

Korištenje neuronskih mreža za uspoređivanje dinamičnih slika

Steiner, Domagoj

Master's thesis / Diplomski rad

2022

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://um.nsk.hr/um:nbn:hr:200:819015>

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

Download date / Datum preuzimanja: **2024-07-15**

Repository / Repozitorij:

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



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA OSIJEK**

Sveučilišni studij

**KORIŠTENJE NEURONSKIH MREŽA ZA
USPOREĐIVANJE DINAMIČNIH SLIKA**

Diplomski rad

Domagoj Steiner

Osijek, 2022

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA **OSIJEK****Obrazac D1: Obrazac za imenovanje Povjerenstva za diplomski ispit**

Osijek, 09.09.2022.

Odboru za završne i diplomske ispite

Imenovanje Povjerenstva za diplomski ispit

Ime i prezime Pristupnika:	Domagoj Steiner
Studij, smjer:	Diplomski sveučilišni studij Automobilsko računarstvo i komunikacije
Mat. br. Pristupnika, godina upisa:	D-52ARK, 11.10.2020.
OIB studenta:	28309588252
Mentor:	Izv. prof. dr. sc. Mario Vranješ
Sumentor:	,
Sumentor iz tvrtke:	Matija Pul
Predsjednik Povjerenstva:	Prof. dr. sc. Marijan Herceg
Član Povjerenstva 1:	Izv. prof. dr. sc. Mario Vranješ
Član Povjerenstva 2:	Izv. prof. dr. sc. Ratko Grbić
Naslov diplomskog rada:	Korištenje neuronskih mreža za uspoređivanje dinamičnih slika
Znanstvena grana diplomskog rada:	Obradba informacija (zn. polje računarstvo)
Zadatak diplomskog rada:	Verifikacija funkcionalnosti aplikacija koje imaju grafičko sučelje jedan je od čestih zadataka u području potrošačke elektronike. To mogu biti aplikacije vezane izravno uz neki uređaj (televizor, set-top-box, mobitel, itd.), a mogu biti i računalne aplikacije. Kod black box testing (BBT) pristupa, ne postoji pristup kodu aplikacije nego se kontrolira ulaz u sustav (klik na ekranu ili pritisak gumba na daljinskom upravljaču) i promatra se prikaz na zaslonu. Drugim riječima, za određen redoslijed akcija korisnika, na zaslonu se mora pokazati određeno stanje grafičkog sučelja. Moderne aplikacije koriste animacije i efekte kako bi učinile svoj sadržaj zanimljivijim i ugodnijim oku korisnika. Stoga sadržaj slike ovjese trenutku u kojem je preuzeta, tj. slike preuzete iz
Prijedlog ocjene pismenog dijela ispita (diplomskog rada):	Izvrstan (5)
Kratko obrazloženje ocjene prema Kriterijima za ocjenjivanje završnih i diplomskih radova:	Primjena znanja stečenih na fakultetu: 3 bod/boda Postignuti rezultati u odnosu na složenost zadatka: 3 bod/boda Jasnoća pismenog izražavanja: 3 bod/boda Razina samostalnosti: 3 razina
Datum prijedloga ocjene od strane mentora:	09.09.2022.
Potvrda mentora o predaji konačne verzije rada:	<i>Mentor elektronički potpisao predaju konačne verzije.</i>
	Datum:

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**IZJAVA O ORIGINALNOSTI RADA**

Osijek, 19.09.2022.

Ime i prezime studenta:

Domagoj Steiner

Studij:

Diplomski sveučilišni studij Automobilsko računarstvo i komunikacije

Mat. br. studenta, godina upisa:

D-52ARK, 11.10.2020.

Turnitin podudaranje [%]:

3

Ovom izjavom izjavljujem da je rad pod nazivom: **Korištenje neuronskih mreža za uspoređivanje dinamičnih slika**

izrađen pod vodstvom mentora Izv. prof. dr. sc. Mario Vranješ

i sumentora ,

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija. Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis studenta:

IZJAVA

o odobrenju za pohranu i objavu ocjenskog rada

kojom ja Domagoj Steiner, OIB: 28309588252, student/ica Fakulteta elektrotehnike,

računarstva i informacijskih tehnologija Osijek na studiju Diplomski sveučilišni studij Automobilsko računarstvo i komunikacije, kao autor/ica ocjenskog rada pod naslovom:

Korištenje neuronskih mreža za uspoređivanje dinamičnih slika,

dajem odobrenje da se, bez naknade, trajno pohrani moj ocjenski rad u javno dostupnom digitalnom repozitoriju ustanove Fakulteta elektrotehnike, računarstva i informacijskih tehnologija Osijek i Sveučilišta te u javnoj internetskoj bazi radova Nacionalne i sveučilišne knjižnice u Zagrebu, sukladno obvezi iz odredbe članka 83. stavka 11. *Zakona o znanstvenoj djelatnosti i visokom obrazovanju* (NN 123/03, 198/03, 105/04, 174/04, 02/07, 46/07, 45/09, 63/11, 94/13, 139/13, 101/14, 60/15).

Potvrđujem da je za pohranu dostavljena završna verzija obranjenog i dovršenog ocjenskog rada. Ovom izjavom, kao autor/ica ocjenskog rada dajem odobrenje i da se moj ocjenski rad, bez naknade, trajno javno objavi i besplatno učini dostupnim:

- a) široj javnosti
- b) studentima/icama i djelatnicima/ama ustanove
- c) široj javnosti, ali nakon proteka 6 / 12 / 24 mjeseci (zaokružite odgovarajući broj mjeseci).

**U slučaju potrebe dodatnog ograničavanja pristupa Vašem ocjenskom radu, podnosi se obrazloženi zahtjev nadležnom tijelu Ustanove.*

Osijek, 19.09.2022.

(mjesto i datum)

(vlastoručni potpis studenta/ice)

SADRŽAJ

1. UVOD.....	1
2. USPOREDBA DINAMIČNIH SLIKA	3
2.1. Pregled postojećih baza dinamičnih slika.....	6
2.2. Postojeća rješenja za usporedbu dinamičnih slika	13
2.2.1. Postojeće rješenje za usporedbu dinamičnih slika zasnovano na ključnim točkama	13
2.2.2. Postojeća rješenja za usporedbu dinamičnih slika zasnovana na neuronskim mrežama	14
3. IZRADA NEURONSKE MREŽE ZA USPOREDBU DINAMIČNIH SLIKA	19
3.1. Podešavanje radnog okruženja i instalacija biblioteka	19
3.2. Kreiranje vlastite baze dinamičnih slika za treniranje neuronske mreže	20
3.2.1. Algoritam za generiranje dinamičnih slika.....	21
3.2.2. Generirana baza dinamičnih slika za treniranje neuronske mreže	26
3.3. Arhitektura sijamske neuronske mreže.....	30
3.4. Treniranje sijamske neuronske mreže.....	32
3.5. Upute za pokretanje i korištenje	36
3.5.1. Upute za pokretanje i korištenje generatora slika	36
3.5.2. Upute za treniranje i testiranje sijamske neuronske mreže	38
4. VERIFIKACIJA RADA NEURONSKE MREŽE ZA USPOREDBU DINAMIČNIH SLIKA	41
4.1. Verifikacija rada neuronske mreže na CGIAD bazi podataka	44
4.2. Verifikacija rada neuronske mreže na bazi slika STB grafičkog sučelja	54
4.3. Usporedba rezultata sijamske neuronske mreže i DISM algoritma	66
5. Zaključak.....	72
LITERATURA.....	74
SAŽETAK	76
ABSTRACT.....	77
ŽIVOTOPIS	78
PRILOZI.....	79

1. UVOD

Razvoj tehnologije doveo je i do razvoja aplikacija koje koristimo u svakodnevnom životu. Zbog velikog broja aplikacija koje su se počele pojavljivati, programerske tvrtke primorane su osmisliti načine kako ostati konkurentni na tržištu i privući što veći broj ljudi da koriste njihove aplikacije. Jedan od načina na koji programerske tvrtke pokušavaju privući korisnike je implantacijom interaktivnijih i animiranih grafičkih sučelja. Takva se sučelja nazivaju dinamična grafička sučelja. Obilježje dinamičnih grafičkih sučelja i izbornika s animacijom je to što se pozadina i objekti koji se nalaze na takvim grafičkim sučeljima ili izbornicima s animacijom mogu mijenjati tijekom vremena. Uz mijenjanje pozadine i objekta tijekom vremena, mogu postojati i efekti koji prekrivaju objekte na grafičkom sučelju ili im mijenjaju oblik na neki način (npr. rotiraju ih, povećavaju ih, smanjuju ih). Dinamična grafička sučelja mogu stvoriti problem tijekom automatiziranih funkcionalnih testiranja kod kojih se provjerava je li grafičko sučelje aplikacije dobro prikazano, odnosno prikazuje li sučelje ono što bi i trebalo nakon što korisnik pritisne određenu tipku ili odabere određeni izbornik. Provjera se izvodi tako što se uzme slika objekta na ispravnoj verziji aplikacije u trenutku kada je određeni izbornik na sučelju izabran i slika istog objekta na testnoj verziji aplikacije nakon što se izabere isti izbornik. Ove se dvije dobivene slike uspoređuju. Zbog dinamičnog grafičkog sučelja, dvije odabrane slike prikazivati će isti objekt, ali objekti mogu biti u različitim pozicijama, različito prekriveni određenim efektom ili s drugačijom pozadinom. Ove dvije slike primjer su onoga što nazivamo dinamične slike. Dinamične su slike one slike koje imaju jednaki kontekst, ali različit sadržaj. Kontekst slike opisuje ono što se prikazuje na slici, a u slučaju grafičkog sučelja to je objekt koji se promatra. Sadržaj slike predstavljaju elementi slike (engl. *pixel*). Oni predstavljaju najmanju osnovnu jedinicu slike. Zbog efekata i različitih pozicija objekata na svakoj od dinamičnih slika, njihove su vrijednosti drugačije na svakoj od odabranih slika. Ljudima, kada uspoređuju dinamične slike, nije problem vidjeti da se na obje izabrane slike nalazi isti objekt. Problem nastaje kada se usporedba dinamičnih slika pokuša automatizirati tako da računalo to obavlja samostalno.

Obično, računala uspoređuju slike tako da uspoređuju element slike jedan po jedan, odnosno računala vrše usporedbu slika na temelju sadržaja slika. Osobina je dinamičnih slika različit sadržaj, a jednak kontekst. Ovo znači da par dinamičnih slika prikazuje isti objekt. Ovaj problem postaje jedan od većih problema koje rješava područje tehnologije zvano računalni vid. Zadatak je računalnog vida automatizacija ljudskog vizualnog sustava na računalima, a ono se postiže obradom digitalnih slika i videozapisa. Jedna od metoda računalnog vida pomoću koje se pokušava

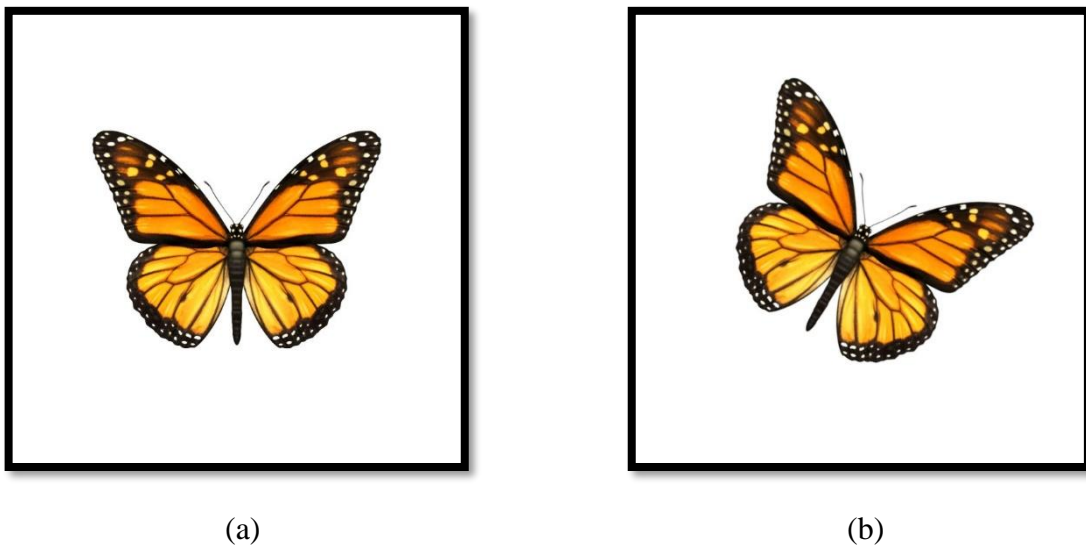
riješiti problem usporedbe dinamičnih slika, a nije zasnovana na strojnom učenju, je pomoću izdvajanja i uparivanja značajki, tj. ključnih točaka. Uspoređivanje dinamičnih slika vrlo se uspješno izvodi pomoću ove metode, ali razvoj tehnologija nudi nam mogućnosti za potencijalno boljim i bržim rješenjima. Kako su trenutno neuronske mreže jedan od najčešćih alata za detekciju objekata na slici, korištenje neuronskih mreža predstavlja jedno od mogućih rješenja za usporedbu dinamičnih slika. Cilj je ovog diplomskog rada provjeriti mogu li se koristiti neuronske mreže za rješavanje problema usporedbe dinamičnih slika i ispitati kakvi se rezultati za rješavanje problema pomoću neuronske mreže dobivaju u usporedbi s rješavanjem problema pomoću izdvajanja i uparivanja značajki.

Zadatak ovog diplomskog rada bio je izraditi alat za usporedbu dinamičnih slika koji je zasnovan na neuronskim mrežama. Za ostvarivanje ovog zadatka bilo je potrebno proučiti postojeća rješenja za usporedbu dinamičnih slika koja su zasnovana na osnovnim tehnikama računalnog vida i rješenja koja su zasnovana na neuronskim mrežama. Nakon proučavanja postojećih rješenja, potrebno je iskoristiti jedno postojeće rješenje zasnovano na neuronskim mrežama, nadograditi ga i utvrditi je li moguće koristiti neuronske mreže za rješavanje problema usporedbe dinamičnih slika. Kako bi se neuronske mreže mogle trenirati, potrebno je osigurati opsežan skup podataka za treniranje. Stoga je u sklopu ovog rada bilo potrebno izraditi alat koji će od jedne ulazne slike primjenom različitih transformacija kreirati set slika za treniranje (ovdje se misli na kreiranje seta slika, od kojih svaka kreirana slika čini par dinamičnih slika sa slikom od koje je izvorno nastala).

U drugom poglavlju ovog rada ukratko su predstavljene postojeće baze dinamičnih slika i postojeća rješenja za usporedbu dinamičnih slika. U trećem poglavlju opisan je postupak izrade neuronske mreže i baze dinamičnih slika potrebnih za treniranje koje su nastale tijekom izrade diplomskog rada. Nadalje, u četvrtom su poglavlju predstavljeni rezultati testiranja i analiza rada neuronske mreže. Također u četvrtom je poglavlju dana usporedba rješenja zasnovanih na izdvajanju i uparivanju značajki i na neuronskoj mreži koja je predložena u ovom radu. Na kraju rada, u petom poglavlju, iznesen je zaključak rada.

2. USPOREDBA DINAMIČNIH SLIKA

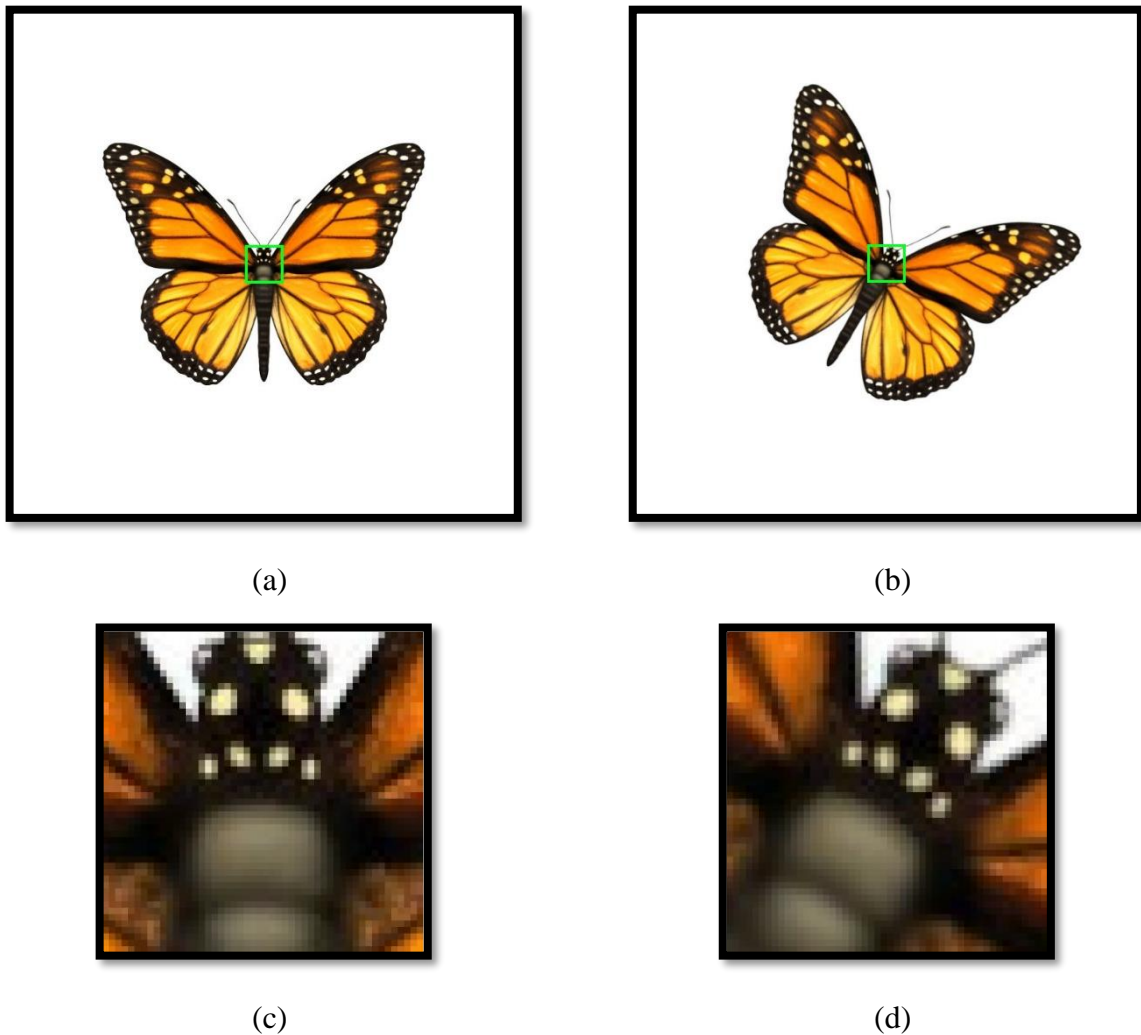
Svaka digitalna slika ima svoj sadržaj i svoj kontekst. Kod statičnih se slika sadržaj i kontekst tijekom vremena ne mijenjaju, dok se kod dinamičnih slika sadržaj tijekom vremena mijenja. Kao primjer dinamičnih slika možemo uzeti slike preuzete s kamere koja je snimala životinje u šumi. Prva slika prikazuje ulazak životinje u kadar. Na prvoj je slici životinja mala jer je daleko od kamere. Također, ne možemo ju vidjeti cijelu jer se nalazi iza rijetkog grmlja. Drugu sliku uzimamo kada se životinja nalazi bliže kameri. U tom je trenutku životinja puno veća u usporedbi s prvom slikom, a i ispred nje se ne nalazi grmlje te ju vidimo bez problema. Kada se slike uspoređuju, vidljivo je da je ista životinja prikazana na obje slike. Ovo bi značilo da je kontekst slika jednak. Obje slike prikazuju istu životinju u istom području šume koju snimamo. Razlike koje nalazimo promatrajući ove dvije slike su mijenjanje položaja životinje, promjena veličine životinje te promjena prekrivenosti životinje okolnom vegetacijom. Spomenute promjene između tih dviju slika utječu na promjenu sadržaja slika, odnosno promjene vrijednosti elemenata slika. Mijenjanje sadržaja slika uzrokuje problem prilikom standardne usporedbe dinamičnih slika, jer uspoređuju se elementi slika za koje je moguće da na istim lokacijama slike nisu jednaki na objema slikama. Slika 2.1 prikazuje jednostavan primjer dviju dinamičnih slika.



Slika 2.1 Dinamične slike na primjeru slike leptira: (a) referentna slika leptira, (b) zarotirana slika leptira

Usporedbom dinamičnih slika (Slika 2.1) pokušava se dokazati da se na objema slikama nalazi isti objekt, odnosno da se na objema slikama nalazi isti leptir, što bi značilo da slike imaju isti kontekst. Ljudi zamjećuju da se radi o slikama koje prikazuju istog leptira, samo što je na slici 2.1 (b) taj leptir zarotiran. Kada bi se vršila usporedba (Slika 2.1 i Slika 2.2) po sadržaju, svaki element slike zasebno, uočilo bi se da će na pojedinim prostornim lokacijama u dvjema slikama

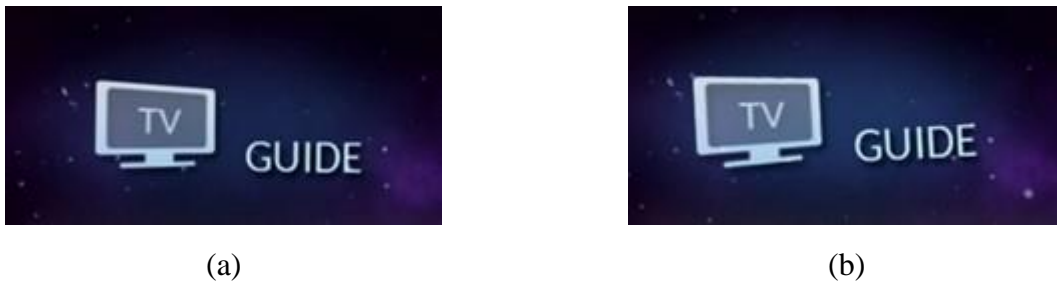
elementi slike biti različiti, što bi značilo da su slike različite po sadržaju. Bilo kakve promjene sadržaja na drugoj slici, koje ne mijenjaju kontekst slike i objekte koje slika prikazuje, stvaraju problem za standardne algoritme usporedbe dinamičnih slika koji su zasnovani na uspoređivanju elemenata slika. Glavni je zadatak ovoga diplomskog rada napraviti neuronsku mrežu koja će uspoređivati dinamične slike, odnosno uspoređivati dinamične slike na temelju njihovog konteksta, a ne sadržaja te za dani par dinamičnih slika kao izlaz dati odgovor jesu li one iste ili različite (u smislu konteksta).



Slika 2.2 (a) Referentna slika leptira, (b) rotirana slika leptira (c) označeni dio slike (a) uvećan 100 puta, (d) označeni dio slike (b) uvećan 100 puta

Tijekom funkcionalnog testiranja na primjeru STB (engl. *Set-Up Box*) dinamičnog korisničkog sučelja, izvodi se uspoređivanje slika preuzetih s dinamičnog grafičkog sučelja aplikacije. Za proces usporedbe slika dinamičnog grafičkog sučelja koriste se dvije slike. Prva odabrana slika predstavlja referentnu sliku. Referentnu sliku uzimamo s ispravne verzije aplikacije. Druga slika koja se uzima predstavlja testnu sliku, a nju se uzima s testne verzije

aplikacije. Primjer referentne i testne slike uzetih s dinamičnog grafičkog sučelja prikazan je na slici 2.3 .



Slika 2.3 Dinamične slike na primjeru STB dinamičnog korisničkog sučelja: (a) referentna slika, (b) testna slika

Kao u prethodnim primjerima, usporedbom ovih slika pokušava se provjeriti prikazuju li slike iste objekte. Tijekom funkcionalnog testiranja pokušava se provjeriti prikazuje li dinamično grafičko sučelje ono što bi trebalo prikazati nakon pritiska određene tipke na daljinskom upravljaču. U slučaju prikazanom na slici 2.3 to je izbornik *Guide* s odgovarajućom slikom. Gledanjem prikazane slike, ljudi vide da se radi o istom objektu, samo je na testnoj slici objekt malo zarotiran. Kako bi se mogle usporediti ispravna verzija aplikacije i testna verzija aplikacije, uzima se nekoliko uzastopnih slika izbornika s grafičkog sučelja testne verzije aplikacije. Zbog animacija, izbornik se na grafičkom sučelju kreće i moguće je da se na jednoj od uzastopnih slika dogodi trenutak kada je sadržaj testne i referentne slike jednak. U ovom će slučaju standardni algoritmi za usporedbu slika tvrditi da se na slikama nalazi isti objekt. Za ostale testne slike, gdje je isti objekt malo pomaknut ili je prekriven manjim animiranim objektima (npr. točkice, pahuljice, kapljice), standardni algoritmi za usporedbu slika pokazat će da se na testnoj slici nalazi različit objekt od onog koji se nalazi na referentnoj slici. Iz tog razloga standardni algoritmi za usporedbu slika nisu pouzdani jer se nikad ne zna u kojem trenutku će biti uzeta neka od testnih slika, odnosno koliko će objekt koji promatramo bit zarotiran ili prekriven manjim animiranim objektima. Na slici 2.4 prikazani su još neki od primjera dinamičnih slika s dinamičnog grafičkog sučelja aplikacije, a u potpoglavlju 2.1 bit će opisane postojeće baze dinamičnih slika.





(c)

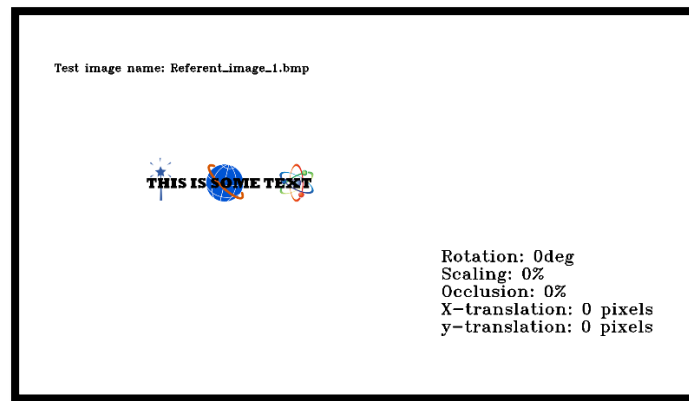


(d)

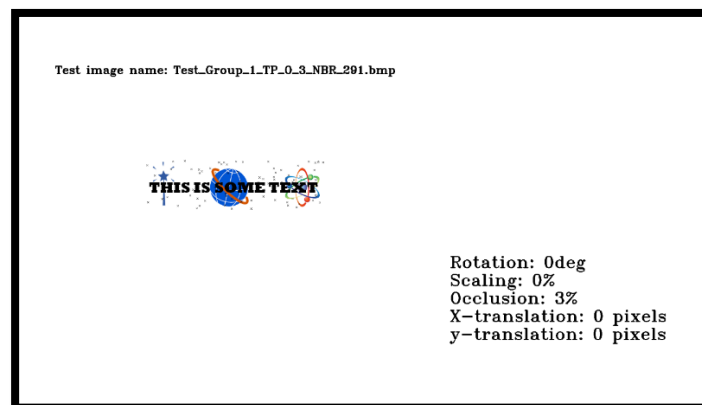
Slika 2.4 Dinamične slike na primjeru STB dinamičnog korisničkog sučelja: (a) referenta slika, (b) testna slika s prekrivenošću sivim točkicama, (c) testna slika s uvećanim objektom i prekrivenošću sivim točkicama, (d) testna slika zarotirana s uvećanim objektom i prekrivenošću sivim točkicama

2.1. Pregled postojećih baza dinamičnih slika

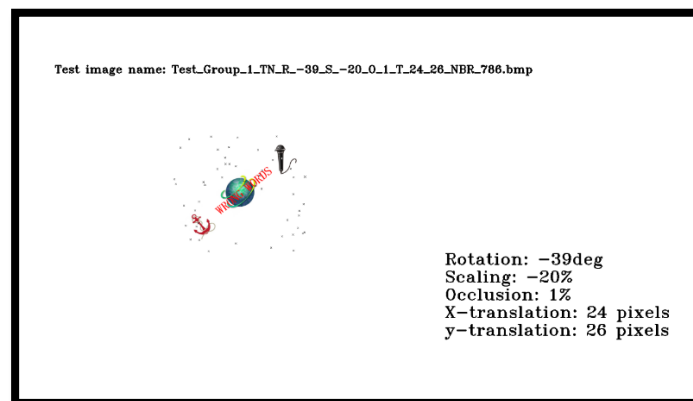
CGIAD (engl. *Computer Generated Image Alteration Database*) [1] je računalno generirana baza dinamičnih slika, čije su dimenzije 1920x1080 elemenata slike. Sastoji se od 10 testnih setova i u svakom setu se nalaze dvije slične slike. Dvije slične slike prikazuju slične objekte koji čine da su slike kontekstom vrlo slične, ali sadržajem potpuno različite. Izvršavanjem transformacija na objema slikama testnog seta dobivaju se testne skupine. Svaka testna skupina sastoji se od jedne referentne slike, 500 pozitivnih slika (objekt na slici jednak kao na pripadnoj referentnoj slici, ali izložen transformacijama) i 500 negativnih slika (objekt na slici sličan objektu na referentnoj slici i izložen transformacijama). Sve pozitivne i negativne slike svih testnih skupina sastoje se 50 slika isključivo s rotacijom objekta, 50 slika isključivo sa skaliranjem objekta, 50 slika isključivo s translacijom objekta, 50 slika gdje postoji prekrivenost objekta na slici i 300 slika na kojima je kombinacija svih navedenih transformacija [1]. To na kraju čini da se svaka od 10 testnih skupina sastoji od 1000 slika. Svaki objekt koji se nalazi na testnoj slici izložen je jednoj ili više transformacija kao što su translacija, rotacija, skaliranje i prekrivenost objekta (engl. *occlusion*) malim sivim točkicama veličine 5x5 elemenata slike. Maksimalne vrijednosti ovih transformacija su 40° za rotaciju objekta u oba smjera, smanjivanje i povećavanje objekta za 20% njegove originalne veličine odnosno skaliranje objekta za 20%, translacija objekta za 30 elemenata slike vertikalno i horizontalno te 10% za prekrivenost objekta. Primjer referentne, pozitivne i negativne slike CGIAD baze dinamičnih slika prikazan je na slici 2.5.



