

Klasifikacija grafikona primjenom plitkih konvolucijskih neuronskih mreža i izdvajanje podatkovnih značajki iz klasificiranih grafikona

Bajić, Filip

Doctoral thesis / Doktorski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:200:000321>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-12-28**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA

Filip Bajić

**Klasifikacija grafikona primjenom plitkih
konvolucijskih neuronskih mreža i izdvajanje
podatkovnih značajki iz klasificiranih grafikona**

Doktorski rad

Osijek, 2024.

Doktorski rad izrađen je na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija Osijek

Mentor: prof. dr. sc. Krešimir Nenadić, redoviti profesor, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek, Sveučilište Josipa Jurja Strossmayera u Osijeku

Sumentor: dr. sc. Marija Habijan, viši asistent, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek, Sveučilište Josipa Jurja Strossmayera u Osijeku

Doktorski rad ima 99 stranica

Doktorski rad br.: 91

Naslov doktorskog rada

- (HRV) Klasifikacija grafikona primjenom plitkih konvolucijskih neuronskih mreža i izdvajanje podatkovnih značajki iz klasificiranih grafikona
- (ENG) Chart image classification using a shallow convolutional neural network and feature extraction from classified charts

Izvorni znanstveni doprinosi

- Metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama
- Metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima
- Metoda generiranja ispitnih skupova za standardiziranu validaciju uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona

Povjerenstvo za ocjenu doktorske disertacije

- izv. prof. dr. sc. Zdravko Krpić, predsjednik, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek
- izv. prof. dr. sc. Ivica Lukić, član, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek
- doc. dr. sc. Mario Brčić, član, Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva
- izv. prof. dr. sc. Josip Balen, zamjena, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek

Povjerenstvo za obranu doktorske disertacije

- izv. prof. dr. sc. Zdravko Krpić, predsjednik, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek
- izv. prof. dr. sc. Ivica Lukić, član, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek
- doc. dr. sc. Mario Brčić, član, Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva
- izv. prof. dr. sc. Tomislav Keser, zamjena, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek

Datum obrane disertacije: 25. lipnja 2024. godine

ZAHVALE

Zahvaljujem mentoru, dr. sc. Krešimiru Nenadiću, na povjerenju, podršci i vodstvu tijekom doktorskog studija. Hvala na upornosti, razumijevanju, prenesenom znanju i dostupnosti. Ova disertacija odražava ne samo moje napore, već i mentorovu predanost izvrsnosti. Hvala Vam što ste mi bili mentor i omogućili mi akademski razvoj.

Posebnu zahvalnost i poštovanje izrazio bih nekadašnjoj prijateljici iz klupe, a danas vrhunskoj mladoj znanstvenici i sumentorici, dr. sc. Mariji Habijan. Hvala na konstruktivnim savjetima, prijedlozima i usmjeravanju tijekom doktorskog studija. Bila mi je iznimna čast surađivati s Vama, programirati, pisati i objavljivati znanstvene radove. Nadam se da ćemo u budućnosti surađivati na još mnogo znanstvenih projekata.

Hvala Sveučilišnom računskom centru (Srce) na pruženoj financijskoj podršci. Vaša velikodušnost ima značajan utjecaj na moj život i obrazovnu karijeru stoga želim iskreno izraziti priznanje za Vašu nesebičnost. Posebno bih istaknuo šefa, dr. sc. Ognjena Orela, koji je pokazao razumijevanje, stručnost i profesionalnost za vrijeme trajanja doktorskog studija.

Na kraju, zahvalio bih se roditeljima, majci Zdenki i ocu Vladi, što su bili podrška i što su me ohrabрили i potaknuli na upis doktorskog studija.

SADRŽAJ

1	Uvod.....	1
2	Metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama.....	6
2.1	Uvod.....	7
2.2	Istraživanja u području klasifikacije slike grafikona.....	9
2.3	Podatkovni skupovi.....	13
2.3.1	Algoritam predobrade slike grafikona.....	14
2.3.2	Opis korištenih podatkovnih skupova.....	18
2.4	Arhitektura umjetne neuronske mreže.....	21
2.4.1	Plitka konvolucijska neuronska mreža.....	21
2.4.2	Sijamska plitka konvolucijska neuronska mreža.....	24
2.5	Dubinska analiza neuronske mreže.....	26
2.6	Eksperimentalna evaluacija – SCNN.....	31
2.6.1	Monte Carlo unakrsno vrednovanje.....	32
2.6.2	Određivanje uspješnosti.....	34
2.6.3	Otvorena klasifikacija.....	35
2.6.4	Ostvareni rezultati modela.....	36
2.7	Eksperimentalna evaluacija – SSCNN.....	41
2.7.1	Ostvareni rezultati modela.....	42
2.8	Zaključak.....	44
3	Metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima.....	46
3.1	Uvod.....	47

3.2	Istraživanja u području izdvajanja podatkovnih značajki iz slike grafikona	48
3.3	Metoda izdvajanja podatkovnih značajki	52
3.3.1	Kružni grafikoni	53
3.3.2	Mrežasti grafikoni	56
3.4	Eksperimentalna evaluacija.....	58
3.4.1	Kružni grafikoni	59
3.4.2	Mrežasti grafikoni	61
3.4.3	Slike prikupljene internet pretraživanjem	63
3.5	Zaključak	64
4	Metoda generiranja ispitnih skupova za standardiziranu validaciju uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona	66
4.1	Uvod.....	67
4.2	Istraživanja u području dostupnih podatkovnih skupova koji sadrže slike grafikona... 68	
4.3	Metoda za generiranje ispitnih skupova.....	72
4.4	Usporedba podatkovnih skupova ChartDS i ICDAR7.....	75
4.5	Zaključak.....	78
5	Zaključak.....	80
	Literatura	84
	Sažetak	94
	Abstract	96
	Životopis.....	98
	Popis slika	99
	Popis tablica	102
	Prilozi	103

1 UVOD

Jedan od najstarijih dokumentiranih pokušaja slikovnog prikazivanja kvantitativnih informacija potječe iz desetog stoljeća. To je vremenski prikaz kretanja sunca, mjeseca i planeta tijekom godine, čiji autor je nepoznat, a više informacija dostupno je u [1]. Prvi slikovni prikaz sličan današnjem linijskom grafikonu, "*ligne de vie*", izradio je Christiaan Huygens 1669. godine, kao što je prikazano u [2]. Slikovni prikaz sadrži kontinuiranu funkciju distribucije i pokazuje kako pronaći medijan za preostali život osobe. Godine 1765. Joseph Priestley izradio je povijesni vremenski okvir koji prikazuje životni vijek 2000 poznatih osoba od 1200. godine pr. Kr. do 1750. godine, kao što je prikazano u [1]. U istaknutim primjerima autori prepoznaju važnost slikovnih prikaza pri prezentiranju složenih numeričkih informacija. Kroz povijest, doprinos Huygensa i Priestleya nije samo promicao razvoj statistike, već je i unaprijedio razvoj slikovnog prikaza podataka.

William Playfair (1759. – 1823.), škotski inženjer i politički ekonomist, donio je revoluciju u statističkoj grafici. Playfair je bio jedan od prvih koji je koristio podatke ne samo za edukaciju, već i za uvjeravanje ljudi. Prema navodima autora u [1], Playfair se smatra izumiteljem prvog linijskog, stupčastog i kružnog grafikona. Playfair je shvaćao kako slikovni prikaz podataka može omogućiti mozgu učinkovitiju obradu informacija i pri tome smanjiti trošak energije potrebne za pamćenje. Korištenje slikovnog prikaza podataka zahtijeva manje pozornosti, koncentracije i usredotočenosti čitatelja za pohranu važnih informacija u dugotrajno pamćenje. U devetnaestom stoljeću statistika je značajno napredovala, a statistički podatci su postali široko dostupni, što je dodatno ukazalo na potrebu za prikazom složenih struktura podataka. U dvadesetom stoljeću, zbog sve veće količine informacija, nastala je potreba za sažimanjem informacija. Slikovni prikazi počinju se koristiti smisleno za predstavljanje važnih informacija relevantnih za provedena

istraživanja. Vizualizacija podataka, jedan od oblika slikovnog prikaza, omogućuje jasno i potpuno razumijevanje odnosa među podacima, pomaže u razumijevanju ostvarenih rezultata, a može se koristiti i za procjenu vrijednosti koje nisu izravno određene mjernim metodama interpolacije i ekstrapolacije. Vizualizacija podataka je ključna za prepoznavanje neobičnih ili neočekivanih rezultata i olakšava usporedbu različitih vrijednosti i odnosa među podacima. Ljudska sposobnost brzog i lakog vizualnog uočavanja i uspoređivanja temelji se na sposobnosti mozga u otkrivanju pravilnosti i nesavršenosti. Navedena sposobnost događa se u podsvijesti. Usporedba se događa prije razmišljanja.

U današnjem svijetu naprednih internetskih tehnologija podatci i informacije imaju značajnu ulogu. Internet je interaktivni medij kroz koji prolazi značajna količina podataka i informacija svake sekunde. Sve je moguće mjeriti i opisivati brojevima, a brojevi se brzo nakupljaju i stvaraju nestrukturirane podatke. Ljudi teško razumiju nagomilane i nestrukturirane podatke. Tablice koje sadrže veliku količinu podataka nisu jednostavno čitljive i zahtijevaju mentalni napor za dobivanje korisne informacije. Statistika je često predstavljena brojevima, koje je ponekad izazovno pročitati i razlikovati važne od nevažnih informacija. Tijekom godina, razvijene su jedinstvene metode i alati koji omogućavaju slikovni prikaz podataka, odnosno omogućavaju prezentiranje podataka kroz vizualizacije podataka, kao što je prikazano u [3]. Vizualizaciji podataka pripadaju grafikoni ili dijagrami stvoreni postupkom sažimanja podataka u jedan jedinstveni vizualni identitet kojeg pojedinac (čitatelj) može jednostavno razumjeti i omogućavaju jednostavnije prenošenje, razumijevanje informacija i donošenje odluka. Prema [4], grafikon je prikaz nekog procesa u prirodi i društvu predočen krivuljom u koordinatnom sustavu. Svaka vrsta grafikona izrađena je s ciljem predstavljanja određene vrste podataka na odgovarajući način.

Vizualizacija podataka prisutna je u svakom životnom pogledu, od matematike, statistike i analitike do bilo kojeg mjesta gdje postoji potreba za identifikacijom obrasca u skupu podataka. Najčešći medij koji sadrži vizualizaciju podataka su papir (primjerice časopisi i novine), digitalne platforme (primjerice društvene mreže i mobilne aplikacije) i razni oblici umjetničkih instalacija prisutni u muzejima, na javnim površinama i društvenim događajima. Tiskani i digitalni dokumenti sadrže slikovne i tekstualne informacije povezane u jednu jedinstvenu vizualnu cjelinu. Svake godine definiraju se pitanja i izazovi u pronalaženju i pohranjivanju dokumenata i informacija koje se nalaze u njima. Izazov nastaje kada internet tražilice u svoje rezultate pretraživanja trebaju uključiti i rezultate koji su na određeni način vizualizirani ili kada tim informacijama pokušavaju pristupiti slijepe i slabovidne osobe. Mnogi autori vizualizacija podataka nisu upoznati s izazovima

s kojima se susreću slijepe i slabovidne osobe stoga ne razvijaju dokumente i pripadajuće vizualizacije po smjernicama pristupačnosti. Kao rezultat navedenog, većina digitalnih dokumenata ne sadrži dodatne informacije (primjerice metapodatke, opisni tekst ili tablicu) ili su pružene informacije vrlo općenite, kratke i neprimjerene. Primjera radi, prema istraživanju provedenom u [5], skup vizualizacija podataka koji se sastoji od 2702 slike prikupljene pomoću Google pretraživanja slika sadržavao je 39 (1.44%) slika s opisnim metapodacima. Metapodatci su korisni internet tražilicama, ali slijepe i slabovidne osobe koje koriste različite čitače zaslona zahtijevaju veću količinu informacija kako bi razumjeli predstavljenu vizualizaciju podataka. Čitači zaslona sadrže ugrađenu programsku logiku koja se oslanja na posebno istaknute informacije vezane za slikovni prikaz i služe kao interaktivno sredstvo između digitalnog svijeta i slijepe ili slabovidne osobe. Tekst u neposrednoj blizini slikovnog prikaza najčešće nije dovoljan za pružanje predodžbe prikaza. Tekst u blizini vizualizacije podataka vrlo rijetko sadrži detalje kao što su vrsta i tip grafikona, minimalna, maksimalna ili istaknuta vrijednost. Osvrćući se na predstavljenu povijesnu analizu, do danas, dodatne informacije koje identificiraju ili opisuju slikovni prikaz, a time i vizualizaciju podataka, su najčešće zaključane unutar samog prikaza. Vizualizacije podataka iskorištavaju ljudsku vizualnu svijest za učinkovit i djelotvoran prijenos informacija. Takvi slikovni prikazi nisu namijenjeni korištenju u digitalnom svijetu. Dok ljudi mogu jednostavno protumačiti vizualizacije podataka i kreirati dodatne tablične prikaze, računala ne mogu, više informacija o izazovima dostupno je u [6,7].

Slijedom predstavljenih izazova s kojima su suočene slijepe i slabovidne osobe, računala, čitači zaslona i internet tražilice, provode se različita znanstvena istraživanja s ciljem povećanja dostupnosti vizualiziranih podataka. Poseban naglasak tih istraživanja je na povećanju točnosti klasifikacije grafikona, kako bi se omogućila točnija i detaljnija interpretacija informacije sadržane u slici grafikona, kao što je prikazano u [7–11]. Klasifikacija je proces organiziranja ili razvrstavanja objekata, to jest grafikona u skupine, to jest klase na temelju zajedničkih svojstava. Točnost klasifikacije je mjera koja predstavlja koliko je objekata, to jest grafikona od ukupnog broja, točno razvrstano u skupine, to jest klase. Prvi korak u uspješnoj interpretaciji grafikona je određivanje pripadajuće klase. Nakon uspješne klasifikacije, grafikon se može obraditi algoritmima za izdvajanje podatkovnih značajki, sustavima za optičko prepoznavanje znakova (engl. *Optical character recognition – OCR*) ili sustavima koji omogućavaju generiranje tekstualnog opisa upotrebom algoritama za obradu prirodnog jezika (engl. *Natural language processing*).

U ovoj disertaciji se predlaže razvoj tri metode koje su opisane u tri poglavlja. U poglavlju 2 predstavljena je metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama. Ulaz modela neuronske mreže je slika grafikona ili bilo koja druga slika koja se može naći u elektroničkom dokumentu. S ciljem postizanja što veće točnosti klasifikacije, slike grafikona su predobrađene upotrebom metode koja sa slike uspješno uklanja značajan broj sadržajnih informacija ostavljajući grafički element čiji je sadržaj dovoljan za razlikovanje klase grafikona. U sklopu evaluacije, od predstavljene metode nastalo je 32 modela koja su učena na podatkovnim skupovima različitih veličina i različitog sadržaja. U cjelokupnom postupku evaluacije sudjeluje šest javno dostupnih podatkovnih skupova. Uz evaluaciju, izrađena je i dubinska analiza predložene metode kojoj je cilj objasniti korake u procesu klasifikacije slike grafikona i dokazati ostvarenu točnost klasifikacije iz poglavlja evaluacije.

Metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima predstavljena je u poglavlju 3. U kružne grafikone ubrajaju se tortni, prstenasti i grafikon sunčevih zraka, a u mrežaste grafikone ubrajaju se toplinske karte i grafikoni napretka. Metoda se sastoji od dva modula: modul za predobradu slike i modul za detekciju strukturnih elemenata grafikona. Rad metode zasnovan je na binarnim maskama, morfologiji, detekciji ključnih točaka i matematičkom modelu. Metoda je evaluirana na umjetno stvorenom podatkovnom skupu gdje ostvaruje visoku točnost izdvajanja podatkovnih značajki, kao što su vrijednosti i lokacije elemenata na slici, korištena paleta boja i struktura slike grafikona. U sklopu evaluacije prikazano je nekoliko izgledom složenih grafikona iz kojih metoda stvara uspješnu rekonstrukciju izvorne slike grafikona na osnovu izdvojenih podatkovnih značajki sadržanih unutar zasebnog dokumenta.

U poglavlju 4 predstavljena je metoda generiranja ispitnih skupova za standardizirano vrednovanje uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona. Metoda je temeljena na slučajnim vrijednostima i lokacijama elemenata koji su strukturno sadržani u slici grafikona. Na osnovu predstavljene metode nastao je podatkovni skup ChartDS koji se koristi u evaluaciji metoda iz poglavlja 2 i 3. Podatkovni skup je javno dostupan u podatkovnim repozitorijima Figshare, Srce Dabar i Mendeley Data, te kao takav uspostavlja zajedničku osnovu za evaluaciju, nudeći standardizirano ishodište za procjenu i usporedbu učinkovitosti različitih sustava u razumijevanju i interpretaciji slika grafikona.

U ovoj disertaciji su razvijene prethodno navedene metode koje bi, integrirane u jednu programsku okolinu, činile funkcionalnu cjelinu. Tako razvijeno aplikativno rješenje bi od ulazne

slike grafikona stvorilo objektno orijentiranu strukturu koja u stvarnim računalnim sustavima ima brojne primjene. Aplikacija bi, upotrebom neuronske mreže, ulaznu sliku klasificirala u dvije klase: grafikon i ostalo. Ako je slika klasificirana kao grafikon, određuje se klasa grafikona. Nadalje, ako je slika klasificirana kao kružni ili mrežasti grafikon, sa slike je moguće napraviti izdvajanje podatkovnih značajki i te značajke pohraniti u objektno orijentiranu strukturu koju mogu konzumirati vanjske aplikacije. Ispravnost rada cjelokupnog sustava može se provjeriti javno dostupnim podatkovnim skupom koji sadrži izvorne podatke korištene za izradu slika grafikona izrađene upotrebom metode za generiranje ispitnih skupova. Realizacija aplikativnog rješenja i produkcijski spremna okolina nisu obuhvaćeni ovom disertacijom, ali predložene metode omogućavaju realizaciju aplikativnih rješenja.

2 METODA ZA OTVORENU KLASIFIKACIJU GRAFIKONA U ELEKTRONIČKIM DOKUMENTIMA ZASNOVANA NA PLITKIM KONVOLUCIJSKIM NEURONSKIM MREŽAMA

U ovom poglavlju biti će opisana metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama. Ovaj doprinos istraživanju objavljen je u radovima [12–14], a provedeno istraživanje povezano je s [5,7,8]. Poglavlje opisuje arhitekture neuronskih mreža za klasifikaciju slike grafikona. Proces učenja i evaluacije neuronskih mreža proveden je na šest javno dostupnih podatkovnih skupova, koji su označeni kao: ReVision, ICDAR7, ChartDS, Linnaeus, CIFAR10 i AT&T DoF. Od navedenih podatkovnih skupova ReVision, ICDAR7 i ChartDS sadrže slike grafikona dok ostali podatkovni skupovi sadrže raznolike slike. Za obje arhitekture neuronske mreže dan je detaljan prikaz strukture mrežnih slojeva, a programska implementacija dostupna je unutar priloga. Metodom su ostvarene visoke točnosti klasifikacije koje dostižu maksimalnu vrijednost od 1.00. Kroz evaluaciju je prikazan i utjecaj kvalitete i kvantitete podatkovnog skupa na konačnu točnost klasifikacije.

Evaluacija je provedena na sedam vrsta grafikona, čije su oznake klasa koje će se koristiti u daljnjem tekstu označene unutar zagrada: kružnom (A), horizontalnom stupčastom (B), horizontalnom kutiji (C), linijskom (D), točkastom (E), vertikalnom stupčastom (F), vertikalnom kutiji (G) i zasebnoj klasi koja ne sadrži slike grafikona (H). Dodatno, u okviru istraživanja predstavljen je algoritam za predobradu slika grafikona, a koji radi na razini točke razlučivosti (engl. *Pixel*). Osnovna svrha algoritma je generiranje slike koja sadrži smanjenu, odnosno minimalnu, količinu informacija neophodnu za daljnje učenje i testiranje neuronskih mreža. Algoritam za predobradu slike temelji se na pažljivo odabranim filtrima i parametrima za obradu slika koji omogućuju identifikaciju ključnih vizualnih elemenata na slikama grafikonima. Ovaj pristup ima dvostruki cilj: smanjenje složenosti slike kako bi se olakšalo učenje neuronskih mreža i naglašavanje važnih informacija koje su relevantne za analizu i interpretaciju grafikona. Izlaz iz algoritma je crno-bijela slika koja zadržava osnovne strukturne karakteristike grafikona, eliminirajući suvišne detalje i šum koji ne pridonose značajnoj informacijskoj vrijednosti. Smanjenje informacija čini ulazne podatke pogodnima za daljnje procese učenja, pridonoseći učinkovitosti i točnosti neuronskih mreža u analizi grafikona.

2.1 Uvod

U digitalnom okruženju, kada se grafikon izrađuje pomoću programskih alata poput Microsoft Excel-a, Matlab-a, D3-a, Plotly-a ili se nacrtaju ručno, pojedinačni elementi i vrijednosti grafikona ostaju pristupačni i mogu se dalje mijenjati i obrađivati. Međutim, kada se grafikon pohrani ili digitalizira kao jedinstvena vizualna cjelina (rasterska slika), strukturne informacije o grafikonu postaju nedostupne korisnicima i računalnim algoritmima koji žele izravno pristupiti pojedinačnim elementima grafikona, kao što su izvorne tablične vrijednosti, osi, legenda, naslovi ili drugi grafički simboli (primjerice stupci, linije ili krugovi). S obzirom na široku primjenu grafikona u različitim područjima, postoji potreba za razvojem računalnih metoda koje omogućavaju dobivanje i obradu informacija iz takvih pohranjenih rasterskih slika. Ove metode trebaju biti sposobne prepoznati i izdvojiti ključne elemente grafikona te omogućiti njihovu daljnju analizu i interpretaciju. Navedena tehnološka potreba proizlazi iz važnosti grafikona u suvremenom digitalnom dobu te činjenice da se sve veći broj informacija prenosi vizualnim putem.

Klasifikacija slike je dobro izučeni proces u području računalnog vida, koji se odnosi na određivanje pripadnosti slike prema vizualnom sadržaju. To je jedan od najkritičnijih zadataka u digitalnoj obradi slike. Kako elementi grafikona nisu strogo standardizirani, odnosno ne postoji

definicija koja određuje obavezne elemente grafikona, nove vrste i tipovi grafikona nastaju svakog dana, kao što je prikazano u [15]. U vrste grafikona ubrajaju se linijski, stupčasti, kružni, točkasti i drugi grafikoni, dok se tipovi odnose na posebne izvedbe grafikona namijenjene određenom zadatku. Na primjer, neki od tipova stupčastog grafikona uključuju vertikalne, horizontalne, složene, grupirane i druge varijacije. Neograničen broj vrsta i tipova grafikona stvara dodatni izazov kod klasifikacije slike grafikona. Tijekom godina, istraživači unutar ovog područja koristili su razne metode za klasifikacije slike grafikona, od metoda koje provjeravaju prisutnost elemenata i grafičkih simbola do metoda koje uključuju upotrebu strojnog učenja.

Algoritmi strojnog učenja predstavljaju ključno područje umjetne inteligencije, omogućujući računalima učenje iz podataka i primjenu stečenog znanja na različitim zadacima. Tradicionalni pristupi strojnog učenja uključuju metode koje se temelje na simboličkom predstavljanju znanja i zaključivanja, kao što je prikazano u [16]. Suvremeni pristupi obuhvaćaju razne arhitekture neuronskih mreža koje su najčešće usmjerene na rješavanje složenih zadataka kao što je klasifikacija slika. Prema [17–19], umjetna neuronska mreža je statistički model koji se sastoji od međusobno povezanih obradbenih elemenata, jedinica ili čvorova i obrađuje podatke prema dostavljenim ulaznim vrijednostima. Konvolucijska neuronska mreža (engl. *Convolutional Neural Network – CNN*), vrsta umjetnih neuronskih mreža smatra se jednom od najboljih tehnika za obradu slikovnog prikaza, a tako i slika grafikona, kao što je prikazano u [7,9,10,20]. U sklopu ovog istraživanja, za klasifikaciju slike grafikona, koriste se posebne izvedbe CNN-a: plitka konvolucijska neuronska mreža (engl. *Shallow Convolutional Neural Network – SCNN*) i sijamska plitka konvolucijska neuronska mreža (engl. *Siamese Shallow Convolutional Neural Network – SSCNN*).

U potpoglavlju 2.2 dan je povijesni pregled istraživanja iz područja klasifikacije slike grafikona. Potpoglavlje 2.3 daje opis korištenih podatkovnih skupova i predstavlja algoritam za predobradu slike grafikona. Arhitektura umjetnih neuronskih mreža predstavljena je u potpoglavlju 2.4, a dubinska analiza u potpoglavlju 2.5. Eksperimentalna evaluacija provedena na predstavljenim arhitekturama dostupna je u potpoglavljima 2.6 i 2.7. Zaključak je napisan u potpoglavlju 2.8.

2.2 Istraživanja u području klasifikacije slike grafikona

U ovom potpoglavlju dan je kratki povijesni pregled objavljenih radova iz područja klasifikacije slike grafikona, a prošireni pregled istraživanja u području objavljen je u [7]. Objavljeni radovi u području istraživanja razvrstani su prema korištenoj metodi za klasifikaciju slike grafikona. U prvom dijelu potpoglavlja predstavljene su metode, a u drugom dijelu nalazi se tablica koja uspoređuje ostvarene rezultate.

Tijekom godina istraživači su koristili različite metode za klasifikaciju slike grafikona. Metode je moguće grupirati u četiri kategorije:

- metode zasnovane na heuristici (engl. *Heuristic-based*),
- metode zasnovane na modelu (engl. *Model-based*),
- metode zasnovane na upotrebi stroja s potpornim vektorima (engl. *Support-Vector Machine – SVM*) i
- metode dubokog učenja (engl. *Deep learning*).

Metode zasnovane na heuristici zahtijevaju analizu i izdvajanje značajki slike grafikona na razini pojedinačnih točaka razlučivosti i grafičkih simbola. Metoda koristi predobradu slike kako bi pripremila sliku za daljnju analizu i izdvajanje značajki. Svi znanstveni radovi, neovisno o korištenoj metodi za klasifikaciju slike grafikona, koriste predobradu slike. Najosnovnije vrste predobrade slike su normalizacija rezolucije slike (engl. *Image resolution normalization*), normalizacija boje slike (engl. *Image color space normalization*) i smanjenje šuma slike (engl. *Image noise reduction*). Šum uzrokuje degradaciju slike, a javlja se prilikom nastanka, pretvorbe ili prijenosa slike. Najčešće je riječ o impulsnom šumu (primjerice sol i papar šum ili šum nasumičnih vrijednosti) koji značajno može promijeniti izgled slike i uzrokovati gubitak važnih informacija. Objavljeni radovi u području istraživanja koji koriste navedene vrste predobrade slike su [5,6,8,19,21–45]. Manji broj radova, [6,22,23,25,28,46–48], koristi napredne vrste predobrade slike koje omogućavaju razdvajanje tekstualnih od grafičkih elemenata slike. Napredne vrste predobrade slike su detekcija rubova (engl. *Edge-detection*), vektorizacija (engl. *Vectorization*) i segmentacija. Najjednostavniji rezultat segmentacije je binarna slika koja je pogodna za daljnju analizu, a posebno za klasifikaciju upotrebom neuronskih mreža. Segmentacija sadrži važne informacije koje određuju broj, oblik i položaj elemenata na slici. Svi elementi koji sadrže važnu

informaciju su jedne boje, a svi ostali elementi dijele boju pozadine. Postupak segmentacije omogućuje smanjenje količine informacija sadržanih u slici. Detekcija rubova omogućuje razlikovanje naglih promjena u svjetlini slike. Promjena svjetline najčešće označava promjenu dubine prostora, promjenu materijala ili promjenu scene. Rezultat detekcije rubova su objekti koje je moguće pojedinačno lokalizirati i klasificirati. Vektorizacija slike je proces koji se može koristiti nakon opisanih procesa, a omogućuje pretvorbu rasterske slike u vektorsku sliku. Vektorska slika omogućuje izdvajanje grafičkih oblika kao što su ravne linije i lukovi što dodatno poboljšava odvajanje važnih od nevažnih elemenata potrebnih za klasifikaciju slike grafikona.

Metode zasnovane na modelu za klasifikaciju slike grafikona temelje se na ručno izrađenim modelima. Model se izrađuje za svaku vrstu grafikona zasebno. Unutar modela je potrebno definirati ključne tekstualne i grafičke elemente koje određena vrsta grafikona treba sadržavati. Slika grafikona neće biti uspješno klasificirana ako su neki od važnih elemenata izostavljeni ili ako je njihova lokacija u prostoru slike drugačija od lokacije koja je definirana modelom. Model napisan za jednu vrstu grafikona neće uspješno raditi za pojedine tipove iste vrste grafikona te ga je potrebno dodatno prilagoditi. Objavljeni radovi u području istraživanja koji za klasifikaciju slike grafikona koriste metode zasnovane na modelima su [49,50].

Metode zasnovane na upotrebi SVM-a predstavljaju skup srodnih metoda strojnog učenja koje se koriste za klasifikaciju slike grafikona, kao što je prikazano u [51]. Značajke nastale upotrebom prethodno navedenih metoda moguće je koristiti kao ulazne vrijednosti SVM-a. Sve značajke postavljaju se u konačno-dimenzionalni prostor unutar kojeg se iscrtava hiperravnina. Hiperravnina dijeli značajke u različite klase te se još naziva i marginom ili granicom odluke. Sve značajke koje se nalaze s jedne strane margine će biti klasificirane kao jedna klasa, a sve značajke koje se nalaze s druge strane margine će biti klasificirane kao druga klasa. SVM-ovi ostvaruju uspješne rezultate i s malim podatkovnim skupovima, međutim, postavljanje margine je izazovno kada klase dijele značajke. Objavljeni radovi u području istraživanja koji koriste SVM za klasifikaciju slike grafikona su [6,47,48,52,53].

Klasifikacija slika grafikona započela je sredinom 2000. godine, a metode zasnovane na heuristici, modelu i upotrebi SVM-ova moguće je objediniti jednim nazivom — tradicionalne metode. U suvremenim istraživanjima, koja obuhvaćaju vremenski period od 2015. godine do danas, naglašena je uporaba metoda dubokog učenja, posebno CNN-ova za klasifikaciju slike grafikona. U radu [29], predstavljena je usporedba tradicionalnih metoda i CNN arhitektura te je vidljivo unaprjeđenje rezultata za 20%. Prvi put u području istraživanja su istaknute prednosti primjene CNN-a, uključujući poboljšane rezultate klasifikacije slike grafikona, mogućnost

korištenja prethodno učenih modela i jednostavnost primjene gotovih modela na novim zadacima. Međutim, istovremeno su istaknuti i nedostaci CNN-a, kao što su visoki računalni zahtjevi, nedostatak usporedivih javno dostupnih podatkovnih skupova i potreba za interpretacijom ostvarenih rezultata klasifikacije. Istraživači koriste razne CNN arhitekture, a najzastupljenije su LeNet (naziv po Yann LeCun i drugima, 1989.) koja je primijenjena u radovima [19,33], AlexNet (naziv po Alex Krizhevsky i drugima, 2012) koja je primijenjena u radovima [21,22,36,44], VGG (naziv po Visual Geometry Group, 2014) koja je primijenjena u radovima [5,8,22,25,27,29,34,36,41,42,44,45,54], GoogLeNet (naziv po Google, 2014) koja je primijenjena u radovima [22,28], ResNet (naziv po rezidualna neuronska mreža, engl. *Residual neural network*, 2015) koja je primijenjena u radovima [21,22,25,29,41,43,44,54–56], Inception (naziv po Szegedy i drugima, 2014) koja je primijenjena u radovima [23,24,29,41,44,54] i MobileNet (naziv po mobilna računalno-vizualna mreža, engl. *Mobile computer vision network*, 2017) koja je primijenjena u radovima [23,31,41,54]. Usporedba navedenih CNN arhitektura dostupna je u [22,41,44,57], a rezultat usporedbe prikazuje razliku ostvarenih rezultata do 5%.

Tablica 2.1: Pregled istraživanja u području klasifikacije slike grafikona. Istaknuti redovi predstavljaju radove izrađene u sklopu ovog istraživanja.

Lit.	God.	Metoda	Br. vrsta grafikona	Točnost klasifikacije	Podatkovni skup
[46]	2001	Heuristički zasnovana	2	0.83	mali
[50]	2011	Modelno zasnovana	5	0.92	mali
[49]	2011	Modelno zasnovana	5	0.92	mali
[47]	2011	SVM	10	0.81	srednji
[48]	2012	Heuristički zasnovana, SVM	3	0.97	mali
[58]	2012	Heuristički zasnovana, SVM	8	0.69 – 0.77	mali
[59]	2013	Heuristički zasnovana	3	0.96	srednji
[32]	2015	Duboko učenje	5	0.75	srednji
[21]	2016	Duboko učenje	7	0.84 – 0.86	veliki
[6]	2017	Duboko učenje, SVM	10	0.94	srednji
[33]	2017	Duboko učenje	10	0.70	srednji
[19]	2017	Duboko učenje	11	0.89	srednji
[28]	2017	Duboko učenje	10	0.91	srednji

Lit.	God.	Metoda	Br. vrsta grafikona	Točnost klasifikacije	Podatkovni skup
[60]	2018	Heuristički zasnovana	24	0.83 – 0.94	veliki
[24]	2018	Duboko učenje	2	0.89 – 0.93	mali
[22]	2018	Duboko učenje	5	0.99	veliki
[25]	2019	Duboko učenje	10	0.96	srednji
[26]	2019	Duboko učenje	28	0.88 – 0.93	veliki
[27]	2019	Duboko učenje	2	0.96	veliki
[55]	2019	Duboko učenje	7	0.88 – 0.99	veliki
[8]	2019	Duboko učenje	10	0.81	srednji
[31]	2020	Duboko učenje	14	0.92	srednji
[5]	2020	Duboko učenje	10	0.73 – 0.89	srednji
[37]	2020	Duboko učenje	1	0.93	srednji
[38]	2020	Duboko učenje	1	0.97	srednji
[61]	2020	Heuristički zasnovana, SVM	5	0.98	veliki
[36]	2020	Duboko učenje	16	0.4 – 0.7	srednji
[56]	2020	Duboko učenje	15	0.93 – 1.00	veliki
[42]	2021	Duboko učenje	1	0.85	srednji
[44]	2021	Duboko učenje	9	0.87 – 0.98	veliki
[12]	2021	Duboko učenje	7	1.00	srednji
[54]	2021	Duboko učenje	25	0.70 – 0.90	srednji
[13]	2022	Duboko učenje	7	0.99 – 1.00	srednji
[62]	2023	Duboko učenje	3	0.98	veliki
[14]	2024	Duboko učenje	7	0.00 – 1.00	srednji

Usporedba rezultata iz tablice 2.1 je izazovna budući da svi autori ne navode ili ne koriste javno dostupne podatkovne skupove, ne definiraju organizaciju podatkovnog skupa i ne navode korištenu programsku podršku i računalne resurse. Iako neki od autora koriste postojeća, gotova rješenja za klasifikaciju slike grafikona, prilagodili su ulazne parametre kako bi odgovarali njihovim potrebama i podatkovnom skupu. Podatkovni skup je najvažniji ulazni podatak koji omogućuje usporedbu različitih metoda. U tablici 2.1, podatkovni skupovi su svrstani u tri

kategorije: mali (broj korištenih slika je manji od 1000), srednji (broj slika je između 1000 i 10000) i veliki (broj slika je veći od 10000). Budući da se svi korišteni podatkovni skupovi kvalitativno i kvantitativno razlikuju, predstavljene okvirne brojke treba promatrati kao dodatnu informaciju, a ne kao referentnu točku u kojoj je za postizanje X rezultata, uz Y metodu, potrebna veličina podatkovnog skupa Z . U rijetkim slučajevima, autori upotrebljavaju iste javno dostupne podatkovne skupove i navode organizaciju podatkovnog skupa, ali čak i u tim slučajevima upotrijebljene slike su i dalje nepoznate.

CNN je moguće koristiti samostalno ili u kombinaciji sa SVM-om, gdje se CNN koristi za izvoz značajki, a SVM kao klasifikator, kao što je prikazano u [24–26,31]. Bilo koji način upotrebe CNN-a rezultira vrlo visokom točnošću klasifikacije. Posljednja tri rada izrađena u sklopu ovog istraživanja za evaluaciju koriste natjecateljski podatkovni skup definiran u [55]. Osim navedenog natjecanja, održana su još dva natjecanja sličnog tipa. Na posljednjem natjecanju, čiji konačni rezultati u trenutku pisanja još nisu dostupni, po prvi puta za klasifikaciju slika grafikona upotrijebljen je transformator vida (engl. *Visual Transformer*). U usporedbi s CNN arhitekturama, transformator vida ostvario je nižu točnost klasifikacije slike grafikona. S obzirom na visoku točnost klasifikacije koju postižu transformatori vida u drugim znanstvenim područjima, očekuje se značajan proboj ovih tehnika i u područje klasifikacije slike grafikona.

2.3 Podatkovni skupovi

Predobrada slike je važan korak koji je potrebno napraviti prije procesa učenja neuronske mreže. U radu [5], koji se bavi klasifikacijom slike grafikona upotrebom CNN-a dokazano je da predobrada slike značajno utječe na ishod učenja neuronske mreže. U radovima [63–65] istraživanje o važnosti predobrade slike provedeno je na raznim CNN arhitekturama i podatkovnim skupovima te je vidljivo povećanje točnosti klasifikacije slike grafikona. Empirijski rezultati pokazuju kako CNN može ostvariti veću točnost klasifikacije slike grafikona sa smanjenim podatkovnim skupom na kojem je napravljena predobrada slike, ali također, invazivna predobrada slike može rezultirati smanjenom točnošću klasifikacije slike grafikona zbog uklanjanja potencijalno značajnih informacija. U objavljenom istraživanju [5], detaljno je opisana primjena tri različite metode predobrade slika, koje su označene kao blaga, srednja i agresivna, na jednom podatkovnom skupu. Primjenom agresivne metode predobrade slike, neuronska mreža ostvarila je povećanje točnosti klasifikacije slike grafikona za 10%. Korištenjem agresivne metode značajno je smanjena količina grafičkih i tekstualnih elemenata koje slika sadrži, odnosno smanjenje su

sadržajne informacije. Algoritam za predobradu slike grafikona, u izvornom obliku, prvobitno je objavljen u radu [5], nakon toga dodatno prilagođen u [12] i primijenjen u [13,66].

