koriste različiti kerneli SVM klasifikatora

	GIST	SIFT	CENTRIST	LBP
50% + rL1 + LIN	45.00 ± 2.46	80.44 ± 2.56	45.80 ± 2.03	49.16 ± 2.18
50% + rL1 + POL	38.08 ± 1.38	53.76 ± 1.66	32.04 ± 1.73	40.24 ± 2.36
50% + rL1 + RBF	36.44 ± 1.58	45.52 ± 0.58	30.08 ± 1.82	40.20 ± 1.95
50% + rL1 + SIG	36.44 ± 1.58	45.52 ± 0.58	30.08 ± 1.82	40.24 ± 1.97
50% + rL1 + LIN-U	63.24 ± 2.32	82.76 ± 1.91	70.72 ± 2.98	91.52 ± 1.37
50% + rL1 + POL-U	37.96 ± 2.83	32.56 ± 2.08	27.00 ± 1.61	43.84 ± 2.41
50% + rL1 + RBF-U	44.76 ± 1.63	65.16 ± 2.11	40.04 ± 2.03	65.84 ± 1.28
50% + rL1 + SIG-U	38.36 ± 1.89	57.00 ± 2.84	35.76 ± 2.29	56.44 ± 1.04

radijalnim (RBF) i sigmoidnim (SIG) kernelom, u slučajevima kada se vrši i ne vrši dodatno uniformno skaliranje deskriptora. Kako bi se bolje naglasile razlike u performansama u slučajevima kada se koriste različiti kerneli, regularizacioni parametar je u ovom setu eksperimentu postavljen na vrijednost $C = 10$. Klasifikacija je izvršena na UIUC bazi, sa 20 trening slika iz svake klase, a rezultati eksperimenata su dati u tabeli 4.3. Pregledom rezultata klasifikacije, može se primjetiti da uniformno skaliranje doprinosi značajnom povećanju tačnosti klasifikacije, osim u slučaju kada se koristi polinomski kernel. Međutim, potrebno je naglasiti da nije izvršeno fino podešavanje parametara koji mogu dodatno poboljšati performanse klasifikatora. Stoga, i pored nekoliko izuzetaka, može se uočiti korist od dodatnog procesiranja deskriptora. Sa druge strane, ukoliko se uporede tačnosti klasifikacije za različite deskriptore, jasno je da najveći uticaj na tačnost klasifikacije ima dobar odabir deskriptora za konkretnu aplikaciju.

4.5. Zaključci korisni za naredne eksperimente

U ovom poglavlju diskutovani su rezultati eksperimenata kroz koje je analiziran uticaj pojedinih koraka prilikom automatske klasifikacije sivih slika na osnovu teksture. Performanse klasifikacije su ispitane za različite slučajeve kada se koriste četiri odabrana deskriptora, kao bi se utvrdilo pod kojim okolnostima oni daju najbolje rezultate klasifikacije sivih slika. Eksperimenti su izvršeni na dvije baze slika, pod istim uslovima (ista podjela na trening i test slike i jednaki parametri deskriptora i klasifikatora), sa ciljem da se diskutuju koraci pretprocesiranja slika i procesiranja deskriptora, pogodni za korišćenje prilikom klasifikacije multispektralnih slika. Pošto u literaturi ne postoji slična analiza, rezultati prethodni eksperimenata imaju veliki značaja za analizu slučajeva kada se koriste slike u boji i multispektralne slike. Na osnovu prethodno datih rezultata može se izvesti nekoliko zaključaka:

- (1) Ukoliko je moguće, korisno je izvršiti multirezolucionu analizu tj. izdvajanje deskriptora na više skala, što obezbjeđuje tačniji uvid u obilježja slike i veću tačnost klasifikacije. Istovremeno, moguće je izvršiti smanjenje dimenzija slike prije izdvajanja deskriptora u kojem slučaju je potrebno obraditi manju količinu podataka što zahtjeva manje memorije i vremena. Ukratko, računanje deskriptora na više skala nakon promjene veličine slika može popraviti rezultate klasifikacije.
- (2) Kombinacija normalizacije deskriptora i uniformnog skaliranja njegovih vrijednosti može u najvećem broju slučajeva povećava tačnost klasifikacije. Korišćenjem normalizacije deskriptori se predstavljaju u prostoru obilježja na način koji omogućava efikasnije obučavanje klasifikatora.
- (3) Redukcijom dimenzionalnosti mogu se sačuvati performanse klasifikacije, naročito u slučaju ako su slike iz različitih klasa vizuelno različite, a klasifikacija sa deskriptorima nižih dimenzionalnosti zahtjeva manje računskih operacija i memorijskih resursa prilikom obučavanja klasifikatora i klasifikacije.
- (4) Za različite deskriptore, najveće tačnosti klasifikacije dobijaju se primjenom različitih tehnika pretprocesiranja slika i procesiranja deskriptora. Uopšte,

različiti deskriptori mogu biti definisani i projektovani za različite vrste praktičnih aplikacija, pa ne postoji opšti pristup koji će garantovati bolje performanse kada se koriste određeni koraci procesiranja prilikom računanja deskriptora. Npr. ukoliko je LBP deskriptor pokazao bolje performanse u odnosu na GIST deskriptor na bazama tekstura, konzistentni rezultati se ne moraju dobiti na drugim bazama slika.

- (5) Poželjno je nakon redukcije dimenzionalnosti deskriptora izvršiti njihovu normalizaciju i uniformno skaliranje, jer se tako može povećati tačnost klasifikacije. Ovaj pozitivan uticaj je naročito izražen kada postoji veća vizuelna razlika između slika iz različitih klasa.
- (6) Na tačnost klasifikacije najveći uticaj ima dobar odabir deskriptora. Konkretno, na osnovu vizuelnih i semantičkih karakteristika i sadržaja slika, kao i prema potrebama praktičnih aplikacija, dizajnirani su različiti deskriptori koji mogu na odgovarajući način predstaviti podatke sa slika. Ukratko, bolji podaci znače veću tačnost klasifikacije.

5 Niskodimenzionalni multispektralni deskriptori tekstone

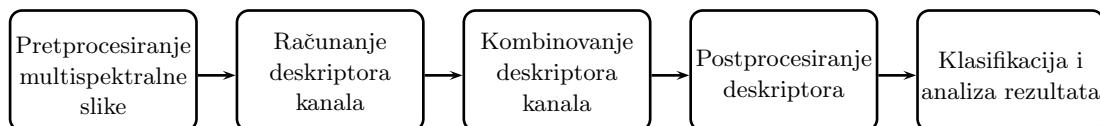
Multispektralne slike sadrže tri ili više spektralnih kanala, koji se mogu tretirati kao sive slike. Izdvajanje deskriptora iz sivih slika i redukcija njihove dimenzionalnosti je analizirana u prethodnom poglavlju, pa će u ovom poglavlju biti izvršena analiza različitih načina za kombinaciju informacija iz različitih spektralnih opsega. Cilj je dobijanje deskriptora multispektralnih slika što niže dimenzionalnosti, koji uključuje informacije iz svih spektralnih opsega i obezbjeđuje što veću tačnost klasifikacije. Uzimajući u obzir rezultate klasifikacije baza sivih slika predstavljene u prethodnom poglavlju, svi naredni eksperimenti će uključivati promjenu veličine slike na pola njene originalne veličine, zatim normalizaciju deskriptora rL1 i uniformno skaliranje. Za klasifikaciju će biti korišćen SVM klasifikator sa linearnim kernelom. Takođe, automatska klasifikacija će biti izvršena isključivo na bazama slika u boji i multispektralnih slika iz UCM i NirScene baza.

5.1. Postavka eksperimenata

Algoritam za klasifikaciju slika u boji i multispektralnih slika se ne razlikuje suštinski od algoritma za klasifikaciju sivih slika, čiji je šematski prikaz dat na Sl. 4.1. Međutim, dostupnost više kanala omogućava više načina pretprocesiranja slika i kombinovanja deskriptora pojedinačnih kanala. Šematski prikaz korišćenog algoritma za automatsku klasifikaciju multispektralnih slika prikazan je na Sl. 5.1. U poređenju sa klasifikacijom sivih slika, postoji dodatni korak u kojem se kombinuju deskriptori izračunati iz pojedinačnih kanala.

U narednim eksperimentima biće ispitani sljedeći koraci:

- (a) **Pretprocesiranje slika:** Kada se vrši obrada slika u boji ili multispektralnih slika, moguće je izvršiti transformacije slika u druge kolor prostore. Pošto je



Slika 5.1: Šematski prikaz algoritma za automatsku klasifikaciju multispektralnih slika

u prethodnom poglavlju pokazana korist od smanjenja dimenzija slika, u narednim eksperimentima će sve slike biti smanjenje na pola svojih originalnih dimenzija prije računanja deskriptora. Takođe, biće ispitana tačnost klasifikacije u slučajevima kada se koriste različiti kolor prostori, slično kao što je razmatrano u [74].

- (b) **Izdvajanje deskriptora:** Kao kod klasifikacije sivih slika, koriste se deskriptori GIST, SIFT, CENTRIST i LBP, međutim oni se računaju za svaki od dostupnih kanala posebno.
- (c) **Kombinovanje deskriptora različitih kanala:** Pošto je cilj dobijanje deskriptora što moguće niže dimenzionalnosti, kombinovanje deskriptora različitih kanala je izvršeno na različite načine u kombinaciji sa redukcijom dimenzionalnosti, kao što je predloženo u [13]. Posebno se obraća pažnja na kreiranje nekoliko varijanti multispektralnog deskriptora baziranog na SIFT obilježjima, s obzirom da se računaju u kombinaciji sa paradigmom vizuelnih riječi.
- (d) **Postprocesiranje deskriptora:** U svim eksperimentima vezanim za baze slika u boji i multispektralne slike koristi se rL1 normalizacija konačnog deskriptora (nakon kombinovanja) i njegovo uniformno skaliranje.
- (e) **Klasifikacija:** U svim eksperimentima koristi se linearni SVM klasifikator sa regularizacionim parametrom $C = 1$, ukoliko to nije drugačije naglašeno.

5.2. Rezultati automatske klasifikacije slika u boji i multispektralnih slika

Slike iz baze UCM su slike u boji, dakle sadrže tri kanala koji predstavljaju crveni (R), zeleni (G) i plavi (B) kanal, dok slike iz baze NirScene pored RGB kanala sadrže i podatke iz bliskog infracrvenog opsega (NIR). Dakle, slike iz baza UCM i NirScene imaju tri i četiri kanala, redom. Najjednostavniji način kombinovanja deskriptora pojedinačnih kanala predstavlja njihovo spajanje u jedan deskriptor multispektralne slike. U tom slučaju dimenzionalnost deskriptora se povećava srazmjerno broju dostupnih kanala. Na primjer, za slike u boji dobija se dimenzionalnost GIST deskriptora 1536, SIFT-1000 deskriptora 4000, CENTRIST deskriptora 762 i LBP deskriptora 162. Svaki novi dostupni kanal nosi nove informacije, ali dodatno povećava dimenzionalnost deskriptora i zahtjeve za računarskim resursima i memorijom.

Slike iz baza UCM i NirScene se razlikuju od slika iz UIUC i UMD baza na osnovu vizuelnog i semantičkog sadržaja. Dok su slike iz baza UIUC i UMD prikazuju isključivo tekture, ravnomjerno raspoređene po cijeloj slici, na slikama koje će biti razmatrane u ovom poglavlju postoji više regiona sa teksturama različitih karakteristika. Takođe, često se veći dio korisnih informacija nalazi u obliku nekog

regiona na slici, npr. kada taj region prikazuje objekat određenog oblika. Iako klasifikacija aero snimaka i scena datih u slikama iz baza UCM i NirScene predstavlja teži zadatak predstavlja teži zadatak u poređenju sa klasifikacijom tekstura iz baza UIUC i UMD, mogu se koristiti zaključci izvedeni na kraju prethodnog poglavlja.

U svim narednim eksperimentima izvršeno je 5 podjela na trening i test slike za svaku klasu iz posmatrane baze, tako da rezultati eksperimenta predstavljaju srednje vrijednosti tačnosti klasifikacije za svaku od 5 podjela. Ukoliko je naglašen broj trening slika iz svake klase, u svakoj podjeli je na slučajan način izabran naznačeni broj slika iz klase za trening, dok su ostale slike iz klase korišćene za testiranje.

Uticaj različitih kolor prostora

Transformacija slika u drugi kolor prostor prije njihove klasifikacije predstavlja jedan od mogućih koraka pretprocesiranja slika. Predstavljanje slika u različitim kolor prostorima može omogućiti drugačije opisivanje vizuelnih i semantičkih obilježja prisutnih na slikama i eventualno povećati tačnost klasifikacije i omogućiti veću redukciju dimenzionalnosti deskriptora. U ovim eksperimentima upoređeni su rezultati klasifikacije u slučajevima kada se deskriptori slika u boji računaju iz RGB, LAB, HSV, oRGB i PCA-RGB kolor prostora. Deskriptori slika u boji i multispektralnih slika za ove eksperimente, dobijeni su jednostavnim spajanjem deskriptora pojedinačnih kanala, bez redukcije dimenzionalnosti. Rezultati eksperimenata su dati na Sl. 5.2. U slučaju kada su dostupni i drugi kanali pored RGB, kao što je NIR posebno se ispituje tačnost klasifikacije za RGB+NIR, oRGB+NIR i PCA-RGB+NIR kolor prostore.

PCA dekokorelacija je izvršena prema transformaciji koja zavisi od podataka iz baze slika koja se klasifikuje. Od ukupnog broja slika iz baze slučajno se bira određeni dio slika iz kojih se uzimaju podaci o boji piksela. Količina podataka se bira da bude dovoljna da uzme se u obzir raznolikost boja iz kompletne baze slika, ali da ne bude prevelika kako bi se izbjegli veliki računski troškovi. Kroz nekoliko eksperimenata pokazano je da se koeficijenti transformacije ne mijenjaju značajno nakon povećanja broja korišćenih slika preko određenog procenta. Konkretno, dovoljno je uzeti do 20% podataka iz baze slika za računanje koeficijenata transformacije. UCM baza sadrži slike u boji, pa se transformacija za dekokorelaciju boja piksela dobija kao:

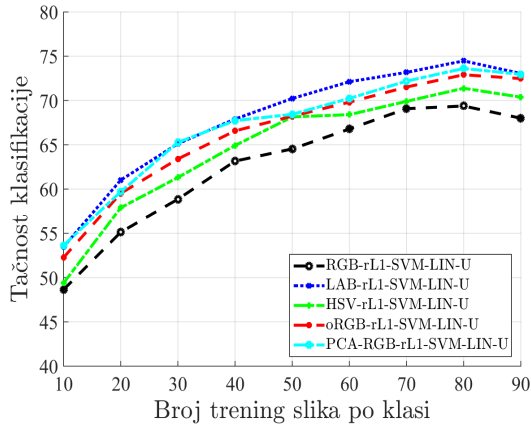
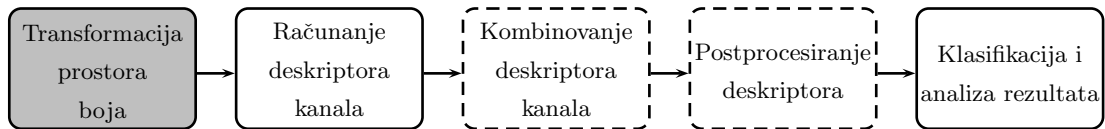
$$\mathbf{P}_{\text{UCM}}^{rgb} = \begin{bmatrix} 0.569 & 0.707 & 0.420 \\ 0.594 & 0.000 & -0.804 \\ 0.569 & -0.707 & 0.420 \end{bmatrix}. \quad (5.1)$$

Slično, za NirScene bazu slika, mogu se dobiti dvije transformacione matrice, u slučaju kada se koriste samo kanali iz vidljivog dijela spektra i u slučaju kada se koristi dodatni infracrveni kanal. Dobija se:

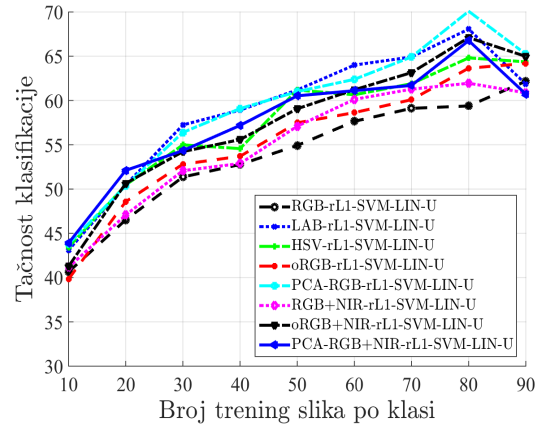
$$\mathbf{P}_{\text{NirScene}}^{rgb} = \begin{bmatrix} 0.576 & 0.707 & 0.411 \\ 0.580 & 0.000 & -0.814 \\ 0.576 & -0.707 & 0.409 \end{bmatrix} \quad (5.2)$$

za dekokorelaciju RGB kolor prostora i

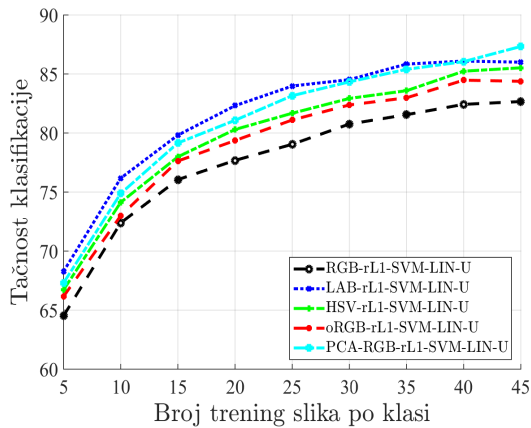
5.2. Rezultati automatske klasifikacije slika u boji i multispektralnih slika



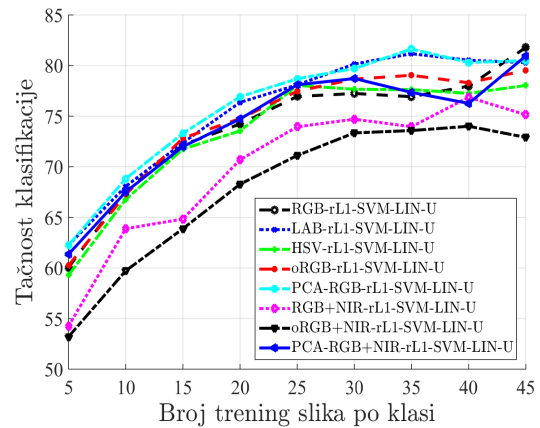
(a) UCM baza, GIST deskriptor



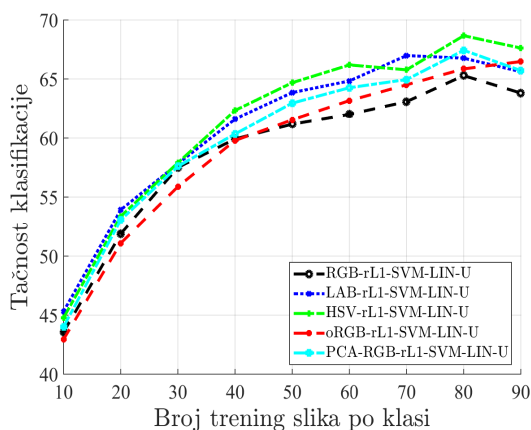
(b) NirScene baza, GIST deskriptor



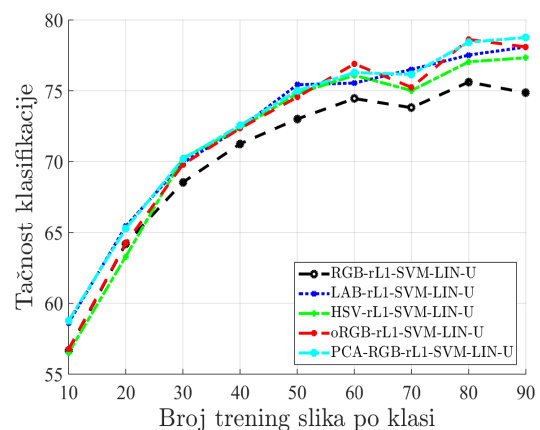
(c) UCM baza, SIFT deskriptor



(d) NirScene baza, SIFT deskriptor



(e) UCM baza, CENTRIST deskriptor



(f) UCM baza, LBP deskriptor

Slika 5.2: Analiza uticaja različitih prostora boja na tačnost klasifikacije, za različite deskriptore i baze slika u boji i multispektralnih slika

$$\mathbf{P}_{\text{NirScene}}^{rgb+nir} = \begin{bmatrix} 0.498 & -0.631 & -0.505 & 0.316 \\ 0.502 & -0.319 & 0.499 & -0.631 \\ 0.502 & 0.316 & 0.497 & 0.633 \\ 0.498 & 0.633 & -0.499 & -0.319 \end{bmatrix} \quad (5.3)$$

za dekorelaciju RGB+NIR kolor prostora. Posebno je interesantno uočiti sličnost između matrica transformacije datih jednačinama (5.1) i (5.2), pošto su izvedene iz različitih baza slika. Može se zaključiti da su transformacije potrebne za dekorelaciju kolor prostora prirodnih slika slične. Rezultati klasifikacije prikazani u zavisnosti od korišćenog prostora boja, za različite deskriptore i baze slika su na Sl. 5.2. Prije analize dobijenih rezultata, potrebno je naglasiti da se drugačije tačnosti klasifikacije mogu postići optimalnim izborom parametara postprocesiranja deskriptora ili podešavanjem parametara klasifikatora.

Poređenjem tačnosti klasifikacije, kada se deskriptori računaju za slike iz RGB i PCA-RGB, odnosno RGB+NIR i PCA-RGB+NIR kolor prostora, može se primjetiti da se dekorelacijom kanala konzistentno popravljaju tačnost klasifikacije do 5%, bez obzira na odabrani deskriptor ili bazu slika. Takođe, mogu se primjetiti bolji rezultati klasifikacije ukoliko se umjesto RGB koristi LAB kolor prostor. Uopšte, poređenjem rezultata klasifikacije za različite kolor prostore, vidi se da postoje odstupanja tačnosti klasifikacije od nekoliko procenata i preporučuje se korišćenje oponentnih prostora boja. Daljim razmatranjem se za svaki deskriptor može naći kolor prostor, dopunjen dodatnim kanalima i/ili dekorelisani, za koji se postiže najbolji rezultati klasifikacije.

Za ovo istraživanje važan je sljedeći zaključak: pretprocesiranje slika koje uključuje konverziju u druge kolor prostore ili dekorelaciju može se popraviti tačnost klasifikacije i eventualno kompenzovati gubitak koji se dobija redukcijom dimenzionalnosti. Taj gubitak je uočen analizom performansi deskriptora redukovanih dimenzionalnosti na bazama sivih slika. Kasnije će biti razmotrena mogućnost za popravljavanje tačnosti klasifikacije uključivanjem podataka iz dodatnih spektralnih opsega van vizuelnog dijela spektra.

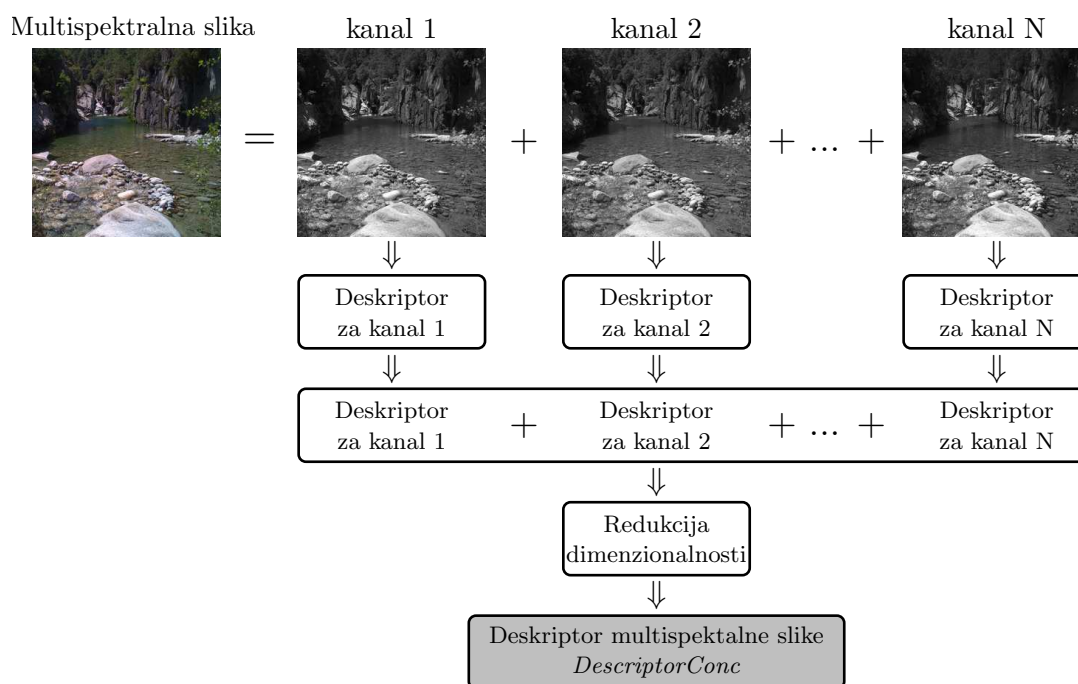
5.3. Predložena rješenja za kombinovanje deskriptora različitih kanala

Kao što je već ranije naglašeno, jedan od ciljeva ovog istraživanja jeste prijedlog algoritma za računanje deskriptora teksture za slike u boji i multispektralne slike, koji kombinuje podatke iz svih dostupnih spektralnih opsega i daje konačni deskriptor što manje moguće dimenzionalnosti. Iako je kratak pregled literature već napravljen, potrebno je skrenuti posebnu pažnju na načine izdvajanja deskriptora iz multispektralnih slika. U literaturi su predloženi i opisani deskriptori multispektralnih slika koji se oslanjaju isključivo na spajanje deskriptora pojedinačnih kanala, iako su načini računanja lokalnih obilježja nešto drugačiji. U [53] je predloženo proširenje SIFT detektora kako bi se računala obilježja multispektralnih slika, ali posljednji korak jeste samo spajanje histograma različitih kanala. mCENTRIST [11] se oslanja na kombinovanje združenih podataka iz dva različita kanala, kako

bi se odredile lokalne karakteristike, što predstavlja efikasan način za izdvajanje lokalnih obilježja. Nedostatak jeste visoka dimenzionalnost deskriptora, naročito ako postoji veći broj kanala. Sličan pristup za proširenje LBP deskriptora na multispektralne slike je iskorišćen u [77], dok je u [78] izvršeno odgovarajuće proširenje kako bi se postigla invarijantnost na promjene osvjetljenja. Razmatrajući načine za kombinovanje podataka iz različitih spektralnih opsega, jedino se u [79] eksplicitno poredi rezultati pretraživanja satelitskih snimaka visoke rezolucije prema deskriptorima panhromatskih verzija slika i slika koje uključuju normalizovani indeks vegetacije dobijen iz dostupnih kanala. U daljem tekstu će biti razmotreni predloženi načini kombinovanja deskriptora pojedinačnih kanala.

