

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
VARAŽDIN

Filip Milohanović

IZRADA KOMPONENTE ZA DETEKCIJU
OZNAČENIH ODGOVORA NA PISMENIM
ISPITIMA

DIPLOMSKI RAD

Varaždin, 2025.

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Filip Milohanović

JMBAG: 0016148270

Studij: Informacijsko i programsко inženjerstvo

**IZRADA KOMPONENTE ZA DETEKCIJU OZNAČENIH ODGOVORA
NA PISMENIM ISPITIMA**

DIPLOMSKI RAD

Mentor :

Doc. dr. sc. Marko Mijač

Varaždin, lipanj 2025.

Filip Milohanović

Izjava o izvornosti

Izjavljujem da je moj diplomski rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor potvrdio prihvaćanjem odredbi u sustavu FOI-radovi

Sažetak

Opsega od 100 do 300 riječi. Sažetak upućuje na temu rada, ukratko se iznosi čime se rad bavi, teorijsko-metodološka polazišta, glavne teze i smjer rada te zaključci.

Ključne riječi: riječ; riječ; ...riječ; Obuhvaća 7+/-2 ključna pojma koji su glavni predmet rasprave u radu.

Sadržaj

1. Uvod	1
2. Metode i tehnike rada	2
3. Razrada teme	3
3.1. Digitalna slika	3
3.1.1. Slike u boji	3
3.1.2. Sive slike	4
3.1.3. Binarne slike	5
3.1.4. Važnost različitih tipova digitalnih slika	7
3.2. Digitalna obrada slike	8
3.2.1. Preprocesiranje slike	9
3.2.1.1. Pribavljanje slike	9
3.2.1.2. Poboljšanje slike	9
3.2.2. Procesiranje slike	12
3.2.2.1. Segmentacija slike	12
3.2.2.2. Morfološka obrada	13
3.2.2.3. Reprezentacija i opis	14
3.2.2.4. Detekcija i prepoznavanje objekata	15
3.2.3. Postprocesiranje slike	16
3.2.3.1. Obrada izvučenih podataka	16
3.2.3.2. Vizualizacija rezultata	17
3.2.3.3. Kompresija slike	17
3.3. Analiza postojećih rješenja i pristupa problemu	18
4. Praktični dio	19
4.1. Opis rješenja	19
4.2. Funkcionalni zahtjevi	20
4.3. Nefunkcionalni zahtjevi	21
4.4. Implementacija	22
4.5. Analiza rezultata	22
4.6. Limitacije rješenja	22
4.7. Manualno testiranje rješenja	22
5. Zaključak	23
Popis literature	25

Popis slika	26
Popis popis tablica	27

1. Uvod

U današnjem obrazovnom sustavu nastavnici provode jako puno vremena na radnim zadacima koji ne podižu kvalitetu obrazovanja već su administrativni i često repetitivni poslovi. Jedan od tih zadataka je ocjenjivanje ispita, pogotovo ako se radi o velikoj količini pristupnika testu. Kod testova s izrazito velikom količinom pristupnika se zato koriste ispitni koji uključuju list za odgovore što olakšava ocjenjivanje. Time se značajno ubrzava cijeli proces ocjenjivanja pošto nije potrebno čitati odgovore koji su u različitim formatima i potencijalno na više stranica, ali i dalje ostaje repetitivni dio posla koji uključuje usporedbu odgovora s rješenjima. Kroz ovaj rad bit će prikazana softverska komponenta otvorenog koda koja rješava ovaj problem tako da naspram slike detektira označene odgovore i ocjenjuje ispit.

Kroz ovaj rad istražuje se razne metode za detektiranje označenih odgovora s ciljem izrade komponente visoke preciznosti, prateći metodologiju znanstvenog oblikovanja.

2. Metode i tehnike rada

S obzirom na to da se radi o temi koja uključuje obradu slike, što zahtijeva opsežno istraživanja, manipulaciju parametrima i testiranje pristupa odlučeno je koristiti inkrementalan pristup razvoju. Prije same izrade rješenja, provedena je minimalno potrebna količina istraživanja, a zatim se ostatak istraživanja odvijao paralelno s razvojem.

Iterativni razvoj rješenja može se podijeliti u pretprocesiranje, izdvajanje relevantnih obilježja slike i na kraju ocjenjivanje testa. Jedna od najvažnijih tehnika tijekom razvoja bila je vizualizacija svakog koraka obrade slike. To je omogućilo izravan uvid u status i smjer razvoja. Od samog početka omogućen je uvid u mane i prednosti korištenih metoda. Ovo je značajno doprinijelo definiranju budućih razvojnih koraka, poput dodatnog istraživanja, promjene pristupa i sl.

Da bi sama vizualizacija bila što učinkovitija korišteno je više različitih ulaznih slika matrice odgovora slikanih u različitim uvjetima. To je omogućilo verifikaciju funkcionalnosti dijelova rješenja na različitim ulaznim podacima od samog početka razvoja, što je značilo da na kraju projekta nije bilo potrebno pokrivati rubne slučajeve jer je rješenje bilo dizajnirano da ih pokriva od početka.

Rješenje je izrađeno pomoću .NET tehnologije i C# programskog jezika u obliku biblioteke, što omogućuje drugim rješenjima jednostavnu integraciju koristeći .dll.

Kroz ovaj rad nije bio cilj razviti vlastitu biblioteku za računalni vid, korištena je jedna od postojećih biblioteka otvorenog koda. EmguCv je odabran pošto se radi o .NET omotaču za jednu od najpopularnijih biblioteka za računalni vid OpenCV. Osim toga, rješenje omogućuje korisnicima da koriste alternativnu biblioteku s uvjetom da izrade svoju implementaciju odgovarajućih metoda.

3. Razrada teme

Kroz teoretski dio rada predočit će se cijeli proces detekcije odgovora na pismenim ispitima koristeći obradu slike pomoću računalnog vida. Analizirat će postojeća rješenja ovog problema koja koriste računalni vidi i alternativne pristupe.

Svim ti pristupi imaju isti cilj, a to je izvlačenje relevantnih podataka iz slike, a zatim kreće obrada tih podataka. U kontekstu ovog rada to bi značilo da se iz slike moraju izvući podaci o označenim odgovorima za određeno pitanje te se ti podaci uspoređuju s listom točnih odgovora.

3.1. Digitalna slika

Za početak važno je razumjeti što su to digitalne slike i koje sve informacije one sadrže. Digitalna slika je zapravo računalna datoteka koja reprezentira fotografiju pomoću piksela. Pixel je zapravo najmanji element slike koji je reprezentiran s 3 kanala boja: Crvena (R), zelena (G), plava (B) u RGB modelu. Postoje i drugi kanali boja sli za potrebe ovog rada fokus će biti RGB modelu. [1].

Sama boja piksela se određuje kombinacijom vrijednosti iz RGB kanala. Dok je sam broj mogućih boja određen brojem bitova koju svaki RGB kanal može poprimiti. Taj koncept se još naziva dubinom boje (eng. color depth). Na primjer, 24-bitna slika može poprimiti do 16,777,216 unikatnih boja pošto svaki kanal ima 8 bitova [1]:

$$\text{Broj boja} = 2^8 \times 2^8 \times 2^8 = 2^{24} = 16,777,216$$

S pomoću piksela se također može definirati veličinu slike to jest rezolucija. Rezolucija je sveukupan broj piksela koje neka slika sadrži i izražava se množenjem širine slike s visinom slike npr. 1920×1080 . Obično veća rezolucija znači više detalja na slici. Što se više detalja nalazi na slici to su bolje šanse za izvući korisne informacije sa slike. No naravno da rezolucija sama ne određuje kvalitetu digitalne slike, već veliku ulogu igraju i fizički uvjeti u kojima je slika izrađena. Pod to spada osvjetljenje, čistoća kamere, kvaliteta senzora i sl.

3.1.1. Slike u boji

Većina digitalnih slika koje ljudi danas izrađuju su slike u boju te sadrže sva 3 kanala boje. Za razumijevanje ovog rada potrebno je razumjeti kakve informacije sadrži pojedini kanal boje. Svaki kanal predstavlja jednu primarnu boju poput crvene, zelene ili plave. I svaki pixel poprima određenu vrijednost za svaki kanal ovisno o kombinaciji bitova. Za 24-bitnu sliku svaki kanal ima 8 bitova to jest 256 mogućih vrijednosti.

