

DETEKTIRANJE ARTEFAKTA SMRZAVANJA SLIKE U VIDEO SIGNALU ZASNOVANO NA STROJNOM UČENJU

Vukadin, 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:200:262075>

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

Download date / Datum preuzimanja: **2025-03-27**

Repository / Repozitorij:

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



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

Sveučilišni studij

**DETEKTIRANJE ARTEFAKTA SMRZAVANJA SLIKE U
VIDEO SIGNALU ZASNOVANO NA STROJNOM
UČENJU**

Diplomski rad

Domagoj Vukadin

Osijek, 2022.

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

Osijek, 15.09.2022.

Odboru za završne i diplomske ispite

Imenovanje Povjerenstva za diplomski ispit

Ime i prezime Pristupnika:	Domagoj Vukadin
Studij, smjer:	Diplomski sveučilišni studij Automobilsko računarstvo i komunikacije
Mat. br. Pristupnika, godina upisa:	D-55ARK, 11.10.2020.
OIB studenta:	01537520837
Mentor:	Prof. dr. sc. Marijan Herceg
Sumentor:	,
Sumentor iz tvrtke:	Danijel Babić
Predsjednik Povjerenstva:	Izv. prof. dr. sc. Mario Vranješ
Član Povjerenstva 1:	Prof. dr. sc. Marijan Herceg
Član Povjerenstva 2:	Izv.prof.dr.sc. Ratko Grbić
Naslov diplomskog rada:	DETEKTIRANJE ARTEFAKTA SMRZAVANJA SLIKE U VIDEO SIGNALU ZASNOVANO NA STROJNOM UČENJU
Znanstvena grana diplomskog rada:	Telekomunikacije i informatika (zn. polje elektrotehnika)
Zadatak diplomskog rada:	U prijenosu video sadržaja mrežom može doći do gubitka ili kašnjenja paketa koji prenose video signal. Gubitak ili kašnjenje paketa uzrokuje artefakte (oštećenja) koji se vizualno manifestiraju pri reprodukciji prenesenog video signala. Jedan od artefakata koji se pri tome javlja je artefakt smrzavanja slike u video signalu. Artefakt se manifestira na način da se pri reprodukciji ponavlja posljednji ispravno preneseni okvir video signala. Umjesto ponavljanja posljednjeg ispravno prenesenog okvira videa, može se ponavljati jednobojni okvir videa (najčešće crn ili zelen), ali takve je slučajeve vrlo lako detektirati. Postoje prijenosni sustavi kod kojih nije poznat izvorni video sadržaj te tada govorimo o prijenosnim sustavima bez referenc. Kod takvih sustava detekcija
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:	15.09.2022.
Potvrda mentora o predaji konačne verzije rada:	Mentor elektronički potpisao predaju konačne verzije.
	Datum:

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

Osijek, 29.09.2022.

Ime i prezime studenta:

Domagoj Vukadin

Studij:

Diplomski sveučilišni studij Automobilsko računarstvo i komunikacije

Mat. br. studenta, godina upisa:

D-55ARK, 11.10.2020.

Turnitin podudaranje [%]:

3

Ovom izjavom izjavljujem da je rad pod nazivom: **DETEKTIRANJE ARTEFAKTA SMRZAVANJA SLIKE U VIDEO SIGNALU ZASNOVANO NA STROJNOM UČENJU**

izrađen pod vodstvom mentora Prof. dr. sc. Marijan Herceg

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:

SADRŽAJ

1. UVOD	1
2. DETEKCIJA ARTEFAKTA SMRZAVANJA U VIDEO SIGNALU	3
2.1. Artefakt smrzavanja	3
2.2. Pristup detekciji artefakta smrzavanja.....	6
2.2.1. Klasični pristup detekcije smrzavanja.....	6
2.2.2. Pristup na temelju strojnog učenja	7
2.3. Baze video sekvenci koje sadrže artefakt smrzavanja	8
2.4. Dodatne video sekvence.....	9
2.5. Postojeća rješenja za detekciju artefakta smrzavanja slike u video signalu.....	9
3. IZRADA RJEŠENJA ZA DETEKCIJU ARTEFAKTA SMRZAVANJA SLIKE U VIDEO SEIGNALU NA TEMELJU STROJNOG UČENJA.....	12
3.1. Radno okruženje.....	12
3.2. Proces izrade baza, arhitektura i treninga.....	12
3.2.1. Stvaranje podataka za treniranje	12
3.2.2. Arhitekture neuronske mreže	14
3.2.3. Proces treniranja neuronskih mreža	17
3.3. Upute za pokretanje.....	19
4. EVALUACIJA I VERIFIKACIJA IZRAĐENIH RJEŠENJA TEMELJENIH NA STROJNOM UČENJU	22
4.1. Verifikacija na LIVE Mobile bazi.....	24
4.2. Verifikacija na FERIT-RTRK UHD VFD bazi.....	25
4.3. Verifikacija na VQEG HDTV bazi	26
4.4. Verifikacija na SMVS FullHD bazi	27
4.5. Analiza rezultata.....	28
5. ZAKLJUČAK	29
LITERATURA.....	30
SAŽETAK.....	32

ABSTRACT	33
ŽIVOTOPIS	34
PRILOZI.....	35

1. UVOD

Razvoj komunikacijskih tehnologija za prijenos multimedijских sadržaja omogućio je pristup video sadržajima putem mreže, zbog čega korisnici očekuju da je taj pristup moguć u bilo kojemu trenutku s bilo kojega mjesta. Moderni pametni telefoni, tableti i televizije odlikuju se zaslonom visoke razlučivosti koji korisnicima pruža kvalitetan prikaz video sadržaja, bilo da se radi o gledanju serija, filmova ili izravnog prijenosa uživo. Gotovo svi korisnici očekuju najbolju kvalitetu u svakom trenutku, osobito oni koji plaćaju pretplatu za gledanje određenog sadržaja. Visoka razlučivost video sadržaja i zaslona čine greške kompresije i prijenosa vidljivijima nego ikada prije. Greške u kompresiji i prijenosu video signala uzrokuju degradaciju video signala, a specifične degradacije nazivaju se artefaktima. Neki od artefakta su artefakt stvaranja blokova, artefakt zamučivanja, artefakt smrzavanja, itd. Artefakti stvaranja blokova i zamučivanja primjer su artefakata koji nastaju zbog kompresije s gubicima, dok je artefakt smrzavanja primjer artefakta koji nastaje zbog prijenosa s gubicima. Naglasak na visoku kvalitetu potaknuo je razvoj metoda za utvrđivanje i kontrolu kvalitete video signala (engl. *Video Quality Assessment*, VQA). Metode za ocjenu kvalitete video signala se oslanjaju na detekciju određenih artefakata prisutnih u video signalu koji narušavaju kvalitetu sadržaja, a izravno utječu na kvalitetu iskustva (eng. *Quality of Experience*, QoE) krajnjeg korisnika. Za ocjenu kvalitete video signala koriste se subjektivne i objektivne metode. Subjektivne metode zahtijevaju ljudske resurse, odnosno ljude koji će pregledati video sekvence koje sadrže različite degradacije te uz svaku video sekvencu pridružiti ocjenu. Objektivne metode za ocjenu kvalitete video signala ne zahtijevaju ljudske resurse jer se ocjena kvalitete video signala donosi na temelju matematičkih modela. Ljudska (subjektivna) ocjena kvalitete video signala smatra se točnijom od matematičkih modela. S druge strane, provođenje subjektivnih metoda je znatno skuplje naspram objektivnih metoda stoga se potiče razvoj i unaprjeđivanje objektivnih metoda sa svrhom ostvarivanja rezultata što sličnijih subjektivnim metodama. Korištenje objektivnih metoda u stvarnome vremenu isključuje pristup izvornom nekomprimiranom video signalu, stoga se najčešće koristi podskupina objektivnih metoda za ocjenu kvalitete video signala bez pristupa referentnom video signalu. Takve metode isključivo se oslanjaju na primljeni video signal.

Zadatak ovog rada je napraviti rješenje koje detektira artefakt smrzavanja slike u video signalu na temelju strojnog učenja. Potrebno je testirati različite arhitekture neuronskih mreža i usporediti njihove rezultat na testnom skupu podataka.

U drugom poglavlju se nalazi opis artefakta smrzavanja i pristupa za detekciju istoga. Dan je pregled javno dostupnih baza video sekvenci koje sadrže artefakt smrzavanja i pregled postojećih algoritama za detekciju artefakta smrzavanja bez pristupa referentnom sadržaju. Treće poglavlje sadrži opis vlastitog rješenja zasnovanog na strojnom učenju. U četvrtom poglavlju predstavljeni su rezultati novih rješenja te usporedba s rezultatima postojećih algoritama. Peto poglavlje donosi zaključak na temelju cjelokupnog rada.

2. DETEKCIJA ARTEFAKTA SMRZAVANJA U VIDEO SIGNALU

U ovom poglavlju opisan je nastanak artefakta smrzavanja čija pojava uzrokuje degradaciju video signala, a samim time utječe na smanjenje kvalitete video signala. Detaljan opis artefakta smrzavanja nalazi se u potpoglavlju 2.1. Smanjena kvaliteta video signala narušava iskustvo krajnjeg korisnika, stoga je nužno detektirati pojavu opisanog artefakta na korisničkoj strani. U svrhu detekcije najčešće se koriste metode koje nemaju pristup referentnom video signalu, već samo primljenom. U potpoglavlju 2.2. opisan je klasičan pristup detekcije artefakta smrzavanja te ideja i obrazloženje pristupa zasnovanog na strojnom učenju. Obrazloženi su i problemi koje je potrebno uzeti u obzir pri razvoju rješenja za detekciju artefakta smrzavanja. U potpoglavlju 2.3. opisane su dostupne baze video sekvenci koje sadrže artefakt smrzavanja na temelju kojih je razvijeno rješenje u 3. poglavlju i obavljena verifikacija i usporedba rezultata u 4. poglavlju. U potpoglavlju 2.5. nalazi se opis postojećih rješenja, a detaljnije su opisani algoritmi predloženi radovima [1] i [2] čija se implementacija koristi u 4. poglavlju za usporedbu rezultata s rezultatima rješenja predloženih ovim radom u 3. poglavlju.

2.1. Artefakt smrzavanja

U sustavima za prijenos multimedijских podataka zahtijeva se obrada podataka u stvarnom vremenu. Kako bi prijenos video signala u stvarnom vremenu bio moguć potrebno je voditi brigu o ograničenjima kao što su: propusnost, raspoloživa memorija, računalna moć i sl. Video signal nužno je obraditi (komprimirati) kako bi bio pogodan za prijenos u stvarnom vremenu. Kompresija video signala u svrhu prijenosa unosi određenu razinu pogreške što dovodi do narušavanja kvalitete video signala [3]. Uz kompresiju, uzrok pogreške može biti i prijenosni kanal u kojem je moguć gubitak paketa tijekom prijenosa. Gubitak paketa može biti uzrokovan ograničenom širinom pojasa, lošom konfiguracijom hardvera, greškama u softveru ili zagušenjem (preopterećenošću) mreže [4]. Različite degradacije video signala, nastale uslijed prijenosa i kompresije, nazivaju se artefaktima.

Komprimirani video signal uobičajeno se šalje prijenosnim kanalom podijeljen u pakete. U slučaju prijenosa uživo, poput video konferencije, neometano reproduciranje video signala primljenog u paketima zahtijeva pravovremeni prijem paketa. Tijekom prijenosa postoji mogućnost da paketi zakasne ili uopće ne stignu na odredište. Današnji video dekoderi u tom slučaju nastavljaju prikazivati zadnji ispravno primljeni okvir video signala sve dok ne prime idući ispravni okvir kao što je prikazano slikom 2.1. U tom slučaju smrznuti okvir (engl. *frozen frame*) predstavlja okvir videa koji je ponovljen odnosno identičan prethodnome okviru, a pojava niza uzastopno smrznutih

okvira naziva se smrzavanje (engl. *freeze event*) [5]. Promatrano sa stajališta kvalitete video signala, takva pojava naziva se artefakt smrzavanja. Svaki okvir koji na odredište nije pristigao na vrijeme smatra se izgubljenim, a opisan tip smrzavanja naziva se smrzavanje zbog gubitka paketa [6].