(a)



(b)



(c)

Slika 2.5 Primjer dinamičnih slika CGAID testne grupe 1 (a) referentna slika, (b) pozitivna slika, (c) negativna slika

Odlika CGIAD baze dinamičnih slika je ta što se testne grupe sastoje od slika koje imaju više različitih konteksta, za razliku od baza slika koje će biti spomenute u nastavku. Baze koje će biti spomenute u nastavku prikazuju samo jedan specifičan kontekst s promjenjivim sadržajem. Iz

tog razloga CGIAD predstavlja dobru bazu za razvijanje algoritama za usporedbu dinamičnih slika te prema vlastitom saznanju je jedina javno dostupna baza dinamičnih slika.

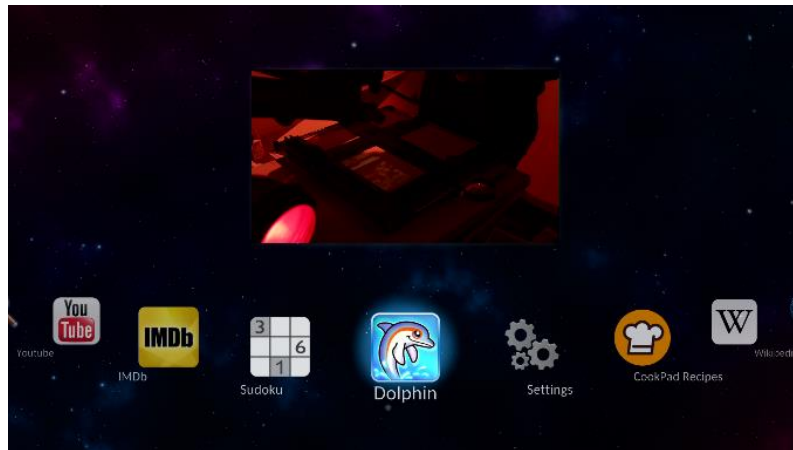
Za provjeru kako algoritmi usporedbe dinamičnih slika rade u uvjetima funkcionalnog testiranja, napravljena je baza slika STB grafičkog sučelja dobivenih iz grafičkog sučelja C-MORE UI/UX [2]. Ovu bazu dinamičnih slika kreirao je Institut RT-RK Osijek d.o.o te nije javno dostupna za korištenje. Promotivni video grafičkog sučelja može se pogledati na [2], a na slici 2.6 se nalazi primjer grafičkog sučelja C-MORE UI/UX.



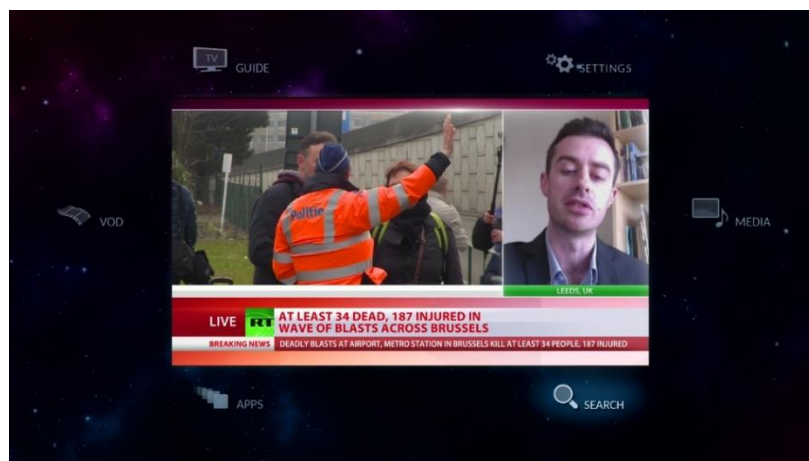
Slika 2.6 Primjer grafičkog sučelja C-MORE UI/UX

Baza slika STB grafičkog sučelja baza je od 50 testnih slika dinamičnog grafičkog sučelja. Ova baza predstavlja prave testne slučajeve koji se koriste tijekom funkcionalnog testiranja aplikacije. Sastoji se od 20 testnih skupina s jednom referentnom slikom i različitim brojem testnih slika po svakoj testnoj skupini. Svaka testna skupina predstavlja jedan objekt na grafičkom sučelju koji je izložen jednoj ili više transformacija, kao što su rotacija, translacija, okluzija i skaliranje. Slike prikazuju tri različita grafička sučelja (grafičko sučelje aplikacija, grafičko sučelje izbornika i grafičko sučelje TV rasporeda). Grafičko sučelje aplikacija prikazano je na dvjema testnim skupinama i ono prikazuje sučelje gdje se nalazi izbor više različitih aplikacija. Grafičko sučelje izbornika prikazano je na 12 testnih skupina. Na 12 testnih skupina grafičkog sučelja izbornika prikazano je 6 izbornika s 2 različita stanja, kada je izbornik izabran i kada izbornik nije izabran. Kada je jedan od izbornika na ovom grafičkom sučelju izabran, izbornik se poveća, posvijetli i počinje se rotirati. Na 6 testnih skupina prikazano je grafičko sučelje TV rasporeda. Grafičko sučelje TV rasporeda prikazuje kada su odabrani različiti dani na TV rasporedu i kada su odabrani

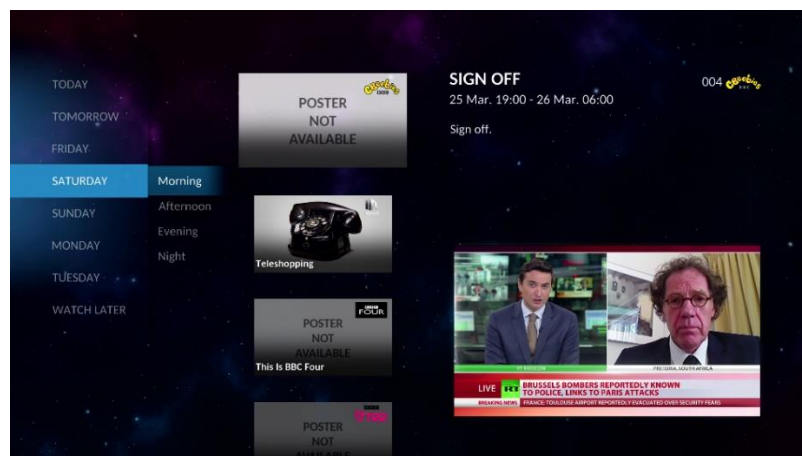
različiti programi u TV rasporedu. Sve su slike ove baze dinamičnih slika rezolucije 1920x1080 elemenata slike. Primjeri svih triju grafičkih sučelja koja se nalaze u bazi slika STB grafičkog sučelja prikazana su na slici 2.7.



(a)



(b)



(c)

Slika 2.7 Sučelja baze slika STB grafičkog sučelja (a) grafičko sučelje aplikacija, (b) grafičko sučelje izbornika, (c) grafičko sučelje TV rasporeda

U nastavku spomenute su baze slika koje se sastoje od jednog specifičnog konteksta s promjenjivim sadržajem. Jedna od takvih baza je FERET baza fotografija lica osoba [3]. FERET baza fotografija lica osoba sastoji se od 14051 portreta prikupljenih od 1209 osoba. Osobe su fotografirane iz profila, poluprofila te sprijeda. Na pojedinim fotografijama od osoba je zahtijevano da naprave tužnu, sretnu ili iznenađenu izraz lica licu, stave naočale ili povežu kosu. Kako bi se dodatno povećala raznolikost baze fotografija, osobe su fotografirane s različitim pozadinama, a pojedine su slike skalirane ili posvijetljene. Na slici 2.8 prikazani su primjeri fotografija FERET baze fotografija lica osoba.



(a)



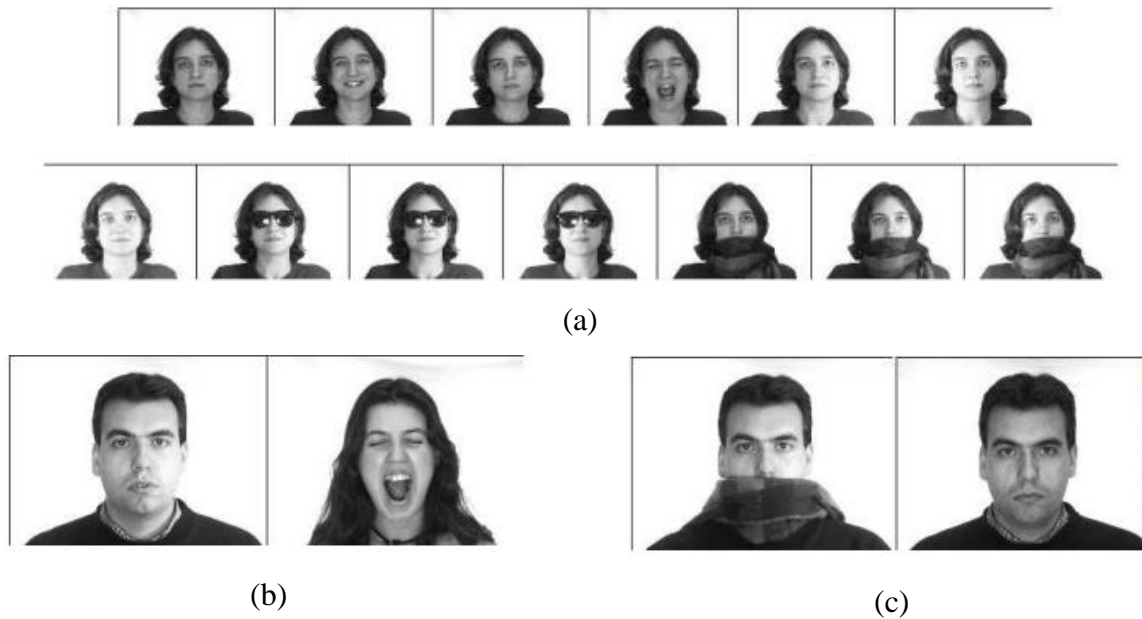
(b)



(c)

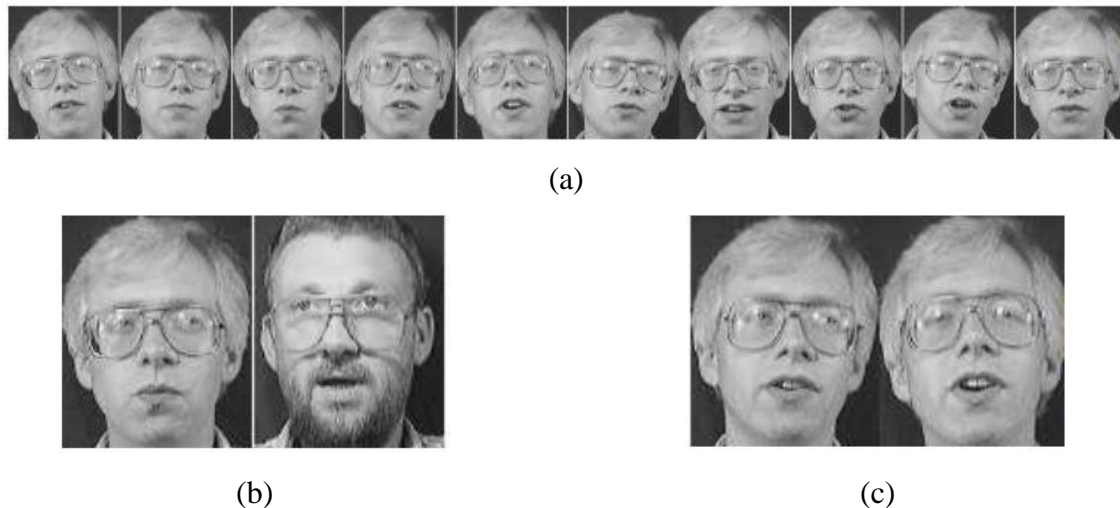
Slika 2.8 Primjer fotografija FERET baze fotografija lica: (a) sve fotografije jedne osobe, (b) jedan par fotografija gdje su različite osobe i (c) jedan par fotografija gdje je na objema fotografijama ista osoba [4]

AR baza fotografija lica osoba [5] sastoji se od 3000 fotografija lica osoba u boji. Sve fotografije osoba su fotografirane sprijeda u različitim uvjetima osvjetljenja te s drugačijim promjenama na licima ljudi, kao što su drugačiji izrazi lica, objekti na licu (npr. šal i naočale). Za potrebe baze fotografija osoba fotografirano je 116 osoba, 63 muške osobe i 53 ženske osobe. Slika 2.9 prikazuje primjere fotografija AR baze fotografija lica osoba



Slika 2.9 Primjer fotografija AR baze fotografija lica: (a) sve fotografije jedne osobe, (b) jedan par fotografija gdje su različite osobe, (c) jedan par gdje je na objema fotografijama ista osoba [4]

AT&T ili drugog naziva ORL baza fotografija lica [6] sastoji se od 400 fotografija lica, od čega je 10 fotografija po osobi. Fotografije se razlikuju u izrazima lica osoba, nosi li osoba naočale te različitim osvjetljenjima prilikom fotografiranja. Sve fotografije su fotografirane sprijeda i sadrže tamnu pozadinu. Na slici 2.10 prikazani su primjeri AT&T ili ORL baze fotografija lica osoba.



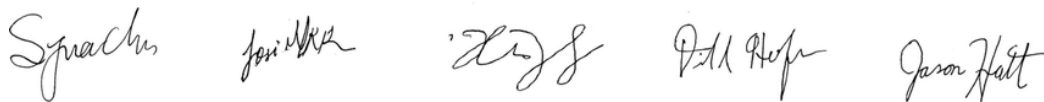
Slika 2.10 Primjer fotografija AT&T baze fotografija lica: (a) sve fotografije jedne osobe, (b) jedan par fotografija gdje su različite osobe i (c) jedan par gdje je na objema fotografijama ista osoba [4]

CEDAR baza potpisa [7], [8] je baza koja sadržava potpise 55 osoba, od kojih je traženo da u razmaku od 20 minuta naprave po 24 potpisa, što ukupno čini 1320 originalnih potpisa. Uz

originalne potpise, u ovoj bazi se nalaze i 1320 krivotvorenih potpisa. Na slici 2.11 je prikazani su primjeri potpisa CEDAR baze potpisa.



(a)



(b)

Slika 2.11 Primjer CEDAR baze potpisa. (a) originalni potpisi, (b) krivotvoreni potpisi [9]

Baza potpisa slična CEDAR bazi potpisa je GPDS-300 baza potpisa [10], [11], a čine ju 7200 originalnih potpisa i 9000 krivotvorenih potpisa od ukupno 300 osoba. Na slici 2.12 prikazani su primjeri potpisa GPDS-300 baze potpisa.



(a)



(b)

Slika 2.12 Primjer GPDS-300 baze potpisa. (a) originalni potpisi, (b) krivotvoreni potpisi [9]

Od svih ranije opisanih baza, jedino CGIAD baza dinamičnih slika i baza slika STB grafičkog sučelja će se koristiti u izradi ovog diplomskog rada. Ostale baze, baze fotografija osoba i slika potpisa, imaju samo jedan specifičan kontekst, prikazuju osobu ili prikazuju nečiji potpis. Jedan specifičan kontekst nije dovoljan ako se želi testirati ispravna klasifikacija negativnih slika, odnosno prepoznavanje nalazi li se različit objekt na slikama. Zbog tog razloga potrebna nam je baza dinamičnih slika s više različitih konteksta i promjenjivim sadržajima. Jedina spomenuta baza dinamičnih slika koja ima više različitih konteksta je CGIAD baza dinamičnih slika. Kako bi se izradila još jedna baza dinamičnih slika s više različitih konteksta, gdje bi imali kontrolu nad transformacijama koje se izvršavaju nad objektima na slikama, u sklopu ovog diplomskog rada izrađen je alat koji će generirati transformirane slike.

2.2. Postojeća rješenja za usporedbu dinamičnih slika

2.2.1. Usporedba dinamičnih slika zasnovana na ključnim točkama

Za uspješnu usporedbu dinamičnih slika potrebno je uspješno detektirati objekte na objema slikama te ih zatim usporediti. Objekti na slikama prepoznaju se po naglašenim rubovima i naglim prijelazima intenziteta. Zbog ovih karakteristika objekta može se pretpostaviti da će se na području gdje se nalazi objekt pojaviti veći broj ključnih točaka. Ključne točke su točke na slici koja su po nečemu jedinstvena za tu sliku, tj. određeni objekt, i pomažu pri detekciji objekata na slici. Korištenje različitih metoda za traženje ključnih točaka nudi otpornost na neke od promjena koje se mogu izvršiti na slici, kao što su skaliranje, rotacija i translacija, a za rad s ključnim točkama potrebno je obaviti dva koraka – detekcija ključnih točaka i opis njihovih okolina. Napredni algoritmi koji rade detekciju ključnih točaka i opis okolina ključnih točaka su SIFT (engl. *Scale-Invariant Feature Transform*) [12], SURF (engl. *Speeded-Up Robust Features*) [13] i ORB (engl. *Oriented FAST and Rotated BRIEF*) [14].

SIFT algoritam služi za detekciju ključnih točaka i opis njihove okoline. Ključne točke određene pomoću ovog algoritma otporne su na promjene u veličini, rotaciji i osvjetljenju slike. SIFT algoritam sastoji se od četiri faze: detekcija ekstrema, lokalizacija ključnih točaka, određivanje orijentacije ključnih točaka te na kraju oblikovanje deskriptora ključne točke. SIFT algoritam efikasno detektira ključne točke, ali je računalno zahtjevan algoritam. Algoritam koji je sličan SIFT algoritmu i brži u pronalasku ključnih točaka je SURF. Kao i kod SIFT-a, SURF prvo detektira ključne točke. Detekcija ključnih točaka izvršava se koristeći Hessian matrice, a određuju se na temelju maksimuma determinante Hessian matrice. Nakon što su ključne točke određene, određuju se njihove orijentacije pomoću Harrovih filtara u x i y smjeru, a na rezultat filtriranja primjenjuju se težine određene Gaussovom funkcijom. Posljednji je korak oblikovanje deskriptora. Oblikovanje deskriptora izvodi se na način da se okolina ključne točke dijeli na manje podregije, a za svaku se podregiju računa vektor deskriptora. SURF metoda izdvajanja značajki osjetljiva je na promjene osvjetljenja i promjenu kuta gledanja. ORB algoritam sastoji se od rBRIEF (engl. *Rotated Binary Robust Independent Elementary Features*) [15] i oFAST (engl. *Oriented Features from Accelerated Segment Test*) [16] algoritama. Algoritam oFAST je algoritam koji se koristi za detektiranje ključnih točaka. Naziv oFAST je dobio zbog dodavanja orijentacijske komponente na detektirane ključne točke pomoću FAST. BRIEF algoritam koristi se za opisivanje okoline ključnih točaka, odnosno dobivanje deskriptora. Sami po sebi, BRIEF deskriptori nisu otporni na rotaciju slike. Iz tog razloga u ORB algoritmu se koristi nadograđeni BRIEF algoritam koji se naziva rBRIEF, koji pruža otpornost na rotaciju slike

Na temelju ORB algoritma napravljen je DISM (engl. *Dynamic Image Similarity Measurement*) algoritam [17], [18] koji služi za uspoređivanje dinamičnih slika. DISM algoritam prvo pomoću ORB algoritma na referentnoj i testnoj slici detektira ključne točke. One ključne točke čiji je odziv slabiji od traženog, odbacuju se, te se na taj način izvršava filtriranje ključnih točaka. Preostale ključne točke se uspoređuju, odnosno uspoređuju se deskriptori koji su izračunati za te ključne točke. Pomoću ovog koraka usporedbe deskriptora ključnih točaka se određuju referentni objekti. Određivanje referentnih objekta se postiže na način da se uzme objekt s referentne slike i traži se s kojim objektom s testne slike ima najviše jednakih deskriptora. Ostali objekti s testne slike s kojima objekt s referentne slike ima jednake deskriptore smatraju se kao neispravno detektirani jednaki deskriptori te su filtrirani. Filtriranjem se postiže da pojedini objekt s referentne slike odgovara samo jednom objektu s testne slike. Nakon izvršenog filtriranja povezuju se referentni objekti s njihovim parovima na testnoj slici. Ako neki od referentnih objekata nije pronađen na testnoj slici, algoritam zaključuje da se radi o različitim slikama. Ukoliko su pronađeni svi referentni objekti, tada se izračunava postotak preklapanja svih objekata te se izračunava ukupni postotak preklapanja pomoću kojega se zaključuje radi li se o jednakim ili različitim slikama. Formula (2-1) označava formulu za postotak preklapanja, a oznaka K.T. u formuli označava riječ ključna točka:

$$P[\%] = \frac{\text{broj K.T. referentne slike detektiranih na testnoj slici nakon primjene algoritma}}{\text{broj K.T. pronađenih na referentnoj slici}} * 100 \quad (2-1)$$

Ako dobiveni postotak poklapanja zadovoljava proizvoljno postavljeni prag preklapanja, tada se može reći da su slike jednake. Kada se DISM algoritam testira na 10000 slika CGIAD baze dinamičnih slika, uz prag preklapanja od 50% algoritam je dvije slike klasificirao kao neispravno pozitivne slike i dvije slike kao neispravno negativne. Ovo označava da se dobiva točnost od 99.96%. DISM algoritam zatim testiran je na 50 slika baze slika STB grafičkog sučelja. Uz prag od 80% poklapanja, dobivena je točnost od 100%.

DISM algoritam jedini je javno dostupni algoritam zasnovan na ključnim točkama koji učinkovito uspoređuje dinamične slike i posljedično, jedini koji je spomenut u ovom diplomskom radu.

2.2.2. Usporedba dinamičnih slika zasnovana na neuronskim mrežama

Razvoj neuronskih mreža omogućio je korištenje neuronske mreže u svrhe uspoređivanja potpisa ljudi te detektiranja je li predstavljeni potpis krivotvoren ili je pravi. Prvi primjer takve

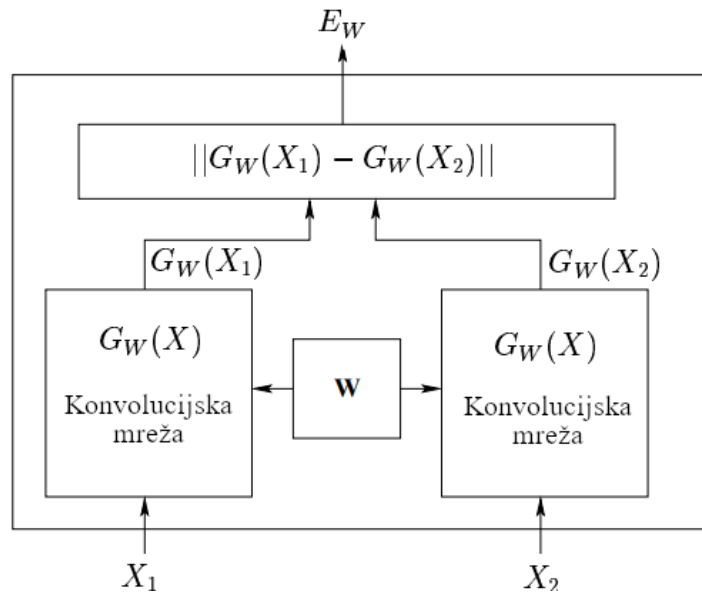
primjene neuronske mreže pružili su Bromley i ostali [19] u svom radu, u kojem uspoređuju prave i krivotvorene potpise. U svom radu koristili su sijamsku neuronsku mrežu. Sijamska neuronska mreža sastoji se od dvije pod-mreže koje imaju spojene izlaze. Tijekom treniranja dvije pod-mreže izvlače značajke iz dvaju potpisa, dok zajednički neuron izračunava udaljenost između dvaju vektora značajki. Ako je udaljenost vektora značajki manja od proizvoljno postavljenog praga udaljenosti dvaju vektora značajki, potpis se prihvaća kao originalan, dok se u suprotnom slučaju potpis smatra krivotvorenim. Sijamsku neuronsku mrežu testirali su na 532 originalna potpisa i 424 krivotvorena potpisa. Kada je postavljen prag da se detektira 80% krivotvorenih potpisa, ispravna identifikacija originalnih potpisa dogodila se u 95,5 % slučajeva. Preostalih 4,5 % slučajeva ukazuje na to da je originalni potpis prepoznat kao da je krivotvoren, odnosno 24 originalna potpisa su odbijena.

Chopra, Hadsell i LeCun [4] iskoristili su ovu ideju korištenja neuronske mreže za usporedbu potpisa i ponudili rješenje za problem otkrivanja nalazi li se na dvjema fotografijama ista osoba. Za rješavanje ovoga problema nije bilo potrebno prepoznati koja osoba se nalazi na fotografiji, nego samo je li na fotografiji ista osoba. Kao prvu fotografiju uzeli su fotografiju jedne osobe, dok su za drugu fotografiju uzeli fotografiju iste te osobe ili druge osobe. Potom su dvije slike uspoređene pomoću sijamske neuronske mreže. Kao rezultat dobivena je skalarna mjera različitosti između ulaznih podataka, odnosno ulaznog para slika, a definirana je kao:

$$E_w(X_1, X_2) = \|G_w(X_1) - G_w(X_2)\| \quad (2-2)$$

U formuli (2-2) X_1 i X_2 predstavljaju par slika, a rezultat ove jednadžbe $E_w(X_1, X_2)$ označava mjeru različitosti ulaznih podataka. Male vrijednosti mjere različitosti označavaju jednake parove slika, dok visoke mjere različitosti označavaju različite parove slika. Oznake $G_w(X_1)$ i $G_w(X_2)$ označavaju obrađene ulazne podatke X_1 i X_2 pomoću konvolucijske mreže

Obilježje sijamske neuronske mreže predložene u ovom radu je da se sastoji od dviju konvolucijskih pod mreža koje dijele težine. Arhitekturu sijamske neuronske mreže predstavljene u radu [4] možemo vidjeti na slici 2.13. Težine koje dijele dvije sijamske neuronske mreže, na slici 2.13 označene su s oznakom W .