To znači da u svakom kanalu postoji vrijednost od 0 do 255 koja definira jačinu svjetlosti za tu primarnu boju. Nula bi značila kompletno crna boja dok bi 255 bila potpuno osvjetljena primarna boja ovisno o kojem se kanalu radi. Kombinacijom tih 3 kanala se zapravo dobivaju

međuboje i time se stvara konačni izgled slike. [2].



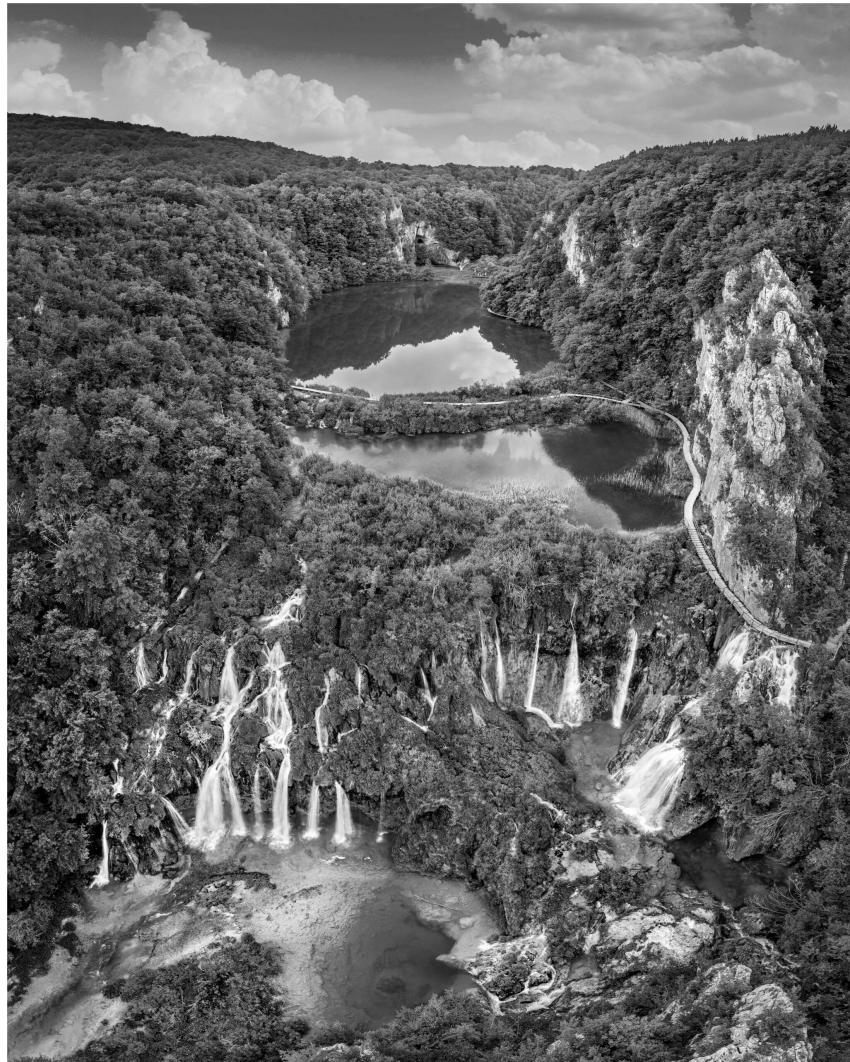
Slika 1: Prikaz slike u boji i pripadajućih kanala boja u RGB modelu (vlastita izrada)

Kao što je prikazano na slici, kanali boja prikazani su kao slike čiji pikseli poprimaju boje samo crnu, bijelu i nijanse sive boje, takozvane sive slike. Također kanali se ponekad znaju vizualizirati pomoću primarne boje kanala kojeg reprezentiraju, ali sami podaci koje digitalna slika sadrži su uvijek jednaki, jedino način njihove vizualizacije varira.

3.1.2. Sive slike

Sive slike (*eng. grayscale*) su poseban tip digitalnih slika koji ima samo jedan kanal boja. U tom kanalu vrijednosti piksela reprezentira svjetlinu piksela. Vizualizacija takvih slika poprima boje od bijele do crne uključujući nijanse sive. Na primjer za sliku sa dubinom boja od 8 bitova, nula bi predstavljala crnu boju, 255 bijelu, a vrijednosti između nijanse sive [2].

Sive slike su jedan od važnijih koncepata obrade slike i računalnog vida. One su zapravo baza mnogih algoritama za procesiranje slika. Često se koriste kao prvi korak obrade. Izrazito su korisne za detekciju rubova, segmentaciju slike, prepoznavanje uzorak i slično [3]. Kroz rad će se detaljnije prikazati njihova važnost i uloga u kontekstu detektiranja označenih odgovora na pismenim ispitima.



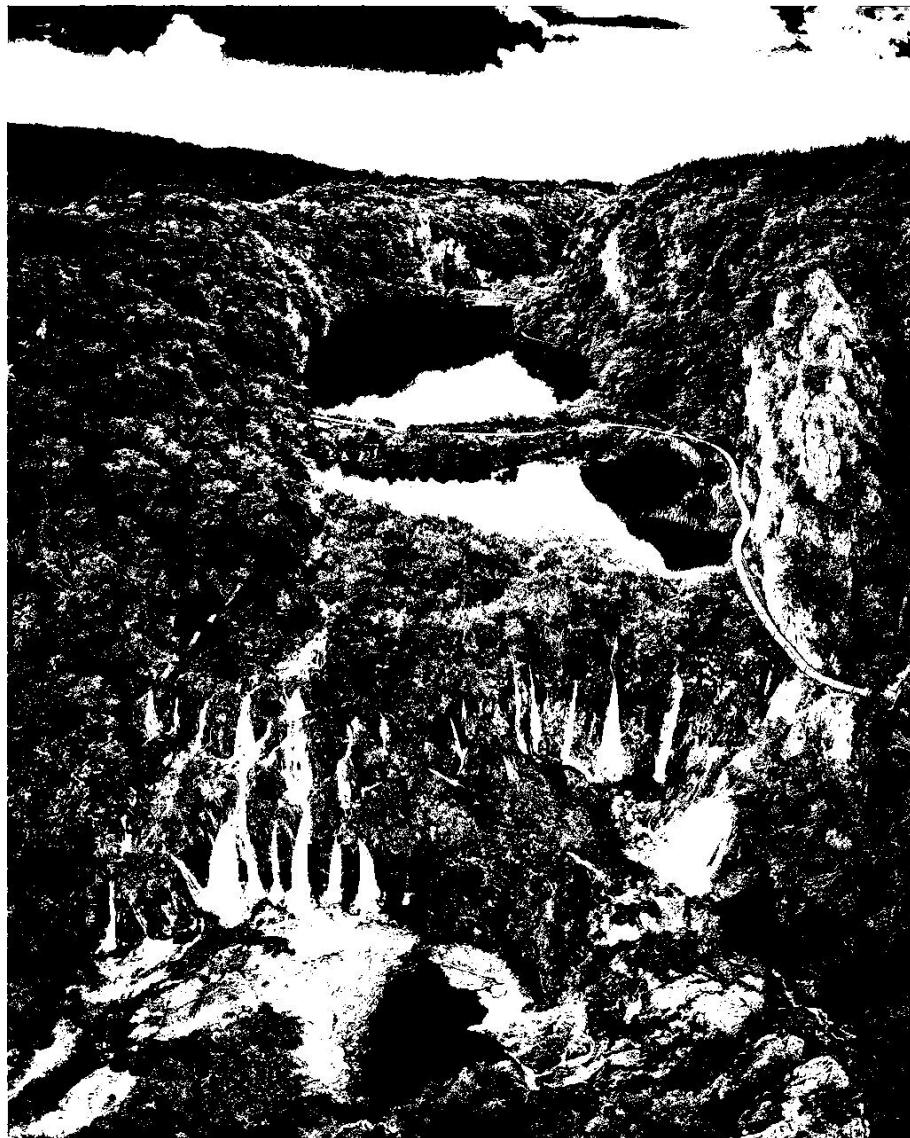
Slika 2: Prikaz sive slike (vlastita izrada)

Na priloženoj sivoj slici se također može vidjeti da je različita od sivih slika kanala prikazanih na slici 1. To je zato što se kod pretvorbe slike u boji u sivu sliku uzimaju u obzir vrijednosti piksela za sva 3 kanala prema formuli:

$$I_{\text{gray}}(x, y) = 0.299 \cdot R(x, y) + 0.587 \cdot G(x, y) + 0.114 \cdot B(x, y) [3]$$

3.1.3. Binarne slike

Osim slika u boji i sivih slika također postoje i binarne slike. Binarne slike su tip slika gdje piksel može imati samo dvije moguće vrijednosti 0 i 1 od kud im dolazi i naziv. Binarne slike sadrže puno manje informacija od ostalih tipova slika, ali isto tako smanjuju kompleksnost same slike i omogućuju fokusiranje na važne značajke slike. Baš zbog toga se jako često koriste kod obrada slika koristeći računalni vid [4].