Referentni video signal:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	X	Y	Z
Degradirani video signal:	A	B	C	D	E	F	G	H	I	I	I	I	I	I	O	P	Q	R	S	T	U	V	X	Y	Z

Slika 2.1. Prikaz smrzavanja nastalog zbog gubitka paketa [5]

Ukoliko se ne radi o prijenosu video signala uživo, već o prijenosu prethodno snimljenog video sadržaja, tada se u slučaju kašnjenja okvira posljednji ispravno primljeni okvir ponavlja kao i kod smrzavanja zbog gubitka paketa. Međutim, prijenos prethodno snimljenog video sadržaja podrazumijeva dvosmjernu komunikaciju te je moguće zatražiti ponovno slanje od pošiljatelja. Tada će nakon smrzavanja biti reproduciran okvir koji u referentnom sadržaju slijedi nakon posljednjeg ispravno prenesenog okvira. Takav tip smrzavanja, prikazan slikom 2.2., naziva se smrzavanje zbog kašnjenja paketa [6].

Referentni video signal:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	X	Y	Z
Degradirani video signal:	A	B	C	D	E	F	G	H	H	H	H	H	H	I	J	K	L	M	N	O	P	Q	R	S	T

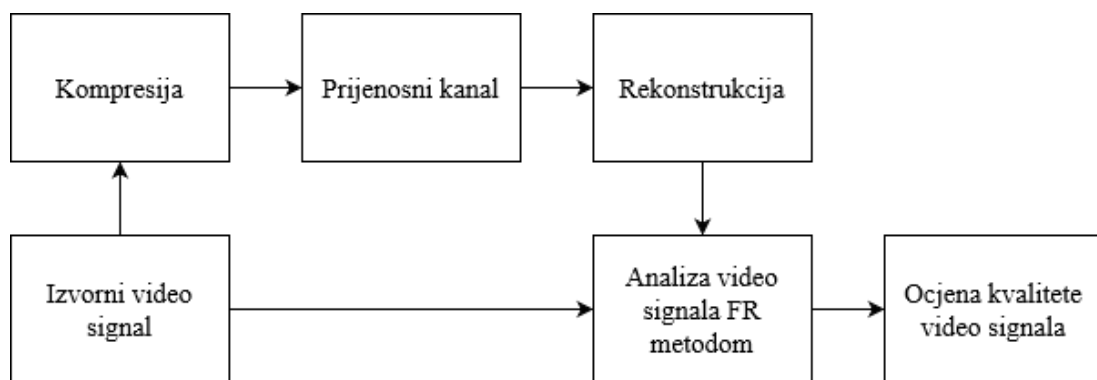
Slika 2.2. Prikaz smrzavanja nastalog zbog kašnjenja paketa [2]

Pojava artefakta smrzavanja smatra se vremenskom degradacijom video signala koja narušava kvalitetu video signala. Narušena kvaliteta video signala utječe na smanjenje iskustvene kvalitete usluge krajnjeg korisnika. Mrežni operater ima cilj svojim korisnicima ponuditi iskustvenu kvalitetu usluge na najvišoj razini što zahtijeva neprestano praćenje mogućih neželjenih pojava na korisničkoj strani. U tu svrhu razvijaju se subjektivne i objektivne metode za vrednovanje kvalitete video signala. Subjektivno vrednovanje kvalitete provodi se istraživanjem u kojemu ljudi ocjenjuju kvalitetu video signala uz različite degradacije u video sekvencama. Smatra se da ljudi mogu najbolje percipirati što narušava kvalitetu video signala i u kojoj mjeri. Objektivne metode podrazumijevaju razvoj algoritama koji detektiraju prostorne i vremenske degradacije u video signalu bez potrebe ljudskog vrednovanja. Rezultati subjektivnih metoda vrednovanja kvalitete video signala ključni su za razvoj objektivnih metoda čiji rezultati trebaju biti što sličniji ljudskoj percepciji [7].

Objektivne metode za ocjenu kvalitete video signala podijeljene su u tri kategorije [8]:

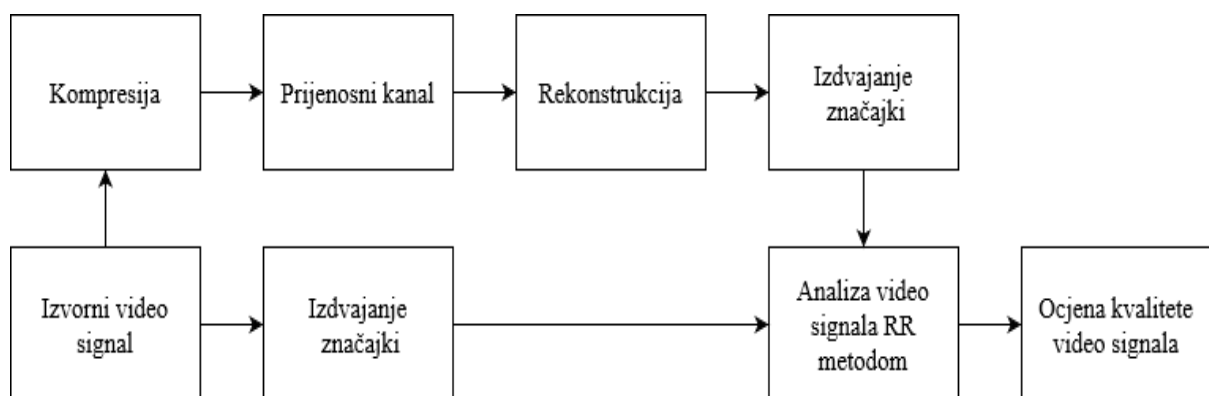
- Metode s pristupom referentnom video signalu (engl. *full reference*, *FR*)
- Metode s nepotpunim pristupom referentnom video signalu (engl. *reduced reference*, *RR*)
- Metode bez pristupa referentnom video signalu (engl. *no-reference*, *NR*)

FR metode za ocjenu kvalitete video signala zahtijevaju pristup izvornom video signalu. Izvorni video signal uspoređuje se s primljenim (rekonstruiranim) video signalom na razini elementa slike, odnosno svaki element slike primljenog video signala uspoređuje se s pripadajućim elementom slike iz referentnog video signala. Ovakav pristup nije moguć pri prijenosu video signala uživo jer podrazumijeva pristup nekomprimiranom izvornom video signalu. Dijagram sustava za prijenos video signala s ocjenom kvalitete FR metodom prikazan je slikom 2.3.



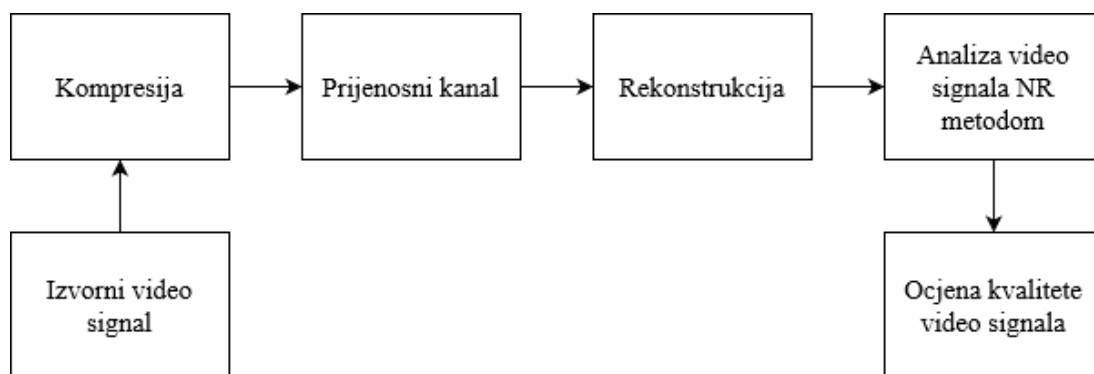
Slika 2.3. Dijagram sustava za prijenos video signala s ocjenom kvalitete FR metodom [4]

RR metode za ocjenu kvalitete video signala imaju nepotpun pristup izvornom video signalu, odnosno imaju pristup samo nekim informacijama o izvornom video signalu. Usporedba izvornog i primljenog video signala temelji se na usporedbi izdvojenih značajki (dio okvira, statističke informacije i sl.). Ovakav pristup brži je od FR metoda jer ne zahtijeva pristup nekomprimiranom izvornom video signalu već samo pristup izdvojenim značajkama. Slikom 2.4. prikazan je dijagram sustava za prijenos video signala s ocjenom kvalitete RR metodom.



Slika 2.4. Dijagram sustava za prijenos video signala s ocjenom kvalitete RR metodom [4]

NR metode podrazumijevaju ocjenjivanje kvalitete video signala bez pristupa izvornom video signalu. Budući da se ocjena kvalitete donosi isključivo na temelju primljenog video signala, ovakav pristup pogodan je za sadržaj koji se prenosi uživo. U slučaju ovakve metode ocjena kvalitete video signala zasniva se na matematičkom modelu. Slikom 2.5. prikazan je dijagram sustava za prijenos video signala s ocjenom kvalitete NR metodom [4].



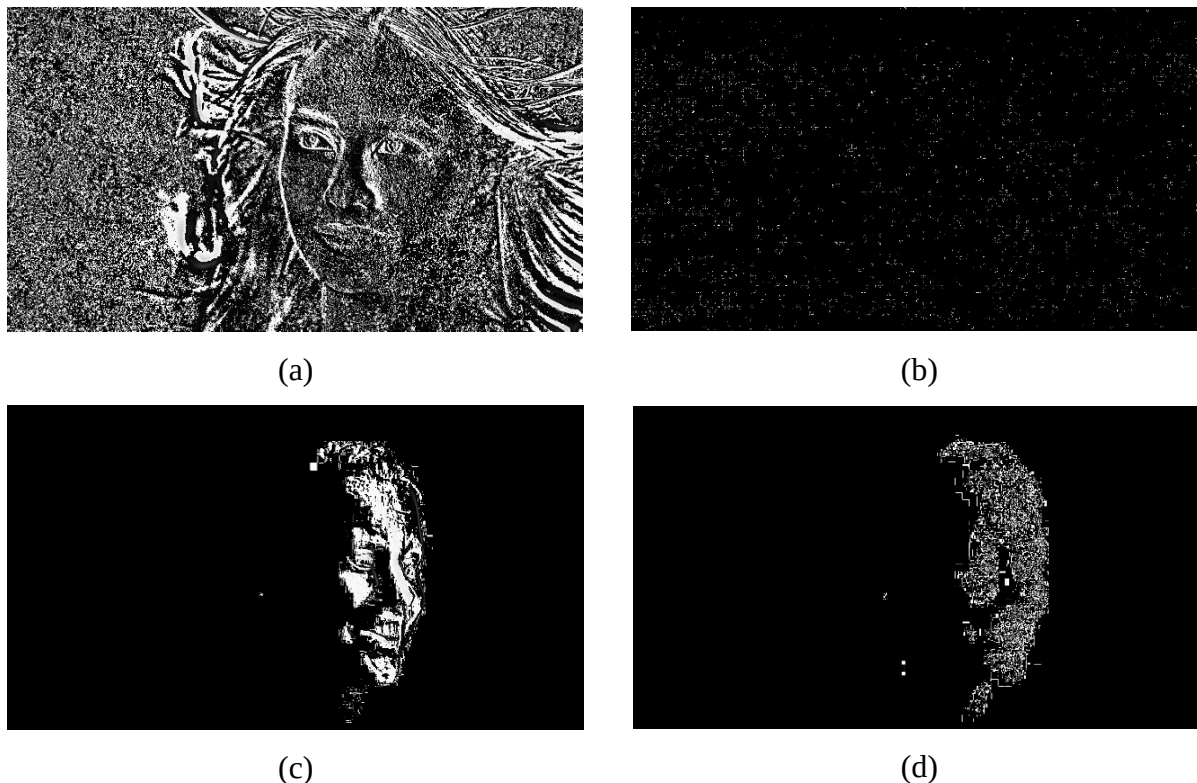
Slika 2.5. Dijagram sustava za prijenos video signala s ocjenom kvalitete NR metodom [4]

