Image Super-Resolution Using Deep Convolutional Networks

Seminarski rad u okviru kursa Računarska inteligencija

Matematički fakultet,univerzitet u Beogradu

Vladimir Mandić mi18465

**Sadržaj:**

1. **Uvod**
2. **Metode zasnovane na primerima**

**3. Algoritam**

**4. Eksperimentalni rezultati**

**5. Doprinos studije**

**6. Literatura**

1. **Uvod**

Super rezolucija slike ima za cilj dobijanje slike visoke rezolucije od slika niske rezolucije.Cilj ovog rada je upoređivanje različitih metoda za dobijanje super rezolucije,korišćenjem vizuelnog prikaza i odgovarajućih merenja.

Razne metode,bez obzira na temeljne korake dobijanja super rezolucije,imaju određene načine stabilizovanja nizova slika,rekonstruisanja,te generisanja slika visoke rezolucije.

Algoritmi super rezolucije za jednu sliku se mogu podeliti na 4 tipa:

1. Modeli predviđanja,
2. Metode zasnovane na ivicama,
3. Slike statističkih metoda,
4. Metode koje su zasnovane na delovima slike(ili primerima)

Ove metode su temeljno istražene i procenjene u Jangovom radu.Među njima,metoda zasnovana na primerima dostiže vrhunske performanse.

Interne metode zasnovane na primerima koriste svojstvo samosličnosti i generišu primere delova slike,a na osnovu ulazne slike.Ovo je prvo predloženo u Glasnerovom radu.Pojam samosličnosti se koristi u matematici,samosličan je objekat koji ima isti oblik kao jedan ili više svojih delova

Sa druge strane,spoljne metode koje su zasnovane na primerima,uče mapiranja između delova (niske/visoke) rezolucije,a iz spoljašnjih skupova podataka.U pionirskom radu Freeman-a,rečnici su predstavljeni kao parovi

delova (jednog dela niske i jednog dela visoke rezolucije) i najbliži sused tog dela koji je nađen u prostoru niske rezolucije,zajedno sa odgovarajućim delom visoke rezolucije se koristi za rekonstrukciju.

Sa druge strane,u Čengovom radu je uvedena tehnika višestrukog ugrađivanja,koja zamenjuje strategiju najbližeg suseda (NN).

Za slike u boji,gore pomenute metode prvo transformišu problem u drugi prostor boja (YcbCr ili YUV),a SR se primenjuje samo na kanal osvetljenosti.Postoje radovi koji pokušavaju da reše sve kanale istovremeno npr Kim,Kwon i Dai.

Predložena metoda za super rezoluciju slike je deep learning.Naša metoda direktno uči preslikavanje od slike sa niskom rezolucijom do slike sa visokom rezolucijom.To preslikavanje je reprezentovano kao duboka konvolutivna neuronska mreža koja uzima sliku sa malom rezolucijom kao ulaz i daje sliku sa visokom rezolucijom kao izlaz.

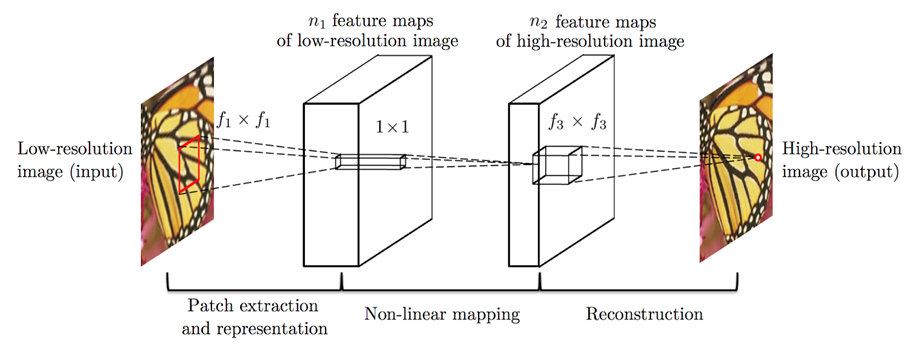
Konvolutivna neuronska mreža je nedavno dobila ogromnu popularnost,pre svega zbog uspeha u klasifikaciji slika,a takođe se uspešno primenjuje i u drugim oblastima računarstva,npr otkrivanju predmeta,prepoznavanju lica itd.Nekoliko bitnih faktora koji su uticali na napredak **CNN** mreža su:

1. Efikasna trening implementacija na modernim GPU
2. Rectifiel Linear Unit (ReLU) koja konvergenciju čini mnogo bržom,dok kvalitet i dalje ostaje dobar
3. Lak pristup ogromnom spektru podataka(poput Image Net-a)
4. **Metode zasnovane na primerima**

Metode zasnovane na primerima mogu biti formulisane za generisanje slika putem super-rezolucije ili mogu biti dizajnirane tako da odgovaraju konkretnim zadacima iz domena.Metoda zasnovana na oskudnom tj retkom kodiranju je jedna od reprezentativnih SR metoda,koja je zasnovana na primerima.Ova metoda uključuje nekoliko koraka u rešavanju problema.Prvi korak je da se izvuku delovi iz ulazne slike.Ovi delovi se zatim enkodiraju koristeći rečnike niske rezolucije.Proređeni koeficijenti se prosleđuju rečniku visoke rezolucije radi rekonstrukcije delova koji imaju visoku rezoluciju.Na kraju,rekonstruisani delovi se objedinjuju kako bi se proizveo konačni izlaz.Ovaj pipeline je podeljen od strane eksternih metoda zasnovanih na primerima.Ono što se pokazuje u ovom radu je da je taj pipeline ekvivalentan dubokoj konvolutivnoj neuronskoj mreži.Dakle,razmatra se konvolutivna neuronska mreža koja direktno uči mapiranje,od slike koja ima nisku rezoluciju do slike koja ima visoku rezoluciju.

1. **Algoritam**

Naš metod se znatno razlikuje od postojećih spoljnih metoda zasnovanih na primerima,po tome što naš metod ne uči eksplicitno rečnike ili pipelinove za modelovanje dela prostora,oni se implicitno postižu preko skrivenih slojeva.Model imenujemo sa SRCNN tj Super-rezoluciona konvolutivna neuronska mreža.



SRCNN-duboka konvolutivna neuronska mreža koja uči mapiranje od slike sa niskom rezolucijom do slike sa visokom rezolucijom.Moj model se može koristiti kako bi poboljšali kvalitet slike koja ima nisku rezoluciju

Za procenu performansi mreže,koristio sam sledeće pokazatelje kvaliteta slike:

1. Peak signal to noise ratio (PSNR) – PSNR je vrednost koja predstavlja logaritamski odnos snage signala ekvivalentnog kvadratu maksimalne moguće vrednosti elementa slike i snage šuma ekvivalentne iznosu srednje kvadratne greške MSE.
2. Mean squared error (MSE)
3. Structural similarity (SSIM) index – Metoda za merenje sličnosti između 2 slike.SSIM je metoda napravljena kako bi se poboljšale tradicionalne metode za određivanje kvaliteta slike kao što su PSNR i MSE.Glavni zadatak ove metode je da se što više približi ljudskoj percepciji.Zavisnost između piksela je veća što su oni međusobno bliži i to je informacija koja opisuje strukturu objekta na slici.Rezultati se dobijaju preko 3 lokalna prozora veličine 8x8.Ti prozori su podeljeni na svetlost,kontrast i strukturu.Rezultati se povezuju u 1 prozor i konačni rezultat je SSIM indeks.SSIM indeks se računa poređenjem originalne i ispitivane.Ukoliko je SSIM indeks 1 tada ne postoji razlika između referentne i testirane slike

Metodologija u izradi projekta:

1. Priprema degradiranih slika menjanjem veličine slike na dole,pa na gore.Nove slike imaju istu rezoluciju kao i početne slike.Kada smanjimo veličinu slike,mi informacije o originalnim pikselima skladištimo na manjem prostoru,tako da mi gubimo te informacije kada povećamo veličinu slike.Ovo je urađeno korišćenjem modula OpenCV
2. Konstruisanje modela:konstruisao sam tri modela sa sledećim parametrima:
3. Prvi model se sastoji od 3 konvolutivna sloja,broj filtera je redom 128,64,3 a odgovarajuće veličine filtera su 9,3 i 5.Korišćena je aktivaciona funkcija relu i prilikom dodavanja prvog sloja navedeno je da je ulaz dimenzije (128,128,3) jer su slike 3D,a veličine (128,128)
4. Drugi model se sastoji od 4 sloja konvolucije,broj filtera po slojevima je redom 64,32,16 i 3,a veličine filtera su 9,1,1 i 5.Korišćena je aktivaciona funkcija gelu.Dimenzija input-a prosleđena prvom sloju je ista kao i u slučaju prethodnog modela.
5. Treći model se sastoji od 3 sloja konvolucije,broj filtera po slojevima je redom 64,32,3,a filteri su veličine 9,1 i 5 respektivno.Korišćena je aktivaciona funkcija elu.Input je istih dimenzija kao u prethodna 2 modela,ovo je zato što smo modele obučavali nad skupom slika General-100 čije su slike u boji i odgovarajuće su veličine.
6. Učitavanje SRCNN mreže:Najpre su učitane slike iz skupa General-100 koje ću koristiti za treniranje modela,zatim sam podešavao parametre za kompiliranje modela,za funkciju gubitka korišćena je metoda srednje kvadratne greške,kao optimizator je odabran Adam,a metrike koje će biti praćene u toku obučavanja modela su psnr i ssim koje sam prethodno implementirao.Samo obučavanje modela se vrši u 100 epoha,a paketi su veličine 32.
7. Testiranje modela – nakon treniranja mreže,mi možemo da primenimo super rezoluciju na svim ulaznim slikama.Nakon obrade možemo izračunati PSNR,MSE i SSIM na slikama koje proizvodimo.
8. **Eksperimentalni rezultati**

MSE – što je veća vrednost ovog parametra to je manja rezolucija slike

PSNR - želimo da bude što je moguće veća vrednost zato što želimo da šum bude što manji,ovaj parametar predstavlja odnos signala i šuma

SSIM:ovaj parametar se kreće u intervalu od 0.0 do 1.0 i prikazuje sličnost između originalne i odgovarajuće slike

**Bilinearna interpolacija** – algoritam gde se koeficijenti interpolacije računaju na osnovu udaljenosti tačaka iz osnovnog rastera od interpolacione tačke

w(1,1) = (1-y)\*(1-x) w(1,2) = (1-y)\*x

w(2,1) = y \* (1-x) w(2,2) = y \* x



Na osnovu ova 2 para slika(1 par čine referentna slika i slika koja je rezultat primene bilinearne interpolacije) možemo zaključiti sledeće:Interpolirana slika drugog para ima veću vrednost parametra PSNR od odgovarajuće slike prvog para što znači da ima manji šum u odnosu na referentnu sliku,takođe vrednost parametra MSE je značajno veća za prvu interpoliranu sliku što nam ukazuje da ona ima manju rezoluciju u odnosu na referentnu sliku.Poslednji parametar tj. SSIM je veći u slučaju druge interpolirane slike što govori da je ona sličnija svojoj referentnoj slici.

**Bikubična interpolacija** – daje bolji kvalitet interpolirane slike u odnosu na bilinearnu interpolaciju,a na račun povećane složenosti izračunavanja.Pri interpolaciji koristi se 16 piksela iz originalnog rastera (4x4 okolina) i vrednosti piksela koji su bliži interpoliranom pikselu imaju veći uticaj na interpoliranu vrednost.

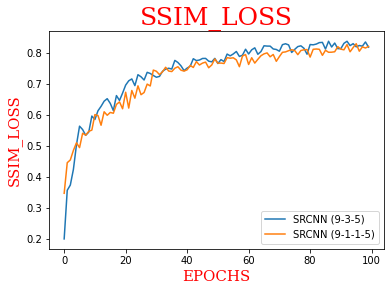
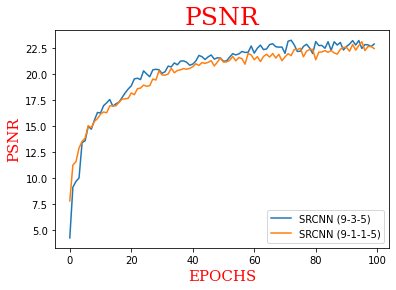


Na osnovu ova 2 para slika(1 par čine referentna slika i slika koja je rezultat primene bilinearne interpolacije) možemo zaključiti sledeće:Interpolirana slika drugog para ima veću vrednost parametra PSNR od odgovarajuće slike prvog para što znači da ima manji šum u odnosu na referentnu sliku,takođe vrednost parametra MSE je značajno veća za prvu interpoliranu sliku što nam ukazuje da ona ima manju rezoluciju u odnosu na referentnu sliku.Poslednji parametar tj. SSIM je veći u slučaju druge interpolirane slike što govori da je ona sličnija svojoj referentnoj slici.