Slika 3: Prikaz binarne slike (vlastita izrada)

Sam proces izrade binarne slike je dosta jednostavan. Za početak potrebno je imati sivu sliku. Zatim se definira granica (*eng. threshold*) u obliku postotka. Ta granica se zatim koristi za određivanje vrijednosti piksela, ako je vrijednost iznad granice onda se vrijednost postavlja na jedan, a inače na nulu. Priložena formula prikazuje osnovni tip kreiranja binarne slike (*eng. thresholding*).

$$I_{\text{gray}}(x, y) = 0.299 \cdot R(x, y) + 0.587 \cdot G(x, y) + 0.114 \cdot B(x, y)$$

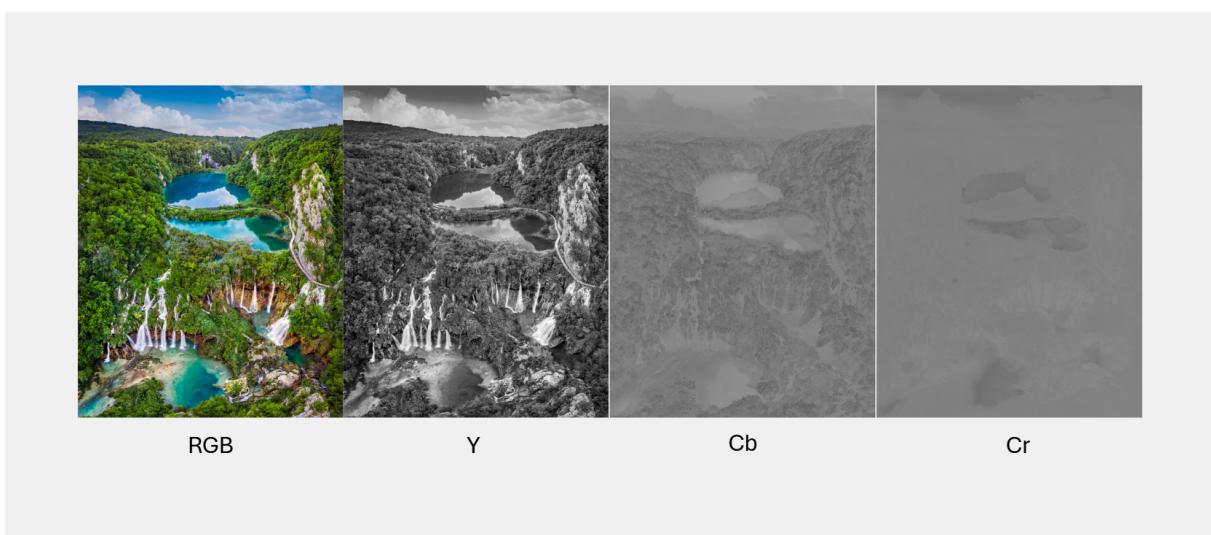
$$I_{\text{binary}}(x, y) = \begin{cases} 1, & \text{if } I_{\text{gray}}(x, y) \geq T \\ 0, & \text{if } I_{\text{gray}}(x, y) < T \end{cases}$$

3.1.4. Važnost različitih tipova digitalnih slika

Do sada smo prošli kroz 3 najvažnija tipa digitalnih slika za ovaj rad. Kroz ovu cjelinu prikazat će se važnost pojedinog tipa slike naspram ostalih.

Zapravo najveća razlika između ovih tipova je broj kanala boje. Slike u boji imaju 3 kanala za razliku od sivih i binarnih slika koje imaju po jedan kanal. To zapravo znači da slike u boji sadrže značajno više informacija.

No ponekad nam te informacije nisu potrebne ovisno o problemu koji pokušavamo riješiti. Na primjer ako pokušavamo detektirati boju svjetla na semaforu onda su nam informacije o bojama izrazito važne. No ako pokušavamo pročitati tekst s prometnog znaka onda nam informacija o boji znaka i teksta ne doprinosi značajno rješavanju problema. Zapravo nam otežava rješavanje problema pošto moramo raditi s nepotrebnim informacijama.



Slika 4: Usporedba svjetlosne komponente i komponenta boja(vlastita izrada)

Na priloženoj slici prikazana je ista slika na različite načine:

- Slika u RGB formatu
- Slika u YCbCr formatu
 - Svjetlosne informacije u Y kanalu
 - Informacije o bojama u Cb i Cr kanalima

Na ovom primjeru je prikazana činjenica da svjetlosni kanali sadrže puno više informacija o struktturnim elementima na slici. baš zbog toga se u raznim algoritmima za detektiranje značajki, filtriranje, segmentiranje i sl. koriste baš sive slike [5].

Sive slike se koriste kada boja nije važan faktor u obradi slike. U tom slučaju je boja jednako korisna kao buka na slici pa je zbog toga uklonimo. Time se veličina slike smanje i time ubrza daljnji proces obrade slike [5].

Ako je potreban još agresivniji pristup otklanjanju nepotrebnih podataka onda binarne slike mogu postati korisne. One dodatno odstranjuju nepotrebne podatke tako da povećavaju razliku u kontrastu [5].

Važno je napomenuti da svaki tip slika ima svoju upotrebu ovisno o problemu koji se rješava. U ovom radu će veliku važnost imati sive i binarne slike pošto se radi o procesu detektiranja strukture fotografije i izvlačenja relevantnih informacija iz nje.

3.2. Digitalna obrada slike

Digitalna obrada slike (*eng. digital image processing*) podrazumijeva obradu digitalne fotografije koristeći računalne algoritme s ciljem unaprjeđivanja slike, izvlačenja korisnih informacija i sl. Predstavlja ključni korak u raznim primjenama računalnog vida temeljenog na dubokom učenju, poput prepoznavanja lica, detekcije objekata i kompresije slika [6].

Sam proces digitalne obrade slike može se podijeliti u sljedeće faze [6]:

- Pribavljanje slike (*eng. image acquisition*)
- Poboljšanje slike (*eng. image enhancement*)
- Kompresija slike (*eng. image compression*)
- Morfološka obrada (*eng. morphological processing*)
- Segmentacija slike (*eng. image segmentation*)
- Reprezentacija i opis (*eng. representation and description*)
- Detekcija i prepoznavanje objekata (*eng. object detection and recognition*)

Naravno nisu sve faze uvijek potrebne, već se radi o generalnim smjernicama koje se prema specifičnostima problema mogu modificirati, smanjiti ili proširiti. U nekim slučajevima čak se redoslijed faza može promijeniti ili se faze mogu provoditi više puta.

Te faze se dodatno podijeliti na **preprocesiranje, procesiranje i postprocesiranje slike**. U fazu **preprocesiranja** spadaju svi koraci koji pripremaju sliku za procesiranje, poput poboljšanja slike. Nakon preprocesiranja slijedi **procesiranje** koje se odnosi na korištenje algoritama za detekciju i prepoznavanje objekata i transformacije. Posljednja faza uključuje korake poput pripreme rezultata za prezentaciju, vizualizaciju i daljnju analizu [7].

3.2.1. Preprocesiranje slike

Preprocesiranje slike je ključan proces koji se provodi prije obrade slike. Kao ulaz prima izvornu, sirovu sliku, a kao izlaz vraća sliku u korisnjem formatu za daljnju obradu i analizu. Omogućuje da se sa slike odstrane ili istaknu područja relevantnih informacija te da se poboljša kvaliteta slike prije daljnje obrade [8].

3.2.1.1. Pribavljanje slike

Pribavljanje slike je proces nastajanja digitalne slike tako da se informacije iz realnog svijeta zapiše u digitalnom obliku kojim računalo zatim može manipulirati. Slika se može pribaviti raznim uređajima poput fotoaparata, pametnog mobitela ali i drugih tehnologija poput skenera, rendgена i drugih uređaja. U ovom radu fokus će biti na fotografije probavljene uređajima koji posjeduju digitalnu kameru [9].