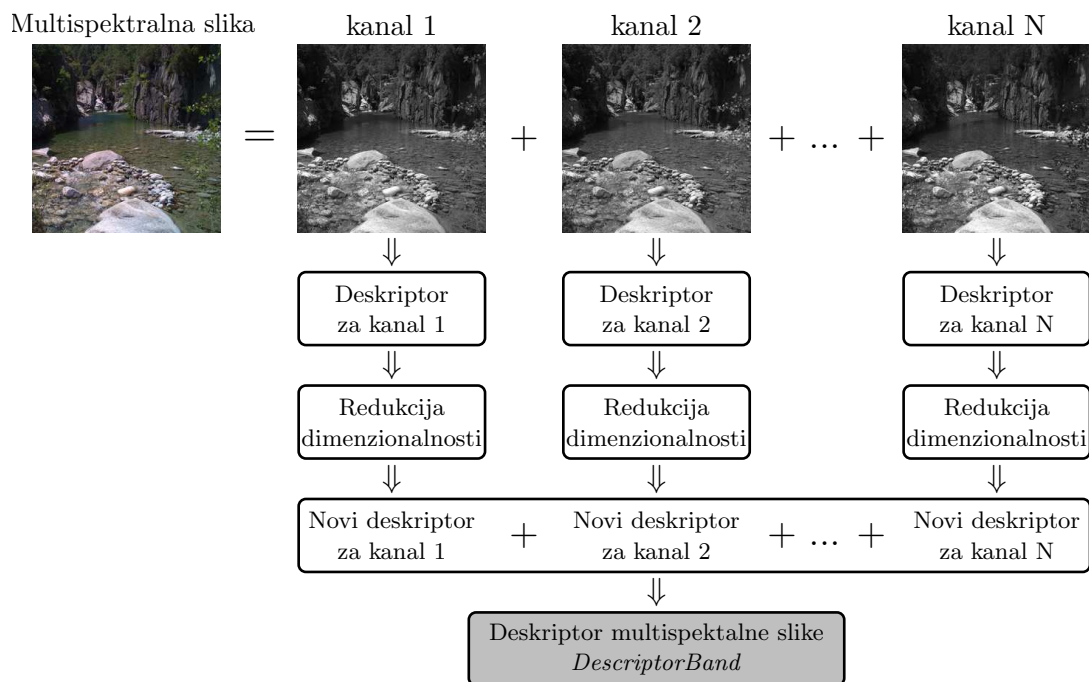
Kombinovanje deskriptora pojedinačnih kanala

U [13] su predložena i razmotrena dva rješenja za kombinovanje deskriptora pojedinačnih spektralnih opsega i tehnika za redukciju dimenzionalnosti. Posmatrajući jednu sliku u boji ili multispektralnu sliku, postoji tri ili više sivih slika koje predstavljaju podatke iz različitih spektralnih opsega ili kanala. Pošto su razmatrani deskriptori teksture (GIST, SIFT, CENTRIST i LBP) definisani isključivo za sive slike, iz svakog od dostupnih kanala se računa po jedan deskriptor. Predložena su dva algoritma za kombinaciju deskriptora pojedinačnih kanala.



Slika 5.3: Šematski prikaz algoritma za računanje deskriptora teksture za multispektralne slike (*DescriptorConc*)

DescriptorConc Prva mogućnost kombinovanja deskriptora pojedinačnih kanala podrazumijeva njihovo spajanje i redukovanje dimenzionalnosti tako dobijenog deskriptora. Odgovarajući šematski prikaz formiranja deskriptora na ovaj način, prikazan je na Sl. 5.3. Nakon što se izdvoje deskriptori iz sivih slika svakog od



Slika 5.4: Šematski prikaz algoritma za računanje deskriptora teksture za multispektralne slike (*DescriptorBand*)

dostupnih kanala, oni se spajaju, a zatim se na rezultujući deskriptor primjenjuje jedna od tehnika za redukciju dimenzionalnosti. Deskriptori korišćeni u prethodnom eksperimentu, prilikom testiranja uticaja kolor prostora na tačnost klasifikacije, predstavljaju varijantu *DescriptorConc* algoritma bez redukcije dimenzionalnosti.

Jednostavnim spajanjem deskriptora nekoliko sivih slika koje predstavljaju pojedinačne kanale jedne multispektralne slike koristimo podatke koji mogu imati visok stepen suvišnih informacija. To se dešava jer sive slike, koje predstavljaju pojedinačne kanale prirodnih slika, često imaju visok stepen međusobne vizuelne sličnosti. Stoga, konačan deskriptor multispektralne slike dobijen spajanjem nekoliko deskriptora sivih slika, pored toga što je visoke dimenzionalnosti, ne garantuje veću tačnost klasifikacije. Pogodno je izabrati onu tehniku redukcije dimenzionalnosti koja će, pored smanjenja broja komponenata deskriptora, izvršiti dekokorelaciju njihovih vrijednosti.

DescriptorBand Preglednom rezultata eksperimenata koji su razmatrali redukciju dimenzionalnosti prilikom klasifikacije sivih slika, može se primjetiti da pojedini deskriptori sivih slika značajno smanjene dimenzionalnosti mogu postići uporedive tačnosti klasifikacije kao i kada nema smanjenja dimenzionalnosti. Stoga, drugi pristup za kombinovanje deskriptora pojedinačnih kanala vrši redukciju dimenzionalnosti deskriptora svakog pojedinačnog kanala, prije spajanja u konačni deskriptor. Šematski prikaz formiranja deskriptora multispektralnih slika prema *DescriptorBand* algoritmu, prikazan je na Sl. 5.4. Pošto se u ovom pristupu prvo obrađuju podaci iz jednog kanala, redukcija dimenzionalnosti primjenom PCA transformacije će izvršiti dekokorelaciju podataka iz jednog kanala.

Pristupi *DescriptorConc* i *DescriptorBand* se razlikuju prema redoslijedu kombi-

novanja deskriptora pojedinačnih kanala i redukcije dimenzionalnosti. Izbor između ovih algoritama najviše zavisi od podataka, tj. tekstura na slikama koje se klasifikuju, od tehnike za redukciju dimenzionalnosti koja se koristi, kao i od konačne dimenzionalnosti deskriptora multispektralnih slika. Ukoliko postoji velika korelisanost između numeričkih vrijednosti deskriptora pojedinačnih kanala na slikama, kao što je slučaj sa prirodnim slikama, možemo očekivati da će algoritam DescriptorConc postići bolje rezultate klasifikacije. Sa druge strane, DescriptorBand će bolje iskoristiti podatke iz kanala koji imaju manji stepen međusobne korelisanosti. Međutim, treba naglasiti da prethodne pretpostavke uveliko zavise od dimenzionalnosti deskriptora na koju je izvršena redukcija.

Kombinovanje vizuelnih riječi kod multispektralnih slika

Paradigma vizuelnih riječi je koncept preuzet iz klasifikacije teksta [80] i na pogodan način proširen na klasifikaciju slika. Koristi se da predstavi koliko je kojih vizuelnih riječi prisutno na nekoj slici. Ukoliko poredimo sadržaj dvije slike, BoVW zapravo prebrojava koliko se sličnih lokalnih karakteristika nalazi na svakoj od slika, što omogućava njihovo poređenje prema vizuelnom sadržaju.

Posmatrajmo način kreiranja deskriptora na osnovu SIFT lokalnih obilježja, gdje se BoVW koncept koristi prilikom računanja broja lokalnih vizuelnih obilježja na sivim slikama. Za početak, potrebno je zadati broj vizuelnih riječi koji će se računati na svakoj slici, a zatim odrediti koje su to vizuelne riječi. U tu svrhu iz skupa trening slika se izdvaja određeni broj lokalnih obilježja $\mathbf{x}_i = [x_{i,1} \ x_{i,2} \ \dots \ x_{i,128}]$, $i = 1 \dots N$, koja se klasterizuju u rječnik vizuelnih riječi $\mathbf{V} = [v_1 \ v_2 \ \dots \ v_M]$, gdje je M broj prethodno zadanih vizuelnih riječi. U posljednjem koraku, SIFT deskriptor se formira kao histogram koji pokazuje koliko je lokalnih obilježja na slici najbliže nekoj od vizuelnih riječi iz rječnika vizuelnih riječi. U eksperimentima sa sivim slikama u *Poglavlju 4*, korišćeno je 1000 vizuelnih riječi.

Pošto se rječnik vizuelnih riječi formira prema lokalnim obilježjima sivih slika, postavlja se pitanje na koji način je pogodno proširiti ovaj koncept na slike u boji i multispektralne slike. U [29] je predložen i opisan koncept multispektralnog SIFT deskriptora (MSIFT), dok je u [13] predložen koncept multispektralnog rječnika vizuelnih riječi (MBoVW). Oba pristupa se mogu koristiti i prilikom klasifikacije slika u boji.

MSIFT Multispektralni SIFT je histogram lokalnih obilježja koji se računa prema rječniku lokalnih obilježja teksture dobijenom SIFT transformacijom na sljedeći način. Svaki od N dostupnih spektralnih opsega se tretira kao jedna siva slika. Sa istih prostornih koordinata, računaju se lokalna obilježja iz svih kanala, koja se spajaju u jedan $128 \cdot N$ -dimenzionalni vektor:

$$\mathbf{x}_{1 \times 128 \cdot N}^{(j)} = \underbrace{\begin{bmatrix} x_{1,1}^{(j)} & x_{1,2}^{(j)} & \dots & x_{1,128}^{(j)} \end{bmatrix}}_{\text{kanal 1}} \underbrace{\begin{bmatrix} x_{2,1}^{(j)} & x_{2,2}^{(j)} & \dots & x_{2,128}^{(j)} \end{bmatrix}}_{\text{kanal 2}} \dots \underbrace{\begin{bmatrix} x_{N,1}^{(j)} & x_{N,2}^{(j)} & \dots & x_{N,128}^{(j)} \end{bmatrix}}_{\text{kanal N}}. \quad (5.4)$$

Iz određenog broja ovako formiranih lokalnih vektora obilježja:

$$\left\{ \mathbf{x}_{1 \times 128 \cdot N}^{(1)}, \mathbf{x}_{1 \times 128 \cdot N}^{(2)}, \dots, \mathbf{x}_{1 \times 128 \cdot N}^{(j)}, \dots \right\} \quad (5.5)$$

klasterizacijom se dobija odgovarajući rječnik vizuelnih riječi $\mathbf{V}_{\text{MSIFT}}$. Dakle, lokalna obilježja multispektralnih slika dobijaju se spajanjem lokalnih obilježja pojedinačnih

kanala sa istih prostornih koordinata. Prema rječniku $\mathbf{V}_{\text{MSIFT}}$ formira se histogram, koji predstavlja MSIFT deskriptor.

MBOVW Multispektralni rječnik vizuelnih riječi koristi drugačiji pristup prilikom formiranja rječnika, pri čemu se ne kreiraju multispektralna lokalna obilježja, kao što je dato jednačinom (5.4). U ovom slučaju, lokalna vizuelna obilježja iz svih dostupnih kanala se zajedno klasterizuju, bez obzira na to kojem kanalu pripadaju. Ukoliko se multispektralno lokalno obilježje, dato jednačinom (5.4) napiše u obliku:

$$\mathbf{x}_{1 \times 128 \cdot N}^{(j)} = \left[\mathbf{x}_{1 \dots 128}^{(j)}; \mathbf{x}_{129 \dots 128 \cdot 2}^{(j)}; \dots \mathbf{x}_{128 \cdot (N-1) + 1 \dots 128 \cdot N}^{(j)} \right], \quad (5.6)$$

formira se skup vizuelnih riječi koji se sastoji iz vizuelnih riječi iz svih dostupnih spektralnih opsega:

$$\left\{ \dots \mathbf{x}_{128 \cdot (N-1) + 1 \dots 128 \cdot N}^{(j-1)}, \mathbf{x}_{1 \dots 128}^{(j)}, \mathbf{x}_{129 \dots 128 \cdot 2}^{(j)}, \dots \mathbf{x}_{128 \cdot (N-1) + 1 \dots 128 \cdot N}^{(j)}, \mathbf{x}_{1 \dots 128}^{(j+1)} \dots \right\}. \quad (5.7)$$

Nakon klasterizacije podataka iz (5.7), dobija se multispektralni rječnik vizuelnih riječi $\mathbf{V}_{\text{MBOVW}}$, na osnovu kojeg se računa histogram tj. deskriptor MBOVW.

5.4. Rezultati klasifikacija multispektralnih slika predloženim deskriptorima

Kako bi se ispitala efikasnost predloženih algoritama, vrši se automatska klasifikacija baza UCM i NirScene, sa različitim dimenzionalnostima konačnih deskriptora slika u boji i multispektralnih slika. Pošto su razni koraci pretprocesiranja slika i postprocesiranja deskriptora već razmatrani u prethodnim eksperimentima, neće biti ispitivan njihov uticaj, iako se performanse različitih deskriptora mogu poboljšati boljim podešavanjem parametara prilikom pretprocesiranja ili parametara klasifikatora. U svakom od narednih eksperimenata, slike su smanjene na pola svojih originalnih dimenzija, a konačni deskriptori su normalizovani rL1 normalizacijom i uniformno skalirani. Korišćen je SVM klasifikator sa linearnim kernelom i regularizacionim parametrom $C = 1$. Konkretno, pored se performanse sljedećih deskriptora:

- (a) GISTCONC je deskriptor formiran tako što se spajaju GIST deskriptori sivih slika svakog od dostupnih kanala multispektralne slike, nakon čega se vrši redukcija dimenzionalnosti prema DescriptorConc algoritmu;
- (b) GISTBAND je deskriptor formiran spajanjem GIST deskriptora pojedinačnih kanala redukovane dimenzionalnosti, prema DescriptorBand algoritmu;
- (c) MSIFTCONC je deskriptor formiran prema rječniku vizuelnih riječi dobijenim klasterizacijom multispektralnih lokalnih obilježja (5.4), nakon čega je izvršena redukcija dimenzionalnosti;
- (d) MSIFTBAND je deskriptor koji se dobija spajanjem SIFT deskriptora sivih slika pojedinačnih kanala, nakon redukcije njihove dimenzionalnosti, pri čemu se rječnici vizuelnih riječi formiraju posebno za svaki od kanala;
- (e) MBOVW je deskriptor koji se formira prema multispektralnom rječniku vizuelnih riječi, nakon čega se izvršava redukcija njegove dimenzionalnosti;

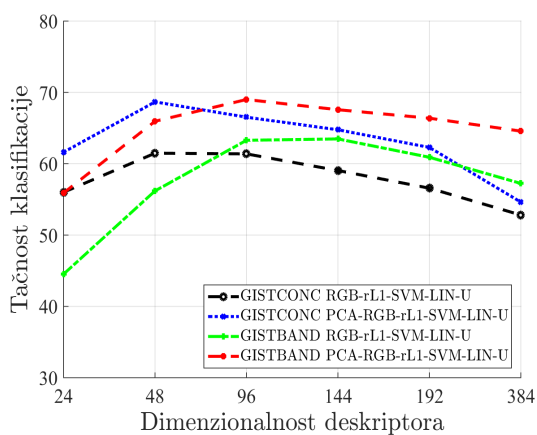
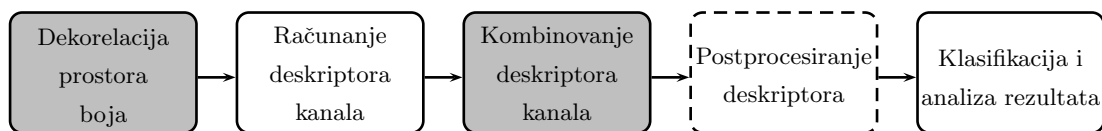
- (f) CENTRISTCONC je deskriptor koji se formira spajanjem CENTRIST deskriptora pojedinačnih kanala, nakon čega se vrši redukcija dimenzionalnosti, prema DescriptorConc algoritmu;
- (g) CENTRISTBAND je deskriptor formiran spajanjem CENTRIST deskriptora pojedinačnih kanala prethodno redukovane dimenzionalnosti, prema DescriptorBand algoritmu;
- (h) LBPCONC je deskriptor koji se formira spajanjem LBP deskriptora pojedinačnih kanala, nakon čega se vrši redukcija dimenzionalnosti, prema DescriptorConc algoritmu;
- (i) LBPBAND je deskriptor formiran spajanjem LBP deskriptora pojedinačnih kanala, prethodno redukovane dimenzionalnosti, prema DescriptorBand algoritmu.

Prilikom klasifikacija slika iz UCM baze, 80 slika iz svake klase je slučajnim izborom korišćeno za trening, a ostale za testiranje. Slično, prilikom klasifikacije slika iz NirScene baze, 40 slika je korišćeno za trening, a ostale za testiranje. U svakom eksperimentu, podjela na trening i test slike je urađena 5 puta. Za svih 5 podjela je izvršena klasifikacija, tako konačan rezultat klasifikacije predstavlja srednju vrijednost tačnosti klasifikacije za svih 5 podjela.

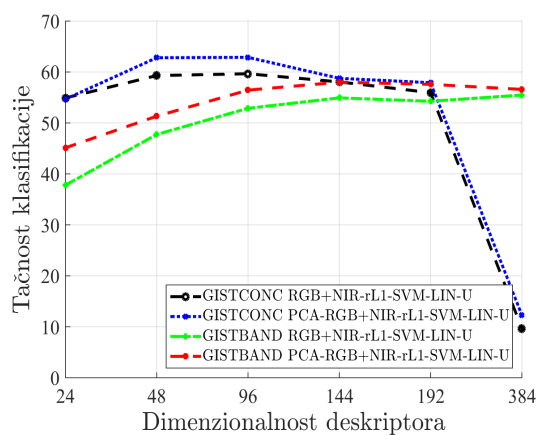
Svi eksperimenti su izvršeni tako da se ispita tačnost klasifikacije za nekoliko različitih dimenzionalnosti. Prilikom izbora dimenzionalnosti deskriptora, vodilo se računa o mogućnostima poređenja performansi deskriptora koji se formiraju na različite načine. Što se tiče deskriptora koji su računati prema DescriptorConc algoritmu nema puno ograničenja u izboru nove dimenzionalnosti, s obzirom da se redukcija dimenzionalnosti vrši nakon spajanja deskriptora pojedinačnih kanala. Međutim, u slučaju deskriptora računatih prema DescriptorBand algoritmu, dimenzionalnost redukovanih deskriptora pojedinačnih kanala treba da bude takva da se njihovim spajanjem dobije željena dimenzionalnost konačnog deskriptora. Npr. u slučaju da je GISTBAND deskriptor dimenzionalnosti 96, a koristi se prilikom klasifikacije slika u boji, deskriptori svakog pojedinačnog kanala se redukuju na dimenzionalnost $96/3 = 32$, prije njihovog spajanja. Ako se želi ispitati GISTBAND deskriptor dimenzionalnosti 96 u slučaju klasifikacije multispektralnih slika sa četiri dostupna kanala, deskriptori svakog pojedinačnog kanala se redukuju na dimenzionalnost $96/4 = 24$ prije njihovog spajanja. Slično važi i multispektralne verzije sa deskriptora sa CENTRIST i LBP deskriptorima sivih slika. Što se tiče varijanti baziranih na SIFT-u, potrebno je razlikovati sljedeće: MSIFTCONC se formira prema multispektralnim lokalnim obilježjima, tako da za slučaj od 1000 kodnih riječi (koji se koristi u svim eksperimentima), dobijamo deskriptor dimenzionalnosti 1000 prije njegove redukcije. MSIFTBAND računa deskriptore za pojedinačne kanale, tako da se za 1000 kodnih riječi po kanalu, formira onoliko deskriptora dimenzionalnosti 1000 koliko ima kanala. Prvo se vrši redukcija njihove dimenzionalnosti, pa zatim njihovo spajanje u konačan multispektralni deskriptor. MBOVW računa deskriptore prema multispektralnoj kodnoj knjizi, tako da je konačna dimenzionalnost deskriptora jednaka broju kodnih riječi, nakon čega se prema potrebi vrši redukcija dimenzionalnosti.

Pregledom rezultata datih na Sl. 5.5, mogu se izvesti zanimljivi zaključci. Interesantno je da se, u većini slučajeva, dobija veća tačnost klasifikacije za Descrip-

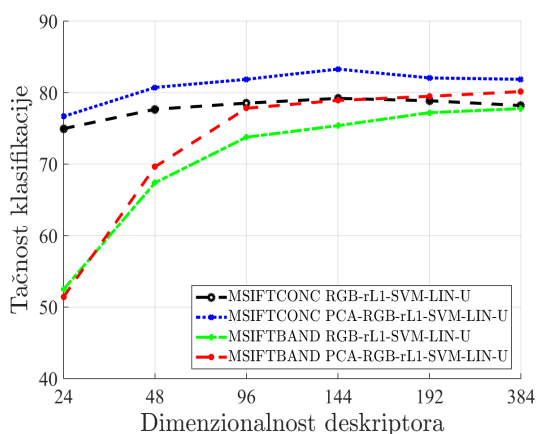
5.4. Rezultati klasifikacija multispektralnih slika predloženim deskriptorima



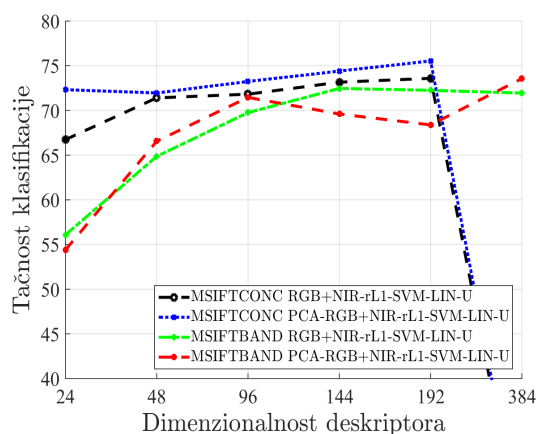
(a) UCM baza, GIST deskriptor



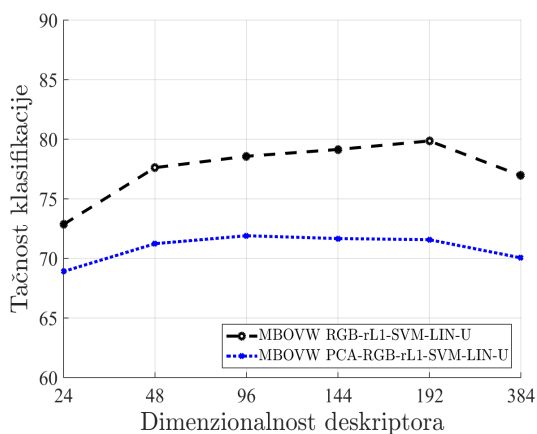
(b) NirScene baza, GIST deskriptor



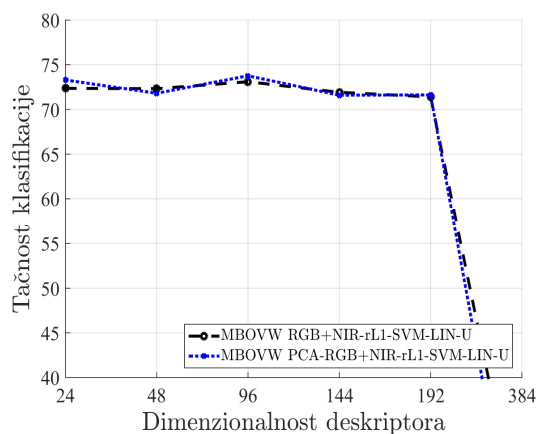
(c) UCM baza, MSIFT deskriptor



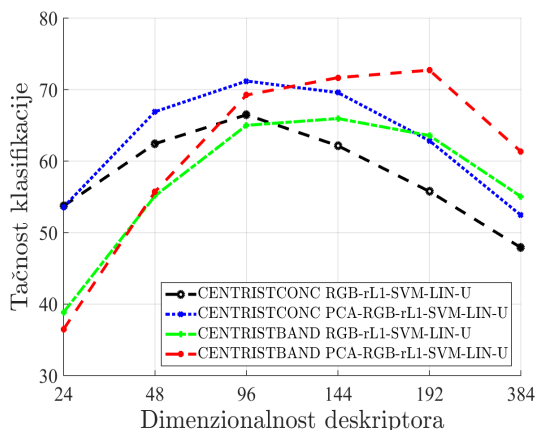
(d) NirScene baza, MSIFT deskriptor



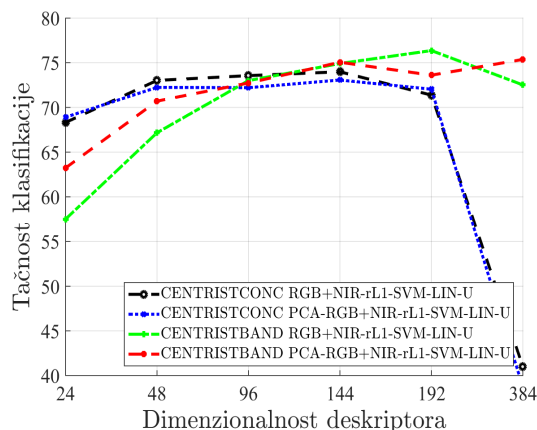
(e) UCM baza, MBOVW deskriptor



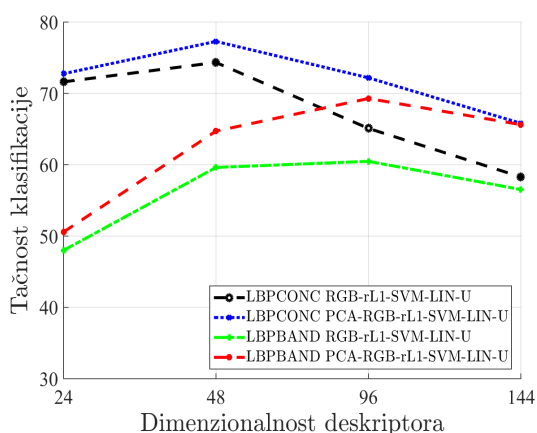
(f) NirScene baza, MBOVW deskriptor



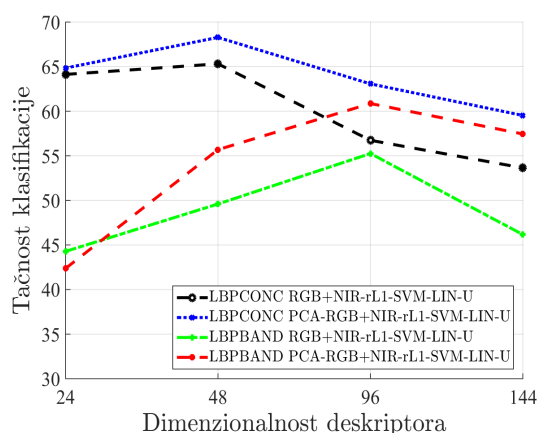
(g) UCM baza, CENTRIST deskriptor



(h) NirScene baza, CENTRIST deskriptor



(i) UCM baza, LBP deskriptor



(j) NirScene baza, LBP deskriptor

Slika 5.5: Analiza uticaja dimenzionalnosti multispektralnih deskriptora na tačnost klasifikacije

torConc za manje dimenzionalnosti konačnog deskriptora, što se moglo i očekivati imajući u vidu da se klasifikuju prirodne slike i aero snimci. Za multispektralne deskriptore bazirane na GIST-u, za bazu slika u boji (rezultati prikazani na Sl. 5.5a), najveća tačnost klasifikacije dobija se za slučaj GISTCONC deskriptora izračunatog u PCA-RGB kolor prostoru i to za relativno malu dimenzionalnost od 48. Poredeći sa slučajem kada se ne radi redukcija dimenzionalnosti, tj. kada je ona jednaka 1536, tačnost klasifikacije je manja za približno 5%, ali je klasifikacija znatno brža. GISTBAND varijanta daje približno iste rezultate kao i GISTCONC, ali za duplo veću dimenzionalnost deskriptora. U oba slučaja, primjetno je poboljšanje kada se koristi prethodno dekokorelisan kolor prostor. Za bazu multispektralnih slika, rezultati klasifikacije sa deskriptorima baziranim na GIST-u su dati na Sl. 5.5b. Može se primjetiti da GISTCONC varijanta ponovo ima bolje performanse u poređenju sa GISTBAND pristupom. Najveća tačnost klasifikacije dobija se za dimenzionalnost 48, što je mnogo manje od početne dimenzionalnosti jednake 2048, međutim, ponovo dolazi do smanjenja tačnosti klasifikacije od nekoliko procenata u odnosu na slučaj kada se koristi deskriptor neredukovane dimenzionalnosti.