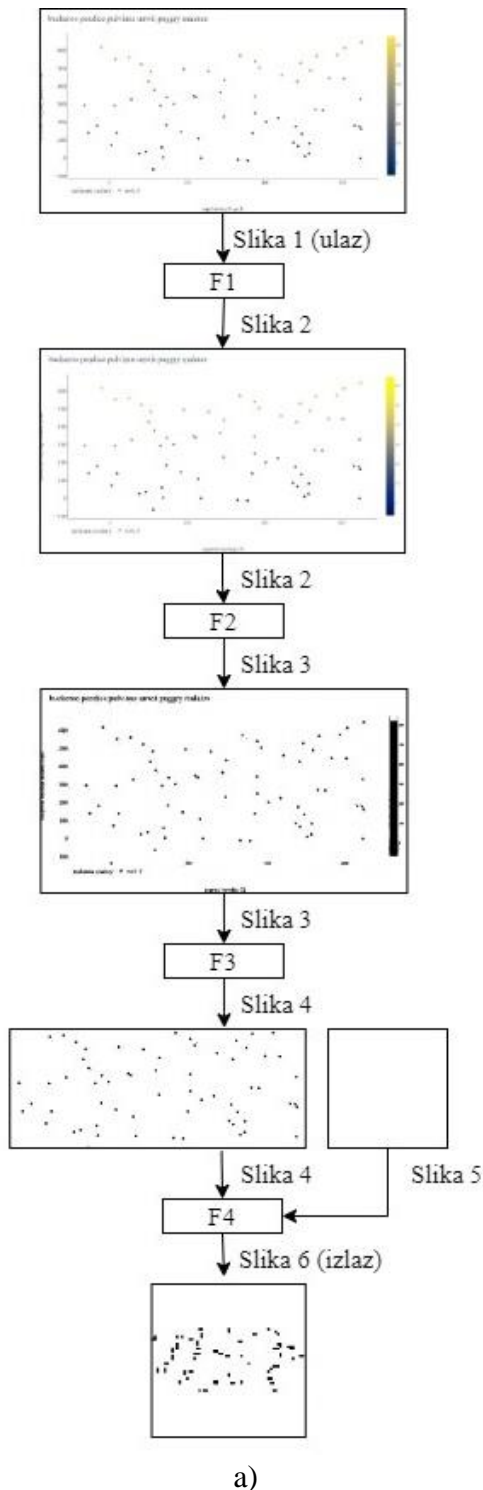
U potpoglavlju 2.3.1 detaljno je prezentiran korišteni algoritam za predobradu slike grafikona koji se primjenjuje na podatkovnim skupovima prikazanim u potpoglavlju 2.3.2.

2.3.1 Algoritam predobrade slike grafikona

Grafikon je prema strukturi heterogena vizualna cjelina koja sadrži tekstualne i grafičke elemente. Iz predstavljenih rezultata istraživanja [7] nije potrebno obraditi i tekstualne i grafičke elemente kako bi se postigla točna klasifikacija slike grafikona. Brojni objavljeni radovi zaključak o klasi grafikona donose na temelju grafičkih elemenata [5,8,12,13,23,26,29,30,46,60], ali postoje i radovi koji analiziraju i tumače tekstualne elemente prije donošenja konačne odluke o klasi grafikona [6,25,28,47].

U sklopu ovog istraživanja, zaključak o klasi grafikona donosi se isključivo korištenjem grafičkih elemenata. Za postizanje što veće točnosti klasifikacije CNN-a, na ulaznoj slici grafikona provodi se agresivna metoda predobrade slike grafikona, prikazana slikom 2.1. Slika prikazuje rad algoritma na slici točkastog grafikona (označeno kao a) i pseudokod algoritma (označeno kao b).

Algoritam predobrade slike računalno stvorenih slika grafikona sastoji se od četiri bloka predstavljena oznakama F1, F2, F3 i F4. Unutar svakog bloka nalazi se grupacija filtera korištenih u procesu obrade slike. Na ulazu algoritma nalazi se umjetno stvorena slika grafikona. Slika s ulaza prosljeđuje se na prvi blok F1 unutar kojeg se vrši izoštravanje i pojačavanje grafičkih elemenata slike. Za navedeno koriste se tri filtra: kontrast, gama-korekcija i zasićenost boja, čije vrijednosti su izložene unutar pseudokoda. Cilj prvog bloka je istaknuti grafičke elemente i simbole koji su najčešće prekriveni bojom, kako bi u drugom bloku F2 mogli biti dalje obrađeni. U drugom bloku, slika iz prethodnog bloka se pojednostavljuje, odnosno količina informacija koju izvorna slika sadrži, se značajno smanjuje. Za smanjenje informacija na slici koriste se tri filtra: gama-korekcija, smanjenje broja boja i medijan filter. Cilj drugog bloka je ukloniti boje sa slike grafikona i smanjiti šum. Boja kao osnovni grafički element grafikona nije potrebna za klasifikaciju grafikona i kao takva ne sadrži značajnu informaciju koja utječe na klasu grafikona. Bilo koja klasa grafikona može sadržavati bilo koju boju, a kružni grafikonu najčešće sadrže i nekoliko desetaka boja. Crno-bijela slika iz drugog bloka prosljeđuje se u treći blok F3. Unutar trećeg bloka sa slike se uklanjaju rubni elementi. Empirijskom analizom umjetno stvorenih slika grafikona, rub slike grafikona najčešće sadrži koordinatne osi, legendu ili opisne oznake. Navedeni elementi ne sadrže značajnu informaciju za klasifikaciju slike grafikona na osnovi grafičkih elemenata, te je iste potrebno



FUNKCIJA F1(slika, kontrast, gamaKorekcija, zasićenost):

```
slika = primjeniKontrast(slika, kontrast)
slika = primjeniGamaKorekcija(slika, gamaKorekcija)
slika = primjeniZasićenost(slika, zasićenost)
```

ISPISI slika

VRATI slika

FUNKCIJA F2(slika, gamaKorekcija, brojBoja, median):

```
slika = primjeniGamaKorekcija(slika, gamaKorekcija)
slika = primjeniBoje(slika, brojBoja)
slika = primjeniMedian(slika, median)
```

ISPISI slika

VRATI slika

FUNKCIJA F3(slika, vrijednost):

```
slika = primjeniIzreziPostotak(slika, vrijednost)
```

ISPISI slika

VRATI slika

FUNKCIJA F4(slika, slikaPrazna, veličina):

```
slika = primjeniRazlučivost(slika, veličina)
```

```
slika = zalijepi(slika, slikaPrazna)
```

ISPISI slika

VRATI slika

Postavljanje ulaza

```
slika1 = "put/do/slike"
slika5 = stvoriSliku(64)
```

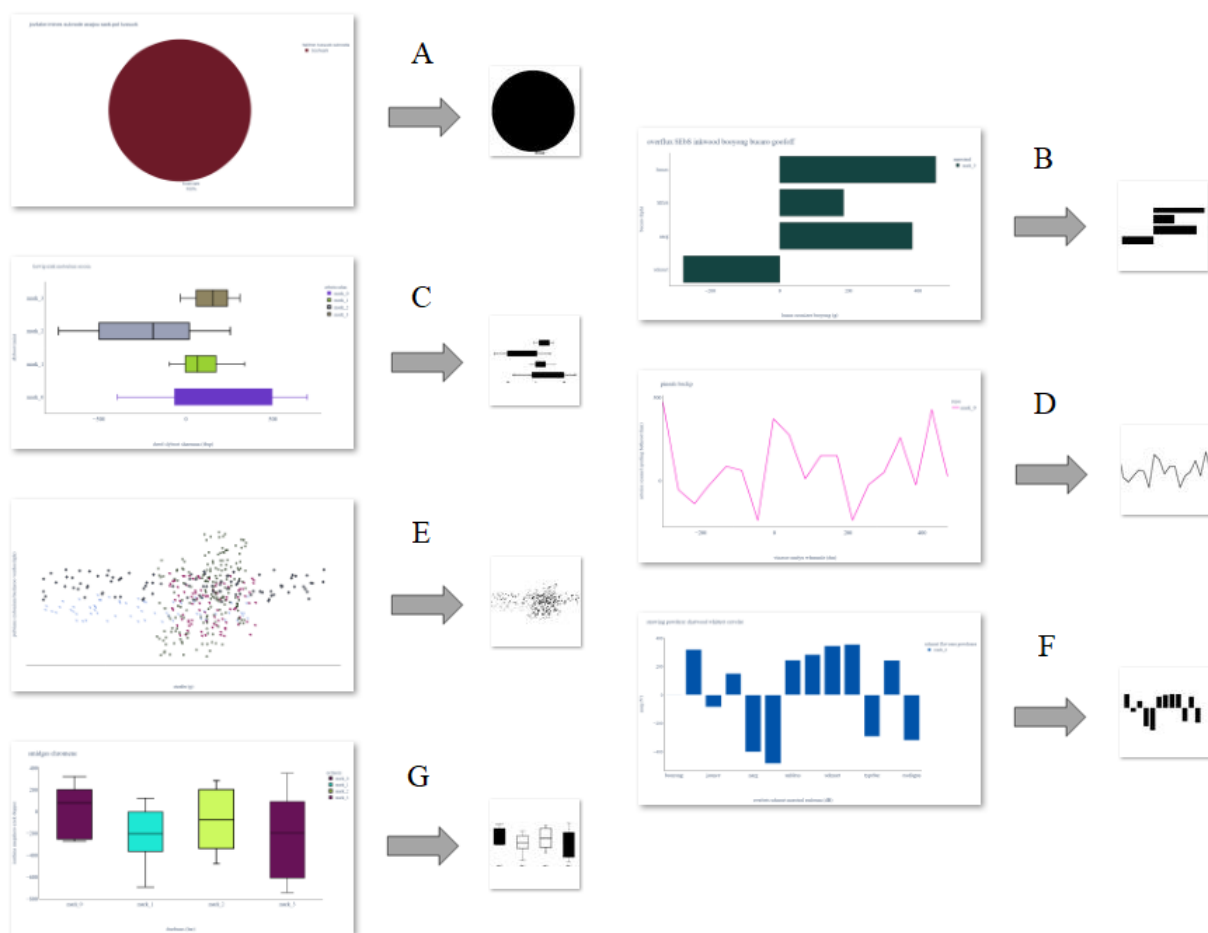
Pozivanje funkcija s parametrima

```
slika2 = F1(slika1, 60, 1.20, 90)
slika3 = F2(slika2, 0.01, 1, 3)
slika4 = F3(slika3, 10)
slika6 = F4(slika4, slika5, 64)
```

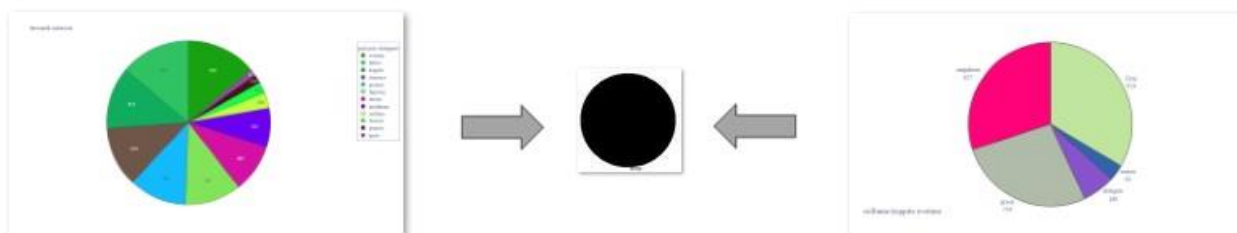
b)

Slika 2.1: Algoritam predobrade slike grafikona: a) prikaz rada algoritma na slici točkastog grafikona, b) pseudokod algoritma

ukloniti. U posljednjem bloku F4, slika se priprema za ulaz u CNN koja zahtijeva sliku grafikona razlučivosti od 64×64 točke. Izlazna rezolucija algoritma za predobradu slike i broj boja jednaki su ulaznim zahtjevima CNN-a. Izlazna slika grafikona zadržava originalni odnos visine i širine, a prazan prostor se popunjava s bijelom pozadinom koja ne sadrži značajnu informaciju. U primjeru

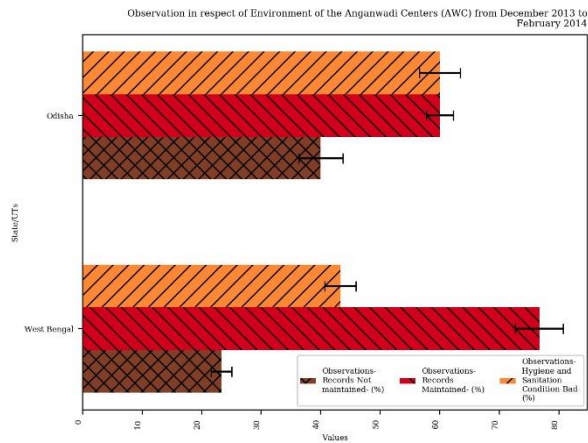


Slika 2.2: Primjeri ulaznih i izlaznih slika predstavljenog algoritma za sedam klasa grafikona (A-G). Navedene slike sadržane se unutar ChartDS podatkovnog skupa.

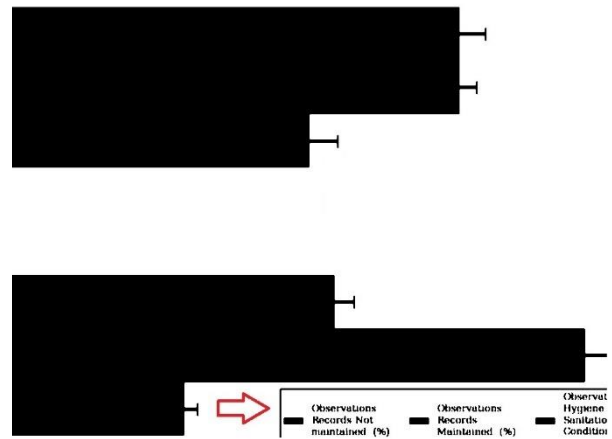


Slika 2.3: Pojava ujednačavanja različitih slika iste vrste i istog tipa grafikona.

koji je prikazan slikom 2.1, ulazna slika grafikona sadrži elemente poput naziva koordinatnih osi, vrijednosti na osima, naziva na osima, legende s pripadajućim vrijednostima i višebojnih grafičkih simbola, a izlazna slika grafikona sadrži samo jednobojne grafičke simbole. Na ulaznoj slici grafikona nalazi se ukupno jedanaest elemenata, a izlazna slika grafikona sadrži jedan element. Omjerom broja elemenata sadržanih na izlaznoj slici grafikona i broja elemenata sadržanih na ulaznoj slici grafikona, algoritam je u navedenom primjeru smanjio količinu informacija za više



a)



b)

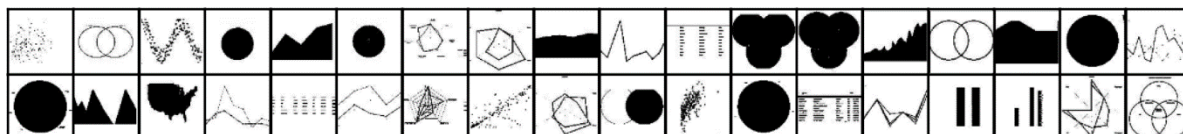
Slika 2.4: Prikaz ograničenja rada algoritma predobrade slike grafikona. Na slici a) slika iz podatkovnog skupa ICDAR7 i b) izlazna slika predstavljenog algoritma.

od 90%. Količina smanjenja informacija ovisi o vrsti i tipu grafikona i nije konstantna vrijednost. U slučaju kružnih grafikona, koji ne sadržavaju elemente vezane za koordinatne osi, smanjenje informacija je najčešće oko 50%. Slikom 2.2 prikazani su primjeri ulaznih i izlaznih slika predstavljenog algoritma za sedam klasa grafikona koji pripadaju ChartDS podatkovnom skupu. Upotrebom predstavljenog algoritma može doći do pojave ujednačavanja različitih slika istog tipa grafikona. Pojava je najizraženija na kružnim grafikonima, prikazana slikom 2.3. Za dvije različite ulazne slike istog tipa grafikona stvaraju se dvije izlazne slike s jednakom informacijom. Navedena pojava može povećati točnost klasifikacije modela jer postoji veća vjerojatnost da se slična slika već nalazila u podatkovnom skupu za učenje, ali isto tako, moguće je i smanjenje točnosti klasifikacije zbog učenja na primjerima koji ne posjeduju znatnu razliku u količini informacija. Algoritam predobrade slike grafikona sadrži i poznata ograničenja koja su predstavljena slikom 2.4. Slika prikazuje ulaznu sliku stupčastog grafikona iz podatkovnog skupa ICDAR7 (označeno kao a) i izlaznu sliku predstavljenog algoritma (označeno kao b). Prvo ograničenje je da ulazna slika ne smije sadržavati legendu u neposrednoj blizini grafičkih simbola. Drugo ograničenje odnosi se na istaknute linije rešetke. Ako su linije rešetke grafikona naglašene (primjerice podebljane i tamne boje), iste neće biti uspješno uklonjene. Posljednjim ograničenjem je određeno da ulazna slika ne smije sadržavati trodimenzionalni oblik grafikona niti bilo kakve trodimenzionalne efekte (primjerice sjene ili svjetla).

2.3.2 Opis korištenih podatkovnih skupova

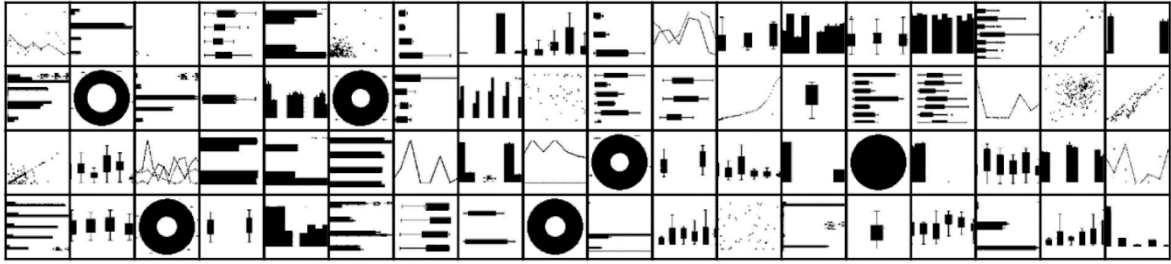
Podatci su ključna komponenta u području strojnog učenja. Bez dovoljne količine podataka, modeli strojnog učenja ne mogu učiti obrasce i odnose među podacima. Osim količine podataka, značaju ulogu u učenju modela strojnog učenja ima i kvaliteta podataka. U području strojnog učenja podatci se najčešće dijele u tri skupine: podatci za učenje, podatci za vrednovanje i podatci za testiranje modela. Podatci za učenje su podatci na osnovu kojih model uči značajke. Podatci za vrednovanje su podatci koji se koriste za periodične provjere modela i ažuriranje parametara za vrijeme učenja. Podatci za testiranje su podatci koji nisu sudjelovali u procesu učenja, odnosno podatci na kojima se provjerava uspješnost modela. Za potrebe učenja, vrednovanja i testiranja CNN modela koji će biti prikazani u potpoglavlju 2.4 primijenjeno je nekoliko različitih podatkovnih skupova.

ReVision [47] je podatkovni skup nastao 2011. godine spajanjem dva tada dostupna podatkovna skupa te je nadopunjen dodatnim slikama prikupljenim putem Google pretrage. Od tada, ReVision je korišten kao glavni podatkovni skup za usporedbu rezultata iz područja digitalne obrade grafikona. Nažalost, podatkovni skup u originalnoj veličini više nije javno dostupan te je korišten manji podskup. U sklopu ovog istraživanja, ReVision podatkovni skup koristi se za testiranje CNN modela te je u potpunosti obrađen algoritmom za predobradu slike iz potpoglavlja 2.3.1. Primjer slika sadržanih u ReVision podatkovnom skupu prikazan je slikom 2.5.



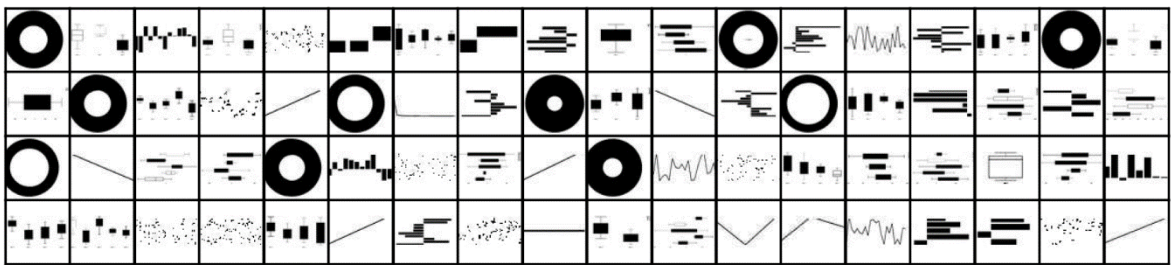
Slika 2.5: Primjer slika sadržanih unutar ReVision podatkovnog skupa.

ICDAR7 [55] je umjetno stvoreni podatkovni skup izrađen upotrebom Python Matplotlib programskog alata. Podatkovni skup je nastao u sklopu Međunarodne konferencije o analizi i prepoznavanju dokumenata (engl. *International Conference on Document Analysis and Recognition – ICDAR*) za potrebe natjecanja iz područja prikupljanja informacija i izvornih tablica iz vizualnih prikaza (engl. *Competition on harvesting answers and raw tables from infographics*). U sklopu ovog istraživanja, ICDAR7 podatkovni skup koristi se za učenje, vrednovanje i testiranje CNN modela te je u potpunosti obrađen algoritmom za predobradu slike iz potpoglavlja 2.3.1. Primjer slika sadržanih u ICDAR7 podatkovnom skupu prikazan je slikom 2.6.



Slika 2.6: Primjer slika sadržanih unutar ICDAR7 podatkovnog skupa.

ChartDS [67] je umjetno stvoreni podatkovni izrađen upotrebom Python Plotly programskog alata. Podatkovni skup je objavljen u [68] i javno je dostupan u podatkovnim repozitorijima Figshare [67], Srce Dabar [69] i Mendeley Data [70]. Sadrži slike grafikona u visokoj razlučivosti koje su organizirane i označene po vrstama grafikona. Svaka slika grafikona dodatno sadrži i opisni tekstualni dokument s izvornim vrijednostima. U sklopu ovog istraživanja, ChartDS podatkovni skup koristi se za učenje, vrednovanje i testiranje CNN modela te je u potpunosti obrađen algoritmom za predobradu slike iz potpoglavlja 2.3.1. Primjer slika sadržanih u ChartDS podatkovnom skupu prikazan je slikom 2.7.



Slika 2.7: Primjer slika sadržanih unutar ChartDS podatkovnog skupa.

Linnaeus [71] je podatkovni skup nastao prikupljanjem slika upotrebom raznih internet tražilica. Slike se koriste kao zasebna klasa, odnosno kao klasa koja ne sadrži slike grafikona te provjeravaju mogu li modeli neuronskih mreža prepoznati grafikon od ostalih slika koje bi se mogle naći u okruženju grafikona. Kako ne bi došlo do gubitka sadržajne informacije, slike su obrađene upotrebom blage predobrade koja uključuje prilagodbu razlučivosti i smanjenje broja boja. Primjer slika sadržanih u Linnaeus podatkovnom skupu prikazan je slikom 2.8.



Slika 2.8: Primjer slika sadržanih unutar Linnaeus podatkovnog skupa.

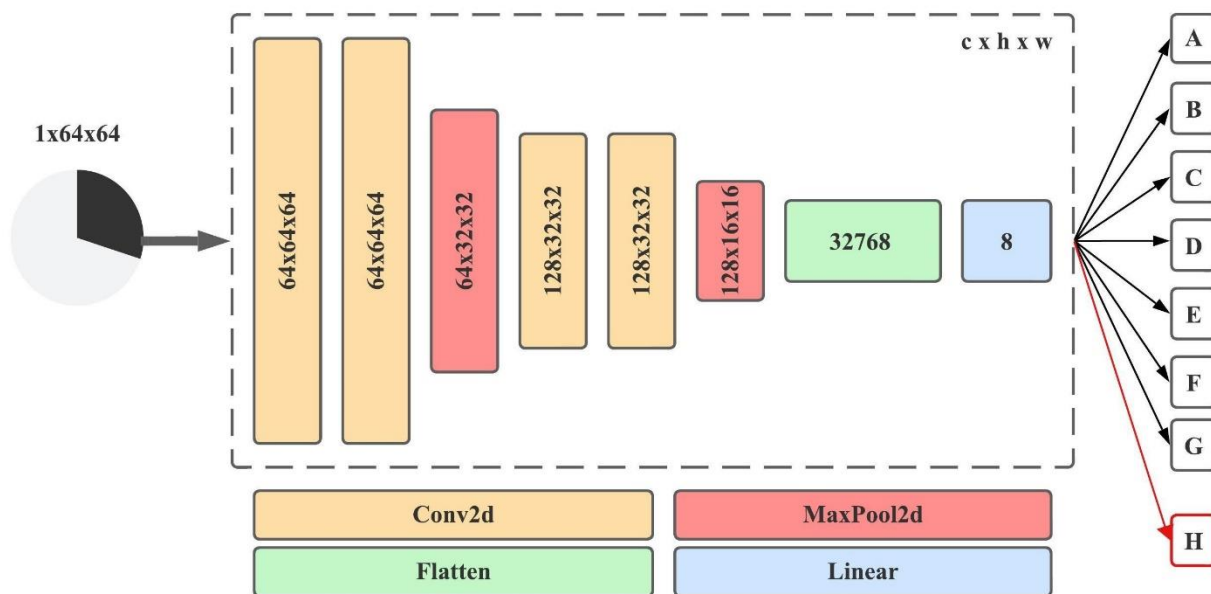
2.4 Arhitektura umjetne neuronske mreže

Svakog dana nastaju iznimne količine podataka koje su dostupne u digitalnom obliku i koje je potrebno obrađivati. Potreba za digitalnom obradom podataka potaknula je razvoj računalnih resursa koji su usko povezani s razvojem dubokih neuronskih mreža. Promatrajući arhitekturu neuronske mreže vidljiva je struktura koja sadrži razne slojeve (primjerice ulazni, skriveni, izlazni sloj), a dubina se odnosi na broj slojeva kojih može biti od nekoliko desetaka do nekoliko stotina (primjerice arhitekture kao DenseNet i EffNet sadrže nekoliko stotina slojeva). Upotrebom dubokih konvolucijskih neuronskih mreža (engl. *Deep Convolutional Neural Network – DCNN*) moguće je rješavati probleme u svim područjima znanosti, a prema predstavljanim istraživanjima u potpoglavlju 2.2, neuronske mreže je moguće koristiti i za klasifikaciju slika grafikona. Kako bi DCNN ostvarila visoku točnost klasifikacije, potrebno je nekoliko tisuća slika po klasi grafikona, što u slučaju novih vrsta i tipova grafikona nije dostupno. Izazov nastaje kada se DCNN koristi s malim (ograničenim) podatkovnim skupom, javlja se pojava prenaučivosti (engl. *Overfitting*). Prenaučeni modeli predobro su naučili podatkovni skup dostupan za vrijeme učenja, što je rezultiralo niskim gubitkom vrijednosti za vrijeme učenja (engl. *Training Loss*) i niskom sposobnošću generalizacije, to jest kada se ovom modelu na ulaz dostavi testni podatkovni skup, model će ostvariti smanjenu točnost klasifikacije ulaznog podatka. S povećanjem dubine DCNN-a javlja se i pojava nestajućih ili eksplodirajućih gradijenata (engl. *Vanishing or exploding gradients*), više informacija u [74]. Uvođenjem dodatnih slojeva u model, gradijenti funkcije gubitka (*sigmoid*) približavaju se nuli. Rezultat je sporije učenje ili potpuni prestanak učenja DCNN-a. Povećanjem dubine modela paralelno raste potreba za povećanjem računalnih resursa i povećanjem podatkovnog skupa. Izrada primjerenog podatkovnog skupa je zahtjevan i dugotrajan posao koji uključuje postupke prikupljanja, analiziranja, organiziranja i označavanja podataka.

Kako bi se riješili predstavljeni izazovi i smanjila potreba za računalnim resursima u sklopu ovog istraživanja izrađena je nova metoda za klasifikaciju slika grafikona zasnovana na arhitekturi sa smanjenim brojem slojeva — SCNN.

2.4.1 Plitka konvolucijska neuronska mreža

Dostupna literatura ne definira što je to SCNN niti od koliko bi se slojeva trebala sastojati. Prema radovima [75,76], savršena SCNN sastoji se samo od ulaznog sloja, jednog skrivenog sloja i izlaznog sloja, što nije izvedivo kada je potrebno koristiti složene podatkovne skupove. Šira definicija SCNN-a je mreža s minimalnim brojem skrivenih slojeva namijenjena izvršavanju točno



Slika 2.10: Prikaz arhitekture plitke konvolucijske neuronske mreže za klasifikaciju ulazne slike u sedam klasa grafikona (A-G) i zasebnu klasu koja ne sadrži slike grafikona (H).

određenog zadatka, kao što je prikazano u [77]. Kada se razmatra dubina mreže, računaju se samo slojevi koji su skriveni između ulaza i izlaza neuronske mreže i sadrže parametre koje neuronska mreža može naučiti. Takvi slojevi nazivaju se težinskim slojevima, a njih čine konvolucijski i potpuno povezani slojevi. Arhitektura, odnosno prikaz slojeva, SCNN-a nastala u sklopu ovog istraživanja prikazana je slikom 2.10.

Arhitektura se sastoji od četiri konvolucijska sloja (engl. *Convolutional layer – Conv2d*), dva sloja s maksimalnim sažimanjem (engl. *Max-pooling layer – MaxPool2d*) i jednog potpuno povezanog sloja (engl. *Fully-connected layer – Linear*). Zbog potpuno povezanog sloja, ulazna slika grafikona je ograničena na razlučivost od 64×64 točke. Prije ulaska u model, slike su obrađene predstavljenim algoritmom za predobradu slike iz potpoglavlja 2.3.1. Zatim, prilagođena slika prolazi kroz dva bloka koja se sastoje od dva konvolucijska sloja i jednog sloja s maksimalnim sažimanjem. Napredovanjem slike kroz konvolucijske blokove neuronska mreža uči složenije uzorke što potvrđuje analiza provedena u potpoglavlju 2.5.

Konvolucijski sloj je temeljni sloj svake CNN, sastoji se od filtra i služi za izdvajanje značajki iz slike. U pravilu, filtri su manjih prostornih dimenzija od ulaza, a broj kanala odgovara ulazu. Filtri se kreću kroz receptivno polje i provjeravaju prisutnost značajki. Rezultat konvolucijskog sloja je dvodimenzionalna aktivacijska matrica koja predstavlja odziv filtra u prostoru slike. Po uzoru na temeljnu ideju VGG arhitekture, prikazanu u [78], korišteni su višestruki konvolucijski slojevi sa smanjenim filtrom receptivnog polja prije sloja maksimalnog

sažimanja. Receptivno polje je ograničeno područje vidnog polja unutar kojeg pojedinačni neuroni reagiraju na podražaje. Autori su dokazali korištenjem dvostrukih filtra receptivnog polja razlučivosti 3×3 točke ima jednaki učinak kao korištenje jednog filtra receptivnog polja razlučivosti 5×5 točki. Dvostruki filtar receptivnog polja razlučivosti 3×3 točke smanjuje broj težina, ali povećava broj slojeva u neuronskoj mreži, dubinu neuronske mreže. Smanjeni broj težina smanjuje potrebu za računalnim resursima, a povećana dubina pomaže neuronskoj mreži učiti složenije nelinearne značajke. Korišteni filtar receptivnog polja razlučivosti 3×3 točke je najmanji filtar koji može dohvatiti informacije i iz susjednih točaka. Parni filtri receptivnog polja razlučivosti 2×2 točke ili 4×4 točke ne uzimaju se u obzir jer stvaraju izobličenja unutar slojeva. Stoga, svi konvolucijski slojevi u predloženoj arhitekturi koriste filtar receptivnog polja razlučivosti 3×3 točke.

U prvom konvolucijskom bloku, konvolucijski korak je konstantna vrijednost koja iznosi jednu točku, a prostorna razlučivost je očuvana s ispunjavanjem okolnih vrijednosti u prostoru od jedne točke (ulazna visina i širina su očuvane). Dvodimenzionalne aktivacijske matrice značajki dobivene konvolucijskim slojevima su ovisne o lokaciji, što znači neuronska mreža uči stvarne lokacije značajki na slici u procesu učenja. Navedena pojava rezultira smanjenom točnošću klasifikacije. Uzimajući u obzir slike grafikona, lokacija značajki na slikama ne predstavlja značajnu informaciju, te zbog toga nakon konvolucijskog sloja primjenjuje se sloj maksimalnog sažimanja.

Sloj maksimalnog sažimanja provodi smanjenje dimenzionalnosti smanjenjem broja parametara na ulazu. Sloj odabire maksimalnu vrijednost točki u receptivnom polju koje je pokriveno filtrom. Prvi sloj maksimalnog sažimanja smanjuje količinu izdvojenih značajki kroz smanjivanje uzorkovanja. Sloj koristi filtar receptivnog polja razlučivosti 2×2 točke s korakom konstantne vrijednosti dvije točke. Veličina filtra maksimalnog sažimanja manja je od konvolucijskog filtra koji smanjuje veličinu izdvojenih značajki na pola (visina i širina su prepolovljene, ali je broj kanala očuvan). Smanjenje dimenzija izdvojenih značajki smanjuje broj težina u neuronskoj mreži.

Drugi konvolucijski blok sličan je prvom, osim što nije potrebno očuvati prostornu razlučivost, to jest prostorno ispunjavanje nije obavezno koristiti. Rubovi slike grafikona uglavnom sadrže bijeli prostor koji nije ispunjen značajnim informacijama, što znači da se dimenzije slike mogu dodatno smanjiti, čime se dodatno smanjuje broj težina u neuronskoj mreži. U radu [13] algoritam za predobradu slike grafikona ostavlja veću količinu praznog prostora na rubu slike te nije potrebno očuvati prostornu razlučivost. U sklopu ovog istraživanja, algoritam za

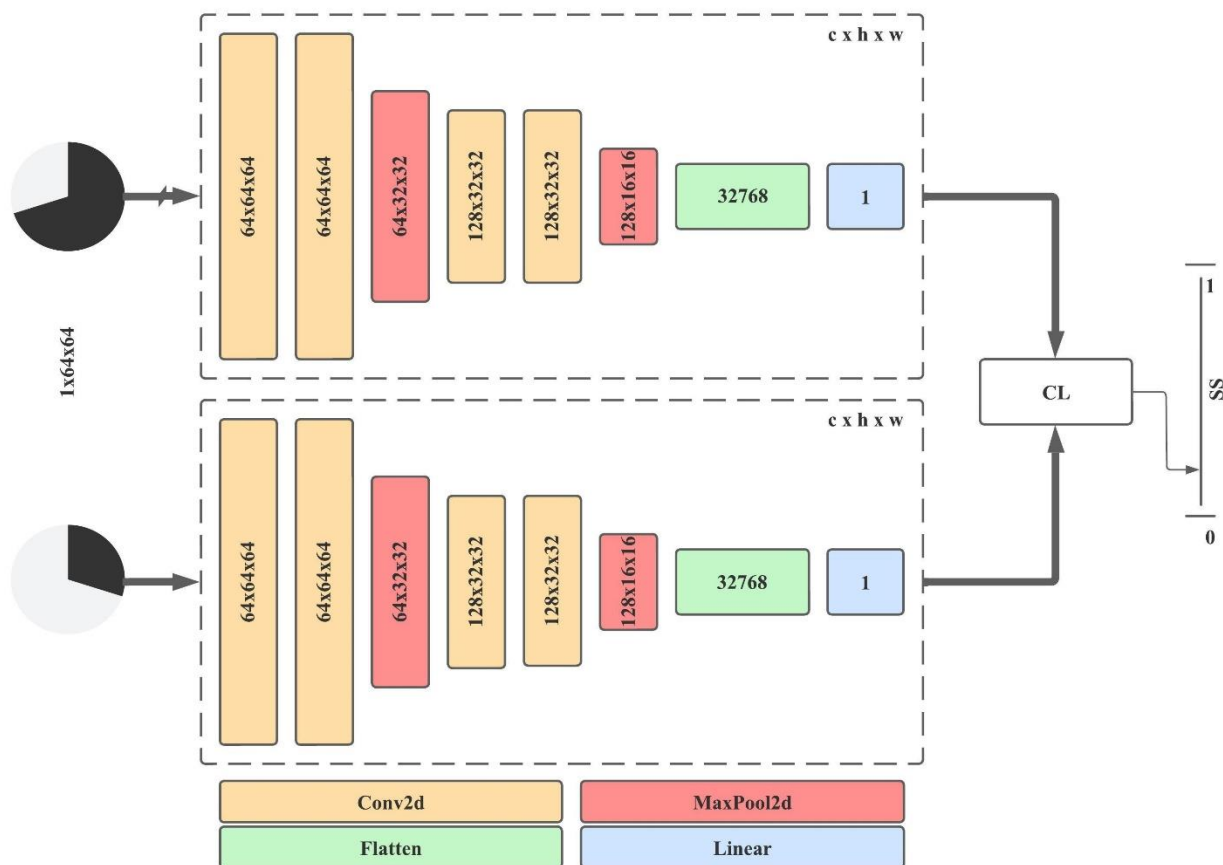
predobradu slike je dodatno poboljšana te je ispravno zadržati prostornu razlučivost, u suprotnom nastati će gubitak značajnih informacija. Na kraju bloka koristi se drugi sloj maksimalnog sažimanja, s istim parametrima kao i prvi sloj maksimalnog sažimanja. Filtar unutar sloja maksimalnog sažimanja radi na istom principu kao i filtar u konvolucijskom sloju, ali razlika je u tome što filtri u slojevima maksimalnog sažimanja ne sadrže težine. Izlaz sloja maksimalnog sažimanja povezan je s ulazom potpuno povezanog sloja.

Potpuno povezani sloj povezuje svaki ulaz iz prethodnog sloja sa svakim izlaznim neuronom i izravna ulaz u jedan vektor veličine $32768 \times 1 \times 1$. Navedeni broj predstavlja broj neurona izračunat množenjem broja kanala (engl. *Channels* - c), visine (engl. *Height* - h) i širine (engl. *Width* - w) iz prethodnog sloja ($128 \times 16 \times 16 = 32768$). Aktivacijska funkcija *softmax* koristi se za konačnu klasifikaciju na kraju predložene arhitekture. *Softmax* pretvara ulazne vrijednosti u raspon od 0 do 1 koje se zatim tumače kao vjerojatnosti. Navedeni model isporučuje vjerojatnosti za osam klasa (A-H). Zbroj svih vjerojatnosti ne može biti veći od 1.

Ako bi se navedeni model koristio nakon programskog alata za strukturnu analizu dokumenta, model bi uspješno klasificirao sedam klasa grafikona (A-G), a ostale prirodne slike raznolikog sadržaja koje se najčešće nalaze unutar istog dokumenta, bile bi izdvojene u zasebnu klasu (H). Arhitektura predstavljene SCNN implementirana je u Python programskom jeziku i nalazi se unutar priloga [P.1](#).

2.4.2 Sijamska plitka konvolucijska neuronska mreža

Sijamske neuronske mreže prvobitno su uvedene zbog izazova rješavanja verifikacije potpisa, kao što je prikazano u [79]. U prošlosti se nisu dodatno koristile zbog potrebe za znatnim računalnim resursima. Razvoj naprednih unificiranih računalnih resursa (engl. *Compute Unified Device Architecture* - *CUDA*) i različitih *CUDA* biblioteka omogućio je ubrzano vrijeme učenja ovih neuronskih mreža. Razvojem tehnologije, sijamske neuronske mreže postale su široko dostupne. Danas se često koriste u sigurnosnim aplikacijama, kao što je prepoznavanje lica [80], verifikacija potpisa [81], klasifikacija ulaznog signala [82], prepoznavanje govora [83] i drugo. Osnovne prednosti sijamskih neuronskih mreža uključuju ostvarivanje visoke točnosti klasifikacije i za klasu koja nije sudjelovala u procesu učenja, dijeljenje težina, objašnjive rezultate i ostvarivanje klasifikacije čak i kada je u procesu učenja sudjelovala samo jedna slika. Također, ističe se važnost kvalitete nad kvantitetom podatkovnog skupa. Međutim, postoji nekoliko nedostataka, uključujući povećanu potrebu za računalnim resursima u odnosu na tradicionalne arhitekture i potrebu za



Slika 2.11: Prikaz arhitekture sijamske pitke konvolucijske neuronske mreže za klasifikaciju ulazne slike grafikona.

odabirom funkcije gubitka. Pri učenju ovih neuronskih mreža, ključno je posvetiti posebnu pozornost kvaliteti slika pri izradi podatkovnog skupa, jer nagomilane slike same po sebi neće pružiti zadovoljavajuće rezultate točnosti klasifikacije.