Da bi bilo kakav procesiranje moglo krenuti prvo je potrebno pribaviti sliku. Moglo bi se reći da je pribavljanje slike čak najvažniji korak u cijelom procesu obrade slike, njegov rezultat ima veliku ulogu za postizanje krajnjeg rezultata obrade i analize slike. Pribavljena slika mora biti kvalitetna da bi ostali koraci bili uspješni. Na kvalitetu slike ne utječe samo rezolucija, osvjetljenje i kut pod kojim je slika uslikana također imaju značajnu ulogu [9].

Razni algoritmi i koraci u preprocesiranju mogu korigirati pribavljenu sliku, no to nikad nije zamjena za sliku dobre kvalitete. Dobra i loša slika mogu biti razlika između uspješnog i neuspješnog detektiranja elemenata na slici. Naravno u većini slučajeva sustavi za obradu slike ne mogu utjecati na kvalitetu slike koje će im korisnik poslati. Zbog toga je važno detaljno testirati sustav u različitim uvjetima i utvrditi što su njegove granice i je li potrebno podržati neki dodatni korisnički slučaj od onih postojećih.

3.2.1.2. Poboljšanje slike

Poboljšanje slike je ključan korak, cilj mu je modificirati sliku tako da ona postane korisnija za daljnju obradu. To se može postići na više načina poput otklanjanja buke, isticanja važnih značajki slike, poboljšanja oštchine i kontrasta slike.

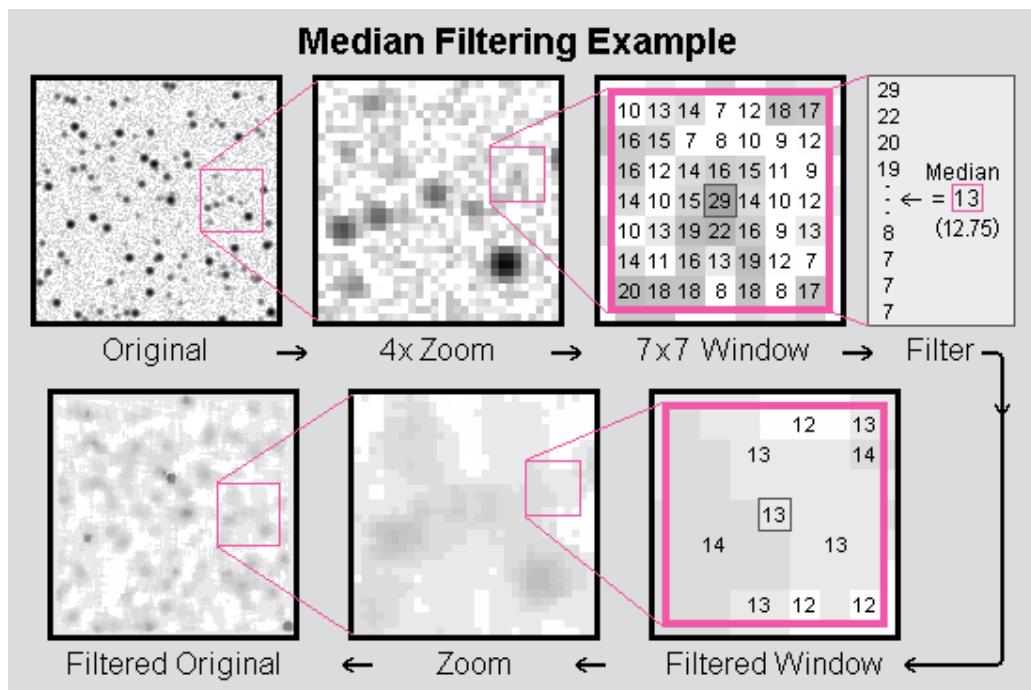
Do sada smo spomenuli da se nepotrebne informacije na slici također mogu biti buka. Baš zbog toga pretvaranje u sivu i binarnu sliku spadaju pod poboljšanje slike. Oba dvije operacije uklanjuju nepotrebne informacije o bojama sa slike što omogućuje bržu i jednostavniju obradu i analizu slike. naravno to ima smisla samo ako nam informacije o bojama nisu potrebne.

Osim nepotrebnih boja, bukom se može smatrati i nasumična razlika u boji između susjednih piksela. U digitalnoj fotografiji buka je praktički neizbjegna i može se definirati kao nasumična varijacija signala u slici. Ta varijacija je reprezentirana različitim nijansama boja susjednih piksela. Buka se može smanjiti manipulacijom parametara digitalne kamere, ali ne može se kompletno ukloniti [10].

Takve informacije se zbog malih razlika u boji piksela većinom mogu smatrati nevažnim, a loše utječu na daljnju obradu slike. Zbog toga postoje razne tehnike za filtriranje slike. Najpopularnije tehnike za filtriranje buke su [11]:

- Gaussov filter (*eng. gaussian filter*)
- Srednji filter (*eng. mean filter*)
- Medijanski filter (*eng. median filter*)
- Bilateralni filter (*eng. bilateral filter*)

Svaki tip filtera ima svoju upotrebu i koristan je za određeni tip buke. Buka u digitalnim slikama je jako opširna tima koje se neće detaljnije objašnjavati u ovom radu.



Slika 5: Primjer medijanskog filtera [12]

Na priloženoj slici je prikazan proces korištenja medijanskog filtera koji naspram matrice susjednih vrijednosti određuje vrijednost piksela koristeći medijan svih vrijednosti.

Važno je napomenuti da su pojmovi poput oštine i mutnoće slike usko povezani s bukom. Jako često mutna slika sadrži manje buke, dok oštra slika može sadržavati više buke. Zbog toga je potrebno naći ravnotežu pri uklanjanju buke. Ako uklonimo previše buke, slika postaje mutna i gubi se previše informacija, obrnuto vrijedi ako uklonimo pre мало buke.

Postoji još razne metode poboljšavanja slike poput povećavanja kontrasta, ali neće biti potrebne u ovom radu. Također u nekim literaturama se spominju tehnike poput smanjenja buke i raznih drugih transformacija u kontekstu restauracije slike. Ali u ovom radi će se o tome govoriti u kontekstu poboljšanja slike.

Druge važne metode poboljšavanja slike su razne geometrijske transformacije slike poput translacije, rotacije i skaliranja. No to sve te osnovne geometrijske operacije se koriste kod ispravljanja perspektive. Baš zbog toga je ispravak perspektive puno zanimljivija operacija za ovaj rad [13].

Ispravljanje perspektive omogućuje da slika prikazuje na načina to da je slikana u nekom drugoj ravnini u prostoru. To omogućuje da se uklone razlike između slika poput kuta kamere naspram subjekta fotografije, u ovom slučaju papira s matricom odgovora.

Korekcija perspektive mapira točku $\mathbf{p} = [x, y, 1]^T$ u originalnoj slici (slikanoj iz nekog neželjenog kuta) $\mathbf{p}' = [x', y', w']^T$ u sliku ispravljene perspektive s pomoću matrice H [14].

$$\mathbf{p}' = H \cdot \mathbf{p}$$

Gdje H predstavlja 3×3 homografsku matricu:

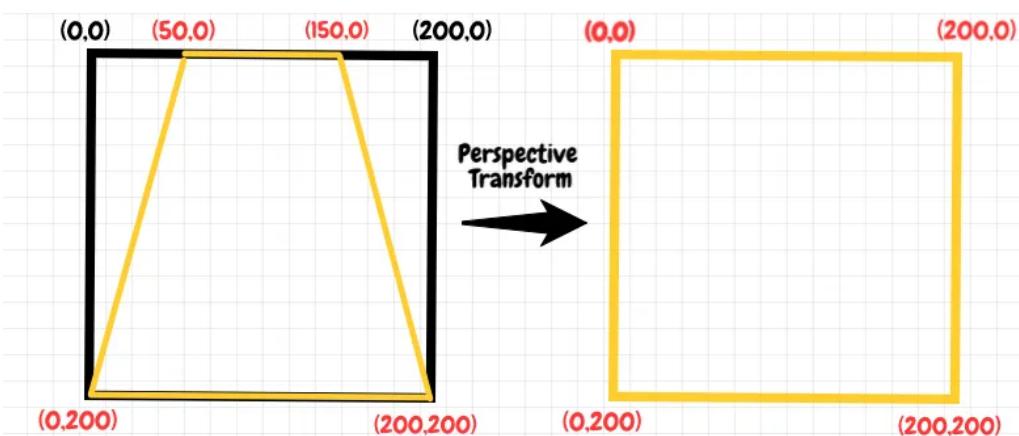
$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

Transformirane koordinate se računaju na sljedeći način:

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Da bi se koordinate bile korisne potrebno ih je normalizirati s w :

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$



Slika 6: Prikaz procesa ispravka perspektive subjekta na slici [14]

Na priloženoj slici je prikazano kako se matrica transformacije koristi da bi se pers-

pektiva ispravila. U praksi to funkcioniра tako da se matrica transformacije računa naspram 4 izvorne i 4 odredišne točke. Te se zatim prema formuli množi sa svakom izvornom točkom da bi se izračunale odredišne točke.