Slika 2.13 Arhitektura sijamske neuronske mreže korištene u za usporedbu osoba na fotografijama [4]

Cilj je sijamske neuronske mreže približiti vektore značajki parova slika koje su označene kao jednake, a udaljiti vektore značajki parova slika koji su označene kao različite. Kako bi ovo postigli i kako bi ispravno trenirali sijamsku neuronsku mrežu, autori su uveli kriterijsku funkciju zvanu kontrastna (engl. *contrastive*) kriterijska funkcija:

$$L(W, (Y, X_1, X_2)^i) = (1 - Y)L_G(E_W(X_1, X_2)^i) + YL_I(E_W(X_1, X_2)^i), \quad (2-3)$$

gdje su:

- $(Y, X_1, X_2)^i$ – i -ti uzorak sastavljen od para slika i oznake Y (0 ako su slike jednake, 1 ako su slike različite)
- L_G - djelomična kriterijska funkcija para jednakih slika
- L_I – djelomična kriterijska funkcija para različitih slika

Navedena mreža trenirana je i testirana pomoću AT&T baze fotografija lica [6] i AR baze fotografija lica osoba [5]. Testiranje je izvedeno na 5000 parova fotografija AT&T baze fotografija i AR baze fotografija. Od 5000 parova fotografija, na 500 parova fotografija se nalazila ista osoba, a na 4500 parova fotografija se nalazila različita osoba. Testovi se izvode tako da se uzima fotografija prve osobe koja se definira kao referentna slika, druga fotografija nasumično se odabire iz iste baze fotografija i definira kao testna slika. Performanse mreže mjerene su računanjem postotka parova fotografija s različitim osobama koje je mreža klasificirala kao da prikazuju jednaku osobu i postotka parova fotografija s istom osobom koje je mreža klasificirala kao da prikazuje različite osobe. Na AT&T bazi fotografija lica osoba, uz postotak od 5% za neispravno

klasificirane parove slika s različitim osobama, dobiven je postotak od 1% za neispravno klasificirane parove slika s istom osobom. Na AR bazi fotografija lica osoba, uz postotak od 5% za neispravno klasificirane parove slika s različitim osobama, dobiven je postotak od 19% za neispravno klasificirane parove slika s istom osobom. [4]

Novu kriterijsku funkciju u svom su radu koristili Dey i ostali [20]. Oni su u svom radu nadogradili neuronsku mrežu za raspoznavanje je li potpis originalan ili je krivotvoren. Iskoristili su sposobnosti sijamske neuronske mreže u kombinaciji s konvolucijskim slojevima te pojednostavili formulu kriterijske funkcije. Ovu mrežu nazvali su *SigNet*, a pojednostavljena formula kriterijske funkcije glasi:

$$L(s_1, s_2, y) = \alpha(1 - y)D_w^2 + \beta y \max(0, m - D_w)^2 \quad (2-4)$$

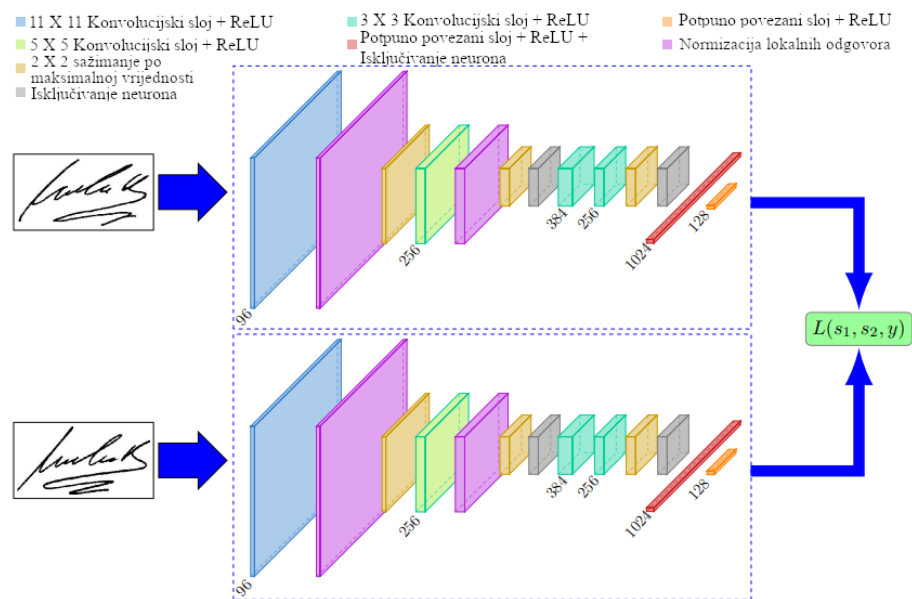
U formuli (2-4) s_1 i s_2 označavaju ulazni par slika. Ulazni par slika predstavljaju dvije slike od kojih prva slika, referentna slika, uvijek prikazuje originalni potpis. Druga slika ulaznog para slika predstavlja testnu sliku, odnosno potpis s kojim se originalni potpis uspoređuje. Potpis s kojim se originalni potpis uspoređuje može biti originalan ili krivotvoren. Oznaka y pokazuje je li druga ulazna slika ona originalnog potpisa ili krivotvorenog potpisa. Konstante su prikazane znakovima α i β , oznaka m označava marginu za koju je uzeta vrijednost 1, dok oznaka D_w označava euklidsku udaljenost značajki dobivenih obradom obiju slika potpisa pomoću konvolucijskog dijela mreže. Konvolucijski dio mreže sastoji se od 3 konvolucijska sloja, ReLU (engl. *Rectified linear activation unit*) aktivacijske funkcije, 3 sloja sažimanja te 2 potpuno povezana sloja (Slika 2.14).

Za izbjegavanje problema kada se na trening podacima dobivaju puno bolji rezultati nego na testnim podacima (engl. *overfitting*), koriste se regularizacijske metode, kao što su isključivanje neurona (engl. *dropout*) i normizacija (engl. *normalization*). Isključivanje neurona je popularna metoda regularizacije, gdje se određeni postotak neurona isključuje iz neuronske mreže te se time sprječava pretjerana usklađenost neuronske mreže na trening podataka. Normizacija pridonosi bržem procesu učenja i većoj stabilnosti mreže [21].

Regularizacijske metode korištene u neuronskoj mreži spomenutoj u radu [20] su metoda isključivanja neurona i normizacija lokalnih odgovora (engl. *Local Response Normalization, LRN*). Unutar konvolucijskog dijela neuronske mreže isključivanje neurona iznosi 30%, a u potpuno povezanom dijelu neuronske mreže isključivanje neurona iznosi 50%.

Mreža je testirana na CEDAR [8] bazi potpisa gdje su uspoređivani originalni i krivotvoreni potpisi te je postignuta točnost od 100%. Uz CEDAR bazu potpisa, *SigNet* mreža testirana je na

GPDS300 [11] skupu potpisa te je na ovom skupu originalnih i krivotvorenih potpisa postigla točnost od 88.79%. Na slici 2.14 prikazana je arhitektura *SigNet* mreže.



Slika 2.14 Arhitektura *SigNet* sijamske neuronske mreže [20]

Uspoređivanje potpisa i osoba problem je blizak uspoređivanju dinamičnih slika. Na prvoj slici može se nalaziti jedna osoba, dok se na drugoj slici može nalaziti ista ta osoba, ali fotografirana iz drugog kuta ili s drugačijim osvjetljenjem. To će i dalje biti ista osoba, dakle kontekst slike će biti isti, ali sadržaj će biti drugačiji. Kada bi se na drugoj slici pojavila fotografija potpuno druge osobe, tada bi se i kontekst i sadržaj na drugoj slici promijenili. Slična stvar događa se i tijekom uspoređivanja potpisa. Potpis na drugoj slici može biti onaj učinjen od strane iste osobe, no zbog drugačijeg pritiska olovke na papir, sadržaj slike na kojoj se potpis nalazi bit će različit, ali kontekst ostaje isti. Uzimanjem potpisa potpuno druge osobe i njegovim prikazivanjem na drugoj slici, mijenjamo i sadržaj i kontekst slike. Iz ovih razloga, za prijedlog rješenja u nastavku rada koristit će se sijamska neuronska mreža kao rješenje za uspoređivanje dinamičnih slika zasnovano na neuronskoj mreži.

3. IZRADA NEURONSKE MREŽE ZA USPOREDBU DINAMIČNIH SLIKA

Za uspješno rješavanje problema usporedbe dinamičnih slika pomoću neuronske mreže potrebna je neuronska mreža koja ima sposobnost primanja dviju slika, te njihovom obradom treba odrediti jesu li slike jednake ili različite. Kako je u prethodnom dijelu 2.2.2 najavljeno, kao tip neuronske mreže korištena je sijamska neuronska mreža. Ovaj tip neuronske mreže sastoji se od dva jednaka konvolucijska dijela koji kao rezultat obrade slika daju vektore značajki pomoću kojih se izračunava sličnost slika. Kako bi se uspješno trenirala sijamska neuronska mreža, bilo je potrebno napraviti novu bazu dinamičnih slika. Pomoću nove baze dinamičnih slika moguće je imati kontrolu nad transformacijama koje se događaju na objektima u slici. Iz tog razloga, napravljen je alat za generiranje dinamičnih slika, a pomoću njega generirana je vlastita baza dinamičnih slika koja je korištena za treniranje sijamske neuronske mreže. Algoritam za generiranje dinamičnih slika i neuronska mreža pisani su u programskom jeziku *Python* korištenjem programskog alata *Visual Studio Code*. Za pisanje algoritma i izradu neuronske mreže korištene su biblioteke *PyTorch* i *OpenCV* (engl. *Open Source Computer Vision Library*).

U nastavku ovog poglavlja prvo je opisano postavljanje radnog okruženja i instalacija potrebnih biblioteka. Zatim u potpoglavlju 3.2 opisan je algoritam za kreiranje baze dinamičnih slika te proces kreiranja baze dinamičnih slika korištenih za treniranje neuronske mreže. U potpoglavljima 3.3 i 3.4 objašnjeni su korištena arhitektura sijamske neuronske mreže te proces treniranja mreže. Posljednje potpoglavlje 3.5 daje upute za korištenje algoritma za kreiranje baze slika za treniranje, kao i upute za korištenje sijamske neuronske mreže.

3.1. Podešavanje radnog okruženja i instalacija biblioteka

Prvi korak koji je bilo potrebno napraviti pri izradi predloženog rješenja za usporedbu dinamičnih slika zasnovanog na neuronskim mrežama je instalirati programski alat *Visual Studio Code*, odgovarajuću verziju *Python* programskog jezika i potrebne biblioteke, *OpenCV* i *PyTorch*. *Python* verzija koja se koristi u radu je 3.8.12.

OpenCV [22] je biblioteka otvorenog koda sastavljena od funkcija koje se koriste u području računalnog vida i strojnog učenja. U biblioteci se nalazi preko 2500 optimiziranih algoritama koji mogu pomoći u obradi slika, a između ostalog u detektiranju i prepoznavanju lica, identificiranju objekata, praćenju pokreta objekata itd. *OpenCV* biblioteka napisana je u *C* i *C++* programsku jeziku. Prevedena je za programske jezike kao što su *Python* i *Java* te ima podršku za *Windows*, *Linux*, *Android* i *Mac OS*. Za potrebe ovog diplomskog rada korištena je *OpenCV* verzija 4.2.0.34,

a instalira se pomoću naredbe koja se unosi u komandnu liniju (engl. *Command Line Interface, CLI*). Naredba za instalaciju *OpenCV-a* glasi :

```
pip install opencv-python==4.2.0.34
```

PyTorch [23] predstavlja biblioteku otvorenog koda namijenjenu za razvoj i treniranje neuronskih mreža zasnovanih na modelima dubokog učenja. Namijenjena je za korištenje pomoću *Python* programskog jezika, ali se može koristiti i pomoću *C++* programskog jezika. Omogućava korištenje preko 200 različitih matematičkih operacija. Pomoću ove biblioteke može se stvoriti vlastita neuronska mreža sa svim potrebnim slojevima, istrenirati je i testirati. Uz *PyTorch*, za proces treniranja koristi se platforma za paralelno izračunavanje koja se naziva *CUDA* (engl. *Compute Unified Device Architecture*). *CUDA* platformu razvila je tvrtka *NVIDIA* kako bi omogućila izračunavanja na grafičkim procesorskim jedinicama. Korištenjem *CUDA* platforme znatno se ubrzava potrebno vrijeme treniranja neuronskih mreža. Verzija *PyTorch* biblioteke koja je korištena u ovom radu je 1.9.0, a instalirana je pomoću upisivanja ove naredbe u komandnu liniju:

```
pip install pytorch==1.9.0 torchvision==0.10.0 torchaudio==0.9.0 cudatoolkit=11.3 -c  
pytorch -c conda-forge
```

3.2. Kreiranje vlastitih baza dinamičnih slika za treniranje neuronske mreže

Za treniranje neuronske mreže potrebna je prikladna baza dinamičnih slika. Prema vlastitom saznanju, trenutno ne postoji baza dinamičnih slika, osim CGIAD baze dinamičnih slika [1], koja bi mogla poslužiti pri razvijanju algoritama i neuronskih mreža za funkcionalno testiranje aplikacija u kojima se pojavljuju dinamične slike. Baze fotografija lica i slika potpisa, spomenute u potpoglavlju 2.1, ne mogu se koristiti za rješavanje problema usporedbe dinamičnih slika jer se na njima nalazi samo jedan specifičan kontekst, lice osobe ili potpis osobe. Zbog tog razloga, u sklopu ovog rada napisan je algoritam pomoću kojega se može napraviti vlastita baza dinamičnih slika koja može sadržavati više različitih konteksta. Još jedan od razloga zbog kojega je kreiran alat za generiranje dinamičnih slika je taj što se pomoću ovog alata može imati bolja kontrola transformacija koje se rade na objektima koji se nalaze na slikama te time imati bolju kontrolu tijekom treniranja neuronske mreže.

3.2.1. Algoritam za generiranje dinamičnih slika

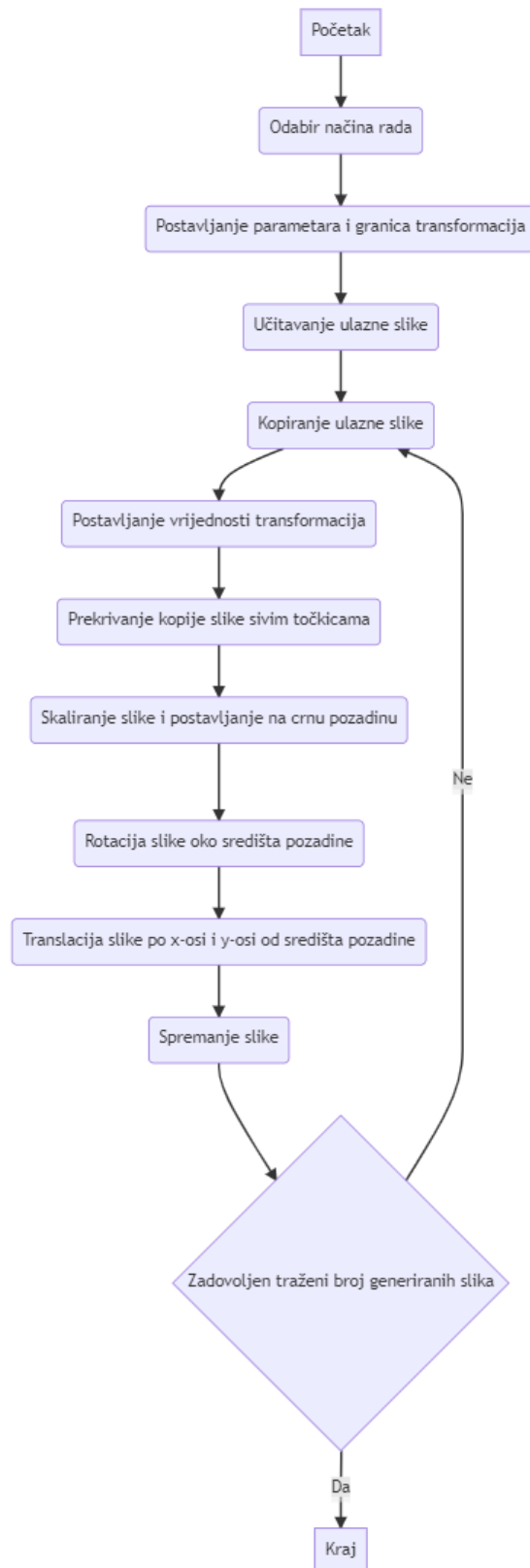
Algoritam za generiranje dinamičnih slika izrađen je u programskom jeziku *Python* i koristi *OpenCV* biblioteku za obradu slika. Algoritam prima jednu referentnu sliku te, izvršavajući različite transformacije na cijeloj slici, stvara nove slike. Transformacije koje se izvršavaju su:

- Rotacija referentne slike za određeni broj stupnjeva
- Skaliranje referentne slike za određeni postotak njezine originalne veličine
- Postavljanje referentne slike u središte pozadine novih većih dimenzija
- Translacija referentne slike od središta slike po x-osi za određeni broj elemenata slike
- Translacija referentne slike od središta slike po y-osi za određeni broj elemenata slike
- Prekrivanje određenog postotka referentne slike pomoću točkica.

Slika na izlazu iz algoritma istovremeno može biti zarotirana, skalirana, translatairana i dio slike može bi zaklonjen pomoću sivih točkica. Ovim transformacijama pokušavaju se simulirati animacije koje bi se mogle pojaviti na grafičkom sučelju aplikacije koja se testira u stvarnom svijetu.

Algoritam može raditi u dva načina rada. Prvi je način rada da transformacije izvodi linearno od donje granice za svaku transformaciju do gornje granice za svaku transformaciju. Algoritam podiže vrijednosti transformacija za određeni broj koraka koji je drugačiji za svaku transformaciju. Na primjeru transformacije rotacije, izvodi se rotacija od -20° do $+20^\circ$, i to svakih 5° . Drugi je način rada da za određeni broj ukupnih slika, algoritam izvede transformacije tako da nasumično odabere parametre transformacija, uz pridržavanje postavljenih granica.

Kada se pokrene algoritam pozivanjem modula *image_generator.py*, čitaju se uneseni parametri i putanja do referentne slike na kojoj korisnik želi da se izvrše transformacije. Referentna slika mora biti veličine manje od 500x500 elemenata slike. Preporučena veličina referentne slike je oko 300x300 elemenata slike kako bi se smanjila vjerojatnost događaja da referentna slika zbog transformacija „izlazi“ s generirane slike. Pročitani podaci spremaju se u varijable koje se zatim predaju kao parametri klase *ImageGenerator*, koja se nalazi u modulu *image_generator_class.py*. Algoritam poziva funkciju *run_gen* i prvo postavlja vrijednosti za svaku od transformacija. Nakon toga počinju se generirati transformirane slike. Na slici 3.1 prikazan je dijagram toka algoritma za generiranje dinamičnih slika.



Slika 3.1 Dijagram toka algoritma za generiranje dinamičnih slika

Referentna se slika prvo kopira te se sve transformacije rade na kopiji referentne slike koja se naziva generirana slika. Redoslijed izvođenja transformacija je okluzija sivim točkicama, skaliranje, rotacija pa translacija. Tijekom koraka prekrivanja slike sivim točkicama, na generiranu

sliku prvo se dodaje okluzija u obliku točkica. Količinu okluzije izražavamo u postotcima i oni govore koliki postotak površine slike je prekriven sivim točkicama. Nakon koraka skaliranja, algoritam postavlja generiranu sliku s okluzijom u središte crne pozadine veličine 500x500 elemenata slike. Pomoću postavljanja slike na takvu pozadinu omogućuje se da sve generirane slike budu jednakih dimenzija nakon izlaska iz algoritma. Nakon postavljanja generirane slike s okluzijom i skaliranjem na crnu pozadinu, dodaje se rotacija. Cijela generirana slika se zarotira za određeni broj stupnjeva oko središta slike. Negativne vrijednosti rotacije rotiraju generiranu sliku u smjeru kazaljke na satu, dok pozitivne vrijednosti rotiraju generiranu sliku u smjeru suprotnom od kazaljke na satu. Korak translacije pomiče generiranu sliku u određenom smjeru od središta generirane slike za određeni broj elemenata slike. Pozitivne vrijednosti translacije po x-osi pomiču generiranu sliku desno od središta generirane slike, dok negativne vrijednosti pomiču generiranu sliku lijevo od središta generirane slike. Isto tako, pozitivne vrijednosti translacije po y-osi pomiču generiranu sliku prema dolje od središta generirane slike, a negativne vrijednosti translacije po y-osi pomiču generiranu sliku prema gore od središta generirane slike. Tijekom svih transformacija nakon skaliranja, generirana slika zadržava svoju veličinu od 500x500 elemenata slike. Nakon izvedenih transformacija slika se sprema na putanju koju je odredio korisnik, a sprema se pod nazivom:

gen_img_S_(postotak skaliranja originalne veličine)_R_(stupnjevi rotacije)_Tx_(vrijednost translacije po X-osi)_Ty_(vrijednost translacije po Y-osi)_O_(postotak prekrivenosti šumom).png.

Naziv *gen_image* označava da se radi o generiranoj slici pomoću ovog algoritma i svim slikama taj dio naziva je jednak. Nazivi generiranih slika razlikuju se u nastavku naziva gdje su opisane vrijednosti izvedenih transformacija.

Transformacije se izvode u granicama koje korisnik postavlja prilikom pokretanja ili ako korisnik ne navede granice, koriste se inicijalne granice koje su postavljene u modulu *generator_config.py*. Vrijednosti inicijalnih granica:

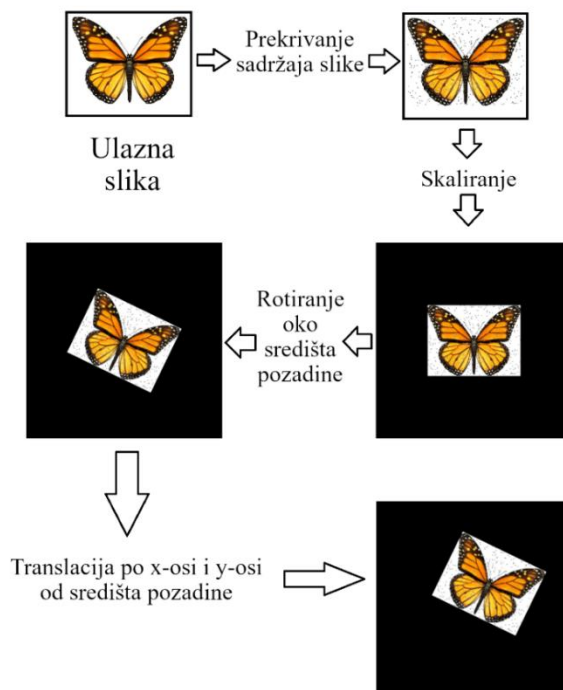
- Širina izlazne slike - 500 elemenata slike
- Visina izlazne slike - 500 elemenata slike
- Boja točkica kojima se postiže okluzija, boja opisana u vrijednostima od 0 do 255 za svaku komponentu crvene, zelene i plave (engl. *red, green, blue, RGB*) - 128, 128, 128
- Način rada generatora slika - broj 0 označava nasumičan način rada, a broj 1 označava linearan način rada. Inicijalna je vrijednost 0
- Broj generiranih slika u nasumičnom načinu rada - 100

- Minimalni kut rotacije slike - -40°
- Maksimalni kut rotacije slike - $+40^\circ$
- Donja granica skaliranja slike - 80%
- Gornja granica skaliranja slike - 120%
- Donja granica površine slike na kojoj se događa okluzija - 0%
- Gornja granica površine slike na kojoj se događa okluzija - 10%
- Donja granica pomicanja slike po x-osi od središta - 40 elemenata slike ulijevo od središta slike
- Gornja granica pomicanja slike po x-osi od središta - 40 elemenata slike udesno od središta slike
- Donja granica pomicanja slike po y-osi od središta - 40 elemenata slike prema gore od središta slike
- Gornja granica pomicanja slike po y-osi od središta - 40 elemenata slike prema dolje od središta slike
- Inkrement povećanja kuta rotacije slike - 20
- Inkrement povećanja postotka skaliranja slike - 20
- Inkrement povećanja površine slike na kojoj se događa okluzija - 2
- Inkrement povećanja pomaka slike po x-osi od središta slike - 20 elemenata slike
- Inkrement povećanja pomaka slike po y-osi od središta slike - 20 elemenata slike

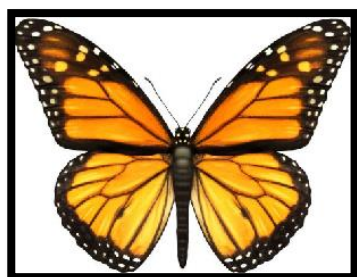
Primjer procesa generiranja dinamične slike i krajnji rezultat generiranja dinamične slike vidljiv je na slici 3.2 i slici 3.3, a njen naziv je *gen_img_S_80.0_R_-15_Tx_22_Ty_11_O_1.7*.

Vrijednosti njenih transformacija su:

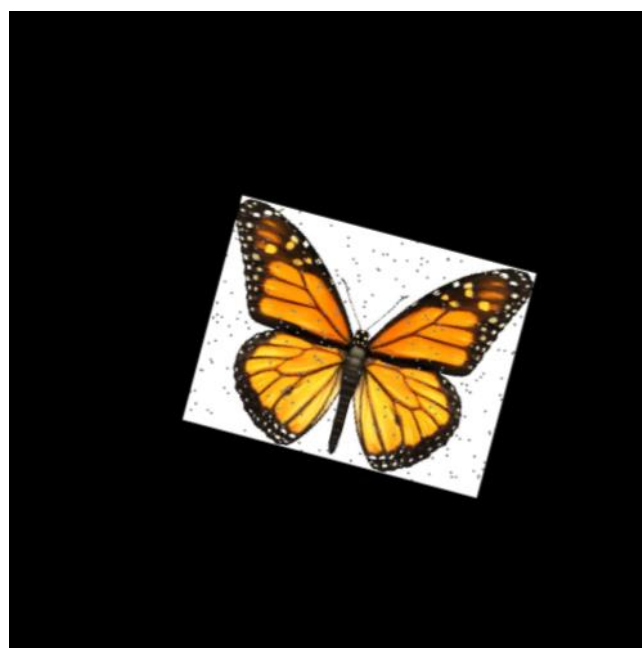
- Skaliranje - generirana slika je skalirana na 80 % od veličine referentne slike
- Rotacija - generirana slika je zakrenuta za 15°
- Pomak slike po x-osi od središta slike - generirana slika je pomaknuta za 22 elemenata slike udesno od središta slike
- Pomak slike po y-osi od središta slike - generirana slika je pomaknuta za 11 elemenata slike dolje od središta slike
- Okluzija - 1,7% površine generirane slike se događa okluzija



Slika 3.2 Primjer procesa generiranja dinamične slika. Generirana slika ima vrijednosti transformacija: okluzija 1,7%, skaliranje 80%, rotacija 15° oko središta pozadine, translacija po x-osi 22 elementa slike prema desno od središta pozadine, translacija po y-osi 11 elemenata slike prema gore od središta pozadine



(a)



(b)


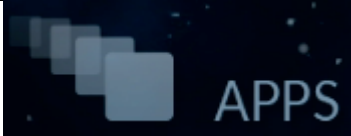



Slika 3.3 Primjer transformirane slike pomoću vlastitog kreiranog algoritma za transformaciju slika: (a) referentna slika, (b) slika generirana na osnovu referentne slike (skaliranje 80%, rotacija 15°, translacija x-os 22 elementa slike, translacija y-os 11 elemenata slike, okluzija 1,7%)








U prilogu P.3.1. na elektroničkom mediju priloženom uz ovaj rad nalazi se programski kod napisan za potrebe rada algoritma za generiranje dinamičnih slika sa svim potrebnim datotekama.

3.2.2. Generirane baze dinamičnih slika za treniranje neuronske mreže

Za treniranje sijamske neuronske mreže generirane su dvije baze dinamičnih slika. Jedna generirana baza dinamičnih slika korištena je kako bi se provjerilo može li se sijamska neuronska mreža koristiti kao univerzalno rješenje za usporedbu dinamičnih slika. Druga baza dinamičnih slika generirana je kako bi se provjerilo može li se sijamska neuronska mreža koristiti kao specifično rješenje za usporedbu dinamičnih slika. Prva generirana baza dinamičnih slika naziva se *CG-STB-DI* (engl. *Custom generated STB dynamic images*). Slike od kojih je napravljena *CG-STB-DI* baza slika izabrane su iz baze slika STB grafičkog sučelja [2]. Izabrano je 12 slika (Tablica 3.1), koje prikazuju po jedan izbornik kada je taj izbornik izabran ((a), (c), (e), (g), (i), (k)) i kada taj izbornik nije izabran ((b), (d), (f), (h), (j), (l)). Pomoću izabranih slika pokušava se postići da neuronska mreža raspoznaje kada je izbornik odabran, a kada isti taj izbornik nije odabran.