SSCNN je podvrsta CNN-a gdje je arhitektura neuronske mreže izgrađena pomoću dvije identične neuronske mreže. SSCNN arhitektura sa slike 2.11 sadrži dvije identične SCNN prikazane u potpoglavlju 2.4.1. Obje SCNN sadrže istu konfiguraciju slojeva te posjeduju zajedničke težine i parametre. Iako neuronske mreže sadrže istu konfiguraciju slojeva, ulazna slika (ulazni vektor) je drugačija i na izlazu se stvaraju različiti vektori. Jedan izlazni vektor je unaprijed izračunat i stvara osnovicu za usporedbu u odnosu na drugi izlazni vektor koji će se uspoređivati.

Izlazni vektori će sadržavati jednake vrijednosti ako se na ulaze neuronskih mreža proslijede dvije identične slike grafikona. Izlazni vektori se koriste za izračun kontrastnog gubitka (engl. *Contrastive loss* – *CL*), funkcije gubitka koja izračunava euklidsku udaljenost između dva vektora. Kontrastni gubitak znatno kažnjava model ako je udaljenost između sličnih instanci ulaza velika ili ako je udaljenost između različitih instanci ulaza mala, učinkovito usmjeravajući model

učenju smislenog prostora unutar kojeg su slične instance ulaza blizu, a različite instance izdvojene. Udaljenost instanci ulaza određuje se hiperparametrom koji predstavlja graničnu vrijednost. U usporedbi s SCNN modelom, koji daje rezultat vjerojatnosti u rasponu od 0 do 1, rezultat izračuna kontrastnog gubitaka je koeficijent sličnosti (engl. *Similarity score – SS*) s vrijednošću u rasponu od 0 do 1. Za dvije identične slike grafikona, koeficijent sličnosti je 0. Navedeno je najosnovniji test kojim se provjerava ispravnost rada naučenog modela. Kada se dvije različite slike, na primjer dva različita kružna grafikona dovedu na ulaz modela, koeficijent sličnosti trebao bi biti bliži nuli, što ukazuje na istu klasu. S druge strane, ako se na ulaz modela dovedu slike primjerice kružnog i stupčastog grafikona, koeficijent sličnosti trebao bi biti dalje od nule, što ukazuje na različitu klasu. Arhitektura predstavljene SSCNN implementirana je u Python programskom jeziku i nalazi se unutar priloga [P.2](#).

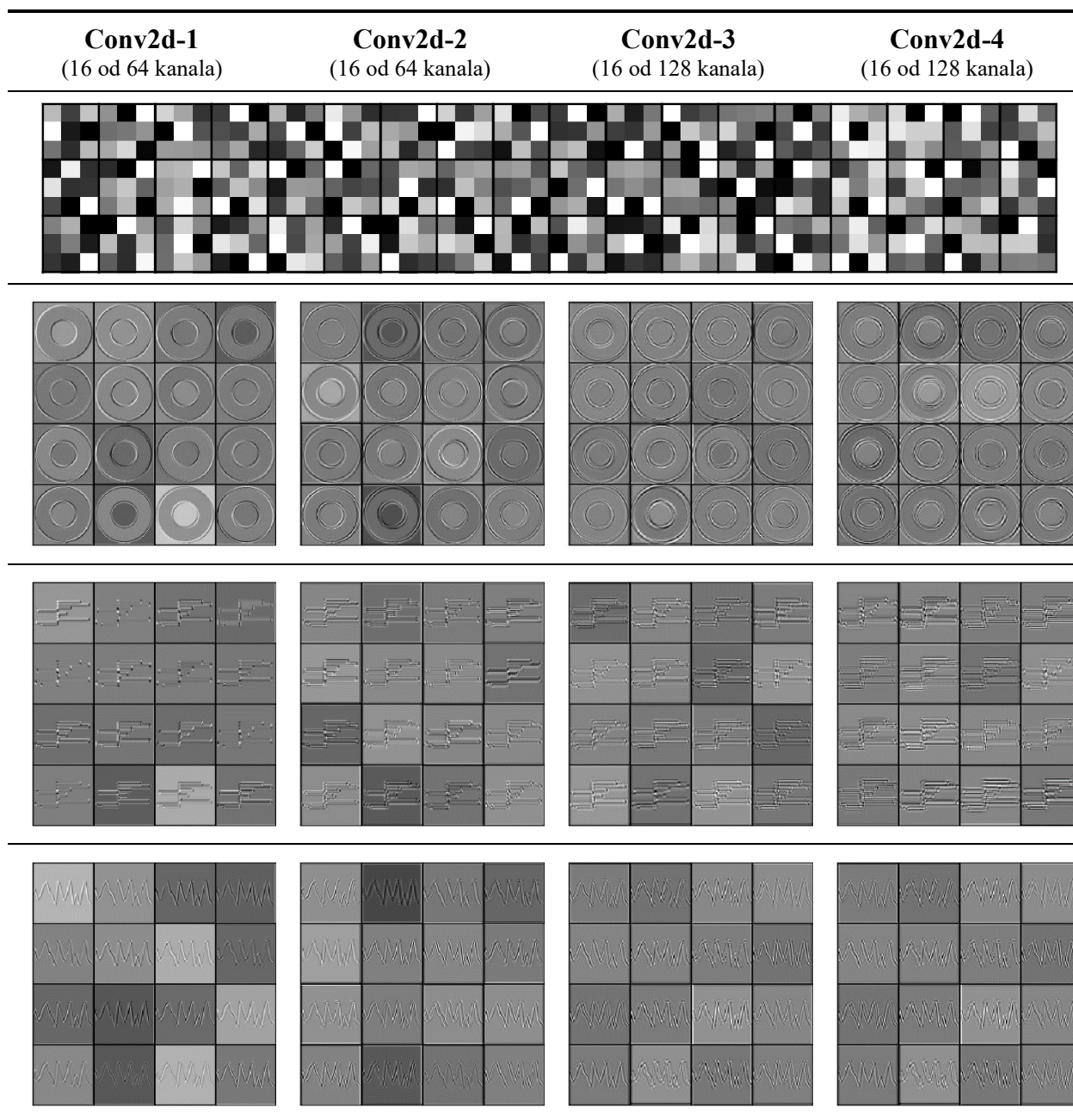
2.5 Dubinska analiza neuronske mreže

Koristeći neuronsku mrežu opisanu u prethodnim poglavljima, prije evaluacije na podatkovnom skupu, potrebno je napraviti dubinsku analizu neuronske mreže i provjeriti imaju li konvolucijski slojevi sposobnost razlikovanja važne od manje važne informacije sadržane u slici. Kod izrade dubinske analize neuronske mreže koristi se metoda izdvajanja značajki iz konvolucijskih slojeva. CNN-ovi uče apstraktne značajke i koncepte iz točaka razlučivosti.

Vizualizacija značajki (engl. *Feature Visualization*) slikovno prikazuje naučene značajke maksimiziranjem funkcije aktivacije. Vizualizacija značajki jednog neurona, kanala ili sloja u neuronskoj mreži postiže se pronalaženjem ulazne vrijednosti koja maksimizira aktivaciju željene jedinice, kao što je prikazano u [84,85]. Pojedinačni neuroni su osnovne građevne jedinice neuronske mreže i najtočnije informacije bi se dobile vizualizacijom značajki svakog neurona, što nije izvedivo. S druge strane, računalno jednostavniji i vremenski optimalan pristup je vizualizacija značajki kanala pojedinog sloja ili vizualizacija cijelog sloja neuronske mreže.

Značajke se sastoje od filtra receptivnog polja razlučivosti 3×3 točke i aktivacijskih matrica gdje svaka matrica odgovara određenom filtru. Svaki element u aktivacijskoj matrici značajki predstavlja aktivaciju određenog neurona u neuronskoj mreži, a njegova vrijednost predstavlja stupanj do kojeg je odgovarajuća značajka prisutna na ulaznoj slici. Broj aktivacijskih matrica značajki u konvolucijskom sloju je hiperparametar na kojeg je moguće utjecati tijekom modeliranja neuronske mreže. Povećanje broja aktivacijskih matrica značajki može pomoći u

Tablica 2.2: Napredak kružnog, stupčastog i linijskog grafikona kroz konvolucijske slojeve. U prvom retku nalazi se primjer filtra receptivnog polja, a u ostalim redovima primjer kanala svakog konvolucijskog sloja.



učenju složenijih značajki, ali također povećanje značajno utječe na potrebu za dodatnim računalnim resursima. Učenje CNN-a odvija se u težinskim, to jest konvolucijskim slojevima i svaki konvolucijski sloj uči drugi skup značajki.

U predstavljenoj SCNN arhitekturi nalaze se četiri konvolucijska sloja, s različitim brojem kanala, ali jednakim filtrom receptivnog polja razlučivosti 3×3 točke. Tablicom 2.2, prikazan je napredak kružnog, horizontalnog stupčastog i linijskog grafikona kroz konvolucijske slojeve

(označene kao Conv2d-1 do Conv2d-4). U prvom retku nalazi se nekoliko primjera korištenih filtara receptivnog polja, dok u ostalim redovima prikazane su izlazne slike grafikona gdje je prikaz ograničen na prvih 16 kanala po konvolucijskom sloju. Filtre receptivnog polja nije jednostavno tumačiti. Bijela boja označava veliku pozitivnu težinu, crna označava veliku negativnu težinu, a siva predstavlja nultu težinu. Velika većina prikazanih filtara sadrži sivu boju iz čega se može zaključiti da se na ulazu nalazi slika s malo sadržajnih informacija, odnosno objekata koje je potrebno promatrati. Navedeno je očekivano jer su sve slike predobrađene algoritmom predstavljenim u potpoglavlju 2.3.1 koji stvara slike sa značajno smanjenim brojem grafičkih i tekstualnih elemenata. Vrijednosti filtara se ponavljaju i jedinstvene su za svaki kanal konvolucijskog sloja. Prvi konvolucijski sloj (Conv2d-1) na ulaz prima sliku s jednim kanalom u kojem se nalaze 64 filtra. Drugi konvolucijski sloj (Conv2d-2) sadrži 4096 filtara jer na ulaz prima sliku iz prvog konvolucijskog sloja (Conv2d-1) koja sadrži 64 kanala. Svaki filter se provodi preko ulazne slike grafikona i stvara drugačiju izlaznu sliku grafikona koja omogućava neuronskoj mreži učenje drugačijih značajki. U prvom konvolucijskom sloju (Conv2d-1) neuronska mreža uči primitivne oblike (primjerice linije, lukove i rubove) kojima se pridjeljuje važnost i jasno su istaknuti. U drugom (Conv2d-2) i trećem (Conv2d-3) konvolucijskom sloju, neuronska mreža uči uzorke i teksture. Na slikama su vidljiva zadebljanja oko pronađenih granica što se može protumačiti kao učenje okoline. U posljednjem konvolucijskom sloju (Conv2d-4), neuronska mreža promatra objekte, elemente omeđene granicama. Kvaliteta slike u posljednjem koraku je očekivano najniža, gledano iz ljudske perspektive, no za neuronsku mrežu nalaze se važne informacije koje će biti proslijeđene u potpuno povezani sloj, kao što je prikazano u [86].

Predstavljena vizualizacija značajki daje jedinstven uvid u rad neuronske mreže. Kroz vizualizaciju značajki, vidljivo je da predstavljena neuronska mreža razlikuje jednostavne elemente koji opisuju klase grafikona i pomoću njih donosi odluke u klasifikaciji. Predobrada slike grafikona pozitivno utječe na uočavanje važnih primitivnih elemenata, dok se pozadina i boja ispunje objekata ne promatraju. Korištene ulazne slike grafikona su slike iz testnog podatkovnog skupa koje su uspješno klasificirane.

Arhitekturu mreža nije moguće uspoređivati temeljem vizualizacije značajki, ali moguće je upotrebom vremenske (engl. *Time complexity*) i prostorne (engl. *Space complexity*) složenosti. Prostorna složenost definirana je brojem parametara koji predstavljaju zbroj svih težina i predrasuda u neuronskoj mreži. Sloj maksimalnog sažimanja se ne uzima u obzir jer ne sadrži parametre već umjesto toga, sadrži hiperparametre. Vremenska složenost uključuje konvolucijske slojeve, slojeve maksimalnog sažimanja i potpuno povezane slojeve. Konvolucijski slojevi koriste

90% računalnog vremena, kao što je prikazano u [87]. Operacije množenja-akumulacije (engl. *Multiply-Accumulate operations – MACs*) predstavljaju temeljne matematičke operacije u izračunima neuronske mreže. Operacije uključuju množenje dvaju brojeva i zbrajanje rezultata. MAC-ovi se obično koriste za određivanje računalnog opterećenja konvolucijskih i potpuno povezanih slojeva u neuronskoj mreži. Broj MAC-ova je mjera broja množenja i zbrajanja dobivena tijekom obrade podatka u pojedinačnom sloju. Povećanjem broja MAC-ova povećava se računalno opterećenje, a veći broj matematičkih operacija zahtjeva veće vrijeme izvršavanja što dovodi do povećanja vremenske složenosti. Povećanjem broja konvolucijskih slojeva značajno se povećava prostorna i vremenska složenost. U predstavljenom području istraživanja, prikazanom tablicom 2.1, autori navode razne arhitekture neuronskih mreža s kojima na različitim podatkovnim skupovima ostvaruje vrlo visoke točnosti klasifikacije. Usporedba CNN arhitektura s obzirom na vremensku i prostornu složenost predstavljena je tablicom 2.3. Za izračun vremenske i prostorne složenosti korištena je javno dostupna biblioteka naziva *PyTorch-OpCounter* [88] primijenjena u [89,90].

Predstavljena SCNN sadrži najmanji broj težinskih slojeva s kojima uspješno obavlja klasifikaciju slika grafikona. S obzirom na smanjeni broj slojeva, značajno je smanjen broj parametara i vrijeme potrebno za učenje modela. SSCNN ima duplo veću prostornu i vremensku složenost jer je izrađen od dva SCNN modela, ali još uvijek značajno manje u usporedbi s drugim arhitekturama. Posljednji stupac tablice 2.3, procjena veličine modela, odnosi se na minimalne procijenjene računalne zahtjeve za radnom memorijom koje je potrebno zadovoljiti prije pokretanja učenog modela. Procjena se odnosi na prostor potreban za učitavanje parametara modela u radnu memoriju računala izračunat prema [91,92]. Izračun predstavlja umnožak visine, širine i broja kanala ulazne slike, broja parametara težinskih slojeva i količinu prostora potrebnog za pohranu jedne točke razlučivosti u radnu memoriju računala. Cjelokupna analiza je provedena korištenjem dobro uspostavljenih programskih okvira otvorenog koda, uključujući TensorFlow, Keras i PyTorch. Svi modeli su bili podvrgnuti učenju koristeći AMD EPYC 7B12 centralnu obradbenu jedinicu (engl. *Central processing unit – CPU*) s 13 GB dostupne memorije, uz podršku NVIDIA T4 grafičke obradbene jedinice (engl. *Graphics processing unit – GPU*) koja raspolaže sa 16 GB dostupne memorije. Obzirom na korišteno NVIDIA grafičko okruženje, prema [93], za pohranu jedne točke razlučivosti potrebno je 4 B. Za postupak učenja modela potrebna je znatno veća količina raspoložive radne memorije računala zbog učenja na seriji slika. Prema [92], za učenje SCNN-a na seriji koja sadrži 64 slike grafikona potrebno je minimalno 300 MB, za VGG16 8 GB, ResNet50 v2 5 GB, a za Inception v3 6 GB. Navedeno su sve procijenjene vrijednosti koje

Tablica 2.3: Usporedba prostorne i vremenske složenosti za razne CNN arhitekture. Istaknuti redovi predstavljaju arhitekture neuronskih mreža za klasifikaciju slike grafikona izrađene u sklopu ovog istraživanja.

Naziv	Broj parametara ($\times 10^6$)	MACs ($\times 10^9$)	Broj težinskih slojeva	Procjena veličine modela [MB]
SCNN	0.50	0.33	5	2
Pojednostavljeni VGG	29.78	0.47	7	120
AlexNet	61.10	0.77	8	250
SSCNN	1.01	0.67	10	4
VGG16	138.36	15.61	16	550
MobileNet v1	4.24	0.56	23	17
Inception v3	27.16	5.75	48	110
ResNet50 v2	25.56	4.14	50	100
Xception	42.80	16.80	71	170
DenseNet121	7.98	2.90	121	32
EffNetB0	5.28	0.42	237	21

promatraju parametre i arhitekturu neuronske mreže. Stvarne vrijednosti ovise o operacijskom sustavu, programskom okviru, upravljačkim programima, GPU arhitekturi, arhitekturi radne memorije i raznim metodama optimizacije i oslobađanja memorijskog prostora.

U tablici 2.4 prikazana su stvarna vremena potrebna za učenje odabranih arhitektura. Parametri korišteni tijekom faze učenja bili su nepromijenjeni među različitim modelima, čime se osigurava konzistentnost i ponovljivost rezultata dobivenih tijekom evaluacije. U svrhu usporedbe postignutih rezultata s drugim arhitekturama, za učenje korišten je ChartDS podatkovni skup s ukupno 500 slika po klasi grafikona, a za testiranje korišteno je 100 slika po klasi grafikona što pridonosi pouzdanosti i relevantnosti dobivenih rezultata. Vremenski podaci dobiveni su računanjem razlike između završnog vremena i početnog vremena procesa učenja modela. Simbol "-" označava da model nije bio u mogućnosti izvršiti proces učenja zbog nedostatka raspoloživih računalnih resursa. Ostvareni rezultati potvrđuju visoku optimiziranost predložene arhitekture, pri čemu su zadržane iznimne performanse točnosti klasifikacije, čiji rezultati su konkurentni s rezultatima postignutim dubokim arhitekturama. Ova promatranja sugeriraju da je predložena arhitektura učinkovita ne samo u pogledu postizanja visoke točnosti klasifikacije, već i u kontekstu

Tablica 2.4: Usporedba ostvarenih rezultata različitih arhitektura s obzirom na potrebno vrijeme učenja. Istaknuti redovi predstavljaju arhitekture neuronskih mreža za klasifikaciju slike grafikona izrađene u sklopu ovog istraživanja.

Naziv	Podatkovni skup	Vrijeme učenja (CPU) [s]	Vrijeme učenja (GPU) [s]	Točnost klasifikacije
SCNN	ChartDS (t500)	4400	40	0.99
AlexNet	ChartDS (t500)	6000	60	0.98
VGG16	ChartDS (t500)	-	800	0.95
MobileNet v1	ChartDS (t500)	13200	240	0.99
Inception v3	ChartDS (t500)	-	200	0.99
ResNet50 v2	ChartDS (t500)	-	600	0.30
Xception	ChartDS (t500)	-	-	-
DenseNet121	ChartDS (t500)	-	700	1.00
EffNetB0	ChartDS (t500)	26000	520	0.99

prilagodljivosti na ograničene računalne resurse, što je ključna karakteristika za praktičnu primjenu u stvarnim okruženjima s ograničenim računalnim kapacitetima.

2.6 Eksperimentalna evaluacija – SCNN

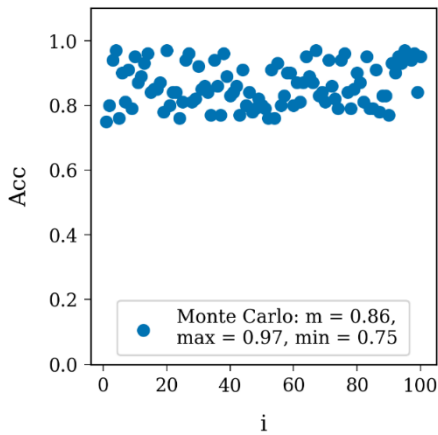
U ovom potpoglavlju prikazana je eksperimentalna evaluacija SCNN modela na dva javno dostupna podatkovna skupa, ChartDS i ICDAR7, izrađena od raznih vrsta i tipova grafikona. Model je zasebno učen na podatkovnom skupu ChartDS i ICDAR7. Postupak učenja modela sastoji se od osam različitih podatkovnih skupova koji sadrže 1, 5, 10, 20, 50, 100, 250 i 500 slika po klasi grafikona. Iz navedenog je nastalo osam uspješnih modela učenih na podatkovnom skupu ChartDS i osam uspješnih modela učenih na podatkovnom skupu ICDAR7. Nakon učenja, modeli su evaluirani na oba podatkovna skupa (ChartDS i ICDAR7). Iz oba podatkovna skupa izuzeti su podatkovni skupovi za testiranje koji su nepromjenjivi i nisu sudjelovali u postupku učenja modela. Podatkovni skup za testiranje podijeljen je na pet podskupova gdje svaki sadrži 20 slika po klasi grafikona (označeno kao S1-5). Modeli su evaluirani na svakom od pet podskupova i prikazane su sumarne vrijednosti. Parametri učenja modela prikazani su u prilogu [P.3](#). Stopa učenja (engl. *Learning rate* – α) je promjenjiv podatak koji se smanjuje kada model tijekom učenja prestane napredovati. Početnu vrijednost, $\alpha = 0.001$, je unaprijed definirao TensorFlow, a kriterij

smanjenja stope učenja je odabran metodom pokušaja i pogrešaka, kao što je prikazano u [13,94,95]. Veličina serije (engl. *Batch size*) odnosi se na broj slika grafikona koji će biti prikazani modelu u postupku učenja. Nakon svake serije slika izvršava se ažuriranje unutarnjih parametara modela. Prema istraživanju provedenom u [96], veličina serije sastoji se od 64 slike grafikona. Za ispitivanje modela tijekom učenja koristi se podatkovni skup za vrednovanje. Modeli su učeni kroz 20 epoha (engl. *Epochs – E*). U neuronskoj mreži, nasumična vrijednost (engl. *Seed*) je vrijednost koja se koristi za početno postavljanje generatora slučajnih brojeva, koji se zatim koristi za nasumično postavljanje težine neuronske mreže. Nasumična vrijednost osigurava različite težine neuronske mreže svaki put kada se postupak učenja ponavlja, kako neuronska mreža ne bi isporučivala lokalne minimume. Nasumična vrijednost također utječe na izlaz neuronske mreže i osigurava isti rezultat ukoliko se koristi ista nasumična vrijednost. Osim kod učenja neuronske mreže, nasumična vrijednost se koristi i za odabir i podjelu slika koje će se nalaziti u podatkovnom skupu za učenje i vrednovanje. Odabir i razvrstavanje slika ima značajnu ulogu u konačnom rezultatu modela. S obzirom da je odabir nasumične vrijednosti usko povezan s težinama i s procesom učenja modela, odabir nasumične vrijednosti je od velikog značaja. Za odabir nasumične vrijednosti korištena je Monte Carlo metoda unakrsnog vrednovanja (engl. *Monte Carlo cross-validation – MCCV*), kao što je prikazano u [97].

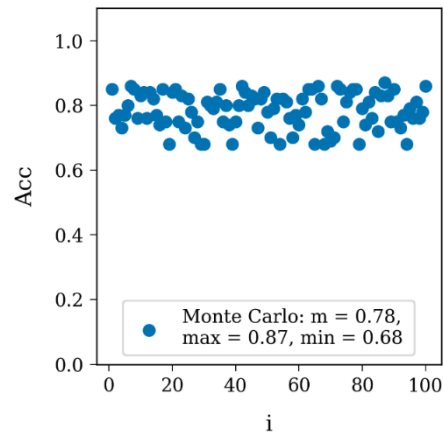
2.6.1 Monte Carlo unakrsno vrednovanje

Unakrsno vrednovanje je statistička metoda koja se koristi u području strojnog učenja za procjenu uspješnosti modela na novim podacima. Metoda unakrsne validacije predstavlja postupak podjele podatkovnog skupa na više podskupova za učenje i vrednovanje, te ponovnog učenja modela. Model uči na podatkovnim podskupovima za učenje i periodično se provjerava upotrebom podskupova za vrednovanje. Uspješnost svakog modela prati se upotrebom unaprijed definirane mjere uspješnosti modela kao što je točnost, preciznost, odziv ili harmonijska sredina preciznosti i odziva, o čemu će više informacija biti prikazano u potpoglavlju 2.6.2. Na kraju, rezultati svih modela se skupno analiziraju i traži se model s najvećom točnošću klasifikacije. Unakrsnim vrednovanjem provjerava se robusnost modela i dokazuje sposobnost učenja i donošenja odluka upotrebom različitih podskupova koji su dio glavnog podatkovnog skupa. Ostvareni rezultati unakrsnog vrednovanja omogućavaju podešavanje hiperparametara modela, odabir modela s najvećom uspješnošću i usporedbu ostvarenih rezultata istraživanja.

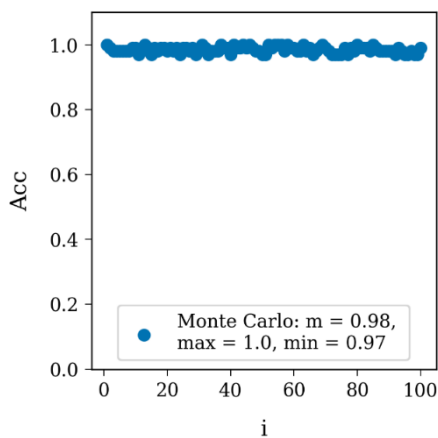
MCCV metoda temeljena je na prethodno opisanoj podjeli podatkovnog skupa. Za razliku od K-podjele metode unakrsnog vrednovanja (engl. *K-folded cross-validation*), podatkovni podskup za vrednovanje generiran ovom metodom može sadržavati dvostruke uzorke, to jest slike



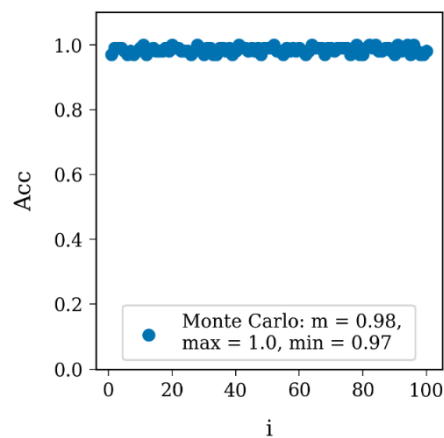
a)



b)



c)



d)

Slika 2.12: MCCV metoda prikazana na podatkovnom skupu s 20 (a, b) i 500 (c, d) slika po vrsti grafikona. Na slikama (a, c) nalazi se postupak učenja na ChartDS, a na slikama (b, d) na ICDAR7 podatkovnom skupu. *Acc* označava točnost klasifikacije, a *i* broj učenja modela.

grafikona. Uzorak se obično pojavljuje u više podskupova za vrednovanje. Ne postoji standardno pravilo za omjer podataka koji će se koristiti u učenju i vrednovanju modela. Omjer podjele podatkovnog skupa je također određen nasumičnom vrijednošću te za učenje iznosi od 70% do 80%, a za vrednovanje od 20% do 30%. Povećanjem broja ponavljanja postupka učenja modela smanjuje se nesigurnost u procjeni izlaznog rezultata i pristranost modela. Odabir broja ponavljanja učenja modela ovisi o veličini podatkovnog skupa, a objavljeni radovi predlažu vrijednost u rasponu od 10 do 1000 ponavljanja, kao što je prikazano u [98–100]. Ako MCCV metoda koristi broj ponavljanja koji je znatno veći od podatkovnog skupa postoji mogućnost pojave istih podatkovnih podskupova što može rezultirati pristranošću modela. S obzirom na veličinu podatkovnog skupa, koji je korišten u procesu učenja modela iz ovog istraživanja i prateći

preporuke objavljenih radova, učenje modela je ponovljeno 100 puta (engl. *Number of iterations* – *i*). Najbolji rezultat MCCV metode za podatkovni skup ChartDS ostvaren je upotrebom nasumične vrijednosti 1529, dok za podatkovni skup ICDAR7 nasumična vrijednost iznosi 635. Postupak učenja modela upotrebom MCCV metode ponovljene 100 puta ($i = 100$) prikazan je slikom 2.12. Slika sadrži empirijsku analizu učenja za podatkovni skup ChartDS (označeno kao a, c) i za podatkovni skup ICDAR7 (označeno kao b, d). Slikom su prikazane ostvarene točnosti klasifikacije na podatkovnim skupovima za vrednovanje. Vrijednosti su raspoređene prema broju slika po klasi grafikona koje su sudjelovale u postupku učenja modela: 20 slika (označeno kao a, b) i 500 slika (označeno kao c, d). Kada se u procesu učenja modela koristi jedna slika po klasi grafikona, model ne uspijeva naučiti značajke. Međutim, korištenjem 20 slika po klasi grafikona model pokazuje napredak u učenju. Vidljiva je izražena oscilacija točnosti klasifikacije od preko 10%, pri čemu je prosječna vrijednost (engl. *Mean – m*) nešto veća za podatkovni skup ChartDS. Korištenjem 500 slika po klasi grafikona, model ostvaruje maksimalnu točnost klasifikacije neovisno o korištenom podatkovnom skupu.

2.6.2 Određivanje uspješnosti

Za određivanje uspješnosti modela koristi se matrica zabune (engl. *Confusion matrix*). Matrica zabune grafički prikazuje vrijednosti stvarnih i predviđenih klasa. Stvarna klasa nalazi se na y-osi, a klasa koju je model predvidio na x-osi. Na dijagonali nalaze se točno predviđene i stvarne vrijednosti klasa ulaznih slika. Cilj je ostvariti što veći broj slika na dijagonali jer se navedeno odražava na visoku točnost klasifikacije. Iz matrice se izračunavaju najčešće korištene mjere uspješnosti modela:

- Točnost/uspješnost (engl. *Accuracy – Acc*) predstavlja udio točno klasificiranih slika u odnosu na cijeli podatkovni skup
- Preciznost (engl. *Precision*) predstavlja udio točno klasificiranih istinitih slika u podatkovnom skupu istinitih slika
- Odziv (engl. *Recall*) predstavlja udio točno klasificiranih istinitih slika i ukupnog broja slika koje su trebale biti točno klasificirane
- F1-mjera (engl. *F1-score*) je harmonijska sredina preciznosti i odziva

Matrica zabune se oslanja na vrijednost vjerojatnosti svake pojedine klase (A-H), pa uz svaku matricu zabune nalazi se i statistički prikaz vjerojatnosti koji je definiran:

- R1 – rezultat uspješnosti prosječne vrijednosti vjerojatnosti gdje je stvarna klasa istinita. Provjeravaju se vjerojatnosti samo na slikama gdje je model trebao ostvariti predikciju (više je bolje).
- R2 – rezultat uspješnosti prosječne vrijednosti vjerojatnosti gdje su stvarna klasa i predviđena klasa istinite. Provjeravaju se vjerojatnosti na slikama gdje je model trebao ostvariti predikciju i gdje je predikcija točna (više je bolje).
- R3 – rezultat uspješnosti prosječne vrijednosti vjerojatnosti gdje je stvarna klasa istinita, a predviđena klasa nije istinita. Provjeravaju se vjerojatnosti na slikama gdje je model trebao ostvariti predikciju, ali nije (manje je bolje).
- R4 – rezultat uspješnosti prosječne vrijednosti vjerojatnosti gdje stvarna klasa nije istinita, a predviđena klasa je istinita. Provjeravaju se vjerojatnosti na slikama gdje je model ostvario predikciju, ali nije trebao (manje je bolje).
- R5 – rezultat uspješnosti prosječne vrijednosti vjerojatnosti gdje nisu istinite ni stvarna klasa ni predviđena klasa. Iznos vjerojatnosti na svim ostalim slikama gdje model nije ostvario predikciju, ali nije ni trebao (manje je bolje).

Uvid u statističke podatke omogućava uvid u rezultate vjerojatnosti modela pri donošenju odluka.

2.6.3 Otvorena klasifikacija

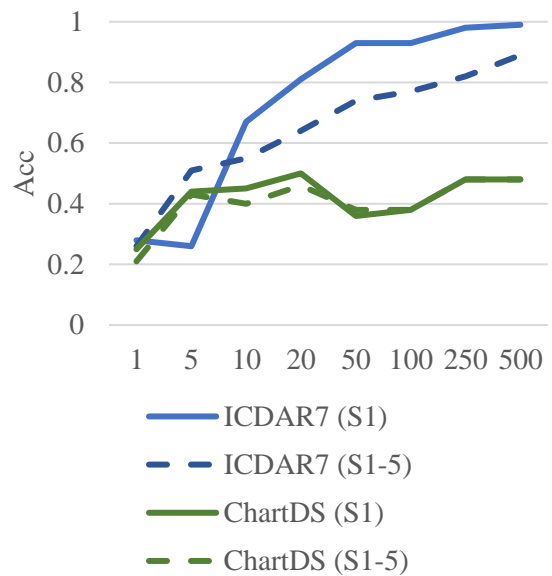
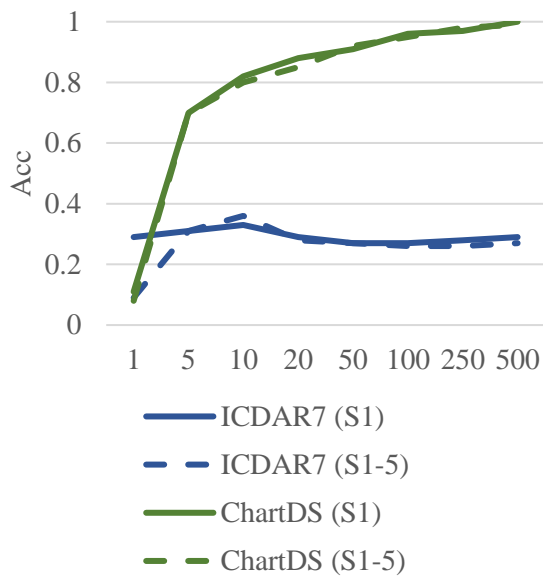
Otvorena klasifikacija (engl. *Open set recognition/classification*) je dobro poznati izazov u području strojnog učenja koji označava sposobnost modela prilagodbi neočekivanim podacima (slikama) na ulazu, kao što je prikazano u [101]. Za ljude je ovo jednostavan izazov gdje je na prvi pogled moguće odrediti pripada li objekt u određenu klasu ili ne, a za računala je ovo posebna kategorija izazova za koju ne postoji jedinstveno rješenje, kao što je prikazano u [102]. U sklopu ovog istraživanja, od modela se očekuje da slike koje ne pripadaju u navedenih sedam klasa grafikona (A-G), budu klasificirane kao zasebna klasa koja ne sadrži slike grafikona (H). Kroz povijest, predloženo je nekoliko različitih rješenja kojima je moguće riješiti izazov otvorene klasifikacije kao što je upotreba binarnog SVM klasifikatora prikazana u [101], upotreba SVM klasifikatora s definiranom granicom prikazana u [103], upotreba statističkog znanja o vjerojatnosti uključivanja ili isključivanja vektora na temelju granične vrijednosti prikazana u [104], odbacivanje određenih ulaza za vrijeme učenja mreže prikazano u [105], izrada zasebne klase koja će sadržavati sve očekivane nepoznate klase ili izrada zasebne klase koja će sadržavati najveći mogući broj poznatih slika prikazana u [106,107] i upotreba učenja iz primjera (engl. *Zero/One/Few shot learning*) prikazana u [108,109]. Nijedna od predloženih metoda nije savršena

i svaka podliježe riziku otvorenog podatkovnog skupa (engl. *Open set risk*) odnosno na jednostavan način nije moguće definirati granice, vrijednosti za podatke koji će se smatrati poznatima od onih nepoznatih. Cilj je model pripremiti na nepoznatu klasu te spriječiti prodiranje nepoznate klase među poznate, kao što je prikazano u [110]. U postupku izrade otvorene klasifikacije primijenjena je metoda izrade zasebne klase (H) koja će sadržavati slike različite od slika grafikona, kao što je prikazano u [111]. Zasebna klasa je konstruirana uključivanjem raznolikih, slučajno odabranih slika iz podatkovnih skupova CIFAR10 i Linnaeus. Strukturno, klasa prati iste smjernice organizacije kao i klase koje sadrže slike grafikona (A-G).

2.6.4 Ostvareni rezultati modela

Na slici 2.13, prikazana je točnost klasifikacije modela učenog na osam podatkovnih skupova koji pripadaju podatkovnom skupu ChartDS (označeno kao a), odnosno pripadaju podatkovnom skupu ICDAR7 (označeno kao b). Modeli su testirani na oba podatkovna skupa. Rezultati testiranja su u skladu s MCCV testiranjem na podatkovnom skupu za vrednovanje. Točnost klasifikacije kada se za učenje modela koristi jedna slika po klasi grafikona je približno 0, kada se za učenje koristi 20 slika po klasi grafikona je približno 0.85, a kada se za učenje koristi 500 slika po klasi grafikona je približno 1. U prijašnjem istraživanju [13] za testiranje je uvijek korišteno 20 slika po klasi grafikona. Taj identični prvobitni podatkovni skup označen je sa S1. Istraživanje je dodatno prošireno na pet podatkovnih skupova, odnosno 100 slika po klasi grafikona te je vidljivo minimalno odstupanje rezultata. Kada je model učen na jednom podatkovnom skupu, a testiran na drugom podatkovnom skupu vidljiv je značajan pad točnosti klasifikacije iako se testiraju iste klase grafikona. Ručnom provjerom ostvarenih rezultata vidljivo je da model većinu slika svrstava u zasebnu klasu (H). O ovoj pojavi napravljeno je dodatno istraživanje koje se nalazi u potpoglavlju 4.4.

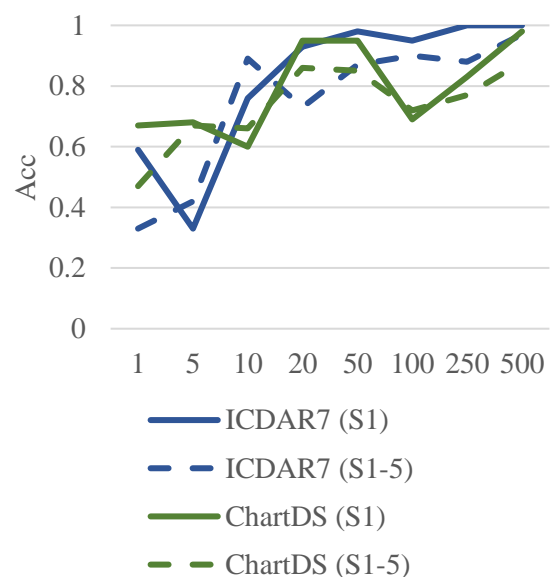
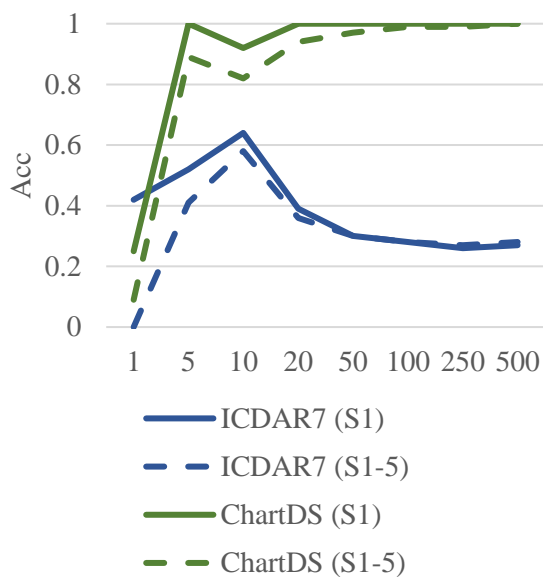
Na slici 2.14, prikazana je točnost klasifikacije modela na zasebnoj klasi (H) te za podatkovni skup ChartDS (označeno kao a) prati sliku 2.13 (a), a za podatkovni skup ICDAR7 (označeno kao b) prati sliku 2.13 (b). Model uspješno razlikuje slike koje sadrže grafikon od slika koje ne sadrže grafikon. Kada se za testiranje modela koji je učen na podatkovnom skupu ChartDS koristi podatkovni skup ICDAR7, model ostvaruje smanjenu točnost klasifikacije. Iako oba podatkovna skupa sadrže zasebnu klasu sastavljenu od CIFAR10 i Linnaeus podatkovnih skupova, u slučaju podatkovnog skupa ICDAR7 model nije uspješan u razlikovanju slika koje sadrže grafikon od slika koje ne sadrže grafikon. Razlog tome je što model, zbog ostvarenih niskih vjerojatnosti, više od 80% slika klasificira u zasebnu klasu. Kada se za testiranje modela koji je učen na podatkovnom skupu ICDAR7 koristi podatkovni skup ChartDS, model uspijeva



a)

b)

Slika 2.13: Prikaz točnosti klasifikacije grafikona (A-G). Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.

