

Super rezolucija slika korišćenjem konvolutivnih neuronskih mreža

Vladimir Mandić
mi18465

SRCNN je duboka konvolutivna neuronska mreža koja uči mapiranje od jednog do drugog kraja (od slike sa niskom rezolucijom do slike sa visokom rezolucijom). Rezultat je taj da se ova mreža može upotrebiti radi poboljšanja kvaliteta slike.

Za računanje performansi mreže korišćene su sledeće metrike:

- 1) Peak to signal noise ratio (PSNR)
- 2) Mean squared error (MSE)
- 3) Structural similarity (SSIM) index

Korišćen je modul OpenCV radi pre/post procesiranja slika. Takođe je korišćeno i konvertovanje između prostora boja (RGB, BGR, YCrCb). Ovo je neophodno jer je SRCNN mreža trenirana na kanalu osvetljenja Y prostora boja YCrCb

```
def model():  
    SRCNN = Sequential([  
        Conv2D(filters=128, kernel_size=(9,9), padding='valid',  
            activation='relu', kernel_initializer = 'glorot_uniform',  
            use_bias=True, input_shape = (None, None, 1)),  
        Conv2D(filters=64, kernel_size=(3,3), padding='same',  
            activation='relu', kernel_initializer = 'glorot_uniform',  
            use_bias=True),  
        Conv2D(filters=1, kernel_size=(5,5), padding='valid',  
            activation='linear', kernel_initializer = 'glorot_uniform',  
            use_bias=True)  
    ], name = 'MyModel')  
    SRCNN.compile(optimizer = Adam(learning_rate=0.0003), loss =  
        'mean_squared_error', metrics = ['mean_squared_error'])  
    return SRCNN
```

Na prethodnoj slici možemo da vidimo od čega se sastoji naša SRCNN, dakle u pitanju su 3 konvolutivna sloja

1. sloj sa 128 filtera koji su veličine (9,9), za aktivaciju je korišćena relu f-ja
2. sloj sa 64 filtera koji su veličine (3,3), za aktivaciju je takođe korišćena relu f-ja
3. sloj sa 1 filterom velčine (5,5), a za aktivaciju je korišćena je linearna f-ja

Ono što je još neophodno napomenuti za samu mrežu je to da je korišćen optimizator Adam, dok je stopa učenja bila 0.0003

Prikaz rezultata za sliku flowers.bmp

Degradirana slika:

MSE: 367.07440518423914

PSNR: 27.25447511864764

SSIM: 0.8690381699992461

Rekonstruisana slika:

MSE: 210.67191712785583

PSNR: 29.665949682056713

SSIM: 0.898887007611708

Originalna slika



Degradirana slika



SRCNN



Objašnjenje toga šta predstavljaju slike sa prethodnog slajda:

Slika levo - originalna (referentna slika) koja pripada skupu Set14

Slika u sredini - degradirana slika, to je slika koja je dobijena tako što smo prvo smanjili sliku i na taj način informacije o originalnim pikselima smestili u manji prostor, time smo izgubili određene informacije koje ostaju izgubljene i nakon povećanja slike koje smo neposredno nakon toga izvršili

Slika desno - primenili smo SRCNN na degradiranu sliku

Objašnjenje samih rezultata za dati primer

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

Dakle za ovaj primer (flowers.bmp) možemo zaključiti sledeće:

MSE kod 3. slike je znatno manji, što je i logično jer to znači da je znatno veća rezolucija slike, a to je i logično budući da je primenjena SRCNN mreža. Dalje, PSNR vrednost je neznatno veća što znači da je šum manji, a takođe i parametar SSIM ima veću vrednost što znači da je slika nakon primene SRCNN sličnija originalnoj tj referentnoj.