Tablica 3.1 Izabrane slike baze slika STB grafičkog sučelja

Ime ulazne slike	Ulazna slika	Veličina slike u broju elemenata slike
(a) <i>apps_selected.png</i>		168x64
(b) <i>apps_not_selected.png</i>		174x70
(c) <i>media_selected.png</i>		207x79
(d) <i>media_not_selected.png</i>		198x77
(e) <i>search_selected.png</i>		165x64

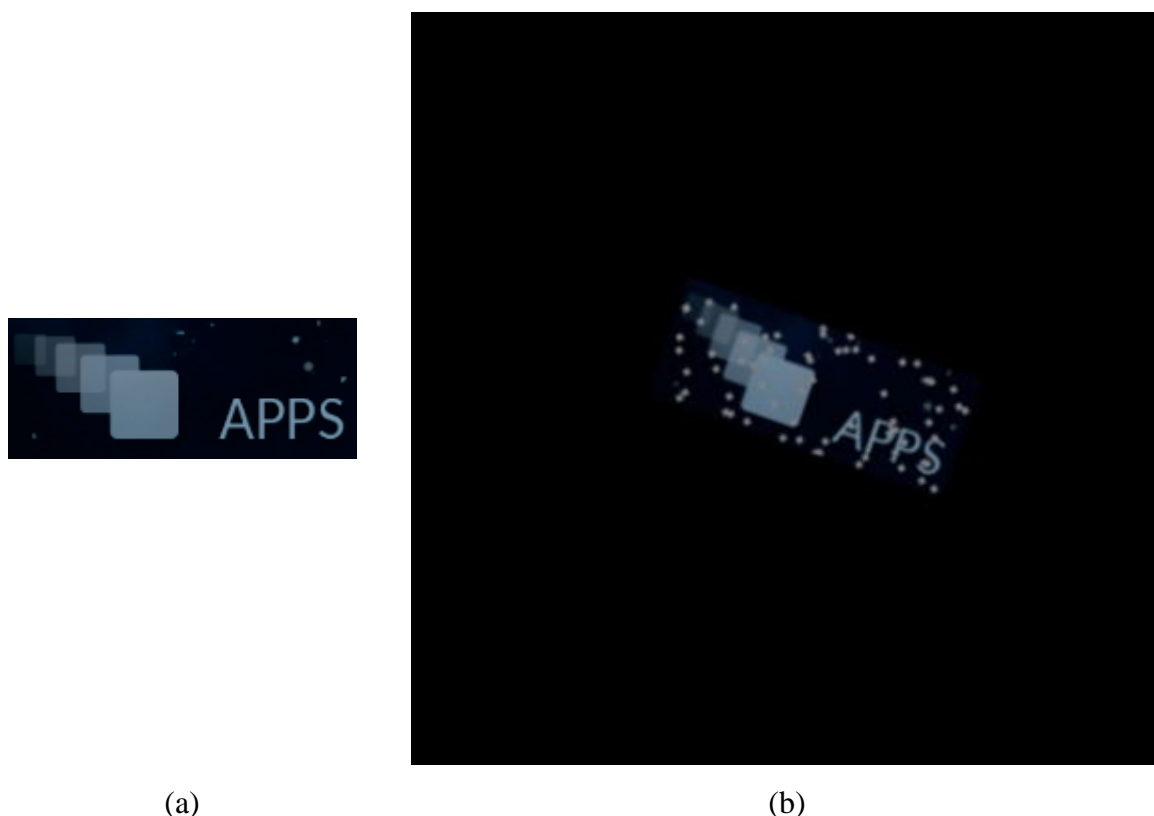
(f) <i>search_not_selected.png</i>		166x66
(g) <i>settings_selected.png</i>		215x53
(h) <i>settings_not_selected.png</i>		190x53
(i) <i>tv_guide_selected.png</i>		191x72
(j) <i>tv_guide_not_selected.png</i>		190x65
(k) <i>vod_selected.png</i>		162x63
(l) <i>vod_not_selected.png</i>		162x56

Sve slike prikazane u tablici 3.1 ubačene su u algoritam za generiranje dinamičnih slika. Algoritam je pokrenut u drugom načinu rada gdje se vrijednosti transformacija uzimaju nasumično te je postavljeno da se generira 170 slika za svaku od prikazanih slika u tablici 3.1. To znači da je za *CG-STB-DI* bazu dinamičnih slika ukupno generirano 2040 slika ($12 \times 170 = 2040$), 1020 slika ($6 \times 170 = 1020$) kada su izbornici odabrani i 1020 slika ($6 \times 170 = 1020$) kada izbornici nisu odabrani. Pomoću ove generirane baze dinamičnih slika provjerava se može li sijamska neuronska mreža raditi kao univerzalno rješenje. Svaka generirana slika na sebi sadrži sve transformacije u isto vrijeme. Razlog malog broja generiranih slika je taj što zbog načina rada sijamske neuronske mreže, njoj nije potrebna velika količina trening podataka za postizanje visoke preciznosti, za razliku od klasičnih neuronskih mreža kada su potrebne velike količine podataka tokom treninga kako bi se postigla visoka preciznost. Generirane slike su veličine 500x500 elemenata slike, a postavljene granice za svaku od transformacija su:

- Skaliranje – od 80% veličine referentne slike do 120% veličine referentne slike;
- Rotacija – rotacija od -40° do $+40^\circ$ oko središta slike;
- Translacija po x-osi – od -40 elemenata slike do +40 elemenata slike od središta;

- Translacija po y-osi – od -40 elemenata slike do +40 elemenata slike od središta;
- Okluzija – od 0% prekrivenosti slike do 10% ukupne prekrivenosti referentne slike točkicama.

Primjer transformirane slike za ulaznu sliku prikazanu u tablici 3.1 (b) nalazi se na slici 3.4. Slika je generirana u načinu rada algoritma gdje se nasumično biraju vrijednosti transformacija koje se nalaze unutar postavljenih granica za svaku od transformacija.






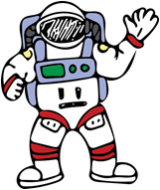





Slika 3.4 Primjer transformirane slike iz tablice 3.1 (slika (b)) pomoću algoritma za transformaciju slika: (a) referentna slika, (b) generirana slika - skaliranje 95,1%, rotacija 19,3°, translacija x-os 14 elemenata slike, translacija y-os – 0 elemenata slike, okluzija 2,2%)

Druga generirana baza dinamičnih slika koristi slike iz prve generirane baze dinamičnih slika uz dodatak 10 slika iz CGIAD baze dinamičnih slika [1]. Naziv druge baze generiranih slika je *CG-STB-CGIAD-DI* (engl. *Custom generated STB CGIAD dynamic images*) Odabrane slike ubačene su u generator dinamičnih slika, a prikazane su u tablici 3.2.

Tablica 3.2 Izabrane slike CGIAD baze dinamičnih slika

Ime ulazne slike	Ulazna slika	Veličina slike
(a) <i>referent_image_1.png</i>		499x116

(b) <i>referent_image_2.png</i>		428x190
(c) <i>referent_image_3.png</i>		137x197
(d) <i>referent_image_4.png</i>		182x140
(e) <i>referent_image_5.png</i>		432x183
(f) <i>referent_image_6.png</i>		352x199
(g) <i>referent_image_7.png</i>		170x199
(h) <i>referent_image_8.png</i>		355x199
(i) <i>referent_image_9.png</i>		181x171

(j) <i>referent_image_10.png</i>		196x136
----------------------------------	--	---------

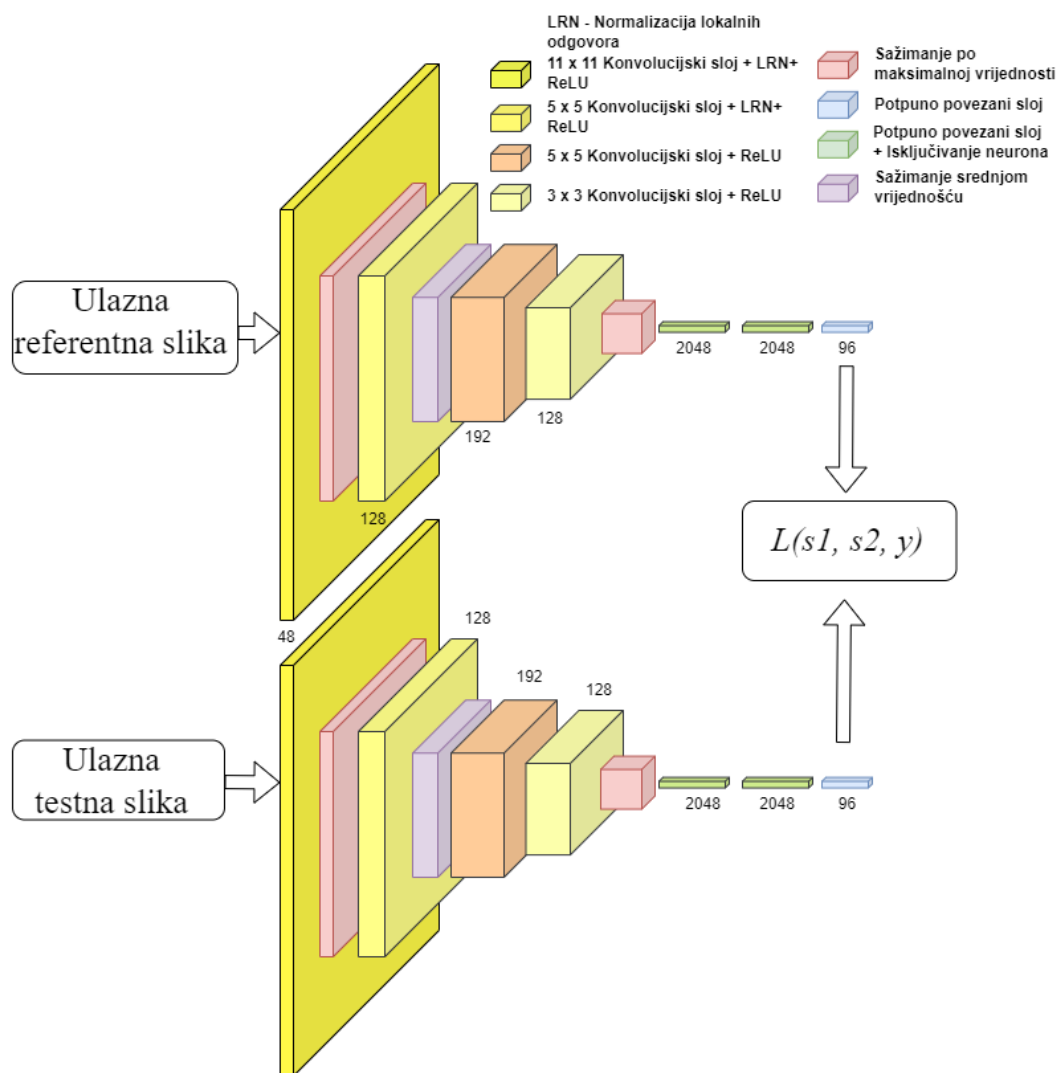
Kao kod *CG-STB-DI* baze dinamičnih slika, za svaku od prikazanih slika u tablici 3.2, generirano je 170 slika ($10 \times 170 = 1700$). Ukupni broj slika druge baze generiranih dinamičnih slika iznosi 3740 ($2040 + 1700$). Pomoću ove baze dinamičnih slika provjerava se može li sijamska neuronska mreža raditi kao specifično rješenje. Slike su generirane u načinu rada algoritma kada se vrijednosti transformacija biraju nasumično unutar postavljenih granica za svaku transformacija. Generirane su slike veličine 500x500 elemenata slika, a granice za svaku od transformacija su sljedeće:

- Skaliranje – od 80% veličine referentne slike do 120% veličine referentne slike
- Rotacija – rotacija od -40° do $+40^\circ$ oko središta slike;
- Translacija po x-osi – od -40 elemenata slike do +40 elemenata slike od središta;
- Translacija po y-osi – od -40 elemenata slike do +40 elemenata slike od središta;
- Okluzija – od 0% prekrivenosti slike do 10% ukupne prekrivenosti referentne slike točkicama

U prilogu P.3.2. na elektroničkom mediju priloženom uz ovaj rad nalaze se *CG-STB-DI* i *CG-STB-CGIAD-DI* baze dinamičnih slika generiranih za treniranje sijamske neuronske mreže.

3.3. Arhitektura sijamske neuronske mreže

Problem uspoređivanja dinamičnih slika zahtjeva da neuronska mreža mora funkcionirati u različitim okruženjima i postoji mogućnost da će često u usporedbi nailaziti na objekte koje nikad nije vidjela. To je jedan od razloga zašto je kao prijedlog za rješavanje problema usporedbe dinamičnih slika odabrana sijamska neuronska mreža. Prednost je sijamske neuronske mreže što tijekom treniranja uči kako razlikovati jesu li slike jednake ili različite. Još je jedna od prednosti sijamske neuronske mreže što ju nije potrebno ponovno trenirati kada se za vrijeme testiranja pojave novi objekti. Iz ovog razloga sijamska neuronska mreža praktična je za korištenje za predmetni problem. Arhitektura sijamske neuronske mreže predložene u ovom radu prikazana je na slici 3.5.



Slika 3.5 Arhitektura sijamske neuronske mreže korištene u ovom radu za problem usporedbe dinamičnih slika

Prikazana arhitektura sijamske neuronske mreže odabrana je nakon testiranja više verzija sijamske neuronske mreže s drugačijim arhitekturama. Izabrana arhitektura sijamske neuronske mreže dala je najbolje rezultate tijekom testiranja.

Ulazi u sijamsku neuronsku mrežu su dvije slike, referentna slika i testna slika. Slike se obrađuju pomoću konvolucijskih i potpuno povezanih slojeva mreže. Na izlazu neuronske mreže dobivaju se vektori značajki za svaku od ulaznih slika. Izlazni vektori značajki na slici arhitekture mreže (Slika 3.5) prikazani su s oznakama $s1$ i $s2$. Oznaka $s1$ vektor je značajki dobiven obradom referentne slike, a oznaka $s2$ vektor je značajki dobiven obradom testne slike. Oznaka y oznaka je koja ima značenje da li slike iz para slika pripadaju istoj klasi ili ne, odnosno prikazuju li jednake objekte ili ne. Ove oznake ulazni su podatci za kriterijsku funkciju korištenu tijekom treninga $L(s1, s2, y)$ (2-4). Tijekom testiranja, oznake $s1$ i $s2$ ubacuju se u formulu za izračun euklidske udaljenosti vektora značajki. Kao izlaz iz sijamske neuronske mreže ne dobivaju se vjerojatnosti

predikcije, kao što je slučaj kod većine konvolucijskih neuronskih mreža, nego se dobiva udaljenost između dviju klasa. Tijekom treniranja sijamsku neuronsku mrežu se uči da normalizira udaljenost između različitih klasa prema 1. Ova udaljenost između dviju klasa predstavlja rezultat različitosti. Tijekom testiranja kada se uspoređuju potpuno drugačije slike može se dobiti udaljenost koja je i puno veća od 1. Iz tog razloga dobivena udaljenost između klasa, odnosno rezultat različitosti može imati vrijednost od 0 do beskonačno. Što je dobiveni rezultat različitosti veći to znači da su slike različitije.

Prvi konvolucijski sloj sastoji se od 48 filtara dimenzije 11x11, s korakom 4. Drugi konvolucijski sloj koristi 128 filtara dimenzije 5x5 s korakom 2. Treći konvolucijski sloj ima 192 filtara dimenzija 3x3 s korakom jedan. Posljednji konvolucijski sloj sastoji se od 128 filtara dimenzija 3x3 s korakom jedan. Broj koraka kod konvolucijskih filtara označava za koliko će se elemenata slike pomaknuti filtar konvolucije. Nakon svakog konvolucijskog sloja koristi se ReLU (engl. *Rectified Linear Activation Unit*) aktivacijska funkcija.

Sijamska neuronska mreža predložena u ovom radu koristi 2 tipa slojeva sažimanja. Prvi tip sloja sažimanja je sažimanje po maksimalnoj vrijednosti. Prvi i treći sloj sažimanja u arhitekturi neuronske mreže predložene u ovom radu koristi sažimanje po maksimalnoj vrijednosti, blok veličine 3x3 s korakom (engl. *stride*) 2. Broj koraka označava za koliko će se elemenata slike pomaknuti operator sažimanja. Drugi tip sloja sažimanja koristi prosjek svih vrijednosti unutar bloka sažimanja, a naziva se sažimanje po srednjoj vrijednosti. Drugi sloj sažimanja koristi sažimanje po srednjoj vrijednosti blok veličine 3 s korakom 2.

Nakon konvolucijskog dijela neuronske mreže, nalazi se potpuno povezani dio sijamske neuronske mreže koji se sastoji od 3 potpuno povezana sloja neurona. Prvi i drugi potpuno povezni sloj sastoje se od po 2048 neurona. Posljednji potpuno povezani sloj sastoji se od 96 izlaznih neurona. Iza svakog potpuno povezanog sloja nalazi se ReLU aktivacijska funkcija.

Za sprječavanje pretreniranja neuronske mreže koriste se metoda isključivanja neurona i metoda normizacije podataka. Metoda isključivanja neurona koristi se u potpuno povezanom dijelu neuronske mreže, a iznosi 50%. Metoda normizacije podataka koja se koristi u ovom radu je normizacija lokalnih rezultata (engl. *Local Response Normalization, LRN*) te se nalazi u konvolucijskom dijelu neuronske mreže, točnije nakon prvog i drugog konvolucijskog sloja.

3.4. Treniranje sijamske neuronske mreže

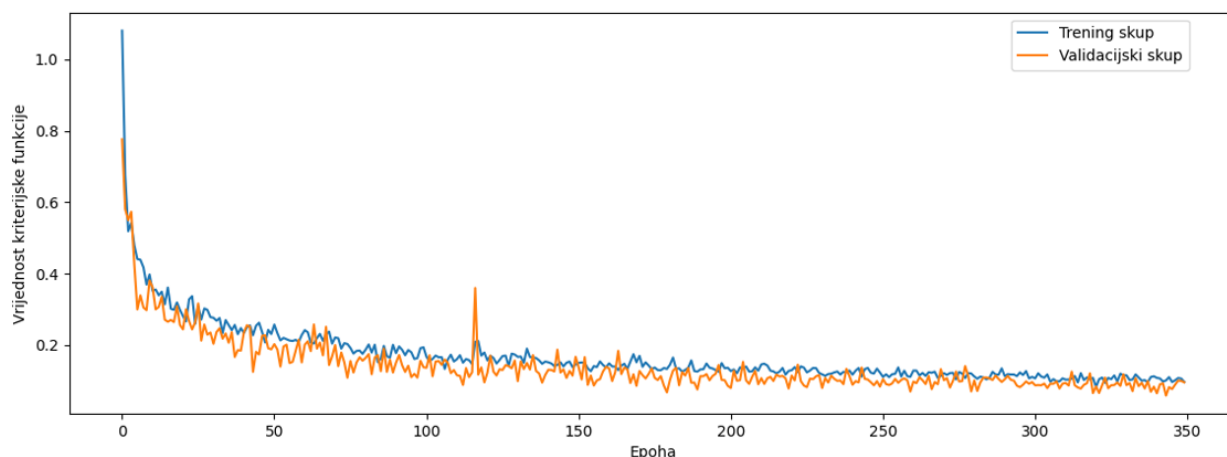
Treniranje sijamske neuronske mreže vršeno je koristeći baze dinamičnih slika napravljenih pomoću alata za generiranje transformiranih slika. Korištene baze transformiranih slika opisana su u dijelu 3.2.2. Na početku procesa treniranja neuronske mreže, baze transformiranih slika

podijeljene su u dva dijela, dio za treniranje i dio koji će biti korišten za validaciju rada sijamske neuronske mreže. Odnos dijela za treniranje i dijela korištenog za validaciju je 80% naprema 20%. Ovaj odnos znači da je za *CG-STB-DI* bazu od ukupno 2040 slika, 1632 korišteno je za treniranje neuronske mreže, a ostalih 408 slika korišteno je za validaciju. Pomoću *CG-STB-DI* baze dinamičnih slika sijamska neuronska mreža trenirana je kao univerzalno rješenje za usporedbu dinamičnih slika. Kako bi se sijamska neuronska mreža koristila kao specifično rješenje trenirana je pomoću *CG-STB-CGIAD-DI* baze dinamičnih slika baze dinamičnih slika. Kod treniranja sijamske neuronske mreže pomoću druge generirane baze dinamičnih slika, 2992 slike su korištene za trening, dok je 748 slika korišteno za validaciju. Za rad sa sijamskom neuronskom mrežom koriste se slike u boji.

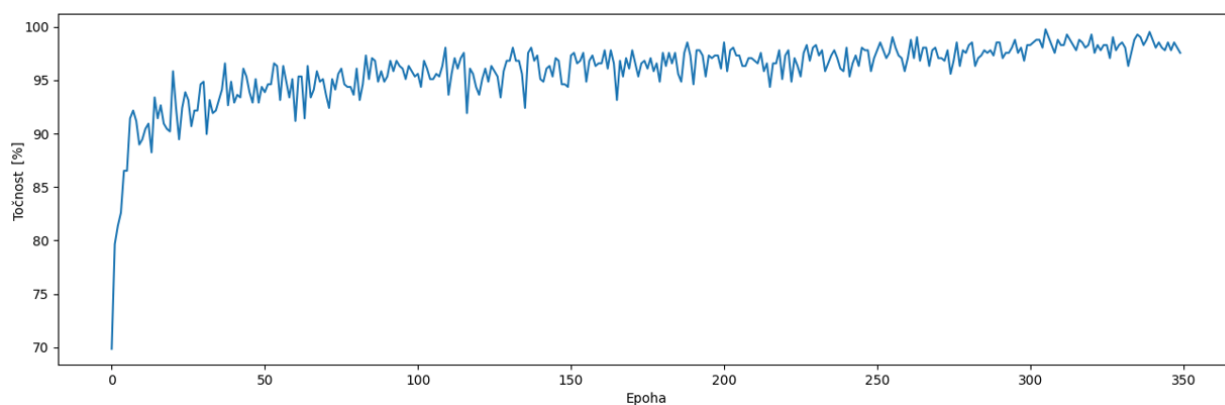
Pri pokretanju algoritma za treniranje sijamske neuronske mreže, od slika koje su dio trening seta prave se parovi slika koje se uspoređuju. Uz napravljene parove, svaki par dobiva oznaku 0 ili 1. Ako obje slike para slika pripadaju istoj klasi dodjeljuje se oznaka 0, a ako jedna slika od para slika pripada različitoj klasi, dodjeljuje se oznaka 1. Pomoću ovih oznaka definira se da je udaljenost između klasa jednaka vrijednosti 1. Razlog ovakvih oznaka je taj što kada se na slici nalaze jednaki objekti, euklidska udaljenost njihovih vektora značajki teži prema 0, a kada su objekti na slikama različiti, euklidska udaljenost njihovih vektora značajki teži prema vrijednostima 1 ili više. Kako je za treniranje potrebna samo informacija jesu li slike iste ili različite, koristit će se vrijednosti 0 i 1. Nakon svake epohe treniranja izvršena je validacija neuronske mreže koristeći parove validacijskog seta slika. Validacijski set slika sadrži slike koje nisu bile korištene za samo treniranje sijamske neuronske mreže, a slične su slikama koje su korištene u trening skupu.

Tijekom treniranja koristi se promjenjiva stopa učenja (engl. *Learning Rate*), koja se mijenja nakon određenog broja epoha. Stopa učenja se mijenja na način da nakon određenog broja epoha, trenutna stopa učenja se množi s postavljenim faktorom. Kao kriterijska funkcija, koristi se kontrastna kriterijska funkcija koja je opisana formulom (2-4), a kao algoritam optimizacije koristi se Adam algoritam [24].

U radu su trenirana dva modela sijamske neuronske mreže. Jedan model sijamske neuronske mreže trenirana je na *CG-STB-DI* bazi dinamičnih slika, a drugi model sijamske neuronske mreže treniran je na *CG-STB-CGIAD-DI* bazi dinamičnih slika. Oba modela sijamske neuronske mreže trenirani su s različitim trening parametrima. Na slici 3.6 i slici 3.7 prikazani su graf vrijednosti kriterijske funkcije i graf točnosti na validacijskom skupu tijekom treninga sijamske neuronske mreže na *CG-STB-DI* bazi dinamičnih slika.



Slika 3.6 Promjena vrijednosti kriterijske funkcije kroz 350 epoha treniranja sijamske neuronske mreže na *CG-STB-DI* bazi dinamičnih slika

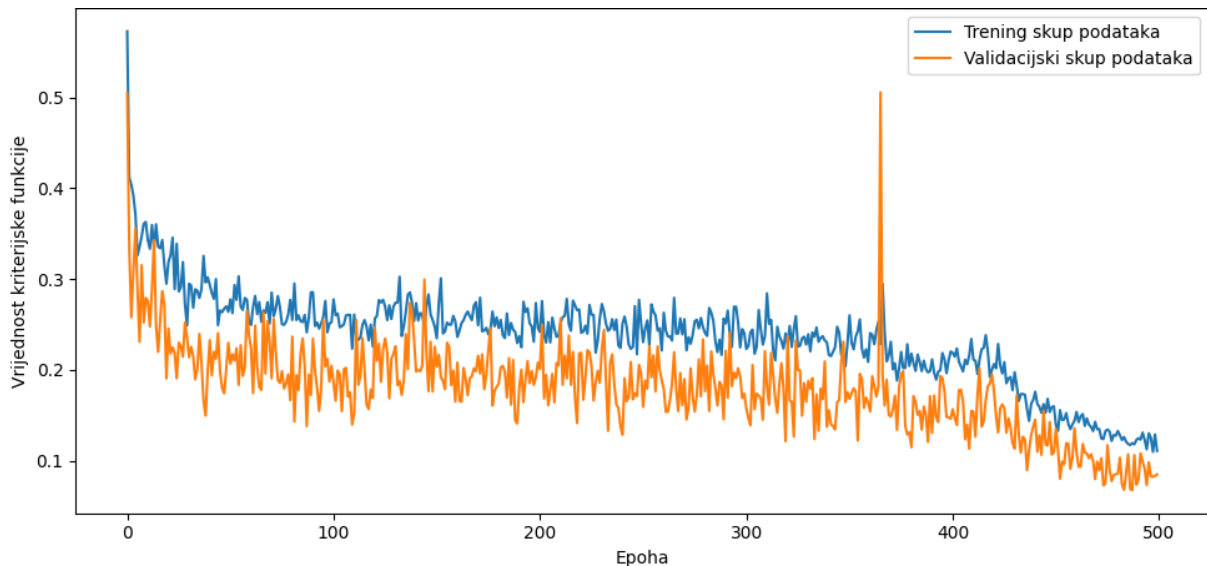


Slika 3.7 Točnost na validacijskom skupu podataka postignuta kroz 350 epoha treniranja sijamske neuronske mreže na *CG-STB-DI* bazi dinamičnih slika

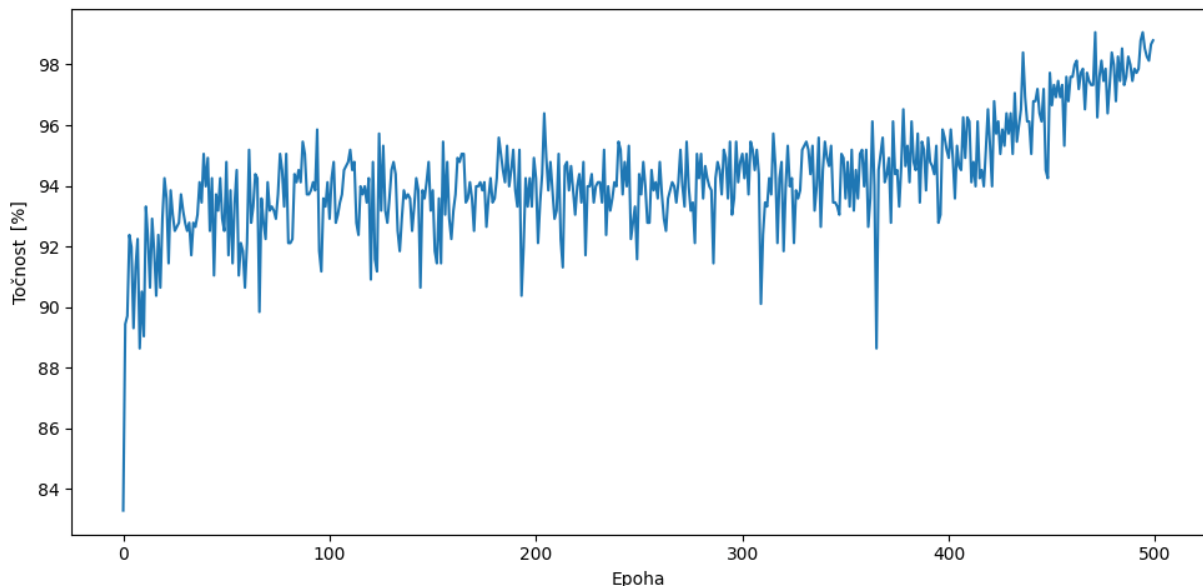
Tijekom treniranja sijamske neuronske mreže na *CG-STB-DI* bazi dinamičnih slika korišteni su ovi trening parametri:

- Stopa učenja - 0,0003
- Veličina serije (engl. *Batch size*) - 16
- Broj epoha - 350
- Smanjenje stope učenja za 0,9 puta nakon svakih 75 epoha
- Isključivanje neurona (engl. *Dropout*) – 50%

Na slici 3.8 i slici 3.9 prikazani su graf vrijednosti kriterijske funkcije i graf preciznosti na validacijskom skupu tijekom treninga sijamske neuronske mreže na *CG-STB-CGIAD-DI* bazi dinamičnih slika.



Slika 3.8 Promjena vrijednosti kriterijske funkcije kroz 500 epoha treniranja sijamske neuronske mreže na *CG-STB-CGIAD-DI* bazi dinamičnih slika



Slika 3.9 Točnost na validacijskom skupu podataka postignuta kroz 500 epoha treniranja sijamske neuronske mreže *CG-STB-CGIAD-DI* bazi dinamičnih slika

Tijekom treniranja sijamske neuronske mreže na *CG-STB-CGIAD-DI* bazi dinamičnih slika korišteni su ovi trening parametri:

- Stopa učenja - 0,0005
- Veličina serije (engl. *Batch size*) - 32
- Broj epoha - 500
- Smanjenje stope učenja za 0,9 puta nakon svakih 75 epoha
- Isključivanje neurona (engl. *Dropout*) – 30%

Ukupan broj parametara sijamske neuronske mreže korištene u ovom radu iznosi 33339544. Tijekom treninga vrijednosti kriterijskih funkcija na trening i validacijskom skupu konvergirale su prema nuli. Za konačne parametre sijamske neuronske mreže trenirane na *CG-STB-DI* bazi dinamičnih slika, uzeti su parametri nakon 324. epohe. Vrijednost kriterijske funkcije na validacijskom skupu nakon 324. epohe je bila najmanja i iznosila je 0.08202. Postotak točnosti na validacijskom skupu iznosio je 94.12%. Konačni parametri sijamske neuronske mreže trenirane na *CG-STB-CGIAD-DI* bazi dinamičnih slika, uzeti su nakon 488. epohe. Nakon 488. epohe vrijednost kriterijske funkcije na validacijskom setu iznosila je 0.06737, dok je točnost na validacijskom skupu iznosila 98.26%.