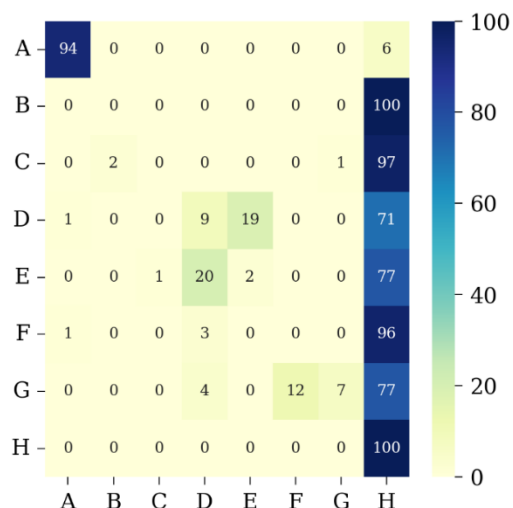


a)

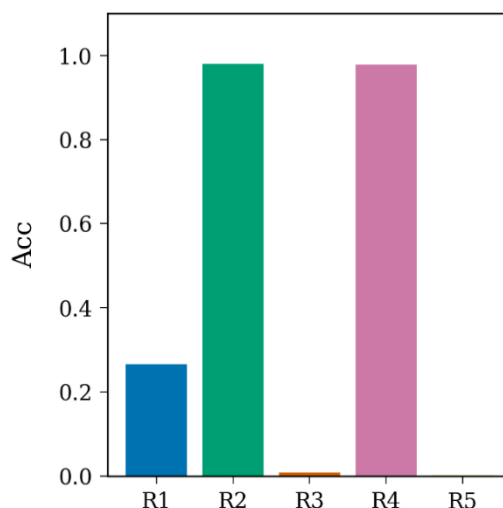
b)

Slika 2.14: Prikaz točnosti klasifikacije zasebne klase (H). Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.

razlikovati slike koje sadrže grafikon od slika koje ne sadrže grafikon. Razlog tome je što je struktura grafikona iz podatkovnog skupa ICDAR7 unificirana i značajno se razlikuje od slika koje ne sadrže grafikon.

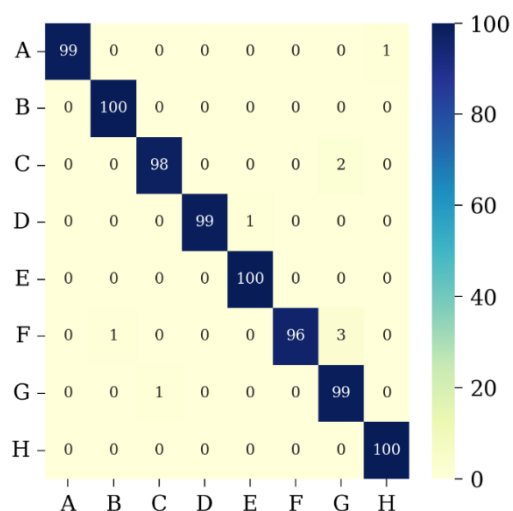


a)

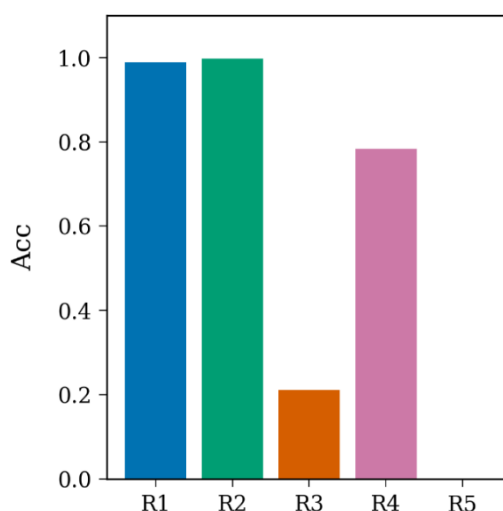


b)

Slika 2.15: Prikaz točnosti klasifikacije za model učen na ChartDS, a testiran na ICDAR7 podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.



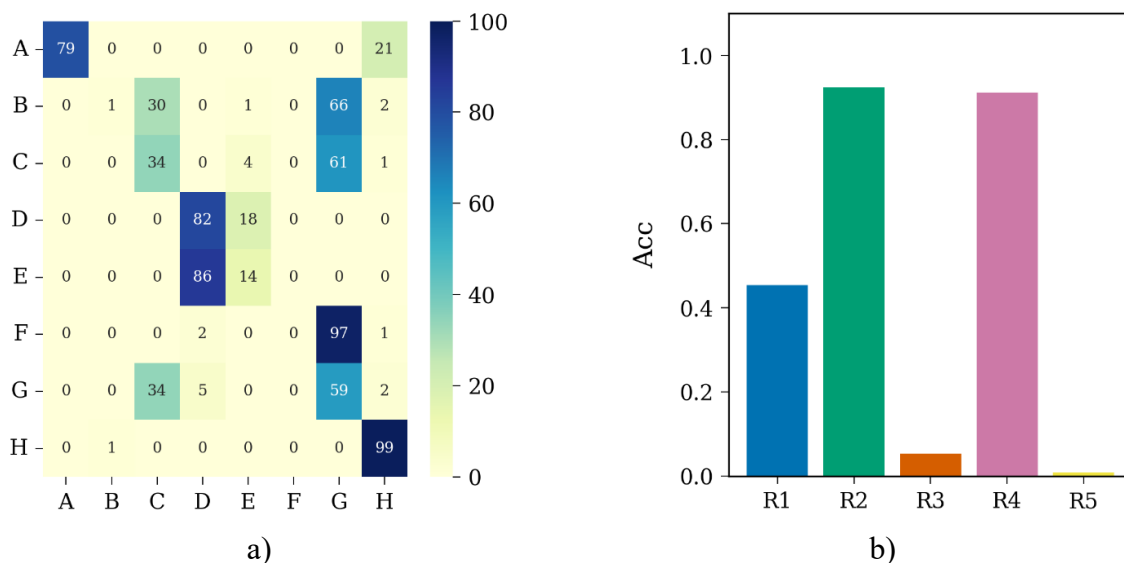
a)



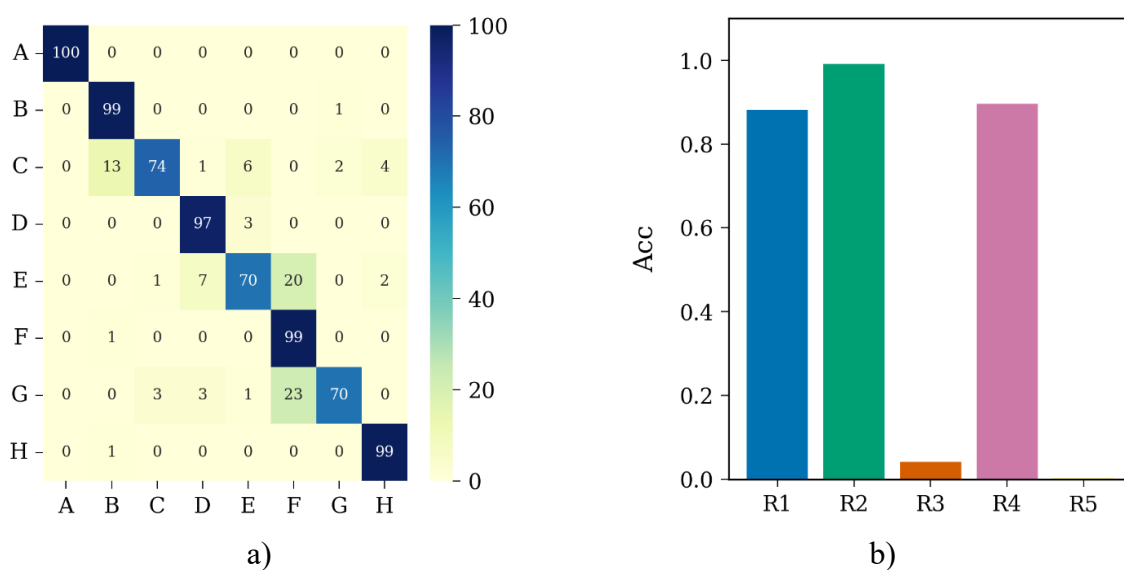
b)

Slika 2.16: Prikaz točnosti klasifikacije za model učen i testiran na ChartDS podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.

Detaljan prikaz točnosti klasifikacije za model učen na podatkovnom skupu ChartDS, a testiran na podatkovni skup ICDAR7 prikazan je slikom 2.15. Ako bi iz matrice zabune (označeno kao a) izolirali zasebnu klasu (H), ostvarena točnost klasifikacije iznosi 0.16, odziv 1.00, a F1-mjera 0.28. Statističkom analizom (označeno kao b) vidljivo je da model ostvaruje vrlo visoku vjerojatnost, no u slučaju ovog podatkovnog skupa model ne uspijeva razlikovati slike koje sadrže grafikon od slika koje ne sadrže grafikon. Kada je model, učen na podatkovnom skupu ChartDS i



Slika 2.17: Prikaz točnosti klasifikacije za model učen na ICDAR7, a testiran na ChartDS podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.



Slika 2.18: Prikaz točnosti klasifikacije za model učen i testiran na ICDAR7 podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.

testiran na istom podatkovnom skupu, rezultati su nadprosječni, prikazano slikom 2.16. Prema matrici zabune (označeno kao a) model ostvaruje točnost klasifikacije veću od 0.99, a pri tome se koristi neuronska mreža sa znatno smanjenom dubinom u odnosu na druga istraživanja iz ovog područja koja su prikazana tablicom 2.1. Dodatno, ovo je prva primjena SCNN-a za otvorenu klasifikaciju u području klasifikacije slike grafikona. Za slike koje ne sadrže grafikon, ostvarena točnost klasifikacije iznosi 0.99, odziv 1.00 i F1-mjera 1.00. Iz statističke analize (označeno kao

b), model ostvaruje predikciju tamo gdje se predikcija i očekuje te je vjerojatnost vrlo visoka. U slučaju netočnih predikcija, koje iznose manje od 0.1, model ostvaruje odgovore s visokom vjerojatnošću (R3), ali vjerojatnost predikcije je i dalje manja u odnosu za istinite predikcije (R1, R2). Postavljanjem granične vrijednosti, ako netočna predikcija treba pripadati zasebnoj klasi, moguće je napraviti dodatno isključenje i tako ostvariti maksimalnu točnost klasifikacije.

S druge strane, detaljan prikaz točnosti klasifikacije za model učen na podatkovnom skupu ICDAR7, a testiran na podatkovni skup ChartDS prikazan je slikom 2.17. Ako bi iz matrice zabune (označeno kao a) izolirali zasebnu klasu (H), ostvarena točnost klasifikacije iznosi 0.79, odziv 0.99, a F1-mjera 0.88. Usporednom analizom s prethodnim testiranjem, model uspijeva razlikovati i A, D i G klasu grafikona. Statističkom analizom (označeno kao b) vidljivo je da model postiže visoku vjerojatnost, no ne uspijeva razlikovati B, C, E i F klasu grafikona. Prema slici 2.18, kada je model učen na podatkovnom skupu ICDAR7 i testiran na istom podatkovnom skupu, rezultati su usporedni s rezultatima istraživanja iz ovog područja koja su prikazana tablicom 2.1. Prema matrici zabune (označeno kao a) model ostvaruje točnost klasifikacije od 0.90. Model ne razlikuje B od C klase i F od G klase grafikona. Provjerom E klase vidljivo je nerazlikovanje od svih ostalih klasa što uzrokuje ukupni gubitak točnosti klasifikacije. Rezultati statističke analize (označeno kao b) su u skladu s prethodnim testiranjima. Model ostvaruje visoku vjerojatnost na mjestima točne i netočne predikcije. Povećanjem raznolikosti slika sadržanih u skupu za učenje, ukupna točnost klasifikacije bi se znatno povećala.

Iako oba modela postižu visoku točnost klasifikacije, model učen na podatkovnom skupu ChartDS pokazuje bolje rezultate. Slike iz podatkovnog skupa ChartDS obično ne sadrže legendu u blizini grafičkih simbola, dok slike iz podatkovnog skupa ICDAR7 često sadrže legendu. Budući da algoritam za predobradu slike grafikona ne može ukloniti takvu legendu sa slike, ona ostaje prisutna nakon predobrade. Očuvanje legende znači da slika sadrži više informacija koje model uči tijekom procesa učenja, kao što je prikazano slikom 2.4. Nadalje, slike iz ICDAR7 podatkovnog skupa često uključuju trake grešaka (engl. *Error bars*) koje nisu prisutne u podatkovnom skupu ChartDS. Trake grešaka su dodatni grafički simboli koji mogu biti prisutni na različitim vrstama grafikona, poput stupčastih ili linijskih. Nakon predobrade slika grafikona, trake grešaka ostaju vidljive, a njihov oblik podsjeća na grafikon kutija, što dodatno uzrokuje pad ukupne točnosti klasifikacije.

2.7 Eksperimentalna evaluacija – SSCNN

Parametri učenja SSCNN modela slični su parametrima učenja SCNN modela, broj epoha povećan je na 100 jer je vidljivo da neuronska mreža nije završila učenje nakon 20 epoha. U učenju SSCNN modela korišten je dodatni AT&T DoF podatkovni skup čije slike ne sudjeluju u procesu testiranja. AT&T DoF podatkovni skup sastoji se od 40 klasa i omogućava modelu kvalitetnije učenje sličnosti između parova slika. Umjesto ovog podatkovnog skupa moguće je koristiti bilo koji drugi označeni i organizirani podatkovni skup. S druge strane, ako se ne koristi dodatni podatkovni skup model, zbog malog broja klasa grafikona, predobro nauči sličnosti između postojećih parova slika i vrijednost funkcije gubitka kreće oscilirati između dvije vrijednosti. Pojava izraženih oscilacija označava nemogućnost modela učenja sličnosti između parova slika.

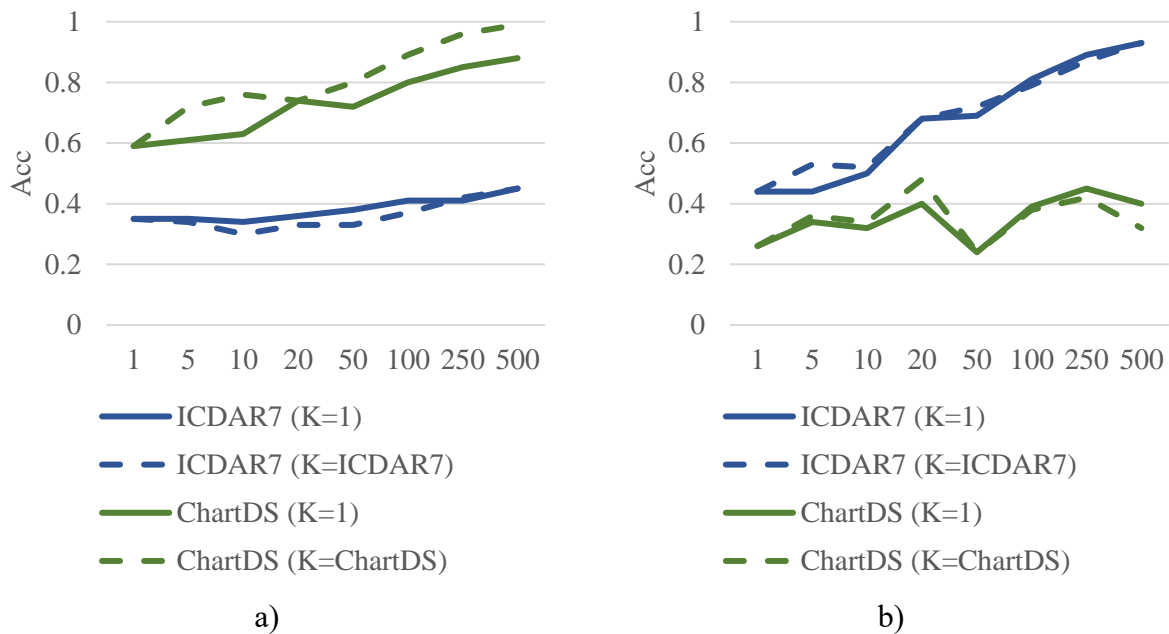
Eksperimentalna evaluacija SSCNN modela znatno se razlikuje od SCNN modela budući da SSCNN model ne daje na izlazu vjerojatnost već koeficijent sličnosti. Za klasifikaciju koja se sastoji od više klasa potrebno je izračunati koeficijent sličnosti ulazne slike za svaku od klasa koje su sudjelovale u procesu učenja modela. Proces izračunavanja koeficijenta sličnosti za svaku klasu naziva se učenje iz primjera (engl. *N-way-K-shot learning*), gdje je N broj klasa, a K broj uzoraka (slika) iz svake klase, kao što je prikazano u [112]. U ovoj evaluaciji koristi se po jedna slika iz sedam klasa koje su sudjelovale u učenju modela i koje sadrže slike grafikona (A-G) te se metoda naziva sedmosmjerno učenje iz jednog primjera (engl. *7-way-1-shot learning*). Prilikom korištenja metode sedmosmjernog učenja iz jednog primjera, na jedan ulaz SSCNN-a dovodi se poznata slika grafikona koju je model vidio i koja je sudjelovala u procesu učenja. Na drugi ulaz SSCNN-a dovodi se nova, nepoznata slika koju je potrebno usporediti sa slikama grafikona koje su sudjelovale u procesu učenja. Ovim postupkom stvara se sedam parova slika za jednu sliku (primjerice parovi slika A-A, A-B, A-C, A-D, A-E, A-F i A-G). Za svaki par slika, SSCNN model izračunava koeficijent sličnosti. Ako je najniži koeficijent sličnosti za ulazni par slika A-A, rezultat se klasificira kao točan, odnosno slike pripadaju istoj klasi. S obzirom da jedan testni podatkovni skup sadrži 20 slika grafikona, postupak se ponavlja 20 puta. Testira se 140 parova slika za jednu klasu, 980 parova slika za svih sedam klasa (oznaka na slikama $K = 1$). Konačni ukupan broj za sve testne podatkovne skupove iznosi 700 parova slika za jednu klasu, 4900 parova slika za svih sedam klasa. Kod SSCNN modela kvaliteta podatkovnog skupa važnija je od kvantitete. Kada se ulazna nepoznata slika uparuje s nasumično odabranom slikom grafikona iz procesa učenja, koeficijent sličnosti može ostvariti visoku ili nisku točnost klasifikacije. Ako model slučajnim odabirom odabere sliku grafikona sličnu ulaznoj slici, koeficijent sličnosti će biti blizu nule; inače

će koeficijent sličnosti biti bliži jedinici. Kako bi se uklonio slučajan odabir, ulazna nepoznata slika se uparuje sa svim slikama grafikona koje su sudjelovale u procesu učenja. Ovom metodom povećava se ukupni broj parova slika na $140 \times \text{broj klase} \times \text{veličina podatkovnog skupa} \times \text{broj podatkovnih skupova}$ (oznaka na slikama $K = \text{naziv_podatkovnog_skupa}$). Rezultat empirijske evaluacije prikazan je u potpoglavlju [2.7.1](#).

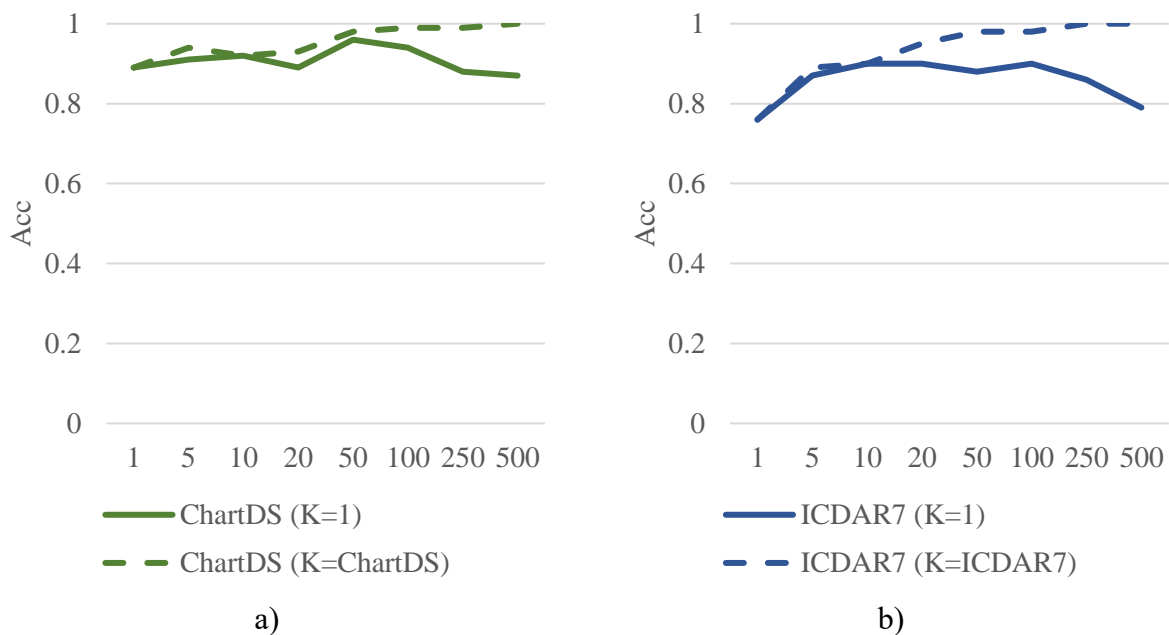
2.7.1 Ostvareni rezultati modela

Za testiranje SSCNN modela koristi se isti podatkovni skup koji je korišten u testiranju SCNN modela. Na slici [2.19](#) prikazana je točnost klasifikacije modela učenog na osam podatkovnih skupova koji sadrže slike grafikona i pripadaju podatkovnom skupu ChartDS (označeno kao a), odnosno pripadaju podatkovnom skupu ICDAR7 (označeno kao b). Modeli su testirani na oba podatkovna skupa. Iz slike je vidljiva prednost SSCNN modela nad SCNN modelom. SSCNN model uspješno radi s vrlo malim brojem slika po klasi grafikona. Kada se za učenje modela koristi samo jedna slika grafikona po klasi, model ostvaruje točnost klasifikacije u rasponu od 0.4 do 0.6. Kada se za učenje modela koristi 20 slika grafikona po klasi vidljiv je rezultat slučajnog odabira, događa se pad točnosti klasifikacije iako se broj primjera u učenju modela povećao. Model dostiže maksimalnu točnost klasifikacije kada se ulazna slika uspoređuje s cijelim podatkovnim skupom, no kada se za usporedbu koristi jedna slika grafikona iz svake klase, $K = 1$, napredak je značajno sporiji te je potrebna veća količina parova slika da bi se ostvario jednaki rezultat. Testiranje modela učenog na jednom podatkovnom skupu s podacima drugog podatkovnog skupa ostvaruje bolje rezultate u odnosu na SCNN model, no rezultati su i dalje ispodprosječni. Ručnom provjerom utvrđeno je da model ne razlikuje B od C klase i F od G klase grafikona

Sijamske neuronske mreže mogu generalizirati i za ulaze koji nisu sudjelovali u procesu učenja. Na slici [2.20](#) prikazana je točnost klasifikacije na klasi koja ne sadrži slike grafikona (H), gdje je model učen na ChartDS podatkovnom skupu (označeno kao a), odnosno ICDAR7 podatkovnom skupu (označeno kao b). SSCNN model, u procesu učenja nije učen na slikama koje ne sadrže grafikon. Točnost klasifikacije zasebne klase, za bilo koji podatkovni skup iznosi više od 0.8, a kod modela koji je učen na 50 ili više slika grafikona po klasi postiže maksimalnu točnost klasifikacije.

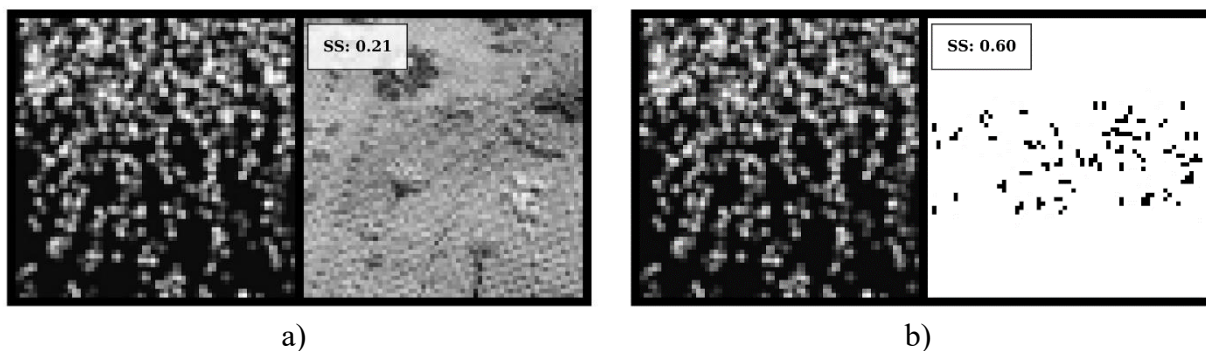


Slika 2.19: Prikaz točnosti klasifikacije grafikona. Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.



Slika 2.20: Prikaz točnosti klasifikacije zasebne klase (H). Na slici a) model je učen na ChartDS podatkovnom skupu, a na slici b) model je učen na ICDAR7 podatkovnom skupu.

Uspješnost SSCNN modela na zasebnoj klasi može se vidjeti i na slici 2.21. Na slici a) uspoređuju se dvije blago predobrađene slike iz podatkovnog skupa Linnaeus, a na slici b) uspoređuje se ista slika iz podatkovnog skupa Linnaeus sa slikom točkastog grafikona iz podatkovnog skupa ChartDS. Na prvi pogled, slika iz Linnaeus podatkovnog skupa je vrlo slična



Slika 2.21: Usporedba slike iz Linnaeus podatkovnog skupa koja ne sadrži grafikon s: a) drugom slikom iz Linnaeus podatkovnog skupa koja ne sadrži grafikon (SS = 0.21) i b) slikom točkastog grafikona iz ChartDS podatkovnog skupa (SS = 0.60)

slici točkastog grafikona, no SSCNN uspješno klasificira sliku s vrijednošću koeficijenta sličnosti od 0.21, dok vrijednost koeficijenta sličnosti s točkastim grafikonom je znatno veća i iznosi 0.60 (vrijednost bliže 0 označava istu klasu).

Dodatno, SSCNN model je testiran i na ReVision podatkovnom skupu. Iz ReVision podatkovnog skupa korišteno je pet vrsta grafikona: prostorni grafikon, stupčasti grafikon, linijski grafikon, radar i Vennov dijagram gdje su sve slike prošle isti proces predobrade kao i slike iz ChartDS i ICDAR7 podatkovnog skupa. Od svake vrste grafikona, slučajnim izborom, odabrano je 100 slika za testiranje. Ukupna postignuta točnost klasifikacije iznosi 0.96. Važno je istaknuti da SSCNN modeli nisu imali pristup ReVision podatkovnom skupu za vrijeme učenja.

2.8 Zaključak

Metode klasifikacije slika grafikona su se mijenjale tijekom godina. Sve metode korištene prije 2015. godine mogu se smatrati tradicionalnim metodama. Iako su neke metode ostvarile točnost klasifikacije iznad 0.80, slike grafikona morale su slijediti unaprijed definirana pravila. ReVision [47] je prvi suvremeni rad koji je uveo višeklasnu klasifikaciju pomoću SVM-a. Autori su koristili različite metode i algoritme za predobradu slike kako bi ostvarili što veću točnost klasifikacije slike grafikona. Pojavom CNN-a u području, zadržane su metode i algoritmi za predobradu slike grafikona, a broj klasa, odnosno vrsta i tipova grafikona je značajno povećan. Sve navedene CNN arhitekture postižu vrlo visoku točnost klasifikacije iznosa većeg od 0.8. Kako bi postigli maksimalnu točnost, autori često kombiniraju različite pristupe koji pomažu CNN-u izdvojiti značajke iz slike i usredotočiti se na područje koje sadrži značajnu informaciju.

U ovom poglavlju predstavljena je SCNN arhitektura kojoj je primarni zadatak klasifikacija slika grafikona. Arhitektura se sastoji od četiri konvolucijska sloja, dva sloja s maksimalnim sažimanjem i jednog potpuno povezanog sloja. Arhitekturu SCNN-a je moguće koristiti samostalno ili povezati dvije neuronske mreže u jednu cjelinu što rezultira stvaranjem SSCNN-a arhitekture. Kada se koristi SCNN arhitektura, eksperimentalna evaluacija pokazuje i učinkovitost i uspješnost u postizanju rezultata. U usporedbi s drugim istraživanjima iz područja klasifikacije slike grafikona, predstavljenom arhitekturom su ostvarene visoke točnosti klasifikacije koje dostižu maksimalnu vrijednost od 1.00, a pri čemu se koristi arhitektura s pet težinskih slojeva. Kada se koristi SSCNN arhitektura, visoka točnost klasifikacije može se postići i s malim podatkovnim skupovima, što predstavlja jedan od ključnih atributa ove arhitekture.

Na osnovu predstavljene arhitekture, stvoreno je sveukupno 32 modela koja su učena na dva javno dostupna podatkovna skupa, ChartDS i ICDAR7, iz kojih je nastalo osam podskupova različitih veličina. Svaki od modela je testiran na oba podatkovna skupa što omogućuje stvaranje referentnog polazišta za buduća istraživanja. Dodatno, svi modeli su testirani i na koncept otvorene klasifikacije. Modeli ostvaruju visoku točnost u razlikovanju slika grafikona od slika koje ne sadrže grafikon. Ipak, u slučaju otvorene klasifikacije, značajnu prednost ostvaruje SSCNN arhitektura i učenje iz primjera. SCNN arhitekturi je potrebna dodatna klasa, inače će ulazi biti generalizirani prema jednoj od postojećih klasa grafikona. Negativna strana je što u konačnici ne postoji dovoljno velik podatkovni skup poznatih slika koji model mogu pripremiti na nepoznatu sliku. SSCNN arhitektura ostvaruje generalizaciju, odnosno koeficijent sličnosti i za slike koje nikada nisu sudjelovale u učenju modela. Upotrebom vrlo malog podatkovnog skupa ostvaruje se prepoznavanje slika koje ne sadrže grafikon s točnošću klasifikacije većom od 0.8, a prilikom upotrebe većeg podatkovnog skupa, postiže se maksimalna točnost klasifikacije.

Visoka učinkovitost SCNN i SSCNN arhitektura ne bi bila ostvariva bez predstavljenog algoritma za predobradu slike grafikona. Predstavljeni algoritam s ulazne slike grafikona uspješno uklanja većinu grafičkih i tekstualnih elemenata koji ne sadrže značajnu informaciju potrebnu za klasifikaciju grafikona i od ulazne slike stvaraju jednostavnu crno-bijelu sliku. Prvobitna inačica algoritma predstavljena je 2020. godine, a do danas je prošla kroz tri iteracije.

3 METODA IZDVAJANJA PODATKOVNIH ZNAČAJKI IZ KRUŽNIH I MREŽASTIH GRAFIKONA ZA PRIMJENU U RAČUNALNIM SUSTAVIMA

U ovom poglavlju biti će opisana metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima. Ovaj doprinos objavljen je u [68], a provedeno istraživanje povezano je s [7]. Poglavlje opisuje metodu izdvajanja podatkovnih značajki iz kružnih grafikona u koje se ubrajaju tortni, prstenasti i grafikoni sunčevih zraka i metodu izdvajanja podatkovnih značajki iz mrežastih grafikona u koje se ubrajaju toplinske karte i grafikoni napretka. Metode dijele značajan broj koraka te ih je zbog toga moguće objediniti jednim nazivom — metoda izdvajanja podatkovnih značajki. Metoda se sastoji od dva modula, modul za predobradu slike grafikona i modul za detekciju elemenata grafikona. Rad metode zasnovan je na binarnim maskama, morfologiji, detekciji ključnih točaka i matematičkom modelu. Metoda je prvenstveno izrađena upotrebom umjetno stvorenog podatkovnog skupa, na kojem je i evaluirana, ali u evaluaciju dodatno je uključen i manji podatkovni skup stvarnih slika grafikona prikupljenim internetskim pretraživanjem. Na umjetno stvorenim slikama grafikona ostvarena je visoka točnost izdvajanja podatkovnih značajki, dok za stvarne slike točnost je očekivano smanjena zbog

heterogenosti strukturnih elemenata korištenih u izradi grafikona. Potpoglavlja su dodatno potkrijepljena priložima u kojima se nalazi nekoliko primjera jednostavnih i složenih grafikona i njihova uspješna rekonstrukcija nastala metodom izdvojenih značajki.

3.1 Uvod

Kontinuirani razvoj predodžbe podataka i raširena uporaba interneta omogućili su stvaranje i prikupljanje ogromne količine podataka, primjerice numeričkih, tekstualnih, slikovnih i drugih oblika podataka. U mnogim slučajevima, ključne informacije, koje je teško prenositi i tumačiti kroz tekst ili tablice, postaju razumljivije kroz grafikone. Međutim, grafikon se često pohranjuje u obliku rasterske slike. Prema strukturnoj definiciji [113], grafikon je skup nedjeljivih grafičkih elemenata prikazanih u jednoj zajedničkoj n -dimenzionalnoj ravnini. Grafikoni posjeduju jaku reprezentativnu snagu, koja nadopunjuje činjenice u glavnom tekstu, te su široko prihvaćeni u akademskim radovima, izvještajima, analizama i vijestima.

Dok osobe s normalnim vidom mogu jednostavno razumjeti slike grafikone, računala i osobe s oštećenim vidom su suočeni s izazovom u pogledu pristupačnosti i dohvata sadržajne informacije slike grafikona. U akademskim radovima eksperimentalni su podatci predstavljeni različitim vrstama grafikona koji su izvor informacija. Jedan od načina na koji osoba s normalnim vidom može dobiti izvorne, neobrađene, podatke grafikona jest korištenjem sposobnosti ljudske vizualne obrade koja višestrukim ponavljanjem postaje iscrpljujuća. Osim navedenog, računala ne mogu obrađivati tekstualne i numeričke informacije sadržane u slikama grafikonima poput ljudi jer prikaz grafikona nije prvobitno namijenjen za računalnu obradu. Nedostatak računalne čitljivosti sprječava daljnju analizu i obradu sadržajne informacije slike grafikona.

Kako bi računala mogla razumjeti sliku grafikona, postaje neophodno pretvoriti sliku u odgovarajuću tabličnu reprezentaciju. Cjeloviti sustav za interpretaciju slike grafikona obično se sastoji od više koraka. U prvom koraku, sustav započinje proces lokalizacije i izdvajanja slike grafikona unutar dokumenta. U drugom koraku, određuje se vrsta grafikona upotrebom raznih metoda za klasifikaciju ulazne slike grafikona. Zatim, nastavlja se obrada slike koja je najčešće podijeljena u dva paralelna koraka: jedan za prepoznavanje tekstualnih elemenata grafikona (primjerice naslov, podnaslov, vrijednosti, oznake na osima i legenda), a drugi za otkrivanje grafičkih elemenata. Po završetku navedenih koraka, sustav objedinjuje tekstualne i grafičke elemente u jednu informacijsku cjelinu koja omogućuje izradu statičke analize, stvaranje generativnih sadržajnih opisa ili konzumaciju podataka u brojnim računalnim sustavima.

Dekodiranje grafičkih elemenata je ključan, ali u usporedbi s tekstualnim elementima, značajno manje razvijen i istraživani korak u procesu digitalne obrade slike grafikona. Jedan od glavnih razloga je heterogenost vrsti, tipova grafikona i strukturnih elemenata čiji grafički izgled nije strogo standardiziran.

Automatizacija cijelog sustava ili dijela sustava doprinosi povećanju dostupnosti sadržaja grafikona ljudima s oštećenim vidom i računalima, te omogućuje korištenje istog u digitalnim okruženjima gdje slike nisu jednostavno podržane.

U potpoglavlju 3.2 dan je povijesni pregled istraživanja iz područja izdvajanja značajki iz slike grafikona. Potpoglavlje 3.3 predstavlja metodu izdvajanja podatkovnih značajki iz slike grafikona. Eksperimentalna evaluacija provedena na umjetno stvorenom podatkovnom skupu i podatkovnom skupu prikupljenih slika s interneta dostupna je u potpoglavlju 3.4. Zaključak i osvrt na metodu napisani su u potpoglavlju 3.5.

3.2 Istraživanja u području izdvajanja podatkovnih značajki iz slike grafikona

U ovom potpoglavlju dan je kratki povijesni pregled objavljenih radova iz područja izdvajanja podatkovnih značajki iz slike grafikona, a prošireni pregled istraživanja u području objavljen je u [7]. Objavljeni radovi u području istraživanja razvrstani su prema korištenoj metodi izdvajanja podatkovnih značajki iz slike grafikona. U prvom dijelu potpoglavlja predstavljene su metode, a u drugom dijelu nalazi se tablica koja uspoređuje ostvarene rezultate.

Tijekom godina istraživači su koristili različite metode izdvajanja podatkovnih značajki iz slike grafikona. Metode je moguće grupirati u tri kategorije:

- izrada općenitih pravila koja prate strukturu grafikona (engl. *Rule-based*),
- detekcija objekta temeljena na ključnim točkama (engl. *Keypoint-based object detection*) i
- metode dubokog učenja.

Autori koji se bave izdvajanjem podatkovnih značajki iz slika grafikona imaju dva osnovna cilja. Prvi cilj obuhvaća izradu izvorne podatkovne tablice iz koje je stvoren grafikon. Drugi cilj je otkrivanje korištenih vizualnih stilova za izradu grafikona (primjerice karta boja, fontovi, dimenzije i položaj elemenata na slici). Za izdvajanje podatkovnih značajki iz slike grafikona, autori koriste različite pristupe u dohvat podataka koji se mogu grupirati u dvije metode:

automatska i interaktivna. U automatskoj metodi, korisnik unosi sliku grafikona u sustav, a sustav obrađuje sliku i vraća objektno orijentiranu strukturu. Ovo je najzastupljenija metoda i ne zahtijeva dodatnu interakciju ili korekciju koju treba izvršiti korisnik. S druge strane, u interaktivnoj metodi od korisnika se očekuje odabir područja interesa iz kojeg će se napraviti izvoz značajki. Odabir najčešće zahtijeva početnu i završnu točku ili skupinu točaka. Interaktivni sustavi su često složeniji od automatskih, a evaluirani su u [28,114]. Iako interaktivni sustavi mogu postići konkurentne rezultate, skloni su pogreškama zbog potrebne ljudske interakcije. Drugi nedostatak interaktivnih sustava je vrijeme potrebno za obradu podatkovnih skupova.