Na narednim slajdovima ću prikazati rezultate za slike iz skupa Set5

Originalna slika
butterfly_GT.bmp



Degradirana slika
butterfly_GT.bmp



MSE:654.5066978272282
PSNR:24.74287520369977
SSIM:0.8789000067908175

SRCNN
butterfly_GT.bmp

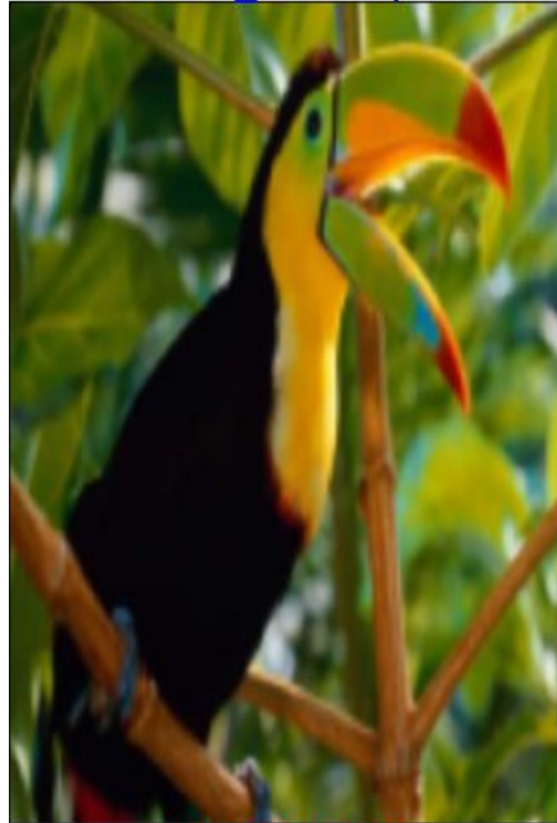


MSE:179.01146505444632
PSNR:30.37320768683507
SSIM:0.9519657744383796

Originalna slika
bird_GT.bmp

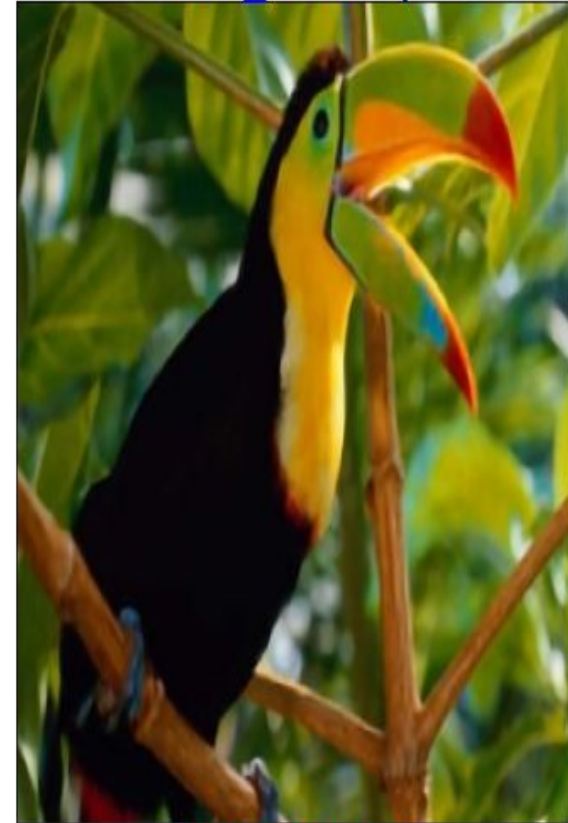


Degradirana slika
bird_GT.bmp



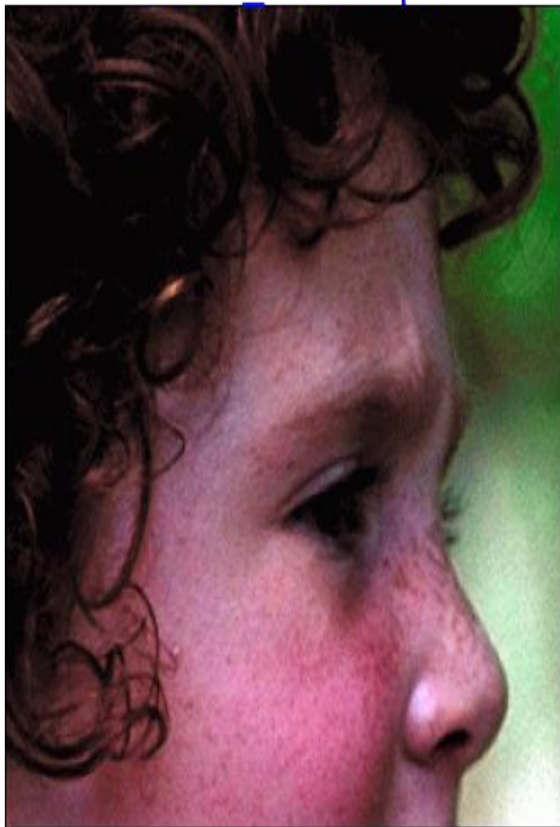
MSE:98.23142459567318
PSNR:32.97951173163997
SSIM:0.9533077338285486