Neuronska mreža trenirana je na grafičkoj kartici NVIDIA GTX 3080 TI. Treniranje sijamske neuronske mreže na *CG-STB-DI* bazi dinamičnih slika od 1632 slika i validacija na bazi od 408 slika, za 500 epoha, trajali su 2 sata i 45 minuta. Na *CG-STB-CGIAD-DI* bazi dinamičnih slika koja se sastoji od 2992 trening slika i 748 slika validacijskog skupa, treniranje sijamske neuronske mreže trajalo je 4 sata i 44 minute.

3.5. Upute za pokretanje i korištenje kreiranog rješenja

U ovom potpoglavlju dane su upute kako koristiti algoritam za generiranje dinamičnih slika i algoritam za treniranje i testiranje sijamske neuronske mreže. Sve naredbe navedene u ovom potpoglavlju upisuju se u komandnu liniju.

3.5.1. Upute za pokretanje i korištenje generatora slika

Za pokretanje algoritma za generiranje transformiranih slika, potrebna nam je jedna referentna slika koju želimo transformirati i izlazna putanja gdje će algoritam spremati generirane slike. Kod pokretanja koriste se obavezni argumenti, a oni su:

- Putanja do izlazne putanje: *-o* ili *--output*
- Putanja do referentne slike: *-ri* ili *-refImage*

Algoritam se pokreće pomoću naredbe:

```
python image_generator.py -ri <putanja_do_slike> -o <mjesto_gdje_će_se_spremati_slike>
```

Ako korisnik ne navede parametre granica transformacija i načina rada algoritma, koristit će se prethodno postavljeni parametri i način rada gdje su parametri slučajno odabrani. Za postavljanje granica transformacija i načina rada algoritma koriste se dodatni argumenti, a oni su:

- Donja granica za rotaciju: *-rmin* ili *--rotationAngleMinimum*

- Gornja granica za rotaciju: *-rmax* ili *--rotationAngleMaximum*
- Donja granica za skaliranje: *-smin* ili *--scalingPercentageMinimum*
- Gornja granica za skaliranje: *-smax* ili *--scalingPercentageMaximum*
- Donja granica za okluziju: *-opmin* ili *--occulsionPercentageMinimum*
- Gornja granica za okluziju: *-opmax* ili *--occulsionPercentageMaximum*
- Granica za translaciju slike u lijevo: *-txneg* ili *--translationXNegativeMinimum*
- Granica za translaciju slike u desno: *-txpos* ili *--translationXPositiveMaximum*
- Granica za translaciju sliku prema gore: *-tyneg* ili *--translationYNegativeMaximum*
- Granica za translaciju sliku prema dolje: *-typos* ili *--translationYPositiveMaximum*
- Vrijednost za koju će se povećavati stupanj rotacije: *-ir* ili *--incrementRotation*
- Vrijednost za koju će se povećavati postotak skaliranja: *-is* ili *--incrementScale*
- Vrijednost za koju će se povećavati postotak okluzije: *-io* ili *--incrementOcculsion*
- Vrijednost za koju će se povećavati vrijednost translacije po x-osi: *-itx* ili *--incrementTranslationX*
- Vrijednost za koju će se povećavati vrijednost translacije po y-osi: *-ity* ili *--incrementTranslationY*
- Odabir načina rada: *-m* ili *--mode*
- Broj slika generiranih u modu slučajnog odabira parametara: *-bs* ili *--batchSize*

Kako bi postavili parametar transformacije koji želimo, pri pozivanju naredbe dodamo gore navedene argumente i upisujemo vrijednost koju želimo postaviti. Za postavljanje načina rada uz odgovarajući argument upisujemo vrijednost 0 ili 1. Ako želimo da se slike generiraju slučajnim odabirom vrijednosti, onda upisujemo broj 0. Ako želimo da se slike generiraju s linearnim povećavanjem vrijednosti od donje granice do gornje granice, upisujemo broj 1. Primjer pokretanja algoritma za generiranje transformiranih slika korištenjem dodatnih argumenata:

```
python image_generator.py -ri <putanja_do_slike> -o <mjesto_gdje_će_se_spremati_slike>
    -m <način_rada> -bs <broj_generiranih_slika>
    -opmin <donja_granica_za_postotak_okluzije>
    -opmax <gornja_granica_za_postotak_okluzije> -rmin <donja_granica_za_rotaciju>
```

Gore napisana naredba generirati će 100 slika u načinu rada slučajnih odabira vrijednosti parametara transformacije. Parametri transformacije se uzimaju u granicama od 0 do 4% za okluziju, od -50° do 50° za rotaciju, od -50 elemenata slike do 50 elemenata slike za translaciju po

x-osi od središta slike, od -50 elemenata slike do 50 elemenata slike za translaciju slike po y-osi od središta slike te od 70% do 120% skaliranja od originalne veličine slike.

3.5.2. Upute za treniranje i testiranje sijamske neuronske mreže

Treniranje, validaciju i testiranje sijamske mreže izvršavalo se pomoću algoritama napisanih u Python programskom jeziku. Svi algoritmi su pokretani pomoću linijskog naredbenog retka.

Treniranje sijamske mreže pokretano je pomoću *network_training.py* datoteke. Ova datoteka će ujedno i pokretati validaciju sijamske neuronske mreže. Za rad algoritma, uz naredbu pozivanja algoritma potrebno je napisati obavezne argumente, a oni su:

- Putanja do trening skupa slika: *-pt* ili *--pathTrain*
- Izlazna putanja gdje će se spremati datoteka nove istrenirane sijamske neuronske mreže: *-o* ili *--outputPath*

Za pokretanje algoritma za treniranje sijamske neuronske mreže koristi se naredba koja se upisuje u komandnu liniju:

```
python network_training.py -ri <putanja_do_trening_baze_slika>  
-o <putanja_do_mjesta_spremanja_modela_neuronske_mreže>
```

Gore navedena naredba pokreće treniranje sijamske neuronske mreže. Uz obavezne argumente postoje i dodatni argumenti s kojima postavljamo parametre treniranja. Kako u gore navedenoj naredbi nisu navedeni dodatni argumenti, koriste se prethodno postavljeni parametri.

Dodatni argumenti koji se mogu postaviti glase:

- Veličina serije (engl. *Batch size*): *-bs* ili *--batchSize*
- Broj epoha: *-ep* ili *--epochs*
- Isključivanje neurona (engl. *Dropout*): *-d* ili *--dropout*
- Stopa učenja (engl. *Learning rate*): *-lr* ili *--learningRate*
- Postavljanje nakon koliko epoha se smanjuje stopa učenja i za koliko puta: *-s* ili *--scheduler*
- Za prikazivanje grafa treniranja: *-p* ili *--showPlot*

Za pokretanje algoritma za treniranje sijamske neuronske mreže s dodatnim argumentima koristi se naredba:

```
python network_training.py -ri <putanja_do_trening_baze_slika>
-o <putanja_do_mjesta_spremanja_modela_neuronske_mreže> -bs <veličina_serije>
-ep <broj_epoha> -d <postotak_isključivanja_neurona> -lr <stopa_učenja>
-s <nakon_koliko_epoha_će_se_stopa_učenja_smanjiti_za_koju_vrijednost>
-p <iscrtavanje_grafova_vrijednosti_kriterijske_funkcije>
```

Gore prikazana naredba će pokrenuti trening sijamske neuronske mreže s veličinom serije od 64. Trening će trajati 500 epoha uz stopu učenja od 0.0001, koja će se smanjivati za 0.7 puta svakih 100 epoha. Isključivanje neurona je postavljeno na 0.3, što znači da će se nasumično gasiti 30% neurona nakon svake obrade slika sijamskom neuronskom mrežom tijekom učenja. Posljednji argument se koristi kako bi se prikazao graf promjene vrijednosti kriterijske funkcije za vrijeme procesa treninga.

Testiranje sijamske neuronske mreže pokrećemo pomoću algoritma *network_batch_testing.py*. Kako bi mogli pokrenuti testiranje moramo napraviti tekstualni dokument koji će sadržavati područje interesa gdje će se nalaziti referentni objekt koji želimo usporediti, ime referentne slike, broj testnih slika te imena testnih slika. Ova tekstualna datoteka se mora nalaziti u istoj mapi kao referentna slika i testne slike, a njen naziv mora biti *test.txt*. Tijekom testiranja koristi se i jedna dodatna slika kao pozadinska slika u slučaju da je postavljeno područje interesa manjih dimenzija od 500 x 500 elemenata slike. Prilikom pokretanja se moraju navesti obavezni argumenti, a oni glase:

- Datoteka sijamske neuronske mreže: *-nf* ili *--netFile*
- Putanja do testnih slika: *-pt* ili *--pathTrain*
- Putanja do pozadinske slike: *-bi* ili *--backgroundImage*

Naredba za pokretanje algoritma za testiranje sijamske neuronske mreže glasi:

```
python network_batch_testing.py -nf <putanja_do_modela_neuronske_mreže>
-pt <putanja_do_testnih_slika> -bi <putanja_do_pozadinske_slike>
```

Pomoću gore navedene naredbe pokreće se testni algoritam sijamske neuronske mreže, a na slici 3.10 prikazano je kako treba izgledati tekstualna datoteka s parametrima za testiranje sijamske neuronske mreže. Parametre redom treba navesti kako je prikazano na slici.

```
861 -Početni stupac elemenata slike za širinu
734 -Početni red elemenata slike za visinu
1077 -Krajnji stupa elemenata slike za širinu
971 -Krajnji red elemenata slike za visinu
5 -Broj testnih slika
Referent_picture_dolphin.bmp -Naziv referentne slike
Dolphin_test_1.bmp -Nazivi testnih slika|
Dolphin_test_2.bmp
Dolphin_test_3.bmp
Dolphin_test_4.bmp
Dolphin_test_5.bmp
```

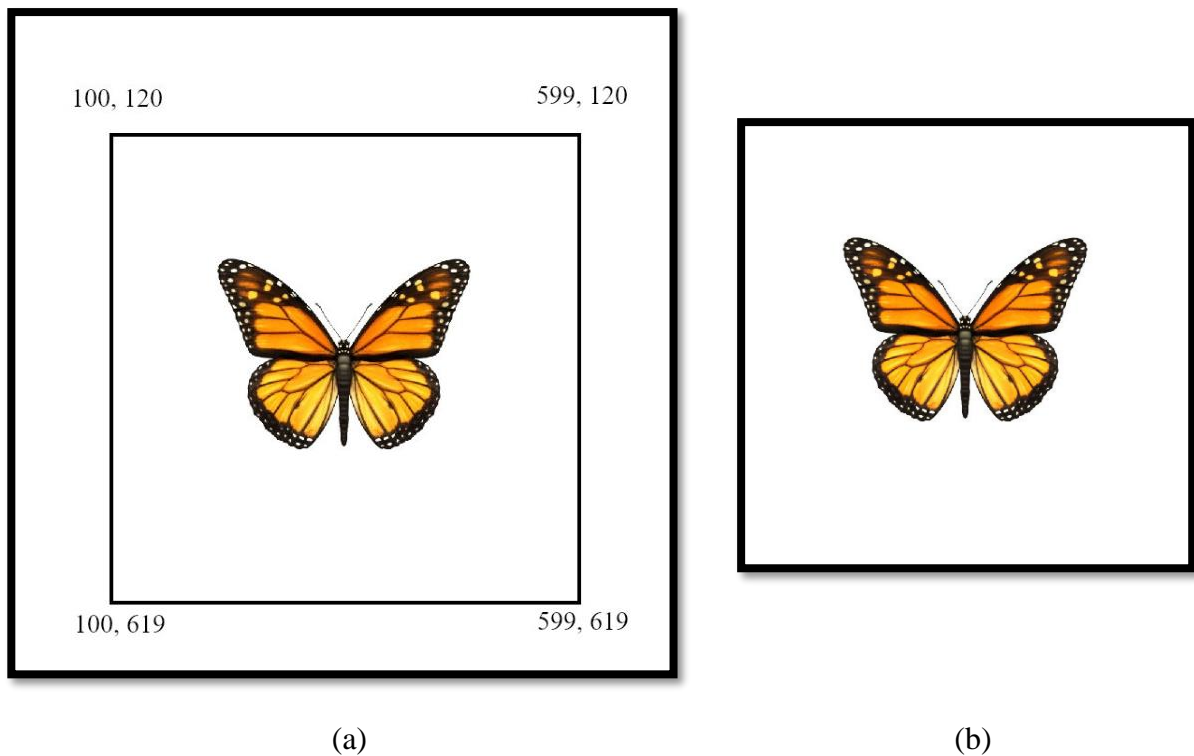
Slika 3.10 Primjer *test.txt* datoteke korištene za testiranje

U prilogu P.3.3. na elektroničkom mediju priloženom uz ovaj rad nalazi se programski kod napisan za potrebe ovog rada pomoću kojega se kreira, trenira i testira sijamska neuronska mreža te sve potrebne dodatne datoteke.

4. VERIFIKACIJA RADA NEURONSKE MREŽE ZA USPOREDBU DINAMIČNIH SLIKA

Verifikacijom rada sijamske neuronske mreže želi se ispitati način rada neuronske mreže trenirane na *CG-STB-DI* i *CG-STB-CGIAD-DI* bazi dinamičnih slika te provjeriti koje performanse postižu oba istrenirana modela sijamske neuronske mreže na određenom skupu podataka. Istrenirani modeli sijamske neuronske mreže testirani su na CGIAD bazi dinamičnih slika i bazi slika STB grafičkog sučelja koje su opisane u potpoglavlju 2.1 ovoga rada. Testiranje se provodi na 10 testnih grupa CGIAD baze dinamičnih slika, a kako se u svakoj testnoj grupi nalazi 500 pozitivnih i 500 negativnih testnih slika, testira se na ukupno 10000 slika CGIAD baze dinamičnih slika. Za testiranje na bazi dinamičnih slika STB grafičkog sučelja koristi se 8 testnih grupa ove baze. Kako svaka testna grupa baze slika STB grafičkog sučelja ima različit broj slika, modeli sijamske neuronske mreže testiraju se na ukupno 68 testnih slika te baze dinamičnih slika. Sve testne grupe na kojima se testirala sijamska neuronska mreža trenirana na *CG-STB-DI* bazi dinamičnih slika nisu korištene u treningu i validaciji sijamske neuronske mreže. Kod testiranja korištenja sijamske neuronske mreže trenirane na *CG-STB-CGIAD-DI* bazi dinamičnih slika, referentne slike CGIAD baze dinamičnih slika korištene su za generiranje vlastite baze dinamičnih slika s kojom je taj model mreže treniran. Testiranje se izvršava tako da se usporede jedna referentna slika i jedna testna slika iz pojedine testne grupe testnih baza. Kao rezultat testiranja na izlazu iz sijamske neuronske mreže dobiva se rezultat različitosti slika koji govori jesu li slike iste ili različite. Ovaj se rezultat različitosti kreće u rasponu od 0 do beskonačno i stoga je potrebno odrediti prag rezultata različitosti. Prag mora biti pozitivan broj i kada rezultat različitosti pređe postavljeni prag, za slike se smatra da su različite. Svi dobiveni rezultati različitosti slika koji su manji od postavljenog praga različitosti označavaju da su dvije slike koje se uspoređuju jednake. Ovaj prag se postavlja proizvoljno i tijekom testiranja njegova vrijednost će biti 0.8. Razlog zašto je prag različitosti postavljen u ovu vrijednost je taj što se tijekom treniranja sijamske neuronske mreže, sama mreža učila da je za parove gdje su slike različite, rezultat različitosti, odnosno udaljenost između dviju klasa iznosila 1. Kako se rezultat različitosti za različite slike tijekom treniranja može nalaziti oko vrijednosti 1, postavljeni prag različitosti je postavljen na vrijednost manju od praga različitosti korištenog tijekom treninga. Ulazne slike za sijamsku neuronsku mrežu moraju biti veličine 500x500 elemenata slike. U slučaju da su ulazne slike manjih dimenzija od traženih 500x500 elemenata slike, ulazne se slike postavljaju u središte crne pozadine koja je veličine 500x500 elemenata slike. Kada su ulazne slike veće od 500x500 elemenata slike, određuje se područje interesa (engl. *Area of Interest, AOI*). Područje interesa je područje slike koje se želi

izdvojiti i iskoristiti za testiranje, a njegova veličina treba biti 500x500 elemenata slike. Unutar područja interesa mora se nalaziti objekt koji se želi usporediti na objema dinamičnim slikama. Tijekom testiranja za referentnu i testnu sliku koristi se ista pozicija područja. Korištenjem iste pozicije područja interesa osigurava se da se uspoređuju objekti koji se nalaze na istom području slike. Ovim ograničavanjem se može dogoditi slučaj kada objekt na jednoj slici od para slika izlazi iz područja interesa. Kako bi se spriječilo izlaženje objekta iz područja interesa zbog većih transformacija (npr. rotacija, translacija i skaliranje), postavlja se područje interesa dovoljno malo da prikazuje jedan objekt i dovoljno veliko da prikazuje cijeli objekt unatoč transformacijama. Primjer određivanja područja interesa prikazan je na slici 4.1.



Slika 4.1 Primjer određivanja područja interesa. (a) originalna slika veličine 700x700 s područjem interesa od 500x500 elemenata slike, (b) odvojeno područje interesa slike veličine 500x500 elemenata slike

Na slici 4.1 prikazano je područje interesa u obliku crnog pravokutnika. Za određivanje područja interesa potrebna su 4 broja, a oni predstavljaju koordinate gornjeg lijevog ugla i donjeg desnog ugla područja interesa. U primjeru sa slike za početni se redak područja interesa koristi 120. redak matrice elemenata slike, a za početni se stupac područja interesa koristi 100. stupac matrice elemenata slike. Kao krajnji redak i stupac područja interesa koriste se 619. redak i 599. stupac matrice elemenata slike. Ovom kombinacijom dobije se područje interesa veličine 500x500 elemenata slike. Kako je dobiveno područje interesa jednako 500x500 elemenata slike, područje interesa se izdvaja i ono se koristi kao ulaz u sijamsku neuronsku mrežu.

Rezultati usporedbe dinamičnih slika klasificiraju se na sljedeći način:

- Ispravno klasificirane pozitivne slike (engl. *True Positive, TP*) - Testne slike koje sadrže isti objekt kao referentna slika i prepoznate su kao da sadrže isti objekt.
- Neispravno klasificirane pozitivne slike (engl. *False Negative, FN*) - Testne slike koje sadrže isti objekt kao referentna slika, a prepoznate su kao da ne sadrže isti objekt.
- Ispravno klasificirane negativne slike (engl. *True Negative, TN*) - Testne slike koje ne sadrže isti objekt kao referentna slika i prepoznate su kao da ne sadrže isti objekt.
- Neispravno klasificirane negativne slike (engl. *False Positive, FP*) - Testne slike koje ne sadrže isti objekt kao referentna slika, a prepoznate su kao da sadrže isti objekt.

Pomoću prethodno navedenih klasifikacija rezultata mogu se izračunati performanse neuronske mreže. Performanse koje se izračunavaju su preciznost (engl. *precision*), odziv (engl. *recall*) i točnost (engl. *accuracy*). Točnost je mjera koja opisuje kakve performanse neuronska mreža ima u odnosu na sve klasifikacije, a dobiva se omjerom broja ispravno prepoznatih slika i ukupnog broja svih testnih prepoznatih slika. Formula za točnost glasi:

$$\text{Točnost} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4-1)$$

Preciznost je mjera koja pokazuje koliko je od svih pozitivno klasificiranih slika njih točno klasificirano. Računa se kao omjer broja ispravno klasificiranih pozitivnih slika i broja svih slika koje su klasificirane kao pozitivne, odnosno onih na kojima je prepoznato da sadrže isti objekt kao referentna slika. Preciznost se izračunava pomoću formule:

$$\text{Preciznost} = \frac{TP}{TP + FP} \quad (4-2)$$

Pomoću mjere odziva dobiva se omjer broja slika koje su klasificirane kao pozitivne i broja svih slika koje su trebale biti klasificirane kao pozitivne. Formula za odziv glasi:









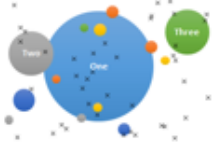



$$\text{Odziv} = \frac{TP}{TP + FN} \quad (4-3)$$




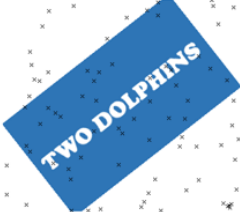







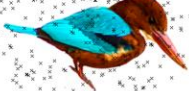



Kako bi se vidjelo kako sijamska neuronska mreža radi u usporedbi s postojećim algoritmom za usporedbu dinamičnih slika zasnovanim na ključnim točkama, svi rezultati testiranja sijamske neuronske mreže na ovim dvjema bazama uspoređeni su s rezultatima koji su dobiveni pomoću DISM algoritma na istim bazama.

4.1. Verifikacija rada neuronske mreže na CGIAD bazi podataka

Oba modela sijamske neuronske mreže testirana su prvo na CGIAD bazi podataka. Testom na ovoj bazi provjerava se prepoznaje li sijamska neuronska mreža nalaze li se isti objekti na dvjema slikama, neovisno o transformacijama koje su napravljene na objektima. Testovi su izvršeni na svih 10 postojećih testnih grupa iz CGIAD baze. Kao što je već navedeno u drugom poglavlju rada, testne grupe sastoje se od po 500 slika koje sadrže jednaki objekt kao referentna slika i po 500 slika koje ne sadrže jednaki objekt kao referentna slika. Slike koje sadrže objekt jednak objektu koji se nalazi na referentnoj slici nazivaju se pozitivne slike. Razlika je između pozitivnih slika i referentne slike što je objekt na pozitivnim slikama izložen transformacijama. Slike koje ne sadrže objekt jednak objektu koji se nalazi na referentnoj slici nazivaju se negativne slike. Na objektima koji se nalaze na negativnim slikama, također su izvršene transformacije. Primjeri referentnih, pozitivnih, i negativnih slika prikazani su u tablici 4.1, a područja interesa za svaku od testnih grupa prikazana su u tablici 4.2.

Tablica 4.1 Primjer slika iz 10 testnih skupina CGIAD baze dinamičnih slika

Naziv grupe	Referentna slika	Pozitivna slika	Negativna slika
Testna grupa 1			
Testna grupa 2			
Testna grupa 3			
Testna grupa 4			

Testna grupa 5			
Testna grupa 6			
Testna grupa 7			
Testna grupa 8			
Testna grupa 9			
Testna grupa 10			

Tablica 4.2 Područja interesa za svaku od testnih grupa CGIAD baze slika

Naziv grupe	Područje interesa (početni redak, početni stupac, krajnji redak, krajnji stupac)
Testna grupa 1	349, 219, 848, 718
Testna grupa 2	334, 243, 833, 742

Testna grupa 3	230, 243, 729, 742
Testna grupa 4	287, 240, 786, 739
Testna grupa 5	307, 240, 806, 739
Testna grupa 6	270, 240, 769, 739
Testna grupa 7	225, 240, 724, 739
Testna grupa 8	280, 240, 779, 739
Testna grupa 9	235, 240, 734, 739
Testna grupa 10	190, 245, 689, 744

Rezultati testiranja sijamske neuronske mreže trenirane na *CG-STB-DI* bazi dinamičnih slika prikazani su u tablici 4.3. U tablici se nalaze rezultati klasifikacije, odziv, točnost i preciznost za svaku od testnih grupa CGIAD baze slika. Posljednji red tablice prikazuje rezultate testiranja za sve testne grupe zajedno.

Tablica 4.3 Prikaz rezultata sijamske neuronske mreže trenirane na *CG-STB-DI* bazi dinamičnih slika testirane na CGIAD bazi podataka s pragom rezultata različitosti 0.8.

Prag različitosti = 0.8	TP	FN	TN	FP	Preciznost [%]	Odziv [%]	Točnost [%]
Testna Grupa 1	349	151	424	76	82.11	69.80	73.30
Testna Grupa 2	486	14	497	3	99.38	97.20	98.30
Testna Grupa 3	399	101	500	0	100.00	79.80	89.90
Testna Grupa 4	299	201	293	207	59.09	59.80	59.20
Testna Grupa 5	325	175	500	0	100.00	65.00	82.50
Testna Grupa 6	351	149	364	136	72.07	70.20	71.50
Testna Grupa 7	490	10	500	0	100.00	98.00	99.00
Testna Grupa 8	362	138	500	0	100.00	72.40	86.20
Testna Grupa 9	235	265	500	0	100.00	47.00	73.50
Test Grupa 10	386	114	311	189	67.13	77.20	69.70
Ukupno	3682	1318	4389	611	85.76	73.64	80.71

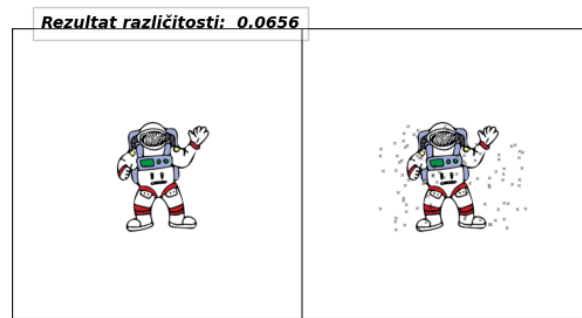
Analizom dobivenih rezultata (Tablica 4.3) nakon testiranja sijamske neuronske mreže kao univerzalnog rješenja na CGIAD bazi dinamičnih slika, može se vidjeti da, kada se uzmu rezultati svih testnih grupa zajedno, točnost sijamske neuronske mreže iznosi 80.71%. Niži postotak odziva od 73.64 % govori o tome kako neuronska mreža prepoznaje veliki broj stvarno pozitivnih slika kao negativne slike. Kada bi se vratili na problem funkcionalnog testiranja aplikacija, neuronska mreža treba imati vrlo dobru sposobnost prepoznavanja i pozitivnih i negativnih slika. Iz tog razloga u analizi ovih rezultata bitna mjera je preciznosti, a ona za ovu neuronsku mrežu iznosi 85.76%. Dobivena preciznost govori da neuronska mreža dio negativnih slika klasificira kao pozitivne, odnosno da sadrže jednaki objekt kao na referentnoj slici, što nije ispravno. Razlozi ovakvih rezultata mogu biti slične boje objekata na referentnoj i negativnoj slici, veliki postotak okluzije, slični izgled i kompozicija objekata na referentnoj i negativnoj slici. U tablici 4.4 prikazane su granice transformacija za CGIAD bazu dinamičnih slika na kojoj je izvršeno testiranje i za obje vlastite baze dinamičnih slika koje su korištene za treniranje.

Tablica 4.4 Granice transformacija za CGIAD bazu dinamičnih slika i obje vlastite baze dinamičnih slika korištene za treniranje sijamske neuronske mreže

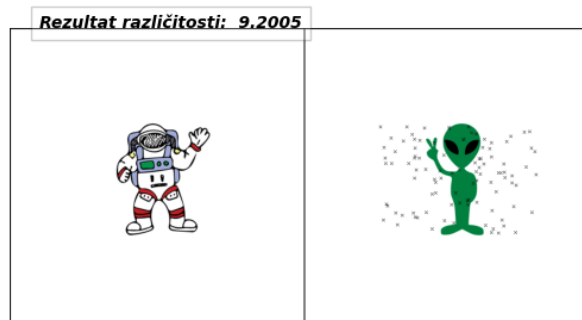
Transformacija	CGIAD baza dinamičnih slika	Vlastite baze dinamičnih slika korištenih za treniranje
Skaliranje	80 % do 120%	80% do 120%
Rotacija	-40° do +40°	-40° do +40°
Translacija po x-osi	-30 do +30 elemenata slike	-40 do +40 elemenata slike
Translacija po y-osi	-30 do +30 elemenata slike	-40 do +40 elemenata slike
Okluzija	0% do 10%	0% do 10%

Pregledom tablice 4.4 može se vidjeti kako su granice transformacija vrlo slične za obje baze. Jedina razlika je granica za koliko elemenata slike će se događati translacija po x-osi i y-osi. Kod vlastitih baza dinamičnih slika translacija po objema osima iznosi od -40 do +40 elemenata slike, a kod CGIAD baze dinamičnih slika translacija po objema osima iznosi od -30 do +30 elemenata slike. Većom vrijednosti granica za translacije po obje osi (od -40 do +40 elemenata slike) pokušava se smanjiti utjecaj translacije objekata na rezultat različitosti. Zbog sličnosti granica transformacija kod CGIAD baze dinamičnih slika i obje vlastitih generiranih baza dinamičnih slika, može se reći kako granice transformacija kod trening baza dinamičnih slika ne igraju ulogu kod dobivanja lošijih performansi na CGIAD bazi dinamičnih slika. Kako bi se otkrili točniji razlozi manje preciznosti u usporedbi dinamičnih slika, trebaju se pogledati rezultati pojedinačno po testnim grupama.

Kada se pogledaju primjeri referentnih, pozitivnih i negativnih slika (Tablica 4.1), može se zaključiti da boja, odnosno komponente boja koje se koriste na objektima imaju utjecaj na rezultat različitosti. Najbolji primjer vidljiv je kod testne grupe 7 na kojoj sijamska neuronska mreža postiže najbolje rezultate klasifikacije pozitivnih i negativnih slika (sve tri mjere performansi su iznad 98%). Na slici 4.2 prikazan je primjer rezultata različitosti koje daje predložena sijamska neuronska mreža na slikama testne grupe 7.