U usporedbi s klasifikacijom slike grafikona, izdvajanje podatkovnih značajki iz slike grafikona podržava znatno manje vrsta grafikona. Vrste grafikona koji se najviše obrađuju su stupčasti, linijski i kružni grafikon. Točkasti grafikon istražen je u [115,116], a površinski i radar grafikon u [28]. Razvijeni algoritmi se često oslanjaju na znanje niske razine (engl. *Low-level*), to jest položaj točaka razlučivosti u prostoru slike grafikona, a heterogenost vrsta grafikona stvara dodatne izazove. Algoritmi koji rade sa značajkama niske razine svrstavaju se u kategoriju općenitog modela, odnosno algoritam prati definiranu strukturu i očekuje pronaći elemente grafikona na unaprijed definiranim mjestima. Pronalaskom elemenata računaju se udaljenosti i omjeri od nepromjenjivih varijabli, kao što su osi koordinatnog sustava (predstavljene konačnim brojem crnih točaka razlučivosti u nizu) ili udaljenosti od središta ili rubova slike. S jedne strane koordinatnih osi nalaze se oznake koje OCR razlikuje, a s druge strane su elementi grafikona detektirani promjenom boje točke razlučivosti. Budući da kružni grafikon ne sadrže koordinatne osi, koriste se drugačije metode. Upotrebom metode slučajnog uzorkovanja (engl. *Random sample consensus – RANSAC*) moguće je detektirati promjenu boje između susjednih točaka razlučivosti, što dovodi do izvoza podatkovnih značajki kružnog isječaka, kao što je prikazano u [23,28,47]. Druga metoda za otkrivanje kružnih isječaka oslanja se na brojanje obojenih točaka razlučivosti između dvije granice koje dijele početno ishodište, kao što je prikazano u [25,117].

Umjesto detekcije svih točaka razlučivosti i raznih metoda brojanja, podatkovne značajke iz slike grafikona je moguće izdvojiti i poznavanjem samo ključnih točaka koje opisuju određene elemente grafikona. Detekcija ključnih točaka najčešće se koristi kod detekcije lica, objekata. položaja u prostoru ili položaja u kojem se nalazi ljudsko tijelo. Za razliku od prethodnog algoritma, detekcija ključnih točaka ne zahtijeva poseban algoritam za svaki tip grafikona. Ako se razmatra tortni grafikon, ključne točke su ishodište grafikona i lokacije na kružnici gdje počinje i završava kružni isječak. Algoritam napisan za tortni grafikon, primjenjiv je i na prstenaste

grafikone i grafikone sunčevih zraka, odnosno na bilo koji kružni grafikon. Korištene metode za detekciju ključnih točaka su raznolike, a uključuju:

- upotrebu binarnih maski u detekciji strukturnih elemenata, kao što je prikazano u [68,118],
- model detekcije objekta koji pronalazi četverokutnu regiju koja okružuje objekt, to jest granični okvir objekta definiran kao par ključnih točaka, gornji lijevi kut i donji desni kut, koristeći neuronsku mrežu, kao što je prikazano u [119],
- model koji koristi tenzorska polja u otkrivanju prostorne lokacije kutova gornje linije kod stupčastih grafikona i točaka u točkastom grafikonu, kao što je prikazano u [120],
- upotrebu CNN modela za detekciju objekata, kao što je prikazano u [121],
- upotrebu enkoder-dekoder mreža (engl. *Encoder-decoder network*), kao što je prikazano u [62,118] ili
- upotrebu transformatora vida, kao što je prikazano u [122].

Zajednička osobina svim navedenim metodama je potreba za dodatnim matematičkim modelom koji će omogućiti izvoz podatkovnih značajki.

Modeli dubokog učenja također mogu otkriti objekte na slikama grafikona. U radu [27] koristi se ponavljajuća neuronska mreža (engl. *Recurrent neural network – RNN*) za otkrivanje kutova u kružnim grafikonima. U radu [25] koriste se tekstualne i grafičke informacije za izdvajanje podataka. Tri različita algoritma stvorena su za tri vrste grafikona: stupčasti, linijski i tortni. Objekti se detektiraju pomoću različitih modela za detekciju objekata u prostoru (*Yolo2*, *Faster R-CNN*, *SSD*) i za svaki objekt stvaraju se granični okviri. Korištenjem različitih metoda predobrade slike unutar graničnih okvira pronalaze se vrijednosti točaka razlučivosti. Upotreba neuronskih mreža i znanja niske razine rezultira visokom točnošću izdvajanja podatkovnih značajki iz navedenih vrsta grafikona. Za izdvajanje značajki iz linijskog grafikona može se koristiti i neuronska mreža za semantičku segmentaciju (engl. *Fast Fully Convolutional Network – FastFCN*) koja predviđa mape vjerojatnosti za svaku liniju na grafikonu, kao što je prikazano u [123]. Sažetak predstavljenog istraživanja u ovom području prikazan je tablicom 3.1.

Slično kao i za klasifikaciju slike grafikona i iz tablice 3.1, usporedba rezultata je izazovna budući da svi autori ne navode ili ne koriste javno dostupne podatkovne skupove, ne definiraju organizaciju podatkovnog skupa i ne navode korištenu programsku podršku i računalne resurse. Svaka navedena metoda je jedinstvena u smislu svoje specifičnosti, a dodatno, kada se razmatraju različite vrste grafikona, često se radi o precizno određenim tipovima ili oblicima grafikona. Svaka

Tablica 3.1: Pregled istraživanja u području izdvajanja podatkovnih značajki iz slike grafikona. Istaknuti red predstavlja rad izrađen u sklopu ovog istraživanja.

Lit.	God.	Metoda	Br. vrsta	Uspješnost izdvajanja značajki	Podatkovni skup
[47]	2011	Izrada modela	2	0.70	mali
[124]	2011	Izrada modela	1	0.67	mali
[125]	2015	Izrada modela	1	0.83	mali
[126]	2016	Izrada modela	1	0.86	mali
[39]	2017	Izrada modela	1	0.98	mali
[115]	2017	Izrada modela, Duboko učenje	1	0.84	mali
[22]	2018	Izrada modela	1	0.83	mali
[23]	2018	Izrada modela	2	0.79	-
[117]	2018	Izrada modela	1	0.99	mali
[25]	2019	Izrada modela, Duboko učenje	3	0.88	mali
[27]	2019	Duboko učenje	2	0.84	-
[120]	2020	Ključne točke	2	-	mali
[45]	2021	Izrada modela	1	0.72	mali
[127]	2021	Duboko učenje	1	0.85	mali
[116]	2021	Izrada modela	3	0.90	srednji
[119]	2021	Ključne točke	3	0.93	mali
[121]	2021	Ključne točke	3	0.86	srednji
[68]	2022	Izrada modela, Ključne točke	2	0.98	mali
[62]	2023	Ključne točke	3	0.99	srednji
[118]	2023	Duboko učenje, Ključne točke	1	0.98	srednji
[122]	2023	Duboko učenje	1	0.99	veliki

metoda izdvajanja podatkovnih značajki koristi specifične pristupe i tehnike, prilagođene za rješavanje jedinstvenih izazova povezanih s određenim tipom grafikona. Ova jedinstvenost

metodologija pridonosi raznolikosti analize, ali istovremeno stvara izazov u usporedbi rezultata zbog nedostatka standardiziranih parametara ili referentnih točaka.

Podatkovni skupovi u tablici 3.1 svrstani su u tri kategorije: mali (broj korištenih slika je manji od 1000), srednji (broj slika je između 1000 i 10000) i veliki (broj slika je veći od 10000). Budući da se svi korišteni podatkovni skupovi kvalitativno i kvantitativno razlikuju, predstavljene okvirne brojke treba promatrati kao dodatnu informaciju, a ne kao referentnu točku u kojoj je za postizanje X rezultata, uz Y metodu, potrebna veličina podatkovnog skupa Z. Empirijskom analizom rezultata utvrđeno je da autori za testiranje metoda uglavnom koriste nekoliko desetaka slika grafikona. Razlog korištenja malog podatkovnog skupa je zbog ručne provjere ostvarenih rezultata i usporedbe sa stvarnim vrijednostima iz koje je nastala slika grafikona, što je vremenski zahtjevan zadatak.

3.3 Metoda izdvajanja podatkovnih značajki

U potpoglavljima 3.3.1 i 3.3.2 predstavljeni su potpuno automatizirani algoritmi niske razine koji na ulazu prihvaćaju rastersku sliku grafikona, a na izlazu isporučuju objektno orijentiranu strukturu grafikona koja se sastoji od parova atributa i vrijednosti (engl. *JavaScript Object Notation – JSON*). Algoritam podržava dvije vrste grafikona, gdje se jedna grupa sastoji od grafikona kružnog oblika, a druga grupa uključuje grafikone koji izgledom podsjećaju na mrežu. Budući da jedna vrsta grafikona može sadržavati nekoliko tipova grafikona koji sadrže različite primijenjene stilove, postoji nekoliko ograničenja. Prvo ograničenje je da ulazna slika ne smije sadržavati trodimenzionalni oblik grafikona niti bilo kakve trodimenzionalne efekte (primjerice sjene ili svjetla). Drugo ograničenje određuje da ulazna slika ne može sadržavati više grafikona. Trećim ograničenjem je određeno da svi elementi grafikona trebaju sadržavati različite boje ili granice između pojedinih elemenata moraju biti jasno definirane. Kao zadnje ograničenje postavljeno je da elementi grafikona ne smiju prelaziti preko drugih elemenata i time ometati vidno polje algoritma (primjerice naziv koordinatne osi ne smije prelaziti preko vrijednosti prisutnih na koordinatnoj osi ili naslov grafikona ne smije prelaziti preko grafičkih elemenata).

Na uspješnost cjelokupne metode, značajnu ulogu ima odabir OCR mehanizma za obradu teksta. Obrada teksta je ključan element algoritma, ali isto tako i najslabija točka. U izradi algoritma korištena je osnovna, javno dostupna, inačica Keras OCR [128] mehanizma koja sadrži unaprijed poznata ograničenja kao što su: čitanje horizontalnih i vertikalnih oznaka, čitanje numeričkih znakova i slova engleske abecede i postojanje visokog kontrasta između pozadine i

teksta. Slike grafikona opisuje heterogena struktura, a isto tako i izgled koji se sastoji od različitih grupa fontova, različitih veličina, orijentacija i usmjerenja fonta, različitih boja fonta i mogućnosti preklapanja više različitih tekstualnih i grafičkih elemenata. U slučajevima kada je tekst jednostavan i jasno pozicioniran, Keras OCR će isti detektirati s visokom pouzdanošću, no u svim ostalim slučajevima granični okvir oko teksta će biti netočno pozicioniran ili u potpunosti izostavljen. Zbog navedenih ograničenja, u ovom radu se ne uzima u obzir kvaliteta obrade teksta upotrebom OCR mehanizma. Promatra se slučaj u kojem OCR mehanizam radi s najvišom pouzdanošću i najtočnijim pozicioniranjem graničnog okvira oko tekstualnih oznaka. Umjesto Keras OCR-a moguće je koristiti bilo koji drugi OCR mehanizam kao što su Google Tesseract, Microsoft OCR ili Amazon Rekognition.

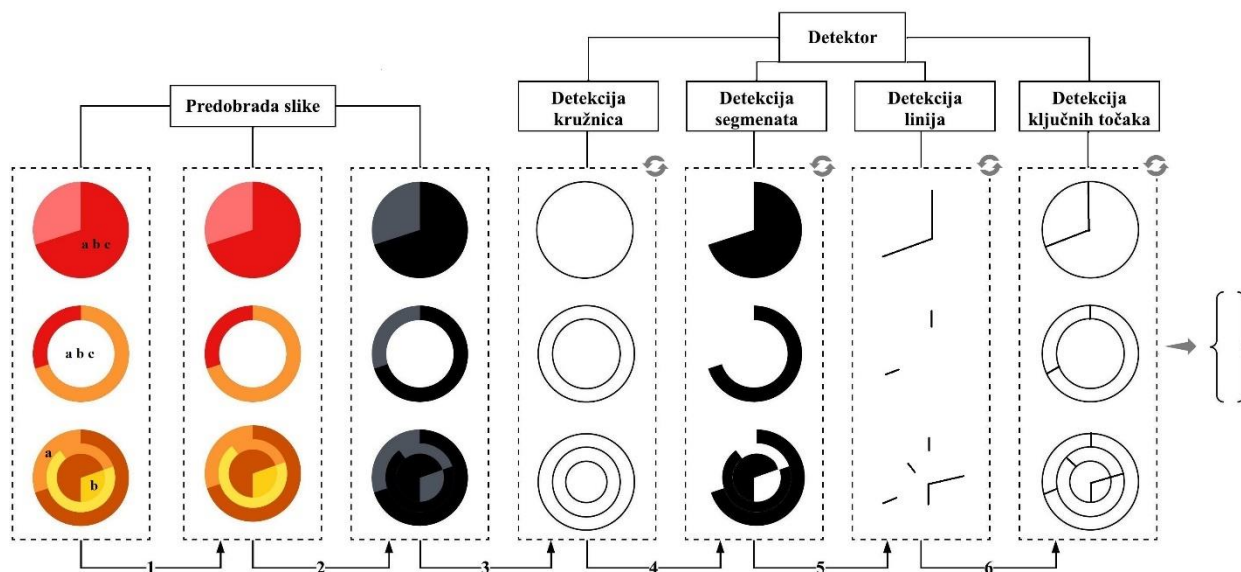
Razvijena metoda uključuje i detekciju legende. Detekcija legende temelji se na rezultatima obrađene slike grafikona. Ostatak neobrađenih obojanih elemenata na slici smatra se legendom. Legenda se nalazi u okruženju grafikona, ali ne može biti dio grafikona (primjerice preklapanje legende i elemenata grafikona nije dozvoljeno). Metoda detekcije legende opisana je u [45], a temelji se na udaljenosti obojanih elemenata od niza graničnih okvira detektiranih pomoću OCR mehanizma. Za niz graničnih okvira čije su koordinatne točke u relativnoj blizini pretpostavlja se da pripadaju legendi i sadrže numeričke ili tekstualne oznake.

Algoritam izdvajanja podatkovnih značajki iz slike grafikona je zasnovan na binarnim slikama i morfološkoj obradi slike. Morfologija označava sredstvo kojim se vrši izmjena slike s ciljem izdvajanja značajki slike korisnih za reprezentaciju i opis objekta (primjerice granice, obrubi i konveksna područja), kao što je prikazano u [129,130].

3.3.1 Kružni grafikoni

Metoda izdvajanja podatkovnih značajki iz kružnih grafikona obuhvaća tortne grafikone, prstenaste grafikone i grafikone sunčevih zraka. Metoda je prikazana slikom 3.1, sastoji se od modula za predobradu slike i detektora, te ukupno šest koraka nakon kojih se stvara objektno orijentirana struktura ulazne rasterske slike grafikona.

Prvi modul algoritma je modul za predobradu slike i sastoji se od dva koraka. Prvi korak prikazanog modula je uklanjanje tekstualnih oznaka sa slike upotrebom OCR mehanizma. Tekstualne oznake je potrebno ukloniti jer, ovisno o boji, veličini i orijentaciji fonta, mogu biti pogrešno protumačene unutar detektor modula. OCR mehanizam pronalazi tekstualna područja oko kojih stvara granični okvir. Unutar graničnog okvira obavlja se zamjena boje točaka razlučivosti s bojom iz okoline. Slika s uklonjenim tekstom se prosljeđuje u drugi korak gdje se



Slika 3.1: Prikaz metode izdvajanja podatkovnih značajki iz kružnih grafikona. Metoda se sastoji od modula za predobradu slike i detektora.

obavlja normalizacija prostora boja slike, smanjuje se količina nastalog impulsnog šuma upotrebom filtra zamućenja i nakon toga se koristi filter za izoštravanje slike kojemu je cilj pojačati granice između pojedinih segmenata (primjerice kružnih isječaka). Slika se zatim pretvara u crno-bijelu sliku s različitim intenzitetom svjetlosti. Proces predobrade ulazne slike grafikona je završen i slika se prosljeđuje u modul detektor.

Modul detektor sastoji se od četiri koraka koji uključuju detekciju kružnice, segmenata, linija i ključnih točaka. Pomoću Houghove transformacije pronalaze se kružnice čiji broj ovisi o tipu kružnog grafikona. Tortni grafikoni sadrže jednu kružnicu, prstenasti dvije, a grafikoni sunčevih zraka dvije ili više. Prednost Houghove transformacije je što može detektirati i djelomične kružnice koje su česta pojava kod grafikona sunčevih zraka. Rezultat svake detektirane kružnice su koordinate ishodišta i iznos radijusa. Nakon detekcije kružnica obavlja se detekcija segmenata. Prvo se stvara binarna maska koja usmjerava detektor na jedan kružni segment. Kvaliteta binarne maske ovisi o Houghovoj transformaciji za detekciju kružnice, te se dodatno koriste morfološke operacije kako bi uklonile nepravilnosti na slici. Svaki kružni segment sastoji se od raznih boja, te se obavlja detekcija prisutnih jedinstvenih boja. Za svaku boju stvara se binarna maska koja isključuje boje koje se trenutno ne promatraju. Binarna maska se proširuje upotrebom morfoloških operacija otvaranja i zatvaranja čime se stvara bijeli linijski obrub oko segmenta. Morfološko otvaranje smanjuje nastali šum i uklanja slabo povezane točke razlučivosti. Morfološko zatvaranje popunjava praznine u promatranom segmentu. S linijskog obruba uzorkuju se koordinatne točke koje će se koristiti u izračunima. Postupak se ponavlja za svaki segment

unutar kružnog segmenta, odnosno za svaki kružni segment. Rezultat spajanja svih binarnih maski je grafikon izrađen od primitivnih bijelih linija i kružnica. Mjesta na kojima se sijeku bijela linija i kružnica su ključne točke potrebne za izdvajanje podatkovnih značajki. Dohvaćanje vrijednosti ključnih točaka nije moguće dobiti jednostavnim iščitavanjem vrijednosti koordinata sa slike, jer gornji lijevi kut na slici predstavlja ishodište, to jest nulte koordinate. Točke razlučivosti prikazuju se s lijeve prema desnoj strani, odozgo prema dolje. Ovim načinom iščitavanja redosljed točaka na kružnom grafikonu neće odgovarati stvarnim vrijednostima na slici, stoga se iščitavanje ključnih točaka obavlja u odnosu na ishodište kružnice, jer će tako ostati zadržana stvarna prostorna raspodjela.

Nakon što su prikupljene sve ključne točke, vrijednosti se prosljeđuju u matematički model koji podatke isporučuju u JSON formatu. U matematičkom modelu koriste se sljedeće jednadžbe:

- Izračun koeficijenta smjera ili nagiba pravca m gdje su x_1, y_1 koordinatne točke ishodišta kružnog segmenta, a x_2, y_2 koordinate ključnih točki iz detektor modula, prikazan jednadžbom 3.1

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad (3.1)$$

- Izračun kuta θ između dvije točke, prikazan jednadžbom 3.2

$$\theta = \arctan(m) \quad (3.2)$$

- Izračun udaljenosti d između dvije točke, prikazan jednadžbom 3.3

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.3)$$

- Izračun površine P_i kružnog isječka, prikazan jednadžbom 3.4

$$P_i = \frac{\theta}{360} \times \pi \times d^2 \quad (3.4)$$

- Izračun kuta φ između dva vektora, prikazan jednadžbom 3.5

$$\varphi = \arccos \left(\frac{x_1 \times y_1 + x_2 \times y_2}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}} \right) \quad (3.5)$$

- Izračun površine kruga P_k , gdje je r vrijednost radijusa kružnog segmenta ostvarena Houghovom transformacijom, prikazan jednadžbom 3.6

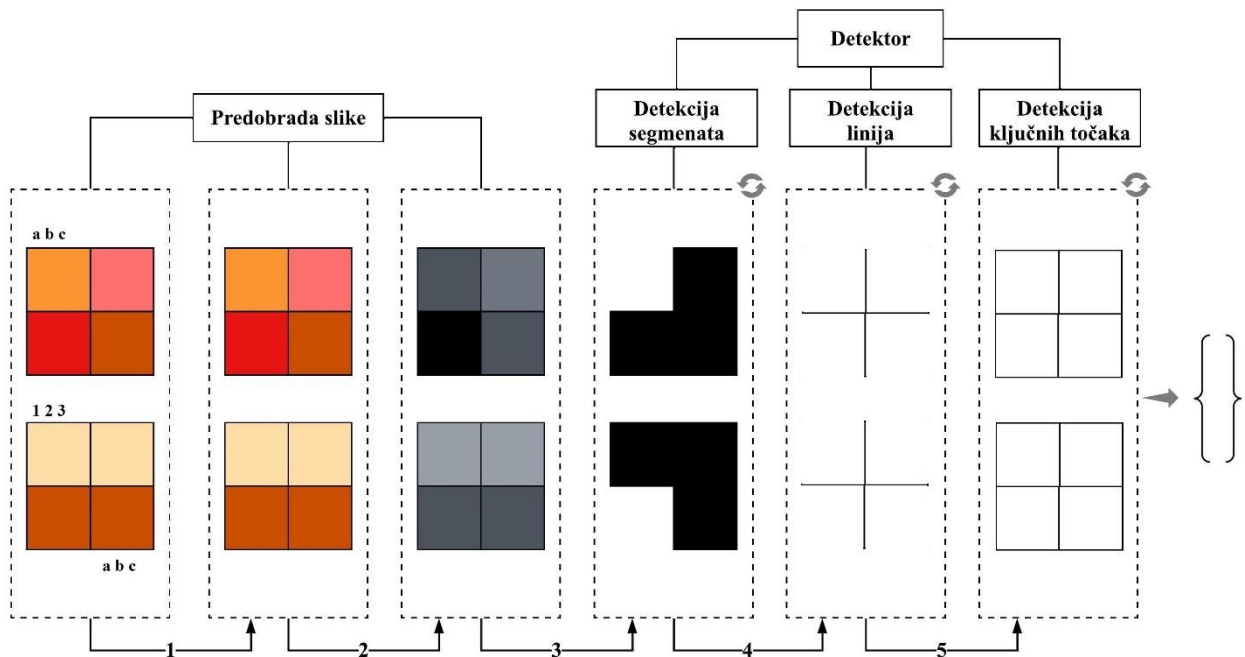
$$P_k = r^2 \times \pi \quad (3.6)$$

Upotrebom navedenih jednačbi izračunavaju se sve vrijednosti koje opisuju kružni grafikon. Izračunate vrijednosti i ključne točke se isporučuju u JSON formatu te omogućuju potpunu rekonstrukciju izvornog grafikona kao što je vidljivo u potpoglavlju evaluacije 3.4.1.

3.3.2 Mrežasti grafikoni

Metoda izdvajanja podatkovnih značajki iz mrežastih grafikona obuhvaća toplinske karte i grafikone napretka. Metoda je prikazana slikom 3.2, a sastoji se od modula za predobradu slike i detektora, te ukupno pet koraka nakon kojih se stvara objektno orijentirana struktura ulazne rasterske slike grafikona.

Prvi modul algoritma je modul za predobradu slike koji se sastoji od dva koraka čija je funkcionalnost jednaka modulu za predobradu slike kružnih grafikona prikazana u potpoglavlju 3.4.1. Strukturna razlika između toplinskih karti i grafikona napretka je zanemariva. Toplinske karte često prikazuju širok spektar boja, gdje svaka boja može zauzeti bilo koje mjesto na karti, dok su kod grafikona napretka boje obično ograničene na dvije i grupirane. Primjera radi, kada se promatra varijabla koja predstavlja popunjenost nekog elementa, donji dio grafikona napretka koji označava popunjenost obično će biti prikazan tamnijim bojama, dok će gornji dio grafikona koji označava prazninu biti prikazan svjetlijim bojama.



Slika 3.2: Prikaz metode izdvajanja značajki iz mrežastih grafikona. Metoda se sastoji od modula za predobradu slike i detektora.

Nakon završetka predobrade, slika grafikona se prosljeđuju u modul detektor.

Modul detektor sastoji se od tri koraka koji uključuju detekciju segmenata, linija i ključnih točaka. Prvi korak je detekcija segmenata. Na početku se stvara binarna maska koja usmjerava detektor na jedan četverokutni segment. Na segmentu se provode morfološke operacije otvaranja i zatvaranja kako bi uklonile nepravilnosti na slici. Detektira se jedinstvena boja segmenta i stvara se binarna maska koja isključuje segmente koji se trenutno ne promatraju. Rezultat navedenih operacija je bijeli linijski obrub oko segmenta. Postupak se ponavlja za svaki segment. Rezultat spajanja svih binarnih maski je grafikon izrađen od primitivnih bijelih linija. Linije se detektiraju upotrebom Houghove transformacije za detekciju linija. Rezultat Houghove transformacije je linijski obrub s kojeg se uzorkuju koordinatne točke koje će se koristiti u izračunima. Mjesto na kojima se nalazi najveći broj bijelih linija (čvorište) je ključna točka potrebna za izdvajanje podatkovnih značajki. Za razliku od kružnih grafikona, ovdje nije potrebna reorganizacija ključnih točaka. Ključne točke se iščitavaju s lijeva na desno, što predstavlja redak, odnosno odozgo prema dolje, što predstavlja stupac mrežastog grafikona.

Nakon što su prikupljene sve ključne točke, vrijednosti se prosljeđuju u matematički model koji podatke isporučuju u JSON formatu. U matematičkom modelu koriste se sljedeće jednadžbe:

- Izračun udaljenosti d između dvije točke gdje su x_1, y_1 koordinatne točke čvorišta od n retka, a x_2, y_2 koordinate točke čvorišta od $n+1$ retka mrežaste strukture grafikona, prikazan jednadžbom 3.3
- Izračun koordinata središnje točke segmenta (x_{12}, y_{12}) , prikazan jednadžbom 3.7

$$\begin{aligned}x_{12} &= \frac{x_1 + x_2}{2} \\y_{12} &= \frac{y_1 + y_2}{2}\end{aligned}\tag{3.7}$$

- Izračun površine segmenta P_s , gdje je d_v visina segmenta, a d_s širina segmenta, prikazan jednadžbom 3.8

$$P_s = d_v \times d_s\tag{3.8}$$

Upotrebom navedenih jednadžbi izračunavaju se sve vrijednosti koje opisuju mrežasti grafikon. Izračunate vrijednosti i ključne točke se isporučuju u JSON formatu te omogućuju potpunu rekonstrukciju izvornog grafikona kao što je vidljivo u potpoglavlju evaluacije 3.4.2.

3.4 Eksperimentalna evaluacija

Proces eksperimentalne evaluacije je izazovan jer ne postoji temeljna osnova pomoću koje je moguće raditi usporedbu empirijskih rezultata s ostalim istraživanjima u području. Razni objavljeni radovi prijavljuju različite vrijednosti koje se mogu smatrati izdvojenim podatkovnim značajkama. Dodatno, evaluacije su rađene na raznim podatkovnim skupovima, a pojedine metode su primjenjive za strogo definirane vrste i tipove grafikona. Predstavljeni izazov dokumentiran je i u radovima [7,10,11]. U sklopu ovog istraživanja, izvoz podatkovnih značajki podrazumijeva dohvat izvornih vrijednosti koje su korištene u procesu izrade grafikona. Za eksperimentalnu evaluaciju predložene metode na kružnim i mrežastim grafikonima koristi se metoda primijenjena u [25,28,127]. Evaluacija je provedena na ChartDS i ICDAR7 podatkovnom skupu gdje se iz ICDAR7 podatkovnog skupa koriste samo tortni i prstenasti grafikoni. U postupku evaluacije koristi se 100 slika grafikona po tipu grafikona. Slike grafikona su ručno provjerene kako bi zadovoljavale uvjete postavljene u potpoglavlju 3.3.

Metoda procjenjuje uspješnost algoritma na temelju koeficijenta uspješnosti. Koeficijent uspješnosti je udio ispravno obrađenih slika grafikona i ukupnog broja slika grafikona u promatranom tipu grafikona. Slika grafikona uspješno je obrađena samo ako je njezin koeficijent pogreške svih izdvojenih podatkovnih značajki unutar postavljenog praga. Točnost izdvojenih podatkovnih značajki grafikona izračunava se pomoću jednadžbe prikazane u 3.9

$$\frac{|g_i - p_i|}{g_i} \leq \varepsilon \quad (3.9)$$

gdje je g_i stvarna vrijednost segmenta (primjerice vrijednost korištena za izradu kružnog isječka ili četverokuta), p_i je metodom izračunata vrijednost segmenta, a ε je vrijednost praga pogreške. Vrijednost praga pogreške kontrolira kvalitetu izdvojenih podatkovnih značajki iz slike grafikona. Niža vrijednost praga pogreške rezultira većom točnošću izdvajanja podatkovnih značajki iz slike grafikona. Ako je izdvajanje podatkovnih značajki iz slike grafikona unutar postavljene vrijednosti praga, slika se smatra ispravno obrađenom; u suprotnom, slika nije ispravno obrađena. Prosječna točnost metode izdvajanja podatkovnih značajki iz slike grafikona P_t se zatim izračunava pomoću jednadžbe prikazane u 3.10

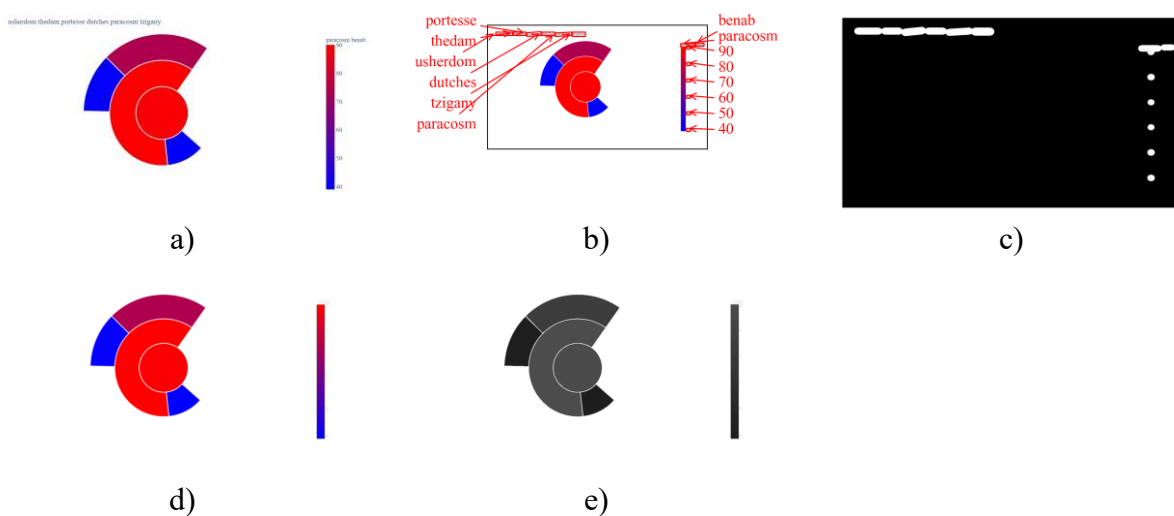
$$P_t = \frac{N_c}{N_t} \quad (3.10)$$

gdje je N_c broj točno obrađenih slika grafikona, a N_t ukupan broj slika po tipu grafikona ($N_t = 100$). Drugim riječima, vrijednost praga pogreške koja je manja ili jednaka od 0.05 ($0.05 \leq \varepsilon$) odgovara koeficijentu pogreške od 0.05. Vrijednost praga pogreške je odabrana s obzirom na provedena istraživanja u [25,127].

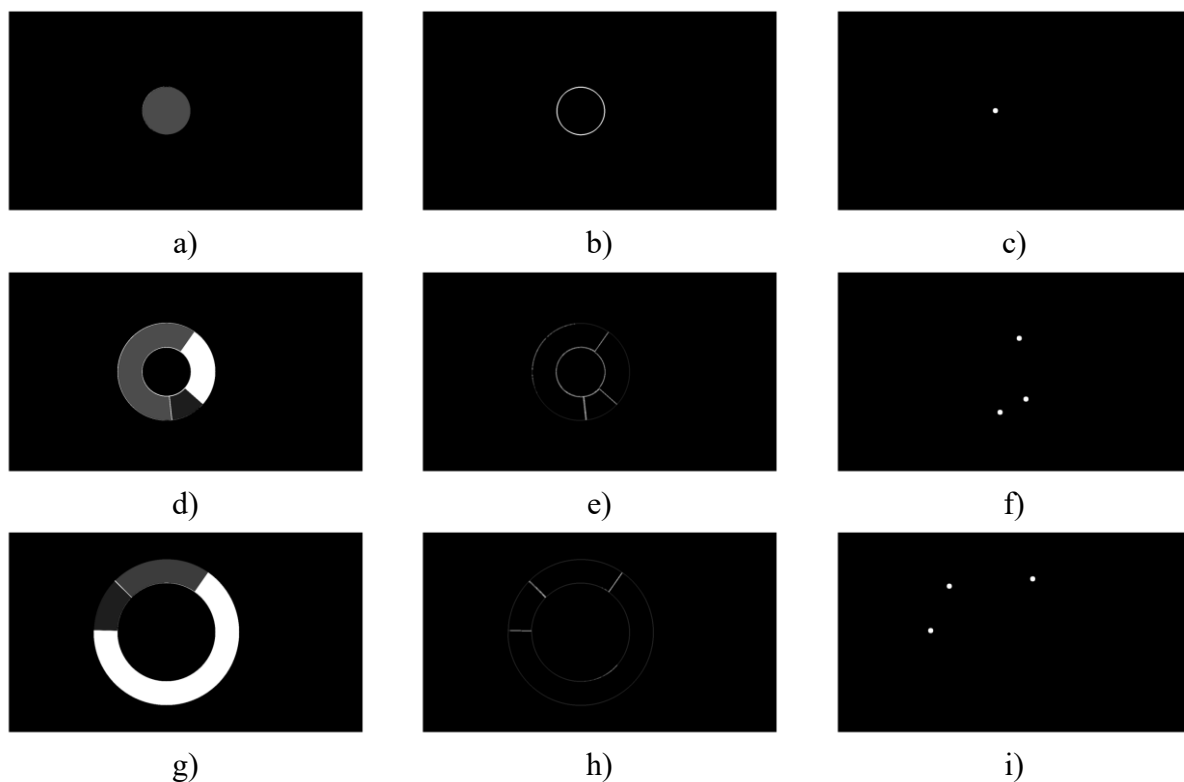
3.4.1 Kružni grafikoni

Detaljan proces evaluacije jednog kružnog grafikona (grafikon sunčevih zraka) prikazan je slikama 3.3 i 3.4. Slike prate korake iz slike 3.1.

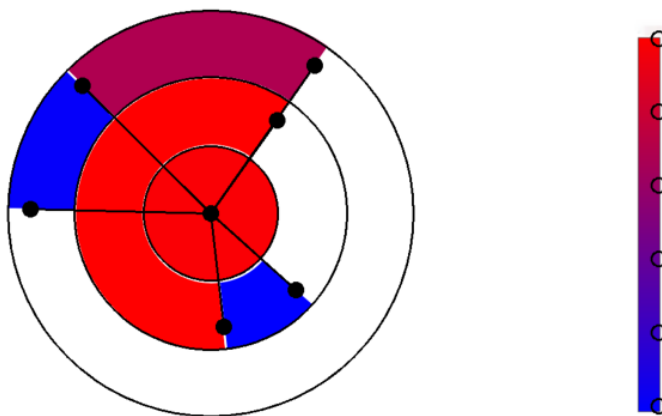
Na ulazu modula za predobradu slike kružnog grafikona, prikazan slikom 3.3, nalazi se grafikon čija struktura se sastoji od jednog cijelog i dva djelomična kružna segmenata, naslova i legende (označeno kao a). Preko ulazne slike grafikona provodi se OCR mehanizam koji stvara granične okvire oko tekstualnih elemenata (označeno kao b). Za svaki detektirani tekstualni element, stvara se binarna maska (označeno kao c), a rezultat je slika grafikona bez tekstualnih oznaka (označeno kao d). Slika bez tekstualnih oznaka se pretvara u crno-bijelu sliku s različitim intenzitetom svijetlosti čime završava modul predobrade slike (označeno kao e). Slika se prosljeđuje u modul detektor koji sadrži ponavljajuće korake čiji broj ovisi o broju kružnih segmenata, prikazan slikom 3.4. Prvo se detektiraju kružnice (označeno kao a, d, g). Potom se provodi detekcija pojedinačnih segmenata i linija (označeno kao b, e, h). Na kraju se provodi detekcija ključnih točaka (označeno kao c, f, i). Prikupljeni podatci se prosljeđuju u matematički model koji podatke isporučuju u JSON formatu. Konačni rezultat, sadržaj datoteke, dostupan je u prilogu P.4.



Slika 3.3: Tijek grafikona sunčevih zraka kroz modul za predobradu slike kružnog grafikona. Na slici a) ulazna slika, b) detekcija tekstualnih elemenata, c) stvaranje binarne maske, d) uklanjanje tekstualnih elemenata i e) izlazna slika modula.



Slika 3.4: Tijek grafikona sunčevih zraka kroz modul detektor kružnog grafikona. Na slici je prikazana detekcija kružnice (a, d, g), detekcija linija (b, e, h) i detekcija ključnih točaka (c, f, i).



Slika 3.5: Rekonstrukcija grafikona sunčevih zraka upotrebom izdvojenih podatkovnih značajki. Ključne točke i linije su označene crnom bojom. Algoritam je uspješno detektirao i pripadajuću legendu.

