

# Machine Learning for Font Classification and Generation

---

Nemanja Mićović - Faculty of Mathematics, University of Belgrade

31.10.2018.



**MLA**  
M A T F

# Uvod

---

- Uvod u tipografiju
- Klasifikacija simbola
  - Pristupi pre dubokog učenja
  - Rad DeepFont koji koristi konvolutivne mreže
  - Deo master rada Fontool - alat za klasifikaciju i generisanje fontova
- Generisanje simbola i fontova
  - Rad koji koristi generativne suparničke mreže (eng. *GAN*)
  - Rad koji uči tipografski stil
  - Deo master rada Fontool - alat za klasifikaciju i generisanje fontova

# Tipografija

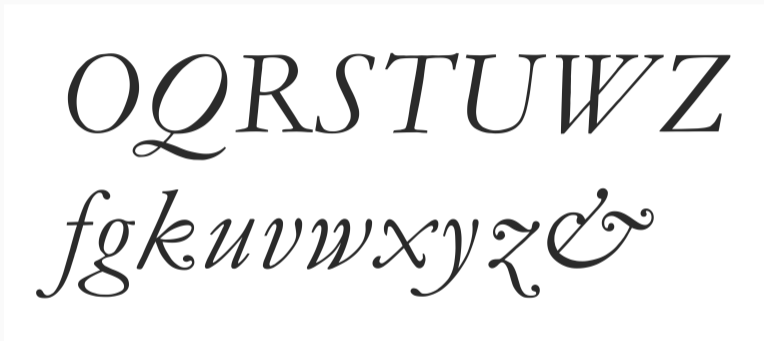
---

## Bavi se:

- grafičkim uređenjem teksta, slova i strane
- vizuelnom percepcijom tekstuelnog sadržaja
- dizajnom i razvojem fontova

A B C D E F G H I  
K L M N O P Q R  
S T U V W X Y Z &  
a b c d e f g h i j k l m n o p q r  
s t u v w x y z.

Slika 1: Prvi font - Black Letter.



**Slika 2:** Prvi font koji je imao kurziv (eng. *italic*).



fi → fi  
fl → fl

**Slika 3:** Reč levo - bez ligature i reč desno - sa ligaturom.



**Slika 4:** Slovo levo poseduje serife dok slovo desno ne.



**Slika 5:** Donji font predstavlja font sa jednakim razmakom. Gornja reč zauzima isto prostora koliko i donja uprkos tome što poseduje 3 slova više.

# Tipografija - vizuelna percepcija sadržaja

## Emily Prosser

E-Mail: emily.prosser@gmail.com  
Tel: 215-837-0553

Address: 542 Byers Lane Philadelphia, PA, 19019

### Academic Details

09.2007 – 07.2008 - University of Pennsylvania, Philadelphia, CA

- MA in Social Communications
- Specialization in PR and Media Marketing

### Languages

- Chinese (Intermediate)
- Spanish (bilingual)

### Work Experience

**Date of Joining:** 2013-03 to now

**Post:** Senior PR Specialist

**Organization:** Berkadia, Ambler, PA

**The scope of work:** Developed and implemented policies and procedures for overall public relations effort in cooperation with sales and customer service teams. Supervised the external communications team in securing leads in top media outlets in the US and in international media. Supervised the preparation and distribution of materials such as news releases, fact sheets, and scripts to media outlets. Reviewed all forms of communication prior release.

**Date of Joining:** 2009-06 to 2013-03

**Post:** PR Specialist

**Organization:** Adidas, Boston, MA

**The scope of work:** Created and maintained lists of media contacts. Researched opportunities across online media channels. Produced product pitches and press kits. Responded to media inquiries and supported event organization. Lead a project team in designing and implementing a comprehensive social media relations strategy for a new line of lifestyle products.

**Date of Joining:** 2008-08 to 2009-06

**Post:** Junior Communications Specialist

**Organization:** SAP, Newton Square, PA

**The scope of work:** During industry events, distributed notes, press credentials, set up press boxes, recorded and transcribed press conferences. Helped keep PR storage areas organized and the resources easily accessible, assisted in office administration activities. Assisted with producing press releases, organizing press conferences, publicity events, and off-day media availabilities. Performed minor translating tasks, including translating press releases into Spanish and assisting live translations into Spanish.