(a)



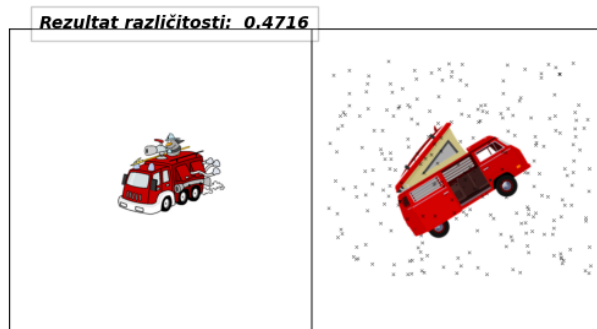
(b)

Slika 4.2 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 7 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku

Sijamska neuronska mreža je prepoznala da se na negativnoj slici nalazi drugačiji objekt i da se na pozitivnoj slici nalazi jednak objekt. Većem rezultatu različitosti doprinijela je razlika u bojama objekata na referentnoj i negativnoj. Na negativnoj slici objekt je skoro jednobojan, dok se na referentnoj slici nalazi objekt koji na sebi ima više različitih boja. Najlošije rezultate sijamska neuronska mreža postigla je na testnoj grupi 4, a jedan primjer je prikazan na slici 4.3.



(a)



(b)

Slika 4.3 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 4 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku

Na slici 4.3 može se vidjeti primjer neispravne klasifikacije pozitivne i negativne slike. Razlog neispravne klasifikacije pozitivne slike može biti nedovoljna istreniranost sijamske neuronske mreže koja uzrokuje da zbog velike količine točkica koje prekrivaju objekt. Neispravna klasifikacija negativne slike mogla se dogoditi zbog korištenja iste komponente boje te slične orijentacije objekta. Najveće razlike između objekata na testnim i referentnim slikama su detalji koji se nalaze na objektima. Tijekom filtriranja slika pomoću konvolucijskih filtara i slojeva sažimanja može doći do gubitka detalja koji se nalaze na slikama. Gubitak detalja objekata koji se nalaze na negativnoj slici mogu učiniti objekte na negativnoj slici vrlo sličnima onima koji se nalaze na referentnoj slici. Ovo može dovesti do toga da sijamska neuronska mreža negativne slike klasificira kao pozitivne, unatoč što su objekti na njima različitih boja od onih s referentnih slika.

Način na koji se može smanjiti utjecaj pojedinih elemenata kao što su boja, oblik i orijentacija objekata je promjena baze slika korištene za treniranje neuronske mreže te nadogradnja ili promjena arhitekture mreže. Bazu slika korištenih za treniranje neuronske mreže može se promijeniti na taj način da se dodaju slike s objektima koji su vrlo slični bojom i oblikom pojedinim objektima na slikama u bazi slika korištenih za treniranje te na taj način predstavljaju negativne slike. Također, mijenjanjem slojeva konvolucije i sažimanja može se utjecati na količinu

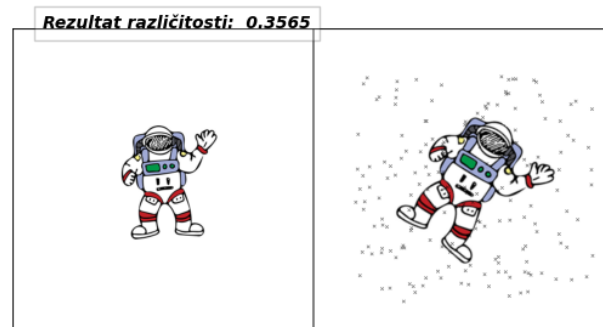
izgubljenih detalja. Ove promjene bi mogle pridonijeti boljim rezultatima prilikom testiranja sijamske neuronske mreže.

Nakon testiranja sijamske neuronske mreže trenirane na *CG-STB-DI* bazi dinamičnih slika, testirana je sijamska neuronska mreža trenirana na *CG-STB-CGIAD-DI* bazi dinamičnih slika. Rezultati testiranja sijamske neuronske mreže kao specifičnog rješenja na CGIAD bazi dinamičnih slika i njene performanse preciznosti, točnosti i odziva prikazane su u tablici 4.5.

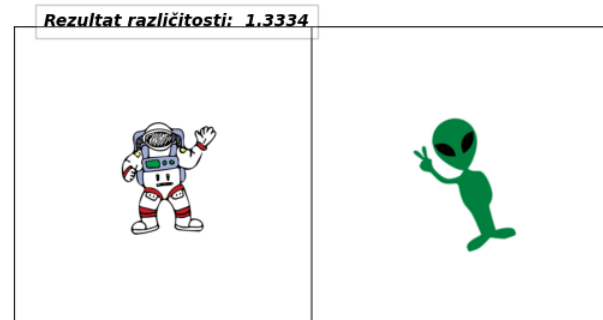
Tablica 4.5 Prikaz rezultata sijamske neuronske mreže trenirane na *CG-STB-CGIAD-DI* bazi dinamičnih slika testirane na CGIAD bazi podataka s pragom rezultata različitosti 0.8.

Prag različitosti = 0.8	TP	FN	TN	FP	Preciznost [%]	Odziv [%]	Točnost [%]
Test Grupa 1	500	0	437	63	88.81	100.00	93.70
Test Grupa 2	500	0	499	1	99.80	100.00	99.90
Test Grupa 3	433	67	500	0	100	86.66	93.30
Test Grupa 4	414	86	442	58	87.71	82.80	85.60
Test Grupa 5	412	88	500	0	100,00	82,40	91,20
Test Grupa 6	459	41	500	0	100.00	91.80	95.90
Test Grupa 7	500	0	500	0	100.00	100.00	100.00
Test Grupa 8	497	3	500	0	100.00	99.40	99.70
Test Grupa 9	500	0	426	74	87.11	100.00	92.60
Test Grupa 10	423	77	424	76	84.77	84.60	84.70
Ukupno	4638	362	4728	272	94.46	92.76	93.66

Pregledom tablice 4.5 može se vidjeti kada je sijamska neuronska mreža trenirana kao specifično rješenje da se postiže ukupna preciznost od 94.46%, odziv od 92.76% i točnost 93.66%. Najbolje rezultate preciznosti, odziva i točnosti sijamska neuronska mreža postiže na testnoj grupi 7 CGIAD baze dinamičnih slika. Najlošiju preciznost i točnost sijamska neuronska mreža postiže na testnoj grupi 10, a najlošiji odziv sijamska neuronska mreža postiže na testnoj grupi 5. Primjer rezultata postignutih na testnoj grupi 7 prikazan je na slici 4.4.



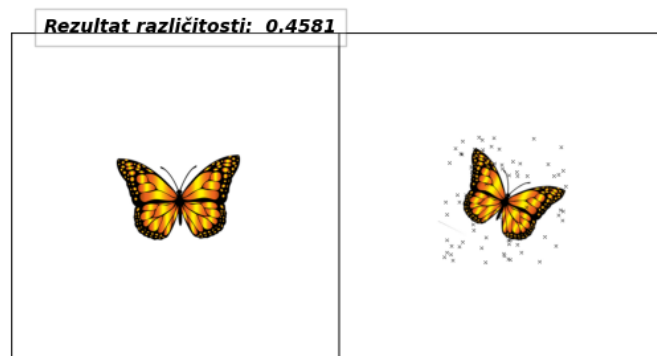
(a)



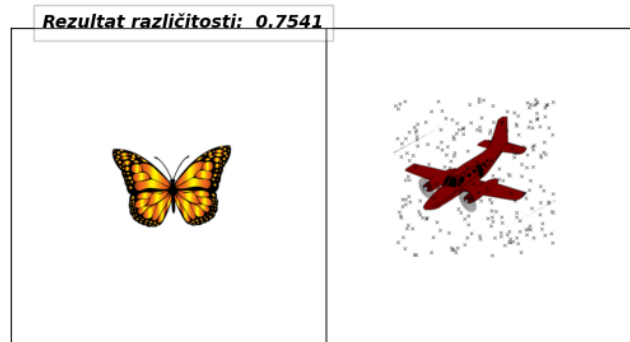
(b)

Slika 4.4 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 7 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku

Na slici 4.4 može se vidjeti kako su boje objekata na referentnoj i testnoj negativnoj slici različite, iako su oblici objekata koje slike prikazuju slični. Na testnoj negativnoj slici se nalazi samo jedna boja, dok na referentnoj slici se nalazi više različitih boja. Najvjerojatniji razlog ispravne klasifikacije je značajnija razlika u boji objekata koji se nalaze na referentnoj i testnoj negativnoj slici. Na slici 4.5 prikazani su rezultati različitosti postignuti na testnoj grupi 10 CGIAD baze dinamičnih slika.



(a)



(b)

Slika 4.5 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 10 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku

Na testnoj grupi 10 CGIAD baze dinamičnih slika sijamska neuronska mreža korištena kao specifično rješenje postigla je najlošije rezultate preciznosti i odziva. Na slici 4.5 (b) prikazana je neispravna klasifikacija negativne slike. Usporedbom referentne i testne negativne slike dobiven je rezultat različitosti od 0.7541. Iako postoji okluzija na slici i na slici se nalazi različit objekt, sijamska neuronska mreža je ove slike klasificirala kao jednake. Zbog konvolucijskih slojeva izgubljeni su detalji koji razlikuju slike te je mreža zbog sličnosti gornjeg desnog kuta svakog od objekta na referentnoj i testnoj negativnoj slici mogla klasificirati slike kao jednake. Također, korištenje sličnih komponenata crvene boje može biti razlog neispravno klasificiranih slika. Neispravne klasifikacije mogu se vidjeti na parovima testnih grupa 5 i 6, čije primjere rezultata različitosti se može vidjeti na slici 4.6 i slici 4.7.



(a)

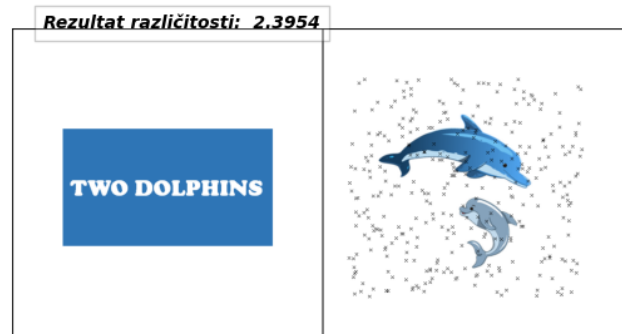


(b)

Slika 4.6 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 5 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku



(a)



(b)

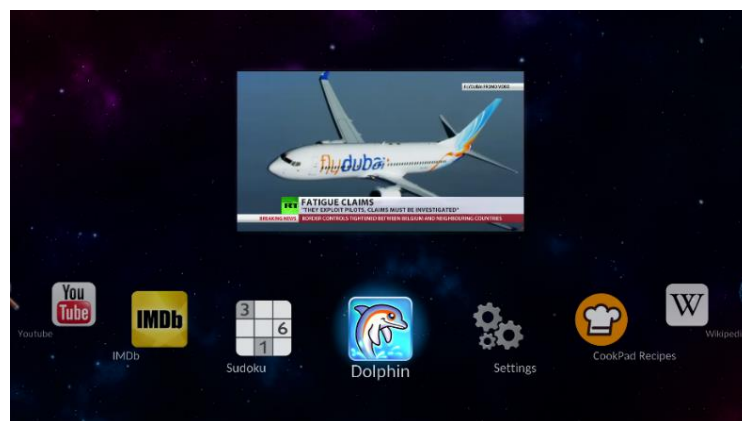
Slika 4.7 Primjer rezultata različitosti koje daje predložena neuronska mreža na slikama iz testne grupe 6 (a) rezultat različitosti za jednu pozitivnu sliku, (b) rezultat različitosti za jednu negativnu sliku

Na slici 4.6 i slici 4.7 mogu se vidjeti neispravne klasifikacije pozitivnih testnih slika. Kao razlog neispravnih klasifikacija može biti nedovoljno velik skup trening podataka na kojima se nalazi maksimalna vrijednost okluzije. Povećavanjem broja slika s maksimalnom okluzijom ili povećavanjem granice za okluziju mogao bi se izbaciti utjecaj okluzije na rezultate usporedbe dinamičnih slika.

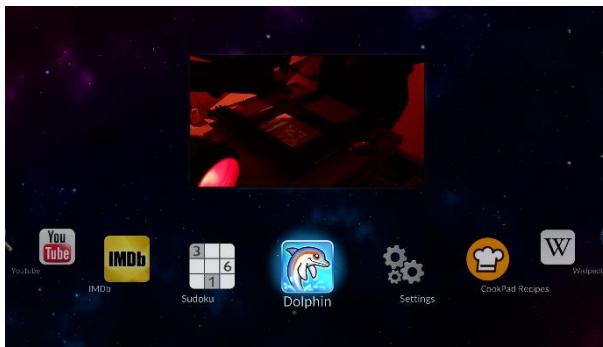
Korištenje sijamske neuronske mreže trenirane na *CG-STB-CGIAD-DI* daje bolje rezultate tijekom testiranja na CGIAD bazi nego kada se koristi sijamska neuronska mreža trenirana na *CG-STB-DI*. Ovakvi rezultati su dijelom i očekivani zato što se za generiranje *CG-STB-CGIAD-DI* baze dinamičnih slika koriste referentne slike CGIAD baze dinamičnih slika. Povećavanjem broja različitih klasa unutar baza dinamičnih slika korištenih za treniranje ove sijamske mreže, mijenjanjem arhitekture sijamske neuronske mreže te dodatnim treniranjem, vjerojatno bi se mogli postići bolji rezultati različitosti na CGIAD bazi dinamičnih slika.

4.2. Verifikacija rada neuronske mreže na bazi slika STB grafičkog sučelja

Testiranje na bazi slika STB grafičkog sučelja [2] je korišteno kako bi provjerili rad sijamske neuronske mreže u situaciji funkcionalnog testiranja jedne realne aplikacije s dinamičnim slikama. U ovom testu koriste se slike izdvojene iz grafičkog sučelja STB-a. Baza slika STB grafičkog sučelja sastoji se od slika izvučenih iz video sekvenci u nekom nasumično odabranom trenutku. Slike su podijeljene u 8 testnih skupina (*Apps_GALLERY*, *Apps_DOLPHIN*, *TV_Guide_POSTER*, *TV_Guide_POSTER_ARTE*, *TV_Guide_SATURDAY*, *TV_Guide_SUNDAY*, *TV_Guide_TODAY*, *TV_Guide_TOMORROW*). Svaka od 8 testnih skupina ima svoju jednu referentnu sliku i različit broj pozitivnih i negativnih testnih slika. Pozitivne slike sastoje se objekata jednakih kao objekt na referentnoj slici. Jedine su razlike mala promjena osvjetljenja pozadine ili pojavljivanje okluzije u obliku bijelih točkica. Negativne slike prikazuju potpuno drugi objekt u odnosu na objekt na referentnoj slici. Drugi način dobivanja negativnih slika je da negativne slike prikazuju jednaki objekt kao na referentnoj slici, ali razlikuju se u tome je li izbornik koji pokazuje objekt izabran ili ne. Primjeri testnih skupina s njihovim referentnim, pozitivnim i negativnim slikama dani su na slici 4.8, slici 4.9, slici 4.10, slici 4.11, slici 4.12, slici 4.13, slici 4.14 i slici 4.15. Slike nisu prikazane tablično kao CGIAD baza dinamičnih slika jer bi tada slike bile premale i ne bi se vidjeli detalji.



(a)

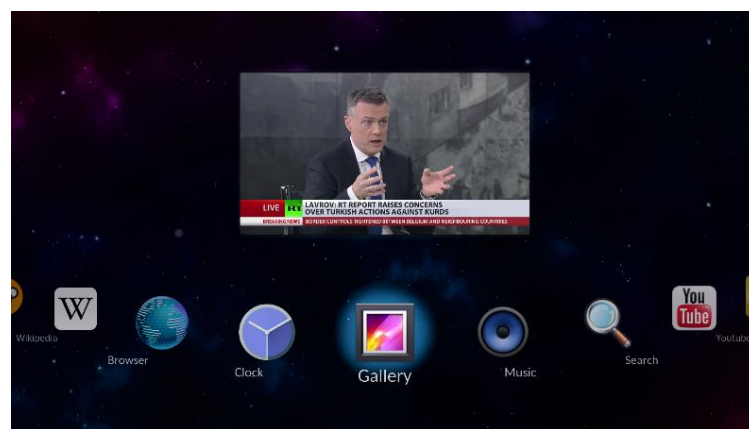


(b)

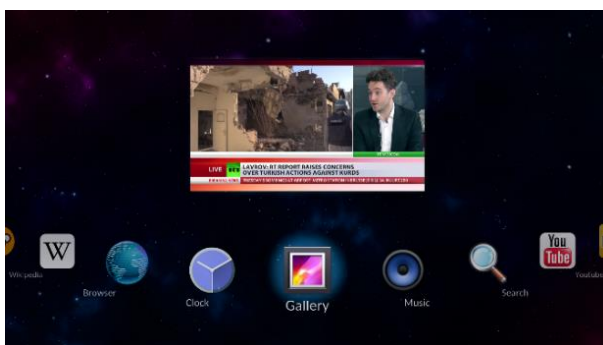


(c)

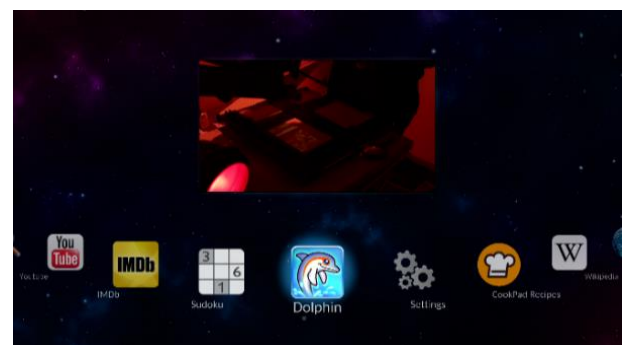
Slika 4.8 Primjer testne skupine *Apps_DOLPHIN*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

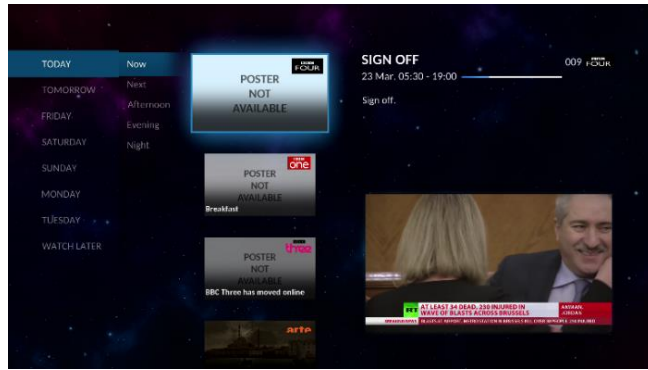


(b)

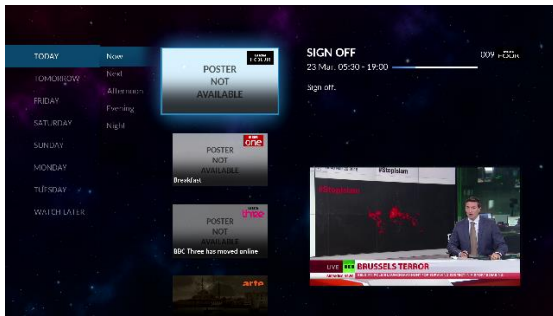


(c)

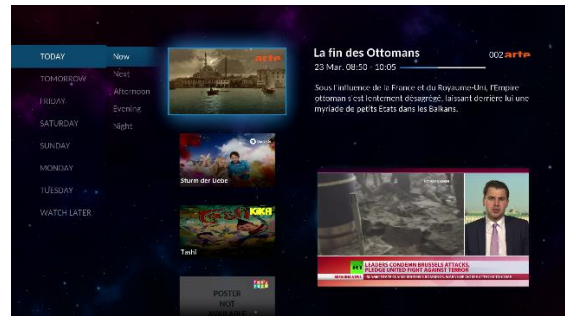
Slika 4.9 Primjer testne skupine *Apps_GALLERY*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

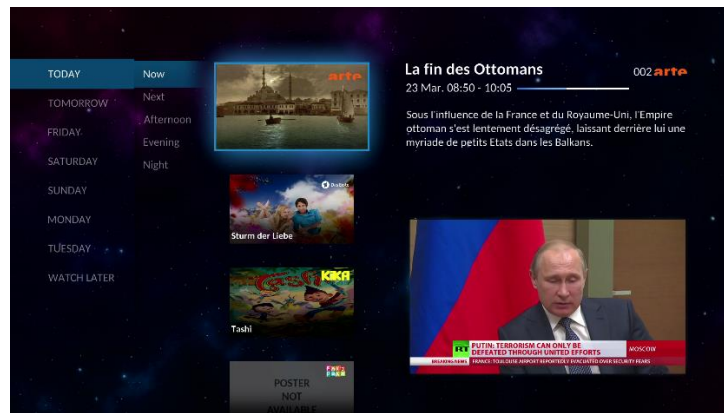


(b)

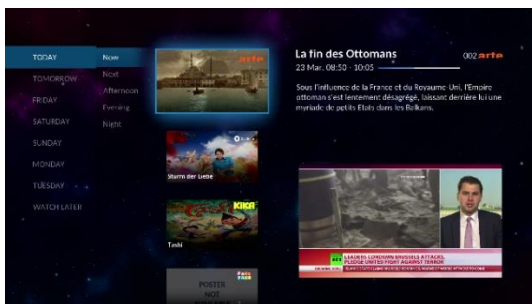


(c)

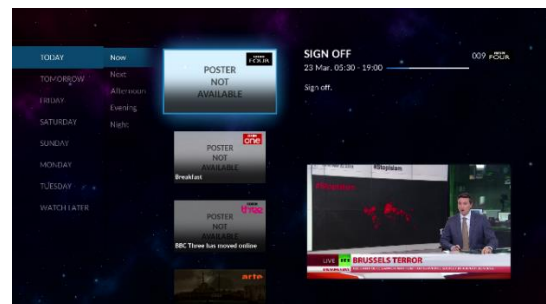
Slika 4.10 Primjer testne skupine *TV_Guide_POSTER*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

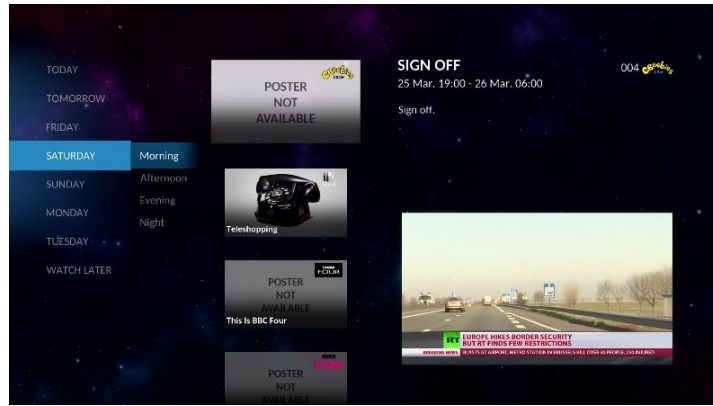


(b)

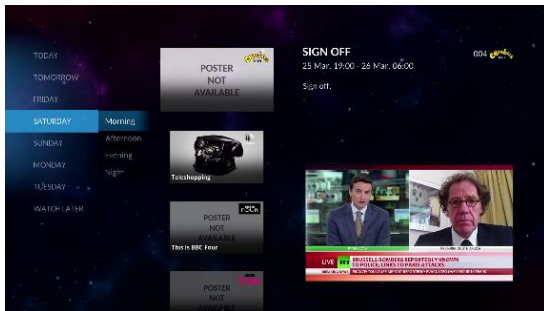


(c)

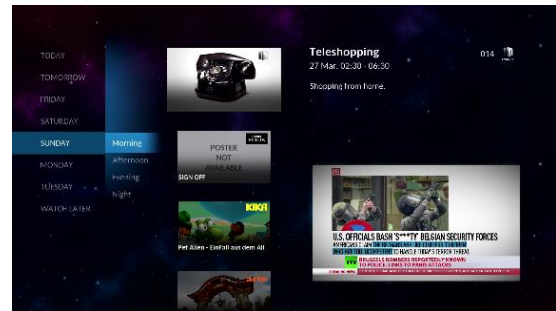
Slika 4.11 Primjer testne skupine *TV_Guide_POSTER_ARTE*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

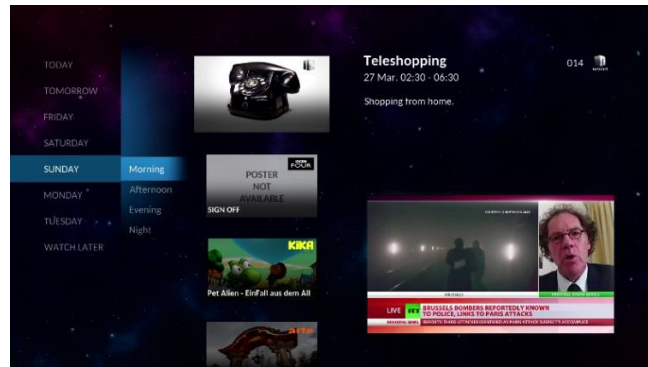


(b)

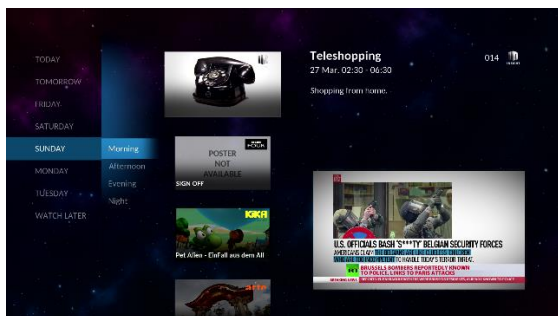


(c)

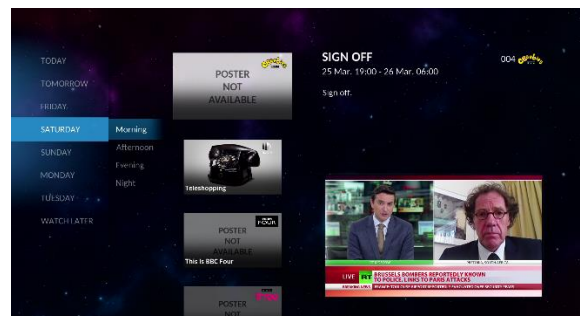
Slika 4.12 Primjer testne skupine *TV_Guide_SATURDAY*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

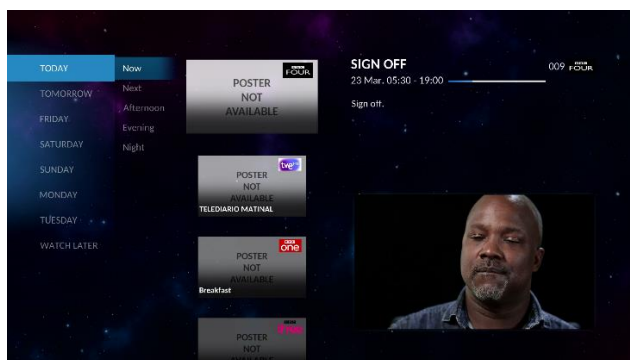


(b)

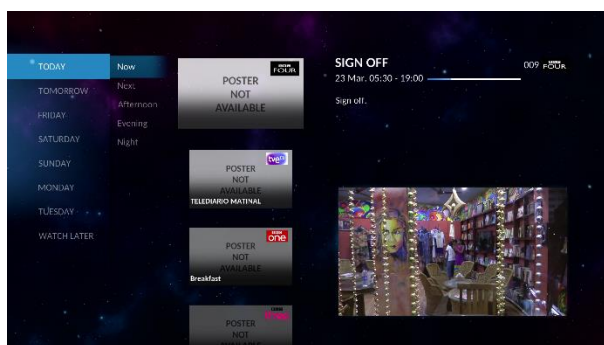


(c)

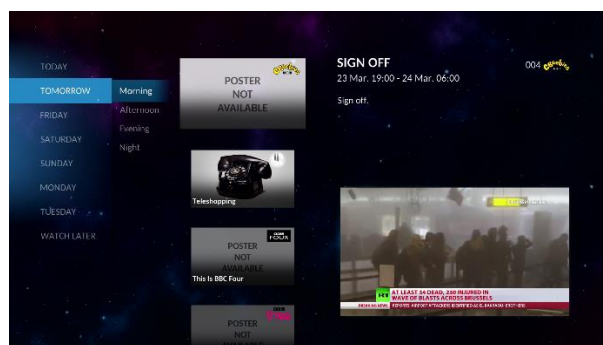
Slika 4.13 Primjer testne skupine *TV_Guide_SUNDAY*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)

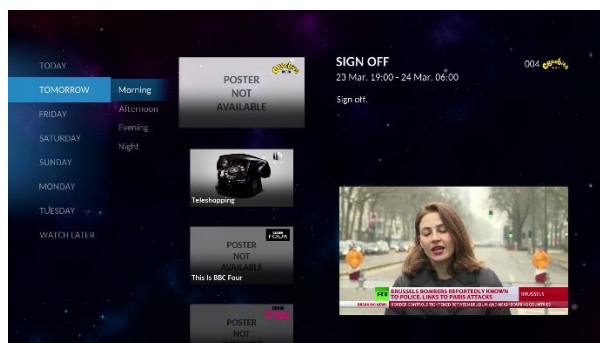


(b)

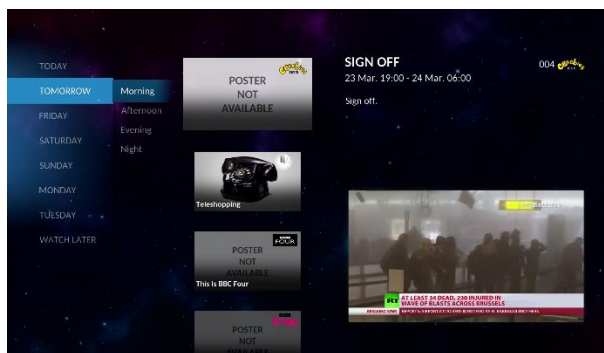


(c)

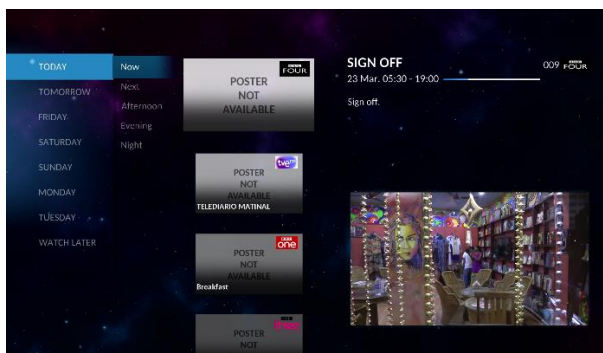
Slika 4.14 Primjer testne skupine *TV_Guide_TODAY*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika



(a)



(b)



(c)

Slika 4.15 Primjer testne skupine *TV_Guide_TOMORROW*: (a) referentna slika, (b) pozitivna slika, (c) negativna slika

Sve su slike testnih skupina veličine 1920x1080 elemenata slike pa je potrebno odrediti područje interesa za svaku testnu skupinu. Unutar područja interesa mora se nalaziti objekt koji se uspoređuje. U tablici 4.6 prikazana su područja interesa i broj pozitivnih i negativnih slika u svakoj testnoj skupini. Kako su područja interesa manja od potrebnih 500x500 elemenata slike, područje interesa se izdvaja i lijepi na crnu pozadinu.

