

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

# **Suparnički generativni modeli za prevođenje slika**

*Krešimir Vukić*

Voditelj: *prof. dr. sc. Siniša Šegvić*

Zagreb, travanj 2018.

# SADRŽAJ

<b>1. Uvod</b>	<b>1</b>
<b>2. Duboke mreže</b>	<b>2</b>
2.1. Konvolucija . . . . .	2
2.2. Primjer konvolucijske mreže . . . . .	3
<b>3. Generativni suparnički modeli (GAN-ovi)</b>	<b>4</b>
3.1. Generativni naspram diskriminativnih algoritama . . . . .	4
3.2. Rad GAN-a . . . . .	5
<b>4. Uvjetne generativne mreže</b>	<b>6</b>
4.1. Primjena . . . . .	6
<b>5. Translacija bez uparenih primjera za učenje pomoću kružnih GAN-ova</b>	<b>7</b>
5.1. Prednosti . . . . .	7
5.2. Rad CycleGAN-a . . . . .	7
5.3. Primjena . . . . .	7
<b>6. Prijenos stila</b>	<b>8</b>
<b>7. Zaključak</b>	<b>9</b>

# 1. Uvod

Generiranje slika na temelju skice, prijenos stilova i rekonstrukcija fotografije samo su neki od problema računalnog vida. Pri njihovom rješavanju uglavnom želimo naučiti funkciju mapiranja između parova slika koju zatim koristimo za efikasnu identifikaciju značajki ili po potrebi, njihovo generiranje.

Među najkorištenijim i najuspješnijim arhitekturama današnjice nalaze se generativni suparnički modeli. Ovim radom istražiti će se različita proširenja kojima se ti modeli prilagođavaju specifičnim problemima, njihove karakteristike i primjene.

## 2. Duboke mreže

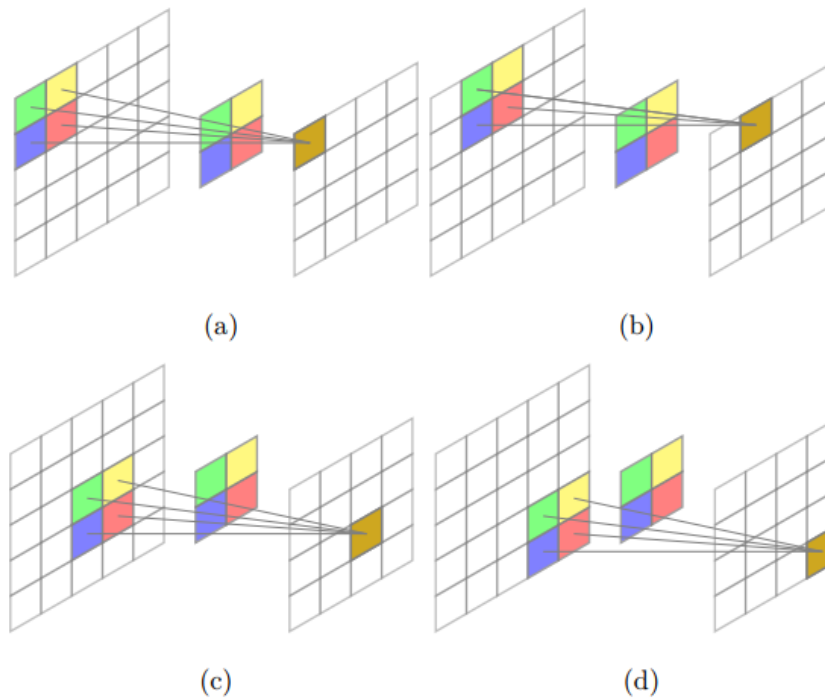
Duboke neuronske mreže zadnjih godina izuzetno su popularne u rješavanju klasifikacijskih problema i pronalaženju uzoraka. Tome je zaslužna prednost koju imaju naspram klasičnih pristupa, a koja se očituje u vidu mogućnosti da samostalno pronalaze bitne značajke u podacima te na temelju njih i tražena rješenja.

Slojevit pristupu rastavljanja podataka na jednostavnije građevne jedinice čini ih posebice pogodnima za radu sa problemima hijerarhijske prirode poput slika.

### 2.1. Konvolucija

Konvolucijske mreže osmišljene su s ciljem dohvata lokalnih relacija među podacima koje se posebice očituju pri radu sa slikama. Naime, ukoliko želimo analizirati malenu značajku na slici, nepoželjno je da njena pozadina ili lokacija unutar slike utječe na nju samu.

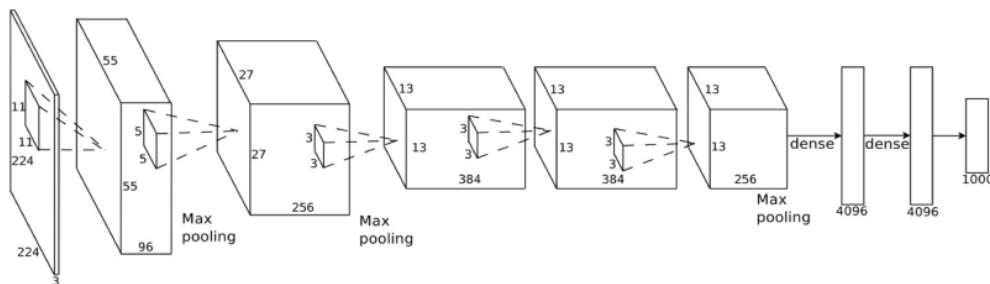
Ideja konvolucije je da se kao ulaz u neurone sljedećeg sloja dovede više vrijednosti s prethodnog što je prikazano slikom 2.1. To se odvija tako da se unaprijed definiran kvadrat pomiče (konvoluirá) po ulaznim podacima i tako određuje mapiranje na sljedeći sloj. Rezultat tog procesa je da pridodajemo veću važnost pikselima u neposrednoj okolini istraživanog područja.