### Hard and Soft Skills

- Press releases
- Media relations
- Crisis management
- Facebook marketing
- Creative writing

## Emily Prosser

PR Specialist

Phone 215 837 0553

WWW emily-insights.com

E-mail emily.prosser@gmail.com

LinkedIn in/wes/in/emilyprosser

Senior PR specialist with 9+ years of experience in marketing and corporate public relations. Managed multiple PR and marketing annual budgets from \$5,000 to \$300,000. Helped increase clients' annual sales by up to 40%.

### Experience

2013-03 - present Senior PR Specialist

Berkadia, Ambler, PA

- Developed and implemented policies and procedures for overall public relations effort in cooperation with sales and customer service teams.
  - Supervised the external communications team in securing leads in top media outlets in the US and in international media.
  - Supervised the preparation and distribution of materials such as news releases, fact sheets, and scripts to media outlets.
  - Reviewed all forms of communication prior release.
- Key achievement:** 100% efficiency in meeting the external communications department's projects' milestones and goals.

2009-06 - 2013-03

PR Specialist

Adidas, Boston, MA

- Created and maintained lists of media contacts.
- Researched opportunities across online media channels.
- Produced product pitches and press kits.
- Responded to media inquiries and supported event organization

**Key achievement:** Lead a project team in designing and implementing a comprehensive social media relations strategy for a new line of lifestyle products - grew Facebook fan base from 0 to 12,000 in 4 months, gained 30,000 Instagram followers in 3 months.

2008-08 - 2009-06

Junior Communications Specialist

SAP, Newton Square, PA

- During industry events, distributed notes, press credentials, set up press boxes, recorded and transcribed press conferences.
- Helped keep PR storage areas organized and the resources easily accessible, assisted in office administration activities.
- Assisted with producing press releases, organizing press conferences, publicity events, and off-day media availabilities.

### Education

2007-09 - 2008-07

MA in Social Communications, specialization in PR and Media Marketing

University of Pennsylvania, Philadelphia, CA

### Languages

Spanish

Bilingual

Chinese

Intermediate

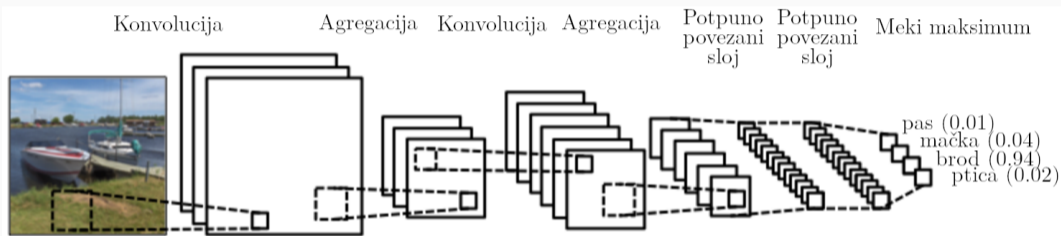
- Vizuelno prepoznavanje fontova (eng. *visual font recognition - VFR*)
- Vizuelno generisanje fontova (eng. *visual font generation - VFG*)
- Vizuelizacija fontova

## Kratak osvrt na neuronske mreže

---

- Vodeći pristup za klasifikaciju slika
- Postigle značajne rezultate na problemima u oblasti računarske vizije (eng. *computer vision*)
- Na skupu ImageNet ostavarile manju grešku u odnosu na čoveka
- Daju delimičnu interpretabilnost za svoj rad (Zeiler and Fergus 2013)

# Konvolutivne neuronske mreže

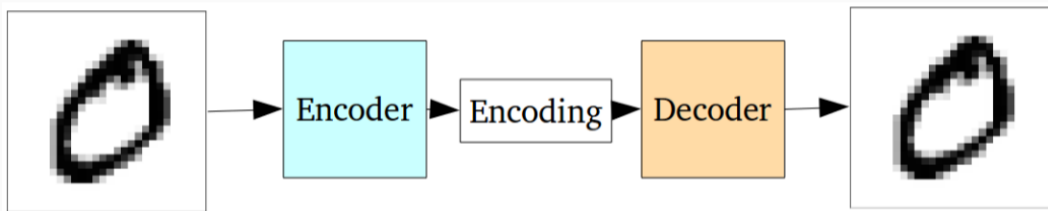


**Slika 6:** Primer arhitekture konvolutivne mreže.



- Autoenkoder je neuronska mreža koja ima cilj da nauči kompresovanu reprezentaciju podataka sa ulaza
- Sastoji iz dva dela, *enkoder* i *dekoder*
- Variacioni autoenkoder dodatno uči raspodelu verovatnoće podataka
- Usled toga, moguće je vršiti generisanje podataka