Tablica 4.6 Područja interesa i broj pozitivnih i negativnih slika za svaku testnu skupinu baze slika STB grafičkog sučelja

Testna skupina	Područje interesa (početni redak, početni stupac, krajnji redak, krajnji stupac)	Broj pozitivnih slika	Broj negativnih slika
Apps_GALLERY	856, 726, 1085, 972	4	4
Apps_DOLPHIN	861, 734, 1077, 971	4	4
TV_Guide_POSTER	528, 126, 973, 391	6	6
TV_Guide_POSTER_ARTE	530, 136, 967, 396	6	6
TV_Guide_SATURDAY	8, 371, 328, 446	3	4
TV_Guide_SUNDAY	8, 445, 328, 525	4	3
TV_Guide_TODAY	11, 140, 323, 213	3	4
TV_Guide_TOMORROW	11, 222, 324, 291	2	4

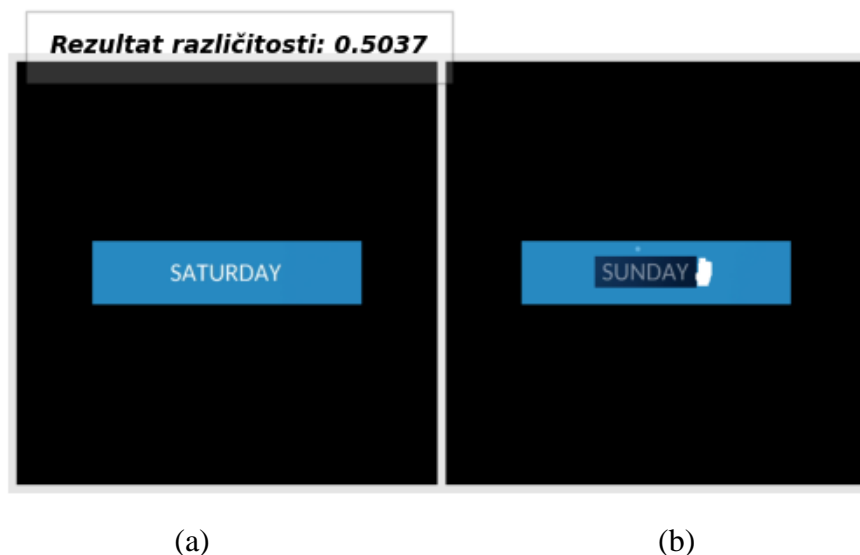
Usporedba slika izvedena je tako da su uspoređene referentna slika sa svakom pozitivnom testnom i negativnom testnom slikom pojedine testne skupine. Rezultati usporedbe i performanse sijamske neuronske mreže na bazi slika STB grafičkog sučelja dane su u tablici 4.7.

Tablica 4.7 Prikaz rezultata sijamske neuronske mreže trenirane na *CG-STB-DI* baze dinamičnih slika testirane na bazi slika STB grafičkog sučelja s pragom rezultata različitosti 0.8.

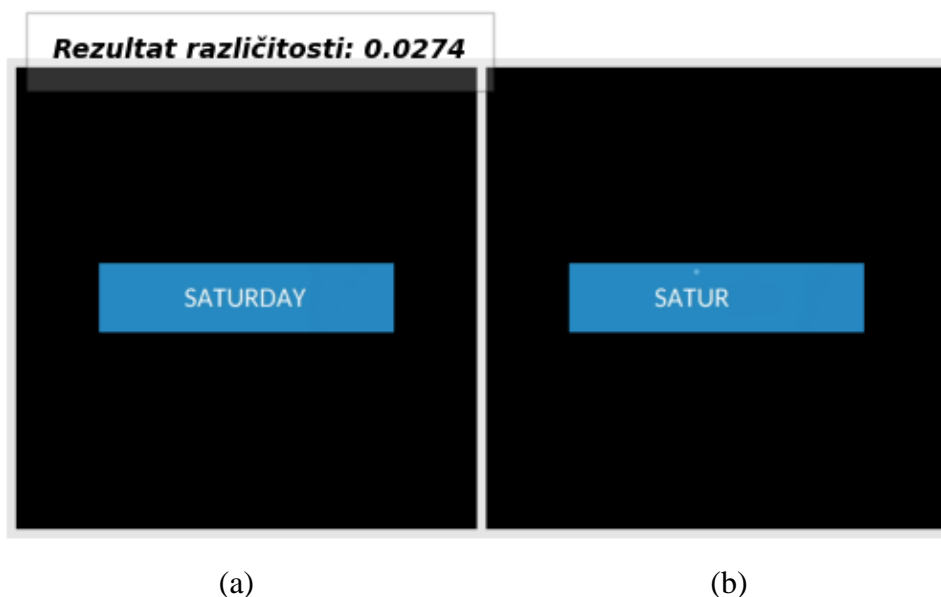
Prag različitosti = 0,8	TP	FN	TN	FP	Preciznost [%]	Odziv [%]	Točnost [%]
Apps_GALLERY	4	0	4	0	100.00	100.00	100.00
Apps_DOLPHIN	4	0	4	0	100.00	100.00	100.00
TV_Guide_POSTER	6	0	6	0	100.00	100.00	100.00
TV_Guide_POSTER_ARTE	6	0	6	0	100.00	100.00	100.00
TV_Guide_SATURDAY	3	0	3	2	60.00	100.00	75.00
TV_Guide_SUNDAY	4	0	3	0	100.00	100.00	100.00

TV_Guide_TODAY	3	0	4	0	100.00	100.00	100.00
TV_Guide_TOMORROW	2	0	4	0	100.00	100.00	100.00
Ukupno	32	0	34	2	94.11	100.00	97.05

Rezultati koji su prikazani u tablici 4.7 prikazuju da je sijamska neuronska mreža za sve testne grupe zajedno postigla točnost od 97.05%. Sijamska neuronska mreža sve je pozitivne slike uspješno označila, što je dovelo do odziva od 100%. U slučaju funkcionalnog testiranja aplikacija gdje je potrebno provjeriti je li grafičko sučelje dobro ispunjeno, potrebno je uspješno klasificirati ne samo pozitivne nego i negativne slike. Tijekom funkcionalnog testiranja ne smije se događati da sijamska neuronska mreža neispravno klasificira negativne slike kao pozitivne. Ovo bi moglo omogućiti da neispravan proizvod bude pogrešno klasificiran kao ispravan (što znači da nema više dodatnih provjera) te da se takav isporuči kupcu. U slučaju neispravnih klasifikacija pozitivnih slika proizvod ide na dodatnu ručnu verifikaciju, gdje ljudska osoba dodatno provjerava proizvod. Ručnom verifikacijom se može otkriti je li problem u proizvodu ili je greška nastala pri neispravnoj klasifikaciji. Zbog tog razloga pomoću mjere preciznosti provjerava se uspješnost sijamske neuronske mreže u zadatku klasificiranja pozitivnih i negativnih slika, a ona za korištenu bazu slika iznosi 94.11%, uzimajući u obzir sve testne grupe. Par slika koji sijamska neuronska mreža nije uspješno klasificirala su parovi koji se nalazi u testnoj grupi *TV_Guide_SATURDAY* te je prikazan na slici 4.16 i slici 4.17 s pripadnim rezultatom različitosti.



Slika 4.16 Neispravno klasificirani par slika testne grupe *TV_Guide_SATURDAY*. (a) referentna slika, (b) negativna testna slika

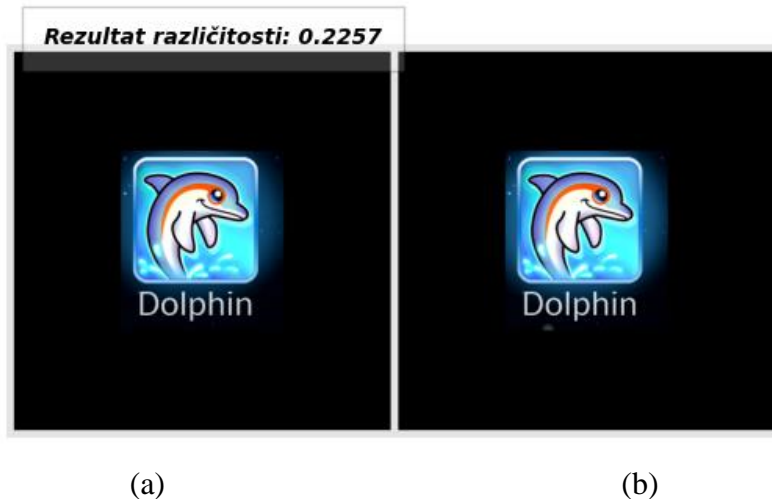


Slika 4.17 Neispravno klasificirani par slika testne grupe TV_Guide_SATURDAY. (a) referentna slika, (b) negativna testna slika

Za usporedbu para slika prikazanih na slici 4.16 i slici 4.17 kao izlaz iz sijamske neuronske mreže dobiven je rezultat različitosti od 0.5037 i 0.0274. Ovaj rezultat ne prelazi postavljeni prag rezultata različitosti koji iznosi 0.8, što znači da će testna slika biti klasificirana kao pozitivna. Kao što je vidljivo na slici 4.16 i slici 4.17, slike se jedino razlikuju u tekstu koji je se nalazi unutar objekta. Oba primjera negativne testne slike predstavljaju neispravno ispunjeno grafičko sučelje te slučajeve kao ovi sijamska neuronska mora uspješno klasificirati, kako bi se moglo reći da je mreža sposobna za funkcionalno testiranje.

Način na koji bi mogli utjecati na poboljšanje rezultata različitosti kod ovakvih je slučajeva koristiti trening slike na kojima je tekst na objektu različit, a njegova okolina jednaka ili dodavanjem slika gdje tekst zauzima veći dio slike. Ovime bi se moglo postići da nedostatak teksta ima veći utjecaj na rezultat različitosti. Još jedan način na koji bi se mogao poboljšati dobiveni rezultat različitosti je dodavanje slika koje imaju jednak dio teksta koji se nalazi na objektu ili se prikazuju jednaki tekst ali drugi objekt.

Na slici 4.18, slici 4.19, slici 4.20, slici 4.21 prikazani su neki od parova referentnih i testnih slike koji su ispravno klasificirani.



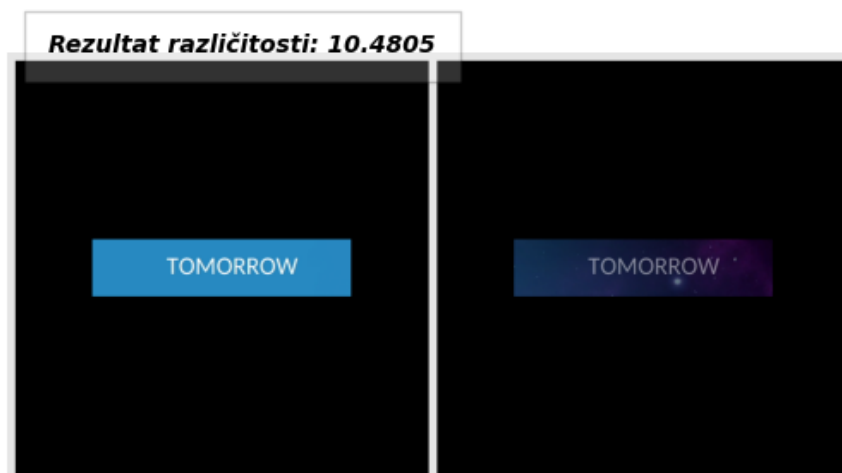
Slika 4.18 Primjer ispravno klasificirane pozitivne testne slike testne grupe *Apps_DOLPHIN*.
(a) referentna slika, (b) pozitivna testna slika

Slika 4.18 prikazuje primjer ispravno klasificirane pozitivne slike. Razlika kod ovih slika je ta što je na testnoj slici pozadina objekta malo osvijetljenija. Ovu razliku u osvijetljenju se može „vidjeti“ i u dobivenom rezultatu različitosti koji je malo povišen i iznosi 0.2257.



Slika 4.19 Primjer ispravno klasificirane negativne testne slike iz testne grupe *Apps_DOLPHIN*.
(a) referentna slika, (b) negativna testna slika

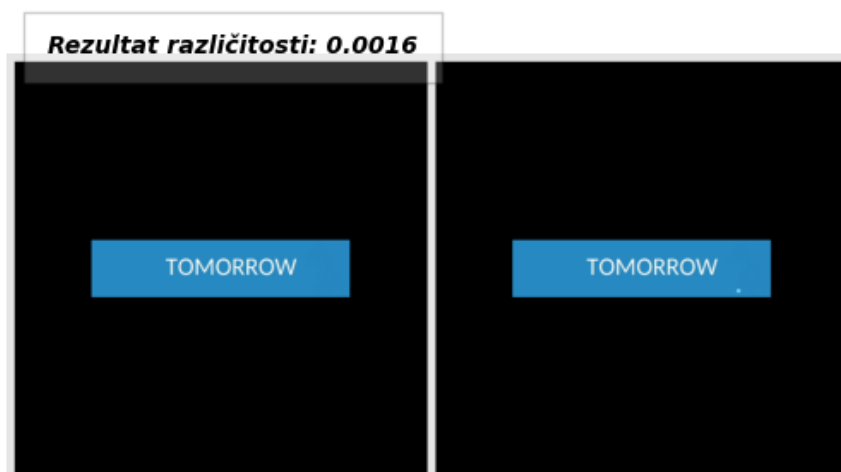
Na slici 4.19 prikazan je primjer kada se referentna slika uspoređuje s testnom negativnom slikom. Sijamska neuronska mreža je usporedila ove slike i kao rezultat različitosti je dobivena vrijednost 4.8519, koja ukazuje na to da su slike različite. Ovaj rezultat je ispravan jer, kao što se može vidjeti, objekti koje slike prikazuju se u potpunosti razlikuju u bojama, tekstu i izborniku koji prikazuju.



(a)

(b)

Slika 4.20 Primjer ispravno klasificirane negativne testne slike iz testne grupe *TV_Guide_TOMORROW*: (a) referentna slika, (b) negativna testna slika



(a)

(b)

Slika 4.21 Primjer ispravno klasificirane pozitivne testne slike iz testne grupe *TV_Guide_TOMORROW* (a) referentna slika, (b) pozitivna testna slika

Na slikama, Slika 4.20 i Slika 4.21, prikazana je usporedba i rezultati različitosti u slučajevima kada je na testnoj slici izabran isti izbornik kao na referentnoj slici (Slika 4.21) i kada na testnoj slici nije izabran isti izbornik kao na referentnoj slici (Slika 4.20). Analizom dobivenih rezultata različitosti može se zaključiti da sijamska neuronska mreža daje dobar rezultat različitosti kada je određeni izbornik izabran i kada isti taj izbornik nije izabran. Kada je na testnoj slici izabran izbornik, a na referentnoj slici isti izbornik nije izabran, dobivamo rezultat različitosti 10.4805. Gledanjem slike 4.20 može se vidjeti kako testna slika ima tamniju pozadinu i slova su sivkaste boje. Također sijamska neuronska mreža dobro uspoređuje slike kada je isti izbornik

odabran na testnoj i referentnoj slici, što je vidljivo na slici 4.21. Rezultat različitosti dobiven usporedbom ove slike iznosi 0.0016, a jedina je razlika između testne i referentne slike bijela točkica koja se nalazi u donjem desnom kutu objekta koji se uspoređuje.

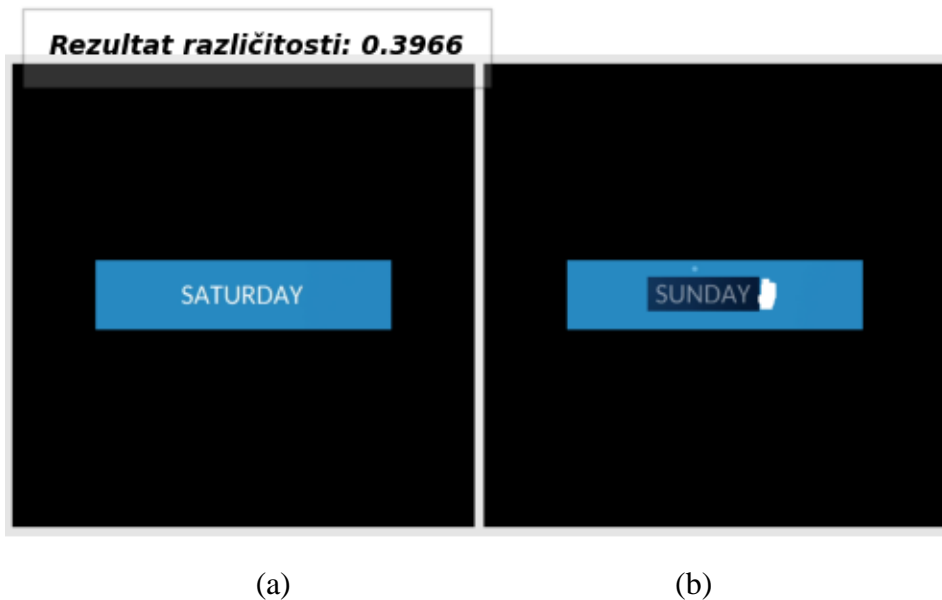
Rezultati dobiveni testiranjem na bazi slika STB grafičkog sučelja pokazuju da sijamska neuronska mreža trenirana na *CG-STB-DI* bazi dinamičnih slika može uspješno klasificirati većinu slika baze slika STB grafičkog sučelja (ako je mali broj slika korišten za testiranje). Jedini parovi slika kod kojih sijamska neuronska mreža ima neispravne klasifikacije negativnih slika, su one slike koje na sebi prikazuju različite tekstove. Dodavanjem dinamičnih slika s tekstom u bazu slika koja se koristi tijekom treniranja, vjerojatno bi se mogle ispraviti neispravne klasifikacije. Slike koje se nalaze u CGIAD bazi dinamičnih slika prikazuju objekte različitih oblika i boja za razliku od baze slika STB grafičkog sučelja gdje su boje slične i objekti koji se uspoređuju su ponajviše četvrtasti. Spomenute razlike između baza i što je sijamska neuronska mreža trenirana na slikama vrlo sličnim testnim slikama iz baze slika STB grafičkog sučelja mogu predstavljati razlog zašto su dobiveni lošiji rezultati testiranja na CGIAD bazi dinamičnih slika.

Kao u potpoglavlju, 4.1 sijamska neuronska mreža trenirana na *CG-STB-CGIAD-DI* testirana je i na bazi slika STB grafičkog sučelja i dobiveni su rezultati koji su prikazani u tablici 4.8.

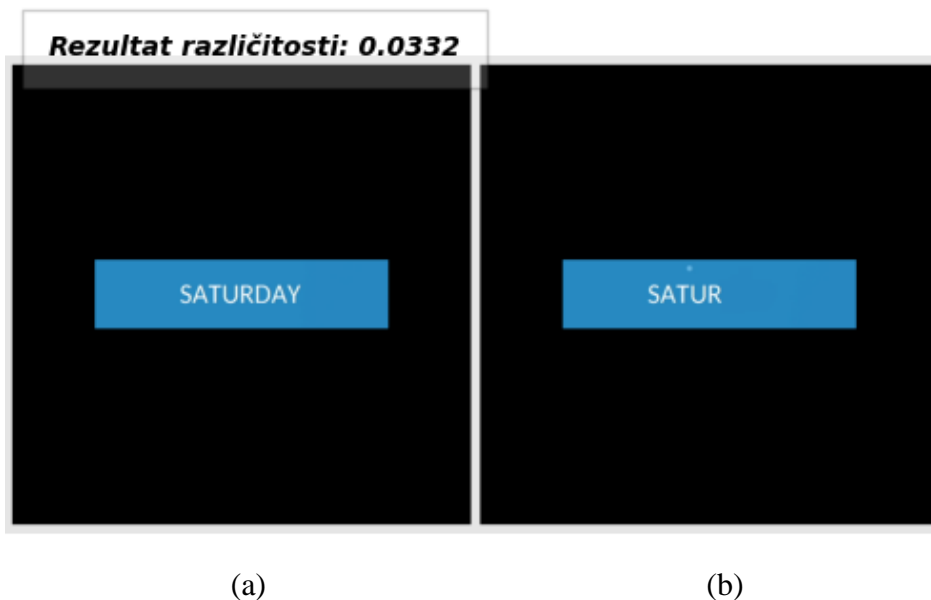
Tablica 4.8 Prikaz rezultata sijamske neuronske mreže trenirane na *CB-STB-CGIAD-DI* bazi slika testirane na bazi slika STB grafičkog sučelja s pragom rezultata različitosti 0.8.

Prag različitosti = 0,8	TP	FN	TN	FP	Preciznost [%]	Odziv [%]	Točnost [%]
Apps_GALLERY	4	0	4	0	100.00	100.00	100.00
Apps_DOLPHIN	4	0	4	0	100.00	100.00	100.00
TV_Guide_POSTER	6	0	6	0	100.00	100.00	100.00
TV_Guide_POSTER_ARTE	6	0	6	0	100.00	100.00	100.00
TV_Guide_SATURDAY	3	0	3	2	60.00	100.00	75.00
TV_Guide_SUNDAY	4	0	3	0	100.00	100.00	100.00
TV_Guide_TODAY	3	0	4	0	100.00	100.00	100.00
TV_Guide_TOMORROW	2	0	4	0	100.00	100.00	100.00
Ukupno	32	0	34	2	94.73	100.00	97.33

Uspoređujući rezultate prikazane na tablici 4.7 i tablici 4.8 može se vidjeti kako nema razlike u dobivenim mjerama preciznosti, odziva i točnosti. Oba modela sijamske neuronske mreže krivo su klasificirale negativne testne slike testne skupine *TV_Guide_SATURDAY*. Neispravno klasificirane slike prikazane se na slici 4.22 i slici 4.23.



Slika 4.22 Neispravno klasificirani par slika testne grupe *TV_Guide_SATURDAY*. (a) referentna slika, (b) negativna testna slika



Slika 4.23 Neispravno klasificirani par slika testne grupe *TV_Guide_SATURDAY*. (a) referentna slika, (b) negativna testna slika

Na slici 4.22 i slici 4.23 prikazane su neispravno klasificirane negativne slike. Slučajevi kao što je ovaj se ne smiju događati tijekom funkcionalnog testiranja grafičkog sučelja nekog proizvoda jer će tada neispravan proizvod biti isporučen kupcu. Razlog neispravnih klasifikacija

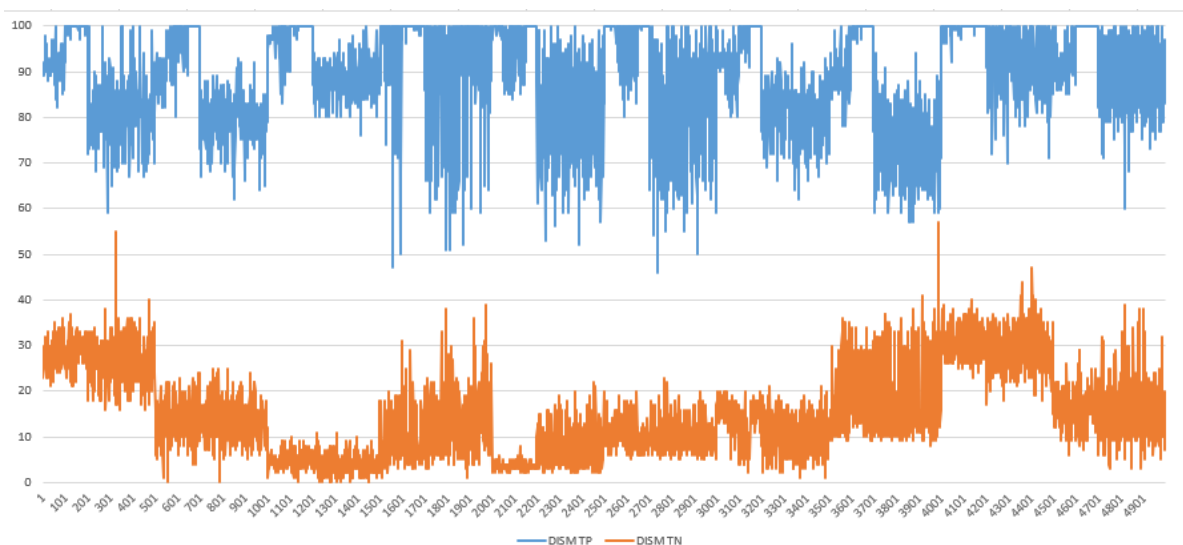
može biti nedovoljno dobra baza dinamičnih slika korištenih za treniranje. Odnosno manjak dinamičnih slika gdje je tekst naglašeniji i gdje različit tekst predstavlja različite slike.

Dobiveni podjednaki rezultati testiranja sijamske neuronske mreže, koja je trenirana na *CG-STB-CGIAD-DI* govore kako baza korištena za treniranje oba modela sijamske neuronske mreže nije dovoljno dobra. Generiranoj bazi dinamičnih slika nedostaje slika gdje je tekst u glavnom fokusu i kada bi se testirale dinamične slike s različitim tekstovima, sijamska neuronska mreža ne bi bila uspješna u njihovoj usporedbi.

4.3. Usporedba rezultata sijamske neuronske mreže i rezultata DISM algoritma

Na temelju rezultata dobivenih testiranjem sijamske neuronske mreže na bazi slika STB grafičkog sučelja i CGIAD bazi dinamičnih slika, može se reći da sijamska neuronska mreža potencijalno predstavlja rješenje za uspoređivanje dinamičnih slika. Promjenama na arhitekturi sijamske neuronske mreže u smislu promjene broja ili veličine filtara, vjerojatno bi se mogli postići još bolji rezultati. Također, poboljšanje u radu sijamske neuronske mreže može se postići korištenjem drugačijih slika za treniranje sijamske neuronske mreže. Kako bi provjerili kako sijamska neuronska mreža radi u usporedbi s postojećim metodama za usporedbu dinamičnih slika, uspoređeni su rezultati sijamske neuronske mreže trenirane na *CG-STB-CGIAD-DI* bazi dinamičnih slika i DISM algoritma dobivenih testiranjem na CGIAD bazi dinamičnih slika i bazi slika STB grafičkog sučelja.

