

TEHNIKE UMETNIČKOG MENJANJA SLIKE PRIMENOM DUBOKOG UČENJA**DEEP LEARNING TECHNIQUES OF ARTISTIC IMAGE MODIFICATION**Milan Keča, *Fakultet tehničkih nauka, Novi Sad***Oblast – ELEKTROTEHNIKA I RAČUNARSTVO**

Kratak sadržaj – U ovom radu su predstavljene metoda umetničkog menjanja stila slika. Metode se baziraju na upotrebi pretreniranih konvolucionih neuronskih mreža za prepoznavanje sadržaja i stila slika. Predstavljene su iterativni algoritmi, prenos stila i Deep Dream, kao i njihova unapređenja. Analizirano je ponašanje algoritama menjajući parametre algoritma i komentarisani su rezultati. Implementacija svih algoritama je izvršena u programskom jeziku Python i paketu Tensorflow.

Ključne reči: Algoritam prenosa stila, Deep Dream, konvolucione neuronske mreže

Abstract – This paper describes the methods of artistic image transformation. The methods rely on convolutional neural networks trained for image recognition to differentiate between image style and content. The paper describes iterative algorithms such as style transfer and Deep Dream, as well as improvements of these original algorithms. The behavior of these algorithms is analyzed by experimenting with different parameter configurations. All algorithms were implemented using the Python programming language and Tensorflow package.

Keywords: Style transfer, Deep Dream, convolutional neural networks

1. UVOD

U ovom radu izvršena je implementacija i poređenje različitih metoda generisanja sadržaja slika i prenos različitih vizuelnih atributa. Vizuelnim atributima se smatraju informacije poput boja, nijanse, teksture i drugi oblici. U radu su predstavljene metode prenosa stila sa jedne slike na drugu, kao i Deep Dream – algoritam koji generiše kreativne, “halucinogene”, efekte na slici iterativnim pojačavanjem šablona koje je mreža trenirana da prepoznaje.

Prikaz semantičkog sadržaja slike u različitim stilovima je veoma izazovan zadatak procesiranja slika. Najteži korak predstavlja eksplicitno razdvajanje semantičkih informacija slike od stila kojim je taj sadržaj predstavljen. Ovaj problem je olakšan upotrebom konvolucionih neuronskih mreža. Tehnika prenosa stila bazirana na konvolucionim neuronskim mrežama podrazumeva korišćenje različitih reprezentacija slike dobijenih iz konvolucionih neuronskih mreža optimizovanih za prepoznavanje objekata.

Na taj način se efikasno mogu izdvojiti primitivnije informacije slike, kao što su linija, nijansa, boja i apstraktnije informacije slike, kao što su objekti koji se nalaze na slici. Algoritam prenosa stila omogućuje razdvajanje informacija slike i njihovo ponovno sklapanje. Rezultat algoritma su slike na kojima se nalazi željeni sadržaj nacrtan željenim stilom [1].

Deep Dream je eksperiment koji vizualizuje šablone naučene od strane neuronske mreže. Slično detetu koje gleda u oblake i pokušava da interpretira nasumične oblike, Deep Dream interpretira i pojačava šablone koje pronalazi na slici. Algoritam je razvijen od strane inženjera Google-a koji su koristili pretreniranu Inception konvolucionu mrežu [2].

Za konstrukciju Deep Dream eksperimenata može se koristiti bilo koja konvoluciona neuronska mreža obučena na slikama. Umesto da eksplicitno tražimo šablone koje želimo mreža da interpretira i pojača, puštamo mrežu da to uradi sama izborom odgovarajućih konvolucionih slojeva.

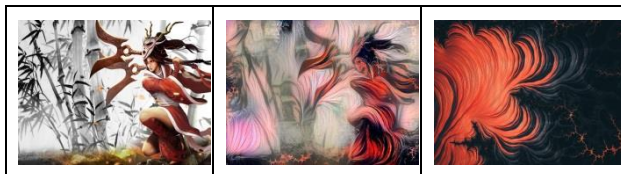
2. UPOTREBA ALGORITMA PRENOSA STILA

Napreci dubokih konvolucionih mreža su doveli do toga da je moguće izdvojiti apstraktnije, semantičke osobine slike. Razdvajanje sadržaja od stila na slici je neophodno za prenos stila sa jedne slike na drugu. U radu [1] je u ovu svrhu korišćena pretrenirana VGG-19 mreža [3], sa napomenom da je za ovaj zadatak moguće koristiti bilo koju konvolucionu neuronsku mrežu istreniranu za prepoznavanje objekata na slikama.