# Autoenkoderi



# Vizuelno prepoznavanje fontova

---

# Vizuelno prepoznavanje fontova (eng. *visual font recognition - VFR*)

## Postavka problema

- Potrebno na slici odrediti koji je font korišćen za kreiranje tekstualnog sadržaja

## VFRa

- Slike koje se prepoznaju su slike nastale na računaru

## VFRb

- Slike koje se prepoznaju su fotografije iz realnog sveta

## Motivacija

- Otkrivanje nepoznatih fontova (Wang et al. 2015; Ramanathan et al. 2009)
- Olakšavanje u procesu dizajna sadržaja





## Raniji pristupi

- Statistička analiza fontova
- Analiza lokalnih delova slova i teksture
- Konstruisanje atributa
- Radovi (Avilés-Cruz et al. 2005; Jung, Shin, and Srihari 1999; Ma and Doermann 2003; Ramanathan et al. 2009)
- Problemi: rad na malom skupu podataka i mali broj dozvoljenih fontova (desetine)

# DeepFont

---



- Rad objavljen 2015. godine (Wang et al. 2015)
- Korišćenje konvolutivnih neuronskih mreža i konvolutivnih autoenkodera
- Sistem razvijen od strane kompanije Adobe
- Podržava preko 2000 fontova (2383)
- Integrisan u alat Adobe Photoshop
- Rešava problem VFRb



Slika 8: Nekoliko slika iz skupa podataka za klasifikaciju fontova.



**Slika 9:** Sintentičke i realne slike za isti font



**Slika 10:** Nekoliko primera iz skupa podataka.

- Skup podataka za VFRb je teško konstruisati
- Nije problem doći do fotografija, ali jeste do informacije o fontu koji je korišćen
- Skup VFRWild325 sadrži 325 slika i 93 fonta (Chen et al. 2014)
- Rad (Chen et al. 2014) odabira 2420 fonta i generiše oko 2.42 miliona instanci
- Adobe izbacuje nekoliko fontova te se broj smanjuje na 2383
- Adobe pravi AdobeVFR skup koji se sastoji iz:
  - VFR\_real\_u: poseduje 197396 nelabeliranih realnih slika
  - VFR\_real\_test: poseduje 4384 labeliranih realnih slika koji pokrivaju 617/2383 fonta
  - VFR\_syn\_train: labelirane sintetičke slike za obučavanje (2 383 000 instanci)
  - VFR\_syn\_val: labelirane sintetičke slike za validaciju (238 300 instanci)

Dataset name	Source	Label?	Purpose	Size	Class
VFRWild325	Real	Y	Test	325	93
VFR_real_test	Real	Y	Test	4, 384	617
VFR_real_u	Real	N	Train	197, 396	/
VFR_syn_train	Syn	Y	Train	2,383, 000	2, 383
VFR_syn_val	Syn	Y	Test	238, 300	2, 383

**Slika 11:** Pregled skupa podataka 'AdobeVFR'.

## Koristi se nekoliko transformacija

1. Šum
2. Zamućivanje
3. Rotacija perspektive
4. Promena boje pozadine
5. Modifikacija razmaka između karaktera
6. Modifikacija razmere širine i visine slike