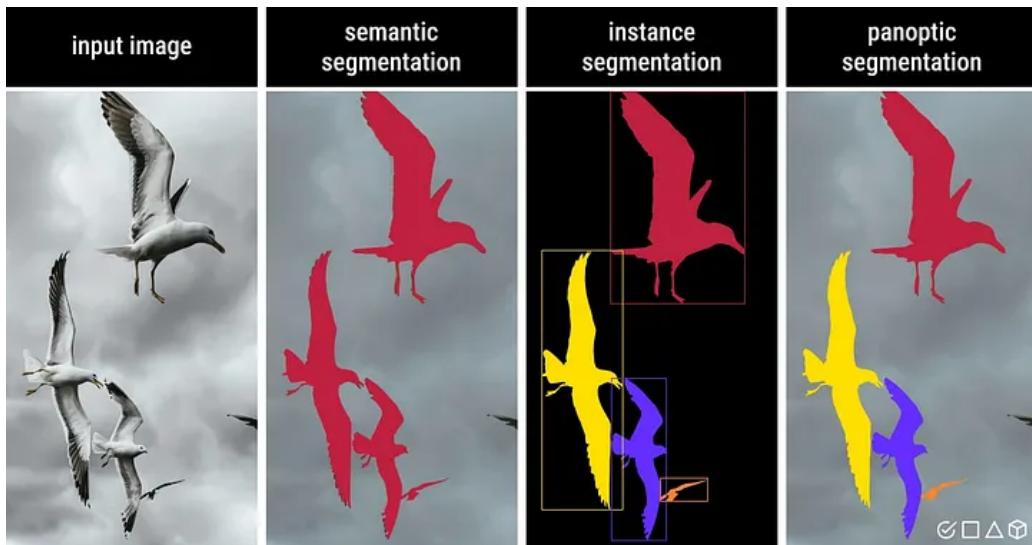
3.2.2. Procesiranje slike

Nakon što je faza preprocesiranja pripremila pretvorila sliku u korisniji format, kreće glavna faza obrade to jest procesiranje slike. Glavne tehnike kod procesiranja slike su segmentacija, detekcija objekata, morfološka obrada te reprezentacija i opis.

3.2.2.1. Segmentacija slike

Segmentacija slike je proces grupiranja piksela prema nekim obilježjima. Za razliku od običnih klasifikatora poput neuronskim mrežama, segmentacija nudi informacije gdje se točno na slici nalazi objekt od interesa. Isto kao u kod klasifikatora potrebno je definirati moguće klase kojima piksel može pripadati. Definiranje tih klasa se može učiniti na 3 načina prema kojima dijelimo tipove segmentacije [15]:

- semantička segmentacija (*eng. semantic segmentation*)
- segmentacija instance (*eng. instance segmentation*)
- panonička segmentacija (*eng. panoptic segmentation*)



Slika 7: Usporedba tipova segmentacije [15]

Za priloženu sliku su definirane dvije klase nebo i ptice. Semantička slika je sve piksela ptica klasificirala kao jednu klasu, što znači da ne znamo točno o kojoj ptici pripada piksel. Da bi to riješili postoji segmentacija prema instanci što odvaja svaku pticu u zasebnu klasu što znači da više ne znamo koji sve pikseli pripadaju roditeljskoj klasi ptica. Da bi to riješili postoji

panonička segmentacija koja kombinira ta dva pristupa i pruža informacije o klasi piksela i kojoj instanci pripada. Tip segmentacije ovisi o specifičnostima problema [15].

3.2.2.2. Morfološka obrada

Morfološka obrada je skup operacija koji procesira sliku naspram oblika. Omogućuje pronalazak elemenata na slici sličnih definiranom obliku. A usporedbu oblika vrši tako da traži piksele koji odgovaraju tom obliku [16].

Morfološka obrada se u nekim segmentima preklapa sa segmentacijom slike, ali ne radi se o istim operacijama. Morfološke operacije se često koriste nakon segmentacije da uklone nedostatke iz detektiranih segmenata [16].

Dvije najvažnije morfološke operacije su erozija i dilatacija. Erozija funkcioniра na način da uklanja piksele s ruba objekta. Vrlo je korisna u slučajevima kada je potrebno odvojiti dva objekta određenog oblika [16].



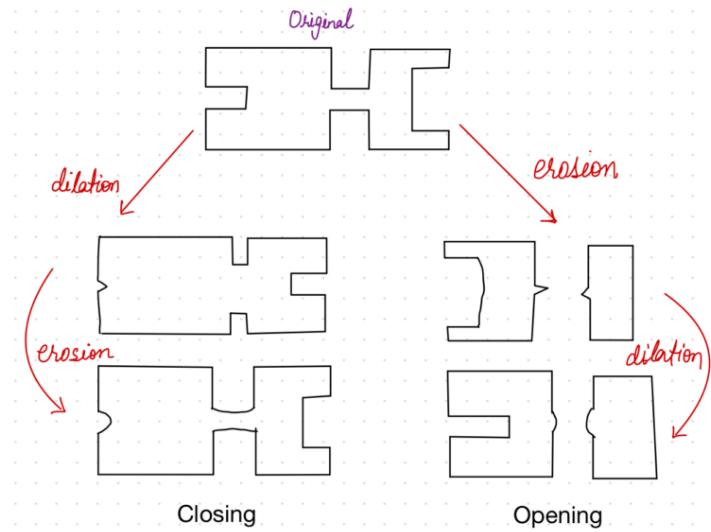
Slika 8: Prikaz erozije kao morfološke operacije [16]

Za razliku od erozije, dilatacija dodaje piksele na oko ruba objekta. Time omogućuje da se popune praznine između objekata.



Slika 9: Prikaz dilatacije kao morfološke operacije [16]

Najčešće se te dvije operacije ne koriste zasebno već zajedno kao kompozitna morfološka operacija. Ako prvo provedemo eroziju zatim dilataciju radi se o otvaranju. Dok obrnuti redoslijed operacija čini zatvaranje [16].



Slika 10: Prikaz kompozitnih morfoloških operacija [16]

3.2.2.3. Reprezentacija i opis

Nakon što je obavljena segmentacija i po potrebi morfološka obrada većinom slijedi reprezentacija i opis segmentiranih dijelova. Cilj reprezentacije i opisa je reprezentirati piksele tako da budu korisniji za daljnju obradu [17].

Dva glavna tipa reprezentacije segmentiranih objekata, ovisno o tome je li fokus na [17]:

- vanjskim karakteristikama, (kontura ili rub objekta)
- unutarnjim karakteristikama (pikseli koji čine objekt)

Reprezentacija prema vanjskim karakteristikama se često koristi kad nam je za obradu važan sam oblik, dok se unutarnje karakteristike koriste kada je potrebna informacija o boji i teksturi objekta. Iz dobivenih reprezentacija je moguće izvući razne opisne vrijednosti (*eng. descriptors*) koje su korisne kod daljne obrade. Važno je napomenuti da je generalno pravilo da bi opisne vrijednosti trebale biti imune na promjenu skaliranja, translacije i rotacije [17].