Na slici 1 prikazan je primer primene prenosa stila.

Slika 2. prikazuje međurezultate ovog iterativnog procesa, odnosno, slike generisane nakon određenog broja iteracija algoritma prenosa stila.

U ovom poglavlju ćemo analizirati kako na algoritam prenosa stila utiču faktori poput izbora algoritma za iterativnu optimizaciju slike (poglavlje 2.1) i izbora sloja konvolucione neuronske mreže za koji se pretpostavlja da efektivno reprezentuje stil kojim je slika nacrtana.



Slika 1. Primer prenosa stila. Iterativno se prenosi stil sa slike stila (desno), na sliku sadržaja (levo) i dobija se rezultujuća slika (sredina).

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bila dr Jelena Slivka, vanr. prof.



Slika 2. Prikaz iterativnih izmena slike prilikom prenosa stila

2.1. Izbor algoritma za iterativnu optimizaciju slike

U nastavku je predstavljeno nekoliko primera prenosa stila u kojima se koriste različiti algoritmi optimizacije za iterativno menjanje ciljne slike.

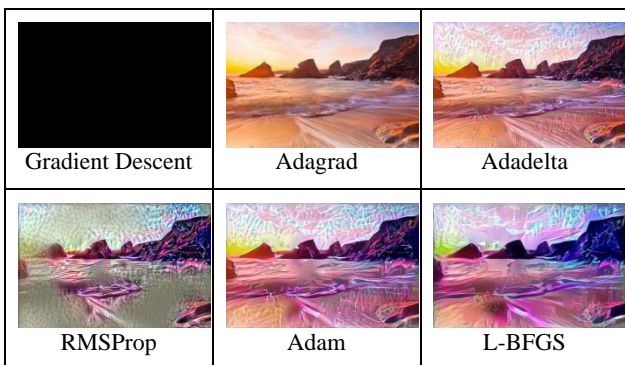
U eksperimentima izvršenim u ovom radu isprobani su sledeći algoritmi: *Gradient Descent*, *Adagrad*, *Adadelta*, *RMSProp*, *Adam* i *L-BFGS*.

Iako se za optimizaciju ciljne slike može koristiti bilo koji algoritam, u eksperimentima isprobanim u ovom radu su najbolje rezultate proizveli *L-BFGS* i *Adam* (slika 3). U prikazanom eksperimentu, svi parametri prenosa stila su fiksirani na iste vrednosti, a menja se samo optimizacioni algoritam. Na slici 3 su prikazane slike korišćene za sprovedene eksperimente.

U sprovedenim eksperimentima prikazanim na slici 4 pokazalo se da *Adam* i *L-BFGS* algoritmi daju najbolje rezultate u najmanjem broju iteracija. U ovom kontekstu, rezultati su najbolji u smislu da su šabloni slike stila u najvećoj prebačeni na sliku sadržaja, uz najmanji šum. Rezultati, naravno, zavise i od podešavanja specifičnih parametara svakog od algoritama optimizacije, kao i ostalih parametara algoritma za prenos stila, koji su u ovom eksperimentu bili fiksni.



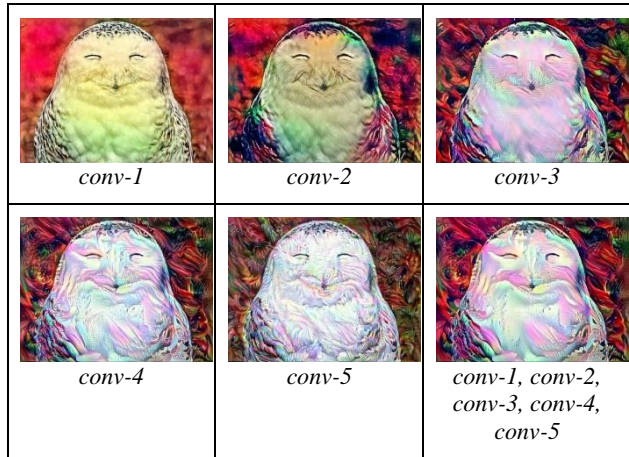
Slika 3. Prikaz slike sadržaja (levo) i slike stila (desno), korišćene za poređenje optimizujućih algoritama.