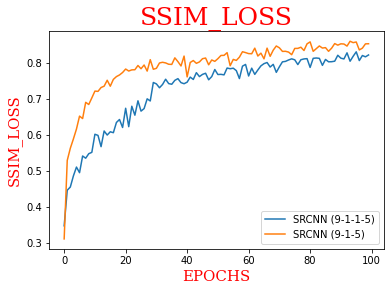
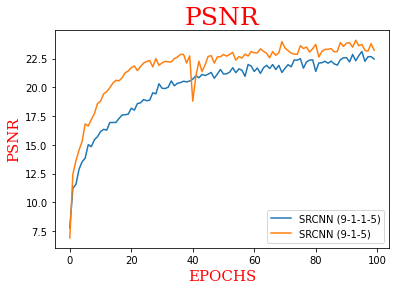
**SRCNN**

**SEKCIJA VEZANA ZA TRENING**

Prilikom obučavanja modela SRCNN(9-3-5) i SRCNN(9-1-1-5) nad slikama iz skupa General-100 dobijeni su rezultati prikazani na slikama.Možemo zaključiti da se nad trening podacima prvi pomenuti model za nijansu bolje pokazao jer su obe metrike (PSNR i SSIM) veće,odnosno manji je šum,a sličnost sa referentnim slikama je veća.



Sa druge strane,prilikom obučavanja modela SRCNN(9-1-1-5) i SRCNN(9-1-5) nad slikama iz istog skupa dobijeni su sledeći rezultati:



Te u ovom slučaju možemo zaključiti da se nad podacima za trening,drugi pomenuti model (tj SRCNN(9-1-5)) značajno bolje pokazao jer su vrednosti za obe metrike značajno veće.

**SEKCIJA VEZANA ZA TESTIRANJE**

Testiranje je izvršeno na slikama iz skupa Set5.Korišćene su metrike PSNR i SSIM.





Za sliku leptira dobijamo da model SRCNN(9-1-1-5) ima bolje performanse od modela SRCNN(9-3-5) jer ima manji šum (veću vrednost parametra PSNR-a),a takođe je i sličnija referentnoj slici jer ima veću vrednost parametra SSIM.Sa druge strane,u slučaju ptice slika koja je dobijena primenom modela SRCNN(9-3-5) ima veći šum jer ima manju vrednost PSNR-a,ali je sličnija referentnoj slici.





Prilikom poređenja rezultata dobijenih primenom modela SRCNN(9-1-5) i SRCNN(9-1-1-5) na slike leptira i ptice iz skupa Set5 dolazimo do sledećeg zaključka:

U oba slučaja dobijamo da je drugi pomenuti model bolji tj. dalje bolje performanse jer su vrednosti PSNR i SSIM veće.

1. Doprinos studije

**Doprinosi ove studije su u 3 aspekta i to su**:

1. **Predstavljam potpunu konvolutivnu neuronsku mrežu za slike sa super rezolucijom.Mreža direktno uči mapiranje od slike sa niskom do slike sa visokom rezolucijom,uz malo preprocesiranja**
2. **Uspostavljam vezu između SR metode koja je zasnovana na dubokom učenju i tradicionalnih SR metoda koje su zasnovane na oskudnom tj retkom kodiranju**
3. **Pokazujem da je duboko učenje korisno u klasičnom računarskom pogledu super rezolucije i da može postići dobar kvalitet**
4. **Literatura**

**Kao pomoć u izradi projekta koristio sam naučni rad Image-Super-Resolution Using Deep Convolutional Networks (autori rada su:Chao Dong,Chen Change Loy and Xiaoou Tang),a rad se nalazi na stranici**

<https://arxiv.org/pdf/1501.00092.pdf>