Neke od korisnik opisnih vrijednosti su [17]:

- Dužina konture objekta
- Promjer konture objekta
- Zaobljenost
- Okvirna kutija
- Površina objekta

- Sličnost željenom obliku
- Odstupanje od simetrije
- Koeficijent istezanja

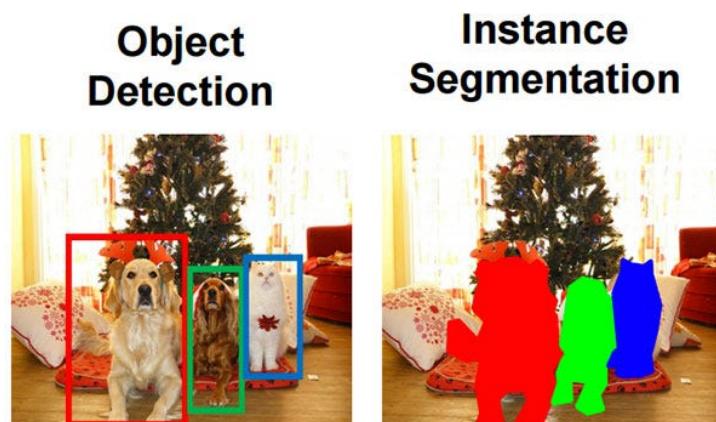
3.2.2.4. Detekcija i prepoznavanje objekata

Detekcija i prepoznavanje objekata je jedna od ključnih tehnika procesiranja slika na tradicionalan način ali i s pomoću neuronskih mreža. Bazirano je na principu pronađaska objekta i crtanjem okvirne kutije oko tog objekta [18], [19].

To mu je ujedno i glavna razlika naspram segmentacije slike, segmentacija dodjeljuje klasu svakom pikselu koji odgovara onome što se traži, dok detekcija objekta to ne radi već samo locira objekt određene klase na slici [18], [19].

Ponekad je samo informacija o lokaciji objekta dovoljna, u slučaju da nije segmentacija nudi puno više informacija o samom objektu.

Da bi detektiranje objekata bilo uspješno često se koriste konture. Konture su ništa drugo nego krivulje koje opisuju rubove objekata, zbog toga su jako korisne kod detektiranja oblika. Također se ponekad koriste i kod segmentiranja.



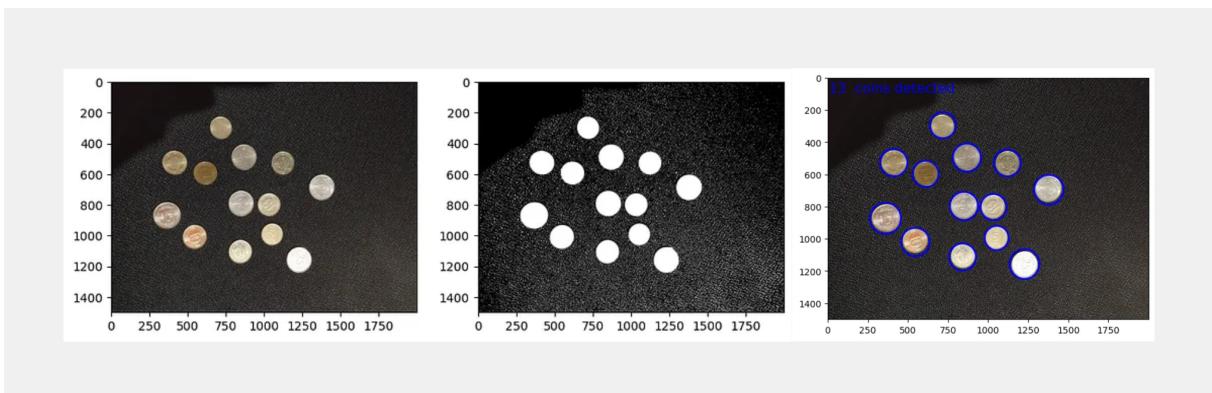
Slika 11: Usporedba segmentacije i detekcije objekta [19]

Da bi se moglo pronaći konture prvo je potrebno sliku pretvoriti u binarnu sliku i zatim provesti operaciju detektiranja kontura. Konture se detektiraju tako da se traži razlika između vrijednosti piksela na binarnoj slici i time se utvrđuje ujedno i granica između objekata. Zatim se primjenjuje jedna od metoda za detekciju objekta nad tim konturama. Te metode zapravo traže konture koje odgovaraju traženom obliku s definiranim tolerancijama za oblik, veličinu i slično [20].

Pošto će se u ovom radu koristiti EmguCv nama su najzanimljivije metode poput:

- **CvInvoke.ApproxPolyDP()** - Aproximacija kontura u jednostavnije poligone (korisno za prepoznavanje oblika).

- **CvInvoke.MinAreaRect()** - Pronalaženje minimalnog pravokutnika koji okružuje konturu (korisno za detekciju objekata u obliku pravokutnika).
- **CvInvoke.FitEllipse()** - Fitting elipse na konturu (korisno za prepoznavanje eliptičnih objekata).
- **CvInvoke.HoughLinesP()** - Houghova transformacija za detekciju linija.
- **CvInvoke.HoughCircles()** - Houghova transformacija za detekciju krugova.
- **CvInvoke.BoundingRectangle()** - Pronalaženje okvira (bounding box) koji okružuje konture objekta.
- **CvInvoke.FindContoursHierarchical()** - Detekcija kontura s informacijama o hijerarhiji (za detekciju unutarnjih i vanjskih objekata).



Slika 12: Prikaz rezultata detekcije krugova pomoću kontura i CvInvoke.MinEnclosingCircle metode [19]

Na priloženoj slici je jasno prikazano kako se izvorna slika pretvara u binarnu sliku i zatim se s pomoću kontura i metode za detektiranje oblika naspram kontura pronalazi lokacija kruga.

3.2.3. Postprocesiranje slike

Nakon što je glavna faza obrade slike zavšrena kreću aktivnosti postprocesiranja. Prijašnja faza je izvukla korisne podatke iz slike i pripremila ih u korisnom formatu. Cilj postprocesiranje je obrada tih podataka na način da se dođe do konačnog rješenja. Osim obrade podataka, postprocesiranje je također povezano sa aktivnostima poput vizualizacije rezultata, kompresije i slično.

3.2.3.1. Obrada izvučenih podataka

Prije nego što se kreće u konačnu obradu izvučenih podataka potrebno ih je pročistiti. Načini pročišćavanja podataka ovise o specifičnostima problema, ali neke od važnih metoda za ovaj rad su pročišćavanja su:

- Filtriranje po veličini objekta
- Filtriranje po obliku objekta
- Filtriranje po boji objekta
- Filtriranje po hijerahijskoj odnusu objekta s ostalim objektima

Spomenute metode bazirane su na usporebi opisnih vrijednosti iz faze reprezentacije i opisa. Nakon što su podaci pročišćeni kreće konačna faza obrade. Način konačne obrade ovisi o specifičnostima problema, ali često se koriste baze znanje.

Baze znanja omogućuju usporedbu traženih podataka sa pročišćenim podatakom koje omogućuje rješenje specifičnog problema. U ovom slučaju baza znajnja bi bila lista odgovora za svako pitanje podijeljenih u dvije klase koje predstavljaju točne i netočne odgovore [6].

3.2.3.2. Vizualizacija rezultata

Naokon što su podaci obrađeni i utvrđen je rezultat kako često ga je potrebno vizualizirati. Vizualizacija nudi uvid u njihovu važnost i omogućuje da se jednostavnije analiziraju i izvode zaključi iz rezultata. Tip vizualizacije ovisi o doemni problema koji se rješava, može biti u raznim oblicima poput pdf izvješća, vizualizacije rezultata nad originalnom slikom i slično [21].

Kod vizualiziranja rezultata važno je uzeti u obzir tip podataka koji se vizualizira, publiku za koju je vizualizacija namjenjena [21].

3.2.3.3. Kompresija slike

U slučaju potrebe za smanjenjem veličine slike, često se koristi kompresija. Kompresija nastoji smanjiti veličinu slike na način da pritom zadrži kvalitetu na zadovoljavajućoj razini. Često se primjenjuje kada je potrebno smanjiti veličinu slike pri prijenosu preko mreže, spremanju u memoriju ili za ubrzanje obrade [6].