SRCNN
bird_GT.bmp

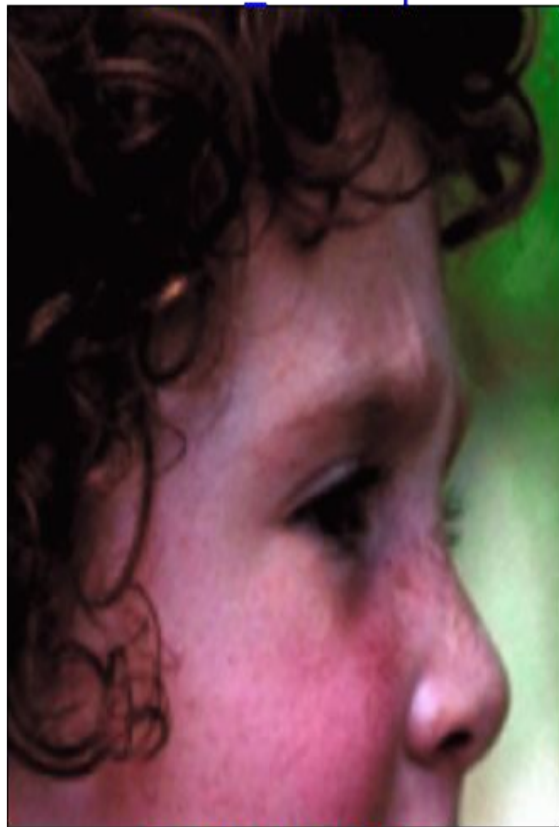


MSE:43.1028670447385
PSNR:36.55695456788854
SSIM:0.9693832103556429

Originalna slika
head_GT.bmp



Degradirana slika
head_GT.bmp



MSE:157.67491478348694
PSNR:30.924390106396494
SSIM:0.8012033642786381

SRCNN
head_GT.bmp



MSE:132.49316163784033
PSNR:31.680081519531733
SSIM:0.8289816946308021

Originalna slika
woman_GT.bmp



Degradirana slika
woman_GT.bmp



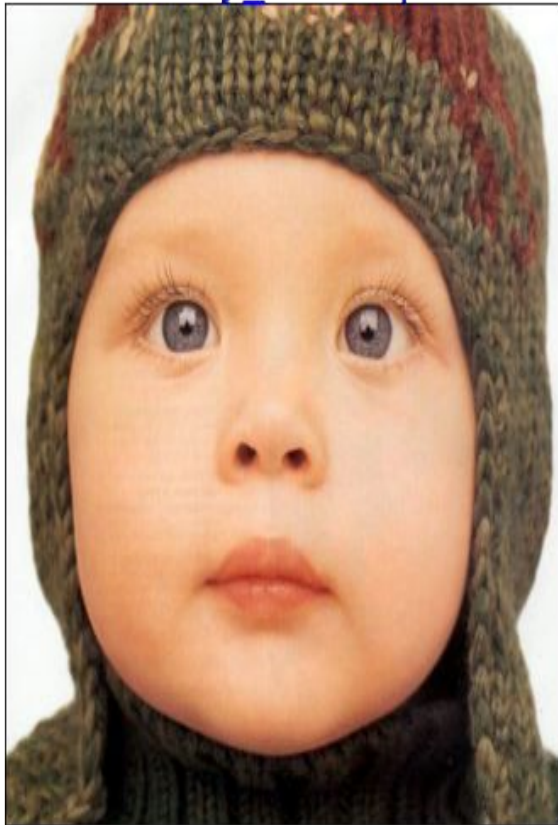
MSE:225.3165263748597
PSNR:29.374085683369664
SSIM:0.9336749260789752

SRCNN
woman_GT.bmp



MSE:84.53644781144781
PSNR:33.631576205958524
SSIM:0.9653791777540882

Originalna slika
baby_GT.bmp



Degradirana slika
baby_GT.bmp



MSE:73.13272366574732
PSNR:34.260898675294186
SSIM:0.9345509319632762

SRCNN
baby_GT.bmp



MSE:46.07448670182739
PSNR:36.26741109977648
SSIM:0.9550654288713559

Bilinearna interpolacija

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

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

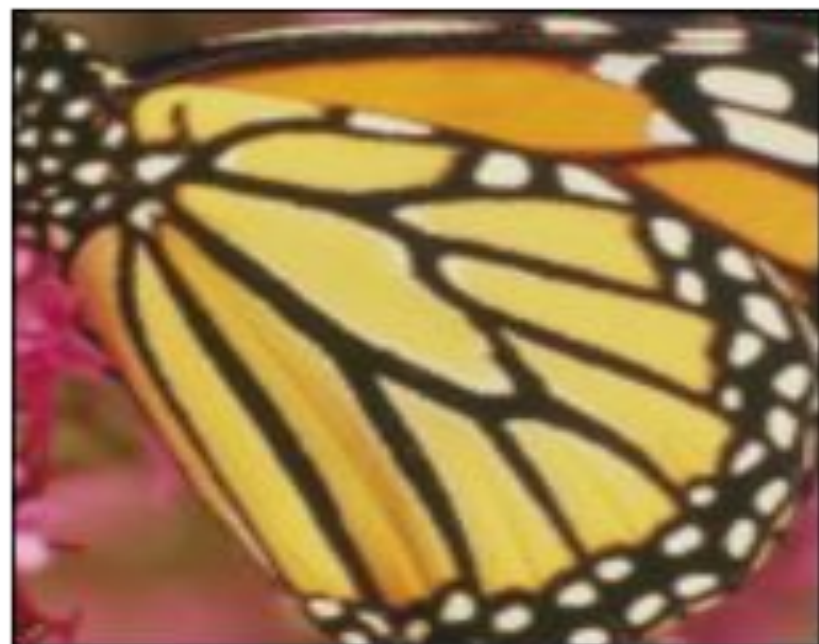
$$w(2,1) = y*(1-x) \quad w(2,2) = y*x$$

Na narednim slajdovima ćemo prikazati rezultate primene bilinearne interpolacije na slike iz skupa Set5