Na slici 3.5, vidljiva je potpuna rekonstrukcija ulazne slike grafikona na osnovu podataka zapisanih u datoteci koja uključuje i legendu. Evaluacija je napravljena s obzirom na

Tablica 3.2: Rezultati evaluacije metode izdvajanja podatkovnih značajki na kružnim grafikonima

Podatkovni skup	Tip grafikona	Broj slika grafikona unutar $\varepsilon = 0.05$
ChartDS	Tortni	99/100
	Prstenasti	98/100
	Sunčeve zrake	97/100
ICDAR7	Tortni	98/100
	Prstenasti	98/100

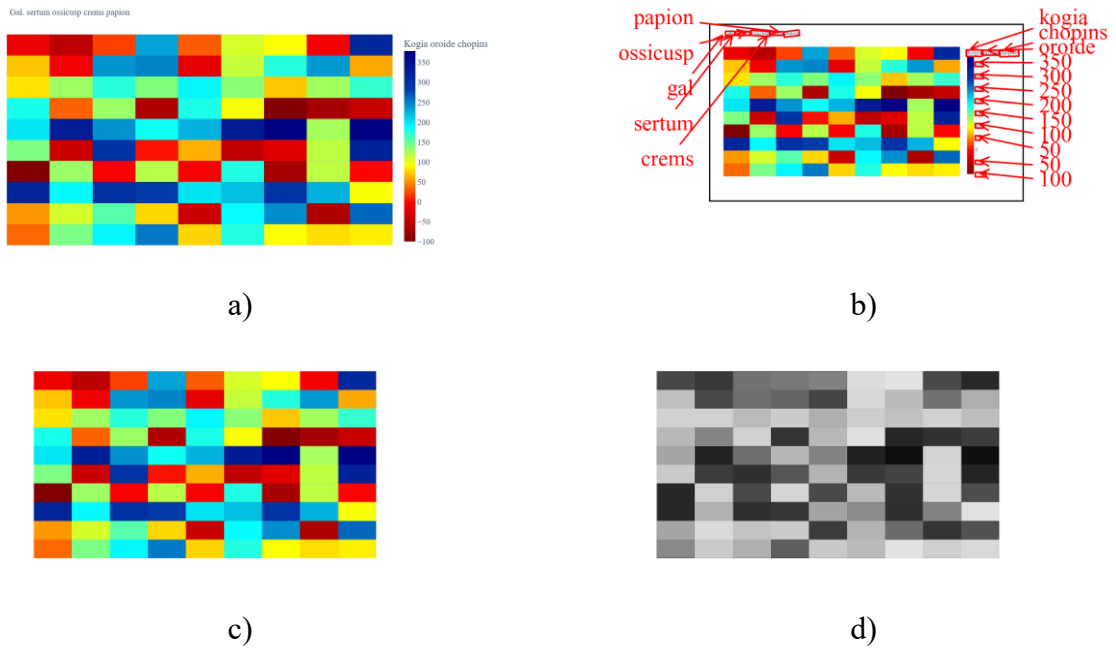
vrijednost praga pogreške $0.05 \leq \varepsilon$. Rezultati evaluacije za ChartDS i ICDAR7, prikazani su tablicom 3.3.

Za umjetno stvorene podatkovne skupove (ChartDS i ICDAR7) broj točno obrađenih grafikona N_c je visok, a prosječna točnost Pt iznosi 0.98. U prilogu P.5 nalaze se dodatni primjeri raznih kružnih grafikona koji prikazuju uspješnost predstavljene metode.

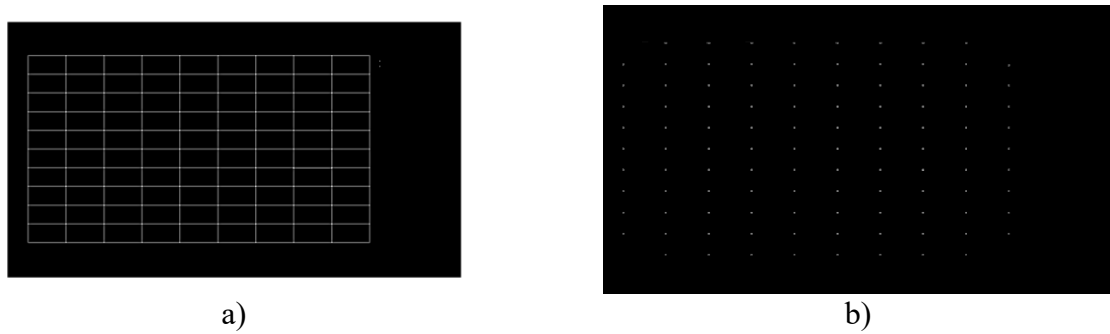
3.4.2 Mrežasti grafikoni

Detaljan proces evaluacije jednog mrežastog grafikona (toplinska karta) prikazan je slikama 3.6 i 3.7. Slike prate korake iz slike 3.2.

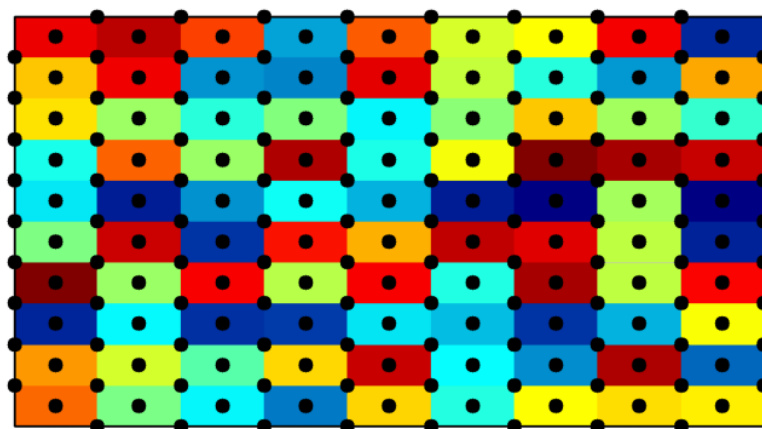
Na ulazu modula za predobradu slike mrežastog grafikona, prikazan slikom 3.6, nalazi se grafikon čija struktura se sastoji od devet stupaca, deset redaka, naslova i legende (označeno kao a). Preko ulazne slike grafikona provodi se OCR mehanizam koji stvara granične okvire oko tekstualnih elemenata (označeno kao b). Za svaki detektirani tekstualni element, stvara se binarna maska, a rezultat je slika grafikona bez tekstualnih oznaka (označeno kao c). Slika bez tekstualnih oznaka se pretvara u crno-bijelu sliku s različitim intenzitetom svjetlosti čime završava modul predobrade slike (označeno kao d). Slika se prosljeđuje u modul detektor koji sadrži ponavljajuće korake čiji broj ovisi o broju segmenata opisanih redcima i stupcima, prikazano slikom 3.7. Na početku se obavlja detekcija segmenata i linija (označeno kao a). Nakon stvaranja mrežaste strukture provodi se detekcija ključnih točaka (označeno kao b). Prikupljeni podatci se prosljeđuju u matematički model koji podatke isporučuju u JSON formatu. Konačni rezultat, sadržaj datoteke, dostupan je u prilogu P.6. Na slici 3.8, vidljiva je potpuna rekonstrukcija ulazne slike grafikona na osnovu podataka zapisanih u datoteci.



Slika 3.6: Tijek grafikona toplinske karte kroz modul za predobradu slike mrežastog grafikona. Na slici a) ulazna slika, b) detekcija tekstualnih elemenata, c) uklanjanje tekstualnih elemenata i d) izlazna slika modula



Slika 3.7: Tijek grafikona toplinske karte kroz modul detektor kružnog grafikona. Na slici a) detekcija linija i b) detekcija ključnih točaka



Slika 3.8: Rekonstrukcija grafikona toplinske karte upotrebom izdvojenih podatkovnih značajki. Ključne točke i linije su označene crnom bojom.

U trenutku izrade metode, niti jedan javno dostupan podatkovni skup ne sadrži mrežaste grafikone, te je evaluacija napravljena na podatkovnom skupu ChartDS s obzirom na vrijednost praga pogreške $0.05 \leq \varepsilon$. Rezultati evaluacije prikazani su tablicom 3.5.

Tablica 3.3: Rezultati evaluacije metode izdvajanja značajki na mrežastim grafikonima

Podatkovni skup	Vrsta grafikona	Broj slika grafikona unutar $\varepsilon = 0.05$
ChartDS	Mrežasti	98/100

Kod umjetno stvorenih mrežastih grafikona ne postoji strukturalna razlika između toplinskih karti i grafikona napretka, te su oba tipa u tablici 3.5 objedinjena s nazivom mrežasti grafikon. Broj točno obrađenih umjetno stvorenih grafikona N_c je visok, a prosječna točnost Pt iznosi 0.98. U prilogu P.7 nalaze se dodatni primjeri raznih mrežastih grafikona koji prikazuju uspješnost predložene metode.

3.4.3 Slike prikupljene internetskim pretraživanjem

Na stvarnim slikama grafikona prikupljenim putem internetskog pretraživanja, točnost izdvajanja podatkovnih značajki značajno opada, kako je prikazano tablicom 3.4. Kružni i mrežasti grafikoni ovise o alatu kojim su stvoreni, pri čemu svaki alat definira strukturu grafikona na svoj način. Stvarne slike često sadrže nečistoće i šum koji se ne može ukloniti predloženom metodom. Osim toga, u neposrednoj blizini grafikona često se pojavljuju dodatni grafički elementi koji nisu dio grafikona ili dodatni tekstualni elementi koji ometaju vidno polje predstavljene metode. Heterogenost vrsta i tipova grafikona dodatno predstavlja izazov. Na primjer, grafikoni sunčevih zraka i toplinskih karata često nemaju standardne oblike te se mogu pojaviti u nepravilnim oblicima poput različitih elipsoida ili raznih poligonalnih elemenata.

Tablica 3.4: Rezultati evaluacije metode izdvajanja značajki na stvarnim grafikonima

Podatkovni skup	Vrsta grafikona	Broj slika grafikona unutar $\varepsilon = 0.05$
Stvarne slike	Tortni	13/20
	Prstenasti	12/20
	Sunčeve zrake	7/20
	Mrežasti	8/20
	Tortni	13/20

Navedene izazove prepoznali su i drugi autori u području istraživanja te ne navode evaluaciju na slikama prikupljenim internetskim pretraživanjem, osim rada [27]. Prema istraživanju provedenom u [27], metoda koja ostvaruje točnost izdvajanja podatkovnih značajki od 88% na umjetno stvorenim slikama grafikona, pokazuje smanjenje točnosti na 60% kada se testira na slikama prikupljenim internetskim pretraživanjem, pri čemu je testiranje provedeno na deset slika grafikona.

Evaluacijom metode na slikama prikupljenim putem internetskog pretraživanja potvrđena je sposobnost rada metode. Međutim, kako bi se postigla veća točnost izdvajanja podatkovnih značajki, potrebno je dodatno optimizirati parametre unutar svakog pojedinog modula, uključujući parametre svakog koraka. Optimizacija parametara nije jednostavan zadatak jer zahtijeva primjenu metode pokušaja i pogrešaka. Proces optimizacije može uključivati prilagodbu parametara pomoću kojih se stvara binarna maska i uklanjaju tekstualne oznake, prilagodbu parametara morfoloških operacija i prilagodbu parametara Houghove transformacije. Važno je napomenuti da optimizacija parametara nije jednokratni postupak, već iterativni proces koji zahtijeva kontinuirano testiranje i prilagođavanje metode kako bi se postigla najveća točnost izdvajanja podatkovnih značajki. Osim toga, potrebno je uzeti u obzir moguće interakcije između parametara, jer promjena jednog parametra može utjecati na druge dijelove metode (primjerice rezultat Houghove transformacije utječe na detekciju ključnih točaka). Stoga, ovaj proces zahtijeva temeljitu analizu i pažljivo eksperimentiranje kako bi se postigli optimalni rezultati.

3.5 Zaključak

Predstavljena metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona dokazano je učinkovita, ali postoje poznata ograničenja. Značajno ograničenje metode je obrada teksta. Obrada teksta koristi se za uklanjanje tekstualnih oznaka i izdvajanje legende te je ključni dio predložene metode i najslabija točka. Razlog je upotreba OCR mehanizma opće namjene. Zbog navedenog, kvaliteta obrade teksta se ne razmatra. U većini slučajeva, kada je tekst jednostavan i ograničen na jednu riječ, OCR mehanizam detektira tekst s visokom pouzdanošću. U slučajevima kada se tekstualni element nalazi preko tamnije obojane pozadine ili prelazi granicu između dva promatrana segmenta, dio teksta ili cijeli tekstualni element ostane izostavljen te je potrebna ručna interakcija. U eksperimentalnoj evaluaciji promatra se primjer u kojem OCR mehanizam radi bez ograničenja prepoznavanja tekstualnih elemenata. OCR mehanizme kao što su Google Tesseract, Microsoft OCR ili Amazon Rekognition moguće upotrijebiti i testirati umjesto Keras OCR. Iako

su navedeni mehanizmi učeni na brojnim slikama, još uvijek može postojati ograničenje kod prepoznavanja tekstualnih elemenata na slikama grafikona koje je uvjetovano heterogenom strukturom i brojnim korištenim fontovima.

Prilikom evaluacije slike grafikona kružnog oblika, cijeli proces ovisi o Houghovoj transformaciji za detekciju kružnice, koja zahtijeva precizno podešavanje. Glavna prednost Houghove transformacije je detektiranje djelomičnih kružnih elemenata, koji su standardni dio strukture grafikona sunčevih zraka. Prvo poznato ograničenje Houghove transformacije je nemogućnost detektiranja više kružnica s istim ishodištem iz početne obrade slike; umjesto toga, potrebno je istu sliku obrađivati više puta. Drugo poznato ograničenje uključuje širinu granice kružnice. Ako je granica kružnice široka, Houghova transformacija može detektirati dvije kružnice (primjerice unutarnju i vanjsku), što utječe na detekciju kružnih segmenata, te je potrebno uključiti dodatne provjere ispravnosti. Kružni segmenti identičnih boja mogu proći nedetektirani ukoliko granica između nije jasno definirana. Za rješenje navedenih izazova, umjesto Houghove transformacije za detekciju kružnice moguće je koristiti posebno učene CNN-ove.

Izazov je vidljiv i u slikama mrežastih grafikona. U slučaju kada dva susjedna segmenta sadrže identičnu boju, granica između segmenata i ključnih točaka ostaje nedetektirana. Provodeći dodatne izračune, granice je moguće umjetno stvoriti sve dok su poznate koordinatne točke prvog i posljednjeg retka i dimenzija (visina i širina) jednog segmenta. Umjesto Houghove transformacije za detekciju linija i ovdje je moguće koristiti CNN-ove i značajno smanjiti broj potrebnih izračuna.

Rezultati pokazuju visoku točnost na umjetno stvorenim podatkovnim skupovima, ali točnost se značajno smanjuje kada se koriste slike prikupljene internetskim pretraživanjem. Detekcija segmenata i linija uspješno se provodi samo u malom podskupu slika prikupljenim internetskim pretraživanjem što onemogućuje detekciju ključnih točaka. Bez ovih elemenata ne mogu se napraviti potrebni izračuni što rezultira nemogućnošću izdvajanja podatkovnih značajki. Evaluirana metoda je prilagodljiva i parametre je moguće prilagoditi sukladno potrebama. Rezultat, datoteka koja sadrži JSON format, sadrži detaljan podatkovni skup koji omogućuje rekonstrukciju ulazne slike grafikona. Osim za potrebe rekonstrukcije grafikona, datoteka se može koristiti i s drugim računalnim sustavima koji mogu omogućiti pretvorbu grafikona u tekstualni opis, stvoriti sažetak sadržajne informacije grafikona ili omogućiti poveznicu s pomagalicama za osobe s oštećenim vidom.

4 METODA GENERIRANJA ISPITNIH SKUPOVA ZA STANDARDIZIRANU VALIDACIJU USPJEŠNOSTI IZDVAJANJA PODATKOVNIH ZNAČAJKI IZ KRUŽNIH I MREŽASTIH GRAFIKONA

U ovom poglavlju biti će opisana metoda generiranja ispitnih skupova za standardizirano vrednovanje uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona. Podatkovni skup nastao upotrebom ove metode objavljen je u podatkovnim repozitorijima Figshare [67] i Srce Dabar [69]. Metoda stvara dvije vrste umjetnih slika grafikona; vrstu na kojima su elementi grafikona vidljivo označeni i vrstu koja pokušava oponašati razne slike grafikona, koje je moguće prikupiti internet tražilicama, gdje nisu svi elementi prisutni ili se nalaze na neočekivanim lokacijama. Sve korištene vrijednosti za izradu grafikona bilježe se u zasebnu datoteku u JSON formatu koja omogućava dodatnu evaluaciju, usporedbu ili rekonstrukciju izvorne slike. Metoda stvara raznolike slike grafikona s naglaskom na slučajni odabir lokacija

ispisa, vrijednosti, boja, simbola, tekstualnih i numeričkih oznaka. Nastali podatkovni skup obuhvaća nekoliko vrsta grafikona, pružajući raznoliku varijaciju izgleda grafikona što ga čini svestranim alatom za procjenu i poboljšanje performansi sustava u različitim područjima koji se bave interpretacijom slike grafikona. Predloženim podatkovnim skupom uspostavljeno je jedinstveno ishodište za evaluaciju, procjenu i usporedbu učinkovitosti metoda i algoritama.

4.1 Uvod

U današnjem svijetu koji se temelji na podacima značajnu ulogu predstavlja razumijevanje informacije sadržane u složenom podatkovnom skupu, kao što je prikazano u [131]. Složeni podatkovni skup može biti izrađen od raznih tipova podataka koji uključuju numeričke oznake, tekstualne oznake, slike ili razni audio/video sadržaj. Tako izrađen podatkovni skup nije jednostavan za pregledavanje, te se za opisivanje istog najčešće koriste sumarne vrijednosti predstavljene vizualizacijom podataka, odnosno grafikonom.

Izdvajanje podatkovnih značajki iz slike grafikona, klasifikacija slike grafikona u jednu od mnogih klasa grafikona ili generiranje tekstualnih opisa i sumarnih sadržaja samo su neka područja istraživanja koja zahtijevaju raznolike podatkovne skupove sadržane od raznih vrsta i tipova grafikona. Navedena područja istraživanja koriste razne metode kako bi ostvarile što veću uspješnost u definiranom zadatku, a za to su potrebni podatkovni skupovi organizirani u dvije primarne skupine; slike za učenje i slike za testiranje. Dodatno, ovisno o korištenoj metodi i namjeni, nije dovoljno posjedovati nagomilane podatke, već je potrebna organizacija i označenost podatkovnog skupa, kao što je prikazano u [11,12,132]. Izrada organiziranog i označenog podatkovnog skupa grafikona je složen i dugotrajan proces. Većina objavljenih radova u području istraživanja koristi male ili srednje, često privatne podatkovne skupove, što je vidljivo iz tablice 2.1, u potpoglavlju 2.2 i tablici 3.1, u potpoglavlju 3.2. Rijetki autori izradili su podatkovni skup i učinili ga javno dostupnim. Dostupnost standardiziranog podatkovnog skupa je od velike važnosti u procesu učenja i testiranja raznih metoda i algoritama. Usporedba rezultata u predstavljanim tablicama je izazovna jer autori nemaju jednako polazište za provedbu predstavljenih analiza.

Budući da su svi korišteni podatkovni skupovi kvalitativno i kvantitativno različiti, predstavljene brojke treba smatrati dodatnim informacijama, a ne čvrstom referentnom točkom koja uvjetuje uspješnost metode na osnovu veličine podatkovnog skupa. Usporedba rezultata istraživanja na podatkovnim skupovima različitih veličina i sadržaja može se promatrati kao usporedba krušaka i jabuka. Iako je preporučeno koristiti javno dostupne podatkovne skupove,

mnogi autori koriste privatne podatkovne skupove, kojima često nedostaje složenost i raznolikost grafikona koji se pojavljuju u stvarnom svijetu. Također, postoji značajna razlika između podatkovnih skupova nastalih ručnim prikupljanjem (primjerice putem internet tražilica) i onih stvorenih korištenjem unaprijed definiranih parametara. Metode i algoritmi prilagođeni za umjetno stvorene podatkovne skupove, često ostvaruje lošije rezultate uspješnosti na slikama nastalim ručnim prikupljanjem te zahtijevaju dodatno podešavanje.

Javno dostupni podatkovni skupovi su vremenski ovisni. Podatkovni skupovi koji su dostupni u trenutku provedbe ovog istraživanja nemaju jamstvo za dostupnost u budućnosti. Većina podatkovnih skupa koja je nastala prije provedbe ovog istraživanja više nije dostupna. Neki od podatkovnih skupova su nedostupni zbog promjene politike pravila pristupa, a neki podatkovni skupovi, koji su se oslanjali na vanjske poveznice u prikupljanju podataka, više nisu potpuni. Svaki podatkovni skup sadrži ograničen broj vrsta i tipova grafikona, što je često nedostatno za određene zadatke koji se bave interpretacijom slike grafikona. Ovisno o tipu obrade slike grafikona, uz originalnu sliku poželjno je imati i originalne parametre iz kojih je nastala slika grafikona.

Za rješenje predstavljenih izazova izrađena je metoda koja omogućava parametarsku izradu različitih vrsta grafikona, a time i polazište za provedbu predstavljenih analiza uspješnosti. Podatkovni skup, ChartDS, korišten u dva prethodna poglavlja 2 i 3 nastao je upotrebom ove metode.

U potpoglavlju 4.2 dan je povijesni pregled istraživanja iz područja metoda za generiranje ispitnih skupova i predstavljen je pregled postojećih podatkovnih skupova. Potpoglavlje 4.3 predstavlja metodu generiranja ispitnih skupova za standardizirano vrednovanje uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona. Usporedba ChartDS i ICDAR7 podatkovnih skupa dostupna je u potpoglavlju 4.4. Zaključak i osvrt na predstavljenu metodu napisani su u potpoglavlju 4.5.

4.2 Istraživanja u području dostupnih podatkovnih skupova koji sadrže slike grafikona

Metode za statističku analizu, usporedbu, klasifikaciju i obradu slike grafikona zahtijevaju podatkovne skupove. U ovom potpoglavlju dan je kratki povijesni pregled objavljenih podatkovnih skupova koji sadrže slike grafikona. Podatkovni skupovi su razvrstani od najstarijeg

prema najnovijem te je uz svaki podatkovni skup pridružen niz opisnih detalja uključujući i javnu dostupnost, prikazano tablicom 4.1.

Empirijskom analizom objavljenih radova koji navode izradu podatkovnog skupa, uočene su tri metode stvaranja istog:

- Ručna metoda koja podrazumijeva u potpunosti ručni proces prikupljanja, analiziranja, organiziranja i označavanja svih slika
- Polu-automatska metoda koja podrazumijeva automatizaciju jednog ili više dijelova procesa, najčešće se organizacija i označavanje provodi automatski upotrebom posebno prilagođenih neuronskih mreža
- Automatska metoda koja predstavlja parametarski pristup izradi podatkovnog skupa koji ne zahtjeva dodatnu interakciju pojedinca

Tablica 4.1: Pregled istraživanja u području. Istaknuti red predstavlja rad izrađen u sklopu ovog istraživanja.

Podatkovni skup	Godina	Br. slika ($\times 10^3$)	Br. vrsta	Metoda prikupljanja	Javno dostupan
ReVision [47]	2011	2	10	Polu-automatska	Ne
View [48]	2012	0.3	3	Ručna	Ne
[59]	2013	1.7	3	Ručna	Ne
[32]	2015	5	5	Ručna	Ne
[32]	2015	5	5	Ručna	Ne
FigureSeer [21]	2016	60	7	Polu-automatska	Ne
ChartSeer [133]	2016	60	7	Polu-automatska	Ne
ChartSense [28]	2017	7	10	Ručna	Ne
DVQA [134]	2018	300	1	Automatska	Ne
ChartText [23]	2018	6	2	Automatska	Ne
ChartDecoder [22]	2018	11	5	Ručna	Ne
FigureQA [135]	2018	100	5	Automatska	Da
DocFigure [26]	2019	30	28	Polu-automatska	Da
ICDAR7 [55]	2019	200	7	Automatska	Da
SimulChart [27]	2019	2.5	2	Polu-automatska	Ne

Podatkovni skup	Godina	Br. slika ($\times 10^3$)	Br. vrsta	Metoda prikupljanja	Javno dostupan
LeafQA[136]	2019	250	5	Polu-automatska	Ne
[137]	2020	21	13	Polu-automatska	Ne
PlotQA[138]	2020	225	3	Polu-automatska	Da
ICPR2020 [56]	2021	25	11	Automatska	Da
ChartOCR [119]	2021	400	3	Polu-automatska	Ne
SciCap[139]	2021	400	5	Polu-automatska	Da
ChartDS [68]	2022	120	13	Automatska	Da
ICPR2022 [140]	2022	33	16	Polu-automatska	Da

U tablici 4.1, nalazi se značajan broj objavljenih radova koji sadrže podatkovni skup. Od ukupno 23 evidentirana podatkovna skupa samo je osam javno dostupno u trenutku provedbe ovog istraživanja. Za svaki od javno dostupnih podatkovnih skupova priložene su dodatne opisne informacije koje uključuju i kratki opis metode nastanka istog.

Podatkovni skup DocFigure [26] sadrži 28 vrsta grafikona s ukupno 30000 slika. Slike su nastale pretragom konferencijskih dokumenata i upotrebom raznih alata za izdvajanje slike iz elektroničkog dokumenta. Za organizaciju i označavanje koristi se neuronska mreža koja provjerava sličnost između novih slika i definiranih vrsta. Na kraju se očekuje ljudska provjera i potvrda uspješnosti.

Podatkovni skup ICDAR7 [55] je prvi podatkovni skup korišten za potrebu natjecanja i po završetku natjecanja učinjen je javno dostupnim. ICDAR7 je neorganiziran i neoznačen podatkovni skup koji sadrži i izvorne vrijednosti iz kojih su grafikoni nastali, no izvorne vrijednosti nisu jednostavno dostupne. Podatkovni skup je umjetno stvoren, a primjer sadržanih slika i dodatne informacije dostupne su u potpoglavlju 2.3.2.

Podatkovni skup ICPR2020 [56] je nastao kao rezultat drugog natjecanja iz područja obrade slike grafikona. Osim drugih vrsta grafikona, nema dodatne razlike u odnosu na ICDAR7 podatkovni skup.

Podatkovni skup ICPR2022 [140] je nastao kao rezultat trećeg natjecanja iz područja obrade slike grafikona. Na prva dva natjecanja fokus je bio na umjetno stvorenim slikama grafikonima. Rezultati oba natjecanja su potvrdili da metode koje ostvaruju vrlo visoku uspješnost na umjetno stvorenim podatkovnim skupovima, istu uspješnost ne ostvaruju na slikama stvarnih

grafikona. Zbog navedenog, prikupljene su slike grafikona iz raznih javno dostupnih objavljenih radova i napravljena je automatska organizacija i označavanje upotrebom neuronske mreže. Cijeli postupak je dodatno ručno provjeren i po potrebi su napravljene korekcije. Za izdvajanje podatkovnih značajki korišteni su razni programski alati koji su bili dodijeljeni određenim grupama zaduženim za ručnu obradu pojedine značajke.

Podatkovni skup FigureQA [135] sadrži preko milijun parova pitanja i odgovora za 100000 različitih slika grafikona. Pitanja su usmjerena sadržaju slike grafikona, a nastaju iz predefiniраниh predložaka. Podatkovni skup je umjetno stvoren pomoću parametara pa je kao takav organiziran, označen i dostupne su izvorne vrijednosti iz koji su slike grafikona nastale. U podatkovnom skupu nalazi se pet osnovnih vrsta grafikona koji sadrže jasno definirane tekstualne i grafičke elemente.

Podatkovni skup PlotQA [138] sadrži oko 30 milijun parova pitanja i odgovora za 225000 različitih slika grafikona organiziranih u tri vrste. Cilj izrade podatkovnog skupa je računalno razumijevanje strukture i informacije sadržane u grafikonu, te je zbog toga uz svaki grafikon priloženo i 70 povezanih pitanja nastalih iz predefiniраниh predložaka. Podatkovni skup je nastao upotrebom raznih internet tražilica i izdvajanjem tabličnih vrijednosti. Na osnovu prikupljenih tablica nastale su umjetno stvorene slike grafikona koje sadrže stvarne tekstualne oznake.

Podatkovni skup SciCap [139] je nastao zbog potrebe obrade naslova sadržanog u slici grafikona. Slike grafikona su preuzete iz raznih digitalnih arhiva i ručno su obrađene upotrebom posebno prilagođenih algoritama. Rezultat je podatkovni skup koji sadrži pet vrsta grafikona organiziranih prema broju riječi sadržanih u naslovu grafikona.

Iz priloženih podatkovnih skupova vidljivo je aktivno područje istraživanja slike grafikona, o čemu svjedoče i tri održana natjecanja (ICDAR2019, ICPR2020 i ICPR2022). Svi podatkovni skupovi su nastali sa zajedničkim ciljem, omogućavanje usporedbe rezultata i izrada polazišta za nova istraživanja. Opisani podatkovni skupovi su predodređeni i prilagođeni za točno određene namjene i kao takvi nisu opće primjenjivi. ICDAR7 podatkovni skup je jedini kojeg je moguće koristiti za različite namjene, no glavni nedostatak je potreba za ručnom organizacijom, a čitanje izvornih podataka zahtijeva posebno prilagođena programska rješenja. Zbog navedenih predstavljenih izazova, izrađen je podatkovni skup ChartDS [68] upotrebom parametarske metode koji sadrži 13 vrsta i 20 tipova grafikona, organiziranih, označenih, podijeljenih na jednostavne i složene oblike, i uz svaki grafikon priložena je datoteka koja sadrži JSON format s izvornim podatcima. Detalji korištene metode za izradu podatkovnog skupa dostupni su u potpoglavlju [4.3](#).

4.3 Metoda za generiranje ispitnih skupova

Iz predstavljenih rezultata provedenog istraživanja vidljiv je niz izazova koji je moguće riješiti izradom automatske metode koja će stvoriti traženi broj vrsta i tipova grafikona za pojedinačnu potrebu. Cilj metode je fleksibilnost i raznolikost u stvaranju slika grafikona. Pomoću predložene metode moguće je stvoriti neograničeno velik podatkovni skup koji sadrži čitljive i visoko-kvalitetne slike grafikona. Empirijskom analizom postojećih podatkovnih skupova uočena je potreba za visoko-kvalitetnim slikama grafikona i slikama grafikona na kojima nedostaje neki element ili je taj element djelomično vidljiv. Takvi djelomični grafikoni se najčešće prikupljaju upotrebom internet tražilica ili obradom literature koja nije akademskog sadržaja (primjerice internet stranice, portali ili novine). Zbog navedenog razloga u metodu su ugrađena dva skupa parametara, jednostavni koji stvara potpune grafikone i složeni koji stvara djelomične grafikone. Metoda zasnovana na parametrima nije ovisna o programskom jeziku.

Fleksibilnost programskog jezika u kontekstu generiranja podataka za slike grafikona predstavlja ključni element koji omogućuje široku primjenjivost metode u različitim okruženjima i među raznolikim korisnicima. Odluka da metoda nije ovisna o programskom jeziku pridonosi njezinoj općenitoj primjenjivosti. Ova fleksibilnost čini implementaciju prilagodljivom i dostupnom raznim korisnicima i istraživačima. Osim navedenog, otvorenost prema različitim programskim jezicima potiče suradnju i razmjenu među istraživačima koji koriste različite tehnološke cjeline. Kroz predstavljenu fleksibilnost, metoda postaje pristupačna i korisnicima izvan područja računalnih znanosti koji možda posjeduju više iskustva u statističkoj analizi ili matematičkom modeliranju nego u određenom programskom jeziku. Prvobitna implementacija izrađena je u Pythonu s korištenjem Plotly biblioteke i dodatno olakšava manipulaciju vizualnim elementima, pružajući mogućnost dinamične promjene parametara i istraživanje različitih konfiguracija.

Parametri metode za stvaranje slike grafikona dostupni su u prilogu [P.8](#), a mogu se tumačiti na sljedeći način. Prvo, potrebno je definirati veličinu slike na kojoj će se iscrtati grafikon. Ovisno o namjeni podatkovnog skupa, preporuka je stvoriti sliku visoke razlučivosti (primjerice razlučivosti 1366×768 točke). Drugo, potrebno je odrediti vrstu grafikona na kojoj će se elementi stvarati. Zatim, pokreće se nekoliko slučajnih generatora koji mogu biti generatori brojeva, oznaka ili riječi. Generatori brojeva, oznaka i riječi imaju ključnu ulogu u određivanju podatkovnih segmenata i strukture grafikona. Korištenje različitih generatora omogućuje kontroliranje količine informacija prisutne na grafikonu što utječe na složenost slike grafikona. Generator brojeva

određuje podatkovni segment, podatke koji će biti prikazani na grafikonu, ali i količinu podataka koja će biti prikazana. Kod jednostavnih oblika, količina informacija je smanjena odnosno ako se promatra linijski grafikon, isti će sadržavati samo jednu liniju, a kod složenih oblika nalaziti će se veća količina informacija, odnosno grafikon može sadržavati više linija. Generator oznaka određuje vrijednosti na koordinatnim osima. Za vrijednosti na osima koriste se mjerne jedinice iz Međunarodnog sustava mjernih jedinica. Generator riječi dohvaća riječi iz željenog rječnika. Dohvaća se veći broj riječi koje se spajaju u jednu rečenicu i postupak se ponavlja za svaki tekstualni element.

Empirijska analiza postojećih podatkovnih skupova pruža uvid u karakteristike slika grafikona i omogućuje prilagodbu parametara metode prema željenim podacima. Analiza se fokusira na identifikaciju karakterističnih značajki naslova, legende, koordinatnih osi i drugih elemenata prisutnih na slikama grafikona. Razumijevanje učestalosti i oblika ovih elemenata pomaže u prilagodbi generiranih podataka. Navedeni pristup osigurava da generirani podatci odražavaju stvarne uvjete i izazove, čime se poboljšava relevantnost rezultata. Provedeno istraživanje otkriva specifičnosti poput ograničenja u duljini naslova, broju riječi u legendi, duljini naslova koordinatnih osi i drugih ograničenja koja su često prisutna na slikama grafikona.

Iz rezultata analize podatkovnih skupa i slika grafikona prikupljenim internetskim pretraživanjem vidljivo je da naslov grafikona rijetko sadrži više od 15 riječi, naslov legende više od 10 riječi, naslov koordinatnih osi više od četiri riječi, a legenda može sadržavati numeričke i tekstualne oznake, ali je zbog preglednosti ograničena na tri riječi ukoliko nije predstavljena brojevima. Kod jednostavnih oblika preklapanje elemenata nije dozvoljeno te ako je naslov pozicioniran u gornjem lijevom kutu, u tom kutu se ne može pojaviti legenda, odnosno legenda se premješta u gornji desni kut. Kod složenih oblika, moguće je preklapanje elemenata i postoji mogućnost pojave više različitih elemenata unutar jednog kuta. Font je definiran na razini pojedinog grafikona, ali moguće ga je definirati na razini pojedinog elementa. Veličina fonta prati najčešći odnos veličina u grafikonu gdje je naslov najveći font, a vrijednosti na koordinatnim osima najmanji font. Podatkovni segment je još definiran i bojom, oznakama, ključnim točkama i ispisanim vrijednostima, gdje su gotovo svi navedeni parametri isključeni kod jednostavnih oblika, a kod složenih oblika moguće su raznolike konfiguracije. Na sam izgled grafikona moguće je utjecati i višestrukim pozadinama, mrežnim strukturama i različitom debljinom linija i granica pojedinih elemenata.

Po završetku stvaranja slike grafikona, svi korišteni parametri se pohranjuju u pripadajuću datoteku koja sadrži JSON format i može se koristiti kao sredstvo provjere napisanih metoda za

Tablica 4.2: Statistika podatkovnog skupa ChartDS.

Vrsta grafikona	Tip grafikona	Broj slika grafikona			
		Za učenje		Za testiranje	
		Jednostavni	Složeni	Jednostavni	Složeni
Prostorni		2000	2000	1000	1000
Stupčasti	Osnovni horizontalni	2000	2000	1000	1000
	Osnovni vertikalni	2000	2000	1000	1000
	Grupirani horizontalni	2000	2000	1000	1000
	Grupirani vertikalni	2000	2000	1000	1000
	Naslagani horizontalni	2000	2000	1000	1000
	Naslagani vertikalni	2000	2000	1000	1000
Kutije	Horizontalni	2000	2000	1000	1000
	Vertikalni	2000	2000	1000	1000
Mjehurićasti	-	2000	2000	1000	1000
Prstenasti	-	2000	2000	1000	1000
Toplinska karta		2000	2000	1000	1000
Učestalost	Horizontalni	2000	2000	1000	1000
	Vertikalni	2000	2000	1000	1000
Linijski		2000	2000	1000	1000
Tortni		2000	2000	1000	1000
Točkasti		2000	2000	1000	1000
Sunčeve zrake		2000	2000	1000	1000
Tablica		2000	2000	1000	1000
Napredak		2000	2000	1000	1000
	Ukupno	40000	40000	20000	20000

obradu slike grafikona. Osim navedenog, JSON format podataka predstavlja ključan element u integraciji informacijskih sustava putem API (engl. *Application Programming Interface*) sučelja. Svojom pristupačnom i jednostavno čitljivom sintaksom, JSON se može koristiti za razmjenu strukturiranih podataka između raznih računalnih sustava. U kontekstu aplikativnih rješenja zasnovanih na API sučelju, JSON se primjenjuje kao standardni format odgovora na zahtjeve.

Podatci se oblikuju u JSON formatu prilikom prijenosa između klijenta i poslužitelja, pojednostavljajući interpretaciju informacija na obje strane.

Upotrebom predstavljenih parametara nastao je ujednačen, organiziran i označen podatkovni skup ChartDS koji je javno dostupan u podatkovnim repozitorijima Figshare [67] i Srce Dabar [69]. Podatkovni skup sadrži 13 vrsta grafikona, 20 tipova grafikona i podijeljen je na podatkovni skup za učenje i podatkovni skup za testiranje. Oba podatkovna skupa sadrže podjelu na jednostavne i složene oblike. U podatkovnom skupu za učenje nalazi se 2000 slika po tipu grafikona, a u podatkovnom skupu za testiranje nalazi se 1000 slika po tipu grafikona. Podatkovni skup sadrži ukupno 120000 slika grafikona. Detaljni sadržaj podatkovnog skupa nastalog pomoću predložene metode dostupan je u tablici 4.2.

Podatkovni skup ChartDS je namijenjen testiranju i razumijevanju različitih metoda i algoritama koji se bave klasifikacijom ili interpretacijom slike grafikona. Predstavljene metode u sklopu ovog istraživanja evaluirane su na predstavljenom podatkovnom skupu.

4.4 Usporedba podatkovnih skupova ChartDS i ICDAR7

Usporedba podatkovnih skupova je često adresiran izazov koji se javlja kada se provodi testiranje modela strojnog učenja na više podatkovnih skupova nastalih iz različitih izvora. Kvantitativna usporedba se odnosi na organizacijsku strukturu, broj elemenata u skupu i na opisne vrijednosti pojedinog elementa. Kvantitativnom usporedbom ChartDS je organiziran i označen podatkovni skup koji sadrži nešto manje slika u odnosu na ICDAR7 podatkovni skup. Također, ChartDS sadrži znatno veći broj vrsta i tipova grafikona te je uz svaki grafikon priložena i datoteka s izvornim podacima. Oba podatkovna skupa sadrže raznolike slike visoke razlučivosti izrađene upotrebom jednog programskog alata. S druge strane, za kvalitativnu usporedbu, najučinkovitija je ocjena promatrača koja je podložna percepciji i razmišljanju pojedinca. Od automatskih metoda izdvajaju se:

- Upotreba srednje kvadratne pogreške (engl. *Mean Squared Error*) koja radi na razini pojedine točke razlučivosti i služi za izračunavanje srednje vrijednosti kvadrata razlika između točaka razlučivosti dviju slika, kao što je prikazano u [141,142],
- Upotreba metoda za detekciju objekata i značajki na slici, kao što je prikazano u [143] i

- Upotreba sijamskih neuronskih mreža jer sve sijamske neuronske mreže rade na principu traženja koeficijenta sličnosti između dva ulazna podataka.