Slika 4. Poređenje različitih optimizera za prenos stila.

2.2. Izbor slojeva za reprezentaciju stila

U radu [1] su iz pretrenirane VGG-19 mreže korišćeni sloj *conv4_2* za reprezentaciju sadržaja i slojevi *conv1_1*, *conv2_1*, *conv3_1*, *conv4_1*, *conv5_1* za reprezentaciju stila. U sledećim eksperimentima će biti prikazani rezultati algoritma prenosa stila upotrebom različitih slojeva za reprezentaciju stila. Ostali parametri algoritma će biti isti za sve eksperimente.



Slika 5. Poređenje različitih reprezentacija stila (sloja konvolucione neuronske mreže) za prenos stila.

Iz postignutih rezultata (slika 5) se može zaključiti da se korišćenjem prvog sloja efektivno prenose samo boje, dok se teksture menjaju u veoma maloj meri. Upotrebom dubljih slojeva postiže se efekat detaljnijih tekstura, ali se gube informacije o boji slike stila. Korišćenjem svih pet slojeva postižu se detaljne teksture i očuvanje boje slike stila.

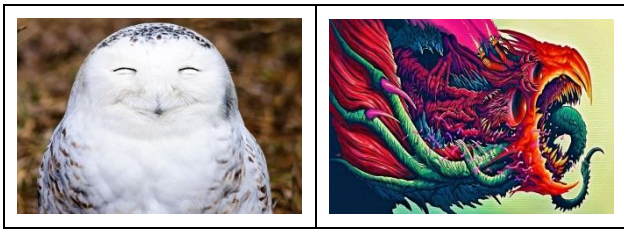
4. OČUVANJE ORIGINALNIH BOJA U PRENOSU STILA

Tehnika prenosa stila će neizbežno, pored tekstura, preneti i boje sa slike stila na sliku sadržaja. U određenim situacijama to može biti nepoželjno. U tom slučaju se kao ekstenzija originalnom algoritmu prenosa stila može koristiti tehnika prenosa boja sa jedne slike na drugu [4].

U ovom poglavlju su predstavljeni eksperimenti prenosa stila, uz očuvanje originalnih boja slike sadržaja. Očuvanje originalnih boja u prenosu stila je motivisano opažanjem da je vizuelna percepcija mnogo osetljivija na promene u osvetljenosti nego na promene u boji [4].

Modifikacija boja rezultujuće slike podrazumeva da se izračunaju kanali osvetljenosti (*luminance*) slike sadržaja L_C i slike stila L_S koristeći YIQ prostor boja. Zatim se primenjuje standardni algoritam prenosa stila kako bi se dobila rezultujuća slika i kanal osvetljenosti L_T . Informacije o boji slike sadržaja su predstavljene kanalima I i Q i kombinuju se sa kanalom Y rezultujuće slike kako bi se dobila krajnja slika [4].

Pored YIQ prostora boja, mogu se koristiti i drugi, kao što su YUV, LUV, Lab, YCrCb, jer na određen način opisuju osvetljenost na slici. Na slici 6 prikazane su ulazne slike sadržaja i stila, dok su na slici 7 prikazani rezultati prenosa stila bez očuvanja boja i sa očuvanjem boja.



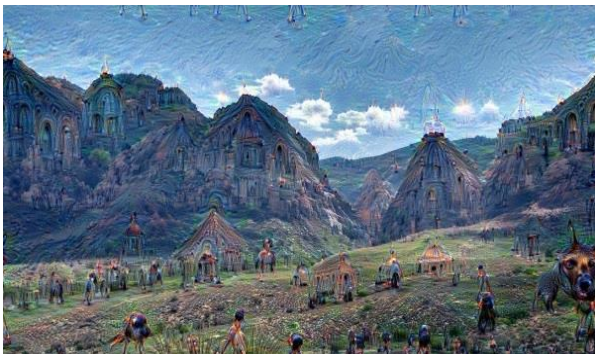
Slika 6. Slika sadržaja (levo) i slika stila (desno)



Slika 7. Rezultat prenosa stila bez očuvanja originalnih boja slike sadržaja (levo) i sa očuvanjem originalnih boja (desno)

5. UPOTREBA DEEP DREAM ALGORITMA

Deep Dream [2] je eksperiment koji vizualizuje šablone naučene od strane neuronske mreže. Naglašava i poboljšava prepoznate šablone na slici. To čini propuštajući sliku kroz mrežu, računajući gradijente aktivacija određenih slojeva i, na kraju, optimizujući sliku tako da se maksimizuju aktivacije istih slojeva. Maksimizacijom aktivacija izabranih slojeva slike, pojačavaju se šablone koje mreža prepoznaje. Na slici 8 prikazan je rezultat *Deep Dream* algoritma.