Rezultati klasifikacije za deskriptore bazirane na SIFT-u, prikazani su na Sl. 5.5c

i 5.5d za MSIFT i na Sl. 5.5e i 5.5f za MBOVW. Poređenjem rezultata, može se primjetiti da MBOVW deskriptor ima nešto lošije performanse u odnosu na MSIFT bazirane varijante. Takođe, interesantno je primjetiti da se dekorelacijom kolor prostora pogoršava tačnost klasifikacije za slučaj slika u boji za MBOVW deskriptor, što se može objasniti većom raznovrsnošću kodnih riječi prilikom računanja kodne knjige. Sa druge strane MSIFTCONC deskriptor ostvaruje bolje rezultate klasifikacije u poređenju sa MSIFTBAND deskriptorom, slično kao u slučaju deskriptora baziranih na GIST-u. Ponovo, može se zaključiti da je baza prirodnih slika uzrok ovakvim rezultatima klasifikacije. U poređenju sa rezultatima klasifikacije koji se dobijaju za deskriptore pune dimenzionalnosti, MSIFTCONC računat iz PCA-**RGB** kolor prostora ostvaruje približno jednake rezultate kao SIFT deskriptor pune dimenzionalnosti računat iz **RGB** kolor prostora, prilikom klasifikacije slika u boji iz UCM baze. Dakle, tačnost klasifikacije od približno 83% se postiže sa deskriptorom dimenzionalnosti 1000 ili sa predloženim deskriptorom dimenzionalnosti 144. Sa druge strane, za NirScene bazu multispektralnih slika, smanjuje se tačnost klasifikacije za nekoliko procenata.

Performanse deskriptora baziranih na CENTRIST-u za obje baze, date su na Sl. 5.5g i 5.5h. U ovom slučaju, može se primjetiti da su performanse CENTRISTCONC deskriptora bolje sa nižim dimenzionalnostima, dok CENTRISTBAND daje bolje rezultate klasifikacije na nešto višim dimenzionalnostima. U slučaju klasifikacije aero snimaka, CENTRISTCONC deskriptor dimenzionalnosti 48 računat prema PCA-**RGB** kolor prostoru, povećava tačnost klasifikacije za nekoliko procenata u poređenju sa CENTRIST deskriptorom pune dimenzionalnosti 762. Neznatno veća tačnost klasifikacije se dobija sa CENTRISTBAND deskriptorom dimenzionalnosti 192. U slučaju klasifikacije multispektralnih slika, mogu se izvesti slični zaključci. Najveća tačnost klasifikacije dobija se sa CENTRISTBAND deskriptorom dimenzionalnosti 192 i to približno 77%.

Kada se koriste deskriptori bazirani na LBP-u, rezultati klasifikacije su dati na Sl. 5.5i i 5.5j za UCM i NirScene baze, redom. Mogu se izvesti slični zaključci kao za deskriptore bazirane na CENTRIST-u. Bolje performanse dobijaju se na niskim dimenzionalnostima kada se koristi LBPCONC deskriptor, dok je LBPBAND nešto bolji za veće dimenzionalnosti. Takođe, važno je primjetiti da se primjenom LBPCONC deskriptora dimenzionalnosti 48 popravljaju rezultati klasifikacije u poređenju za LBP deskriptorom pune dimenzionalnosti 162, za slučaj slika u boji iz baze UCM. Dobri rezultati klasifikacije se mogu objasniti rotacionom invarijantnošću ovog deskriptora i činjenicom da su aero snimci često imaju rotiran prikaz u zavisnosti od ugla pod kojima se vrši snimanje.

5.5. Poređenje rezultata klasifikacije i diskusija

Iako se analizom rezultata prethodnih eksperimenata došlo do nekoliko važnih zaključaka, interesantno je izvršiti direktno poređenje rezultata klasifikacije za slučajeve kada se koriste različiti parametri prilikom računanja deskriptora. U tabeli 5.1 izvršena je komparativna analiza tačnosti klasifikacije za različite dimenzionalnosti deskriptora. Prilikom klasifikacije slika iz UCM baze, 80 slika iz svake klase je

5.5. Poređenje rezultata klasifikacije i diskusija

Tabela 5.1: Poređenje tačnosti klasifikacije multispektralnih slika u slučajevima kada se deskriptori računaju i procesiraju na različite načine. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama prosječni vremenski intervali potrebno za izvršavanje jedne klasifikacije.

UCM baza				
	GIST	MSIFT	CENTRIST	LBP
dimenzionalnost	1536	1000	762	162
100% + LIN	67.76 (8.7)	83.28 (9.8)	73.81 (3.3)	78.95 (1.2)
50% + LIN	67.81 (6.8)	81.23 (7.1)	72.00 (3.9)	74.43 (0.9)
50% + rL1 + LIN-U	69.38 (6.7)	82.42 (7.2)	65.28 (4.5)	75.61 (0.6)
PCA 100% + LIN	72.05 (5.7)	86.00 (9.5)	77.81 (3.0)	80.62 (0.8)
PCA 100% + rL1 + LIN-U	72.19 (6.0)	90.81 (9.6)	76.42 (4.0)	81.67 (0.6)
PCA 50% + rL1 + LIN-U	73.62 (5.7)	86.04 (6.5)	67.42 (5.6)	78.42 (3.4)
	GISTCONC	MSIFTCONC	CENTRISTCONC	LBPCONC
dimenzionalnost	192	256	96	96
100% + LIN	65.33 (1.9)	80.67 (1.3)	70.76 (1.2)	78.38 (1.2)
50% + LIN	62.52 (1.9)	79.62 (1.2)	69.95 (1.1)	74.28 (1.0)
50% + rL1 + LIN-U	55.57 (3.1)	78.38 (2.0)	66.47 (1.8)	65.76 (0.6)
PCA 100% + LIN	67.52 (2.0)	83.47 (1.5)	75.00 (1.2)	80.23 (0.8)
PCA 100% + rL1 + LIN-U	67.67 (2.4)	85.85 (2.0)	75.23 (1.2)	77.47 (0.5)
PCA 50% + rL1 + LIN-U	62.28 (2.4)	82.14 (2.0)	74.23 (1.0)	77.28 (0.5)

NirScene baza				
	GIST	MSIFT	CENTRIST	LBP
dimenzionalnost	2048	1000	1016	216
100% + LIN	63.78 (0.6)	72.74 (0.5)	70.17 (0.2)	66.77 (0.2)
50% + LIN	63.73 (0.6)	68.40 (0.8)	68.95 (0.2)	62.61 (0.2)
50% + rL1 + LIN-U	61.94 (0.6)	76.87 (0.7)	74.65 (0.2)	65.32 (0.1)
PCA 100% + LIN	66.33 (0.9)	75.84 (0.4)	69.14 (0.2)	66.58 (0.1)
PCA 100% + rL1 + LIN-U	66.83 (0.7)	78.59 (0.4)	72.92 (0.3)	71.25 (0.1)
PCA 50% + rL1 + LIN-U	66.76 (0.6)	76.26 (0.6)	73.00 (0.3)	71.23 (0.1)
	GISTCONC	MSIFTCONC	CENTRISTCONC	LBPCONC
dimenzionalnost	192	256	128	128
100% + LIN	60.28 (2.3)	71.20 (0.9)	69.91 (0.3)	65.17 (0.2)
50% + LIN	58.06 (2.4)	72.23 (1.0)	68.09 (0.3)	63.45 (0.2)
50% + rL1 + LIN-U	55.95 (2.3)	72.22 (1.1)	73.20 (0.3)	53.54 (0.1)
PCA 100% + LIN	62.30 (2.4)	74.41 (1.0)	70.58 (0.3)	64.90 (0.2)
PCA 100% + rL1 + LIN-U	56.92 (2.3)	70.57 (0.4)	68.17 (0.3)	70.23 (0.1)
PCA 50% + rL1 + LIN-U	57.89 (2.6)	72.76 (0.4)	72.92 (0.3)	58.59 (0.1)

korišćeno za trening, a ostale za testiranje. Korišćena je ista podjela na trening i test slike kao i u prethodnim eksperimentima, kako bi poređenje rezultata bilo korektno. Slično, za NirScene bazu za trening je korišćeno 40 slika iz svake klase. Takođe, rezultati dati u tabeli 5.1 su značajni jer u literaturi ne postoji detaljno poređenje performansi klasifikacije za različite baze slika i za različite vrijednosti parametara.

Konačni rezultati klasifikacije predstavljaju srednju vrijednost tačnosti za 5 različitih podjela na trening i test slike, dok brojevi u zagradama predstavljaju prosječni vremenski interval u sekundama potreban za izvršavanje klasifikacije prema jednoj trening/test podjeli¹. Koristi se SVM klasifikator sa linearnim kernelom, pri čemu je regularizacioni parametar $C = 1$. Bez redukcije dimenzionalnosti, multispektralni SIFT deskriptor je korišćen u varijanti MSIFTCONC, tako da je njegova ukupna dimenzionalnost jednaka 1000 za slike u boji i za multispektralne slike.

Razmatrajući potrebne vremenske intervale za izvršavanje klasifikacije, treba skrenuti pažnju da dimenzionalnost deskriptora nije jedini parametar. Veći dio vremena se utroši na obučavanje klasifikatora, pa je za brzinu izvršavanja važno koliko se trening slika koristi za obučavanje. Konkretno, u eksperimentima čiji su rezultati dati u tabeli 5.1, za UCM bazu koristi se ukupno 1680 trening slika, dok je za NirScene bazu broj trening slika jednak 450. Sa aspekta učenja klasifikatora veći broj trening slika znači više iteracija prilikom obučavanja. S obzirom na to, nije pogodno porediti brzine klasifikacije za različite baze, ali je moguće procijeniti eventualna ubrzanja kada se koriste deskriptori redukovane dimenzionalnosti. Ipak, treba naglasiti da se prilikom klasifikacije pomoću deskriptora redukovane dimenzionalnosti izvršava nekoliko računanja transformacija za redukciju dimenzionalnosti. Za NirScene bazu vidimo da postoji povećanje ili neznatno smanjenje vremena klasifikacije u slučaju korišćenja redukcije dimenzionalnosti, pa se može zaključiti da je proces redukcije dimenzionalnosti vremenski zahtjevniji od obučavanja klasifikatora.

U opštem slučaju, može se primjetiti da se redukcijom dimenzionalnosti smanjuje tačnost klasifikacije za nekoliko procenata, ali za nekoliko slučajeva, tj. za posebne kombinacije upotrebljenih koraka procesiranja i vrijednosti parametara, povećava se tačnost klasifikacije kada se koriste predloženi niskodimenzionalni deskriptori. Ti slučajevi su posebno naznačeni u tabeli 5.1.

5.6. Mogućnosti poboljšanja tačnosti klasifikacije

Pored različitih metoda za pretprocesiranje slika i procesiranje deskriptora, kao i varijacije parametara klasifikatora, moguće je na druge načine poboljšati tačnost klasifikacije. Veliki broj metoda oslanja se na SIFT detektor obilježja i paradigmu vizuelnih riječi, ali koristi drugačije načine za računanje rječnika vizuelnih riječi [63, 62] i različite načine za formiranje deskriptora [81, 82, 42, 83, 84, 51]. Iz prethodno navedene literature, može se zaključiti da je jedan istraživački pravac za poboljšanje tačnosti klasifikacije prirodnih tekstura zasnovan na boljem matematičkom modelovanju vektora obilježja. Drugi način za postizanje boljih rezultata jeste upotreba

¹Ovi eksperimenti su izvršeni na računaru sa Intel Core i5-4300M procesorom na 2.6 GHz i sa 8 GB RAM-a, bez paralelizacije ili upotrebe GPU procesora.

konvolucionih neuronskih mreža [69, 70, 41, 16]. U suštini, konvolucione neuronske sadrže veći broj jednostavnih struktura za učenje, tako da je mreža sposobna naučiti i iskoristiti veliki broj lokalnih obilježja u slučaju da postoji dovoljno veliki broj trening podataka.

Konvolucione neuronske mreže su neprikosnovene ukoliko postoji dovoljno veliki broj trening slika, ali ne mogu se obučiti na osnovu malog broja slika. Tako da možemo uočiti treći pristup za poboljšanje tačnosti klasifikacije, koji zapravo predstavlja kompromis između modelovanja i konvolucionih neuronskih mreža. U ovom pristupu koristi se što više podataka o lokalnim obilježjima, sa malog broja trening slika, kako bi se izračunali deskriptori koji bolje modeluju teksturalne karakteristike prikaza u slikama. Dugo su memorijski resursi bili ograničavajući faktor prilikom računanja SIFT deskriptora, ali posljednjih nekoliko godina razvoj hardvera uspjeva da nadomjesti potrebe za velikom količinom memorije. Dovoljni memorijski resursi omogućavaju manipulaciju sa velikom količinom lokalnih obilježja, što može pomoći boljem modelovanju vizuelnog sadržaja na slikama, što je diskutovano u [15].

Da bi se ilustrovale mogućnosti prethodno opisanog načina računanja deskriptora, u [15] je izvršena analiza performansi SIFT deskriptora prilikom klasifikacije sivih slika u slučajevima kada se koristi veliki broj lokalnih obilježja teksture i kada se formiraju rječnici sa velikim brojem kodnih riječi. U skladu sa rezultatima prezentovanim u [15], postavljeni su eksperimenti koji ispituju efikasnost predloženog metoda u slučaju klasifikacije slika u boji i multispektralnih slika. Prije svega, ispitano je koliko povećanje dimenzionalnosti SIFT deskriptora može uticati na tačnost klasifikacije, a zatim su eksperimenti ponovljeni u slučaju kada se koristi mnogo više podataka prilikom računanja SIFT deskriptora. Ispituje se uticaj povećanja broja kodnih riječi i količine podataka koji opisuju lokalna obilježja. Za MSIFTBAND deskriptor, tačnosti klasifikacije za povećan broj kodnih riječi su dati u tabeli 5.2.

Poređenjem tačnosti klasifikacije datih u tabeli 5.2 sa rezultatima datih u tabeli 5.1, vidimo da povećanjem količine lokalnih obilježja i dimenzionalnosti može poboljšati tačnost klasifikacije. Konkretno, tačnost klasifikacije UCM baze slika u

Tabela 5.2: Poređenje tačnosti klasifikacije multispektralnih slika SIFT deskriptorom, za različite veličine kodnih knjiga. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama standardne devijacije za 5 različitih trening/test podjela

UCM baza				
dimenzionalnost	3000	6000	15000	30000
50% + rL1 + LIN-U	85.47 ± 1.08	84.62 ± 1.54	85.33 ± 2.20	83.28 ± 1.46
PCA 50% + rL1 + LIN-U	90.52 ± 2.05	87.04 ± 2.81	86.42 ± 1.71	86.42 ± 2.91

NirScene baza				
dimenzionalnost	4000	8000	20000	40000
50% + rL1 + LIN-U	79.90 ± 0.99	77.41 ± 2.41	77.64 ± 1.97	76.50 ± 2.81
PCA 50% + rL1 + LIN-U	79.08 ± 3.58	78.49 ± 2.86	79.31 ± 2.82	77.86 ± 3.26

boji izvedene pod istim uslovima (pretprocesiranje slika i normalizacija deskriptora) se povećava za otprilike 5%, dok je povećanje tačnosti klasifikacije za NirScene bazu otprilike 3%. Takođe, može se primjetiti da dalje povećanje dimenzionalnosti ne garantuje veću tačnost klasifikacije, odakle se može zaključiti da više dostupnih podataka ne garantuje veću tačnost klasifikacije, ali može da je poboljša. Takođe, više podataka zahtjeva više računarskih i memorijskih resursa. Stoga je potrebno naći kompromis između optimalne količine podataka za reprezentaciju slike i brzine izvršavanja klasifikacije.

Za bolje opisivanje vizuelnog ili semantičkog sadržaja na slikama, koriste se i razni drugi pristupi. Na primjer, moguće je iskoristiti podatke o prostornom rasporedu lokalnih obilježja na slikama. U literaturi je predloženo nekoliko načina za eksploataciju prostornog rasporeda obilježja tekstura [8, 56, 49, 10, 57, 61, 11, 14], što će biti detaljno diskutovano u narednom poglavlju. Autori u [11] su iskoristili detektor ivica primjenjen na R kanal, kako bi definisali novi kanal u multispektralnim slikama i obogatili količinu informacija o ivicama na slikama.

5.7. Zaključci važni za klasifikaciju multispektralnih slika

Eksperimenti čiji su rezultati predstavljeni u ovom poglavlju su osmišljeni tako da ispituju mogućnosti računanja niskodimenzionalnih deskriptora teksture multispektralnih slika i njihove performanse prilikom klasifikacije slika. Ispitani su uticaji različitih kolor prostora, kao i dekorelacije RGB prostora izvedene iz podataka konkretne baze slika i uticaji različitih algoritama za kombinovanje tehnika redukcije dimenzionalnosti i deskriptora pojedinačnih kanala multispektralnih slika. Takođe, navedene su mogućnosti za dalje poboljšanje tačnosti klasifikacije. Na osnovu prethodnih rezultata, moguće je izvesti nekoliko bitnih zaključaka:

- (1) Tačnost klasifikacije slika u boji i multispektralnih slika zavisi od izbora kolor prostora iz kojeg se računaju deskriptori. Dekorelacijom RGB kolor prostora povećava se tačnost klasifikacije za nekoliko procenata, bez obzira na izbor deskriptora ili bazu slika koja se klasifikuje. Ukoliko je moguće, poželjno je izračunati transformaciju za dekorelisanje kolor prostora, a moguće je koristiti neki od drugih oponentnih kolor prostora, kao što je Lab ili oRGB.
- (2) Dodatne informacije iz spektralnih opsega van vidljivog dijela spektra doprinose povećanju tačnosti klasifikacije. Npr., prilikom klasifikacije aero snimaka i scena koje sadrže vodene površi ili vegetaciju, podaci iz NIR kanala su naročito korisni jer je refleksija talasa u NIR opsegu karakteristična. Pored uticaja na tačnost klasifikacije, korišćenjem dodatnih kanala povećava se dimenzionalnost deskriptora i potrebni memorijski resursi, pa je vremenski interval izvršavanja klasifikacije duže.
- (3) Kombinovanjem podataka iz više kanala i tehnika za redukciju dimenzionalnosti moguće je značajno smanjiti dimenzionalnost deskriptora i kreirati niskodimenzionalni deskriptor multispektralnih slika. Eksperimentalno je pokazano

da takvi deskriptori postižu slične ili bolje rezultate klasifikacije kao i multispektralni deskriptori punih dimenzionalnosti.

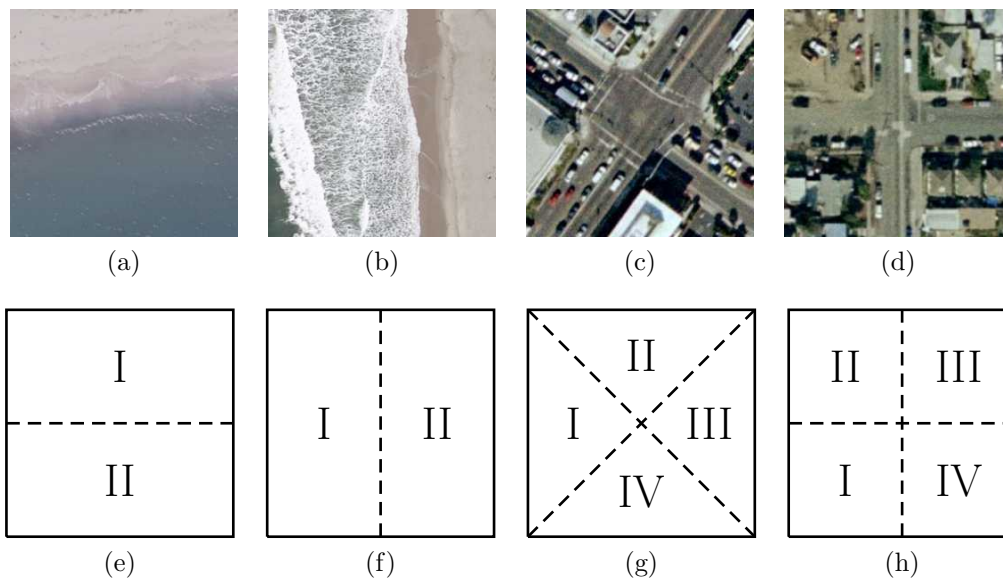
- (4) Za svaki od deskriptora tekstone sivih slika, podešavanjem parametara klasifikacije, načina pretprocesiranja slika, kombinovanja podataka iz različitih kanala i procesiranja deskriptora, mogu se optimizovati performanse prilikom klasifikacije multispektralnih slika. Svaki deskriptor je dizajniran tako da opisuje sadržaj slike na različit način, tako da svakom deskriptoru odgovaraju različiti parametri klasifikacije.
- (5) Više podataka znači bolji model i veću tačnost klasifikacije. Ukoliko je moguće, potrebno je izdvojiti što veću količinu lokalnih obilježja, kako bi se omogućilo tačnije obučavanje klasifikatora.

6 Prostorni raspored lokalnih obilježja teksture

Značaj prostornog rasporeda lokalnih obilježja je diskutovan u uvodnom poglavlju i prilikom razmatranja metodologije istraživanja, ali nije uzet u obzir prilikom postavke eksperimenata u prethodnim poglavljima. Poznato je da prostorni raspored regiona različitih karakteristika na slikama ima značajan uticaj na vizuelnu sličnost slika, pa je poželjno razmatranje mogućnosti kako iskoristiti podatke o prostornom rasporedu. Potrebno je iskoristiti informacije o tome da li se lokalna obilježja sličnih vizuelnih karakteristika nalaze na istom mjestu na različitim slikama. Moguće je iskoristiti slučaj kada su lokalna obilježja na jednoj slici pomjerena za isto odstojanje u odnosu na položaj sličnih lokalnih obilježja na drugoj slici.

Prilikom eksploatacije prostornih karakteristika obilježja, u većini slučajeva koriste se lokalni deskriptori, pošto se za njihovo računanje koristi grupa od nekoliko susjednih piksela koje je lako prostorno lokalizovati. Međutim, prostorni raspored teksture je prvi put upotrebljen prilikom računanja globalnog deskriptora, jer GIST se sastoji od koeficijenata Gaborove vejevlet transformacije na 16 nepreklapajućih pravougaonih regiona na slici [8]. Razmatranje prostorne podjele lokalnih deskriptora na više nivoa, prvi put je diskutovano u [56] kada je zaključeno da to može doprinijeti povećanju tačnosti klasifikacije. Kompletan CENTRIST deskriptor vrši prostornu podjelu koeficijenata dobijenih nakon Census transformacije na 31 pravougli preklapajući region na tri nivoa [10, 11]. Posebno se razmatra centralni dio slike sa 10 preklapajućih pravougaonih regiona na drugom i trećem nivou.