2.2. Pristup detekciji artefakta smrzavanja

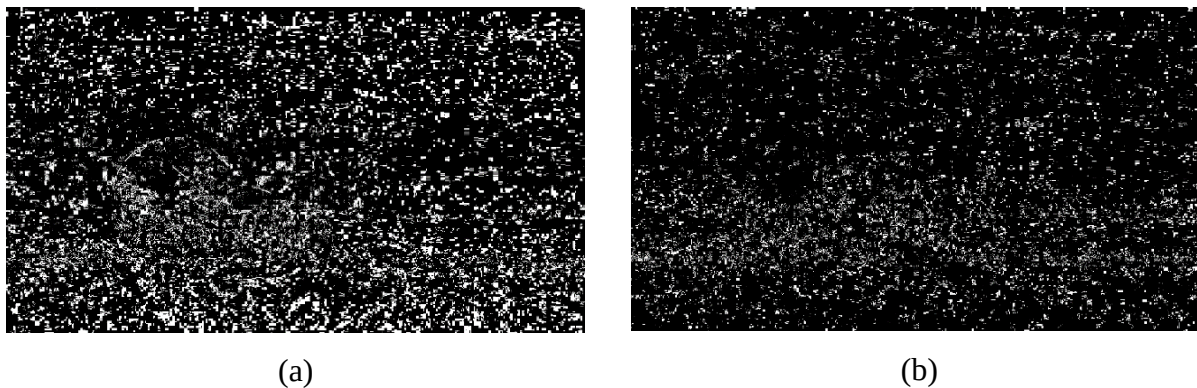
2.2.1. Klasični pristup detekcije smrzavanja

Gledano sa stajališta metode bez pristupa referentnom video signalu, jedina poznata stavka jest da se u slučaju pojave smrzavanja zadnji ispravno primljeni okvir ponavlja sve dok ne pristigne idući ispravni okvir. Osnovna pretpostavka pri detekciji smrzavanja jest da između prethodnog okvira i ponovljenog okvira ne postoji razlika. Ako su okviri isti, tada bi izračun srednje kvadratne pogreške između dva uzastopna okvira trebao biti jednak 0. U stvarnim uvjetima sustav za reprodukciju video signala unosi šum ili blage promjene u svjetlini u ponovljeni okvir. Slikom 2.6. (b, d) prikazana je apsolutna razlika luminantne komponente dvaju uzastopnih okvira gdje je jasno vidljiv problem pojave šuma u ponovljenim okvirima pri smrzavanju koji unosi sustav za reprodukciju (dohvaćanje) video signala. U tom slučaju ponovljeni okviri nisu potpuno isti te se zahtijeva uvođenje kriterija za donošenje odluke radi li se o ponovljenom okviru koji nije potpuno identičan prethodnome ili o stvarnoj promjeni okvira. Kako bi se smrzavanje detektiralo unatoč maloj razlici koju unosi šum, srednju kvadratnu razliku između dva okvira moguće je usporediti s unaprijed postavljenim pragom različitim od 0. Takav pristup rješava problem prisutne razlike u ponovljenom okviru, ali može uzrokovati pogrešno detektirano smrzavanje ako se radi o video signalu s iznimno sporim promjenama jer su u tom slučaju razlike u promjeni između uzastopnih okvira izrazito male [5]. U slučaju izrazito sporih promjena apsolutna razlika luminantne

komponente dvaju uzastopnih okvira video signala, prikazana slikom 2.7., prilično je slična neovisno radi li se o pojavi smrzavanja ili normalnom tijeku video sekvence.



Slika 2.6. Apsolutna razlika luminantne komponente dvaju uzastopnih okvira: (a) bez smrzavanja, (b) pri smrzavanju, (c) bez smrzavanja, (d) pri smrzavanju



Slika 2.7. Apsolutna razlika luminantne komponente dvaju uzastopnih okvira: (a) bez smrzavanja (b) pri smrzavanju

2.2.2. Pristup na temelju strojnog učenja

Postoje razne definicije pojma strojno učenje. Sveukupna bit raznih definicija svodi se na činjenicu da su današnja računala u mogućnosti učiti na velikom broju danih primjera te na temelju stečenog iskustva donijeti odluku poput ljudske inteligencije. Nadzirano strojno učenje podrazumijeva učenje na označenim podacima (uz svaki ulazni podatak poznat je izlazni podatak) s ciljem da se

za novi neoznačeni ulazni podatak dobije točan rezultat [9]. Ovakav pristup moguće je primijeniti u svrhu detekcije artefakta smrzavanja. Ako se u svrhu nadziranog strojnog učenja koristi apsolutna razlika luminantne komponente svakog elementa slike dvaju uzastopnih okvira video signala, ulazni podaci za učenje izgledaju kao na slici 2.6.

2.3. Baze video sekvenci koje sadrže artefakt smrzavanja

Proces nadziranog strojnog učenja podrazumijeva korištenje baza podataka čiji su podaci označeni kao negativni i pozitivni primjeri. U slučaju video sekvenci koje sadrže artefakt smrzavanja pozitivnim primjerom se smatra svaki okvir video sekvence koji sadrži artefakt smrzavanja, a negativnim primjerom svaki okvir koji ne sadrži artefakt smrzavanja.

U svrhu učenja, testiranja i usporedbe rezultata zasnovanih na strojnom učenju te usporedbe s rezultatima ostvarenim postojećim algoritmima korištene su četiri različite baze video sekvenci: LIVE Mobile, VQEG HDTV, UHD VFD i SMVS FullHD.

LIVE Mobile [10]–[12] baza video sekvenci utemeljena je na 10 različitih video sekvenci rezolucije 1280 x 720. Od svake temeljne video sekvence napravljene su po 4 video sekvence koje sadrže različite artefakte smrzavanja čime je ukupno nastalo 40 video sekvenci sa prisutnim artefaktom smrzavanja. Od svake temeljne video sekvence generirane su 3 video sekvence u kojima je artefakt smrzavanja posljedica kašnjenja paketa i generirana je jedna video sekvenca u kojoj je artefakt smrzavanja posljedica gubitka paketa. Video sekvence mogu sadržavati smrzavanja u trajanju od 1, 2 ili 4 sekunde koje se pojavljuje po 2, 4 ili 8 puta. Sve video sekvence imaju brzinu izmjene okvira od 30 okvira u sekundi i trajanje od 15 sekundi.

VQEG HDTV [13] baza video sekvenci sastoji se od 28 video sekvenci koje sadrže artefakt smrzavanja, a nastale su od 7 izvornih video sekvenci. Sve sekvence imaju rezoluciju 1920 x 1280 i trajanje od 10 sekundi uz brzinu izmjene okvira od 25 okvira u sekundi. Pojavljivanje artefakta smrzavanja u svakoj sekvenci je nasumično, a oponaša gubitak paketa prilikom prijenosa putem mreže.

FERIT-RTRK UHD VFD [14] baza video sekvenci sastoji se od 13 izvornih sekvenci rezolucije 3840 x 2160. 6 izvornih video sekvenci ima brzinu izmjene okvira 23,98 okvira u sekundi, a preostalih 7 izvornih video sekvenci 20 okvira u sekundi. Od svake izvorne video sekvence napravljene su po dvije video sekvence koje sadrže artefakt smrzavanja, ukupno 26 video sekvenci s prisutnim artefaktom smrzavanja. 13 video sekvenci sadrži artefakt smrzavanja uzrokovan

gubitkom paketa, a u preostalih 13 artefakt smrzavanja uzrokovan je kašnjenjem paketa pri mrežnom prijenosu podataka.

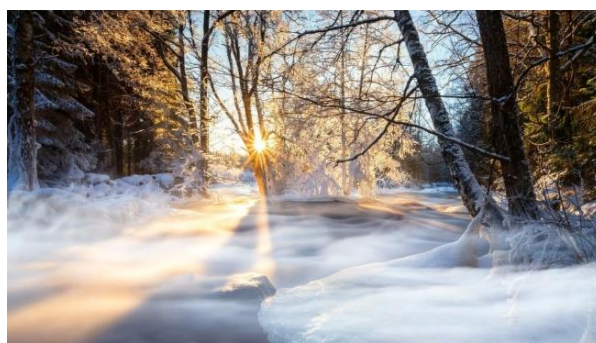
SMVS FullHD (engl. *Slow Motion Video Sequences Database*) baza video sekvenci sadrži 10 izvornih video sekvenci. Brzina izmjene okvira i trajanje različito je za svaku video sekvencu. Zajednička obilježja video sekvenci ove baze su FullHD rezolucija i statični kadrovi u kojima su promjene između okvira izrazito male. Od svake izvorne video sekvence generirane su po dvije degradirane; s jednim smrzavanjem trajanja 1s i s dva smrzavanja u trajanju od 1 i 2s. Prema tome, ova baza sadrži 20 video sekvenci s artefaktom smrzavanja. Pristup video sekvencama baze moguće je ostvariti na upit, a poveznica na bazu nalazi se u prilogu P.2.1.

2.4. Dodatne video sekvence

Osim video sekvenci iz dostupnih baza video sekvenci s prisutnim artefaktom smrzavanja u svrhu izrade rješenja za detekciju smrzavanja zasnovanog na strojnom učenju koriste se dvije dodatne video sekvence preuzete s web stranice Pixabay [15]. Video sekvence preuzete su u rezoluciji 1280 x 720. Slikom 2.8. prikazan je po jedan okvir iz svake video sekvence. Preuzete video sekvence ne sadrže artefakt smrzavanja, ali sadrže statične kadrove s malim vremenskim aktivnostima kao u slučaju SMVS FullHD baze. Razlog korištenja dvije dodatne video sekvence pri izradi rješenja za detekciju smrzavanja zasnovanog na strojnom učenju je ostvarivanja boljih rezultata u slučaju testnih podataka koji sadrže spore vremenske promjene između uzastopnih okvira.



(a)



(b)

Slika 2.8. Jedan okvir iz dodatnih video sekvenci: (a) *Sunrise sunbeams* [16] (b) *Sunrise fog* [17]

2.5. Postojeća rješenja za detekciju artefakta smrzavanja slike u video signalu

Jedan od najjednostavnijih pristupa ka detekciji artefakta smrzavanja bez pristupa referentnom video signalu opisan je u radu [5], primljeni okvir video signala prebacuje se u YUV format te se

za svaku komponentu Y, U i V računa srednja kvadratna pogreška između trenutnog i prethodnog okvira video signala. Ako je srednja kvadratna pogreška između dva uzastopna okvira manja od unaprijed postavljenog praga, smatra se da je došlo do pojave smrzavanja. Autor u radu [18], osim računanja srednje kvadratne pogreške, uvodi računanje energije pokreta između okvira video signala kako bi mogao postaviti promjenjivi prag za detekciju artefakta smrzavanja. Autori u radu [19] navode da je pri detekciji smrzavanja dovoljno koristiti Y komponentu YUV formata jer luminantna komponenta sadržava podatke o teksturi koji se nalaze u okviru. U tom radu prvi korak ka detekciji smrzavanja je računanje apsolutne razlike luminantne komponente trenutnog i prethodnog okvira za svaki element okvira. Potom se iz izračunate razlike stvaraju dva vektora; jedan koji sadrži zbroj svakog retka te jedan koji sadrži zbroj svakog stupca matrice (okvira). Za svaki vektor računaju se srednja vrijednost i standardna devijacija. Izračunate vrijednosti uspoređuju se s unaprijed postavljenim pragom kako bi se ustanovilo je li došlo do smrzavanja. Algoritam za detekciju smrzavanja [2] pod nazivom HBFDA (engl. *Histogram-Based Freezing artifacts Detection Algorithm*) temelji se na nešto kompleksnijoj usporedbi dvaju uzastopnih okvira i koristi promjenjivi prag za donošenje odluke o postojanju smrzavanja. Na učitani okvir primjenjuje se Gaussov filtar te se računa apsolutna razlika između izvornog i filtriranog okvira. Matrica apsolutne razlike iz prethodnog koraka dijeli se u regije od kojih svaka ima identičan broj elemenata slike, pri čemu dimenzije regija zadržavaju razmjer proporcija izvornog okvira te se za svaku regiju računa histogram. Razina aktivnosti između dva uzastopna okvira dobije se usporedbom histograma svake regije trenutnog okvira s histogramom pripadajuće regije iz prethodnog okvira korištenjem Hi-kvadrat testa. Na temelju razine aktivnosti računa se prilagodljivi prag i donosi odluka je li došlo do smrznutog okvira. Ako je detektirano više uzastopnih smrznutih okvira, razina aktivnosti se dodatno računa za trenutni okvir i posljednji ispravno preneseni okvir kako bi se izbjegla pogrešna detekcija ako se radi o sadržaju s izrazito sporim promjenama. U radu [1] predstavljen je algoritam za detekciju smrzavanja namijenjen radu u stvarnom vremenu pod nazivom RTFDA (engl. *Real-Time NR Freezing Detection Algorithm*). Rad algoritma temelji se na usporedbi apsolutne razlike svakog poduzorka trenutnog okvira s pripadajućim poduzorkom iz jednog od prethodnih okvira. Razmak između dva okvira koji se uspoređuju je unaprijed definiran, a temelji se na istraživanju utjecaja smrzavanja na iskustvenu kvalitetu usluge (eng. Quality of Experience, QoE) [20]. Rezultati istraživanja ukazuju da smrzavanja čije je trajanje duže od 200 ms imaju značajan utjecaj na iskustvenu kvalitetu usluge. Iz tog razloga RTFDA koristi usporedbu okvira čiji je razmak 200ms. U slučaju brzine izmjene okvira od 25 okvira u sekundi uspoređuju se prvi i šesti okvir u nizu video signala. Poduzorcji nastaju iz luminantne komponente okvira korištenjem filtra pomične srednje vrijednosti (engl.