RESTAURANT RESTAURANT

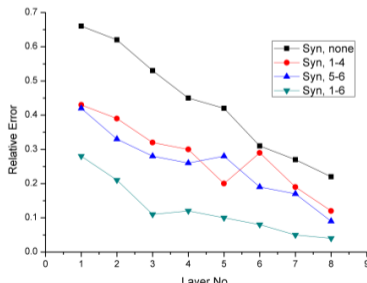
(a) Synthetic, none

(c) Synthetic, 5-6

RESTAURANT RESTAURANT

(b) Synthetic, 1-4

(d) Synthetic, 1-6

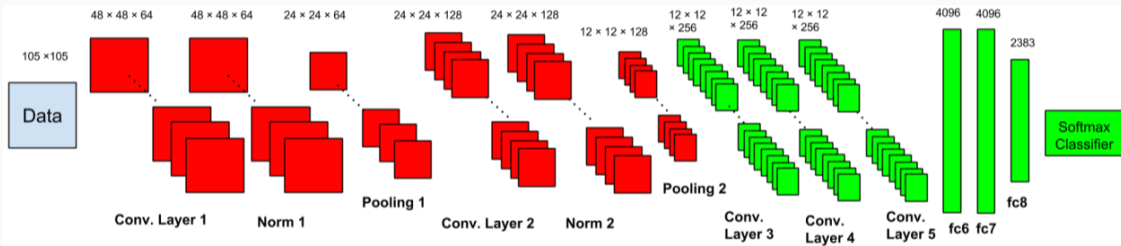


## DeepFont - Kako obuhvatiti sintentičke i realne slike?

- Javlja se različitost domena (eng. *domain mismatch*)
  - Sintentičke slike
  - Slike iz realnog sveta
- Pristup proširuje originalni pristup iz rada (Glorot, Bordes, and Bengio 2011)
- Klasifikator se deli u dva dela i  $N$  slojeva:
  - Mreža koja ulaz preslikava u zajednički domen sa  $K$  slojeva -  $C_u$
  - Mreža koja vrši klasifikaciju za ulaz iz zajedničkog domena sa  $N - K$  slojeva -  $C_s$

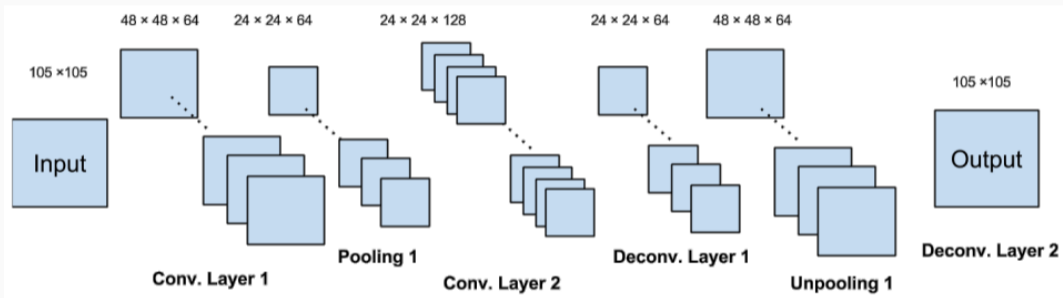


# DeepFont - Predložena arhitektura



**Slika 12:** Mreža  $C_U$  je označena crvenom bojom, a mreža  $C_S$  zelenom,  $N = 8$ ,  $K = 2$ .

- Predstavlja **enkoder** od konvolutivnog autoenkodera
- Vrší nelinearno smanjivanje dimenzionalnosti podataka
- Obučava se nezavisno od ostatka mreže
- Koristi se **srednjekvadratna greška** (eng. *mean squared error - MSE*)
- Krajnji cilj je preslikati realne i sintetičke podatke u isti domen



Slika 13: Arhitektura za konvolutivni autoenkoder.

- Obučavanje se vrši u dve faze
  - Obučava se autoenkoder
  - Nakon obučavanja autoenkodera, odgovarajući slojevi (koji čine  $C_u$ ) se ubacuju u kompletnu mrežu  $C$
  - Mreža  $C$  se obučava nadgledano, ali sa fiksiranim težinama za  $C_u$
- fc6 i fc7 koriste regularizaciju izostavljanjem (eng. *dropout*)
- Veličina podskupa za obučavanje (eng. *batch size*): 128
- Parametar učenja za  $C_s$  inicijalno 0.01 i deli se sa 10 kada greška na validacionom skupu prestane da opada

**Neka je data slika  $X$  dimenzija  $a \times b$**

- Skalira se na  $105 \times c$  tako da se očuva odnos širine i visina
- Potom se prave tri nove slike tako što im se širina menja slučajnom promenljivom iz raspodele  $U[1.5, 3.5]$
- Za tri dobijene slike, uzima se 5 podslika dimenzija  $105 \times 105$  na nasumično odabranim pozicijama
- Time se dobija 15 slika
- Predikcija se vrši uprosečavanjem 15 dobijenih vektora mekog maksimuma (eng. *softmax*) na izlazu klasifikatora

- Prikazana arhitektura klasifikatora poseduje veliki broj parametara
- Smanjenje broja parametara i zadržavanje kvaliteta je od velikog značaja u praksi
- Za konvolutivne mreže, najveći deo memorije zauzimaju gusto povezani slojevi
- Sloj fc6 zauzima oko 85% memorije pri skladištenju
- Parametri mreže se mogu zapisati matrično

## Singularna dekompozicija

- Neka je  $W \in R^{m \times n}$  matrica parametara modela (ili matrica težina za sloj fc6)
- Koristeći singularnu dekompoziciju (eng. *SVD*) važi:

$$W = U\Sigma V^T$$

gde su:

- $U \in R^{m \times n}$  i  $V \in R^{n \times n}$  ortogonalne matrice
- $\Sigma \in R^{n \times n}$  dijagonalna matrica
- $\sigma_1, \dots, \sigma_n$  singularne vrednosti (elementi na dijagonali matrice  $\Sigma$ )

## Signularna dekompozicija - aproksimacija matrica

- Ukoliko se uzme:
  - prvih  $k$  levih singularnih vektora u matrici  $U$  - matrica  $\tilde{U}$
  - prvih  $k$  desnih singularnih vektora u matrici  $V$  - matrica  $\tilde{V}$
  - prvih  $k$  sopstvenih vektora u matrici  $\Sigma$  - matrica  $\tilde{\Sigma}$

aproksimacija matrice  $W$  se može dobiti kao:

$$W \approx \tilde{W} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$$

- Odabir parametra  $k$  kontroliše koliko se gubi na kvalitetu aproksimacije
- Matrica  $\tilde{W}$  je aproksimacija matrice  $W$  ranga  $k$



- Jedan od pristupa bi bio upravo aproksimacija matrice  $W$  nakon obučavanja
- Postoje i drugi načini za kompresiju modela
- Oblast istraživanja kompresije modela je vrlo bitna za praktične primene
- Kompresija modela je vrlo značajna za rad na uređajima sa ograničenim hardverom

Methods	Training Data	Train	Test	
			N	R
SCAE N	<b>N</b> : VFR_syn_train, no data augmentation	0.02	3.54	31.28
SCAE S	<b>S</b> : VFR_syn_train, standard augmentation 1-4	0.21	2.24	19.34
SCAE F	<b>F</b> : VFR_syn_train, full augmentation 1-6	1.20	1.67	15.26
SCAE R	<b>R</b> : VFR_real_u, real unlabeled dataset	9.64	5.73	10.87
SCAE FR	<b>FR</b> : Combination of data from <b>F</b> and <b>R</b>	6.52	2.02	14.01

**Slika 14:** Poređenje greške (%) na skupovima za obučavanje i testiranje ( $K = 2$ )

- $N$  označava skup sintetičkih podataka
- $R$  označava skup realnih podataka



(a) original (b) SCAE N (c) SCAE S



(d) SCAE F (e) SCAE R (f) SCAE FR

**Slika 15:** Ilustracija rekonstrukcije realne slike sa autoenkoderom

$K$	0	1	2	3	4	5
Train	8.46	9.88	11.23	12.54	15.21	17.88
VFR_real_test	20.72	20.31	18.21	18.96	22.52	25.97



(a)  $K=1$

(b)  $K=2$

(c)  $K=4$

(d)  $K=5$

**Slika 16:** Kvalitet modela pri variranju parametra  $K$ .

## DeepFont - rezultati

- F - svih 6 koraka pretprocesiranja na VFR\_syn\_train
- R - VFR\_real\_u - nelabelirani skup realnih slika
- FR - kombinacija podataka iz F i R

Methods	Training Data		Training Error	VFR_syn_val		VFRWild325		VFR_real_test	
	C <sub>u</sub>	C <sub>s</sub>		Top-1	Top-5	Top-1	Top-5	Top-1	Top-5
LFE	/	/	/	26.50	6.55	44.13	30.25	57.44	32.69
DeepFont S	/	F	0.84	1.03	0	64.60	57.23	57.51	50.76
DeepFont F	/	F	8.46	7.40	0	43.10	22.47	33.30	20.72
DeepFont CAE_FR	FR	F	11.23	6.58	0	<b>38.15</b>	<b>20.62</b>	<b>28.58</b>	<b>18.21</b>
DeepFont CAE_R	R	F	13.67	8.21	1.26	44.62	29.23	39.46	27.33

**Slika 17:** Rezultati evaluacije.

	fc6 size	Total size	Ratio	<b>Method</b>	Error
default	150,994,944	177,546,176	NA	NA	18.21
$k=5$	204,805	26,756,037	6.64	Lossy	20.67
				Lossless	19.23
$k=10$	409,610	26,960,842	6.59	Lossy	19.25
				Lossless	18.87
$k=50$	2,048,050	28,599,282	6.21	Lossy	19.04
				Lossless	18.67
$k=100$	4,096,100	30,647,332	5.79	Lossy	18.68
				Lossless	18.21

**Slika 18:** Rezultati kompresije mreže.

# Alat Fontool

---

## Predloženi pristup problemu

- Inspirisano radom (Wang et al. 2015)
- Rešavan problem VFRA
- Cilj: razviti sličan sistem za alat Gimp
- Problem sveden na klasifikaciju slika