Pogodnosti uzimanja u obzir prostornog rasporeda teksture mogu se razmotriti sa primjera datih na Sl. 6.1. Prikazane su po dvije slike iz klase *beach* i klase *intersection*, koje pripadaju bazi UCM. Aero snimci dati na Sl. 6.1a i 6.1b sastoje se od dva disjunktna regiona, koja prikazuju plažu i more. Na Sl. 6.1a snimak je takve orijentacije da je plaža prikazana u gornjem dijelu slike, a more u donjem dijelu slike, dok je na Sl. 6.1b orijentacija snimka takva da je more prikazano na lijevom dijelu slike, a plaža na desnom. Iako su slike iz iste semantičke klase, njihova međusobna vizuelna sličnost nije velika, jer je more u jednom slučaju mirno a u drugom slučaju

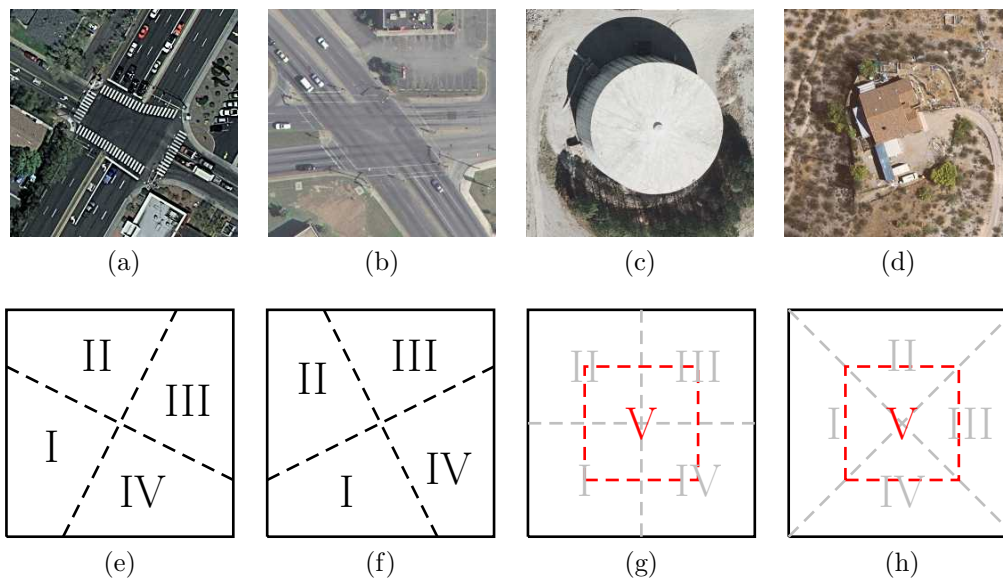


Slika 6.1: Primjeri različitih prostornih raspodjela tekstura na aero snimcima i odgovarajući prostorni rasporedi regiona

uzburkano, što je predstavljeno značajno drugačijom teksturom. Međutim, dijelovi slike koji prikazuju plažu imaju mnogo veću vizuelnu sličnost. Na Sl. 6.1c i 6.1d dati su snimci koji prikazuju raskršća, koja su različitih orijentacija. Prilikom poređenja ovih snimaka na osnovu vizuelne sličnosti mogli bi se porediti regioni snimaka koji prikazuju objekte pored raskrsnica. Ovi regioni se nalaze na različitim pozicijama na snimcima, ali su sličnog vizuelnog sadržaja. Na Sl. 6.1e, 6.1f, 6.1g i 6.1h su prikazane prostorne podjele slika koje se odnose na regione različitog vizuelnog sadržaja.

Pored uspravnih podjela prostora slike, kao što je dato na primjerima prikazanim na Sl. 6.1e, 6.1f i 6.1h, na aero snimcima se često pojavljuju i rotirane podjele, kao što je prikazano na Sl. 6.1g, gdje je podjela rotirana za 45° . Međutim, mogu se naći primjeri gdje rotacija za 45° nije optimalna. Takođe, prilikom detekcije usamljenih objekata na aero snimcima, treba voditi računa da su oni često prikazani na sredini slike. Dodatni primjeri podjela na regione su prikazani na Sl. 6.2. Na Sl. 6.2a i 6.2b vidimo aero snimke sa raskrsnicama, čija je rotacija takva da im ne odgovara uspravna podjela niti podjela rotirana za 45° , a odgovarajuće podjele su date na slikama 6.2e i 6.2f. U ovom slučaju, radi se o rotiranim podjelama prostora slika za 22.5° u smjeru kazaljke na satu i u suprotnom smjeru kazaljke na satu, redom. Kada su u pitanju usamljeni objekti na snimcima, kao što je prikazano na slikama 6.2c i 6.2d, pored prostorne podjele na disjunktne regione, korisno je iskoristiti dodatni region koji je centriran na slici, a približno je istih dimenzija kao i objekat koji se treba opisati. Pošto je određivanje dimenzija objekta usko vezano sa segmentacijom slika, često se problem pojednostavljuje usvajanjem pravougaonog centralnog regiona fiksnih dimenzija. Centralni region se često koristi zajedno sa disjunktne regionima cijele slike, kao što prikazano na slikama 6.2g i 6.2h.

Ovakav način podjele prostora slike može se posmatrati kao multirezoluciona analiza sadržaja slika, pri čemu kreiranje histograma lokalnih obilježja (tj. deskriptora) za cijelu sliku daje histogram nultog nivoa. Disjunktne regioni i centralni

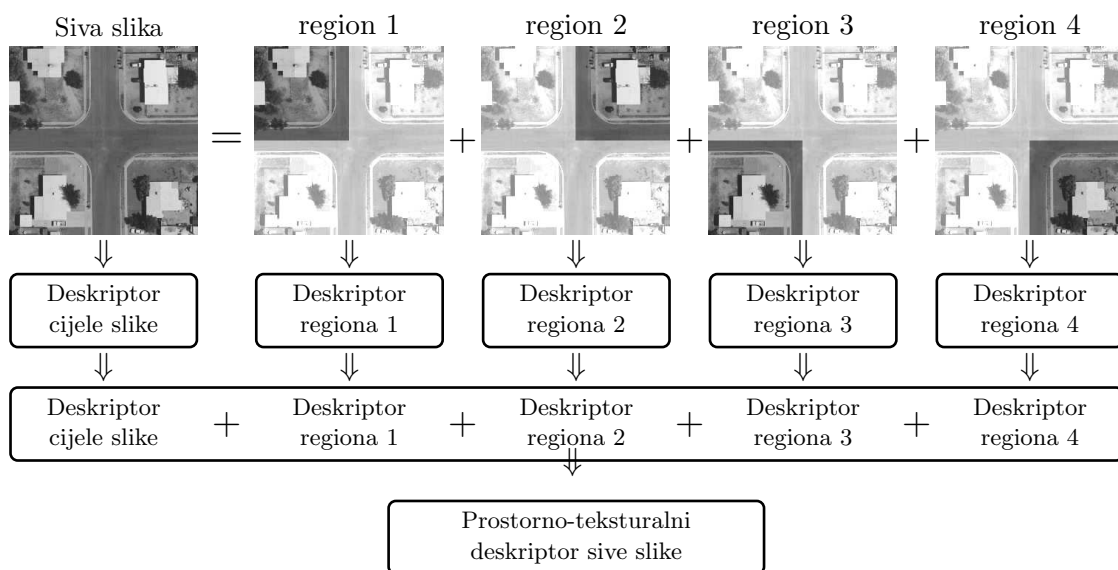


Slika 6.2: Primjeri različitih prostornih raspodjela tekstura na aero snimcima i odgovarajući prostorni rasporedi regiona

regioni prikazani na slikama 6.1 i 6.2 koriste se za kreiranje histograma prvog nivoa, a podjela se može nastaviti na višim nivoima za svaki od pravougaonih regiona iz prvog nivoa. Tako se regionima drugog nivoa dobijaju podjelom prostora jednog od regiona prvog nivoa, itd. Podjela prostora slike na više nivoa uzrokuje značaj porast broja regiona i porast dimenzionalnosti prostorno-teksturalnih deskriptora. Postoje i drugi načini za podjelu prostora slike. Npr., u [57] su razmatrano učenje optimalne podjele na osnovu sadržaja slike i verifikacija rezultata na osnovu jednog dijela trening slika. Međutim, u ovom slučaju proces učenja zahtjeva značajne računarske resurse.

6.1. Predloženo rješenje

Na osnovu analize prostornog rasporeda tekstura na aero snimcima (Sl. 6.1) može se zaključiti da korišćenje informacija o prostornom rasporedu lokalnih obilježja može povećati tačnost klasifikacije. Prema algoritmu predloženom u [14], lokalna obilježja teksture se računaju u svakom pikselu na slici, da bi se izvršilo kreiranje histograma prema unaprijed izabranom prostornom rasporedu. Ukoliko se koriste četiri regiona, kreiraju se četiri dodatna histograma koja se pridružuju histogramu izračunatom prema lokalnim obilježjima sa cijele slike. Cilj je da se u konačnom deskriptoru slike posebno istaknu regionima sa što uniformnijom teksturom ili semantičkim sadržajem, kako bi se smanjila udaljenost između deskriptora semantički sličnih slika. Za slučaj sive slike i uspravne podjele prostora na četiri nepreklapajuća regiona, šematski prikaz računanja proširenog deskriptora prikazan je na Sl. 6.3. Ukoliko se radi sa slikama u boji ili multispektralnim slikama, algoritam za računanje proširenog deskriptora se primjenjuje na sive slike svakog od pojedinačnih kanala, tako da se konačan prostorno-teksturalni deskriptor višekanalne



Slika 6.3: Šematski prikaz algoritma za računanje deskriptora teksture koji uključuje podatke o prostornom rasporedu lokalnih obilježja

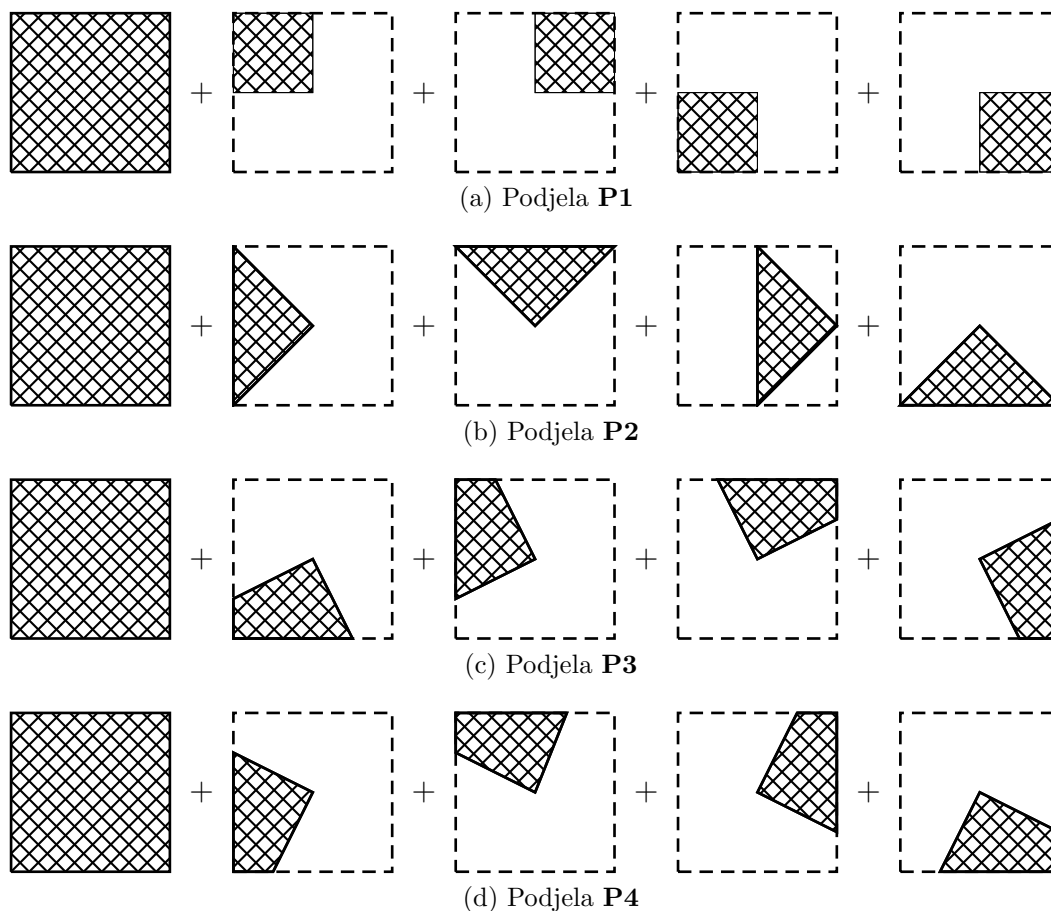
slike dobija spajanjem proširenih deskriptora pojedinačnih kanala.

Ovakav način kreiranja prostorno-teksturalnog deskriptora ima svoje prednosti i nedostatke. Prednost je to što je lokalna obilježja potrebno računati jednom, nakon čega se ona koriste potreban broj puta za računanje deskriptora regiona, ukoliko se regioni preklapaju. Tako je moguće naglasiti regione slike koji imaju poseban značaj za prepoznavanje semantičkog sadržaja prikazanog na slici. Očigledan nedostatak jeste značajno povećana dimenzionalnost ovako kreiranog deskriptora, što može usporiti klasifikaciju.

6.2. Postavka eksperimenta

Potrebno je ispitati uticaj proširenja lokalnih deskriptora teksture kada se uzimaju u obzir podaci o prostornom rasporedu lokalnih obilježja na slikama. Ovom prilikom neće biti razmatran uticaj redukcije dimenzionalnosti, već isključivo efikasnost prostorno-teksturalnih deskriptora. Predloženi model računanja prostorno-teksturalnih deskriptora slike, prikazan na Sl. 6.3, biće ispitan za različite podjele prostora slika, za različite lokalne deskriptore teksture i to za različite baze sivih slika. U ovom poglavlju se razmatraju isključivo baze sivih slika, kako bi se utvrdile performanse niskodimenzionalnih prostorno-teksturalnih deskriptora. Prostorno-teksturalni deskriptori će u narednom poglavlju biti prošireni na adekvatan način i iskorisćeni za klasifikaciju multispektralnih slika.

Da bi se izbjeglo preveliko povećanje dimenzionalnosti, vrši se samo podjela prostora slike prvog nivoa. Porede se tačnosti klasifikacije kada se kao lokalni deskriptori koriste SIFT, CENTRIST i LBP. U ovom eksperimentu, računaju se numeričke vrijednosti lokalnih obilježja za sve piksele osim graničnih, bez obzira na tip deskrip-



Slika 6.4: Načini podjele prostora slike i slučajevi predloženog algoritma za računanje prostorno-teksturalnog deskriptora

tora. Detekcijom obilježja se svakom pikselu dodjeljuje nova numerička vrijednost, što praktično znači da lokalni deskriptori vrše transformaciju slike u novi prostor u kojem vrijednosti piksela odgovaraju određenim numeričkim vrijednostima pripisanim lokalnim obilježjima. U zavisnosti od prostornog položaja tih obilježja, ona se koriste za formiranje histograma odgovarajućeg regiona. Četiri osnovna načina podjele koji se ispituju ilustrovani su na Sl. 6.4 i redom su označeni sa **P1**, **P2**, **P3** i **P4**. Za svaku od podjela, konačni prostorno-teksturalni deskriptor se računa na sljedeći način: računaju se lokalna obilježja na cijeloj slici, odgovarajući histogram za cijelu sliku (nulti nivo) i histogrami za svaki od regiona prema određenoj podjeli prostora slike (prvi nivo). Konačni prostorno-teksturalni deskriptor dobija se spajanjem svih histograma lokalnih obilježja sa nultog i prvog nivoa. Prema tome, za podjele **P1**, **P2**, **P3** i **P4** dobija se po pet histograma čijim se spajanjem dobija konačni deskriptor. Posebno će biti ispitan doprinos dodatnog centralnog histograma, što će u daljem tekstu biti označeno sa $\mathbf{P}_i + \mathbf{C}$, $i = 1 \dots 4$. Konačno, prostorno-teksturalni deskriptor označen sa **P5** čini spoj histograma cijele slike te histograma svih 16 regiona prikazanih na Sl. 6.4, a deskriptor **P5+C** sadrži i histogram centralnog regiona.

U narednim eksperimentima, prostorno-teksturalni deskriptori su kreirani prema lokalnim deskriptorima SIFT, CENTRIST i LBP, da bi se ispitalo postoji li po-

boljšanje tačnosti klasifikacije za predložene načine podjela prostora slike. Predložena su dva načina kombinovanja histograma regiona, s obzirom na to da histogram računat na osnovu obilježja sa cijele slike ima četiri puta više podataka nego svaki od regiona posebno.

SPN¹ Svaki od histograma na svim korišćenim nivoima se normalizuje posebno, nakon čega se spajaju u prostorno-teksturalni deskriptor za sivu sliku. Pošto se ovako dobijaju histogrami koji su predstavljeni vektorima obilježja jednake norme, nema uticaja različitog broja lokalnih obilježja. U slučaju multispektralnih slika, konačni prostorno-teksturalni deskriptori se dobijaju spajanjem deskriptora sivih slika.

SPR² Normalizacija konačnog deskriptora se može izvesti i nakon spajanja histograma pojedinačnih regiona, prilikom čega treba obezbijediti jednak broj lokalnih obilježja iz pojedinačnih regiona. U ovom pristupu, histogram cijele slike se formira nakon što se njene dimenzije smanje na dimenzije pojedinačnih regiona. Ako se radi o multispektralnim slikama, svaki kanal se obrađuje pojedinačno, nakon čega se spajaju deskriptori pojedinačnih kanala.

6.3. Klasifikacija prostorno-teksturalnim deskriptorima

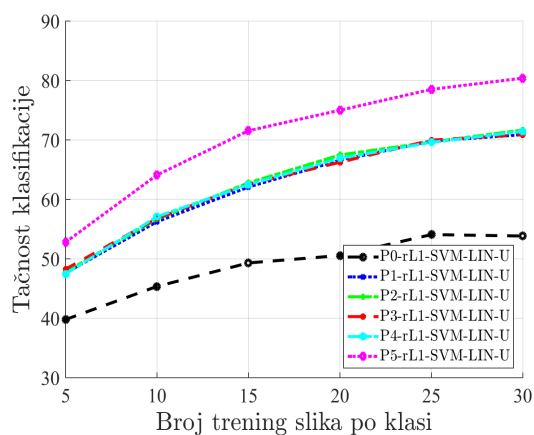
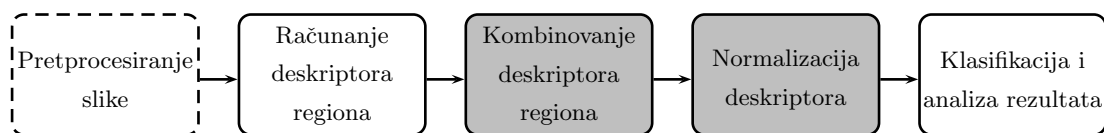
Pošto se broj regiona za koje se računa odgovarajući deskriptor teksture, srazmjerno povećava povećanjem broja dostupnih kanala, uticaj načina normalizacije će biti ispitan na bazama sivih slika, UIUC i UMD. Analizirajući rezultate posljednjeg eksperimenta iz prethodnog poglavlja, vidi se da povećanje količine podataka može popraviti tačnost klasifikacije, međutim uzrokuje znatno sporiju klasifikaciju. Naredni eksperimenti treba da pokažu da li je i u kojoj mjeri moguće popraviti rezultate klasifikacije korišćenjem podataka o prostornom rasporedu lokalnih obilježja, dok će razvoj niskodimenzionalnih prostorno-teksturalnih deskriptora biti predstavljen u narednom poglavlju.

Da bi se obezbijedilo korektno poređenje rezultata klasifikacije, koristi se rL1 normalizacija u oba slučaja. Konkretno, ukoliko se koristi SPN algoritam, histogram svakog pojedinačnog regiona se normalizuje prije spajanja, a ukoliko se koristi SPR algoritam, normalizuje se konačni prostorno-teksturalni deskriptor. Eksperimenti su ponovljeni za **Pi** i **Pi+C** načine podjele prostora slike. Za različit broj trening slika, korišćena je ista trening/test podjela, kao i kod eksperimenta iz Poglavlja 4. Osim normalizacije, koristi se uniformno skaliranje konačnog prostorno-teksturalnog deskriptora. Slično kao i u prethodnim eksperimentima koristi se SVM klasifikator sa linearnim kernelom, međutim ovaj put je regularizacioni parametar postavljen na vrijednost $C = 2^{-5}$. Manja vrijednost regularizacionog parametra se koristi kada se smanjuje forsiranje linearne separabilnosti u korist postizanja veće margine između regiona koji razdvajaju klase. Strožiji uslovi se koriste kako bi se bolje naglasila razlika između predloženih pristupa.

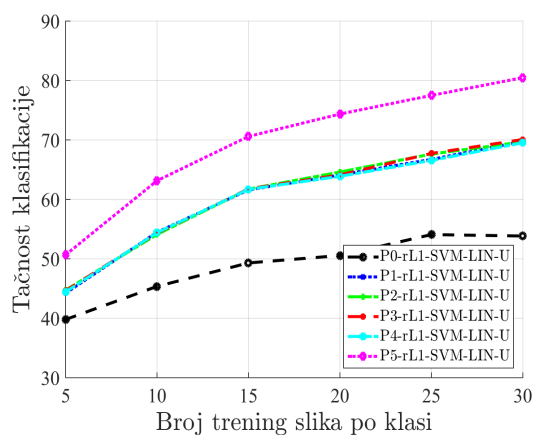
¹*SPatial partitioning with Normalization* (eng.)

²*SPatial partitioning with equal Regions* (eng.)