Kompresija slike dijeli se na dva glavna tipa [22]:

- Kompresija sa gubitkom podataka (*eng. lossy*)
- Kompresija bez gubitka podataka (*eng. lossless*)

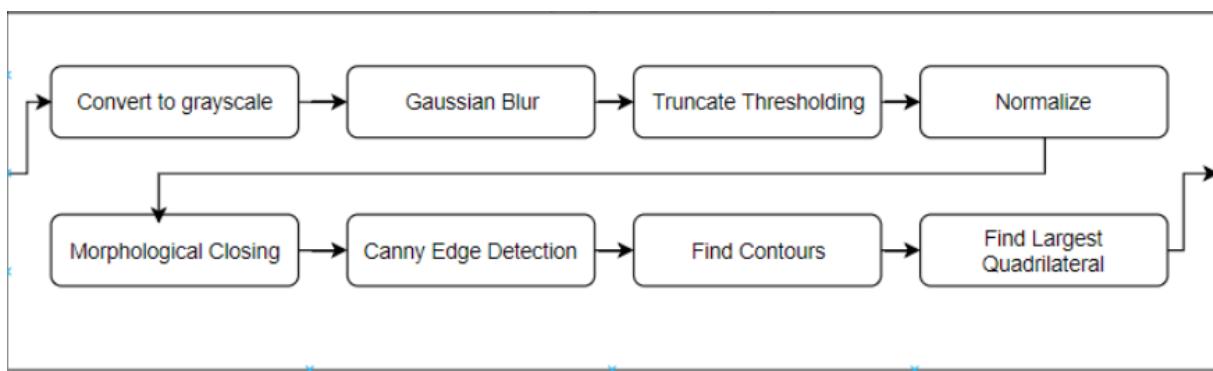
Oba tipa smanjuju veličinu slike, ali se razlikuju u načinu na koji to postižu. Prvi pristup uklanja određene informacije iz slike kako bi smanjio njezinu veličinu (*lossy*), dok drugi pristup efikasnije zapisuje podatke bez gubitka informacija (*lossless*). Važno je napomenuti da je kod kompresije uvijek potrebno balansirati veličinu slike i kvalitetu. Što je slika više komprimirana, to je izraženija distorzija slike. Glavna razlika između ovih dvaju pristupa je u tome može li se originalna slika potpuno rekonstruirati iz komprimirane slike [22].

3.3. Analiza postojećih rješenja i pristupa problemu

Prije praktičnog dijela rada važno je analizirati posotjeća rjepenja na tržištu. Ova analiza fokusirat će se samo na detekciju odgovora na pismenim ispitima koristeći matricu odgovora to jest OMR tehnikom (*eng. optical mark recognition*). Pod OMR spadaju sve tehnike koje su do sada obrađene. Fokus analize je samo na rješenja visoke razine točnosti.

Osim korištenja računalnog vida uz OMR sve češće su pristupi koji koriste elemente dubokog učenja samostalno ili u kombinaciji sa tradicionalnim računalnim vidom [23].

Analizirana su rješenja OMRChecker i OpenOMR, radi so rješenjima izrazito visoke točnosti koji tu točnost zadržavaju i kada je slika uslikana u izrazito izazovnim uvjetima. Također oba rješenja podržavaju određeni stupanj konfiguracije liste sa odgovorima [23].



Slika 13: Prikaz dijela toka procesiranja slike kod OMRChecker-a [24]

Sa priložene slike vidljivo je da OMRChecker koristi veliku većinu metoda i tehnika koje su do sada obrađene u ovom radu, za OpenOMR je situacija također dosta slična. Kroz praktični dio rada prikazat će se koji je cilj rješenja izrađenog u skolpu ovog rada.

4. Praktični dio

Kroz daljni dio rada fokus će biti na izradi i prikazu komponente otvorenog koda za detekciju odgovora na pismenim testovima. Prikazati će se svi važni koraci i dizajnerske odluke koje utječe na konačni dizajn i funkcionalnost rješenja. Izrađena komponenta će također biti ugrađena u jednostavnu stolnu aplikaciju da bi se olakšalo razvoj i vizualizirao sam rad komponente.

4.1. Opis rješenja

Kao što je već spomenuto cilj je izraditi komponentu otvorenog koda za detekciju odgovora na pismenim ispitima. Fokus će biti na detektiranju odgovora na matrici odgovora. Prednost toga pristupa je to što nije potrebno slikati više stranica da bi se ocijenio jedan ispit, već je dovoljno ocijeniti samo matricu odgovora to jest obraditi samo jednu stranicu.

Detekcija odgovora na matrici odgovora je puno jednostavniji problem nego detekcija na standardnim ispitima. Zbog toga što matrica može imati unaprijed defininaru i standardiziranu strukturu to jest predložak. Unatoč tome matrica limitira rješenje samo na pitanja na višestruki ili jednostruki odgovor. No to nije problem za ovaj rad pošto drugi tipovi pitanja nisu u opsegu ovog rada.

Jedan od glavnih ciljeva ovog rješenja je da kozistentno pruža pozdane rezultate, čak i ako uvijeti u kojima je slika uslika nisu idealni. Ta pouzdanost se treba postići za predloške koji mogu sadržavati varijabilan broj pitanja i odgovora. Da bi izrada predložaka bila jednostavnija rješenje bi također trebalo biti posebno prema određenim parametrima generirati predložak matrice odgovora u obliku PDF-a.

Također da bi se ubrzao rad sa aplikacijom ideja je da korsnik ne unosi ručno podatke o točnim odgovorima za sva pitanja, već jednostavno slika ispravno rješen ispit te da komponente izvuće te informacije sama iz fotografije. Da bi se dodatno ubrzao proces ocjenjivanja sustav treba biti dizajniran na način da detektirane odgovore usporedi sa ispravnim odgovorima i ujedno prema korisničko definiranoj skali ocijeni ispit. Ocjenjeni ispit mora biti vizualiziran da ocijenjivač ipak može uvidjeti je li došlo do greške kod ocjenjivanja.

Za svrhe razvoja, pronaska grešaka i slično sustav bi trebao imati mogućnost vizualiziranja svakog koraka obrade od samog unosa slike pa sve do rješenja. Ta opcija bi se trebala moći uključiti ili isključiti po potrebi.

Pošto se radi o komponenti otvorenog koda želja je da se ostavi mogućnost zamjene važnih biblioteka koje će sustav koristiti, pšoput one za računalni vid. Time će se omogućiti da neko po potrebi napiše svoju implementaciju i zamjeni zadalu EmguCv biblioteku.

4.2. Funkcionalni zahtjevi

Identifikator	FZ-01
Zahtjev	Sustav mora detektirati odgovore s matrice odgovora.
Obrazloženje	Detekcija se temelji isključivo na obradi jedne stranice – matrice odgovora – čime se pojednostavljuje i ubrzava ocjenjivanje.
Način provjere	Prilikom unosa slike matrice, sustav treba pravilno označiti sve odgovore koji su označeni od strane korisnika.
Prioritet [1–5]	
Identifikator	FZ-02
Zahtjev	Sustav mora omogućiti generiranje PDF predložaka za matricu odgovora.
Obrazloženje	Korisnici trebaju moći jednostavno izraditi matricu s unaprijed definiranim brojem pitanja i odgovora.
Način provjere	Iz sustava treba biti moguće izvesti generirani predložak u PDF formatu.
Prioritet [1–5]	
Identifikator	FZ-03
Zahtjev	Sustav mora omogućiti automatsko prepoznavanje točnih odgovora iz slike ispravno riješenog ispita.
Obrazloženje	Korisnik ne mora ručno unositi odgovore; slika ispravnog rješenja bit će dovoljna za generiranje ključa.
Način provjere	Učitavanjem slike ispravno riješenog ispita sustav mora automatski prepoznati sve točne odgovore.
Prioritet [1–5]	
Identifikator	FZ-04
Zahtjev	Sustav mora usporediti odgovore s točnim rješenjem i izračunati ocjenu prema definiranoj skali.
Obrazloženje	Automatsko ocjenjivanje ubrzava proces ispravljanja i smanjuje mogućnost pogreške.
Način provjere	Nakon što se ispiti obrade, mora biti prikazana ocjena za svaki ispit u skladu sa skalom.
Prioritet [1–5]	

Identifikator	FZ-05
Zahtjev	Sustav mora omogućiti vizualizaciju rezultata ocjenjivanja za svakog korisnika.
Obrazloženje	Omogućava ljudsku kontrolu rezultata i ispravnost prepoznavanja i ocjenjivanja.
Način provjere	Vizualna prezentacija ispita s označenim točnim i netočnim odgovorima mora biti dostupna.
Prioritet [1–5]	
Identifikator	FZ-06
Zahtjev	Sustav mora omogućiti prikaz svih koraka obrade slike.
Obrazloženje	Korisnicima i developerima je korisno za praćenje grešaka i ponašanja sustava.
Način provjere	U postavkama sustava mora postojati opcija kojom se uključuje ili isključuje prikaz koraka obrade.
Prioritet [1–5]	
Identifikator	FZ-07
Zahtjev	Sustav mora omogućiti unos varijabilnog broja pitanja i odgovora u matricu.
Obrazloženje	Prilagodljivost broju pitanja i odgovora ključna je za podršku različitim tipovima ispita.
Način provjere	Tijekom generiranja predloška korisnik može definirati broj pitanja i odgovora, koji se zatim reflektira u matrici.
Prioritet [1–5]	

Tablica 1: Funkcionalni zahtjevi

4.3. Nefunkcionalni zahtjevi

Identifikator	NFZ-01
Zahtjev	Sustav mora pružati pouzdane rezultate čak i u slučaju loših uvjeta pri snimanju (npr. sjenke, loš kontrast, zakrenutost).
Obrazloženje	U praksi slike ispita neće uvijek biti idealne kvalitete, pa sustav mora biti robustan na takve uvjete.
Način provjere	Sustav se testira na uzorku slika s različitim uvjetima osvjetljenja, orijentacije i kvalitete.
Prioritet [1–5]	
Identifikator	NFZ-02
Zahtjev	Sustav mora omogućiti jednostavnu zamjenu vanjskih biblioteka poput one za računalni vid.