Svaka od navedenih metoda je podložna greškama, a usporedba je vremenski i računalno zahtjevan zadatak jer se istovremeno uspoređuju samo dvije slike. Za usporedbu ChartDS i ICDAR7 podatkovnih skupova koristi se predloženi algoritam napisan pseudokodom i prikazan slikom 4.1.

```

FUNKCIJA stvoriPrazniDirektorij(putanja):
  AKO putanja postoji ONDA
    ISPIŠI "Direktorij već postoji."
  INAČE
    STVORI putanja
    ISPIŠI "Direktorij uspješno stvoren."

FUNKCIJA kopirajDatoteke(izvor, odredište):
  ZA SVAKU datoteku U izvor:
    AKO datoteka je direktorij ONDA
      novilzvor = odredište + "/" + datoteka
      kopirajDatoteke(novilzvor, odredište)
    INAČE
      KOPIRAJ datoteku U odredište
      ISPIŠI "Datoteke uspješno kopirane."

FUNKCIJA pokreniUčenjeNeuronskeMreže(izvor):
  # SCNN
  SPREMI modelUsporedba
  ISPIŠI "Učenje uspješno završeno."

FUNKCIJA provjeriTočnostModela(model, podatci):
  # Testiranje SCNN
  ISPIŠI "Rezultati testiranja."

# Postavljanje varijabli
izvorniDirektorij = "put/do/direktorija"
direktorijChartDS = "put/do/ChartDS"
direktorijICDAR7 = "put/do/ICDAR7"
direktorijDS1 = izvorniDirektorij + "/" + "DS1"
direktorijDS2 = izvorniDirektorij + "/" + "DS2"
direktorijTest = "put/do/test"

# Stvori direktorije DS1 i DS2 ako ne postoje
stvoriPrazniDirektorij(direktorijDS1)
stvoriPrazniDirektorij(direktorijDS2)
stvoriPrazniDirektorij(direktorijTest)

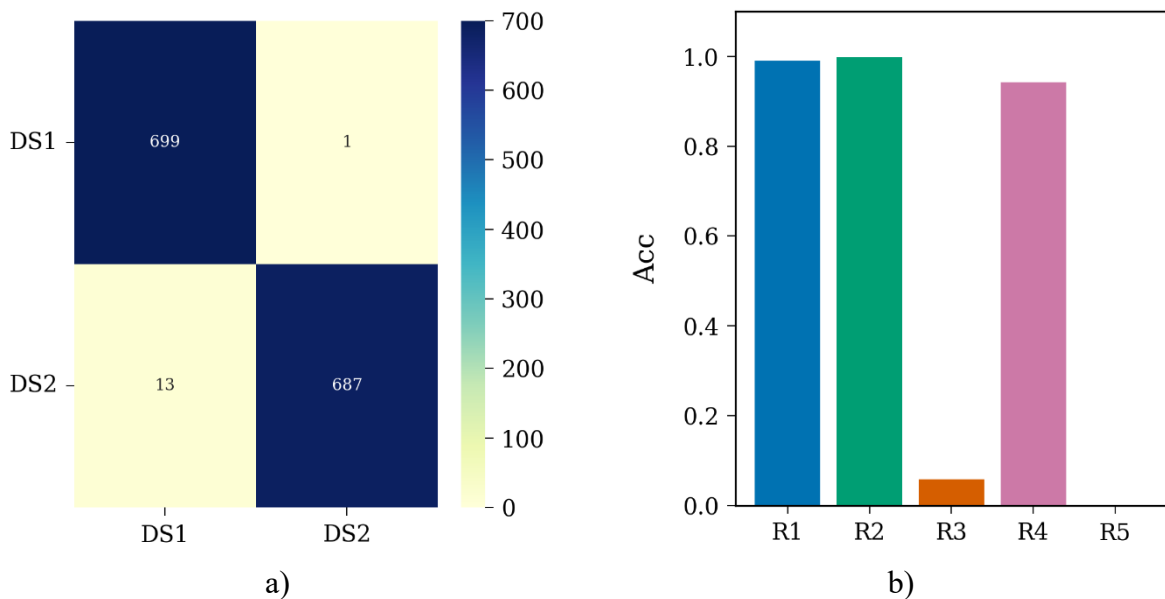
# Kopiraj sadržaje ChartDS i ICDAR7 u DS1 i DS2
kopirajDatoteke(direktorijChartDS, direktorijDS1)
kopirajDatoteke(direktorijICDAR7, direktorijDS2)
kopirajDatoteke("put/do/testChartDS", direktorijTest)
kopirajDatoteke("put/do/testICDAR7", direktorijTest)

# Pokreni učenje SCNN-a na izvornom direktoriju koji sadrži klase DS1 i DS2
pokreniUčenjeNeuronskeMreže(izvorniDirektorij)

# Pokreni testiranje modela
provjeriTočnostModela("put/do/modelUsporedba", direktorijTest)

```

Slika 4.1: Pseudokod algoritma za usporedbu podatkovnih skupova ChartDS i ICDAR7



Slika 4.2: Prikaz točnosti klasifikacije za model učen i testiran na slikama grafikona iz ChartDS (DS1) i ICDAR7 (DS2) podatkovnih skupova. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune.

Osnovna ideja algoritma je stvoriti dva direktorija DS1 i DS2. Direktorij DS1 sadrži sve slike koje su sadržane u ChartDS podatkovnom skupu ali bez dodatne organizacije na poddirektorije, gdje svaki poddirektorij sadrži jednu klasu grafikona. Na isti način je stvoren i DS2 direktorij koji sadrži sve slike koje su sadržane u ICDAR7 podatkovnom skupu. Tako stvorena dva direktorija predstavljaju dvije klase na kojima će biti učen SCNN model. Korišteni SCNN model je isti kao i za pojedinačnu provjeru navedena dva podatkovna skupa, odnosno samo je zadnji sloj prilagođen za klasifikaciju u dvije klase. Podatkovni skup za testiranje sadrži sve slike sadržane u ChartDS i ICDAR7 testnim podatkovnim skupovima. Rezultati testiranja prikazani su slikom 4.2.

Ukupna ostvarena točnost klasifikacije za 1400 slika (700 slika po podatkovnom skupu) iznosi 0.99. Pri tome, statistička analiza potvrđuje vrlo visoku vjerojatnost u donošenju odluka. Ručnom provjerom 14 netočno ostvarenih predikcija vidljiva je netočna klasifikacija kružnih grafikona jer isti posjeduju najmanju varijaciju oblika nakon procesa predobrade slike. Iako oba podatkovna skupa sadrže iste vrste grafikona, iz provedene empirijske analize vidljivo je da su slike u navedenim podatkovnim skupovima sadržajno drugačije. SCNN model sa visokom točnošću klasifikacije određuje pripadnost slika iz testnog podatkovnog skupa. Kada se SCNN model uči na jednom podatkovnom skupu, a zatim testira na drugom, primjećuje se smanjenje točnosti klasifikacije. SCNN model većinski klasificira slike drugog podatkovnog skupa u klasu koja ne sadrži slike grafikona (H). Navedeno potvrđuje postojanje razlika između slika iz različitih

podatkovnih skupova. Također, provedeno testiranje ukazuje na činjenicu da neuronska mreža uči promatrajući objekte prisutne na slikama, što u slučaju ICDAR7 podatkovnog skupa često znači veća količina objekata, kao što je prikazano slikom 2.4 nego kada se promatra ChartDS podatkovni skup.

Navedeni postupak usporedbe dokazuje heterogenost u strukturi slike grafikona. Slike grafikona stvorene upotrebom istog programskog alata ne moraju nužno sadržavati iste elemente. Upotrebom samo jednog programskog alata moguće je stvoriti neograničen broj izgledom raznolikih grafikona. Metode za obradu slike grafikona trebaju posjedovati zalihost i robusnost u algoritmu te trebaju biti spremne na nove vrste i tipove grafikona što stavlja dodatni naglasak na važnost metode koja može generirati neograničen broj slika grafikona. Ako bi neuronsku mrežu učili na oba podatkovna skupa, očekivalo bi se postizanje jednake točnosti klasifikacije na oba testna podatkovna skupa. Dodatno, ako bi se neuronska mreža učila na dodatnim podatkovnim skupovima poput ReVision-a i drugih skupova navedenih u tablici 4.1, očekivalo bi se stvaranje modela koji je sposoban klasificirati slike grafikona neovisno o njihovoj strukturi ili alatu koji je korišten za njihovu izradu.

4.5 Zaključak

Iz provedenog istraživanja vidljiva je promjena dostupnosti podatkovnih skupova. Objavljeni podatkovni skupovi najčešće se nalaze u sklopu konferencija, natjecanja ili privatnih istraživanja te su pohranjeni na raznim nestandardiziranim lokacijama čija dostupnost se mijenja. Objavljeni radovi prije jednu, dvije ili pet godina, koriste podatkovne skupove koji danas ne postoje, nisu više potpuni ili im je pristup ograničen. Autori za vlastite istraživačke ciljeve izrađuju podatkovne skupove upotrebom različitih metoda koje mogu uključivati potpuno ručni pristup prikupljanja, označavanja i organiziranja podatkovnog skupa, polu-automatski pristup gdje je jedan dio procesa automatiziran ili koriste parametarski pristup koji je automatiziran i kojemu nije potrebna ljudska interakcija ili provjera. Nažalost, tako nastali podatkovni skupovi su najčešće ograničeni jer su namijenjeni točno određenom zadatku koji može biti primjerice klasifikacija grafikona, obrada tekstualnog sadržaja grafikona, obrada grafičkih elemenata grafikona i drugo. Osim navedenog, sadrže ograničen broj vrsta, tipova i pripadajućih slika grafikona, koji je nepromjenjiv. Dodatni izazov nastaje kada je potrebno napraviti usporedbu različitih metoda i algoritama, a ne postoji jedinstveno polazište.

Sukladno predstavljanim izazovima izrađena je metoda koja nije ovisna o programskom jeziku i omogućava izradu podatkovnih skupova s neograničenim brojem slika grafikona. Jedinstven ispitni podatkovni skup ima ključnu ulogu u znanstvenim istraživanjima, posebno u kontekstu razvoja i evaluacije metoda, modela ili algoritama. Njegova važnost proizlazi iz višestrukih aspekata koji pridonose integraciji i unaprjeđenju raznih znanstvenih istraživanja. Postojanje jedinstvenog ispitnog skupa ključno je za osiguravanje pouzdanosti rezultata istraživanja. Reprodukcijska istraživanja od strane drugih znanstvenika postaje moguća, omogućujući potvrdu dobivenih rezultata i poticanje znanstvene reproduktivnosti. Jedinstveni ispitni podatkovni skup doprinosi generalizaciji modela i metoda. Raznolikost podataka omogućuje proučavanje stabilnosti i robusnosti algoritama ili modela, posebno identificirajući izazove koji se javljaju unutar specifičnih domenskih zahtjeva. Transparentnost i javna dostupnost jedinstvenog ispitnog podatkovnog skupa omogućava drugima testiranje i usporedbu novih pristupa s postojećima, podatci postaju temelj za kritičko vrednovanje i usklađivanje s visokim standardima u znanstvenom istraživanju. Dodatno, potiču se inovacije u području vizualizacije podataka. Raznolikost podataka motivira istraživače u razvijanju novih metoda prilagođenih različitim izazovima.

U konačnici, jedinstveni ispitni podatkovni skup predstavlja ključnu komponentu za istraživanje i razvoj u području vizualizacije podataka. Njegova raznolikost i dostupnost potiču znanstvenu zajednicu na razumijevanje, prilagodbu i unaprjeđenje metoda, čime se osigurava relevantnost, pouzdanost i primjenjivost znanstvenih istraživanja u širem kontekstu vizualizacije podataka.

5 ZAKLJUČAK

Cilj ove disertacije bio je razvoj automatiziranih metoda strojnog učenja za obradu digitalne slike grafikona što bi implementacijom u stvarnoj produkcijskoj okolini omogućilo kompletan proces od učenja i prepoznavanja vrsta grafikona do izdvajanja podatkovnih značajki za primjenu u brojnim računalnim sustavima. Takav kompletan sustav bi nakon lokaliziranja i izdvajanja vizualnog sadržaja iz elektroničkog dokumenta prvo izvršio klasifikaciju svakog vizualnog elementa zasebno. Klasifikacija omogućava razlikovanje grafikona od ostalog vizualnog sadržaja koji se izdvaja u zasebnu klasu. Ako je slika klasificirana kao grafikon, pristupa se određivanju vrste grafikona, a potom, ako je klasificirana određena vrsta grafikona, pristupa se izdvajanju podatkovnih značajki. Podatkovne značajke uključuju izvorne podatke iz kojih je promatrani grafikon nastao. Podatkovne značajke pohranjuju se u računalno pristupačnom obliku kojeg mogu konzumirati brojni računalni sustavi, a sama organizacija je razumljiva i ljudima. Takav kompletan sustav od iznimne je važnosti internet tražilicama, pretraživačima slikovnog sadržaja, znanstvenicima, istraživačima, ali i osobama s oštećenim vidom koji pristup informacijama ostvaruju upotrebom čitača ekrana. Ovakav sustav, računalno i cjenovno je optimiziran te ne zahtijeva nikakva posebna ulaganja u računalne resurse već bi ga bilo moguće pokrenuti na vlastitom osobnom računalu.

U sklopu ovog istraživanja predstavljene su tri metode koje u postupku evaluacije koriste podatkovne skupove označene kao: ReVision, ICDAR7, ChartDS, Linnaeus, CIFAR10 i AT&T DoF. Ostvareni rezultati metoda su prikazani usporedno s drugim rezultatima istraživanja u ovom području.

Prvi doprinos, metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama, sadrži općenito prvu primjenu SCNN-a u području istraživanja koje se bavi klasifikacijom rasterske slike grafikona. Osim prve primjene SCNN-a, predstavljena je i prva primjena SSCNN-a koja je izrađena upotrebom dvije SCNN arhitekture. SCNN je dubinski analizirana čime je dokazana djelotvornost u konstrukciji mrežnih slojeva i optimiziranost cjelokupne neuronske mreže pri čemu sadrži samo pet težinskih slojeva. Vremenska i prostorna složenost smanjene su nekoliko desetaka puta u odnosu na arhitekture kao što su: VGG16, Inception v3 ili Xception. SCNN je predodređena za klasifikaciju grafikona i kao takva ostvaruje približno maksimalnu točnost klasifikacije na sedam klasa grafikona čiji rezultati su evaluirani na nekoliko stotina slika grafikona. SSCNN je evaluirana na identičnom podatkovnom skupu kao i SCNN te je ostvarena maksimalna točnost klasifikacije. U slučaju otvorene klasifikacije, odnosno kada se na ulazu modela ne nalazi slika grafikona, predstavljena su rješenja i za SCNN i za SSCNN model. U evaluaciji otvorene klasifikacije, prednost ostvaruje SSCNN model koji kada je učen s podatkovnim skupom od 50 ili više slika po klasi grafikona, ostvaruje maksimalnu točnost klasifikacije, odnosno raspoznavanja što pripada klasama grafikona, a što ne. Tijekom istraživanja i evaluacije razvijena je i metoda za predobradu slike grafikona koja s ulazne rasterske slike uklanja značajan broj informacija, ostavljajući crno-bijelu sliku koja sadrži osnovne grafičke elemente pomoću kojih je moguće razlikovati klase grafikona. Upotrebom predstavljene metode nastao je organizirani i označeni podatkovni skup koji sadrži sve slike korištene u procesu učenja i testiranja modela. Podatkovni skup je javno dostupan što omogućava ponovljivost i usporedivost predstavljenih rezultata.

Drugi doprinos, metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima, sadrži prvu poznatu obradu kružnih i mrežastih grafikona. Znanstvenici u području istraživanja uglavnom se bave izdvajanjem podatkovnih značajki iz linijskih, stupčastih i jednostavnih tortnih grafikona, te je metodom predstavljeno prvo izdvajanje podatkovnih značajki iz složenih tipova grafikona kao što su grafikoni sunčevih zraka i toplinske karte. Metoda se sastoji od dva modula: modul za predobradu slike i modul za detekciju raznih strukturnih elemenata grafikona. Metoda je nastala na osnovu postojećeg iskustva istraživanja unutar ovog područja i koristi pristup obrade slike niske razine. Rad metode zasnovan je na binarnim maskama, morfologiji, detekciji ključnih točaka i matematičkom modelu. Modulu za predobradu slike uključuju prepoznavanje tekstualnih elemenata i uklanjanje istih sa slike. S obzirom da je OCR mehanizam zasebno područje istraživanja, metoda ne razmatra kvalitetu detekcije tekstualnih elemenata već podrazumijeva idealnu situaciju kada OCR mehanizam radi

bez ograničenja. Detektor modul uključuje obradu segmenata, detekciju linija, kružnica i ključnih točaka. Sve detektirane ključne točke prosljeđuju se u matematički model gdje se izračunavaju izvorni podatci korišteni za izradu slike grafikona. Metoda je evaluirana na ChartDS i ICDAR7 podatkovnim skupovima, gdje prosječna točnost izdvajanja podatkovnih značajki iznosi 0.98. Metoda je evaluirana i na slikama prikupljenim internetskim pretraživanjem, gdje se zbog heterogenosti u strukturi grafikona prosječna točnost izdvajanja podatkovnih značajki značajno smanjuje i iznosi približno 0.5. Metoda je u potpunosti automatizirana i isporučuje datoteku koja sadrži JSON format. Navedenu datoteku je moguće implementirati s brojnim računalnim sustavima koji bi mogli omogućiti dodatne usluge na temelju prikupljenih podataka.

Posljednji doprinos, metoda generiranja ispitnih skupova za standardizirano vrednovanje uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona, predstavlja jedinstveno ishodište za autore i istraživače u području interpretacije slike grafikona. Metoda je nastala zbog nedostatka javno dostupnih podatkovnih skupova koji su organizirani, označeni, raspodijeljeni na podatkovne skupove za učenje i testiranje i sadrže izvorne podatke iz kojih je slika grafikona nastala. Metoda prati opće smjernice nastale na osnovu postojećeg iskustva istraživanja unutar ovog područja i omogućava stvaranje jednostavnih i složenih oblika grafikona. Jednostavni oblici podrazumijevaju grafikone koji sadrže najčešće jedan podatkovni skup, a svi elementi na grafikonu su jasno vidljivi i označeni. Suprotno, složeni oblici podrazumijevaju grafikone koji mogu sadržavati nekoliko različitih podatkovnih skupova, a elementi na grafikonu ne moraju biti prisutni ili je moguće preklapanje istih. Upotrebom navedene metode nastao je podatkovni skup ChartDS koji je iskorišten u svim predstavljenim metodama i evaluacijskim postupcima. ChartDS je javno dostupan podatkovni skup što uspostavlja zajedničku osnovu za evaluaciju, nudeći standardiziranu ishodište za procjenu i usporedbu učinkovitosti različitih sustava u razumijevanju i interpretaciji slika grafikona.

Nastavak ovog istraživanja moguć je u nekoliko smjerova koji uključuju proširenje klasifikacijske metode na znatno veći broj klasa grafikona, poopćenje i proširenje metode za izdvajanje podatkovnih značajki na veći broj vrsta grafikona, izrada metode za generiranje tekstualnog opisa ili sadržajne informacije grafikona i obrada tekstualne informacije sadržane u slici grafikona. Osim toga, većina istraživanja u ovom području ograničena je na korištenje umjetno stvorenih slika grafikona, koje su pažljivo pripremljene i često optimizirane kako bi se postigla maksimalna točnost metoda. Međutim, kada se iste metode primjene na stvarne slike grafikona prikupljene putem internetskog pretraživanja, postižu se znatno slabiji rezultati. Navedeno ukazuje na potrebu daljnjeg istraživanja i razvoja metoda koje su otpornije na

raznolikost stvarnih slika grafikona i koje mogu učinkovito raditi na slikama gdje je prisutan raznoliki šum i nepravilnosti koje nisu prisutne u umjetno stvorenim slikama grafikona.

LITERATURA

- [1] C. Chen, W. Härdle, A. Unwin, M. Friendly, A Brief History of Data Visualization, in: Handbook of Data Visualization, Springer Berlin Heidelberg, 2008: pp. 15–56
- [2] I. Spence, William Playfair and the Psychology of Graphs, JSM - Proceedings of the American Statistical Association, Section on Statistical Graphics (2006) 2426–2436
- [3] D. Chester, S. Elzer, Getting Computers to See Information Graphics So Users Do Not Have to, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 3488 LNAI (2005) 660–668
- [4] grafikon | Hrvatska enciklopedija, (n.d). <https://www.enciklopedija.hr/Natuknica.aspx?ID=23039>
- [5] F. Bajić, J. Job, K. Nenadić, Data Visualization Classification Using Simple Convolutional Neural Network Model, International Journal of Electrical and Computer Engineering Systems 11 (2020) 43–51
- [6] J. Poco, J. Heer, Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images, Computer Graphics Forum 36 (2017) 353–363
- [7] F. Bajić, J. Job, Review of chart image detection and classification, International Journal on Document Analysis and Recognition 26 (2023) 453–474
- [8] F. Bajić, J. Job, K. Nenadić, Chart Classification Using Simplified VGG Model, International Conference on Systems, Signals, and Image Processing 2019-June (2019) 229–233
- [9] Y. Liu, X. Lu, Y. Qin, Z. Tang, J. Xu, Y. Liu, X. Lu, Y. Qin, Z. Tang, J. Xu, Review of chart recognition in document images, SPIE 8654 (2013)
- [10] K. Davila, S. Setlur, D. Doermann, B.U. Kota, V. Govindaraju, Chart Mining: A Survey of Methods for Automated Chart Analysis, Transactions on Pattern Analysis and Machine Intelligence (2020)

- [11] K.C. Shahira, A. Lijiya, Towards Assisting the Visually Impaired: A Review on Techniques for Decoding the Visual Data from Chart Images, *IEEE Access* 9 (2021) 52926–52943
- [12] F. Bajić, J. Job, Chart Classification Using Siamese CNN, *J. Imaging* 7 (2021) 220–238.
- [13] F. Bajić, O. Orel, M. Habijan, A Multi-Purpose Shallow Convolutional Neural Network for Chart Images, *Sensors* 2022, Vol. 22, Page 7695 22 (2022) 7695
- [14] F. Bajić, M. Habijan, K. Nenadić, Evaluation of Shallow Convolutional Neural Network in Open-World Chart Image Classification, *Informatika* 48 (2024) 185–198
- [15] H. Weihua, Scientific chart image recognition and interpretation, National university of Singapore, 2008.
- [16] F. Hofer, J. Juraj, Klasificiranje grafičkih prikaza korištenjem plitkih neuronskih mreža, *Fakultet elektrotehnike, računarstva i informacijskih tehnologija*, 2022
- [17] R. Hecht-Nielsen, *Neurocomputing: The Technology Of Non-Algorithmic Information Processing*, (1989)
- [18] B. Bašić Dalbelo, M. Čupić, J. Šnajder, *Umjetne neuronske mreže*, Zagreb, 2008
- [19] J. Amara, P. Kaur, M. Owonibi, B. Bouaziz, Convolutional neural network based chart image classification, *25th International Conference in Central Europe on Computer Graphics* (2017).
- [20] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, *Artificial Intelligence Review* 2020 53:8 53 (2020) 5455–5516
- [21] N. Siegel, Z. Horvitz, R. Levin, S. Divvala, A. Farhadi, FigureSeer: Parsing Result-Figures in Research Papers, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9911 LNCS (2016) 664–680
- [22] W. Dai, M. Wang, Z. Niu, J. Zhang, Chart decoder: Generating textual and numeric information from chart images automatically, *J Vis Lang Comput* 48 (2018) 101–109
- [23] A. Balaji, T. Ramanathan, V. Sonathi, Chart-Text: A Fully Automated Chart Image Descriptor, *ArXiv Preprint* (2018)
- [24] A.Y. Lin, J. Ford, E. Adar, B. Hecht, VizByWiki: Mining data visualizations from the web to enrich news articles, *The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018* (2018) 873–882
- [25] J. Choi, S. Jung, D.G. Park, J. Choo, N. Elmqvist, Visualizing for the non-visual: Enabling the visually impaired to use visualization, *Computer Graphics Forum* 38 (2019) 249–260.
- [26] K. V. Jobin, A. Mondal, C. V. Jawahar, DocFigure: A Dataset for Scientific Document Figure Classification, *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)* (2019) 74–79.
- [27] X. Liu, D. Klabjan, P. NBless, Data Extraction from Charts via Single Deep Neural Network, *ArXiv* (2019). <https://arxiv.org/abs/1906.11906v1>

- [28] D. Jung, W. Kim, H. Song, J.I. Hwang, B. Lee, B. Kim, J. Seo, ChartSense: Interactive data extraction from chart images, Conference on Human Factors in Computing Systems - Proceedings 2017
- [29] P. Chagas, R. Akiyama, A. Meiguins, C. Santos, F. Saraiva, B. Meiguins, J. Morais, Evaluation of Convolutional Neural Network Architectures for Chart Image Classification, Proceedings of the International Joint Conference on Neural Networks 2018
- [30] K.C. Shahira, A. Lijiya, Document Image Classification: Towards Assisting Visually Impaired, IEEE Region 10 Annual International Conference, Proceedings/TENCON 2019-October (2019) 852–857.
- [31] P. Kaur, D. Kiesel, Combining image and caption analysis for classifying charts in biodiversity texts, VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications 3 (2020) 157–168
- [32] X. Liu, B. Tang, Z. Wang, X. Xu, S. Pu, D. Tao, M. Song, Chart classification by combining deep convolutional networks and deep belief networks, Proceedings of the International Conference on Document Analysis and Recognition, ICDAR 2015-November (2015) 801–805
- [33] P. Chagas, A. Freitas, R. Daisuke, B. Miranda, T. Araújo, C. Santos, B. Meiguins, J. Morais, Architecture proposal for data extraction of chart images using convolutional neural network, Proceedings - 2017 21st International Conference Information Visualisation, IV 2017 (2017) 318–323
- [34] I. Kavasidis, S. Palazzo, C. Spampinato, C. Pino, D. Giordano, D. Giuffrida, P. Messina, A Saliency-based Convolutional Neural Network for Table and Chart Detection in Digitized Documents, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11752 LNCS (2018) 292–302.
- [35] Altan Gokhan, DeepGraphNet: grafiklerin sınıflandırılmasında derin öğrenme modelleri, Avrupa Bilim ve Teknoloji Dergisi (2019) 319–329.
- [36] S. Huang, An Image Classification Tool of Wikimedia Commons, Humboldt-Universität zu Berlin, 2020.
- [37] C. Kosemen, D. Birant, Multi-label classification of line chart images using convolutional neural networks, SN Appl Sci 2 (2020) 1–20
- [38] T. Ishihara, K. Morita, N.C. Shirai, T. Wakabayashi, W. Ohyama, Chart-Type Classification Using Convolutional Neural Network for Scholarly Figures, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12047 LNCS (2020) 252–261.
- [39] R.A. Al-Zaidy, C.L. Giles, A Machine Learning Approach for Semantic Structuring of Scientific Charts in Scholarly Documents, Twenty-Ninth IAAI Conference (2017).
- [40] P. Vougiouklis, L. Carr, E. Simperl, Pie Chart or Pizza: Identifying Chart Types and Their Virality on Twitter, Proceedings of the International AAAI Conference on Web and Social Media 14 (2020) 694–704

- [41] T. Araújo, P. Chagas, J. Alves, C. Santos, B.S. Santos, B.S. Meiguins, A Real-World Approach on the Problem of Chart Recognition Using Classification, Detection and Perspective Correction, *Sensors* 2020 20 (2020)
- [42] K. Dadhich, S. Daggubati, J. Sreevalsan-Nair, BarChartAnalyzer: Digitizing Images of Bar Charts, *IMPROVE* (2021) 17–28.
- [43] W. Ma, H. Zhang, S. Yan, G. Yao, Y. Huang, H. Li, Y. Wu, L. Jin, Towards an efficient framework for Data Extraction from Chart Images, *ArXiv Preprint* (2021).
- [44] J. Thiyam, S.R. Singh, P.K. Bora, Challenges in chart image classification, *Proceedings of the 21st ACM Symposium on Document Engineering* (2021) 1–4.
- [45] C. Rane, S. Subramanya, D. Endluri, J. Wu, C. Giles, ChartReader: Automatic Parsing of Bar-Plots, *IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)* 1 (2021) 318–325.
- [46] I. Redeke, Image & Graphic Reader., in: *IEEE International Conference on Image Processing*, IEEE, Thessaloniki, Greece, 2001: pp. 806–809.
- [47] M. Savva, N. Kong, A. Chhajta, F.F. Li, M. Agrawala, J. Heer, ReVision: Automated classification, analysis and redesign of chart images, *UIST’11 - Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (2011) 393–402.
- [48] J. Gao, Y. Zhou, K.E. Barner, View: Visual Information Extraction Widget for improving chart images accessibility, *Proceedings - International Conference on Image Processing, ICIP* (2012) 2865–2868.
- [49] A. Mishchenko, N. Vassilieva, Model-based chart image classification, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6939 LNCS (2011) 476–485
- [50] A. Mishchenko, V. Natalia, Model-Based Recognition and Extraction of Information from Chart Images., *J. Multim. Process. Technol.* 2 (2011) 76–89
- [51] L.; Auria, R.A. Moro, Support Vector Machines (SVM) as a technique for solvency analysis, *DIW Berlin Discussion Paper No. 811* (2008)
- [52] R.R. Nair, N. Sankaran, I. Nwogu, V. Govindaraju, Automated analysis of line plots in documents, *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR 2015-November* (2015) 796–800.
- [53] Y. Shi, Y. Wei, T. Wu, Q. Liu, Statistical graph classification in intelligent mathematics problem solving system for high school student, *ICCSE 2017 - 12th International Conference on Computer Science and Education* (2017) 645–650
- [54] J. Thiyam, S.R. Singh, P. Kumar, P.K. Bora, Effect of Attention and Triplet loss on Chart Classification: A study on Noisy Charts and Confusing Chart Pairs Effect of Attention and Triplet loss on Chart Classification: A study on Noisy Charts and Confusing Chart Pairs, (2022)
- [55] K. Davila, B.U. Kota, S. Setlur, V. Govindaraju, C. Tensmeyer, S. Shekhar, R. Chaudhry, *ICDAR 2019 Competition on Harvesting Raw Tables from Infographics (CHART-*

- Infographics), in: 2019 International Conference on Document Analysis and Recognition (ICDAR), IEEE, Sydney, Australia, 2019: pp. 1594–1599.
- [56] K. Davila, C. Tensmeyer, S. Shekhar, H. Singh, S. Setlur, V. Govindaraju, ICPR 2020 - Competition on Harvesting Raw Tables from Infographics, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12668 LNCS (2021) 361–380
- [57] A. Dhote, M. Javed, D.S. Doermann, A Survey and Approach to Chart Classification, Document Analysis and Recognition – ICDAR 2023 Workshops. ICDAR 2023. Lecture Notes in Computer Science 14193 (2023) 67–82.
- [58] V. Karthikeyani, S. Nagarajan, Machine Learning Classification Algorithms to Recognize Chart Types in Portable Document Format (PDF) Files, Int J Comput Appl 39 (2012) 1–5
- [59] B. Cheng, R.J. Stanley, S. Antani, G.R. Thoma, Graphical figure classification using data fusion for integrating text and image features, Proceedings of the International Conference on Document Analysis and Recognition, ICDAR (2013) 693–697
- [60] L. Battle, P. Duan, Z. Miranda, D. Mukusheva, R. Chang, M. Stonebraker, Beagle: Automated Extraction and Interpretation of Visualizations from the Web, Conference on Human Factors in Computing Systems - Proceedings 2018-April (2018) 1–8
- [61] J. Zhu, J. Ran, R.K. Lee, K. Choo, Z. Li, AutoChart: A Dataset for Chart-to-Text Generation Task, ArXiv Preprint 1 (2021) 1–9
- [62] W. Xue, D. Chen, B. Yu, Y. Chen, S. Zhou, W. Peng, ChartDETR: A Multi-shape Detection Network for Visual Chart Recognition, ArXiv (2023) 1–8
- [63] K.S. Sudeep, K.K. Pal, Preprocessing for image classification by convolutional neural networks, 2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTEICT 2016 - Proceedings (2017) 1778–1781
- [64] L.-R. Barrios-Roqueme, H. Altamar-Mercado, A. Patiño-Vanegas, Effect of Preprocessing in Deep Learning Systems to Aid the Diagnosis Using X-Ray Pulmonary Images, Figueroa-García, J.C., Hernández, G., Villa Ramirez, J.L., Gaona García, E.E. (Eds) Applied Computer Sciences in Engineering. WEA 2023. Communications in Computer and Information Science 1928 (2023) 185–194
- [65] Y. Liu, L.-K. Soh, E. Lorang, Investigating coupling preprocessing with shallow and deep convolutional neural networks in document image classification, J Electron Imaging 30 (2021) 1–30
- [66] F. Bajić, M. Habijan, K. Nenadić, A synthetic dataset of different chart types for advancements in chart identification and visualization, Data Brief 53 (2024) 110233
- [67] F. Bajić, ChartDS - chartDataset.zip. figshare., (2022).
- [68] F. Bajić, J. Job, Data Extraction of Circular-Shaped and Grid-like Chart Images, Journal of Imaging 2022, Vol. 8, Page 136 8 (2022) 136
- [69] F. Bajić, ChartDS 2022 - chartDataset.zip Srce Dabar, (2022)

- [70] F. Bajić, M. Habijan, K. Nenadić, ChartDataset2023: Introducing a Synthetic Dataset Featuring Various Chart Types for Chart Identification and Visualization, 1 (2023)
- [71] G. Chaladze, Kalatozishvili L, Linnaeus 5 Dataset for Machine Learning, 2017
- [72] A. Krizhevsky, Learning Multiple Layers of Features from Tiny Images, University of Toronto, 2009
- [73] Stan Z. Li, Anil K. Jain, Handbook of Face Recognition, Handbook of Face Recognition 1 (2011)
- [74] S. Khan, H. Rahmani, S.A.A. Shah, M. Bennamoun, A Guide to Convolutional Neural Networks for Computer Vision, (2018)
- [75] M. Bianchini, F. Scarselli, On the complexity of neural network classifiers: A comparison between shallow and deep architectures, IEEE Trans Neural Netw Learn Syst 25 (2014) 1553–1565
- [76] C.C. Aggarwal, Machine Learning with Shallow Neural Networks, Neural Networks and Deep Learning (2018) 53–104
- [77] J. Li, S. Xie, Z. Chen, H. Liu, J. Kang, Z. Fan, W. Li, A Shallow Convolutional Neural Network for Apple Classification, IEEE Access 8 (2020) 111683–111692
- [78] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks For Large-Scale Image Recognition, in: 3rd International Conference on Learning Representations, San Diego, CA, USA, 2015: pp. 1–14
- [79] J. Bromley, I. Guyon, Y. Lecun, E. Sickinger, R. Shah, A. Bell, L. Holmdel, Signature Verification using a “Siamese” Time Delay Neural Network, Adv Neural Inf Process Syst 6 (1993).
- [80] L. Song, Di. Gong, Z. Li, C. Liu, W. Liu, Occlusion Robust Face Recognition Based on Mask Learning with PairwiseDifferential Siamese Network, Proceedings of the IEEE International Conference on Computer Vision 2019-October (2019) 773–782.
- [81] S. Dey, A. Dutta, J.I. Toledo, S.K. Ghosh, J. Lladós, U. Pal, SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification, ArXiv 1 (2017) 1–7.
- [82] Z. Langford, L. Eisenbeiser, M. Vondal, Robust Signal Classification Using Siamese Networks, WiseML 2019: Proceedings of the ACM Workshop on Wireless Security and Machine Learning (2019) 1–5.
- [83] Z. Lian, Y. Li, J. Tao, J. Huang, Speech Emotion Recognition via Contrastive Loss under Siamese Networks, ASMMC-MMAC 2018 - Proceedings of the Joint Workshop of the 4th Workshop on Affective Social Multimedia Computing and 1st Multi-Modal Affective Computing of Large-Scale Multimedia Data, Co-Located with MM 2018 (2019) 21–26.
- [84] C. Molnar, Interpretable machine learning: a guide for making black box models explainable, Independently published, 2022.
- [85] J. Heaton, Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016, 800 pp, ISBN: 0262035618 Article, Genet Program Evolvable Mach 19 (2018) 305–307.