Moving average) veličine $D \times D$ uz korak pomaka veličine D . Svaka apsolutna razlika poduzoraka dvaju okvira uspoređuje se s unaprijed postavljenim pragom. Ukoliko je broj apsolutnih razlika poduzoraka koje su veće od praga manji od unaprijed definirane količine, smatra se da je prisutno smrzavanje.

3. IZRADA RJEŠENJA ZA DETEKCIJU ARTEFAKTA SMRZAVANJA SLIKE U VIDEO SEIGNALU NA TEMELJU STROJNOG UČENJA

Unutar ovoga poglavlja nalazi se opis vlastitih rješenja za detekciju artefakta smrzavanja slike u video signalu koja su zasnovana na strojnom učenju. Izrađene su tri različite arhitekture neuronskih mreža. Prva arhitektura zasniva se na konvolucijskoj neuronskoj mreži (engl. *convolutional neural network*, *CNN*), druga na LSTM-u (engl. *Long short-term memory*, *LSTM*), a treća arhitektura je spoj prve dvije po nazivom CNN-LSTM. Implementacija svih rješenja izrađena je u *Python* programskom jeziku uz pomoć javno dostupnih biblioteka namijenjenih strojnom učenju. U potpoglavlju 3.2. nalazi se opis nastanka podataka koji se koriste u svrhu treniranja predloženih modela, potom su opisane tri različite arhitekture neuronskih mreža te proces treniranja. Cilj je svaku arhitekturu trenirati na jednoj bazi, a testirati na drugoj te usporediti rezultate testiranja i pronaći koja arhitektura daje najbolje rezultate.

3.1. Radno okruženje

Za razvoj predloženog rješenja korišten je *Python* programski jezik koji omogućuje jednostavno korištenje postojećih biblioteka namijenjenih za strojno učenje. Radi jednostavnosti treniranja neuronski mreža pomoću grafičke kartice korišten je operacijski sustav *Ubuntu 20.04.4 LTS*. Uz biblioteku *TensorFlow* dolazi *Keras* aplikacijsko programsko sučelje visoke razine koje omogućuje jednostavno stvaranje, treniranje i testiranje različitih arhitektura neuronskih mreža te je u tu svrhu i korišteno [21]. Korištena je i biblioteka *OpenCV* (engl. *Open Source Computer Vision Library*) koja je orijentirana ka obradi slike, strojnom učenju i računalnom vidu [22].

3.2. Proces izrade baza, arhitektura i treninga

3.2.1. Stvaranje podataka za treniranje

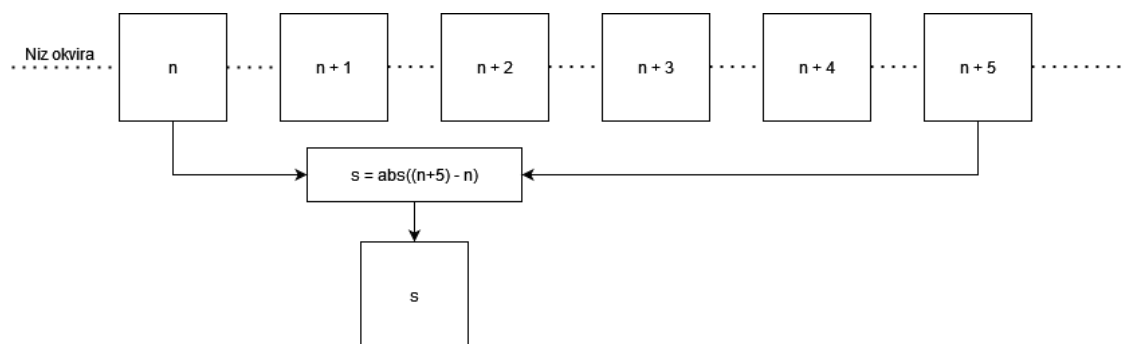
Prvi korak u nadziranom strojnom učenju je priprema podataka koji će se u tu svrhu koristiti. Uz dostupne baze video sekvenci koje sadrže artefakt smrzavanja postoje meta podaci u kojima je za svaku video sekvencu označen redni broj okvira od trenutka kada je smrzavanje započelo do trenutka kada prestaje. Na temelju meta podataka moguće je izdvojiti okvire koji sadrže artefakt smrzavanja i one koji ga ne sadrže te ih označiti pripadajućim oznakama.

Za označavanje okvira izdvojenih iz video sekvenci koje sadrže smrzavanje korištene su cjelobrojne oznake:

- 0 – okviri koji ne sadrže artefakt smrzavanja

- 1 – okviri koji sadrže artefakt smrzavanja

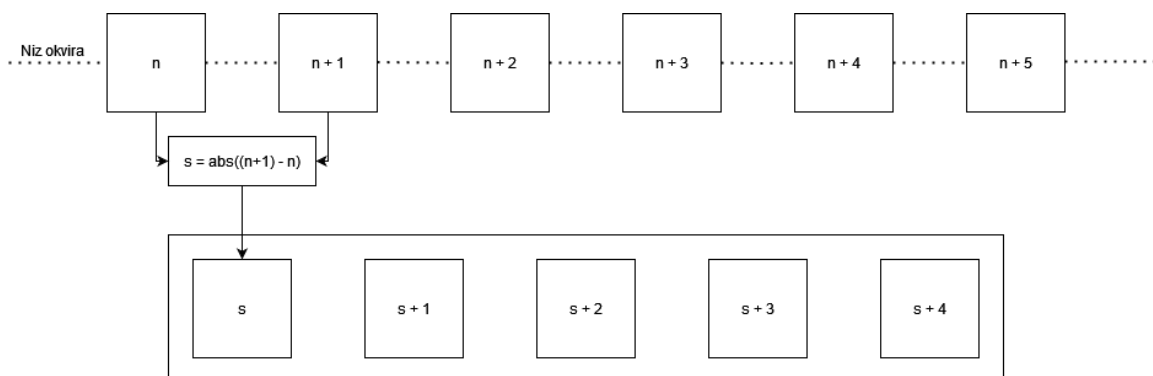
Prvi tip ulaznog podatka koji se koristi za treniranje neuronske mreže je apsolutna razlika luminantne komponente dva okvira odnosno dvodimenzionalna matrica dimenzija 1280 x 720. Odabrana rezolucija je ujedno i najmanja rezolucija među bazama video sekvenci koje sadrže artefakt smrzavanja. Za izračun apsolutne razlike ne koriste se dva uzastopna okvira, već šesti i prvi okvir u nizu po uzoru na pristup primijenjen u RTFDA kao što je prikazano slikom 3.1.



Slika 3.1. Dvodimenzionalni ulazni podatak

Rezolucije video sekvenci u dostupnim bazama video sekvenci s prisutnim artefaktom smrzavanja različite su za svaku bazu; UHD VFD sadrži video sekvence rezolucije 3840 x 2160, VQEG HDTV i SMVS FullHD 1920 x 1080, a LIVE Mobile 1280 x 720. Svaki izdvojeni okvir potrebno je skalirati na rezoluciju 1280 x 720. Kako bi se postigla željena rezolucija u slučaju rezolucije 3840 x 2160, svaki treći element slike ulaznog okvira vertikalno i horizontalno se pohranjuje, a ostala dva odbacuju, dok se za rezoluciju 1920 x 1280 svaki treći element slike vertikalno i horizontalno odbacuje, a prva dva pohranjuju.

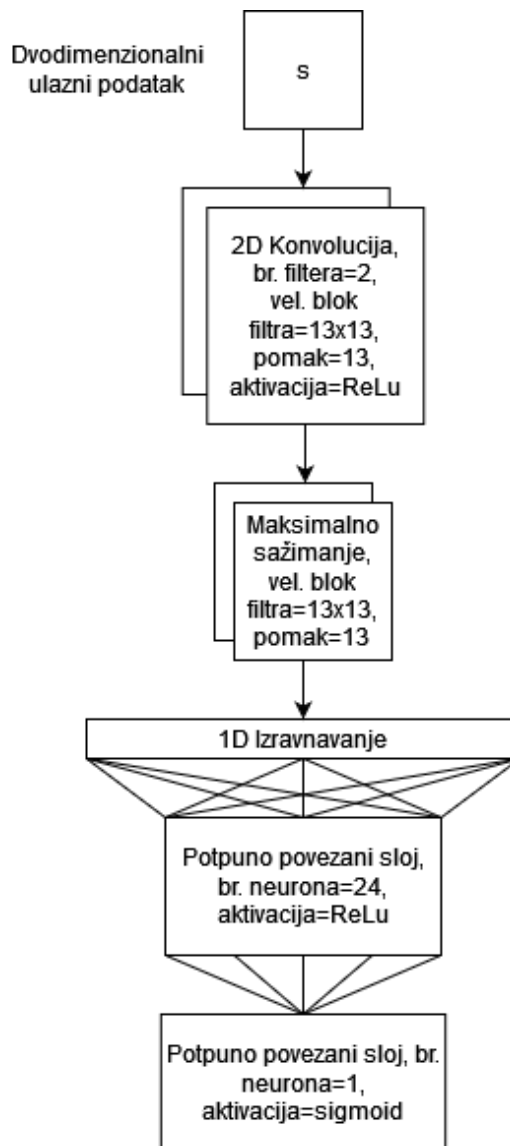
Drugi tip ulaznog podatka korišten u svrhu nadziranog strojnog učenja je trodimenzionalna matrica koja se sastoji od 5 dvodimenzionalnih matrica. Svaka dvodimenzionalna matrica je apsolutna razlika luminantne komponente dvaju uzastopnih okvira dimenzija 1280 x 720.