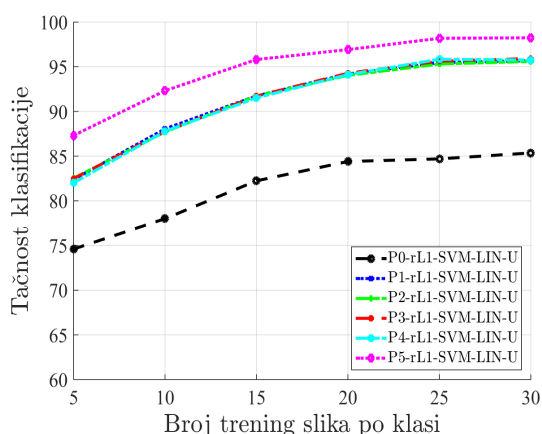
6.3. Klasifikacija prostorno-teksturalnim deskriptorima



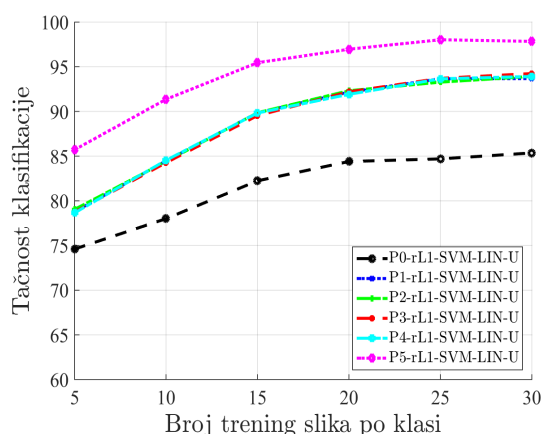
(a) UIUC baza, SIFT deskriptor, SPR metod



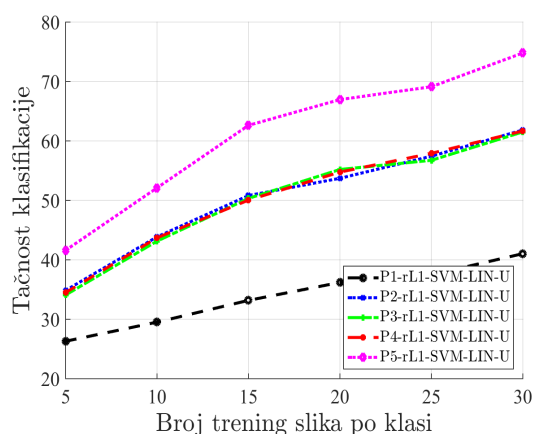
(b) UIUC baza, SIFT deskriptor, SPN metod



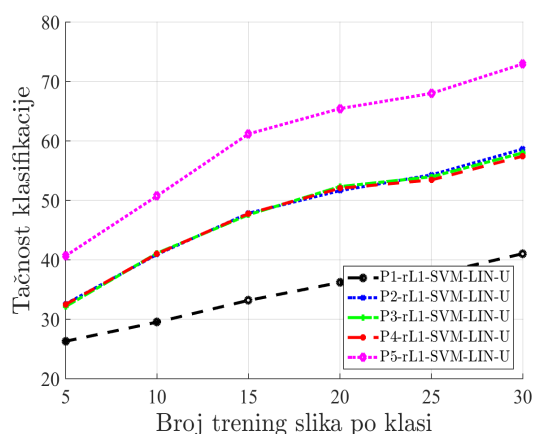
(c) UMD baza, SIFT deskriptor, SPR metod



(d) UMD baza, SIFT deskriptor, SPN metod

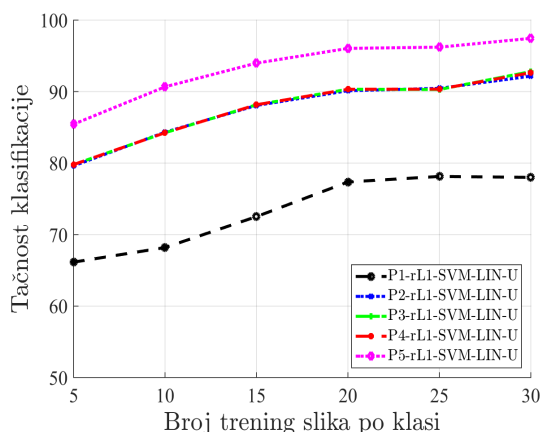


(e) UIUC baza, CENTRIST deskriptor, SPR metod

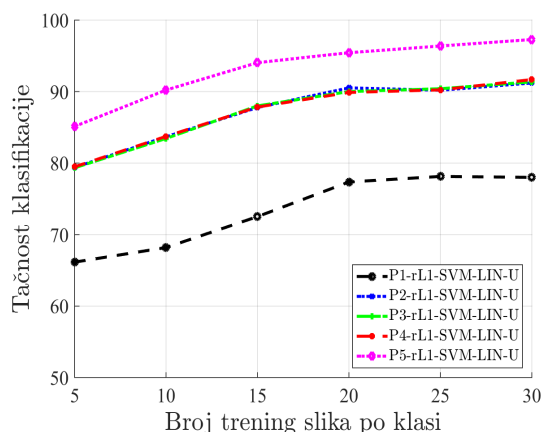


(f) UIUC baza, CENTRIST deskriptor, SPN metod

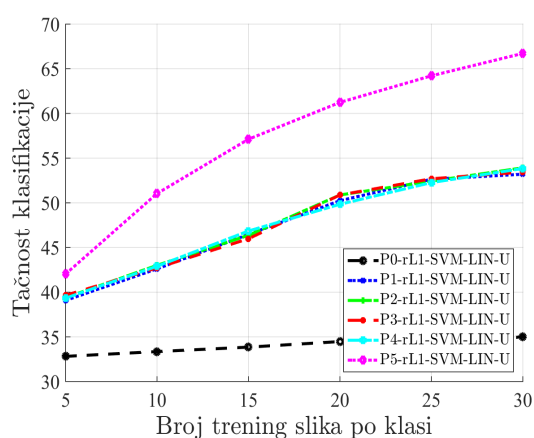
6.3. Klasifikacija prostorno-teksturalnim deskriptorima



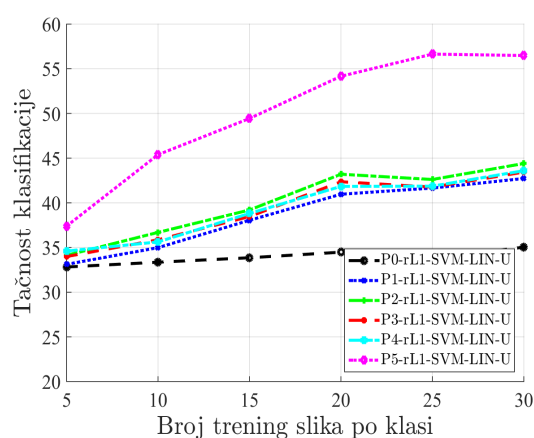
(g) UMD baza, CENTRIST deskriptor, SPR metod



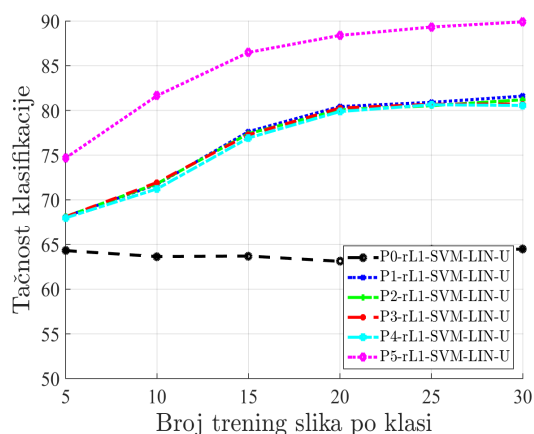
(h) UIUC baza, CENTRIST deskriptor, SPN metod



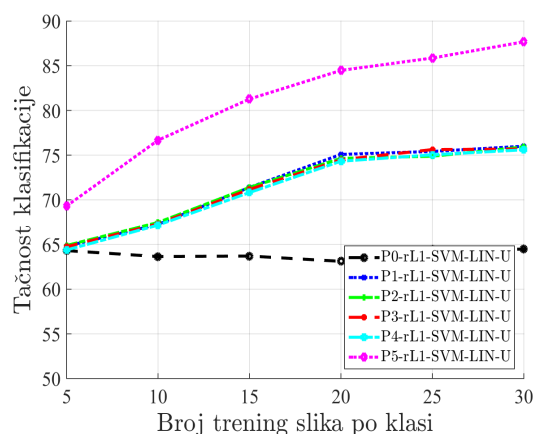
(i) UIUC baza, LBP deskriptor, SPR metod



(j) UMD baza, LBP deskriptor, SPN metod



(k) UMD baza, LBP deskriptor, SPR metod

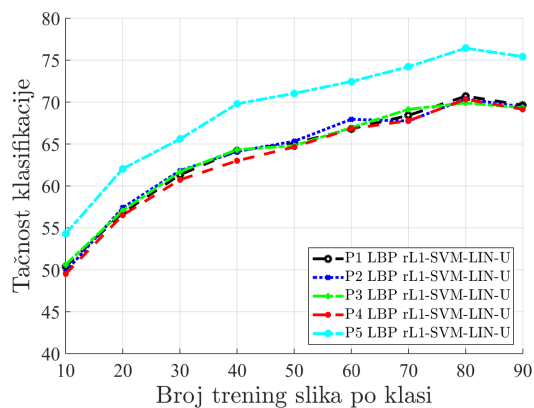
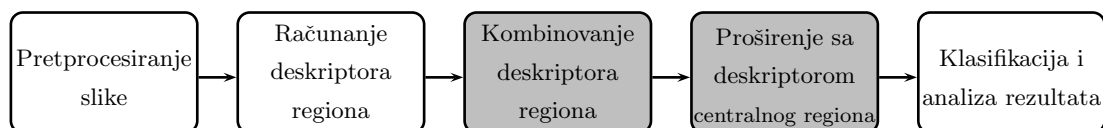


(l) UMD baza, LBP deskriptor, SPN metod

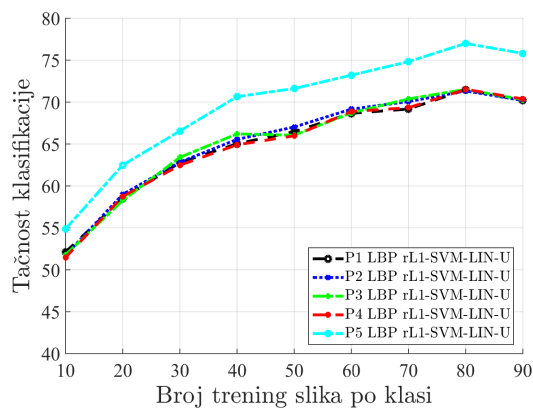
Slika 6.4: Analiza uticaja načina normalizacije prostorno-teksturalnih deskriptora sivih slika na tačnost klasifikacije različite deskriptore i baze sivih slika

Kao što je naglašeno, dimenzionalnost prostorno-teksturalnog deskriptora zavisi od korišćenog broja regiona za koje se računaju dodatni histogrami. Na primjer, za prostorne podjele $\mathbf{P}_i + \mathbf{C}$, $i = 1 \dots 4$ kreira se ukupno pet histograma, dok se za \mathbf{P}_5 podjelu kreira ukupno 17 histograma. Dodatni histogram se kreira ukoliko se

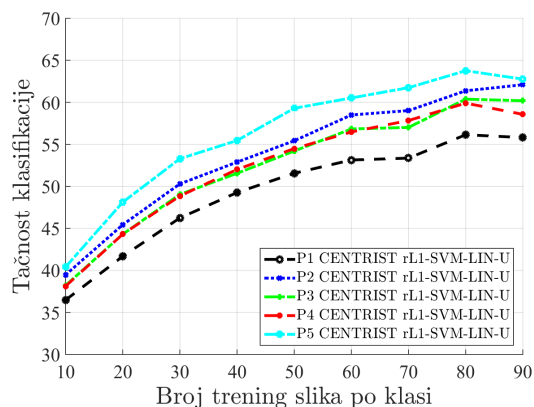
6.3. Klasifikacija prostorno-teksturalnim deskriptorima



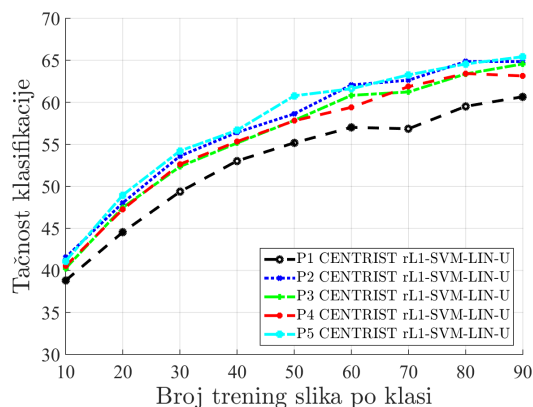
(a) UCM baza, SIFT deskriptor, prostorna podjela bez centralnog regiona



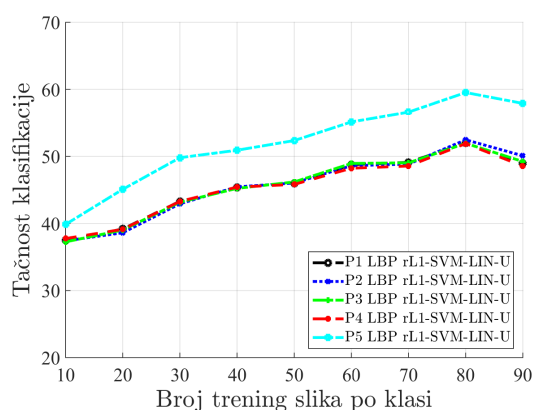
(b) UCM baza, SIFT deskriptor, prostorna podjela sa centralnim regionom



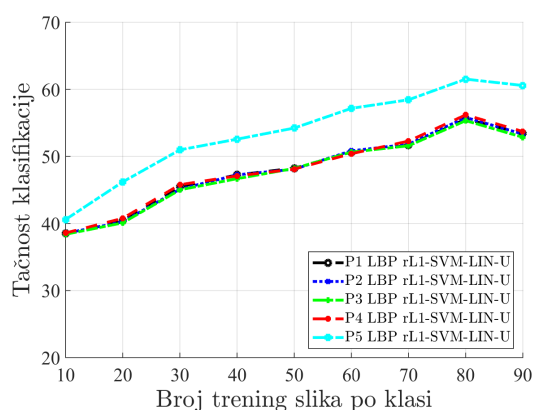
(c) UCM baza, CENTRIST deskriptor, prostorna podjela bez centralnog regiona



(d) UCM baza, CENTRIST deskriptor, prostorna podjela sa centralnim regionom



(e) UCM baza, LBP deskriptor, prostorna podjela bez centralnog regiona



(f) UCM baza, LBP deskriptor, prostorna podjela sa centralnim regionom

Slika 6.5: Uticaj dodatnog histograma centralnog regiona kod prostorno-teksturalnih deskriptora sivih slika na tačnost klasifikacije za UCM bazu sivih slika

koristi dodatni centralni region. Da bi se dobila ukupna dimenzionalnost prostorno-teksturalnog deskriptora, broj regiona se množi sa dimenzionalnošću korišćenog lokalnog deskriptora (sa 254 za CENTRIST i 54 za LBP). U narednim eksperimentima, koristi se BoVW kreiran na osnovu 200 kodnih riječi, pa je dimenzionalnost osnovnog SIFT deskriptora jednaka 200.

Rezultati eksperimenata (Sl. 6.4) pokazuju konzistentno povećanje tačnosti klasifikacije sa proširenjem deskriptora dodatnim histogramima, bez obzira na izbor deskriptora ili bazu slika koja se klasifikuje. Može se primjetiti da varijante prostorno-teksturalnog deskriptora sa podjelama različite orijentacije imaju slične performanse, dok uključivanje svih 16 različitih regiona značajno povećava tačnost klasifikacije. Invarijantnost na ugao rotacije prostorne podjele se lako može objasniti pogledom na vizuelni sadržaj slika koje se klasifikuju.

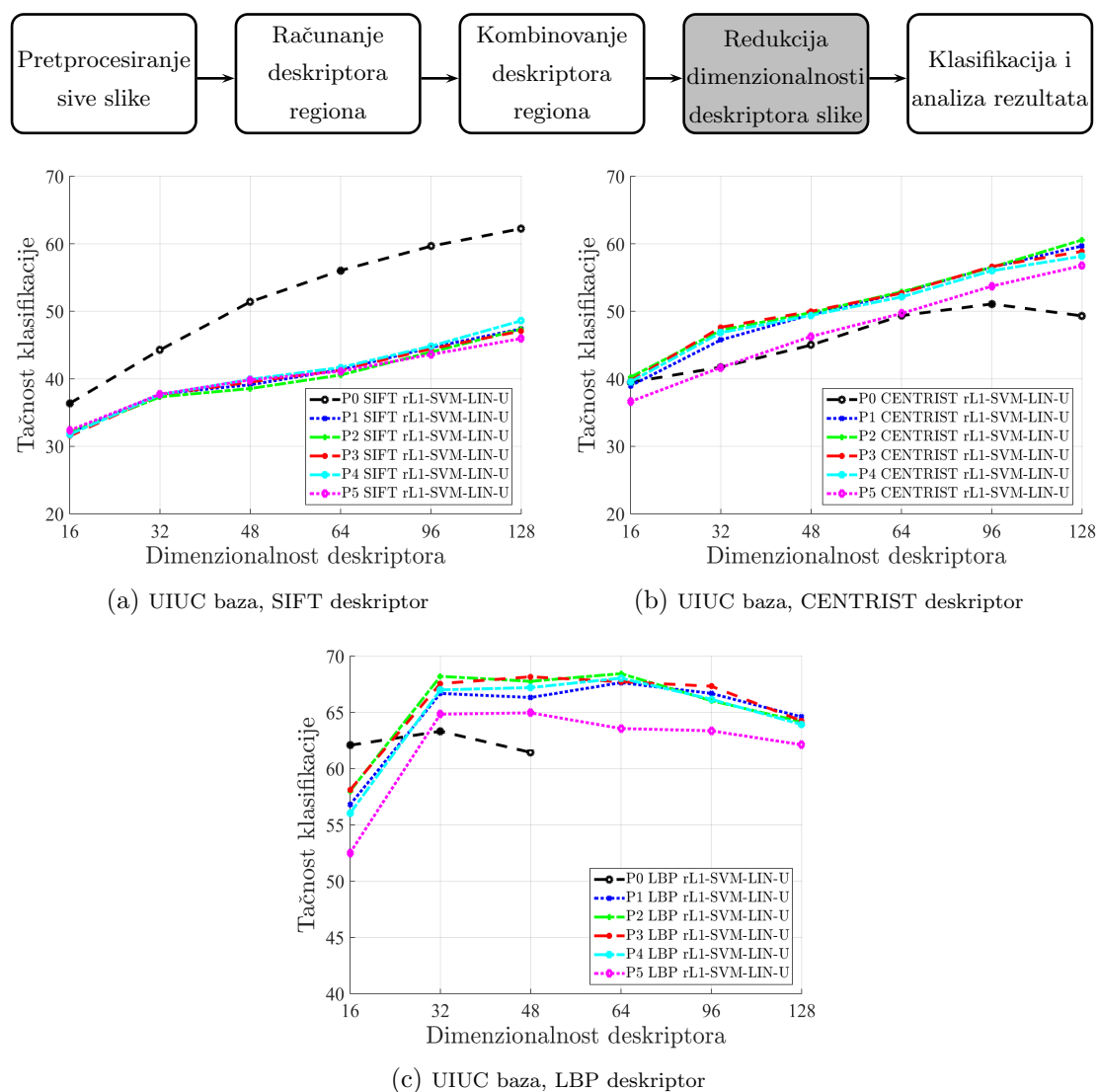
Iz primjera koji su dati u Poglavlju 3 vidi se da je raspored tekture sa velikim stepenom ujednačenosti po prostoru slika. Bez obzira na to, povećanjem broja histograma, povećava se tačnost klasifikacije. Dakle, prostorno-teksturalni deskriptori najveće dimenzionalnosti nose najveću količinu informacija i imaju najbolje rezultate klasifikacije na bazama sivih slika. Slični rezultati su dobijeni i za baze slika u boji u [14]. Algoritam SPR u kojem se normalizacija deskriptora izvršava nakon spajanja histograma svih regiona, pokazuje nešto bolje rezultate klasifikacije u poređenju sa SPN algoritmom. Može se zaključiti da se normalizacijom svakog histograma pojedinačno gubio dio diskriminativnih informacija, pa će se u narednim eksperimentima koristiti isključivo SPR algoritam.

Razmatranje uticaja dodatnog histograma računatog prema centralnom regionu je izvršeno u eksperimentu sa varijantama sivih slika iz baze UCM, jer pojedine klase iz ove baze sadrže objekte na centralnom dijelu slike. Sve slike su smanjenih dimenzija, a deskriptori su normalizovani i uniformno skalirani. Rezultati klasifikacije dati su na Sl. 6.5, gdje su prvoj koloni date zavisnosti tačnosti klasifikacije od broj trening slika iz klase, u slučaju kada se ne koristi centralni region, dok su u drugoj koloni slučajevi kada se koristi centralni region. Može se primjetiti poboljšanje tačnosti klasifikacije za nekoliko procenata kada se uključi histogram centralnog regiona. Takođe, analizom rezultata može se primjetiti da su prostorno-teksturalni deskriptori bazirani na SIFT-u i LBP-u rotaciono invarijantni, što je bilo očekivano.

6.4. Redukcija dimenzionalnosti prostorno-teksturalnih deskriptora

Iz rezultata prethodnih eksperimenata može se primjetiti korist od dodavanja histograma, ali uz cijenu značajno povišene dimenzionalnosti prostorno-teksturalnih deskriptora. Naredni eksperimenti treba da pokažu u kojoj mjeri je moguće smanjiti dimenzionalnost prostorno-teksturalnih deskriptora sivih slika, a pri tome sačuvati ili minimalno smanjiti tačnost klasifikacije kao u slučaju pune dimenzionalnosti.

Eksperimenti su postavljeni tako da se poredi tačnost klasifikacije prostorno-teksturalnih deskriptora, izračunatih prema SPR algoritmu, baziranih na SIFT-u, CENTRIST-u i LBP-u. Slično kao prilikom smanjenja dimenzionalnosti deskriptora sivih slika, kao u Poglavlju 4, vrši se skaliranje dimenzija slika na pola od njihovih



Slika 6.6: Poređenje tačnosti klasifikacije za prostorno-teksturalne deskriptore sivih slika redukovane dimenzionalnosti

originalnih dimenzija, kao i postprocesiranje deskriptora, tj. rL1 normalizacija i uniformno skaliranje. Takođe, izvršeno je 5 klasifikacija prema različitim podjelama na trening i test slike. Po 20 slika iz klase je izabrano za trening, a ostale su bile testne slike. Korišćene su iste podjele na trening/test slike, kao i u eksperimentima iz Poglavlja 4. Regularizacioni parametar linearnog SVM klasifikator je postavljen na $C = 2^{-5}$, kao u prethodnim eksperimentima sa prostorno-teksturalnim deskriptorima. U ovom eksperimentu, korišćen je SIFT deskriptor izračunat prema kodnoj knjizi sa 200 kodnih riječi. Rezultati klasifikacije prikazani su na Sl. 6.6.

Rezultati eksperimenata prikazani su na Sl. 6.6. Mogu se primjetiti različite performanse za prostorno-teksturalne deskriptore bazirane na SIFT-u, CENTRIST-u i LBP-u. Kao što vidimo sa Sl. 6.6a, osnovna varijanta SIFT deskriptora za sive slike ima znatno bolje performanse nego prostorno-teksturalna varijanta. Razlog možemo naći u tome da redukcija dimenzionalnosti odbacuje mnogo više podataka iz prostorno-teksturalnog deskriptora čija dimenzionalnost prije redukcije iznosi 1000

ili 3400 (u zavisnosti od korišćene podjele) u odnosu na obični deskriptor sive slike čija je dimenzionalnost prije redukcije jednaka 200. Interesantno je primjetiti da svaka prostorno-teksturalna varijanta SIFT deskriptora ima slične performanse nakon redukcije dimenzionalnosti.

Kada se pogledaju rezultati za prostorno-teksturalne CENTRIST deskriptore (Sl. 6.6b), vidimo da nakon redukcije dimenzionalnosti oni zadržavaju bolje performanse u odnosu na obični CENTRIST deskriptor sivih slika. Međutim, nakon redukcije dimenzionalnosti prostorno-teksturalnog deskriptora baziranog na **P5** podjeli prostora slike, dolazi do smanjenja tačnosti klasifikacije. Pošto je njegova originalna dimenzionalnost jednaka 4318, može se zaključiti da je značajno umanjeње dimenzionalnosti razlog za lošije performanse. Poredeći sa prostorno-teksturalnim SIFT deskriptorom dobijaju se nešto veće tačnosti klasifikacije. Bolji rezultati klasifikacije prostorno-teksturalnog CENTRIST deskriptora se mogu objasniti činjenicom da se Census transformacija računa u svakom pikselu slike, dok se u ovim eksperimentima SIFT transformacija računa u svakom drugom, što uzrokuje nešto lošije modelovanje prostorno-teksturalnih karakteristika slike.

Posmatrajući rezultate prostorno-teksturalnog LBP deskriptora redukovane dimenzionalnosti, vidimo da su tačnosti klasifikacije bolje u poređenju sa osnovnim LBP deskriptorom sive slike. Varijanta bazirana na prostornoj podjeli **P5** opet ima nešto lošije performanse. Prije redukcije, dimenzionalnost prostorno-teksturalnog LBP deskriptora je 270, odnosno 918, pa se redukcijom uspijevaju sačuvati i naglasiti podaci koji opisuju razliku između slika iz različitih klasa.

Pregledom svih rezultata klasifikacije za prostorno-teksturalne deskriptore može se zaključiti da povećanje dimenzionalnosti deskriptora sivih slika doprinosi povećanju tačnosti klasifikacije, ali da se nakon redukcije dimenzionalnosti mogu sačuvati i eventualno poboljšati performanse samo ukoliko dimenzionalnost prostorno-teksturalnog deskriptora nije prevelika.

6.5. Komparativna analiza rezultata

Prostorno-teksturalni deskriptori predstavljaju proširenje deskriptora sivih slika, tako da se na pogodan način uzmu u obzir i informacije o prostornom rasporedu lokalnih obilježja. S obzirom da imaju povećanu dimenzionalnost u poređenju sa deskriptorima sivih slika, interesantno je porediti tačnost klasifikacije u slučajevima kada se koriste ova dva pristupa za klasifikaciju sivih slika. Dimenzionalnost histograma za svaki pojedinačni region, tj. dimenzionalnost deskriptora sive slike bez prostorne podjele, jeste 200 za SIFT, 254 za CENTRIST i 54 za LBP varijante. U svim eksperimentima koriste se iste trening/test podjele slika, kao u prethodnim eksperimentima. Broj testnih slika po klasi za UIUC bazu je 20, a za UCM bazu je 80. U ovim eksperimentima koristi se linearni SVM klasifikator sa regularizacionim parametrom $C = 2^{-5}$. Rezultati su dati u tabeli 6.1.

Rezultati klasifikacije za obje baze, u slučaju kada se ne vrši redukcija dimenzionalnosti, pokazuju konzistentno povećanje tačnosti klasifikacije sa proširivanjem prostorno-teksturalnog deskriptora dodatnim histogramima regiona. Na ovaj način, može se postići povećanje tačnosti klasifikacije preko 15%. Posmatrajući rezultate

6.5. Komparativna analiza rezultata

Tabela 6.1: Poređenje tačnosti klasifikacije sivih slika u slučajevima kada se koriste obični deskriptori sivih slika i prostorno-teksturalni deskriptori sivih slika. Brojevi predstavljaju tačnosti klasifikacije, dok su vrijednosti u zagradama dimenzionalnosti deskriptora.

UCM baza			
	SIFT	CENTRIST	LBP
50% + LIN-U	57.05 (200)	33.71 (254)	39.28 (54)
50% + rL1 + LIN-U	64.12 (200)	32.09 (254)	38.33 (54)
P1 50% + rL1 + LIN-U	70.81 (1000)	56.14 (1270)	52.00 (270)
P1+C 50% + rL1 + LIN-U	72.57 (1200)	59.52 (1524)	55.57 (324)
P5 50% + rL1 + LIN-U	75.57 (3400)	63.76 (4318)	59.52 (918)
P5+C 50% + rL1 + LIN-U	76.71 (3600)	64.57 (4572)	61.52 (972)
50% + LIN-U	51.52 (128)	36.38 (128)	42.95 (48)
50% + rL1 + LIN-U	62.38 (128)	38.23 (128)	47.38 (48)
P1 50% + rL1 + LIN-U	63.85 (256)	46.38 (384)	49.47 (96)
P1+C 50% + rL1 + LIN-U	66.28 (512)	51.95 (512)	55.04 (128)
P5 50% + rL1 + LIN-U	66.09 (512)	45.47 (1024)	47.14 (256)
P5+C 50% + rL1 + LIN-U	67.33 (512)	48.57 (1024)	49.95 (256)

UIUC baza			
	SIFT	CENTRIST	LBP
50% + LIN-U	46.56 (200)	32.84 (254)	43.92 (54)
50% + rL1 + LIN-U	49.44 (200)	34.28 (254)	44.28 (54)
P1 50% + rL1 + LIN-U	58.24 (1000)	55.48 (1270)	62.00 (270)
P1+C 50% + rL1 + LIN-U	60.08 (1200)	57.52 (1524)	65.72 (324)
P5 50% + rL1 + LIN-U	66.24 (3400)	67.20 (4318)	73.72 (918)
P5+C 50% + rL1 + LIN-U	66.48 (3600)	68.36 (4572)	75.16 (972)
50% + LIN-U	57.04 (128)	49.88 (128)	62.88 (48)
50% + rL1 + LIN-U	62.24 (128)	49.32 (128)	61.44 (48)
P1 50% + rL1 + LIN-U	55.60 (256)	62.40 (384)	66.68 (96)
P1+C 50% + rL1 + LIN-U	55.56 (256)	64.40 (384)	52.00 (256)
P5 50% + rL1 + LIN-U	55.48 (256)	64.92 (384)	53.20 (256)
P5+C 50% + rL1 + LIN-U	55.16 (256)	66.00 (384)	54.56 (256)

klasifikacije nakon izvršene redukcije dimenzionalnosti, mogu se dati dva zanimljiva zapažanja. Ukoliko deskriptor sive slike nije prevelike dimenzionalnosti (otprilike preko 1000), redukcijom dimenzionalnosti mogu se popraviti tačnosti klasifikacije. Takvi slučajevi su posebno naznačeni u tabeli 6.1 boldovanim brojevima. U suprotnom, ukoliko je početna dimenzionalnost prostorno-teksturalnog deskriptora veoma velika (nekoliko hiljada), nakon redukcije dimenzionalnosti mogu se očekivati značajno lošiji rezultati klasifikacije. Na osnovu izloženih rezultata može se preporučiti da prostorno-teksturalni deskriptori sivih slika sa manjim brojem regiona, što je pogodno za klasifikaciju slika sa manjim brojem objekata ili semantički

različitih površi. Slike koje prikazuju veći broj površi sa sadržajem različite semantike bi zahtjevale veći broj regiona i znatno veću dimenzionalnost prostorno-teksturalnih deskriptora.