**Slika 2.1:** Proces konvolucije

## 2.2. Primjer konvolucijske mreže

Nizanjem konvolucijskih slojeva dobivamo arhitekturu u kojoj je svaki sloj zadužen za jednostavniji potproblem sljedećeg sloja. Na slici 2.2 prikazana je duboka neuronska mreža s 5 konvolucijskih slojeva među kojima se odvija sažimanje maksimalnom vrijednošću te dva sloja potpune povezanosti.



**Slika 2.2:** Primjer duboke neuronske mreže

## 3. Generativni suparnički modeli (GAN-ovi)

GAN-ovi su podvrsta dubokih mreža čija arhitektura se bazira na dvije podmreže koje uče kroz međusobno natjecanje. Tako nespecifični model, ovisno o implementaciji, u stanju je kopirati distribuciju širokog spektra podataka: slika, govora, muzike, čak i pisanih zapisa.

### 3.1. Generativni naspram diskriminativnih algoritama

Dvije temeljne skupine kojima dijelimo algoritme za učenje su diskriminativni i generativni. Njihovi nazivi dolaze izravno iz vrste problema koji rješavaju:

Diskriminativni modeli direktno mapiraju ulazne varijable na ciljne vrijednosti i time uče uvjetnu vjerojatnost  $p(Y|X)$ . Dobiveni model svrstava podatke u ciljne klase s određenom vjerojatnošću.

Generativni modeli, kao što im i naziv govori, generativne su prirode. Problem koji rješavaju je kako dobiti  $X$  uz dani  $Y$ , tj. pokušavaju naučiti distribuciju pojedinih klasa  $p(X | Y)$ .

Drugim riječima, dok diskriminativni modeli pokušavaju pronaći granicu između klasa, generativni modeliraju distribucije pojedinih klasa na temelju kojih mogu uzorkovati nove primjere.

## 3.2. Rad GAN-a

Pola mreže bavi se stvaranjem novih podataka dok druga polovica pokušava evaluirati koliko su slični stvarnim podacima. Cijeli proces usporedljiv je s poslovima detektiva i kriminalca. Jedna strana u bavi se falsifikacijom dokaza, a druga analizom njihove autentičnosti.

Koraci rada GANa:

1. generator iz nasumičnih brojeva stvara sliku
2. ta slika se dovodi na ulaz diskriminatora uz skup stvarnih slika iz podatkovnog skupa
3. diskriminator određuje postotak vjerojatnosti da je slika prava

## 4. Uvjetne generativne mreže

Uvjetne generativne mreže pokazale su se kao generalno rješenje problema prevođenja slike . Svojstvene su po učenju ne samo funkcije prijenosa između slika već i funkcije gubitka nad kojom se ona trenira. Dok se u drugim modelima ona ručno zadaje, pristup u kojem se ona uči omogućuje generalizaciju nad setovima problema koji bi u suprotnom zahtijevali vrlo drugačije formulacije gubitka.

### 4.1. Primjena

Neke od mogućnosti uvjetnih GAN-ova su sinteza fotorealističnih slika iz skica, generiranje objekata na temelju njihovih obruba te bojanje crno bijelih fotografija.



## 5. Translacija bez uparenih primjera za učenje pomoću kružnih GAN-ova

### 5.1. Prednosti

Većina postojećih algoritama korištenih u računalnom vidu ovisi o postojanju velikih podatkovnih setova označenih parova slika. No, za mnoge probleme takvi podaci nisu dostupni, a u takvim situacijama primjenu nalaze kružni GANovi.

### 5.2. Rad CycleGAN-a

Temelj njihovog rada je uz da uz učenje funkcije  $G : X \rightarrow Y$ , istovremeno uče njen inverz  $F : Y \rightarrow X$  pomoću gubitka cikličke dosljednosti:  $F(G(X)) \approx X$  i obrnuto.

### 5.3. Primjena

## 6. Prijenos stila

## 7. Zaključak

## **Suparnički generativni modeli za prevođenje slika**

### **Sažetak**

Mnogi problemi u procesiranju slika odnose se na translaciju ulazne slike u željenu izlaznu. Većina današnjih arhitektura bazira se na generativnim suparničkim mrežama, nad kojima su osmišljena mnoga proširenja kako bi se prilagodila specifičnim problemima. Uvjetnim generativnim mrežama nije potrebno definirati funkciju gubitka jer ju mogu same naučiti, kružne generativne mreže sposobne su učenju translacije ne samo ulaznih u izlazne primjere nego i u drugom smjeru, a ograničavanjem mreže na translaciju unutar prostora boja obavljat će prijenos stila bez distorzije sadržaja slike. Ovim radom usporedit će različite arhitekture, njihove karakteristike i primjene na problemima translacije slike u sliku.

**Ključne riječi:** image-to-image generation, style transfer, GANs, CycleGANs

## **Image to image generation using CycleGANs**

### **Abstract**

**Keywords:** image-to-image generation, style transfer, GANs, CycleGANs