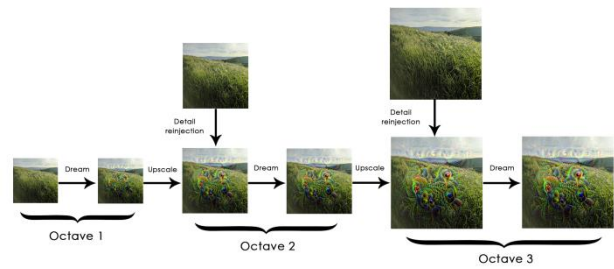
Slika 8. Rezultat *Deep Dream* algoritma [2].

Osnovni oblik *Deep Dream* algoritma [2] proizvodi zadovoljavajuće rezultate, ali postoji i nekoliko problema. Rezultujuća slika je puna šuma i rezolucija je niska. Pored toga, izgenerisani šablone se javljaju na istom nivou granularnosti [5].

To znači da svi dodatni šablone na slici imaju sličan nivo detalja i slične su boje i veličine. Navedeni problemi se mogu rešiti optimizacijom na različitim skalama slike i ubacivanjem izgubljenih detalja iz originalne slike.

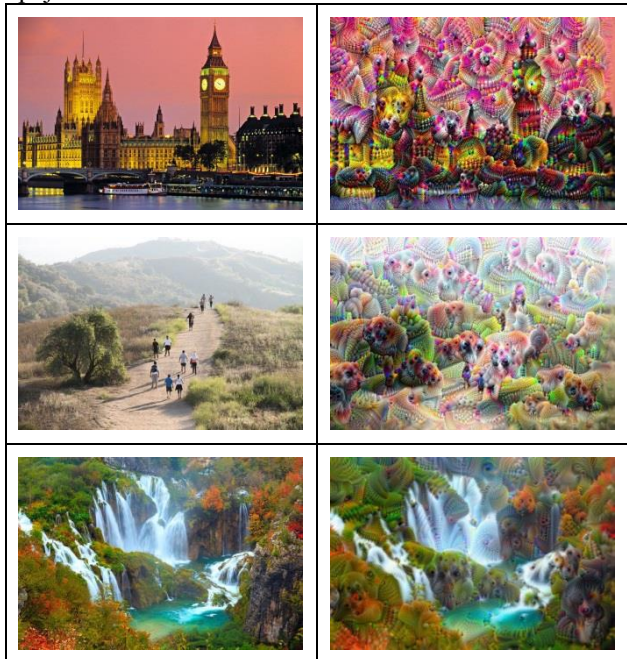
Svaka skala predstavlja određenu rezoluciju originalne slike. To omogućava da se šablone izgenerisani na nižim skalama budu inkorporirani u šablone na višim skalama i popunjeni dodatnim detaljima (slika 9).

To se postiže tako što se uradi *gradient ascent* za modifikaciju slike i zatim poveća veličina slike, što predstavlja jednu oktavu. Proces se ponavlja za određeni broj oktava [2].

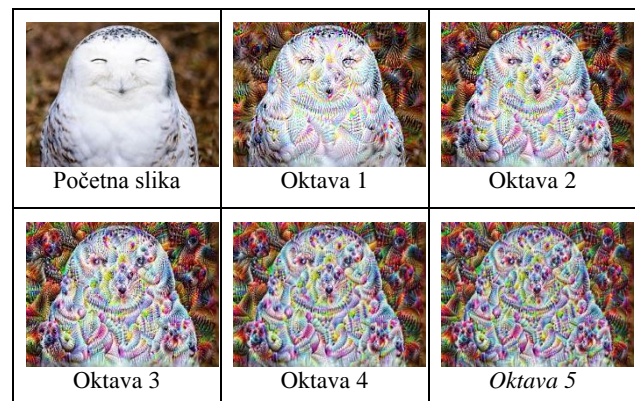


Slika 9. Oktave *Deep Dream* algoritma.