6.6. Zaključci o klasifikaciji prostorno-teksturalnim deskriptorima

U ovom poglavlju je predstavljeno i analizirano rješenje koje dopunjuje deskriptore sivih slika sa podacima o prostornom rasporedu lokalnih obilježja. Predloženo rješenje ne zahtjeva dodatno računanje lokalnih obilježja, već proširuje deskriptor sive slike novim histogramima lokalnih obilježja, računatim prema različitim regionima. Analizom rezultata i uticaja različitih parametara predloženog prostorno-teksturalnog deskriptora, mogu se izvesti sljedeći zaključci:

- (1) Podaci o prostornom rasporedu imaju veliki uticaj na tačnost automatske klasifikacije slika, naročito u slučajevima kada se objekti sličnih karakteristika mogu očekivati na sličnim mjestima u slikama iz iste klase. Eksperimenti su pokazali da korišćenje prostorno-teksturalnih deskriptora sivih slika visokih dimenzija, povećava tačnost klasifikacije u poređenju sa običnim teksturalnim deskriptorima sivih slika.
- (2) Povećanje tačnosti klasifikacije može se postići proširivanjem prostorno-teksturalnog deskriptora sa histogramom lokalnih obilježja centralnog regiona, u zavisnosti od vizuelnog sadržaja slika koje se klasifikuju. Ukoliko postoje klase sa usamljenim objektima na centralnom dijelu slike, korisno je upotrebiti histogram tog regiona. Ovaj pristup može biti od koristi prilikom prepoznavanja objekata i/ili događaja.
- (3) Prilikom redukcije dimenzionalnosti prostorno-teksturalnih deskriptora sivih slika, rezultati klasifikacije se mogu povećati ako je originalna dimenzionalnost deskriptora relativno niska. Dimenzionalnost koja se može smatrati relativno niskom zavisi od izbora osnovnog deskriptora sivih slika. Ukoliko je originalna dimenzionalnost prostorno-teksturalnih deskriptora veoma velika, korišćenje deskriptora redukovane dimenzionalnosti smanjuje tačnost klasifikacije.
- (4) Na osnovu prethodnog zaključka, može se preporučiti redukcija dimenzionalnosti na nivou regiona prije nego redukcija dimenzionalnosti cijelog prostorno-teksturalnog deskriptora.

7 Niskodimenzionalni prostorno-teksturalni deskriptori

U prethodnim poglavljima detaljno su razmotreni uticaj različitih koraka i vrijednosti parametara prilikom automatske klasifikacije sivih slika, slika u boji i multispektralnih slika. Zaključci izvedeni iz rezultata klasifikacije sivih slika su iskorišćeni prilikom obrade pojedinačnih kanala slika u boji i multispektralnih slika. Takođe, analiziran je uticaj ubacivanja podataka o prostornom rasporedu lokalnih obilježja, na tačnost klasifikacije.

Pošto su u Poglavljima 4, 5 i 6 analizirana predložena rješenja za izdvajanje niskodimenzionalnih deskriptora tekstone za sive i multispektralne slike i za izdvajanje prostorno-teksturalnih deskriptora sivih slika, slijedi analiza načina za kombinovanje prostorno-teksturalnih deskriptora iz pojedinačnih kanala multispektralnih slika. Ranije se moglo zaključiti da smanjenje dimenzionalnosti deskriptora sivih slika može značajno uticati na performanse klasifikacije, a takođe, da kombinovanje podataka iz različitih spektralnih opsega i podešavanje koraka i parametara klasifikacije može nadomjestiti gubitak podataka usljed smanjenja dimenzionalnosti. Potrebno je naglasiti i da deskriptori niske dimenzionalnosti obezbjeđuju brže obučavanje klasifikatora.

7.1. Predloženo rješenje

Kao što je već pomenuto, u slučaju kada se koriste prostorno-teksturalni deskriptori za multispektralne slike, potrebno je posebno analizirati sadržaj svakog od kanala. Spajanjem odgovarajućih prostorno-teksturalnih deskriptora sivih slika, dobija se konačni prostorno-teksturalni deskriptor multispektralne slike. U Poglavlju 5, predložena su dva načina kombinovanja podataka iz različitih kanala multispektralnih slika i tehnike za redukciju dimenzionalnosti, koja su ranije predložena u [13].

Takođe, važno je uzeti u obzir i zaključak izveden na osnovu rezultata eksperimenta kada se vrši redukcija dimenzionalnosti prostorno-teksturalnih deskriptora: tačnost klasifikacije je veća ukoliko se koriste prostorno-teksturalni deskriptori niže dimenzionalnosti.

U literaturi je opisan veliki broj algoritama za izdvajanje deskriptora multispektralnih slika [79, 78, 27, 29, 54, 50, 53, 11, 85]. Međutim, samo je u [10, 11] korišćena kombinacija redukcije dimenzionalnosti i deskriptora teksture za pojedinačne regione. Naime, CENTRIST deskriptor koji je opisan u [10] koristi histograme izračunatih nakon Census transformacije za 31 pravougaoni preklapajući region na slici. Dimenzionalnost svakog od tih histograma se smanjuje sa 254 na 40, nakon čega se oni spajaju u konačni CENTRIST deskriptor dimenzionalnosti 1240. Za slučaj multispektralnih slika u [11] je izvršeno prilagođenje originalnog CENTRIST deskriptora pri čemu je predloženi mCENTRIST $N - 1$ puta veće dimenzionalnosti od originalnog CENTRIST deskriptora, gdje je N broj dostupnih kanala.

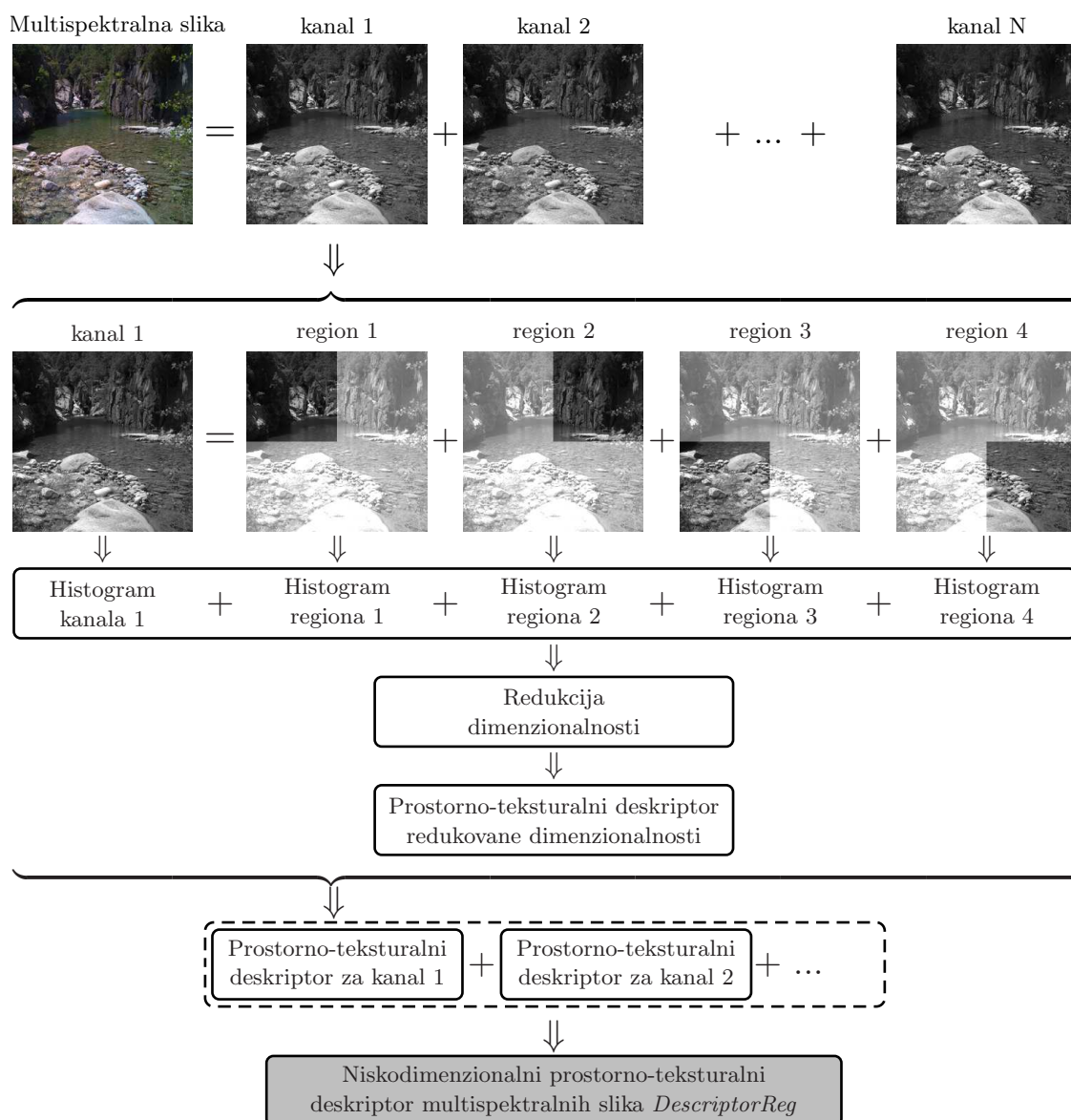
Da bi se ispitalo predloženo rješenje, pretpostavimo da su dostupne multispektralne slike sa četiri kanala i da se koristi prostorna podjela **P1** za svaki od kanala pojedinačno. U ovom slučaju, dimenzionalnost prostorno-teksturalnog deskriptora za jednu multispektralnu sliku je za 20 puta veća od dimenzionalnosti osnovnih deskriptora sivih slika, koji su analizirani u Poglavlju 4. Za redukciju dimenzionalnosti, mogu se koristiti algoritmi DescriptorBand i DescriporConc, koji su detaljno opisani u Poglavlju 5, ali pošto postoji prostorna podjela na regione po svakom od kanala, može se izvesti sljedeći algoritam:

DescriptorReg Korak redukcije dimenzionalnosti može se izvršiti po svakom od regiona posebno, prije kombinovanja deskriptora regiona (tj. odgovarajućih histograma) u deskriptor jednog kanala. Međutim, u tom slučaju je potrebno značajno redukovati dimenzionalnost deskriptora regiona, pri čemu se gubi veći dio informacija. Stoga, predlaže se da se redukcija dimenzionalnosti izvede za cijeli prostorno-teksturalni deskriptor svakog pojedinačnog kanala, nakon čega se vrši spajanje deskriptora različitih kanala. Može se primjetiti da DescriptorReg predstavlja varijantu DescriptorBand algoritma, koja koristi proširene deskriptore pojedinačnih kanala.

S obzirom da predloženi algoritam za računanje prostorno-teksturalnog deskriptora sive slike ne koristi dodatna lokalna obilježja, prilikom kreiranja DescriptorReg varijante bazirane na SIFT-u, nema potrebe da se računaju kodne knjige za svaki od regiona posebno, već se histogrami regiona računaju prema kodnim knjigama za svaki od kanala posebno, kao prilikom računanja MSIFTBAND deskriptora.

7.2. Analiza performansi predloženog deskriptora

Prilikom poređenja performansi niskodimenzionalnih multispektralnih deskriptora, predloženih i opisanih u prethodnim poglavljima, postavljeni su odgovarajući eksperimenti za baze slika u boji i multispektralnih slika. Koriste se iste trening/test podjele slika iz svake klase, kao u prethodnim eksperimentima, pri čemu se koristi 80 trening slika po klasi za bazu UCM i 40 trening slika po klasi za bazu NirScene. Za 5



Slika 7.1: Šematski prikaz algoritma za računanje deskriptora tekstre za multispektralne slike (*DescriptorReg*)

trening/test podjela računaju se tačnosti klasifikacije i konačna tačnost klasifikacije se računa kao srednja vrijednost tačnosti za tih pet slučajeva.

Prilikom poređenja performansi tri različita pristupa za kreiranje niskodimenzionalnih multispektralnih deskriptora, *DescriptorConc*, *DescriptorBand* i *DescriptorReg*, parametre je potrebno postaviti tako da se za sva tri slučaja dobiju konačni deskriptori jednake dimenzionalnosti. *DescriptorBand* i *DescriptorReg* algoritmi na isti način vrše redukciju dimenzionalnosti deskriptora pojedinačnih kanala, sa tom razlikom što je stepen redukcije kod *DescriptorReg* algoritma znatno veći, jer on koristi proširene prostorno-teksturalne deskriptore. Na primjer, ukoliko se redukcija dimenzionalnosti izvrši tako da je nova dimenzionalnost deskriptora jednog kanala jednaka 64, *DescriptorConc* algoritam treba izvršiti tako je konačna dimenzionalnost

deskriptora jednaka 192 za slike u boji, odnosno 256 za četvorokanalne multispektralne slike. Slično kao u eksperimentima iz prethodnih poglavlja, klasifikacija za tri predložena algoritma izvršena je za slučaj kada se veličina slike smanji za pola i kada se koristi dekorelacija kolor prostora. Takođe, nad konačnim deskriptorima izvršena je rL1 normalizacija i uniformno skaliranje. Klasifikacija je izvršena SVM klasifikatorom sa linearnim kernelom i regularizacionim parametrom $C = 1$, a redukcija je izvršena tako da konačni deskriptori budu niske dimenzionalnosti. Broj vizuelnih riječi u kodnoj knjizi MSIFTREG deskriptora jednaka je 200, dok je za MSIFTBAND i MSIFTCONC jednaka 1000. Rezultati eksperimenta su dati na Sl. 7.2.

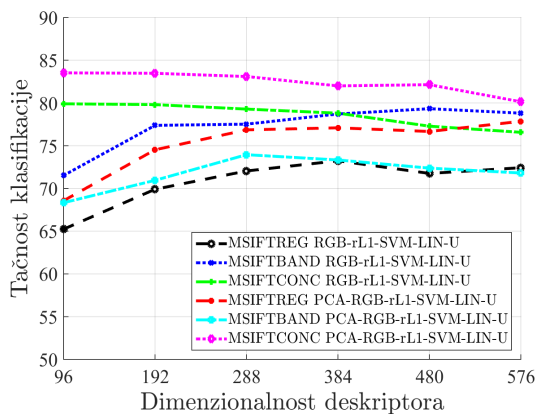
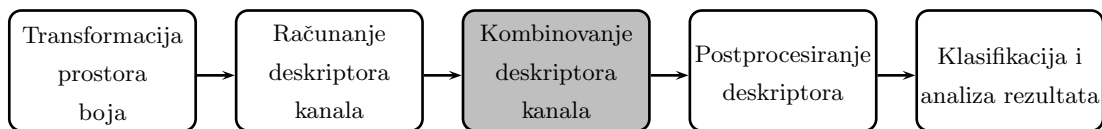
Pregledom rezultata klasifikacije, može se primjetiti da različiti algoritmi za kreiranje deskriptora multispektralnih slika daju najveće tačnosti klasifikacije koristeći različite deskriptore teksture. Multispektralni SIFT deskriptor MSIFTCONC daje znatno bolje rezultate od MSIFTBAND i MSIFTREG varijante, pogotovo za veoma niske dimenzionalnosti. Što se tiče multispektralnih deskriptora baziranih na CENTRIST-u, algoritam deskriptor CENTRISTREG daje najbolje rezultate klasifikacije (Sl. 7.2c i 7.2d). Takođe, može se primjetiti da njegove performanse mnogo manje zavise od ukupne dimenzionalnosti, u poređenju sa CENTRISTCONC i CENTRISTBAND pristupom, kod kojih tačnost klasifikacije naglo opada za nešto veće dimenzionalnosti. Za bilo koju dimenzionalnost ili bazu slika, CENTRISTREG ima veću tačnost klasifikacije i preko 10%. Za multispektralne deskriptore bazirane na LBP-u (Sl. 7.2e i 7.2f), za izrazito niske dimenzionalnosti, najbolje performanse pokazuje LBPCONC deskriptor, dok za nešto veće dimenzionalnosti LBPREG daje najveće tačnosti klasifikacije.

7.3. Komparativna analiza rezultata

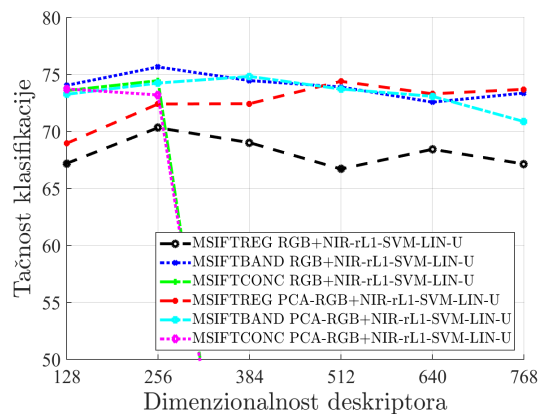
Kako bi se napravilo korektno poređenje tačnosti klasifikacije za različite predložene algoritme formiranja niskodimenzionalnih prostorno-teksturalnih multispektralnih deskriptora, izvršena je klasifikacija za varijante deskriptora različitih dimenzionalnosti. Porede se tačnosti klasifikacije za bazu slika u boji UCM, pri čemu su rezultati dati u tabeli 7.1 i za bazu multispektralnih slika NirScene za koju su rezultati dati u tabeli 7.2. Za obučavanje klasifikatora, prilikom klasifikacije slika iz UCM baze, koristi se 80 slika po klasi, a prilikom klasifikacije slika iz NirScene baze, koristi se 40 slika po klasi. Trening/test podjele slika su iste kao i u svim prethodnim eksperimentima. Koristi se SVM klasifikator sa linearnim kernelom i regularizacionim parametrom $C = 1$.

MSIFT deskriptor koji ne koristi redukciju dimenzionalnosti je računat prema kodnoj knjizi od 500 vizuelnih riječi, kako bi imao približno istu dimenzionalnost kao i predloženi deskriptori multispektralnih slika. Cilj je da se ispita koliko dodatne informacije koje se koriste u predloženim varijantama mogu popraviti tačnost klasifikacije. Ranije je utvrđeno da povećavanjem broja vizuelnih riječi u kodnoj knjizi možemo povećati tačnost klasifikacije po određene vrijednosti, ali u ovom eksperimentu želimo da izbjegnemo potrebu za većim memorijskim resursima i računarsku kompleksnost.

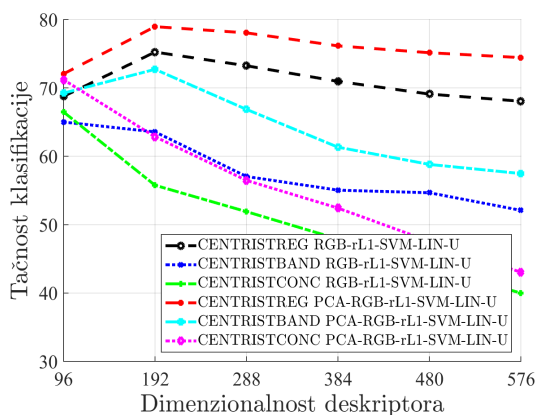
7.3. Komparativna analiza rezultata



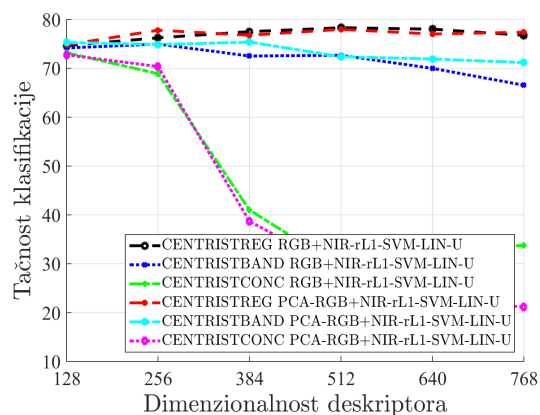
(a) UCM baza, SIFT deskriptor



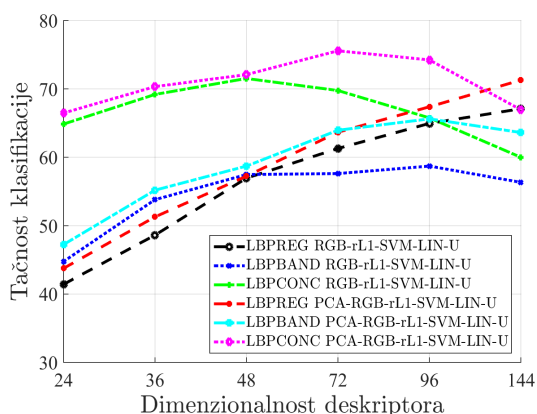
(b) NirScene baza, SIFT deskriptor



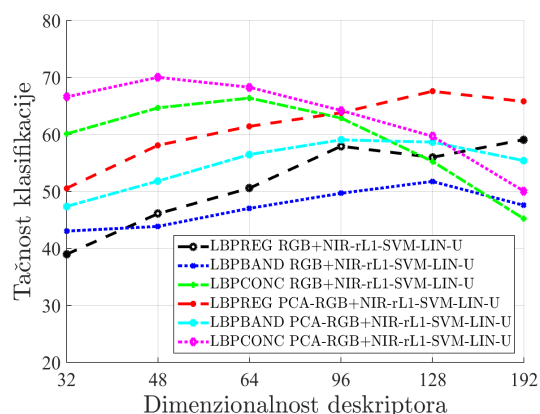
(c) UCM baza, CENTRIST deskriptor



(d) NirScene baza, CENTRIST deskriptor



(e) UCM baza, LBP deskriptor



(f) NirScene baza, LBP deskriptor

Slika 7.2: Poređenje tačnosti klasifikacije različitih niskodimenzionalnih deskriptora multispektralnih slika

7.3. Komparativna analiza rezultata

Tabela 7.1: Poređenje tačnosti klasifikacije slika u boji u slučajevima kada se koriste predloženi niskodimenzionalni prostorno-teksturalni deskriptori. Brojevi predstavljaju tačnosti klasifikacije, a dimenzionalnosti deskriptora su date u zagradama.

UCM baza			
	MSIFT	MCENTRIST	MLBP
100% + LIN	80.52 (500)	73.81 (762)	78.95 (162)
100% + rL1 + LIN-U	84.47 (500)	74.43 (762)	69.10 (162)
PCA 100% + rL1 + LIN-U	79.57 (500)	76.47 (762)	73.76 (162)
P1 100% + rL1 + LIN-U	82.62 (3000)	73.85 (3810)	80.52 (810)
P1+C 100% + rL1 + LIN-U	84.33 (3200)	77.61 (4046)	84.15 (864)
P1+C PCA 100% + rL1 + LIN-U	88.90 (3200)	76.57 (4064)	83.71 (864)
	MSIFTCONC	CENTRISTCONC	LBPCONC
100% + LIN	82.43 (512)	74.28 (512)	78.38 (96)
100% + rL1 + LIN-U	79.67 (512)	44.95 (512)	71.33 (96)
PCA 100% + rL1 + LIN-U	83.81 (512)	50.81 (512)	77.52 (96)
	MSIFTBAND	CENTRISTBAND	LBPBAND
100% + LIN	78.14 (512)	73.76 (512)	60.04 (96)
100% + rL1 + LIN-U	82.04 (512)	56.57 (512)	66.38 (96)
PCA 100% + rL1 + LIN-U	83.04 (512)	62.76 (512)	66.38 (96)
	MSIFTREG	CENTRISTREG	LBPREG
P1 100% + rL1 + LIN-U	76.81 (512)	72.19 (512)	55.47 (96)
P1+C 100% + rL1 + LIN-U	80.38 (512)	74.85 (512)	63.00 (96)
P1+C PCA 100% + rL1 + LIN-U	83.57 (512)	81.52 (512)	67.47 (96)

Rezultati klasifikacije za bazu slika u boji UCM, dati su u tabeli 7.1. Posmatrajući tačnosti klasifikacije za slučajeve multispektralnih deskriptora koji ne koriste redukcije dimenzionalnosti, možemo primjetiti konzistentnu korist od dodatnih podataka i povećane dimenzionalnosti deskriptora, bez obzira na izbor osnovnog deskriptora sive slike. Dalje, vidimo da različiti predloženi metodi za računanje multispektralnih deskriptora mogu popraviti ili pokvariti tačnost klasifikacije u zavisnosti od izbora vrijednosti parametara i osnovnog deskriptora. Slučajevi kada su predloženi niskodimenzionalni deskriptori postigli bolje rezultate klasifikacije u poređenju sa deskriptorima pune dimenzionalnosti su obilježeni boldovanim brojevima.