Slika 3.2. Trodimenzionalni ulazni podatak

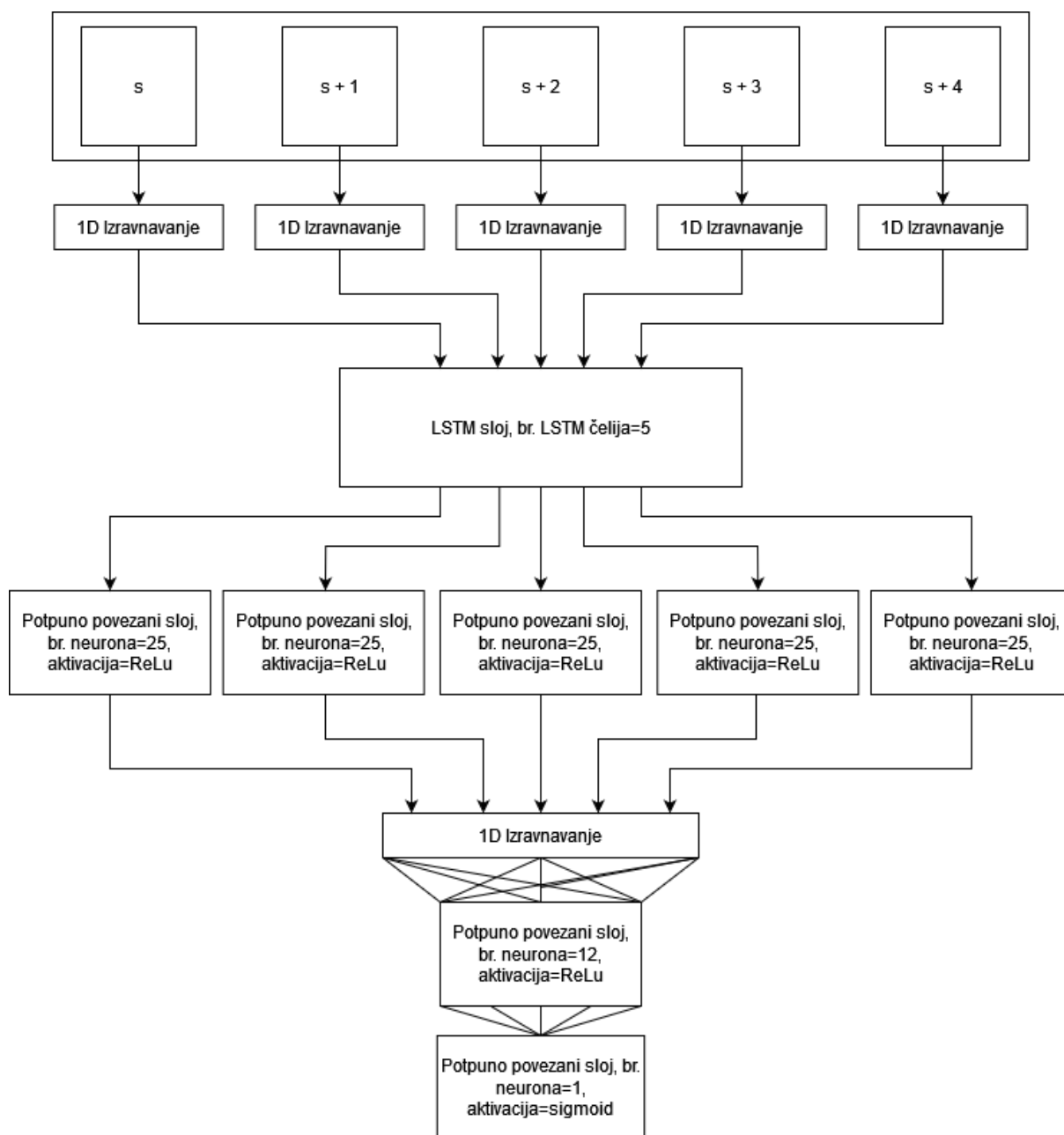
3.2.2. Arhitekture neuronske mreže

U svrhu detekcije artefakta smrzavanja slike u video signalu zasnovane na strojnom učenju odabrane su tri različite arhitekture neuronskih mreža. Prva arhitektura (slika 3.3.) zasnovana je na konvolucijskoj neuronskoj mreži. Predložena CNN arhitektura koristi dvodimenzionalni tip ulaznog podatka navedenog ranije u tekstu. Sadrži jedan konvolucijski sloj u kojemu se koristi ReLu (engl. *Rectified Linear Unit*) aktivacijska funkcija, veličina kvadratnog filtera iznosi 13 x 13 uz pomak 13. Idući slijedi sloj maksimalnog sažimanja koji koristi filter veličine 13 x 13 uz pomak 13. Nakon sažimanja slijedi potpuno povezani sloj sa 24 neurona pri čemu se koristi ReLu aktivacijska funkcija. Na izlazu se nalazi 1 neuron pri čemu se koristi sigmoidna aktivacijska funkcija.



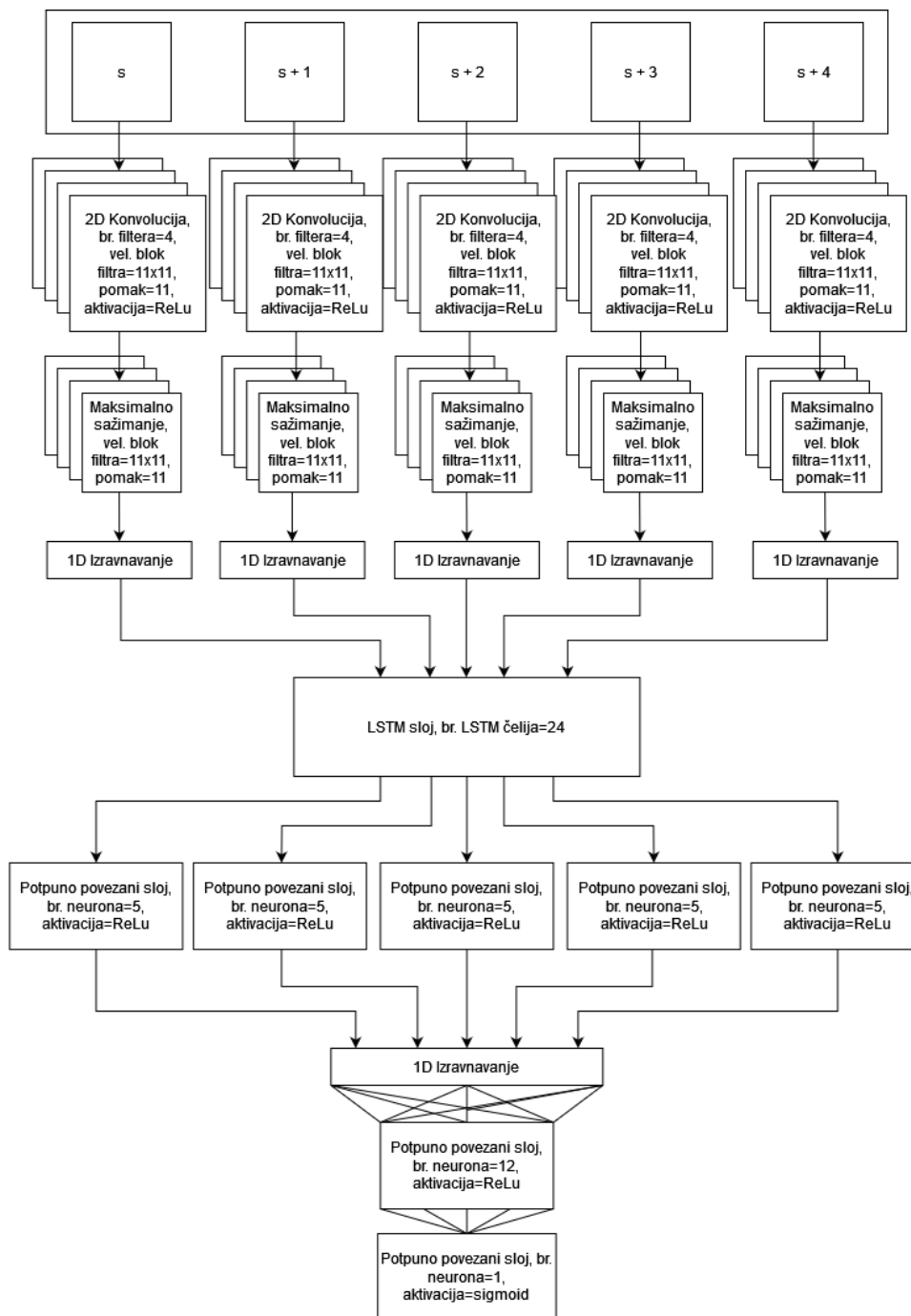
Slika 3.3. Prikaz CNN arhitekture

Druga arhitektura koja je zasnovana na LSTM-u, prikazana slikom 3.4., koristi trodimenzionalni ulazni podatak ranije nadveden u tekstu. Ulazni podatak predstavlja niz povezanih podataka (sekvencu) stoga se koristi *TimeDistributed* sloj prije LSTM sloja. Nakon LSTM sloja dolazi *TimeDistributed* potpuno povezani sloj sa 25 neurona. Potom slijedi potpuno povezani sloj od 12 neurona te isključivanje čvorova (engl. *Dropout*). Za izlaz se koristi 1 neuron potpuno povezan sa prethodnim slojem uz korištenje sigmoidne aktivacijske funkcije.



Slika 3.4. Prikaz LSTM arhitekture

Treća arhitektura neuronske mreže, slika 3.5., temelji se na kombinaciji CNN i LSTM arhitektura, stoga će se u dalje u radu označavati kao CNN-LSTM. Kao i kod LSTM arhitekture ulazni podatak je trodimenzionalni. Kao i kod ranije navedene CNN arhitekture koriste se konvolucija i maksimalno sažimanje, ali se u ovom slučaju koristi *TimeDistributed* sloj. Odabrana veličina blok filtra iznosi 11 x 11 uz pomak 11. Odabrani broj LSTM ćelija je 24. Dok je u idućem *TimeDistributed* potpuno povezanom sloju po 5 neurona. Predzadnji potpuno povezani sloj sadrži 12 neurona te isključivanje čvorova.



Slika 3.5. Prikaz CNN-LSTM arhitekture

3.2.3. Proces treniranja neuronskih mreža

Treniranje predloženih neuronskih mreža zahtijeva stvaranje baza podataka iz dostupnih baza video sekvenci s prisutnim artefaktom smrzavanja. Za svaku video sekvencu izdvajaju se podaci u obliku ovisnom od arhitekture mreže navedeni ranije u tekstu. Nakon izdvajanja ulaznih podataka za treniranje potrebno je izvršiti podjelu na skup za treniranje i validaciju u omjeru približnom 75:25. Podjela se vrši na razini video sekvenci stoga je omjer za pojedinu bazu približan. Tablica 3.1 prikazuje imena video sekvenci koje se koriste za validacijski skup dok se sve ostale video sekvence iz pripadajuće baze koriste u skupu za treniranje. Ako se u skupu za treniranje koriste 3 baze, onda se za validacijski skup koriste pripadajuće video sekvence triju odabranih baza prema tablici 3.1.

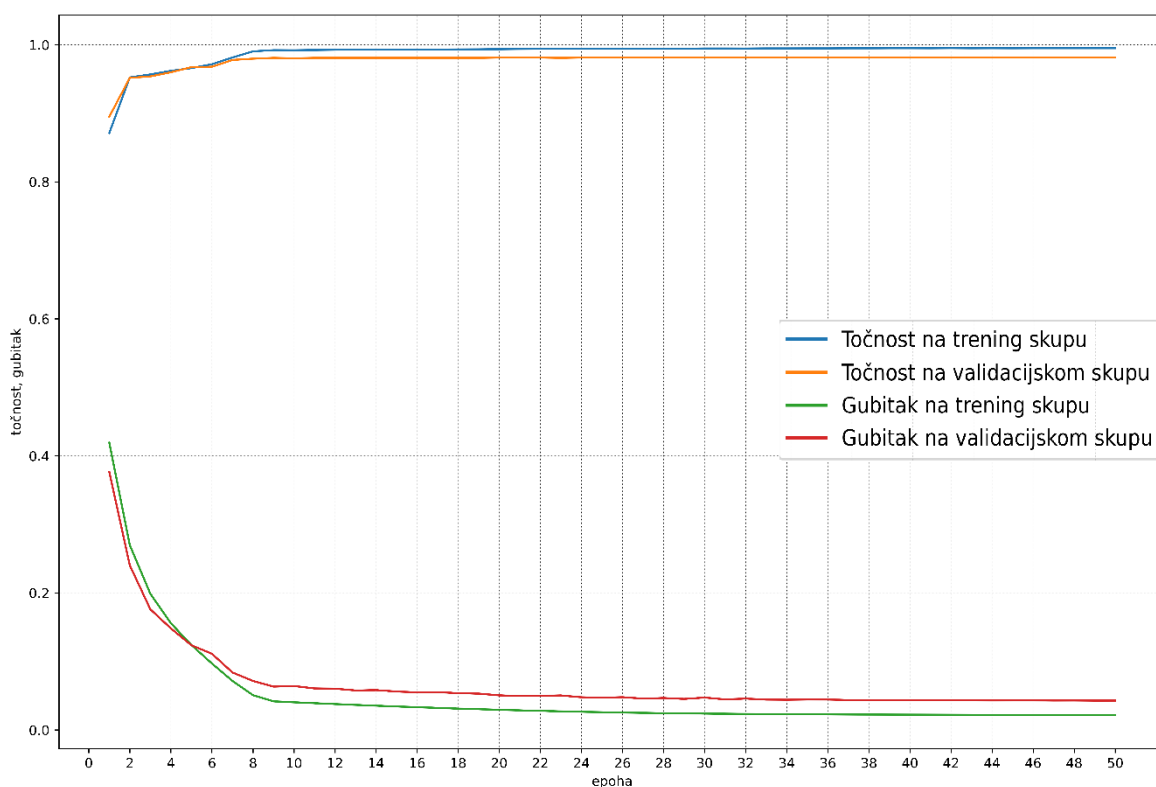
Tablica 3.1. *Video sekvence korištene u skupu za validaciju*