Referentna slika



**Interpolirana slika
(bilinearna interpolacija)**



**file:butterfly_GT.bmp
PSNR:24.782076560337416
MSE:648.6254119873047
SSIM:0.8791344763843051**

Referentna slika



**Interpolirana slika
(bilinearna interpolacija)**



**file:bird_GT.bmp
PSNR:32.896644728720005
MSE:100.12375819830247
SSIM:0.9533644866026473**

Referentna slika



**Interpolirana slika
(bilinearna interpolacija)**



**file:head_GT.bmp
PSNR:31.020502848237534
MSE:154.2237755102041
SSIM:0.8011121330733371**

Referentna slika



**Interpolirana slika
(bilinearna interpolacija)**



**file:woman_GT.bmp
PSNR:29.326236280817465
MSE:227.812729498164
SSIM:0.9335397280466592**

Referentna slika



**Interpolirana slika
(bilinearna interpolacija)**



**file:baby_GT.bmp
PSNR:34.371806409661986
MSE:71.28874588012695
SSIM:0.9356987872724932**

Bikubična interpolacija

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 (4 x 4 okolina) i vrednosti piksela koji su bliži interpoliranom pikselu imaju veći uticaj na interpoliranu vrednost.

Na narednim slajdovima ćemo prikazati rezultate primene bikubične interpolacije na slike iz skupa Set5

Referentna slika



**Interpolirana slika
(bikubična interpolacija)**



**file:butterfly_GT.bmp
PSNR:26.984580538658133
MSE:390.6101837158203
SSIM:0.9154365370881684**

Referentna slika



**Interpolirana slika
(bikubična interpolacija)**



**file:bird_GT.bmp
PSNR:36.19699766428704
MSE:46.82759452160494
SSIM:0.9749036436249959**

Referentna slika



**Interpolirana slika
(bikubična interpolacija)**



**file:head_GT.bmp
PSNR:31.502885310138158
MSE:138.01080357142857
SSIM:0.8310886917931465**

Referentna slika



**Interpolirana slika
(bikubična interpolacija)**



**file:woman_GT.bmp
PSNR:31.70507985630281
MSE:131.7327111383109
SSIM:0.9575809733826678**

Referentna slika



**Interpolirana slika
(bikubična interpolacija)**



**file:baby_GT.bmp
PSNR:36.09787721994766
MSE:47.90864562988281
SSIM:0.9584238177318317**

Na kraju ćemo još samo da uporedimo rezultate za sliku butterfly.bmp između ove 3 metode (SRCNN, bilinearna i bikubična interpolacija)

1) SRCNN

MSE 179.01146505444632
PSNR 30.37320768683507
SSIM 0.9519657744383796

3) Bikubična interpolacija

MSE 390.6101837158203
PSNR 26.984580538658133
SSIM 0.9154365370881684

2) Bilinearna interpolacija

MSE 648.6254119873047
PSNR 24.782076560337416
SSIM 0.8781344763843051

Ukoliko pogledamo 1. parametar MSE vidimo da SRCNN daje najmanju vrednost, a bilinearna interpolacija najveću, to je upravo ono što smo i očekivali budući da što je veća vrednost ovog parametra to je manja rezolucija slike, dakle u našem primeru slika na koju je primenjena SRCNN ima najveću, a slika na koju je primenjena bilinearna interpolacija ima najmanju rezoluciju.

Ukoliko pogledamo 2. parametar PSNR vidimo da najveću vrednost daje primena SRCNN, a najmanju primena bilinearne interpolacije na sliku. To je takođe očekivano budući da što je veća vrednost PSNR (peak-to-signal-noise-ratio) to je manji šum.

Na kraju, ukoliko pogledamo 3. parametar SSIM vidimo da SRCNN daje najveću, a bilinearna interpolacija najmanju vrednost, to znači da je slika na koju je primenjena SRCNN mreža najsličnija referentnoj slici, a slika na koju je primenjena bilinearna interpolacija najmanje slična.

Dakle, mogli bismo zaključiti da najbolje rezultate daje SRCNN, zatim bikubična interpolacija i na kraju bilinearna