Deskriptori bazirani na SIFT-u zasigurno gube korisne informacije redukcijom dimenzionalnosti, ali ipak mogu sačuvati tačnost klasifikacije u poređenju sa varijantama koje ne uključuju redukciju. Što se tiče predloženih multispektralnih deskriptora baziranih na CENTRIST-u, ponovo možemo uočiti velike varijacije u odnosu na vrijednosti parametara klasifikacije, ali vidimo da predloženi algoritam CENTRISTREG povećava tačnost klasifikacije u odnosu na varijante koje ne koriste redukciju, iako ima znatno manju dimenzionalnost. Algoritmi bazirani na LBP-u nisu pogodni za korišćenje sa značajno redukovanim dimenzionalnostima, iz razloga

7.3. Komparativna analiza rezultata

Tabela 7.2: Poređenje tačnosti klasifikacije multispektralnih slika u boji u slučajevima kada se koriste predloženi niskodimenzionalni prostorno-teksturalni deskriptori. Brojevi predstavljaju tačnosti klasifikacije, a dimenzionalnosti deskriptora su date u zagradama

NirScene baza			
	MSIFT	MCENTRIST	MLBP
100% + LIN	71.01 (500)	70.17 (1016)	66.77 (216)
100% + rL1 + LIN-U	77.14 (500)	74.21 (1016)	65.67 (216)
PCA 100% + rL1 + LIN-U	78.60 (500)	72.92 (1016)	68.54 (216)
P1 100% + rL1 + LIN-U	76.48 (3000)	76.91 (3810)	67.12 (810)
P1+C 100% + rL1 + LIN-U	76.91 (3200)	76.74 (4064)	69.35 (864)
P1+C PCA 100% + rL1 + LIN-U	79.51 (3200)	76.20 (4064)	74.33 (864)
	MSIFTCONC	CENTRISTCONC	LBPCONC
100% + LIN	71.20 (256)	70.36 (256)	65.09 (96)
100% + rL1 + LIN-U	71.22 (256)	68.98 (256)	71.75 (96)
PCA 100% + rL1 + LIN-U	76.17 (256)	68.51 (256)	71.08 (96)
	MSIFTBAND	CENTRISTBAND	LBPBAND
100% + LIN	73.76 (256)	68.93 (256)	47.41 (96)
100% + rL1 + LIN-U	76.61 (256)	74.17 (256)	52.27 (96)
PCA 100% + rL1 + LIN-U	77.71 (256)	72.96 (256)	63.28 (96)
	MSIFTREG	CENTRISTREG	LBPREG
P1 100% + rL1 + LIN-U	72.12 (256)	77.92 (256)	60.72 (96)
P1+C 100% + rL1 + LIN-U	72.56 (256)	76.31 (256)	58.35 (96)
P1+C PCA 100% + rL1 + LIN-U	77.42 (256)	74.91 (256)	66.44 (96)

što je njihova dimenzionalnost (za uniformni i rotaciono-invarijantni oblik deskriptora) već nešto niža. Ipak, može se primjetiti da LBPCONC varijanta može popraviti tačnost klasifikacije u odnosu na multispektralni LBP deskriptor neredukovane dimenzionalnosti.

Posmatrajući rezultate klasifikacije za bazu multispektralnih slika (tabela 7.2), može se primjetiti da predloženi niskodimenzionalni deskriptori mogu postići približno slične tačnosti klasifikacije u poređenju sa varijantama multispektralnih deskriptora koje ne uključuju redukciju dimenzionalnosti. Treba istaći da predloženi deskriptori imaju i do 15 puta manju dimenzionalnost u odnosu na varijante multispektralnih deskriptora neredukovane dimenzionalnosti, a da se tačnost klasifikacije smanjuje najviše za 2.5%. Za SIFT bazirane multispektralne deskriptore, ponovo možemo primjetiti da gubitaka podataka umanjuje diskriminativna svojstva deskriptora za 1 do 2%. Posmatrajući rezultate deskriptora baziranih na CENTRIST-u, može se primjetiti da predloženi deskriptor CENTRISTREG, zapravo popravljajući tačnost klasifikacije, iako je znatno niže dimenzionalnosti u odnosu na neredukovani varijantu. LBPCONC popravljajući rezultate klasifikacije u odnosu na neredukovani multispektralni LBP, ali multispektralna prostorno-teksturalna varijanta LBP deskriptora ipak ima najbolje performanse.

Bez obzira na to koji osnovni deskriptor se koristi prilikom računanja multispektralnog deskriptora, ili koja baza slika se klasifikuje, može se primjetiti da proširivanje multispektralnog deskriptora dodatnim podacima o prostornom rasporedu lokalnih obilježja, povećava tačnost klasifikacije za nekoliko procenata. U slučaju kada se koristi LBP deskriptor, to povećanje može biti i preko 10%. Međutim, ukoliko je potrebno kreirati niskodimenzionalne multispektralne deskriptore, mora se napraviti kompromis između tačnosti klasifikacije i dimenzionalnosti (odnosno brzine izvršavanja).

7.4. Zaključci o klasifikaciji niskodimenzionalnim prostorno-teksturalnim deskriptorima

U prethodnim eksperimentima izvršeno je poređenje tačnosti klasifikacije za tri predložena algoritma računanja niskodimenzionalnih prostorno-teksturalnim deskriptorima multispektralnih slika. Na osnovu rezultata klasifikacija, izvedenih sa različitim vrijednostima parametara, mogu se izvesti određeni zaključci:

- (1) Prilikom računanja deskriptora multispektralnih slika, ukoliko se vrši kombinovanje tehnika za redukciju dimenzionalnosti i podataka iz različitih spektralnih opsega slika, tačnost klasifikacije zavisi od redosljed izvršavanja koraka. Razmatrana su tri različita algoritma koja računaju deskriptore multispektralnih slika na različite načine (tj. koriste različit redosljed izvršavanja koraka).
- (2) Različiti lokalni deskriptori tekture mogu se iskoristiti za deskripciju sivih slika pojedinačnih kanala ili određenih regiona slike. U zavisnosti od toga koji lokalni deskriptor tekture se koristi kao osnovni, različiti algoritmi za kreiranje multispektralnih slika postižu najbolje rezultate klasifikacije, bez obzira na parametre klasifikatora ili korišćene metode pretprocesiranja i postprocesiranja.
- (3) Proširivanje deskriptora sa podacima iz različitih kanala i podacima o prostornom rasporedu lokalnih obilježja povećava se tačnost klasifikacije slika u boji i multispektralnih slika.
- (4) Eksperimenti su pokazali da se postižu slični ili bolji rezultati klasifikacije kada se koriste niskodimenzionalni prostorno-teksturalni deskriptori u poređenju sa multispektralnim deskriptorima dimenzionalnosti od nekoliko hiljada.
- (5) Za svaki osnovni deskriptor (u zavisnosti od njegove osnovne dimenzionalnosti), moguće je izabrati odgovarajući algoritam za računanje niskodimenzionalnog multispektralnog deskriptora koji će dati najbolje rezultate klasifikacije.

8 Automatska klasifikacija slika konvolucionim neuronskim mrežama

U uvodnim poglavljima ove disertacije, opisani su istraživački trendovi u oblasti automatske klasifikacije slika, detekcije objekata ili događaja. Napomenuto je da ukoliko postoji dovoljno velika baza trening slika, metode zasnovane na konvolucionim neuronskim mrežama daju značajna poboljšanja u pogledu tačnosti klasifikacije u poređenju sa metodama zasnovanim na deskriptorima. Ipak, ukoliko to nije slučaj, nije moguće valjano obučiti i iskoristiti konvolucione neuronske mreže za automatsku klasifikaciju slika. U ovom poglavlju će biti razmotrene mogućnosti korišćenja konvolucione neuronske mreže za automatsku klasifikaciju aero snimaka i detekciju objekata u aero snimcima visoke rezolucije, kao što je predloženo u [16].

Konvolucione neuronske mreže (CNN) sadrže uzastopne slojeve koji konvolucijom detektuju lokalna obilježja. Arhitektura CNN je takva da se objedinjuje detekcija obilježja nižeg nivoa (slično kao lokalni deskriptori koji su razmatrani u ovoj disertaciji), podaci o njihovom prostornom rasporedu i njihova kombinacija u obilježja višeg nivoa [67]. Tokom obučavanja mreže, podešavaju se težinski faktori u svakom konvolucionom sloju, kako bi se izvršilo što tačnije modelovanje podataka iz trening skupa. Na ovaj način moguće je izvršiti izdvajanje deskriptora slika, a dodavanjem dodatnih slojeva u mrežu je moguće uključiti i korak klasifikacije.

Navedene prednosti CNN-a omogućile su postizanje visokih tačnosti klasifikacije u poređenju sa metodama baziranim na deskriptorima [70, 71, 41, 68, 86]. Ipak, kombinacijom deskriptora i konvolucionih neuronskih mreža mogu se dodatno poboljšati rezultati klasifikacije [51], naročito ukoliko ne postoji dovoljno veliki trening skup za obučavanje neuronskih mreža. Još jedna značajna prednost CNN-a jeste mogućnost da se dio mreže obučen na jednoj bazi slika efikasno iskoristi za klasifikaciju druge baze slika, jer su CNN-e sposobne da nauče lokalne karakteristike bez obzira na raznolikost vizuelnog sadržaja slika. U literaturi se najčešće koriste tri različite arhitekture konvolucionih neuronskih mreža, AlexNET [67], koja se sastoji

od 5 konvolucionih slojeva i dodatnih potpuno povezanih slojeva za klasifikaciju, CaffeNET [87] koja ima sličnu arhitekturu kao i AlexNET i GoogleNET [88] koja ima 22 konvoluciona sloja.

Pored navedenih prednosti, postoji i nekoliko nedostataka. Obučavanje konvolucione mreže od početka za određenu bazu podataka je vremenski i računski veoma zahtjevno i kao što je već rečeno, nedostatak trening podataka može značajno uticati na tačnost klasifikacije. U [41] su date tačnosti klasifikacije za UCM bazu koristeći pristup obučavanja mreže od početka, za CaffeNET i GoogleNET mreže. Za isti broj trening slika po klasi, kao što je korišćeno u ovoj disertaciji, prijavljena je tačnost klasifikacije od 85.71% za CaffeNET mrežu i 92.86% za GoogleNET mrežu. Očigledno, veći broj slojeva i “dublje” učenje koje koristi GoogleNET mreža povećava tačnost klasifikacije. Poredeći sa rezultatima koji su dobijeni u prethodnim eksperimentima, gdje se dobija tačnost klasifikacije $\approx 90\%$, može se primjetiti da metode zasnovane na deskriptorima daju slične ili bolje rezultate klasifikacije iz razloga što CNN-e nemaju dovoljno veliku količinu trening slika.

8.1. Predloženo rješenje za obučavanje konvolucione neuronske mreže

Što se tiče kolekcije slika aero snimaka UCM, koja se razmatra u ovoj disertaciji, u većini eksperimenta uzima se 80 trening slika iz svake klase, što ukupno daje 1680 trening slika. U poređenju sa bazama slika koje sadrže i do nekoliko miliona slika visoke rezolucije, ovo predstavlja relativno mali skup trening podataka. Da bi se prevazišao problem nedostatka trening podataka, moguće je iskoristiti mreže pre-trenirane na drugim bazama slika, koje su dovoljno “bogate” lokalnim obilježjima. Kako bi se postigli bolji rezultati klasifikacije, u [16] predloženo je podešavanje GoogleNET arhitekture obučene pomoću ImageNET baze slika [89] u dvije nezavisne faze. U prvoj fazi se mreža obučena na ImageNET bazi prilagođava UCM bazi slika, tj. ranije naučena obilježja prilagođavaju se za upotrebu na bazi slika sa manjim brojem klasa. Druga faza se sastoji od dodatnog podešavanja adaptivnim algoritmom prema određenom setu trening slika iz UCM baze. Rezultat ovog pristupa obučavanju mreže bolje prilagođavanje obilježja naučenih metodom dubokog učenja na mreži sa mnogo više slojeva, na bazu slika UCM. Mreža sa više slojeva kombinuje obilježja nižeg reda, učeći obilježja višeg reda. Pored mogućnosti automatske klasifikacije slika, u [16] je prikazan metod za automatsku detekciju objekata na slikama visoke rezolucije.

Poređenje rezultata klasifikacije predloženog metoda sa rezultatima koji su dati u literaturi, dato je u tabeli 8.1. Pored navedene metode dati su podaci o godini kada je metoda predložena i koja tačnost klasifikacije je priložena. Treba naglasiti da su svi rezultati dobijeni za 80 trening slika po klasi, ali da nisu izvršene iste podjele na trening/test skupove kao u ovoj disertaciji. Interesantno je primjetiti da metode zasnovane na deskriptorima generalno daju slabije tačnosti klasifikacije u poređenju sa konvolucionim neuronskim mrežama. Međutim, povećanjem količine podataka [13], uzimanjem u obzir lokalnih obilježja višeg reda [33] ili boljim modelovanjem [90,

Tabela 8.1: Poređenje tačnosti klasifikacije za UCM bazu aero snimaka, za metode zasnovane na deskriptorima i metode zasnovane na konvolucionim neuronskim mrežama.

deskriptor	godina	tačnost
BoVW + SCK [49]	2010	77.71%
SPCK+ [91]	2011	76.05%
SPCK++ [91]	2011	77.38%
BRSP [57]	2012	77.80%
Unsupervised feature learning [65]	2014	81.67%
mCENTRIST [11]	2014	89.90%
MSIFT [13]	2014	90.81%
VLAD [33]	2014	92.50%
VLAT [33]	2014	94.30%
PSR [90]	2015	89.10%
UFL-SC [85]	2015	90.26%
neuronska mreža	godina	tačnost
GoogleNet (fino podešavanje) [41]	2015	97.10 %
ConvNet [68]	2015	89.39%
IFK+VGG [71]	2015	98.49%
Multiview deep learning [70]	2015	93.48%
Predloženi metod (dvostruko podešavanje) [16]	2015	98.61%

85] mogu se povećati tačnosti klasifikacije. Sa druge strane, ukoliko neuronske mreže nemaju dovoljno veliku količinu trening podataka [68], performanse klasifikacije će biti nešto lošije.

Takođe, mogu se primjetiti veoma visoke tačnosti klasifikacije kada se koriste neuronske mreže koje su prethodno naučene na ImageNET bazi, pa zatim dodatno podešene za UCM bazu [41, 16] odakle se može primjetiti jedna značajna osobina konvolucionih neuronskih mreža. Naime, ImageNET baza slika sadrži preko 15 miliona slika visoke rezolucije, obilježenih i podijeljenih u preko 22000 klasa. CNN-e su sposobne da nauče veliku količinu lokalnih obilježja višeg i nižeg reda iz te baze i da ih na odgovarajući način prilagode drugim bazama slika, kao što je UCM.

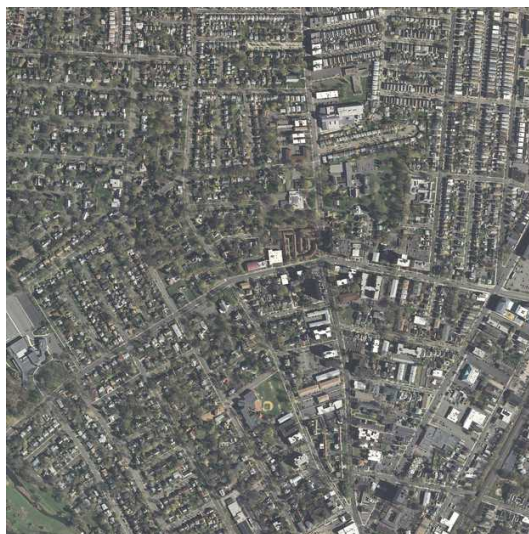
8.2. Automatska detekcija objekata na aero snimcima visoke rezolucije

U [16] je opisana jedna praktična mogućnost korišćenja CNN-a za automatsku detekciju objekata na aero snimcima visoke rezolucije. U ovom eksperimentu korišćeni su aero snimci visoke rezolucije 5000×5000 iz kolekcije USGS nacionalnih mapa Sjedinjenih Američkih Država [49], tj. iz iste kolekcije iz koje je izdvojena baza UCM. Nekoliko primjera aero snimaka dato je na Sl. 8.1.

Snimak prikazan na Sl. 8.1a prikazuje seoski predio sa usamljenim objektima, ob-



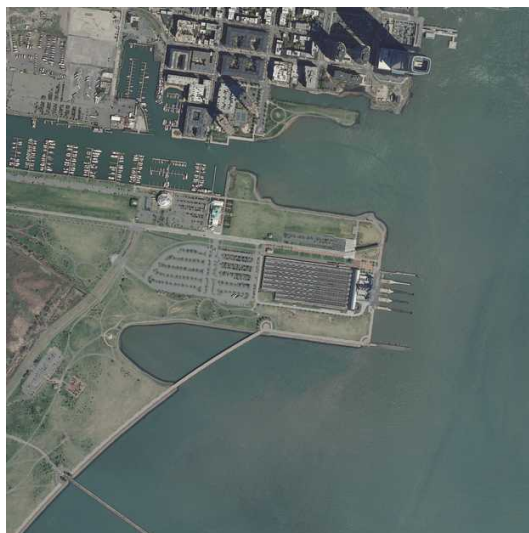
(a) Snimak iz USGS baze koji pokazuje predio obradivih površina sa izolovanim objektima.



(b) Snimak iz USGS baze koji pokazuje gusto naseljenu gradsku oblast



(c) Snimak iz USGS baze koji pokazuje srednje naseljeni seoski predio



(d) Snimak iz USGS baze gdje se vidi luka i dio gradske oblasti sa parkingom

Slika 8.1: Primjeri aero snimaka visoke rezolucije

radivim površinama i šumom. Ovakav snimak može da se iskoristi za procjenu stanja usjeva, nadzor poljoprivrednih resursa i očuvanje šumskih dobara. Pošto su objekti prikazani na slici usamljeni, možemo smatrati da se radi o jednostavnom zadatku detekcije objekata za koji nije potrebna pomoć računarskih programa. Međutim, kada je dostupan veliki broj ovakvih snimaka, procjena broja poljoprivrednih dobara može se uraditi automatskim prebrojavanjem ovakvih objekata. Snimak prikazan na Sl. 8.1b prikazuje gusto naseljenu gradsku oblast, gdje postoje objekti specifične namjene koje je i pažljivim pregledom teško uočiti. Takvi objekti su teniski i bejzbol tereni, a takođe, može se uočiti i određeni broj većih objekata koji nisu stambeni. Ovaj primjer možemo smatrati težim zadatkom za detekciju objekata bilo koje na-



(a) Gornji desni dio aero snimka sa Sl. 8.1a. Na snimku su jasno značeni usamljeni objekti



(b) Donji desni dio aero snimka sa Sl. 8.1b. Označene su veće zgrade i bjezbol igralište



(c) Donji lijevi dio aero snimka sa Sl. 8.1c. Posebno su označeni regioni sa srednjom gustinom naseljenosti, a posebno regioni koji sadrže usamljene objekte



(d) Gornji lijevi dio aero snimka sa Sl. 8.1d. Označeni su brodovi u luci i parking sa popunjem mjestima

Slika 8.2: Rezultati predloženog metoda za automatsku detekciju objekata na slikama visoke rezolucije

mjene. Snimak sa Sl. 8.1c pokazuje seoski pejzaž sa naseljem srednje gustine, gdje možemo uočiti regione srednje i manje gustine. Možemo smatrati da je procjena gustine naseljenosti srednje težak zadatak. Poslednji primjer, prikazan na Sl. 8.1d je snimak koji pokazuje luku, parking mjesta, dio grada i zelene površine.

Eksperiment je postavljen tako da se ispita da li je moguće iskoristiti CNN-u tako da se automatski detektuju objekti određenog semantičkog sadržaja. U tu svrhu, iskorišćena je mreža obučena na UCM bazi koja je upotrebljena prilikom klasifikacije UCM baze, pošto su semantičke klase slika iz UCM definisane upravo

na osnovu USGS snimaka. Detekcija je izvršena tako što se cijeli snimak visoke rezolucije podijelio na preklapajuće regione veličine 128×128 , pri čemu je korak između centara regiona podešen na 128. Za svaki region posmatra se sličnost sa klasama iz UCM baze, pri čemu se praktično izvodi klasifikacija prema predloženoj neuronskoj mreži. Ukoliko je sličnost za jedan region veća od predefinisano praga, taj region se obilježava kao odgovarajuća detekcija. Rezultati automatske detekcije za snimke sa Sl. 8.1 dati su na Sl. 8.2. Da bi se omogućilo bolje uočavanje detalja, prikazani su rezultati detekcije objekata samo za određene dijelove aero snimaka sa Sl. 8.1. Pregledom rezultata, može se primjetiti da su detektovani blokovi semantički slični odgovarajućim klasama iz UCM baze slika. Vizuelnom inspekcijom može se utvrditi da su detektovani svi objekti od interesa. Predloženi pristup, baziran na CNN-a može pronaći objekte na slikama visoke rezolucije ukoliko je na odgovarajući način mreža obučena za njihovo prepoznavanje.

8.3. Zaključci o klasifikaciji konvolucionim neuronskim mrežama

Konvolucione neuronske mreže predstavljaju moćan alat za analizu sadržaja digitalnih slika i ukoliko su obučene na odgovarajući način mogu dati dobre rezultate u raznim praktičnim aplikacijama.

- (1) Kombinovanjem lokalnih obilježja nižeg nivoa, konvolucione neuronske mreže mogu preciznije modelovati vizuelni sadržaj na slikama.
- (2) Njihova arhitektura i način izdvajanja obilježja, inherentno koristi i prostorni raspored lokalnih obilježja.
- (3) Treniranje konvolucionih mreža na relativno malom trening skupu podataka neće dati značajno poboljšanje tačnosti klasifikacije. Međutim, konvolucione neuronske mreže mogu biti pretrenirane na drugoj bazi "bogatoj" obilježjima, i dodatno prilagođene za bazu slika koja se klasifikuje ako bi se omogućila klasifikacija visoke tačnosti. Takođe, CNN-e se mogu koristiti za izdvajanje deskriptora koji služe za obučavanje klasifikatora.

9 | Zaključak

Automatska klasifikacija slika predstavlja važan dio mnogih srodnih inženjerskih oblasti koje objedinjuju praktične aplikacije bazirane na prepoznavanju vizuelnog i semantičkog sadržaja na slikama. Da bi se moglo prepoznati koja slika prikazuje određeni objekat ili koje dvije slike prikazuju istu ili sličnu scenu, potrebno je poznavati određene metode “čitanja” podataka sa slika, koje daju reprezentaciju sadržaja na slikama pogodnu za primjenu u praktičnim aplikacijama. Razumijevanje načina na koji ljudski vizuelni sistem prepoznaje objekte je omogućilo kvalitetniju reprezentaciju vizuelnog i semantičkog sadržaja na slikama. Uočeno je da tekstura ima presudnu ulogu prilikom prepoznavanja scena, tako da je istraživanje bazirano na deskriptorima teksture.

Klasifikacija slika iz određene baze je od velikog praktičnog značaja, jer omogućava razvoj velikog broja korisnih praktičnih alata u industriji, poljoprivredi, nadzoru i sigurnosti, medicini, itd. Da bi se omogućio razvoj praktičnih aplikacija, treba obezbijediti načine na koje se iz slika izdvajaju podaci koji su potrebni. Klasifikacija slika nam služi da razvijemo jednostavnije aplikacije koje “prepoznaju” određene događaje, objekte, scene, itd.

U ovoj disertaciji najveća pažnja posvećena je načinima kako teksturalne deskriptore sivih slika na odgovarajući način proširiti na multispektralne slike, a pri tome upotrebiti podatke o prostornom rasporedu lokalnih karakteristika na slikama. Postavljen je veliki broj eksperimenata koji je analizirao pojedinačni uticaj svakog od koraka koji se koristi prilikom nadgledane automatske klasifikacije slika. Posebno su analizirani deskriptori sivih slika i njihove karakteristike. Analizirano je kako je najpogodnije obraditi deskriptore sivih slika i koliko je moguće redukovati njihovu dimenzionalnost, kako bi se sačuvala tačnost klasifikacije. Na osnovu zaključaka donesenih na osnovu rezultata klasifikacije sivih slika, postavljeni su određeni parametri prilikom klasifikacije slika u boji i multispektralnih slika. Predložena su dva načina kako kombinovati deskriptore pojedinačnih kanala multispektralnih slika i tehnike za redukciju dimenzionalnosti, pri čemu se detaljno analizirane njihove performanse za iste trening podatke i jednake parametre klasifikacije.

Kako je prostorni raspored lokalnih karakteristika, tj. lokalnih obilježja teksture, od velikog značaja za prepoznavanje sadržaja prikazanog na slikama, posebna pažnja

tokom predstavljenog istraživanja posvećena je uključivanju podataka o prostornom rasporedu obilježja u deskriptore sivih slika. Predloženo je originalno rješenje za koje je eksperimentalnim putem utvrđeno da povećava tačnost klasifikacije u poređenju sa osnovnim deskriptorima sivih slika. Rezultati klasifikacije korišćenjem predloženih niskodimenzionalnih deskriptora multispektralnih slika i prostorno-teksturalnih deskriptora sivih slika pokazali su mogućnosti razvoja novih algoritama za računanje niskodimenzionalnih prostorno-teksturalnih deskriptora multispektralnih slika. Na osnovu tih rezultata je u Poglavlju 7 predložen algoritam za kreiranje niskodimenzionalnih deskriptora multispektralnih slika koji koristi podatke o prostornom rasporedu lokalnih obilježja u svakom dostupnih kanala.