- [86] M.D. Zeiler, R. Fergus, Visualizing and Understanding Convolutional Networks, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8689 LNCS (2013) 818–833.
- [87] F. Lei, X. Liu, Q. Dai, · Bingo, W.-K. Ling, Shallow convolutional neural network for image classification, SN Appl Sci 2 (2019) 1–8.
- [88] GitHub - Lyken17/pytorch-OpCounter: Count the MACs / FLOPs of your PyTorch model., (n.d.). <https://github.com/Lyken17/pytorch-OpCounter>
- [89] C. Douwes, P. Esling, J.-P. Briot, Energy Consumption of Deep Generative Audio Models, (2021)
- [90] S. Wang, Z. Pu, Q. Li, Y. Guo, M. Li, Edge Computing-Enabled Crowd Density Estimation based on Lightweight Convolutional Neural Network, 2021 IEEE International Smart Cities Conference, ISC2 2021 (2021)
- [91] K. Siu, D.M. Stuart, M. Mahmoud, A. Moshovos, Memory Requirements for Convolutional Neural Network Hardware Accelerators, 2018 IEEE International Symposium on Workload Characterization, IISWC 2018 (2018) 111–121
- [92] Y. Gao, Y. Liu, H. Zhang, Z. Li, Y. Zhu, H. Lin, M. Yang, Estimating GPU memory consumption of deep learning models, ESEC/FSE 2020 - Proceedings of the 28th ACM Joint Meeting European Software Engineering Conference and Symposium on the Foundations of Software Engineering (2020) 1342–1352
- [93] W. Jeon, G. Ko, J. Lee, H. Lee, D. Ha, W.W. Ro, Deep learning with GPUs, Advances in Computers 122 (2021) 167–215
- [94] R. Chauhan, K.K. Ghanshala, R.C. Joshi, Convolutional Neural Network (CNN) for Image Detection and Recognition, ICSCCC 2018 - 1st International Conference on Secure Cyber Computing and Communications (2018) 278–282
- [95] J. Jepakoch, D.M. Mugo, B.K. Kenduiywo, E.C. Too, The Effect of Adaptive Learning Rate on the Accuracy of Neural Networks, International Journal of Advanced Computer Science and Applications 12 (2021) 736–751
- [96] R. Lin, Analysis on the Selection of the Appropriate Batch Size in CNN Neural Network, Proceedings - 2022 International Conference on Machine Learning and Knowledge Engineering, MLKE 2022 (2022) 106–109
- [97] M. Kuhn, K. Johnson, Applied predictive modeling, Applied Predictive Modeling (2013) 1–600
- [98] P. Zhang, Model Selection Via Multifold Cross Validation, <https://doi.org/10.1214/Aos/1176349027> 21 (1993) 299–313.
- [99] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, 2nd ed., Springer New York, New York, NY, 2009.
- [100] P. Burman, A Comparative Study of Ordinary Cross-Validation, v-Fold Cross-Validation and the Repeated Learning-Testing Methods, Biometrika 76 (1989) 503

- [101] F. De O. Costa, M. Eckmann, W.J. Scheirer, A. Rocha, Open set source camera attribution, *Brazilian Symposium of Computer Graphic and Image Processing* (2012) 71–78.
- [102] J. Huang, D. Prijatelj, J. Dulay, W. Scheirer, Measuring Human Perception to Improve Open Set Recognition, *IEEE Trans Pattern Anal Mach Intell* 45 (2022) 11382–11389.
- [103] W.J. Scheirer, A. Rocha, A. Sapkota, T.E. Boult, Towards Open Set Recognition, *IEEE Trans Pattern Anal Mach Intell* X (2012).
- [104] E.M. Rudd, L.P. Jain, W.J. Scheirer, T.E. Boult, I. Fellow, The Extreme Value Machine, *ArXiv* 1 (2017) 1–12.
- [105] T.E. Boult, S. Cruz, A.R. Dhamija, M. Gunther, J. Henrydoss, W.J. Scheirer, Learning and the Unknown: Surveying Steps Toward Open World Recognition, *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (2019).
- [106] H. Zhang, S. Member, V.M. Patel, S. Member, Sparse Representation-based Open Set Recognition, *IEEE Trans Pattern Anal Mach Intell* (2016) 1.
- [107] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Trans Pattern Anal Mach Intell* 39 (2015) 1137–1149.
- [108] Z. Chen, Y. Fu, Y. Zhang, Y.-G. Jiang, X. Xue, L. Sigal, Multi-level Semantic Feature Augmentation for One-shot Learning, *IEEE Transactions on Image Processing* 28 (2018) 4594–4605
- [109] Y. Fu, T. Xiang, Y.G. Jiang, X. Xue, L. Sigal, S. Gong, Recent Advances in Zero-Shot Recognition: Toward Data-Efficient Understanding of Visual Content, *IEEE Signal Process Mag* 35 (2018) 112–125
- [110] T.G. Dietterich, Steps Toward Robust Artificial Intelligence, *AI Mag* 38 (2017) 3–24.
- [111] C. Geng, S.-J. Huang, S. Chen, Recent Advances in Open Set Recognition: A Survey, *IEEE Trans Pattern Anal Mach Intell* 43 (2018) 3614–3631
- [112] B.M. Lake, R. Salakhutdinov, J. Gross, J.B. Tenenbaum, One shot learning of simple visual concepts, *Proceedings of the Annual Meeting of the Cognitive Science Society* 33 (2011) 1–6
- [113] H.C. Purchase, Twelve years of diagrams research, *J Vis Lang Comput* 25 (2014) 57–75
- [114] L. Yang, W. Huang, C.L. Tan, Semi-automatic Ground Truth Generation for Chart Image Recognition, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 3872 LNCS (2006) 324–335
- [115] M. Cliche, D. Rosenberg, D. Madeka, C. Yee, Scatteract: Automated extraction of data from scatter plots, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10534 LNAI (2017) 135–150
- [116] L. Chen, K. Zhao, An Approach for Chart Description Generation in Cyber–Physical–Social System, *Symmetry (Basel)* 13 (2021) 1552

- [117] P. De, Automatic Data Extraction from 2D and 3D Pie Chart Images, Proceedings of the 8th International Advance Computing Conference, IACC 2018 (2018) 20–25
- [118] J. Lal, A. Mitkari, M. Bhosale, D. Doermann, LineFormer: Rethinking Line Chart Data Extraction as Instance Segmentation, ArXiv 1 (2023) 1–15.
- [119] J. Luo, Z. Li, J. Wang, C.-Y. Lin, ChartOCR: Data Extraction from Charts Images via a Deep Hybrid Framework, IEEE Winter Conference on Applications of Computer Vision (WACV) (2021) 1–9.
- [120] J. Sreevalsan-Nair, K. Dadhich, S.C. Daggubati, I. Hotz, T. Bin Masood, F. Sadlo, J. Tierny, E.) Springer-Verlag, Tensor Fields For Data Extraction From Chart Images: Bar Charts And Scatter Plots, ArXiv (2020) 1–17.
- [121] J. Shtok, S. Harary, O. Azulai, A.R. Goldfarb, A. Arbelle, L. Karlinsky, CHARTER: heatmap-based multi-type chart data extraction, ArXiv (2021) 1–6.
- [122] S. V P, M. Yusuf Hassan, M. Singh, LINEEX: Data Extraction from Scientific Line Charts, IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2023) 1–9.
- [123] H. Kato, M. Nakazawa, H.-K. Yang, M. Chen, B. Stenger, Parsing Line Chart Images Using Linear Programming, IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2022) 1–10
- [124] A. Mishchenko, N. Vassilieva, Chart image understanding and numerical data extraction, 6th International Conference on Digital Information Management, ICDIM 2011 (2011) 115–120
- [125] Al-Zaidy, Rabah A., C.L. Giles, Automatic extraction of data from bar charts, Proceedings of the 8th International Conference on Knowledge Capture, K-CAP 2015 (2015) 1–4
- [126] R.A. Al-Zaidy, Sagnik Ray Choudhury, C. Lee Giles, Automatic summary generation for scientific data charts, Workshops at the Thirtieth Aai Conference on Artificial Intelligence 16 (2016) 658–663
- [127] F. Zhou, Y. Zhao, W. Chen, Y. Tan, Y. Xu, Y. Chen, C. Liu, Y. Zhao, Reverse-engineering Bar Charts Using Neural Networks, J Vis (Tokyo) (2021) 419–435
- [128] keras-ocr — keras_ocr documentation, (n.d.). <https://keras-ocr.readthedocs.io/en/latest/>
- [129] R.M. Haralick, S.R. Sternberg, X. Zhuang, Image Analysis Using Mathematical Morphology, IEEE Trans Pattern Anal Mach Intell PAMI-9 (1987) 532–550
- [130] G.J. Awcock, R. Thomas, Applied Image Processing, Applied Image Processing (1995)
- [131] T. Gebu, J. Morgenstern, Black AI, B. Vecchione, J. Wortman Vaughan, K. Crawford, Datasheets for Datasets, Commun ACM 65 (2021) 86–92
- [132] G. Singh, S. Agrawal, B.S. Sohi, A robust methodology for creating large image datasets using a universal format, Advances in Intelligent Systems and Computing 1166 (2021) 279–288
- [133] J. Zhao, M. Fan, M. Feng, ChartSeer: Interactive Steering Exploratory Visual Analysis with Machine Intelligence, IEEE Trans Vis Comput Graph 28 (2022) 1500–1513

- [134] K. Kafle, B. Price, S. Cohen, C. Kanan, DVQA: Understanding Data Visualizations via Question Answering, in: *Computer Vision and Pattern Recognition*, arXiv, 2018
- [135] S. Ebrahimi Kahou, V. Michalski, A. Atkinson, Á. Kádár, A. Trischler, Y. Bengio, FigureQA: An Annotated Figure Dataset for Visual Reasoning, *ArXiv 1 (2018) 1–20*
- [136] R. Chaudhry, S. Shekhar, A. Research, U. Gupta, P. Bansal, A. Joshi, LEAF-QA: Locate, Encode & Attend for Figure Question Answering, *ArXiv 1 (219AD) 1–10*
- [137] T. Araújo, P. Chagas, J. Alves, C. Santos, B.S. Santos, B.S. Meiguins, A Real-World Approach on the Problem of Chart Recognition Using Classification, Detection and Perspective Correction, *Sensors 2020, Vol. 20, Page 4370 20 (2020) 4370*
- [138] N. Methani, P. Ganguly, M.M. Khapra, P. Kumar, PlotQA: Reasoning over Scientific Plots, *Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020 (2019) 1516–1525*
- [139] T.Y. Hsu, C.L. Giles, T.H. Huang, SciCap: Generating Captions for Scientific Figures, *Findings of the Association for Computational Linguistics, Findings of ACL: EMNLP 2021 (2021) 3258–3264*
- [140] K. Davila, F. Xu, S. Ahmed, D.A. Mendoza, S. Setlur, V. Govindaraju, ICPR 2022: Challenge on Harvesting Raw Tables from Infographics (CHART-Infographics), *Proceedings - International Conference on Pattern Recognition 2022-August (2022) 4995–5001*
- [141] C.C. Beckner, Jr., C.L. Matson, Using mean-squared error to assess visual image quality, *Proceedings of the SPIE 6313 (2006) 129–137*
- [142] K. Gu, S. Wang, G. Zhai, W. Zhang, S. Ma, · Xiaokang Yang, B.K. Gu, Content-weighted mean-squared error for quality assessment of compressed images, *Signal Image Video Process 10 (2016) 1–8*
- [143] E. Karami, S. Prasad, M. Shehata, Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images, *ArXiv 1 (2017) 1–5*

SAŽETAK

Korištenje slikovnog prikaza podataka omogućuje jednostavniju pohranu informacija u dugotrajno pamćenje, zahtijevajući manju količinu pažnje čitatelja. Povijesni razvoj statistike potaknuo je potrebu za slikovnim prikazima kako bi se predstavile složene strukture podataka. U 20. stoljeću, s porastom informacija, nastala je potreba za sažimanjem podataka, a slikovni prikazi postaju učinkovito sredstvo za prikaz važnih informacija u istraživanjima. Vizualizacija podataka, poput grafikona, omogućuje jasno razumijevanje odnosa među podacima, olakšava usporedbu vrijednosti te pomaže u identifikaciji neočekivanih rezultata. U današnje vrijeme, kada količina podataka ubrzano raste, vizualizacija postaje ključna za pojednostavljenje interpretacije i donošenje odluka. Unatoč ostvarenom napretku u području koje se bavi interpretacijom slike grafikona, izazovi i dalje postoje. S izazovima su suočene osobe s oštećenim vidom i računala koja pokušavaju pristupiti informacijama sadržanim u slici grafikona. Brojna istraživanja usmjerena su na poboljšanje pristupačnosti i dostupnosti vizualizacija podataka.

Osnovni cilj ove disertacije bio je razviti automatizirane metode strojnog učenja za obradu digitalnih slika grafikona, omogućujući njihovu implementaciju u stvarnoj okolini. Razvijene metode obuhvaćaju cjelokupni proces, počevši od učenja i prepoznavanja vrsta grafikona do izdvajanja podatkovnih značajki, s mogućnošću primjene u drugim računalnim sustavima.

Tijekom provedenog istraživanja razvijene su sljedeće tri metode:

1. **Metoda za otvorenu klasifikaciju grafikona u elektroničkim dokumentima zasnovana na plitkim konvolucijskim neuronskim mrežama.** Predložena metoda prikazuje dvije arhitekture neuronske mreže — tradicionalnu i sijamsku. Ulaz u model neuronske mreže je slika grafikona ili bilo koja druga slika prisutna u elektroničkom dokumentu. Slike grafikona prolaze

kroz proces predobrade upotrebom metode koja uspješno uklanja značajan broj sadržajnih informacija, zadržavajući grafički element čiji sadržaj omogućuje razlikovanje vrste grafikona. U okviru evaluacije, razvijena su 32 modela, učena na šest javno dostupnih podatkovnih skupova.

2. **Metoda izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona za primjenu u računalnim sustavima.** Kružni grafikoni obuhvaćaju tortne, prstenaste i grafikone sunčevih zraka, dok mrežasti grafikoni uključuju toplinske karte i grafikone napretka. Istraživanjem je razvijena prva metoda usmjerena na grafikone sunčevih zraka i toplinske karte. Metoda se sastoji od dva ključna modula: modula za predobradu slike i modula za detekciju strukturnih elemenata grafikona. Metoda se oslanja na binarne maske, morfologiju, detekciju ključnih točaka i matematički model.

3. **Metoda generiranja ispitnih skupova za standardiziranu validaciju uspješnosti izdvajanja podatkovnih značajki iz kružnih i mrežastih grafikona.** Metoda se temelji na slučajnim vrijednostima i lokacijama strukturalnih elemenata u slikama grafikona. Razvijeni podatkovni skup, ChartDS, proizlazi iz ove metode te se koristi u evaluaciji dviju prethodno predstavljenih metoda. ChartDS je javno dostupan u različitim podatkovnim repozitorijima, pružajući standardiziranu osnovu za procjenu i usporedbu učinkovitosti različitih sustava u analizi i interpretaciji slika grafikona.

Razvijene metode potkrijepljene su detaljnim evaluacijama čiji rezultati su javno objavljeni u raznim znanstvenim radovima.

Ključne riječi: grafikon, izdvajanje značajki, otvorena klasifikacija, plitka konvolucijska neuronska mreža, podatkovni skup, sijamska neuronska mreža, vizualizacija podataka

ABSTRACT

Chart image classification using a shallow convolutional neural network and feature extraction from classified charts

The use of graphical representation of data allows for easier storage of information in long-term memory, requiring less attention from the reader. The historical development of statistics has driven the need for graphical representations to represent complex data structures. In the 20th century, with the increase of information, there was a need to summarize data, and graphical representations became an effective means of presenting important information in scientific research. Data visualization, such as charts or graphs, provides a clear understanding of the relationships between the data, facilitates comparison of values, and helps identify unexpected results. Nowadays, when the data is growing rapidly, visualization becomes essential for simplifying interpretation and decision-making. Despite the progress made in the field of chart image interpretation, challenges remain. Challenges are faced by visually impaired people and computers trying to access the information contained in the chart image. Numerous researches are aimed at improving the accessibility and availability of data visualizations.

The main goal of this dissertation was to develop automated machine learning methods for processing digital chart images, enabling their implementation in a real environment. The developed methods cover the entire process, starting from learning and recognizing chart-types to extracting data features, with the possibility of intergration in other computer systems.

During the conducted research, the following three methods were developed:

1. **A method for accessing information contained within charts in electronic documents based on shallow convolutional neural networks.** The proposed method enables the dual application of the neural network, both as traditional and Siamese. The input to the neural network model is a chart image or any other image present in the document. Chart images undergo a pre-processing using a method that successfully removes a significant amount of information while retaining the graphic element. This retained content allows distinguishing of the chart-type. As part of the evaluation, 32 models were developed and trained on six publicly available datasets.
2. **A method for extracting features from circular and grid charts for application in computer systems.** Circular charts encompass pie, donut, and sunburst diagrams, while grid charts comprise heat maps and waffle charts. The research developed the first method focused on sunburst diagrams and heat maps. The method consists of two modules: one for image pre-processing and another for detecting structural elements of the chart. The method relies on binary masks, morphology, keypoint detection, and a mathematical model.
3. **Validation of the performance of the proposed methods on the newly created chart database.** The method is based on random values and locations of structural elements in chart images. The developed dataset, ChartDS, derives from this method and is used in the evaluation of the two previously presented methods. ChartDS is publicly available in various data repositories, providing a standardized basis for evaluating and comparing the performance of different systems in the analysis and interpretation of chart images.

The developed methods are supported by detailed evaluations, the results of which have been publicly published in various scientific works.

Keywords: chart, data visualization, dataset, feature extraction, open classification, shallow convolutional neural network, Siamese neural network

ŽIVOTOPIS

Filip Bajić rođen je u Požegi u Hrvatskoj. Diplomirao je 2015. godine na Elektrotehničkom fakultetu u Osijeku gdje je stekao zvanje magistra računarstva. Živi i radi u Zagrebu. Od 2016. godine zaposlen je u Sveučilišnom računskom centru (Srce) gdje trenutno radi kao voditelj Informacijskog Sustava Studentskih Prava (ISSP), Informacijskog Sustava Akademskih Kartica (ISAK) i državnih studentskih stipendija. U okviru poslovnih aktivnosti radi na projektima Ministarstva znanosti i obrazovanja, član je nekoliko Povjerenstava Ministarstva znanosti i obrazovanja za podizanje studentskog standarda, te radi na međunarodnim projektima povezivanja studenata (*MyAcademicID, European Digital Student Service Infrastructure, European University of Post-Industrial Cities, European Student Card*). Krajem 2017. godine upisuje Poslijediplomski sveučilišni studij na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija Osijek. U svojem znanstveno istraživačkom radu bavi se računalnom obradom i analizom različitih vrsta grafikona metodama strojnog i dubokog učenja, s fokusom na upotrebu plitkih konvolucijskih neuronskih mreža. Autor je devet radova u znanstvenim časopisima te jednog rada na međunarodnim znanstvenim skupovima.

POPIS SLIKA

Slika 2.1: Algoritam predobrade slike grafikona: a) prikaz rada algoritma na slici točkastog grafikona, b) pseudokod algoritma.....	15
Slika 2.2: Primjeri ulaznih i izlaznih slika predstavljenog algoritma za sedam klasa grafikona (A-G). Navedene slike sadržane se unutar ChartDS podatkovnog skupa.	16
Slika 2.3: Pojava ujednačavanja različitih slika iste vrste i istog tipa grafikona.	16
Slika 2.4: Prikaz ograničenja rada algoritma predobrade slike grafikona. Na slici a) slika iz podatkovnog skupa ICDAR7 i b) izlazna slika predstavljenog algoritma.	17
Slika 2.5: Primjer slika sadržanih unutar ReVision podatkovnog skupa.	18
Slika 2.6: Primjer slika sadržanih unutar ICDAR7 podatkovnog skupa.....	19
Slika 2.7: Primjer slika sadržanih unutar ChartDS podatkovnog skupa.	19
Slika 2.8: Primjer slika sadržanih unutar Linnaeus podatkovnog skupa.....	19
Slika 2.9: Primjer slika sadržanih unutar CIFAR10 podatkovnog skupa.....	20
Slika 2.10: Prikaz arhitekture plitke konvolucijske neuronske mreže za klasifikaciju ulazne slike u sedam klasa grafikona (A-G) i zasebnu klasu koja ne sadrži slike grafikona (H).	22
Slika 2.11: Prikaz arhitekture sijamske plitke konvolucijske neuronske mreže za klasifikaciju ulazne slike grafikona.....	25
Slika 2.12: MCCV metoda prikazana na podatkovnom skupu s 20 (a, b) i 500 (c, d) slika po vrsti grafikona. Na slikama (a, c) nalazi se postupak učenja na ChartDS, a na slikama (b, d) na ICDAR7 podatkovnom skupu. <i>Acc</i> označava točnost klasifikacije, a <i>i</i> broj učenja modela.	33

Slika 2.13: Prikaz točnosti klasifikacije grafikona (A-G). Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.	37
Slika 2.14: Prikaz točnosti klasifikacije zasebne klase (H). Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.	37
Slika 2.15: Prikaz točnosti klasifikacije za model učen na ChartDS, a testiran na ICDAR7 podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.	38
Slika 2.16: Prikaz točnosti klasifikacije za model učen i testiran na ChartDS podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.	38
Slika 2.17: Prikaz točnosti klasifikacije za model učen na ICDAR7, a testiran na ChartDS podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.	39
Slika 2.18: Prikaz točnosti klasifikacije za model učen i testiran na ICDAR7 podatkovnom skupu. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune. Model je učen korištenjem 500 slika, a testiran korištenjem 100 slika po klasi.	39
Slika 2.19: Prikaz točnosti klasifikacije grafikona. Na slici a) model je učen na ChartDS i testiran na ChartDS i ICDAR7 podatkovnom skupu, a na slici b) model je učen na ICDAR7 i testiran na ChartDS i ICDAR7 podatkovnom skupu.	43
Slika 2.20: Prikaz točnosti klasifikacije zasebne klase (H). Na slici a) model je učen na ChartDS podatkovnom skupu, a na slici b) model je učen na ICDAR7 podatkovnom skupu.	43
Slika 2.21: Usporedba slike iz Linnaeus podatkovnog skupa koja ne sadrži grafikon s: a) drugom slikom iz Linnaeus podatkovnog skupa koja ne sadrži grafikon ($SS = 0.21$) i b) slikom točkastog grafikona iz ChartDS podatkovnog skupa ($SS = 0.60$).	44
Slika 3.1: Prikaz metode izdvajanja podatkovnih značajki iz kružnih grafikona. Metoda se sastoji od modula za predobradu slike i detektora.	54
Slika 3.2: Prikaz metode izdvajanja značajki iz mrežastih grafikona. Metoda se sastoji od modula za predobradu slike i detektora.	56

Slika 3.3: Tijek grafikona sunčevih zraka kroz modul za predobradu slike kružnog grafikona. Na slici a) ulazna slika, b) detekcija tekstualnih elemenata, c) stvaranje binarne maske, d) uklanjanje tekstualnih elemenata i e) izlazna slika modula.	59
Slika 3.4: Tijek grafikona sunčevih zraka kroz modul detektor kružnog grafikona. Na slici je prikazana detekcija kružnice (a, d, g), detekcija linija (b, e, h) i detekcija ključnih točaka (c, f, i).	60
Slika 3.5: Rekonstrukcija grafikona sunčevih zraka upotrebom izdvojenih podatkovnih značajki. Ključne točke i linije su označene crnom bojom. Algoritam je uspješno detektirao i pripadajuću legendu.	60
Slika 3.6: Tijek grafikona toplinske karte kroz modul za predobradu slike mrežastog grafikona. Na slici a) ulazna slika, b) detekcija tekstualnih elemenata, c) uklanjanje tekstualnih elemenata i d) izlazna slika modula.....	62
Slika 3.7: Tijek grafikona toplinske karte kroz modul detektor kružnog grafikona. Na slici a) detekcija linija i b) detekcija ključnih točaka.....	62
Slika 3.8: Rekonstrukcija grafikona toplinske karte upotrebom izdvojenih podatkovnih značajki. Ključne točke i linije su označene crnom bojom.	62
Slika 4.1: Pseudokod algoritma za usporedbu podatkovnih skupova ChartDS i ICDAR7.....	76
Slika 4.2: Prikaz točnosti klasifikacije za model učen i testiran na slikama grafikona iz ChartDS (DS1) i ICDAR7 (DS2) podatkovnih skupova. Na slici a) prikazana je matrica zabune, a na slici b) prikazana je statistička analiza matrice zabune.....	77

POPIS TABLICA

Tablica 2.1: Pregled istraživanja u području klasifikacije slike grafikona. Istaknuti redovi predstavljaju radove izrađene u sklopu ovog istraživanja.....	11
Tablica 2.2: Napredak kružnog, stupčastog i linijskog grafikona kroz konvolucijske slojeve. U prvom retku nalazi se primjer filtra receptivnog polja, a u ostalim redovima primjer kanala svakog konvolucijskog sloja.	27
Tablica 2.3: Usporedba prostorne i vremenske složenosti za razne CNN arhitekture. Istaknuti redovi predstavljaju arhitekture neuronskih mreža za klasifikaciju slike grafikona izrađene u sklopu ovog istraživanja.....	30
Tablica 2.4: Usporedba ostvarenih rezultata različitih arhitektura s obzirom na potrebno vrijeme učenja. Istaknuti redovi predstavljaju arhitekture neuronskih mreža za klasifikaciju slike grafikona izrađene u sklopu ovog istraživanja.	31
Tablica 3.1: Pregled istraživanja u području izdvajanja podatkovnih značajki iz slike grafikona. Istaknuti red predstavlja rad izrađen u sklopu ovog istraživanja.	51
Tablica 3.2: Rezultati evaluacije metode izdvajanja podatkovnih značajki na kružnim grafikonima	61
Tablica 3.3: Rezultati evaluacije metode izdvajanja značajki na mrežastim grafikonima	63
Tablica 3.4: Rezultati evaluacije metode izdvajanja značajki na stvarnim grafikonima.....	63
Tablica 4.1: Pregled istraživanja u području. Istaknuti red predstavlja rad izrađen u sklopu ovog istraživanja.	69
Tablica 4.2: Statistika podatkovnog skupa ChartDS.	74

PRILOZI

Prilog P.1. Python programsko rješenje SCNN modela

```
model = Sequential([
    Input(shape=(Config.img_height, Config.img_width, 1)),
    Rescaling(1./255),
    Conv2D(64, (3, 3), strides=1, padding="same"),
    Conv2D(64, (3, 3), strides=1, padding="same"),
    MaxPooling2D(pool_size=(2, 2), strides=2, padding="valid"),
    Conv2D(128, (3, 3), strides=1, padding="same"),
    Conv2D(128, (3, 3), strides=1, padding="same"),
    MaxPooling2D(pool_size=(2, 2), strides=2, padding="valid"),
    Flatten(),
    Dense(num_classes),
    Activation("softmax")
])
```

Prilog P.2. Python programsko rješenje SSCNN modela

```
class SSCNN(nn.Module):
    def __init__(self):
        super(SSCNN, self).__init__()
        self.cnn1 = nn.Sequential(
            nn.Conv2d(in_channels=1, out_channels=64,
                    kernel_size=3,
                    stride=1, padding=1, bias=False),
            nn.Conv2d(in_channels=64, out_channels=64,
                    kernel_size=3,
                    stride=1, padding=1, bias=False),
            nn.MaxPool2d(kernel_size=2, stride=2, padding=0),
            nn.Conv2d(in_channels=64, out_channels=128,
                    kernel_size=3,
                    stride=1, padding=1, bias=False),
            nn.Conv2d(in_channels=128, out_channels=128,
                    kernel_size=3,
                    stride=1, padding=1, bias=False),
            nn.MaxPool2d(kernel_size=2, stride=2, padding=0),
        )

        self.fc1 = nn.Sequential(
            nn.Flatten(),
            nn.Linear(32768, 1)
        )

    def forward_once(self, x):
        output = self.cnn1(x)
        output = output.view(output.size()[0], -1)
        output = self.fc1(output)
        return output

    def forward(self, input1, input2):
        output1 = self.forward_once(input1)
        output2 = self.forward_once(input2)
        return output1, output2
```

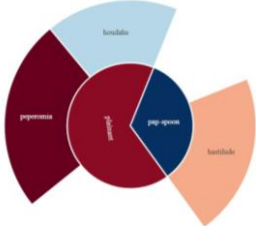
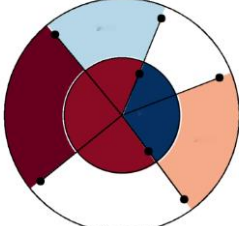
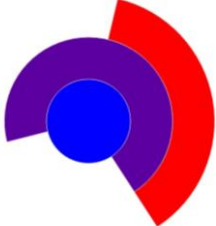
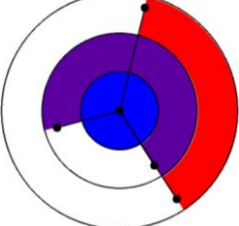

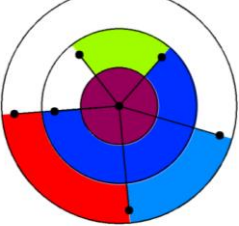

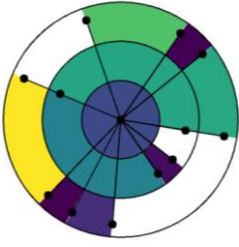



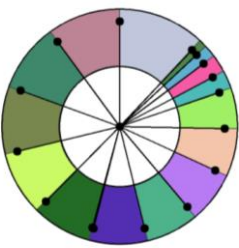

Prilog P.3. Parametri učenja modela

```
class Config():
    img_width = 64
    img_height = 64
    seed = 1529 // 1529 ChartDS // 635 ICDAR7
    epochs = 20
    batch = 64
    learning_rate = ReduceLROnPlateau(monitor='val_loss', factor=np.sqrt(0.1), patience=4)
```

Prilog P.4. Izdvojene podatkovne značajke kružnog grafikona iz slike 3.5

```
{
  "Name": "Sunburst diagram",
  "Path": "datasetBuild/train/sunburstSimple/sunburstSimple_7.png",
  "BackgroundColorHex": "#ffffff",
  "BackgroundColorName": "white",
  "CenterPoint": [ 608, 383 ],
  "Levels": 3,
  "Level_0": {
    "Location": "Inner",
    "Segments": 1,
    "Radius": 93,
    "TransitionPoints": [[ 608, 383 ]],
    "2PointAngle": [ 360 ],
    "Sorted2PointAngle": [ 360 ],
    "SegmentAreaPercent": [ 100 ],
    "SegmentColorHex": [ "#fa0005" ],
    "SegmentColorNames": [ "red" ],
    "NoValueEmptyColor": "None",
    "NumberOfEmptySegments": 0
  },
  "Level_1": {
    "Location": "Middle",
    "Segments": 3,
    "Radius": 189,
    "TransitionPoints": [[ 626, 540 ], [ 726, 489 ], [ 700, 254 ]],
    "2PointAngle": [ 83, 41, 306 ],
    "Sorted2PointAngle": [ 41, 83, 306 ],
    "SegmentAreaPercent": [ 11.67, 61.94, 26.39 ],
    "SegmentColorHex": [ "#0000ff", "#ff0000", "#ffffff" ],
    "SegmentColorNames": [ "blue", "red", "white" ],
    "NoValueEmptyColor": "#ffffff",
    "NumberOfEmptySegments": 1
  },
  "Level_2": {
    "Location": "Outter",
    "Segments": 3,
    "Radius": 281,
    "TransitionPoints": [[ 752, 178 ], [ 430, 206 ], [ 358, 377 ]],
    "2PointAngle": [ 306, 225, 182 ],
    "Sorted2PointAngle": [ 182, 225, 306 ],
    "SegmentAreaPercent": [ 11.94, 22.5, 65.56 ],
    "SegmentColorHex": [ "#0500fa", "#af0050", "#ffffff" ],
    "SegmentColorNames": [ "blue", "crimson", "white" ],
    "NoValueEmptyColor": "#ffffff",
    "NumberOfEmptySegments": 1
  },
  "Legend": {
    "ColorHex": [ "#0300fb", "#3500c9", "#680096", "#9a0064", "#cc0032", "#fe0000" ],
    "ColorName": [ "blue", "mediumblue", "purple", "purple", "crimson", "red" ],
    "ColorValue": [ 40, 50, 60, 70, 80, 90 ],
    "ColorLocations": [[1229,650], [1229,548], [1229,446], [1229,344], [1229,242], [1229,141]],
    "Type": "Linear Sequence"
  }
}
```

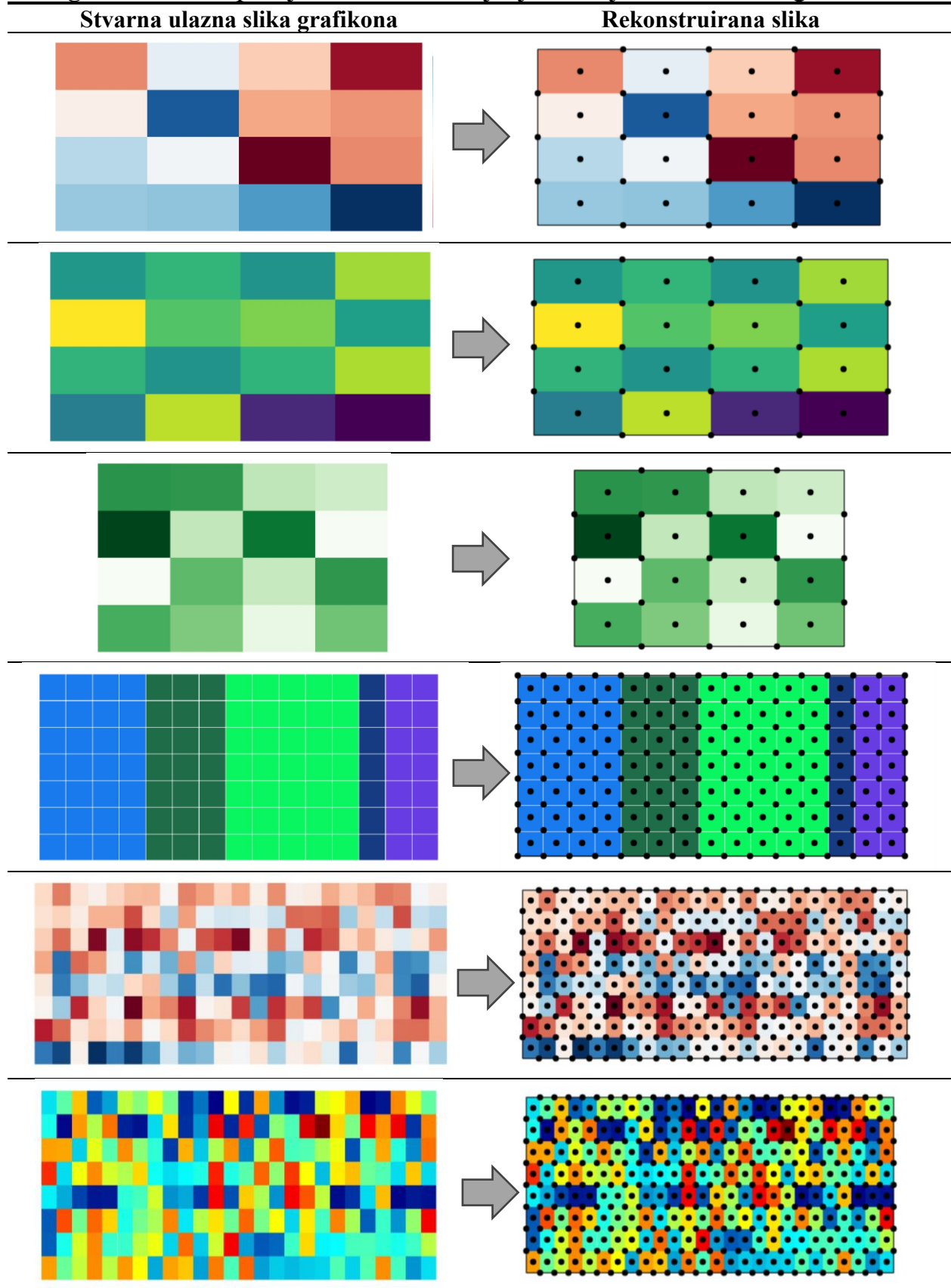
Prilog P.5. Dodatni primjeri metode izdvajanja značajki iz kružnih grafikona

Stvarna ulazna slika grafikona	Rekonstruirana slika
	
	
	
	
	
	

Prilog P.6. Izdvojene podatkovne značajke mrežastog grafikona iz slike 3.8

```
{
  "Name": "Heatmap/Waffle",
  "Path": "datasetBuild/train/heatmapComplex/heatmapComplex_7.png",
  "BackgroundColorHex": "#ffffff",
  "BackgroundColorName": "white",
  "Rows": 10,
  "Columns": 9,
  "Segments": 90,
  "OneSegmentHeight": 55,
  "OneSegmentWidth": 115,
  "OneSegmentArea": 6325,
  "HeatmapWaffleArea": 569250,
  "Row_0": {
    "Segments": 9,
    "TransitionPoints": [
      [175, 100], [290, 100], [404, 100], [518, 100], [633, 100], [747, 100], [861, 100], [976, 100]
    ],
    "MidPoints": [
      [118, 127], [232, 127], [347, 127], [461, 127], [575, 127], [690, 127], [804, 127], [918, 127], [1032, 127]
    ],
    "SegmentColorHex": [
      "#ec0000", "#bb0000", "#fb4000", "#03a2d7", "#fc5b00", "#d7ff29", "#feff01", "#f40000", "#00289d"
    ],
    "SegmentColorNames": [
      "red", "darkred", "orangered", "darkturquoise", "orangered", "greenyellow", "yellow", "red", "darkblue"
    ],
    "NoValueEmptyColor": "None",
    "NumberOfEmptySegments": 0
  },
  "Row_1": { ... },
  "Row_2": { ... },
  "Row_3": { ... },
  "Row_4": { ... },
  "Row_5": { ... },
  "Row_6": { ... },
  "Row_7": { ... },
  "Row_8": { ... },
  "Row_9": { ... },
  "Legend": {
    "ColorHex": ["#bc0000", "#fc8700", "#efff0f", "#6dff94", "#04e8f5", "#0184c9", "#0032a3", "#00128f"],
    "ColorName": [
      "darkred", "darkorange", "yellow", "lightgreen", "cyan", "darkcyan", "darkblue", "darkblue"
    ],
    "ColorValue": [0, 50, 100, 150, 200, 250, 300, 350],
    "ColorLocations": [
      [1150, 601], [1150, 496], [1150, 442], [1150, 389], [1150, 335], [1150, 283], [1150, 229], [1150, 175]
    ]
  },
  "Type": "Linear Sequence"
}
```

Prilog P.7. Dodatni primjeri metode izdvajanje značajki iz mrežastih grafikona



Prilog P.8. Parametri metode za stvaranje slika grafikona

Naziv parametra	Vrijednost	
	Jednostavni oblik	Složeni oblik
Broj točaka razlučivosti	1366 x 768	1366 x 768
Naziv fonta	Slučajno odabrana vrijednost iz javno dostupnih fontova	
Naslov – sadržaj	Slučajno odabrana engleska riječ (imenica ili glagol) između 2 i 8 znakova, minimalno 2, maksimalno 15 riječi	
Naslov – vidljivost	Da	Da, ne
Naslov – lokacija	Lijevo-gore	Lijevo-gore, desno-gore, lijevo-dolje, desno dolje
Naslov – veličina fonta	Slučajno odabrana vrijednost između 20 i 30	
Naslov – udaljenost od grafikona	Slučajno odabrana vrijednost između 20 i 80	
Legenda-naslov – sadržaj	Slučajno odabrana engleska riječ (imenica ili glagol) između 2 i 8 znakova, minimalno 1, maksimalno 10 riječi	
Legenda – sadržaj	Ovisno o podatkovnom segmentu, ako je opisni tekst onda slučajno odabrana engleska riječ (imenica ili glagol) između 2 i 8 znakova, minimalno 1, maksimalno 3 riječi	
Legenda – vidljivost	Da	Da, ne
Legenda – lokacija	Desno-gore	Lijevo-gore, desno-gore, lijevo-dolje, desno dolje
Legenda – orijentacija	Vertikalna	Vertikalna, horizontalna
Legenda – debljina granice	0	Slučajno odabrana vrijednost između 0 i 2
Legenda – veličina fonta	Slučajno odabrana vrijednost između 16 i 24	
Koordinatna os – naslov	Slučajno odabrana engleska riječ (imenica ili glagol) između 2 i 8 znakova, minimalno 1, maksimalno 4 riječi	
Koordinatna os – mjerna jedinica	Slučajno odabrana mjerna jedinica iz Međunarodnog sustava mjernih jedinica	
Koordinatna os – broj oznaka	Slučajno odabrana vrijednost između 2 i 10	
Koordinatna os – debljina linije	Slučajno odabrana vrijednost između 0 i 2	

Naziv parametra	Vrijednost	
	Jednostavni oblik	Složeni oblik
Koordinatna os – vidljivost	Da	Da, ne
Koordinatna os – vidljivost oznaka	Da	Da, ne
Koordinatna os – orijentacija oznaka	0°	-90°, -45°, 0°, 45°, 90°
Koordinatna os – raspon vrijednosti	Od najmanje do najveće vrijednosti iz podatkovnog skupa	
Koordinatna os – veličina fonta	Slučajno odabrana vrijednost između 16 i 22	
Grafikon – vidljivost mreže	Ne	Da, ne
Grafikon – boja mreže	Bijela	Bijela, siva, crna
Grafikon – boja pozadine	Bijela	Naziv boje počine sa „svijetla“
Grafikon – boja okoline	Bijela	Naziv boje počine sa „svijetla“
Grafikon – rotacija	Nema	0° ≤ rotacija ≤ 360°
Podatkovni segment – broj podatkovnih segmenata	≤ 5	5 ≤ broj segmenata ≤ 20
Podatkovni segment – vrijednost	Slučajno odabrana vrijednost između -1000 i 1000	
Podatkovni segment – boja granice	Slučajno odabranih 6 znakova iz „0123456789ABCDEF“	
Podatkovni segment – tip granice	Neprekinut	Neprekinut, točke. crtice
Podatkovni segment – boja ispune	Slučajno odabranih 6 znakova iz „0123456789ABCDEF“	
Podatkovni segment – oznaka ključne točke	Nema	Nema, krug, trokut, kvadrat, dijament, strelica, zvijezda, križ
Podatkovni segment – ispis vrijednosti ključne točke	Ne	Da, ne

Naziv parametra	Vrijednost	
	Jednostavni oblik	Složeni oblik
Podatkovni segment – tip ispisa vrijednosti ključne točke	Nema	Nema, numerička, tekstualna, numeričko-tekstualna, postotak
Podatkovni segment – orijentacija ispisa vrijednosti ključne točke	Nema	Horizontalna, vertikalna
Podatkovni segment – veličina ključne točke	Slučajno odabrana vrijednost između 2 i 8	
Podatkovni segment – debljina granice	Slučajno odabrana vrijednost između 0 i 2	
Podatkovni segment – udaljenost segmenata	Preporučena	Preporučena, preporučena – 10, preporučena + 10
Dozvoljeno preklapanje elemenata grafikona	Ne	Da, ne