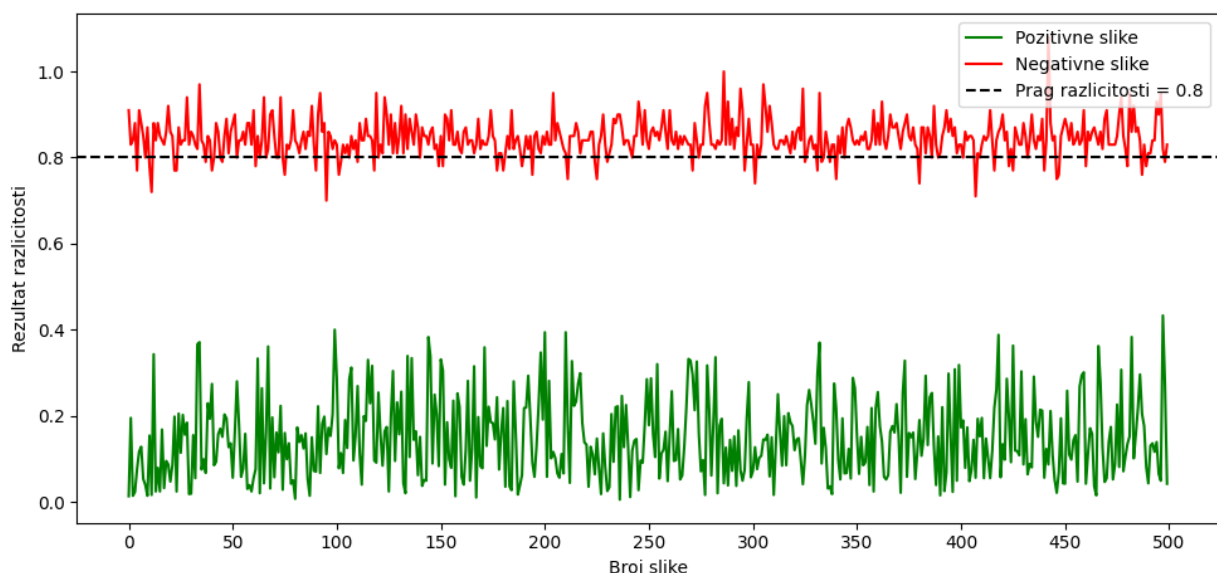
Rezultate koje je DISM algoritam postigao na CGIAD bazi dinamičnih slika mogu se vidjeti na slici 4.24.



Slika 4.24 Rezultati usporedbe slika CGIAD baze dinamičnih slika pomoću DISM algoritma [17]

DISM algoritam je na skupu od 10000 slika CGIAD baze dinamičnih slika postigao točnost od 99.66%, preciznost od 100% i odziv od 99.31%. Prag postotka sličnosti slika koji je morao biti premašen iznosi 60%. Uz tu vrijednost praga postotka sličnost, DISM algoritam 34 slike klasificirao je kao FN. Rezultati koje je sijamska neuronska mreža postigla na CGIAD bazi dinamičnih slika prikazani su u tablici 4.3. U tablici se može vidjeti da je sijamska neuronska mreža na skupu od 10000 slika s pragom rezultata različitosti postavljenim na vrijednost 0.8, 362 slike klasificirala kao FN i 272 slike klasificirala kao FP. Kao što je spomenuto u potpoglavlju 4.1 prilikom funkcionalnog testiranja, vrlo je bitno da ne bude slika koje su klasificirane kao FP. Jer ako se pojave slike koje su klasificirane kao FP, dogodit će se slučajevi kada će se kupcu isporučiti neispravan proizvod. Sama ideja funkcionalnog testiranja je da se ispitaju sve situacije i da se kupcu isporučuje ispravan proizvod.

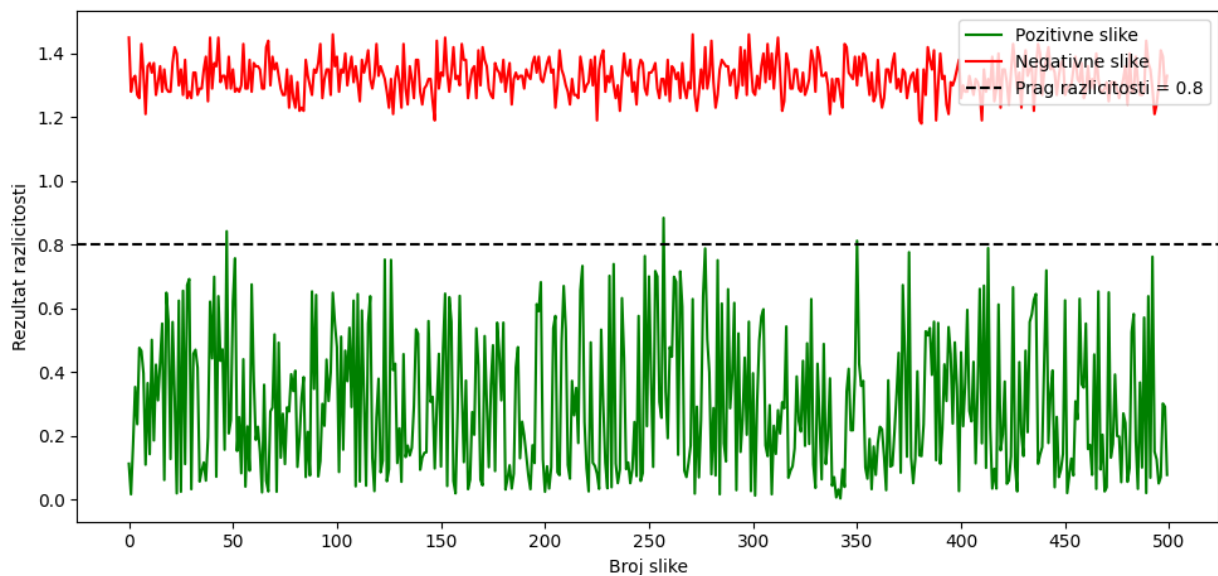
Samom usporedbom dobivenih rezultata može se uočiti kako je DISM algoritam uspješniji od sijamske neuronske mreže. Ako se uzme u obzir da se svaka od 10 testnih grupa CGIAD baze sastoji od 500 pozitivnih i 500 negativnih slika, otkriva se za koje testne grupe DISM daje bolje rezultate u odnosu na sijamsku neuronsku mrežu. Za testnu grupu 1 DISM algoritam za većinu negativnih slika dobiva postotak sličnosti koji prelazi 20%, što bi značilo da su negativne slike blago slične referentnoj slici, ali ipak nedovoljno da bi bile klasificirane kao pozitivne. Na slici 4.25 prikazan je graf rezultata različitosti dobiven pomoću sijamske neuronske mreže za testnu grupu 1 CGIAD baze dinamičnih slika.



Slika 4.25 Graf rezultata različitosti dobiven sijamskom neuronskom mrežom za testnu grupu 1 CGIAD baze dinamičnih slika

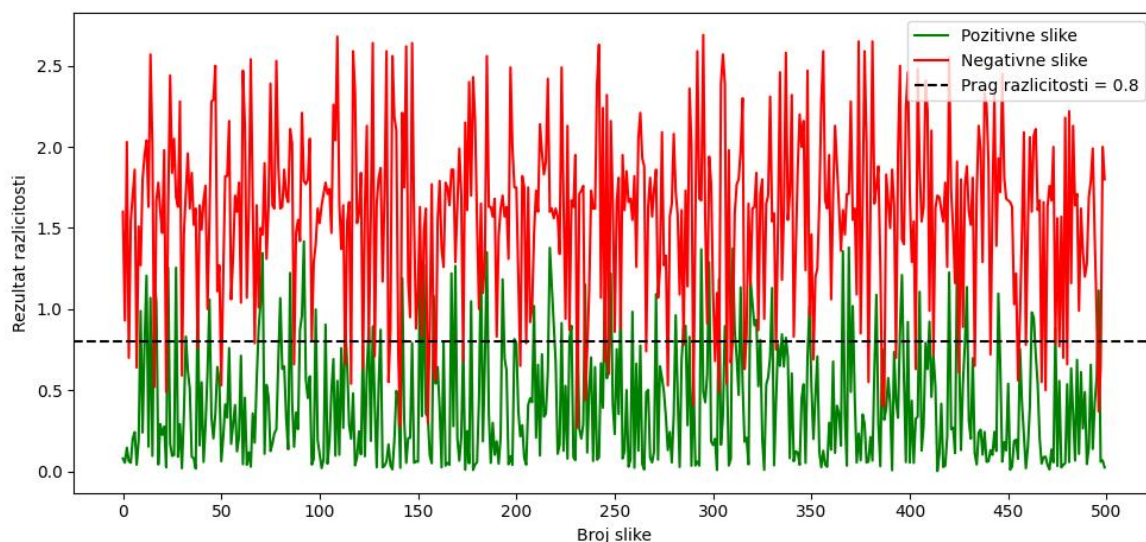
Sličan slučaj (Slika 4.25) se za istu testnu grupu događa i kod sijamske neuronske mreže, gdje je 63 slike klasificirano kao FP. Negativne testne slike klasificirane su kao FP ako se nalaze

ispod crne isprekidane linije, koja predstavlja postavljeni prag različitosti. Isto tako, pozitivne testne slike su klasificirane kao FN, ako se nalaze iznad isprekidane crne linije. Na slici crvena linija označava negativne testne slike, zelena linija označava pozitivne testne slike. Kao što je spomenuto u potpoglavlju 4.1., razlog pogrešnog klasificiranja je sličnost oblika i rasporeda objekata na referentnoj i negativnoj slici. Ovo je vidljivo u rezultatima različitosti za negativne slike koji se najviše nalaze oko rezultata različitosti od 0,85, ali za nekoliko testnih negativnih slika rezultat različitosti iznosi oko 0,75. Na slici 4.26 prikazani su rezultati koje sijamska neuronska mreža postiže na testnoj grupi 8 CGIAD baze dinamičnih slika.



Slika 4.26 Graf rezultata različitosti dobiven sijamskom neuronskom mrežom za testnu grupu 8 CGIAD baze dinamičnih slika

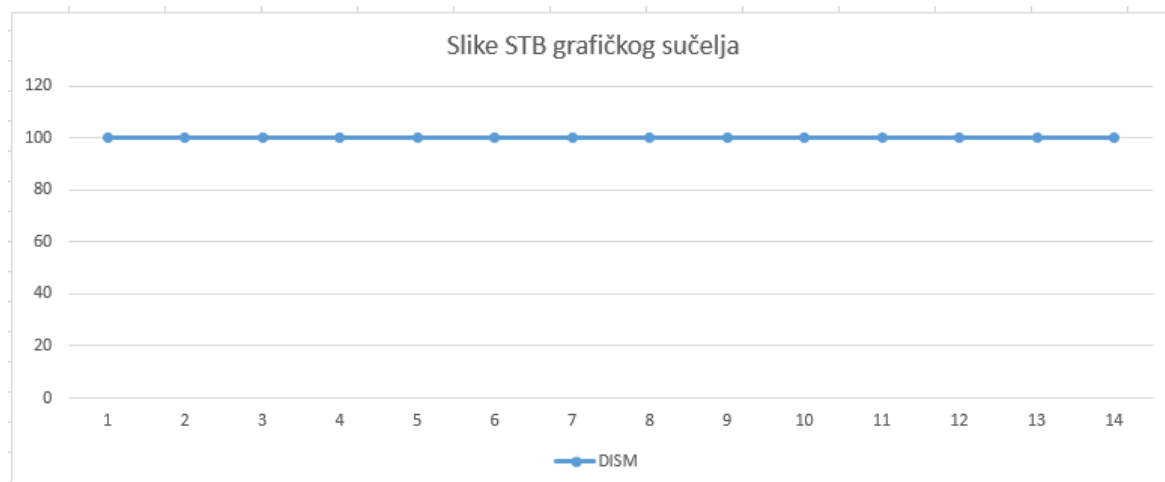
Sijamska neuronska mreža za testnu grupu 8 CGIAD baze dinamičnih slika, za mjere preciznosti dobiva vrijednost od 100%, za mjeru odziva vrijednost od 99.40%, a za mjeru točnosti vrijednost od 99.70%. DISM algoritam za istu testnu grupu dobiva preciznost od 100%, točnost od 99.91% i odziv od 98.19%. DISM algoritam je 9 pozitivnih testnih slika od svih pozitivnih testnih slika testne grupe 8 klasificirao kao negativne, a sijamska neuronska mreža je 3 pozitivne testne slike od svih pozitivnih testnih slika testne grupe 8 ispravno klasificirala. Iz ovoga se može izvući zaključak kako sijamska neuronska mreža i DISM algoritam podjednako uspoređuju slike ove testne grupe. Na slici 4.27 prikazan je graf rezultata različitosti dobiven sijamskom neuronskom mrežom za testnu grupu 4 CGIAD baze dinamičnih slika.



Slika 4.27 Graf rezultata različitosti dobiven sijamskom neuronskom mrežom za testnu grupu 4 CGIAD baze dinamičnih slika

Kod testne grupe 4 CGIAD baze dinamični slika, DISM algoritam za negativne slike dobiva postotke sličnosti s referentnom slikom veće od 20%, što opet govori da su objekti na negativnim slikama u nekoj manjoj mjeri slični objektu na referentnoj slici. Sijamska neuronska mreža 58 negativnih testnih slika i 86 pozitivnih testnih slika neispravno klasificira. DISM algoritam je 10 testnih pozitivnih slika neispravno klasificirao. Broj neispravno klasificiranih slika kod sijamske neuronske mreže i DISM algoritma te povećana postotak sličnosti za negativne slike kod DISM algoritma, govore kako je testna grupa 4 najizazovnije za oba načina usporedbe dinamičnih slika. Ipak DISM algoritam na ovoj testnoj grupi ima bolje mjere preciznosti, odziva i točnosti te time predstavlja bolji algoritam za usporedbu dinamičnih slika.

Druga baza na kojoj su DISM algoritam i sijamska neuronska mreža uspoređeni je baza slika STB grafičkog sučelja. Dio slika baze STB grafičkog sučelja korištenih u testiranju DISM algoritma korišten je za treniranje sijamske neuronske mreže korištene u ovom radu. Iz tog razloga rezultati dobiveni uspoređivanjem tih slika pomoću DISM algoritma se ne mogu koristiti. Slike koje se mogu koristiti pripadaju testnim grupama *TV_Guide_SATURDAY*, *TV_Guide_SUNDAY*, *TV_Guide_POSTER*. Za testnu grupu *TV_Guide_SATURDAY* korištene su 2 testne slike, *TV_Guide_SUNDAY* 3 testne slike, *TV_Guide_POSTER* 9 testnih slika, što ukupno čini 14 slika. Graf rezultata dobivenih DISM algoritmom prikazan je na slici 4.28.



Slika 4.28 Graf postotka sličnosti slika dobivenih DISM algoritmom na STB bazi slika grafičkog sučelja

Postotci sličnosti su poredani redom, 2 slike za testnu grupu *TV_Guide_SATURDAY*, 3 za testnu grupu *TV_Guide_SUNDAY* i 9 za testnu grupu *TV_Guide_POSTER*. Kao što se vidi na slici 4.28, za sve slike ovih testnih grupa DISM algoritam prepoznao je da se radi o jednakim slikama. Kada se iste slike usporede pomoću sijamske neuronske mreže dobiju se rezultati prikazani na slici 4.29.



Slika 4.29 Rezultati različitosti slika dobivenih pomoću sijamske neuronske mreže na STB bazi slika grafičkog sučelja

Analizom rezultata sa slike 4.29 može se vidjeti kako je sijamska neuronska mreža, također za sve testne slike STB grafičkog sučelja dala rezultate različitosti koji govore da se na referentnim i testnim slikama nalaze isti objekti. Rezultati različitosti su poredani istim redom kao na slici 4.28.

Rezultati testiranja DISM algoritma i sijamske mreže su jednaki te su testirani na malom skupu podataka. Iz tih razloga ih se ne može pouzdano usporediti pomoću baze slika STB grafičkog sučelja, ali korištenjem rezultata testiranja na CGIAD bazi dinamičnih slika može se zaključiti kako sijamska neuronska mreža predložena u ovom radu ne daje bolje rezultate nego postojeći DISM algoritam. Korištenjem sijamske neuronske mreže kao specifičnog rješenja za

usporedbu dinamičnih slika mogu se dobiti precizniji i bolji rezultate, ali još uvijek ne bolji rezultati nego DISM algoritam. Pravljenjem ranije predloženih izmjena arhitekture i proširenja baze slika korištene za treniranje, mogli bi se postići rezultati usporedbe dinamičnih slika koji su slični rezultatima usporedbe dinamičnih slika dobivenih pomoću DISM algoritma. Još jedna prednost koju DISM algoritam ima nad sijamskom neuronskom mrežom je ta što se sijamska neuronska mreža prvo mora istrenirati na određenoj bazi i mora se čuvati datoteka koja sadrži model istrenirane sijamske neuronske mreže. DISM algoritam ne treba trenirati niti čuvati nikakve dodatne datoteke. Ako bi trenirali sijamsku neuronsku mrežu kao specifično rješenje, kada bi se pojavili novi objekti bilo bi potrebno ponovno trenirati neuronsku mrežu. Ovaj postupak dodatnog treniranja zahtjeva još resursa i čini sijamsku neuronsku mrežu kao specifično rješenje, nepraktičnom. Iz ovih razloga DISM algoritam još uvijek predstavlja najbolje rješenje za usporedbu dinamičnih slika od svih javno dostupnih rješenja.

5. Zaključak

Kako bi se ubrzalo funkcionalno testiranje STB aplikacija, proces testiranja pokušava se automatizirati. Jedan od testova funkcionalnog testiranja aplikacija je provjera ispunjava li se grafičko sučelje kako treba. Klasično uspoređivanje referentnog i testnog grafičkog sučelja više nije dovoljno zbog raznih animacija koje su se počele pojavljivati te učinile grafička sučelja dinamičnima. U ovom diplomskom radu izrađeno je jedno od mogućih rješenja za automatizaciju funkcionalnog testiranja aplikacija s dinamičnim grafičkim sučeljem. Izrađena je sijamska neuronska mreža koja će primati sliku s referentnog i s testnog grafičkog sučelja te ih usporediti i kao rezultat usporedbe dati rezultat različitosti. Pomoću dobivenog rezultata različitosti moći će se reći radi li grafičko sučelje ispravno. Kako bi mogli istrenirati sijamsku neuronsku mrežu, napravljen je alat za generiranje dinamičnih slika pomoću kojega je napravljene su dvije baze dinamičnih slika koje su korištene za treniranje sijamske neuronske mreže. Ove baze nazivaju se *CG-STB-DI* i *CG-STB-CGIAD-DI*.

Sijamska neuronska mreža koja je trenirana na *CG-STB-DI* i sijamska neuronska mreža koja je trenirana na *CG-STB-CGIAD-DI* testirane su na dvjema bazama dinamičnih slika. Prvi test izveden je na CGIAD bazi dinamičnih slika. Sijamska neuronska mreža trenirana na *CG-STB-DI* bazi, na CGIAD bazi dinamičnih slika dobila je preciznost od 85.76%, odziv od 73.64% i točnost od 80.71%. Kao mogući razlozi neispravnih klasifikacija pozitivnih i negativnih slika mogu biti korištenje slične komponente boja i slični oblici na objektima koji se nalaze na testnim i referentnim slikama. Kada je testirana sijamska neuronska mreža trenirana na *CG-STB-CGIAD-DI* dobivena je preciznost od 94.44%, odziv od 92.76% i točnost od 93.66%. Oba modela sijamske neuronske mreže zatim su testirana na bazi slika STB grafičkog sučelja na kojoj su oba modela postigla preciznost od 94.73%, odziv od 100% i točnost od 97.33%. Od ukupno 32 pozitivne slike, sve su slike uspoređene i ispravno klasificirane kao pozitivne, dok od ukupno 36 negativnih slika koje su uspoređene, dvije su neispravno klasificirane kao pozitivne.

Dopunjavanjem baze korištene za trening i mijenjanjem arhitekture sijamske neuronske mreže pretpostavlja se da bi se mogli dobiti bolji rezultati. Dobiveni rezultati na testiranjima sijamske neuronske mreže na CGIAD bazi dinamičnih slika i bazi slika STB grafičkog sučelja uspoređeni su s rezultatima dobivenih pomoću DISM algoritma. Usporedbom rezultata može se zaključiti kako je sijamska neuronska mreža manje precizna od DISM algoritma na CGIAD bazi dinamičnih slika. Na bazi slika STB grafičkog sučelja sijamska neuronska mreža i DISM algoritam daju podjednake rezultate. Uzimanjem u obzir svih dobivenih rezultata testiranja DISM algoritma, sijamska neuronska mreža nije bolje rješenje od postojećeg DISM algoritma koji ima veću

preciznost, odziv i točnost od sijamske neuronske mreže. Uz veću preciznost, odziv i točnost DISM algoritam nije potrebno trenirati niti čuvati datoteku istrenirane sijamske neuronske mreže. Sijamska neuronska mreža predstavlja potencijalno rješenje za uspoređivanje dinamičnih slika. Kako bi se za uspoređivanje dinamičnih slika komercijalno mogla koristiti predložena sijamska neuronska mreža, potrebna su znatna unaprjeđenja u obliku promjene arhitekture neuronske mreže i nadogradnje baze dinamičnih slika koje se koriste za treniranje. Sva spomenuta unaprjeđenja sijamske neuronske mreže biti će dio budućeg rada.

LITERATURA

- [1] M. Pul, V. Perković, M. Vranješ i R. Grbić, »Computer Generated Image Alterations Database,« [Mrežno]. Available: <https://www.rt-rk.com/other/CGIAD.html>. [Pokušaj pristupa Studeni 2021].
- [2] »C-MORE UI/UX,« [Mrežno]. Available: <https://www.youtube.com/watch?v=Y6xv8wL0zDU>. [Pokušaj pristupa 25 Lipanj 2022].
- [3] P. J. Phillips, H. Wechsler, J. Huang i P. J. Rauss, »The FERET database and evaluation procedure for face-recognition algorithms,« *Image and Vision Computing Journal*, svez. 16, br. 5, pp. 295-306, 1998.
- [4] S. Chopra, R. Hadsell i Y. LeCun, »Learning a Similarity Metric Discriminatively, with Application to Face Verification,« u *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, 2005.
- [5] A. Martinez i R. Benavente, »The AR Face Database,« *CVC Technical Report*, br. 24, 1998.
- [6] »The Database of Faces,« AT&T Laboratories Cambridge, 2001. [Mrežno]. Available: <https://cam-orl.co.uk/facedatabase.html>. [Pokušaj pristupa 25 Lipanj 2022].
- [7] University at Buffalo, »CEDAR,« [Mrežno]. Available: <http://www.cedar.buffalo.edu/NIJ/data/signatures.rar>. [Pokušaj pristupa 19 Srpanj 2022].
- [8] R. Plamondon i S. N. Srihari, »Online and off-line handwriting recognition: a comprehensive survey,« *IEEE Transactions on Pattern Analysis and Machine Intelligence*, svez. 22, br. 1, pp. 63-84, 2000.
- [9] M. Ajij, S. Pratihari, S. R. Nayak, T. Hanne i D. S. Roy, »Off-line signature verification using elementary combinations of directional codes from boundary pixels,« 2021.
- [10] Universidad de Las Palmas de Gran Canaria, »GPDS,« [Mrežno]. Available: <https://gpds.ulpgc.es/>. [Pokušaj pristupa 18 Srpanj 2022].
- [11] D. Impedovo i G. Pirlo, »Automatic Signature Verification: The State of the Art,« *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, svez. 38, br. 5, pp. 609-635, 2008.
- [12] D. G. Lowe, »Distinctive Image Features from Scale-Invariant Keypoints,« *International Journal of Computer Vision*, svez. 60, pp. 91-110, 2004.
- [13] H. Bay, T. Tuytelaars i L. Van Gool, »SURF: Speeded Up Robust Features,« *Computer Vision and Image Understanding*, svez. 110, br. 3, pp. 346-359, 2008.
- [14] E. Rublee, V. Rabaud, K. Konolige i G. Bradski, »ORB: an efficient alternative to SIFT or SURF,« u *International Conference on Computer Vision*, Barcelona, 2011.
- [15] M. Calonder, V. Lepetit, C. Strecha i P. Fua, »BRIEF: Binary Robust Independent Elementary Features,« u *11th European Conference on Computer Vision*, Heraklion, Crete, Greece, 2010.
- [16] E. Rosten i T. Drummond, »Machine Learning for High-Speed Corner Detection,« u *European Conference on Computer Vision*, Berlin, 2006.
- [17] M. Pul, V. Perković, M. Vranješ i R. Grbić, »Automatic Functionality Verification of Hybrid Set-Top Boxes with Dynamic User Interface,« u *IEEE Transactions on Consumer Electronics*.

- [18] M. Pul, »Usporedba dinamičnih slika zasnovana na izdvajanju i usporedbi ključnih točaka,« Osijek, 2016.
- [19] J. Bromley, I. Guyon, LeCun Yann, E. Sackinger i R. Shah, »Signature Verification using a "Siamese" Time Delay Neural Network,« u *Neural Information Processing Systems Conference*, Denver, 1994.
- [20] S. Dey, A. Dutta, J. I. Toledo, S. K. Ghosh, J. Lladós i U. Pal, »SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification,« *arXiv*, 2017.
- [21] R. Grbić i M. Vranješ, »predavanje sa kolegija Strojno učenje u sustavima automobila i umreženih vozila "Uvod u duboke mreže. Konvolucijske neuronske mreže",« u *FERIT*, Osijek, 2022.
- [22] OpenCV.org, [Mrežno]. Available: <https://opencv.org/about/>. [Pokušaj pristupa 28 Lipanj 2022].
- [23] »towardsdatascience.com,« [Mrežno]. Available: <https://towardsdatascience.com/introduction-to-py-torch-13189fb30cb3>. [Pokušaj pristupa 28 Lipanj 2022].
- [24] D. P. Kingma i J. Ba, »Adam: A Method for Stochastic Optimization,« 29 Siječanj 2017. [Mrežno]. Available: <http://arxiv.org/abs/1412.6980>. [Pokušaj pristupa 11 Srpanj 2022].

SAŽETAK

Dinamične slike predstavljaju one slike koje imaju jednaki kontekst, ali različit sadržaj. To znači da dinamične slike prikazuju isti objekt, samo je na jednoj od slika objekt zarotiran, pomaknut, smanjen, povećan ili se ispred objekta nalazi nešto što ga djelomično prekriva. Za potrebe ovoga rada napisan je algoritam za generiranje dinamičnih slika. Ovaj generator generira slike tako da vrši neku od transformacija nad objektom. Pomoću ovog generatora kreirane su baze dinamičnih slika korištene za treniranje sijamske neuronske mreže koja je izrađena u sklopu ovog rada. Sijamska neuronska mreža je vrsta neuronske mreže koja ima sposobnost uspoređivanja dinamičnih slika. Algoritam za generiranje dinamičnih slika i algoritmi za treniranje i testiranje sijamske neuronske mreže napisani su u programsku jeziku *Python*. Sijamska neuronska mreža testirana je na CGIAD bazi dinamičnih slika te su dobiveni rezultati preciznosti od 94.46%, odziva od 92.76% i točnosti od 93.66%. Testiranjem sijamske neuronske mreže pokazano je da sijamska neuronska mreža predstavlja potencijalno rješenje za usporedbu dinamičnih slika, iako trenutno nije bolja u tom zadatku od najboljeg javno dostupnog DISM algoritma. Modifikacijama arhitekture sijamske neuronske mreže i baze dinamičnih slika korištenih za treniranje potencijalno bi se mogle povećati performanse mreže.

Ključne riječi: dinamične slike, usporedba dinamičnih slika, neuronska mreža, DISM algoritam

DYNAMIC IMAGES COMPARISON USING NEURAL NETWORKS

ABSTRACT

Dynamic images represent those images that have the same context but different content. This means that dynamic images show the same object, but in one of the images object is rotated, moved, shrunken, enlarged or there is something in front of the object that partially covers it. For the purposes of this work, an algorithm for generating dynamic images was designed. This generator generates images in a way that makes some of the transformations on the object. With this generator, databases of dynamic images used for training the siamese neural network were made. Siamese neural network is a type of neural network that has the ability to compare dynamic images. Algorithms for generating dynamic images and algorithms for training and testing siamese neural network were written in the programming language Python. Siamese neural network was tested on CGIAD dynamic image database and results of 94.46% for precision, 92.76% for recall and 93.66% for accuracy were obtained. Testing of the siamese neural network showed that it could be potentially used for dynamic image comparison, but it currently isn't better in that task than the DISM algorithm. Modifications on siamese neural network architecture and dynamic image database, which was used for training, would increase network performance.

Key words: dynamic images, dynamic image comparison, neural network, DISM algorithm

ŽIVOTOPIS

Domagoj Steiner rođen je u Požegi 11.08.1998. Osnovnu školu završio u Orahovici 2013. godine. Nakon osnovne škole upisuje Opću gimnaziju u Orahovici. Nakon završene srednje škole upisuje se na preddiplomski studij Računarstva na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku. Godine 2020. završava preddiplomski studij Računarstva i upisuje diplomski smjer Automobilskog računarstva i komunikacija. Tijekom završne godine diplomskog studija, postaje stipendist tvrtke TTTechAuto d.o.o.

Potpis autora

PRILOZI

Prilog P.3.1. Programski kod algoritma za generiranje dinamičnih slika sa svim potrebnim datotekama (elektronički prilog).

Prilog P.3.2. Baza *CG-STB-DI* i baza *CG-STB-CGIAD-DI* korištene za treniranje sijamske neuronske mreže predložene u ovom radu (elektronički prilog).

Prilog P.3.3. Programski kod pomoću kojega se kreira, trenira i testira sijamska neuronska mreža te sve potrebne dodatne datoteke (elektronički prilog)