Ime baze:	Validacijske sekvence:
LIVE Mobile	bf_f1, bf_f2, bf_f4, bf_fr4, dv_f1, dv_f2, dv_f4, dv_fr4
VQEG HDTV	vqeghd5_src01_hrc10, vqeghd5_src01_hrc11, vqeghd5_src01_hrc12, vqeghd5_src01_hrc15, vqeghd5_src02_hrc10, vqeghd5_src02_hrc11, vqeghd5_src02_hrc12, vqeghd5_src02_hrc15
UHD VFD	BeautyPD, BeautyPL, CactusPD, CactusPL, JockeyPD, JockeyPL
SMVS FullHD	PaintDrop2x, Pears2x

Pri stvaranju ulaznih podataka za treniranje u slučaju LIVE Mobile, VQEG HDTV ili UHD VFD baze koriste se sve sekvence koje sadržavaju artefakt smrzavanja, dok se kod SMVS FullHD baze koriste samo sekvence koje sadrže 2 smrzavanja (u nazivu označene s 2x).

Ukoliko se treniranje neuronske mreže vrši korištenjem jedne od dostupnih baza, tada se testiranje vrši na preostalim bazama koje nisu korištene u svrhu treniranja. Osim treniranja na pojedinačnoj bazi, koriste se i različite kombinacije dostupnih baza. Tada se tri baze koriste za treniranje, a jedna preostala baza za testiranje. Svako treniranje neuronske mreže dodatno je ponovljeno uz dodavanje dvije dodatne video sekvence izrazito sporog sadržaja u skup za treniranje.

Pri treniranju neuronske mreže koristi se promjenjiva stopa učenja (engl. *learning rate*), a veličina serije (engl. *batch size*) iznosi 32. U slučaju CNN arhitekture broj epoha je 50, dok je kod LSTM i CNN-LSTM arhitektura 25. Kako bi testiranje bilo moguće, potrebno je pohraniti parametre modela pri treningu neuronske mreže. Nakon svake epohe pohranjuju se četiri datoteke koje sadrže parametre modela. Svaki od pohranjenih parametara modela ovisi o jednom uvjetu koji se uspoređuje sa vrijednostima dobivenim iz prethodnih epoha. Uvjeti hoće li parametri modela biti pohranjeni su: veća točnost modela na skupu za trening, veća točnost na skupu za validaciju, niža vrijednost funkcije gubitka na skupu za trening i niža vrijednost funkcije gubitka na skupu za validaciju. Nakon završenog treninga dostupan je graf, kao na sa slike 3.6., koji prikazuje točnost i gubitak za validacijski i trening skup podataka u odnosu na broj proteklih epoha. Na temelju grafa se odabire konačne parametre modela.



Slika 3.6. Graf rezultata treniranja CNN arhitekture, skup za trening: LIVE Mobile uz dvije dodatne sekvence

Pri treniranju neuronskih mreža u ovome radu najčešći izgled grafa rezultata treniranja je kao na slici 3.6. U tom slučaju točnost na validacijskom skupu stagnira nakon određenog broja epoha, ali gubitak na validacijskom skupu opada nakon svake epohe te se za konačne parametre modela odabiru parametri koji su dali najmanji gubitak na validacijskom skupu.

3.3. Upute za pokretanje

Prvi korak za uspješno pokretanje razvijenih rješenja je stvaranje programskog okruženja. U tu svrhu je korištena *Anaconda* platforma za upravljanje paketima i okruženjima [23]. Za stvaranje programskog okruženja koristi se datoteka *requirements.yml* koja je dostupna na priloženom DVD-u u direktoriju P.3.1. Okruženje se stvara naredbom:

```
conda env create -n <ime_okruženja> --file <putanja_do_requirements.yml>
```

Stvoreno programsko okruženje potrebno je aktivirati naredbom:

```
conda activate <ime_okruženja>
```

Najprije je potrebno pokrenuti *data_generator.py* skriptu za stvaranje podataka (P.3.2.) koji se koriste za treniranje neuronskih mreža naredbom:

```
python data_generator.py
```

Opcije pri pokretanju:

- *-d <putanja_do_direktorija_sa_video_sekvencama>*
- *-n <putanja_do_direktorija_gdje_će_podaci_biti_pohranjeni>*
- *-m <putanja_do_json_datoteke>* - datoteka koja sadrži meta podatke o smrzavanjima za sve video sekvence (P.3.3.)
- *-s <tip_podataka>* - 1 za stvaranje dvodimenzionalnih ulaznih podataka, 0 za stvaranje trodimenzionalnih ulaznih podataka

Nakon izvršavanja nastaju direktoriji koji sadržavaju stvorene podatke. Svaki direktorij nosi naziv video sekvence iz koje su podaci stvoreni. Stvorene direktorije potrebno je podijeliti u dva direktorija pod nazivima *train* i *valid* prema uputama iz potpoglavlja 3.2.3. Nakon podjele direktorija potrebno je stvoriti *hdf5* datoteku koja se koristi za treniranje neuronskih mreža pokretanjem skripte *hdf5_dataset_maker.py* (P.3.2.) naredbom:

```
python hdf5_dataset_maker.py
```

Opcije pri pokretanju:

- *-d <putanja_do_direktorija_sa_stvorenim_podacima>*
- *-n <ime_hdf5_datoteke_koja_će_biti_stvorena>*
- *-s <tip_podataka>* - 1 ako su stvoreni podaci dvodimenzionalni, 0 ako su trodimenzionalni

Za treniranje neuronskih mreža potrebno je pokrenuti skriptu *hdf5_train.py* (P.3.4.) naredbom:

python hdf5_train.py

Opcije pri pokretanju:

- *-d <putanja_do_hdf5_datoteke>*
- *-n <putanja_do_direktorija_gdje_će_podaci_biti_pohranjeni>*
- *-e <broj_epoha>* - preporučena vrijednost za CNN je 50, a za ostale 25
- *-b <veličina_serije>* - preporučena vrijednost je 32
- *-n <vrsta_arhitecture>* - 1 za CNN, 2 za LSTM, 3 za CNN-LSTM
- *-flr <početna_stopa_učenja>* - preporučena vrijednost 0.0001
- *-fep <epoha_promjene>* - broj epohe kada dolazi do promjene početne stope učenja, preporučena vrijednost za CNN je 8, a za ostale 4
- *-slr <druga_stopa_učenja>* - preporučena vrijednost 0.00001
- *-sep <epoha_promjene>* - broj epohe kada dolazi do promjene druge stope učenja, preporučena vrijednost za CNN je 24, a za ostale 7

Tijekom treninga, nakon svake epohe, pohranjuju se 4 *h5* datoteke koje sadržavaju parametre modela te se pohranjuju i podaci o treniranju u *train_log.csv* datoteku. Graf s podacima o treniranju moguće je iscrtati pokretanjem skripte *log.py* (P.3.4.) naredbom:

python log.py

Opcije pri pokretanju:

- *-d <putanja_do_csv_datoteke>*
- *-n <putanja_do_direktorija_gdje_će_graf_biti_pohranjen>* - graf se pohranjuje kao *png* datoteka

Testiranje istreniranog modela moguće je pokretanjem skripte *freezing_detection_all_videos.py* (P.3.5.) naredbom:

python freezing_detection_all_videos.py

Opcije pri pokretanju:

- *-a <vrsta_neuronske_mreže>* - *cnn*, *cnn_lstm*
- *-d <putanja_do_direktorija_sa_video_sekvencama>*

- *-m <putanja_do_json_datoteke>* - datoteka koja sadrži meta podatke o smrzavanjima za sve video sekvence (P.3.3.)
- *-p <putanja_do_h5_datoteke>* - putanja do *h5* datoteke istreniranog modela

Moguće je i testiranje istreniranih modela nastalih u ovom radu koji se nalaze na priloženom DVD-u u direktoriju P.3.6.

4. EVALUACIJA I VERIFIKACIJA IZRAĐENIH RJEŠENJA TEMELJENIH NA STROJNOM UČENJU

Unutar ovoga poglavlja predstavljeni su rezultati detekcije smrzavanja korištenjem tri različite neuronske mreže te usporedba s rezultatima dobivenim pomoću HBFDA i RTFDA algoritama. Rezultati algoritama ostvareni na LIVE Mobile, FERIT-RTRK UHD VFD i VQEG HDTV bazama prikazani su u radovima [2] i [1]. Pomoću dostupnih rezultata provjerena je ispravnost dostupne implementacije oba algoritma.

Radi lakšeg prikaza i razumijevanja rezultata ostvarenih testiranjem, u tablici 4.1. prikazan je način imenovanja istreniranih modela ovisno o arhitekturi neuronske mreže i podacima koji su korišteni u skupu za trening. Naziv arhitekture neuronske mreže nalazi se na prvom mjestu (C, L, CL), nakon naziva arhitekture slijedi podvlaka iza koje se navodi baza koja je korištena u skupu za trening (LIVE, UHD, VQ, SM). Ukoliko je korišteno više baza nazivi su odvojeni podvlakom (pr. VQ_UHD_SM). Ukoliko su u skup za trening uključene dodatne dvije sekvence nakon podvlake slijedi broj 2.

Tablica 4.1. Imenovanje modela

Naziv modela:	Vrsta arhitekture:	Skup za trening:	Dodatne video sekvence:
C_LIVE	CNN	LIVE Mobile	ne
L_LIVE_VQ_SM_2	LSTM	LIVE Mobile, VQEG HDTV, SMVS FullHD	da
CL_UHD	CNN-LSTM	FERIT-RTRK UHD VFD	ne

U svrhu verifikacije promatraju se omjeri parametara dobiveni testiranjem. Parametri dobiveni testiranjem na razini okvira video sekvence su:

- Istinito pozitivan primjer (engl. *True positive, TP*) - okvir koji je detektiran kao smrzavanje, a on to je,
- Lažno pozitivan primjer (engl. *False positive, FP*) - okvir koji je detektiran kao smrzavanje, a on to nije,
- Istinito negativan primjer (engl. *True negative, TN*) - okvir koji nije detektiran kao smrzavanje, a on to nije,

- Lažno negativan primjer (engl. *False negative, FN*) - okvir koji nije detektiran kao smržavanje, a on to je.

Gore navedeni parametri koriste se za računanje mjera pomoću kojih se vrednuju različita rješenja, a mjere vrednovanja su: točnost, preciznost, odziv i F1 mjera. Točnost (engl. *accuracy*) se računa kao omjer svih istinitih primjera i zbroja svih istinitih i lažnih primjera (4-1). Preciznost (engl. *precision*) predstavlja omjer između detektiranih istinito pozitivnih primjera i ukupnog broja svih pozitivnih primjera (4-2). Odziv (engl. *recall*) je definiran kao omjer svih istinito pozitivnih primjera i zbroja istinito pozitivnih i lažno negativnih primjera (4-3). F1 mjera definirana je kao harmonijska sredina mjera preciznosti i odziva (4-4).

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

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

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

$$F1 = 2 * \frac{preciznost*odziv}{preciznost+odziv} \quad (4-4)$$

Mjere vrednovanja računaju se na razini baze video sekvenci, odnosno koristi se ukupan broj TN, TP, FN i FP na razini baze.