S obzirom da je *Inception V3* [6] mreža trenirana na slikama veličine 299 x 299, najbolji rezultati dobijaju se na slikama u opsegu veličine 300 x 300 ili 400 x 400, jer algoritam podrazumeva i skaliranje slike. Bez obzira na to, algoritam se može upotrebiti na slikama bilo koje veličine. Na rezultujućim slikama će se na kraju najviše videti šablone očiju pasa ili perja ptica, jer su slike sa takvim šablonima najzastupljenije u skupu slika na kojima je trenirana mreža. Na slici 10 prikazani su rezultati *Deep Dream* algoritma dobijeni koristeći 5 oktava. Na slici 11 prikazano je kako svaka oktava dalje pojačava šablone naučene od strane mreže.



Slika 10. Primeri upotreba algoritma *Deep Dream*



Slika 11. Prikaz rezultujuće slike nakon svake oktave

6. ZAKLJUČAK

U ovom radu predstavljene su metode za kreativno generisanje slika pomoću konvolucionih neuronskih mreža. Pored toga, predstavljeno je kako eksplicitno razdvajanje semantičkih informacija slike od stila kojim je taj sadržaj predstavljen omogućuje bolje razumevanje šablona koje neuronske mreže nauče prilikom treniranja, ali i otvara nov prostor za manipulisanje slikama. Motivacija za razvoj ovih algoritama je, s jedne strane, bila istražiti i bolje razumeti na koji način funkcionišu konvolucione neuronske mreže i šta mogu da nauče prilikom treniranja. S druge strane, pokazano je na koji način se mogu koristiti neuronske mreže za umetničke potrebe.

U radu su pokazani primeri slika generisani algoritmom prenosa stila [1] i *Deep Dream* algoritmom [2]. Oba algoritma se zasnivaju na upotrebi različitih slojeva konvolucionih mreža za dobijanje različitih reprezentacija slika. Svaka reprezentacija slike se koristi kako bi se uneli novi detalji u sliku. Prikazani su rezultati oba algoritma, kao i kako njihove različite konfiguracije utiču na rezultate.

Arhitekture mreža korišćene kod oba algoritma su konvolucione mreže trenirane na *Imagenet* [7] skupu podataka. Za algoritam prenosa stila, korišćena je VGG-19 [3] mreža, dok je za *Deep Dream* korišćena *Inception-V3* mreža [6]. Obe mreže mogu da prepoznaju različite nivoe kompleksnosti šablona slike, počevši od primitivnih (linija, tekstura i osnovnih geometrijskih oblika), do čitavih objekata na slici. Iz tog razloga su ove mreže veoma zahvalne za manipulisanje slikama.

Osnovni nedostatak ovog rada jeste korišćenje relativno skromnog broja primera uz relativno mali broj iteracija algoritma. Da bi se proverilo da li izvedeni zaključci važe u opštem slučaju, potrebno je testirati algoritme na većem broju slika uz veći broj iteracija. Nemogućnost takvih eksperimenata u ovom istraživanju proizilazi iz zahtevnosti samih algoritama, kako memorijskih, tako i procesorskih.

7. LITERATURA

- [1] L. A. Gatys, A. S. Ecker, M. Bethge, A Neural Algorithm of Artistic Style, 2015.
- [2] Google AI Blog. 2015. Inceptionism: Going Deeper Into Neural Networks. [online] Available at: <<https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>>
- [3] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2015.
- [4] L. A. Gatys, M. Berthge, A. Hertzmann, E. Shechtman, Preserving Color in Neural Artistic Style Transfer, 2016.
- [5] TensorFlow. 2020. Deepdream | Tensorflow Core. [online] Available at: <<https://www.tensorflow.org/tutorials/generative/deepdream>>
- [6] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Rethinking the Inception Architecture for Computer Vision, 2015.
- [7] Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) ImageNet Large Scale Visual Recognition Challenge. IJCV, 2015.

Kratka biografija:



Milan Keča rođen je u Subotici 1993. god. Master rad na Fakultetu tehničkih nauka iz oblasti Elektrotehnike i računarstva – Tehnike umetničkog menjanja slike odbranio je 2020.god. kontakt: vonum.mk@gmail.com