Obrazloženje	Komponenta je otvorenog koda, pa korisnici mogu koristiti druge biblioteke (npr. zamijeniti EmguCV).
Način provjere	Biblioteke se mogu zamijeniti putem konfiguracijskih postavki bez mijenjanja osnovnog koda.
Prioritet [1–5]	

Tablica 2: Nefunkcionalni zahtjevi

4.4. Implementacija

4.5. Analiza rezultata

4.6. Limitacije rješenja

4.7. Manualno testiranje rješenja

5. Zaključak

Popis literature

- [1] *What Is a Digital Image? - Ricoh Scanners*, [Online; accessed 13. Apr. 2025], travanj 2025. adresa: <https://www.pfu-us.ricoh.com/blog/what-is-a-digital-image>.
- [2] *Grayscale Image - an overview | ScienceDirect Topics*, [Online; accessed 20. Apr. 2025], travanj 2025. DOI: 10.1016/B978-0-12-822271-3.00013-X.
- [3] linkgit, „Why use Grayscale Conversion when Processing Images?” *Grayscale Images*, svibanj 2024. adresa: <https://www.grayscaleimage.com/why-use-grayscale-conversion-when-processing-images>.
- [4] *Grayscale Image - an overview | ScienceDirect Topics*, [Online; accessed 16. Apr. 2025], travanj 2025. DOI: 10.1016/B978-0-12-822271-3.00013-X.
- [5] *why we should use gray scale for image processing*, [Online; accessed 20. Apr. 2025], travanj 2025. adresa: <https://stackoverflow.com/questions/12752168/why-we-should-use-gray-scale-for-image-processing>.
- [6] *Image Processing: Techniques, Types, & Applications [2024]*, [Online; accessed 21. Apr. 2025], travanj 2025. adresa: <https://www.v7labs.com/blog/image-processing-guide>.
- [7] [Online; accessed 21. Apr. 2025], travanj 2025. adresa: <https://dl.icdst.org/pdfs/files4/01c56e081202b62bd7d3b4f8545775fb.pdf>.
- [8] M. Patel, „The Complete Guide to Image Preprocessing Techniques in Python,” *Medium*, listopad 2023., ISSN: 3080-4550. adresa: <https://medium.com/@maahip1304/the-complete-guide-to-image-preprocessing-techniques-in-python-dca30804550c>.
- [9] *Image Acquisition | Cloudinary*, [Online; accessed 22. Apr. 2025], travanj 2025. adresa: <https://cloudinary.com/glossary/image-acquisition>.
- [10] *What is noise in photography*, [Online; accessed 22. Apr. 2025], travanj 2025. adresa: <https://www.adobe.com/creativecloud/photography/hub/guides/what-is-noise-in-photography.html>.
- [11] A. Swain, „Noise filtering in Digital Image Processing - Anisha Swain - Medium,” *Medium*, srpanj 2023. adresa: <https://medium.com/@anishaswain/noise-filtering-in-digital-image-processing-d12b5266847c>.

- [12] *Caption for Median Filtering Illustration*, [Online; accessed 22. Apr. 2025], listopad 2022. adresa: <https://visns.neocities.org/MedianFiltering>.
- [13] *Image preprocessing and enhancement | Geospatial Engineering Class Notes | Fiveable*, [Online; accessed 24. Apr. 2025], travanj 2025. adresa: <https://fiveable.me/geospatial-engineering/unit-4/image-preprocessing-enhancement/study-guide/L5fa2eI0Xk8chrKd>.
- [14] R. Shaikh, „OpenCv Perspective Transformation - Analytics Vidhya - Medium,” *Medium*, veljača 2024. adresa: <https://medium.com/analytics-vidhya/opencv-perspective-transformation-9edffefb2143>.
- [15] R. Pulapakura, „Image Segmentation — A Beginner’s Guide | Medium,” *Medium*, veljača 2024. adresa: <https://medium.com/@raj.pulapakura/image-segmentation-a-beginners-guide-0ede91052db7>.
- [16] P. Chhikara, „Understanding Morphological Image Processing and Its Operations | Towards Data Science,” *Towards Data Science*, siječanj 2025. adresa: <https://towardsdatascience.com/understanding-morphological-image-processing-and-its-operations-7bcfled11756>.
- [17] [Online; accessed 25. Apr. 2025], studeni 2022. adresa: https://cmrcet.ac.in/files/ECE/ececoursefile/13.pdf?utm_source=chatgpt.com.
- [18] M. Abramov, „Semantic Segmentation vs Object Detection: Differences | Keymakr,” *Keymakr*, veljača 2025. adresa: <https://keymakr.com/blog/semantic-segmentation-vs-object-detection-understanding-the-differences>.
- [19] P. Sharma, „Image Classification vs. Object Detection vs. Image Segmentation,” *Medium*, prosinac 2021. adresa: <https://medium.com/analytics-vidhya/image-classification-vs-object-detection-vs-image-segmentation-f36db85fe81>.
- [20] siromer, „Contour Detection using OpenCV - siromer - Medium,” *Medium*, siječanj 2025. adresa: <https://medium.com/@siromermer/contour-detection-using-opencv-detecting-and-counting-coins-2d5192597e3c>.
- [21] *Post-processing and visualization | Applications of Scientific Computing Class Notes | Fiveable*, [Online; accessed 28. Apr. 2025], travanj 2025. adresa: <https://library.fiveable.me/applications-of-scientific-computing/unit-8/post-processing-visualization/study-guide/6jrpX4V64bjrgSLQ>.
- [22] R. Sheldon, „image compression,” *WhatIs*, srpanj 2022. adresa: <https://www.techtarget.com/whatis/definition/image-compression>.
- [23] MenyePy, „2 free open-source OMR grading projects on github,” *DEV Community*, veljača 2022. adresa: <https://dev.to/menyePy/2-free-open-source-omr-projects-on-github-2d13>.
- [24] *[v1] Diagrams · Udayraj123/OMRChecker Wiki*, [Online; accessed 29. Apr. 2025], travanj 2025. adresa: <https://github.com/Udayraj123/OMRChecker/wiki/%5Bv1%5D-Diagrams>.

Popis slika

1.	Prikaz slike u boji i pripadajućih kanala boja u RGB modelu (vlastita izrada)	4
2.	Prikaz sive slike (vlastita izrada)	5
3.	Prikaz binarne slike (vlastita izrada)	6
4.	Usporedba svjetlosne komponente i komponenta boja(vlastita izrada)	7
5.	Primjer medijanskog filtera [12]	10
6.	Prikaz procesa ispravka perspektive subjekta na slici [14]	11
7.	Usporedba tipova segmentacije [15]	12
8.	Prikaz erozije kao morfološke operacije [16]	13
9.	Prikaz dilatacije kao morfološke operacije [16]	13
10.	Prikaz kompozitnih morfoloških operacija [16]	14
11.	Usporedba segmentacije i detekcije objekta [19]	15
12.	Prikaz rezultata detekcije krugova pomoću kontura i CvInvoke.MinEnclosingCircle metode [19]	16
13.	Prikaz dijela toka procesiranja slike kod OMRChecker-a [24]	18

Popis tablica

1. Funkcionalni zahtjevi	21
2. Nefunkcionalni zahtjevi	22