U svrhu testiranja korišteno je računalo sa Linux operacijskim sustavom (Ubuntu 20.04.4 LTS), grafička kartica NVIDIA GeForce RTX 3080Ti, Intel Core i9-11900F uz 32 GB RAM.

4.1. Verifikacija na LIVE Mobile bazi

Tablica 4.2. Rezultati testiranja nad video sekvencama LIVE Mobile baze

METODA DETEKCIJE / IME MODELA	TOČNOST	PRECIZNOST	ODZIV	F1
HBFDA	99,13%	98,96%	98,60%	98,78%
RTFDA	99,81%	99,47%	100,00%	99,73%
C_UHD	93,12%	85,58%	97,08%	90,97%
C_UHD_2	98,38%	99,27%	96,19%	97,70%
C_VQ	84,62%	97,23%	58,60%	73,13%
C_VQ_2	90,52%	100,00%	73,46%	84,70%
C_SM	90,65%	81,76%	95,02%	87,89%
C_SM_2	92,96%	86,23%	95,56%	90,66%
C_UHD_VQ_SM	95,90%	91,62%	97,44%	94,44%
C_UHD_VQ_SM_2	97,07%	94,56%	97,39%	95,95%
L_UHD	89,21%	91,72%	76,70%	83,54%
L_UHD_2	64,33%	100,00%	0,11%	0,22%
L_VQ	77,96%	100,00%	38,29%	55,38%
L_VQ_2	88,98%	100,00%	69,16%	81,77%
L_SM	80,69%	100,00%	45,92%	62,94%
L_SM_2	82,50%	100,00%	51,01%	67,56%
L_UHD_VQ_SM	89,81%	98,95%	72,22%	83,50%
L_UHD_VQ_SM_2	93,35%	99,02%	82,19%	89,82%
CL_UHD	94,27%	87,41%	98,08%	92,44%
CL_UHD_2	95,53%	91,92%	95,91%	93,87%
CL_VQ	82,76%	100,00%	51,73%	68,19%
CL_VQ_2	78,92%	100,00%	40,98%	58,13%
CL_SM	92,81%	83,50%	99,56%	90,82%
CL_SM_2	92,22%	83,16%	98,08%	90,00%
CL_UHD_VQ_SM	94,08%	86,70%	98,53%	92,24%
CL_UHD_VQ_SM_2	97,35%	94,46%	98,36%	96,37%

Iz tablice rezultata 4.2. vidljiva su velika odstupanja između mjera preciznosti i odziva, stoga je najbolje promatrati njihovu harmonijsku sredinu, F1 mjeru. U slučaju CNN modela, pri treningu neuronske mreže na bilo kojoj dostupnoj bazi, u svim slučajevima rezultati postaju bolji ukoliko se u trening skup dodaju dvije dodatne video sekvence izrazito sporog sadržaja. Neuronske mreže trenirane s podacima iz VQEG HDTV baze podataka daju najlošije rezultate u slučaju testiranja na LIVE Mobile bazi. Uzrok takvih rezultata su „jednostavna“ smrzavanja u VQEG HDTV bazi, odnosno smrzavanja koja ne sadrže šum u ponovljenim okvirima. LSTM modeli pokazali su se kao najlošiji, osim u slučaju treninga na podacima iz tri preostale baze uz dvije dodatne video sekvence.

4.2. Verifikacija na FERIT-RTRK UHD VFD bazi

Tablica 4.3. Rezultati testiranja nad video sekvencama FERIT-RTRK UHD VFD baze

METODA DETEKCIJE / IME MODELA	TOČNOST	PRECIZNOST	ODZIV	F1
HBFDA	99,22%	99,92%	97,10%	98,49%
RTFDA	100,00%	100,00%	100,00%	100,00%
C_LIVE	90,76%	73,78%	100,00%	84,92%
C_LIVE_2	93,87%	80,92%	100,00%	89,46%
C_VQ	74,20%	100,00%	0,84%	1,66%
C_VQ_2	74,13%	100,00%	0,57%	1,14%
C_SM	88,50%	81,34%	72,39%	76,61%
C_SM_2	99,96%	100,00%	99,85%	99,92%
C_LIVE_VQ_SM	98,81%	95,63%	100,00%	97,77%
C_LIVE_VQ_SM_2	98,87%	95,84%	100,00%	97,88%
L_LIVE	78,92%	94,62%	20,11%	33,17%
L_LIVE_2	82,28%	96,35%	33,17%	49,35%
L_VQ	73,98%	0,00%	0,00%	0,00%
L_VQ_2	74,73%	100,00%	2,86%	5,55%
L_SM	74,70%	100,00%	2,74%	5,34%
L_SM_2	74,88%	100,00%	3,47%	6,70%
L_LIVE_VQ_SM	83,12%	97,62%	35,99%	52,59%
L_LIVE_VQ_SM_2	84,94%	98,00%	42,99%	59,77%
CL_LIVE	97,73%	92,04%	99,92%	95,82%
CL_LIVE_2	98,33%	93,95%	100,00%	96,88%
CL_VQ	73,98%	0,00%	0,00%	0,00%
CL_VQ_2	73,98%	0,00%	0,00%	0,00%
CL_SM	97,71%	91,91%	100,00%	95,79%
CL_SM_2	97,00%	89,85%	99,73%	94,53%
CL_LIVE_VQ_SM	98,47%	94,46%	100,00%	97,15%
CL_LIVE_VQ_SM_2	98,18%	93,45%	100,00%	96,62%

Iz rezultata ostvarenih testiranjem na UHD VFD bazi, prikazani tablicom 4.3., vidljivo je kako LSTM modeli daju izrazito lošije rezultate naspram ostalih arhitektura neuronskih mreža. Sve arhitekture trenirane s podacima iz VQEG HDTV baze daju izrazito loše rezultate kao i u slučaju testiranja na LIVE Mobile bazi podataka. Modeli L_VQ, CL_VQ i CL_VQ_2 imaju F1 mjeru 0% što ukazuje da smrzavanja UHD VFD baze sadrže znatno više šuma naspram smrzavanja iz VQEG HDTV baze. Najbolji rezultat ostvario je C_SM_2 model koji je prema svim mjerama premašio rezultate HBFDA algoritma.

4.3. Verifikacija na VQEG HDTV bazi

Tablica 4.4. Rezultati testiranja nad video sekvencama VQEG HDTV baze

METODA DETEKCIJE / IME MODELA	TOČNOST	PRECIZNOST	ODZIV	F1
HBFDA	99,76%	99,20%	99,83%	99,51%
RTFDA	99,96%	100,00%	99,83%	99,91%
C_LIVE	77,10%	51,96%	100,00%	68,39%
C_LIVE_2	82,41%	58,48%	100,00%	73,80%
C_UHD	98,67%	94,91%	100,00%	97,39%
C_UHD_2	99,90%	99,60%	100,00%	99,80%
C_SM	99,54%	98,19%	100,00%	99,09%
C_SM_2	99,87%	99,77%	99,71%	99,74%
C_LIVE_UHD_SM	87,23%	65,98%	100,00%	79,50%
C_LIVE_UHD_SM_2	87,83%	67,05%	100,00%	80,28%
L_LIVE	98,37%	93,97%	99,83%	96,81%
L_LIVE_2	100,00%	100,00%	100,00%	100,00%
L_UHD	73,07%	41,94%	22,66%	29,43%
L_UHD_2	75,23%	0,00%	0,00%	0,00%
L_SM	99,74%	100,00%	98,96%	99,48%
L_SM_2	99,80%	100,00%	99,19%	99,59%
L_LIVE_UHD_SM	99,93%	100,00%	99,71%	99,86%
L_LIVE_UHD_SM_2	99,80%	99,20%	100,00%	99,60%
CL_LIVE	94,26%	81,18%	100,00%	89,61%
CL_LIVE_2	95,60%	84,92%	100,00%	91,84%
CL_UHD	99,76%	99,03%	100,00%	99,51%
CL_UHD_2	99,83%	99,31%	100,00%	99,66%
CL_SM	96,06%	86,27%	100,00%	92,63%
CL_SM_2	97,56%	91,02%	100,00%	95,30%
CL_LIVE_UHD_SM	97,79%	91,79%	100,00%	95,72%
CL_LIVE_UHD_SM_2	96,60%	87,93%	100,00%	93,58%

U slučaju testiranja nad video sekvencama VQEG HDTV baze, prikazanih u tablici 4.4., najbolje rezultate ostvario je model L_LIVE_2 koji je sve detektirao potpuno točno. Gotovo svi modeli zasnovani na CNN i CNN-LSTM arhitekturi imaju izrazito visok odziv, što znači da točno detektiraju okvire koji sadrže smrzavanje. Visok odziv ukazuje da baze korištene u skupu za treniranje sadrže „kompleksnija“ smrzavanja, odnosno smrzavanja koja sadrže šum za razliku od smrzavanja iz VQEG HDTV baze koja ne sadrže nikakav šum. C_UHD_2 model ostvaruje F1 mjeru identičnu kao RTFDA algoritam te ostvaruje bolje mjere odziva i preciznosti nego HBFDA algoritam.

4.4. Verifikacija na SMVS FullHD bazi

Tablica 4.5. Rezultati testiranja nad video sekvencama SMVS FullHD baze

METODA DETEKCIJE / IME MODELA	TOČNOST	PRECIZNOST	ODZIV	F1
HBFDA	47,13%	22,50%	95,93%	36,45%
RTFDA	83,89%	49,42%	93,72%	64,71%
C_LIVE	22,44%	16,72%	98,50%	28,59%
C_LIVE_2	41,30%	20,82%	97,17%	34,29%
C_UHD	79,60%	43,23%	94,16%	59,26%
C_UHD_2	95,08%	80,23%	91,24%	85,38%
C_VQ	85,59%	100,00%	8,58%	15,81%
C_VQ_2	86,03%	100,00%	11,33%	20,35%
C_LIVE_UHD_VQ	50,49%	23,60%	95,75%	37,87%
C_LIVE_UHD_VQ_2	83,57%	48,88%	92,92%	64,06%
L_LIVE	91,14%	93,50%	47,08%	62,63%
L_LIVE_2	93,24%	96,01%	59,56%	73,51%
L_UHD	92,58%	69,07%	95,84%	80,28%
L_UHD_2	84,84%	100,00%	3,81%	7,33%
L_VQ	84,24%	0,00%	0,00%	0,00%
L_VQ_2	84,39%	100,00%	0,97%	1,93%
L_LIVE_UHD_VQ	95,43%	88,41%	81,68%	84,91%
L_LIVE_UHD_VQ_2	95,13%	95,67%	72,39%	82,42%
CL_LIVE	90,86%	64,30%	94,51%	76,53%
CL_LIVE_2	95,68%	82,80%	91,59%	86,97%
CL_UHD	97,68%	94,96%	90,09%	92,46%
CL_UHD_2	96,76%	96,00%	82,92%	88,98%
CL_VQ	84,24%	0,00%	0,00%	0,00%
CL_VQ_2	84,24%	0,00%	0,00%	0,00%
CL_LIVE_UHD_VQ	96,99%	88,60%	92,83%	90,67%
CL_LIVE_UHD_VQ_2	97,60%	95,88%	88,58%	92,09%

Algoritmi HBFDA i RTFDA pri testiranju nad video sekvencama SMVS FullHD baze daju velik broj FP, odnosno detektiraju da je smrzavanje prisutno u okvirima koji ga ne sadrže. Velik broj FP očituje se niskom mjerom preciznosti (HBFDA 22.5%, RTFDA 49.42%) kao što je vidljivo u tablici 4.5. Neki od modela zasnovanih na strojnom učenju ostvaruju bolji odnos preciznosti i odziva, a samim time i bolju F1 mjeru nego predloženi algoritmi. Najlošije rezultate ostvaruju modeli koji su trenirani pomoću podataka iz VQEG HDTV baze. U slučaju CNN arhitekture modeli koji su trenirani uz dvije dodatne sekvence ostvaruju bolje rezultate od ostalih CNN modela. Modeli zasnovani na CNN arhitekturi, izuzev onih koji su trenirani korištenjem VQEG HDTV baze, ostvaruju slične rezultate poput RTFDA i HBFDA algoritama, imaju visok odziv (>90%) i nižu preciznost. Model C_UHD_2 ostvaruje visok odziv (>90%) i veću preciznost u odnosu na oba testirana algoritma. Najbolju F1 mjeru ostvaruju modeli zasnovani na CNN-LSTM

arhitekturi. Model CL_UHD ostvaruje najbolju F1 mjeru od 92,46%, a CNN-LSTM modeli trenirani na 3 baze također ostvaruju F1 mjeru veću od 90%.