Izvršen je veliki broj eksperimenata na četiri različite baze slika, kako bi se izveli zaključci potrebni za dalji razvoj predloženih algoritama. Posebno su razmotreni uticaji pojedinačnih koraka i vrijednosti parametara na tačnost klasifikacije baza sivih slika, kao i uticaj primjene tehnika za redukciju dimenzionalnosti deskriptora sivih slika. Izvedeni zaključci su iskorišćeni prilikom eksperimenata sa bazama slika u boji i multispektralnih slika. Pošto su kao deskriptori sivih slika razmatrana četiri deskriptora koja su u literaturi pokazala najbolje rezultate, ispitana je tačnost klasifikacije za ukupno 11 varijanti multispektralnih deskriptora. Rezultati klasifikacije pokazali su da, iako se redukcijom gube diskriminativni podaci, da se odgovarajućim odabirom koraka i parametara mogu sačuvati ili povećati tačnosti klasifikacije korišćenjem niskodimenzionalnih deskriptora.

Razmatranje aktuelnih istraživačkih načina za klasifikaciju slika je kompletirano predstavljanjem metode zasnovane na konvolucionim neuronskim mrežama. Predloženi način za obučavanje neuronskih mreža je omogućio postizanje visoke tačnosti klasifikacije slika u boji. Takođe, mogućnosti primjene neuronskih mreža demonstrirane su implementacijom algoritma za automatsku detekciju objekata na aero snimcima visoke rezolucije.

Budući rad treba da uključi razmatranje načina za proširivanje predloženih metoda za slučaj klasifikacije hiperspektralnih slika, kao i implementaciju praktičnih aplikacija koje koriste algoritme predložene u ovoj disertaciji.

Literatura

- [1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.
- [2] I. Reljin and A. Gavrovska, *Telemedicina*. Akademska misao, 2013.
- [3] L. W. Renninger and J. Malik, “When is scene identification just texture recognition?,” *Vision Research*, vol. 44, pp. 2301–2311, September 2004.
- [4] W.-Y. Ma and B. S. Manjunath, “A texture thesaurus for browsing large aerial photographs,” *J. Am. Soc. Inf. Sci.*, vol. 49, no. 7, pp. 633–648, 1998.
- [5] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proceedings of the International Conference on Computer Vision*, vol. 2 of *ICCV ’99*, pp. 1150–1157, IEEE Computer Society, 1999.
- [6] B. Manjunath, P. Wu, S. Newsam, and H. D. Shin, “A texture descriptor for browsing and similarity retrieval,” *JOURNAL OF SIGNAL PROCESSING: IMAGE COMMUNICATION*, vol. 16, pp. 33–43, 2000.
- [7] T. Sikora, “The mpeg-7 visual standard for content description-an overview.,” *IEEE Trans. Circuits Syst. Video Techn.*, vol. 11, no. 6, pp. 696–702, 2001.
- [8] A. Oliva and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope,” *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, 2001.
- [9] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 971–987, Jul 2002.
- [10] J. Wu and J. Rehg, “Centrist: A visual descriptor for scene categorization,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, pp. 1489–1501, Aug 2011.
- [11] Y. Xiao, J. Wu, and J. Yuan, “mcentrist: A multi-channel feature generation mechanism for scene categorization,” *Image Processing, IEEE Transactions on*, vol. 23, pp. 823–836, Feb 2014.
- [12] A. Avramović and V. Risojević, “Descriptor dimensionality reduction for aerial image classification,” in *Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on*, pp. 1–4, June 2011.

-
- [13] A. Avramović and V. Risojević, “Block-based semantic classification of high-resolution multispectral aerial images,” *Signal, Image and Video Processing*, pp. 75–84, 2014.
- [14] A. Avramović and V. Risojević, “Analysis of spatial partitioning approaches for image classification,” in *Neural Network Applications in Electrical Engineering (NEUREL), 2014 12th Symposium on*, pp. 211–216, Nov 2014.
- [15] A. Avramović, I. Ševo, and I. Reljin, “Comparative analysis of texture classification based on low and high order local features,” in *Telecommunications Forum Telfor (TELFOR), 2015 23rd*, pp. 799–802, Nov 2015.
- [16] I. Ševo and A. Avramović, “Convolutional neural network based automatic object detection on aerial images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, pp. 740–744, May 2016.
- [17] M. Popović, *Digitalna obrada slike*. Akademska misao, 2006.
- [18] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, “Color local texture features for color face recognition,” *IEEE Transactions on Image Processing*, vol. 21, pp. 1366–1380, March 2012.
- [19] Z. Wei, Y. Dong, F. Zhao, and H. Bai, “Face detection based on multi-scale enhanced local texture feature sets,” in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 953–956, March 2012.
- [20] X. Xie and K.-M. Lam, “Efficient human face recognition based on shape and texture,” in *Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference on*, vol. 2, pp. 938–941 vol.2, Aug 2004.
- [21] G. K. Verma and B. K. Singh, “Emotion recognition based on texture analysis of facial expression,” in *Image Information Processing (ICIIP), 2011 International Conference on*, pp. 1–6, Nov 2011.
- [22] S. H. Liu, Y. Cao, Y. Li, M. Li, and S. Hu, “Semantic image classification for medical videos,” in *Semantic Computing, 2009. ICSC '09. IEEE International Conference on*, pp. 648–653, Sept 2009.
- [23] I. Ševo, A. Avramović, I. Balasingham, O. J. Elle, J. Bergsland, and L. Aabakken, “Edge density based automatic detection of inflammation in colonoscopy videos,” *Computers in Biology and Medicine*, vol. 72, pp. 138–150, 2016.
- [24] F. Xu and Y.-J. Zhang, “Evaluation and comparison of texture descriptors proposed in mpeg-7,” *Journal of Visual Communication and Image Representation*, vol. 17, no. 4, pp. 701 – 716, 2006.
- [25] B. Manjunath, J.-R. Ohm, V. Vasudevan, and A. Yamada, “Color and texture descriptors,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, pp. 703–715, Jun 2001.

-
- [26] L. Q. Zhu and Z. Zhang, "Auto-classification of insect images based on color histogram and glm," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*, vol. 6, pp. 2589–2593, Aug 2010.
- [27] U. Bayram, G. Can, S. Duzgun, and N. Yalabik, "Evaluation of textural features for multispectral images," in *Proc. SPIE 8180, Image and Signal Processing for Remote Sensing*, pp. 81800I–81800I–14, 2011.
- [28] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, Nov. 2004.
- [29] M. Brown and S. Süssstrunk, "Multi-spectral SIFT for scene category recognition," in *Proc. CVPR*, pp. 177–184, 2011.
- [30] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *in IEEE Conference on Computer Vision and Pattern Recognition(CVPR, 2009)*.
- [31] R. Arandjelovic and A. Zisserman, "All about vlad," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1578–1585, June 2013.
- [32] Z. Ji, J. Theiler, R. Chartrand, G. Kenyon, and S. P. Brumby, "Sift-based sparse coding for large-scale visual recognition," in *SPIE Defense, Security, and Sensing*, 2013.
- [33] R. Negrel, D. Picard, and P. H. Gosselin, "Evaluation of second-order visual features for land-use classification," in *Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop on*, pp. 1–5, June 2014.
- [34] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000.
- [35] V. Crnojević, *Prepoznavanje oblika za inženjere*. FTN izdavaštvo, Nosi Sad, 2014.
- [36] A. McCallum and K. Nigam, "A comparison of event models for naive bayes text classification," in *IN AAAI-98 WORKSHOP ON LEARNING FOR TEXT CATEGORIZATION*, pp. 41–48, AAAI Press, 1998.
- [37] E. Frank and R. R. Bouckaert, "Naive bayes for text classification with unbalanced classes," in *Proceedings of the 10th European Conference on Principle and Practice of Knowledge Discovery in Databases, PKDD'06, (Berlin, Heidelberg)*, pp. 503–510, Springer-Verlag, 2006.
- [38] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, pp. 81–106, Mar. 1986.
- [39] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Sept. 1995.

-
- [40] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [41] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, “Land use classification in remote sensing images by convolutional neural networks,” *CoRR*, vol. abs/1508.00092, 2015.
- [42] H. Jégou, M. Douze, C. Schmid, and P. Pérez, “Aggregating local descriptors into a compact image representation,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 3304–3311, June 2010.
- [43] H. Jegou, F. Perronnin, M. Douze, J. Sánchez, P. Perez, and C. Schmid, “Aggregating local image descriptors into compact codes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, pp. 1704–1716, Sept. 2012.
- [44] I. Jolliffe, *Principal Component Analysis*. Springer Verlag, 1986.
- [45] G. Balakrishnama, “Linear Discriminant Analysis - A brief Tutorial,” 1998.
- [46] S. Lazebnik, C. Schmid, and J. Ponce, “A sparse texture representation using local affine regions,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, pp. 1265–1278, Aug 2005.
- [47] B. Julesz, “Texton gradients: The texton theory revisited,” *Biological Cybernetics*, vol. 54, no. 4, pp. 245–251, 1986.
- [48] S. Alvarez and M. Vanrell, “Texton theory revisited: A bag-of-words approach to combine textons,” *Pattern Recognition*, vol. 45, no. 12, pp. 4312 – 4325, 2012.
- [49] Y. Yang and S. Newsam, “Bag-of-visual-words and spatial extensions for land-use classification,” in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS ’10*, (New York, NY, USA), pp. 270–279, ACM, 2010.
- [50] G. Bradbury, K. Mitchell, and T. Weyrich, “Multi-spectral material classification in landscape scenes using commodity hardware,” in *Computer Analysis of Images and Patterns*, vol. 8048, pp. 209–216, Springer Berlin Heidelberg, 2013.
- [51] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, “Describing textures in the wild,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR ’14*, (Washington, DC, USA), pp. 3606–3613, IEEE Computer Society, 2014.
- [52] M. A. Hoang, J.-M. Geusebroek, and A. W. Smeulders, “Color texture measurement and segmentation,” *Signal Processing*, vol. 85, no. 2, pp. 265 – 275, 2005. {SI} on Content Based Image and Video Retrieval.
- [53] S. Saleem and R. Sablatnig, “A robust sift descriptor for multispectral images,” *IEEE Signal Processing Letters*, vol. 21, pp. 400–403, April 2014.

-
- [54] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, “Multi-spectral image feature points,” *Sensors*, vol. 12, no. 9, p. 12661, 2012.
- [55] A. E. Abdel-Hakim and A. A. Farag, “Csift: A sift descriptor with color invariant characteristics,” in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, pp. 1978–1983, 2006.
- [56] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in *Computer Vision and Pattern Recognition*, vol. 2, pp. 2169–2178, 2006.
- [57] Y. Jiang, J. Yuan, and G. Yu, “Randomized spatial partition for scene recognition,” in *Computer Vision – ECCV 2012*, pp. 730–743, Springer Berlin Heidelberg, 2012.
- [58] P. F. Liu Li and G. Kuang, “Generalized local binary patterns for texture classification,” in *Proceedings of the British Machine Vision Conference*, pp. 123.1–123.11, BMVA Press, 2011. <http://dx.doi.org/10.5244/C.25.123>.
- [59] S. Banerji, A. Verma, and C. Liu, *Cross Disciplinary Biometric Systems*, ch. LBP and Color Descriptors for Image Classification, pp. 205–225. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [60] S. Banerji, A. Sinha, and C. Liu, *Computer Analysis of Images and Patterns: 15th International Conference, CAIP 2013, York, UK, August 27-29, 2013, Proceedings, Part I*, ch. A New Bag of Words LBP (BoWL) Descriptor for Scene Image Classification, pp. 490–497. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.
- [61] W.-L. Zhao, H. Jégou, and G. Gravier, “Oriented pooling for dense and non-dense rotation-invariant features,” in *BMVC - 24th British Machine Vision Conference*, (Bristol, United Kingdom), Sep 2013.
- [62] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, “Locality-constrained linear coding for image classification,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 3360–3367, June 2010.
- [63] J. C. Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders, *Computer Vision – ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III*, ch. Kernel Codebooks for Scene Categorization, pp. 696–709. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.
- [64] J. C. Caicedo, A. Cruz, and F. A. Gonzalez, *Artificial Intelligence in Medicine: 12th Conference on Artificial Intelligence in Medicine, AIME 2009, Verona, Italy, July 18-22, 2009. Proceedings*, ch. Histopathology Image Classification Using Bag of Features and Kernel Functions, pp. 126–135. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

-
- [65] A. Cheriyyadat, “Unsupervised feature learning for aerial scene classification,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, pp. 439–451, Jan 2014.
- [66] V. Risojević and Z. Babić, “Unsupervised quaternion feature learning for remote sensing image classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, pp. 1521–1531, April 2016.
- [67] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25* (P. Bartlett, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, eds.), pp. 1106–1114, 2012.
- [68] K. Nogueira, W. O. Miranda, and J. A. D. Santos, “Improving spatial feature representation from aerial scenes by using convolutional networks,” in *Graphics, Patterns and Images (SIBGRAPI), 2015 28th SIBGRAPI Conference on*, pp. 289–296, Aug 2015.
- [69] O. Penatti, K. Nogueira, and J. Dos Santos, “Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, pp. 44–51, June 2015.
- [70] F. Luus, B. Salmon, F. van den Bergh, and B. Maharaj, “Multiview deep learning for land-use classification,” *Geoscience and Remote Sensing Letters, IEEE*, vol. PP, no. 99, pp. 1–5, 2015.
- [71] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, “Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery,” *Remote Sensing*, vol. 7, no. 11, p. 14680, 2015.
- [72] Y. Xu, H. Ji, and C. Fermüller, “Viewpoint invariant texture description using fractal analysis,” *International Journal of Computer Vision*, vol. 83, no. 1, pp. 85–100, 2009.
- [73] M. Bratkova, S. Boulos, and P. Shirley, “orgb: A practical opponent color space for computer graphics,” *IEEE Comput. Graph. Appl.*, vol. 29, pp. 42–55, Jan. 2009.
- [74] K. van de Sande, T. Gevers, and C. Snoek, “Evaluating color descriptors for object and scene recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, pp. 1582–1596, Sept. 2010.
- [75] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, pp. 886–893 vol. 1, June 2005.
- [76] R. Arandelović and A. Zisserman, “Three things everyone should know to improve object retrieval,” in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 2911–2918, June 2012.

- [77] T. Maenpaa, M. Pietikainen, and J. Viertola, “Separating color and pattern information for color texture discrimination,” in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 1, pp. 668–671 vol.1, 2002.
- [78] C. H. Chan, J. Kittler, and K. Messer, “Multispectral local binary pattern histogram for component-based color face verification,” in *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pp. 1–7, Sept 2007.
- [79] S. D. Newsam and C. Kamath, “Retrieval using texture features in high-resolution multispectral satellite imagery,” in *Proc. SPIE*, vol. 5433, pp. 21–32, 2004.
- [80] M. F. Caropreso and S. Matwin, *Advances in Artificial Intelligence: 19th Conference of the Canadian Society for Computational Studies of Intelligence, Canadian AI 2006, Québec City, Québec, Canada, June 7-9, 2006. Proceedings*, ch. Beyond the Bag of Words: A Text Representation for Sentence Selection, pp. 324–335. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
- [81] F. Perronnin and C. R. Dance, “Fisher kernels on visual vocabularies for image categorization,” in *CVPR*, IEEE Computer Society, 2007.
- [82] F. Perronnin, J. Sánchez, and T. Mensink, “Improving the fisher kernel for large-scale image classification,” in *Proceedings of the 11th European Conference on Computer Vision: Part IV, ECCV’10*, (Berlin, Heidelberg), pp. 143–156, Springer-Verlag, 2010.
- [83] G. Sharma, S. ul Hussain, and F. Jurie, *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII*, ch. Local Higher-Order Statistics (LHS) for Texture Categorization and Facial Analysis, pp. 1–12. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [84] L. Sharan, C. Liu, R. Rosenholtz, and E. H. Adelson, “Recognizing materials using perceptually inspired features,” *International Journal of Computer Vision*, vol. 103, no. 3, pp. 348–371, 2013.
- [85] F. Hu, G. S. Xia, Z. Wang, X. Huang, L. Zhang, and H. Sun, “Unsupervised feature learning via spectral clustering of multidimensional patches for remotely sensed scene classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, pp. 2015–2030, May 2015.
- [86] O. A. B. Penatti, K. Nogueira, and J. A. dos Santos, “Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 44–51, June 2015.
- [87] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in *Proceedings of the 22Nd ACM International Conference on Multimedia, MM ’14*, (New York, NY, USA), pp. 675–678, ACM, 2014.

- [88] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” *CoRR*, vol. abs/1409.4842, 2014.
- [89] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 248–255, June 2009.
- [90] S. Chen and Y. Tian, “Pyramid of spatial relations for scene-level land use classification,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, pp. 1947–1957, April 2015.
- [91] Y. Yang and S. Newsam, “Spatial pyramid co-occurrence for image classification,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 1465–1472, Nov 2011.

Biografija autora:

Aleksej Avramović je rođen 16.2.1984. godine u Mostaru, Bosna i Hercegovina. Usljed ratnih dejstava i izbjeglištva, osnovnu školu je pohađao u Konjicu, Vukovaru i Brčkom, gdje je završio 1998. godine, odličnim uspjehom. U osnovnoj školi učestvovao je na nekoliko opštinskih i regionalnih takmičenja iz fizike i matematike. U Brčkom je upisao i završio gimnaziju "Vaso Pelagić", 2002. godine odličnim uspjehom, a iste godine je upisao Elektrotehnički fakultet u Banjoj Luci, odsjek za Elektroniku i komunikacije. Petogodišnje studije elektrotehnike je završio u predviđenom roku odbranom diplomskog rada 9.7.2007. sa temom "*Segmentacija digitalnih slika primenom Mean shift algoritma*" sa srednjim prosjekom ocjena 8,82. Studije trećeg ciklusa je upisao krajem 2008. godine na Elektrotehničkom fakultetu u Beogradu na smjeru za Telekomunikacije. Sve ispite je položio sa prosječnom ocjenom 10,0.

Od 25.9.2007. je zaposlen kao saradnik na Elektrotehničkom fakultetu u Banjoj Luci, a od 8.3.2008. je stalno zaposlen sa izbor u zvanje *Asistenta* za užu naučnu oblast *Opšta elektrotehnika*. Od 25.5.2012. dobio je izbor u zvanje *Višeg asistenta* na užoj naučnoj oblasti *Opšta elektrotehnika*. Nastavna djelatnost se odnosi na izvođenje auditornih i laboratorijskih vježbi iz predmeta: Osnovi elektrotehnike 1, Osnovi elektrotehnike 2, Električna mjerenja, Teorija električnih kola 1, Teorija električnih kola 2 i Računarska grafika.

Pored nastavnih zaduženja na fakultetu, autor je bio angažovan kao saradnik na nekoliko naučno-istraživačkih projekata finansiranih od strane Ministarstva nauke i tehnologije Republike Srpske, kao i na nekoliko međunarodnih projekata. Od projekata nacionalnog značaja, potrebno je izdvojiti angažovanje na projektima: Arhiviranje i pretraživanje baza radioloških slika (2006-2008.), RFID tehnologija (2008-2010.), Kompresija slika bez gubitaka (2010.) i Automatska klasifikacija pokrivenosti i načina korištenja zemljišta (2012-2013.). Međunarodna naučno-istraživačka saradnja je ostvarena kroz angažovanje na projektima bilateralne saradnje između Republike Bosne i Hercegovine i Republike Slovenije: Kompresija medicinskih slika (Slovenian Research Agency and Ministry of Civil Affairs, Bosnia and Herzegovina, grant BI-BA/10-11-026, Bilateral Collaboration Project, 2010-2011.) i Automatska anotacija medicinskih video zapisa (Slovenian Research Agency and Ministry of Civil Affairs, Bosnia and Herzegovina, grant BI-BA/14-15-035, Bilateral Collaboration Project, 2014-2015.), projekat u toku. Takođe, autor je išao na sedmodnevnu studijsku posjetu Univerzitetu u Paderbornu, Njemačka 2010. godine, dvonedeljnu studijsku posjetu Buskerud koledžu u Kongsbergu, Norveška 2013. godine, kao i sedmičnu radna posjetu IVS istraživačkom centru u Oslu, 2014. godine. Od 2008. godine angažovan je u akademskoj mreži Sarnet, gdje je zadužen za videokonferencijske sisteme. Član je udruženja IEEE, sekcije za kola i sisteme.

Istraživačka djelatnost se odnosi na oblasti Digitalne obrade slike, konkretno na kompresiju slika bez gubitaka, pretraživanje baza slika i klasifikaciju slika, kao i na oblast kompjuterske aritmetike. Do sada je kao autor ili koautor objavio 33 naučna rada, od čega 5 u međunarodnim časopisima sa SCI liste. Autor je dva praktikuma

za laboratorijske vježbe: “Osnovi elektrotehnike 1: praktikum za laboratorijske vježbe” i “Osnovi elektrotehnike 2: praktikum za laboratorijske vježbe”. Radovi u međunarodnim časopisima sa SCI liste:

1. Zdenka Babić, **Aleksej Avramović**, Patricio Bulić: “An Iterative Logarithmic Multiplier”, *Microprocessors and Microsystems - Embedded Hardware Design*, Elsevier, Vol. 35, Issue 1, pp. 23-33, ISSN 0141-9331, DOI 10.1016/j.micpro.2010.07.001, February 2011, Impact Factor 2011: **0.575**
2. **Aleksej Avramović**, Zdenka Babić, Dušan Raič, Drago Strle, Patricio Bulić: “An approximate logarithmic squaring circuit with error compensation for DSP applications”, *Microelectronics Journal*, Elsevier Vol. 45, Issue 3, pp. 263-271, ISSN 0026-2692, DOI 10.1016/j.mejo.2014.01.005, March 2014, Impact Factor 2014: **0.924**
3. **Aleksej Avramović**, Vladimir Risojević: “Block-based semantic classification of high-resolution multispectral aerial images”, *Signal, Image and Video Processing*, Springer, Vol. 10, Issue 1, pp. 75-84, ISSN 1863-1711, DOI 10.1007/s11760-014-0704-x, January 2016, Impact Factor 2015: **0.872**
4. Igor Ševo, **Aleksej Avramović**, Ilangko Balasingham, Olle Jakob Elle, Jacob Bergsland, Lars Aabakken: “Edge density based automatic detection of inflammation in colonoscopy videos”, *Computers in Biology and Medicine*, Elsevier, Vol. 72, pp. 138-150, ISSN 0010-4825, DOI 10.1016/j.combiomed.2016.03.017, May 2016, Impact Factor 2015: **1.521**
5. Igor Ševo, **Aleksej Avramović**: “Convolutional Neural Network Based Automatic Object Detection on Aerial Images”, *IEEE Geoscience and Remote Sensing Letters*, Vol. 13, Issue 5, pp. 740-744, ISSN 1545-598X, DOI 10.1109/LGRS.2016.2542358, May 2016, Impact Factor 2015: **2.228**

Prilog 1.

Izjava o autorstvu

Potpisani Aleksej Avramović,

IZJAVLJUJEM

da je doktorska disertacija pod naslovom “Niskodimenzionalni prostorno-teksturalni deskriptori multispektralnih slika”

- rezultat sopstvenog istraživačkog rada,
- da predložena disertacija u cjelini ni u dijelovima nije predložena za dobijanje bilo koje diplome prema studijskim programima drugih visokoškolskih ustanova,
- da su rezultati korektno navedeni,
- da nisam kršio autorska prava i koristio intelektualnu svojinu drugih lica.

Potpis doktoranda:

U Beogradu, _____

Prilog 2.

Izjava o istovjetnosti štampane i elektronske verzije rada

Ime i prezime autora:	Aleksej Avramović
Naslov rada:	Niskodimenzionalni prostorno-teksturalni deskriptori multispektralnih slika
Mentor:	dr Irini Reljin, redovni profesor

Potpisani Aleksej Avramović,

Izjavljujem da je štampana verzija mog doktorskog rada istovjetna elektronskoj verziji koju sam predao za objavljivanje na portalu Digitalnog repozitorijuma Univerziteta u Beogradu. Dozvoljavam da se objave moji lični podaci vezani za dobijanje akademskog zvanja doktora nauka, kao što su ime i prezime, godina i mjesto rođenja i datum odbrane rada. Ovi lični podaci mogu se objaviti na mrežnim stranicama digitalne biblioteke, u elektronskom katalogu i u publikacijama Univerziteta u Beogradu.

Potpis doktoranda:

U Beogradu, _____

Prilog 3.

Izjava o korišćenju

Ovlašćujem Univerzitetsku biblioteku "Svetozar Marković" da u Digitalni repozitorijum Univerziteta u Beogradu unese moju doktorsku disertaciju pod naslovom "Niskodimenzionalni prostorno-teksturalni deskriptori multispektralnih slika", koja je moje autorsko djelo.

Disertaciju sa svim prilogima predao sam u elektronskom formatu pogodnom za trajno arhiviranje. Moju doktorsku disertaciju pohranjenu u Digitalni repozitorijum Univerziteta u Beogradu mogu da koriste svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons) za koju sam se odlučio.

1. Autorstvo.
2. Autorstvo - nekomercijalno.
3. Autorstvo - nekomercijalno - bez prerade.

4. Autorstvo - nekomercijalno - dijeliti pod istim uslovima.
5. Autorstvo - bez prerade.
6. Autorstvo - dijeliti pod istim uslovima.

(Kratak opis licenci dat je na sljedećoj strani.)

Potpis doktoranda:

U Beogradu, _____

Prilog 3.

1. Autorstvo. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence, čak i u komercijalne svrhe. Ovo je najslobodnija od svih licenci.
2. Autorstvo - nekomercijalno. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu djela.
3. Autorstvo - nekomercijalno - bez prerade. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, bez promena, preoblikovanja ili upotrebe djela u svom djelu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu djela. U odnosu na sve ostale licence, ovom licencom se ograničava najveći obim prava korišćenja djela.
4. Autorstvo - nekomercijalno - dijeliti pod istim uslovima. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca ne dozvoljava komercijalnu upotrebu djela i prerada.
5. Autorstvo - bez prerade. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, bez promena, preoblikovanja ili upotrebe djela u svom djelu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca dozvoljava komercijalnu upotrebu djela.
6. Autorstvo - dijeliti pod istim uslovima. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje djela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca dozvoljava komercijalnu upotrebu djela i prerada. Slična je softverskim licencama, odnosno licencama otvorenog koda.