4.5. Analiza rezultata

Iz prethodno opisanih rezultata testiranja nad video sekvencama pojedine baze može se zaključiti da modeli trenirani pomoću podataka iz VQEG HDTV baze daju lošije rezultate naspram ostalih. Pri testiranju nad video sekvencama iz VQEG HDTV baze većina modela, izuzev modela zasnovanih na LSTM arhitekturi, daje izrazito dobre rezultate. Ovakav odnos rezultata ukazuje da smrzavanja prisutna u VQEG HDTV bazi sadrže najmanju odnosno nikakvu količinu šuma u ponovljenim okvirima.

Modeli zasnovani na LSTM arhitekturi pokazali su se najlošijima iako u pojedinim slučajevima ostvaruju izuzetne rezultate. Model L_LIVE_2 pri testiranju na VQEG HDTV bazi ostvaruje potpuno točnu detekciju, što premašuje rezultate postojećih algoritama, međutim pri testiranju na ostale dvije baze ostvaruje lošije rezultate.

Pojedini CNN modeli ostvaruju slične rezultate kao postojeći algoritmi. Svi CNN modeli kod kojih je skup za trening proširen sa dvije dodatne sekvence ostvarili su bolje rezultate naspram modela bez dodatnih sekvenci u skupu za trening. Model C_UHD_2 pokazao se kao najbolji među ostalim CNN modelima prema rezultatima testiranja na tri različite baze.

CNN-LSTM modeli koji su trenirani na jednom skupu podataka, a testirani na ostalima, ostvaruju slične rezultate neovisno o testnom skupu, čime su se pokazali robusnijima naspram LSTM i CNN modela. Osim toga, CNN-LSTM modeli ostvaruju bolju F1 mjeru nego postojeći algoritmi u slučaju testiranja na SVMS FullHD bazi. Korištenje triju baza u skupu za trening kod CNN-LSTM modela u svim kombinacijama trening/test podataka daje F1 mjeru veću od 90%.

5. ZAKLJUČAK

U ovom radu opisana je pojava artefakta smrzavanja slike u video signalu te problemi koje treba uzeti u obzir pri razvoju rješenja za detekciju artefakta smrzavanja slike bez pristupa referentnom sadržaju. Opisano je nekoliko postojećih rješenja te postojeće baze video sekvenci koje sadrže artefakt smrzavanja.

U sklopu rada razvijena su rješenja za detekciju artefakta smrzavanja zasnovana na strojnom učenju. Stvorene su tri različite arhitekture neuronskih mreža koje su trenirane na temelju četiri dostupne baze video sekvenci koje sadrže artefakt smrzavanja. Sveukupno je stvoreno 48 različitih modela. Svaki model nosi naziv ovisan o arhitekturi i bazi/bazama podataka koji su korišteni u skupu za treniranje. Rezultati svih modela prikazani su posebno za svaki testni skup te je dana usporedba sa rezultatima dva postojeća algoritma.

Rezultati testiranja ukazuju kako je detekciju artefakta smrzavanja moguće ostvariti pomoću rješenja zasnovanog na strojnom učenju. Prema rezultatima testiranja istreniranih modela nad video sekvencama LIVE Mobile baze pet modela ostvaruje preciznost i odziv veću od 90%, a jedan model zasnovan na konvolucijskoj neuronskoj mreži (C_UHD_2) ostvaruje preciznost i odziv veću od 96%. Iako nijedan model ne nadmašuje rezultate postojećih algoritama pri testiranju na LIVE Mobile bazi razlika u F1 mjeri između modela C_UHD_2 i HBFDA algoritma iznosi 1,08%, odnosno 2,03% naspram RTFDA algoritma. Pri testiranju istreniranih modela na UHD VFD i VQEG HDTV bazama video sekvenci najbolji se pokazao C_SM_2 model koji ostvaruje bolju F1 mjeru od HBFDA algoritma za obje baze, dok za LIVE Mobile bazu ostvaruje 8,12% manju F1 mjeru nego HBFDA. Pri testiranju istreniranih modela nad video sekvencama SMVS FullHD baze šest modela ostvaruje F1 mjeru veću od 85%, a najbolju F1 mjeru od 92,46% ostvaruje CL_UHD model čime je ostvarena uspješnija detekcija artefakta smrzavanja naspram HBFDA algoritma koji ostvaruje F1 mjeru od 36,45% i RTFDA algoritma koji ostvaruje F1 mjeru od 64,71%.

LITERATURA

- [1] R. Grbić, D. Stefanović, M. Vranješ, and M. Herceg, "Real-time video freezing detection for 4K UHD videos," *J. Real-Time Image Process.*, vol. 17, no. 5, pp. 1211–1225, 2020.
- [2] D. Babić, D. Stefanović, M. Vranješ, and M. Herceg, "Real-time no-reference histogram-based freezing artifact detection algorithm for UHD videos," *Multimed. Tools Appl.*, vol. 78, no. 13, pp. 17949–17971, 2019.
- [3] D. C. Mocanu, J. Pokhrel, J. P. Garella, J. Seppänen, E. Liotou, and M. Narwaria, "No-reference video quality measurement: added value of machine learning," *J. Electron. Imaging*, vol. 24, no. 6, p. 061208, 2015.
- [4] M. A. Usman, "PERFORMANCE EVALUATION OF VIDEO QUALITY ASSESSMENT METHODS BASED ON FRAME FREEZING," 2014.
- [5] Quan Huynh-Thu and M. Gha, "No-reference temporal quality metric for video impaired by frame freezing artefacts," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, Kairo, Egipt, 2009, pp. 2221–2224.
- [6] Y. Xue, B. Erkin, and Y. Wang, "A Novel No-Reference Video Quality Metric for Evaluating Temporal Jerkiness due to Frame Freezing," *IEEE Trans. Multimed.*, vol. 17, no. 1, pp. 134–139, 2015.
- [7] M. A. Usman, M. R. Usman, and S. Y. Shin, "The Impact Of Temporal Impairment On Quality Of Experience (Qoe) In Video Streaming: A No Reference (Nr) Subjective And Objective Study," 2015.
- [8] M. Shahid, A. Rossholm, B. Lövsström, and H.-J. Zepernick, "No-reference image and video quality assessment: a classification and review of recent approaches," *EURASIP J. Image Video Process.*, vol. 2014, no. 1, p. 40, 2014.
- [9] I. El Naqa and M. J. Murphy, "What Is Machine Learning?," in *Machine Learning in Radiation Oncology*, I. El Naqa, R. Li, and M. J. Murphy, Eds. Cham: Springer International Publishing, 2015, pp. 3–11.
- [10] A. K. Moorthy, L. K. Choi, A. C. Bovik, and G. de Veciana, "Video Quality Assessment on Mobile Devices: Subjective, Behavioral and Objective Studies," *IEEE J. Sel. Top. Signal Process.*, vol. 6, no. 6, pp. 652–671, 2012.
- [11] A. K. Moorthy, L. K. Choi, G. deVeciana, and A. C. Bovik, 'Mobile Video Quality Assessment Database,' IEEE ICC Workshop on Realizing Advanced Video Optimized Wireless Networks, Ottawa, Kanada, 2012.
- [12] A. K. Moorthy, L. K. Choi, G. deVeciana, and A. C. Bovik, 'Subjective Analysis of Video Quality on Mobile Devices,' Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM) (invited article), Scottsdale, Sjedinjene Američke Države, 2012.
- [13] Webster, A., Speranza, F.: VQEG report on the validation of video quality models for high definition video content. Technical report, Video Quality Experts Group (2010).
- [14] Babić, D., Vranješ, M.: FERIT-RTRK UHD VFD. <http://www.rtrk.com/other/UHDVideoDBReadme.html> (2018).
- [15] "PIXABAY." <https://pixabay.com/>, pristup ostvaren 11.7.2022.
- [16] "Sunrise, Sunbeams, Trees, Woods," PIXABAY. <https://pixabay.com/videos/sunrise-sunbeams-trees-woods-83880/>, pristup ostvaren 11.7.2022.
- [17] "Sunrise, Fog, Morning, Nature," PIXABAY. <https://pixabay.com/videos/sunrise-fog-morning-nature-trees-31549/>, pristup ostvaren 11.7.2022.
- [18] S. Wolf, "A No Reference (NR) and Reduced Reference (RR) Metric for Detecting Dropped Video Frames," *Natl. Telecommun. Inf. Adm. NTIA*, p. 6, 2008.

- [19] V. Skobic, U. Maric, S. Tomic, and I. Resetar, "Real time video freezing detection implementation on FPGA," in *2016 24th Telecommunications Forum (TELFOR)*, Beograd, Srbija, 2016, pp. 1–3.
- [20] S. van Kester, T. Xiao, R. E. Kooij, K. Brunnström, and O. K. Ahmed, "Estimating the impact of single and multiple freezes on video quality," San Francisco Airport, Sjedinjene Američke Države, 2011, p. 786500.
- [21] "Keras: the Python deep learning API." <https://keras.io/>, pristup ostvaren 14.7.2022.
- [22] "OpenCV," *OpenCV*. <https://opencv.org/>, pristup ostvaren 14.7.2022.
- [23] "Anaconda," *Anaconda*. <https://www.anaconda.com/products/distribution>, pristup ostvaren 15.7.2022.

SAŽETAK

U radu je opisana i prikazana manifestacija artefakta smrzavanja slike u video signalu. Opisan je temeljni princip detekcije artefakta smrzavanja bez pristupa referentnom sadržaju, problemi pri detekciji koje potrebno uzeti u obzir te ideja rješenja zasnovana na strojnom učenju. Dan je pregled postojećih baza video sekvenci koje sadrže artefakt smrzavanja i pregled postojećih algoritama. Predstavljena su tri vlastita rješenja za detekciju artefakta smrzavanja zasnovana na strojnom učenju. Rezultati postojećih algoritama za detekciju artefakta smrzavanja uspoređeni su s vlastitim rješenjima.

Ključne riječi: artefakt smrzavanja slike, strojno učenje, CNN, LSTM

MACHINE LEARNING BASED FREEZING ARTIFACTS DETECTION IN VIDEO SIGNAL

ABSTRACT

This master's thesis describes and presents manifestation of freezing artefact. The basic principle of no-reference freezing artefact detection, detection problems that need to be considered and the idea of a solution based on machine learning are described. An overview of the existing databases with video sequences containing the freezing artifact is given, as well as an overview of the existing algorithms. Three proposed solutions for freezing artifact detection based on machine learning are presented. The results of two existing no-reference freezing detection algorithms are compared with our own solutions.

Keywords: freezing artefact, machine learning, CNN, LSTM

ŽIVOTOPIS

Domagoj Vukadin rođen je 22. ožujka 1997. godine u Đakovu. 2016. godine po završetku matematičko-prirodoslovne gimnazije Antun Gustav Matoš u Đakovo upisuje preddiplomski sveučilišni studij na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku, smjer računarstvo. 2020. godine uspješno završava preddiplomski studij te upisuje diplomski sveučilišni studij automobilsko računarstvo i komunikacije.

Potpis:

PRILOZI

- P.2.1. poveznica na SMVS FullHD bazu video sekvenci: <http://materijali.ferit.hr/>
- P.3.1. konfiguracijska datoteka potrebna za stvaranje programskog okruženja (priloženo na DVD-u uz rad)
- P.3.2. *data_generator.py* i *hdf5_dataset_maker.py* datoteke (priloženo na DVD-u uz rad)
- P.3.3. *input_video_config.json* datoteka koja sadržava meta podatke o video sekvencama (priloženo na DVD-u uz rad)
- P.3.4. *hdf5_train.py* i *log.py* datoteke (priloženo na DVD-u uz rad)
- P.3.5. *freezing_detection_all_videos.py* datoteka (priloženo na DVD-u uz rad)
- P.3.6. direktorij koji sadrži sve istrenirane modele predložene u ovom radu (priloženo na DVD-u uz rad)