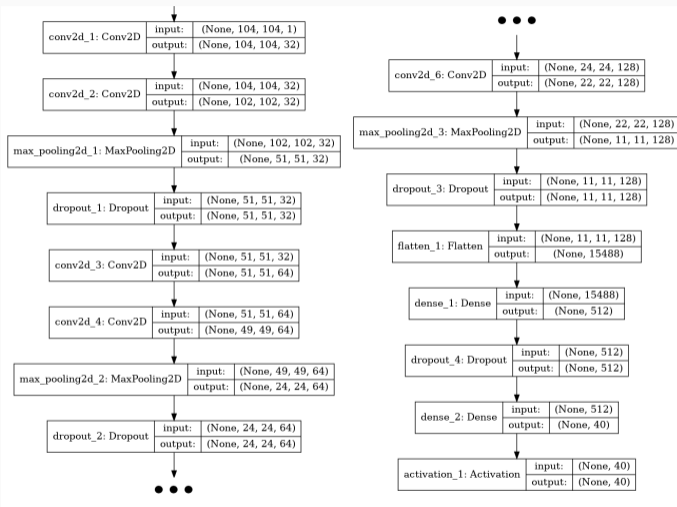
- Razvijen skup podataka f<sub>t</sub>W (eng. *Font Tool Words*)
- Odabrano:
  - 1000 čestih reči engleskog jezika
  - 40 raznovrsnih fontova
- Ukupno 120 000 instanci
- Instanca poseduje informaciju o ligaturama fontova

## Predloženi pristup problemu - uзорak iz skupa podataka



- Razvijen klasifikator razvijen na konvolutivnim neuronskim mrežama
- Korišćene biblioteke Keras i TensorFlow
- Obučavan na grafičkoj karti NVIDIA GEFORCE 1060 GTX 6GB
- Ostvarena preciznost od 0.9993

# Predloženi pristup problemu - arhitektura



# Predloženi pristup problemu - demonstracija

fontool [Home](#) [Font explorer](#) [Classify font](#) [Generate fonts](#)

## Font classifier



Choose File linearstd\_stolica.png

Submit

## Problem VFRa

- Obučavanje nad drastično većim skupom podataka koji ima 18000 fontova

## Problem VFRb

- Generisanje sintetičkih realnih fotografija
- Moguće učiniti koristeći radni okvir za razvoj video igara
- Omogućava se varijacija ugla kamere, osvetljenja, pozadine. . .

## Cilj

- Integracija razvijenog sistema u alat Gimp

# Vizuelno generisanje fontova

---

# Vizuelno generisanje fontova (eng. *visual font generation - VFG*)

## Postavka problema

- Potrebno izvršiti automatsko generisanje jednog ili više simbola za font bez informacije o tome kako izgledaju ostali simboli

## Motivacija

- Generisanje novih simbola za font
- Generisanje novih fontova
- Kreiranje nacрта i ideja pri kreiranju fontova
- Podrška grafičkom tipografu ili dizajneru



## Rad (Campbell and Kautz 2014)

- Simbole fonta posmatra kao Bezierove krive
- Traži površ manjih dimenzija na kojima leže dostupni fontovi

## Rad (Radford, Metz, and Chintala 2015a)

- Koristi duboko učenje i generativne suparničke mreže
- Svodi problem na problem generisanja slika

## Rad (Baluja 2016)

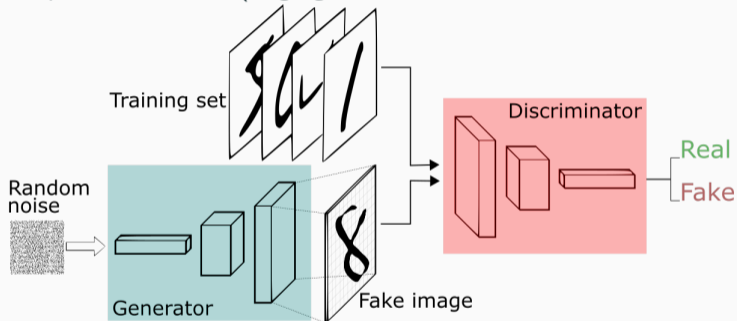
- Koristi duboko učenje i generativne suparničke mreže
- Generiše *nedostajuće* simbole za font

# Generisanje simbola koristeći GAN

---

# Generativne suparničke mreže - GAN

- Generativne suparničke mreže (eng. *generative adversarial networks - GAN*)

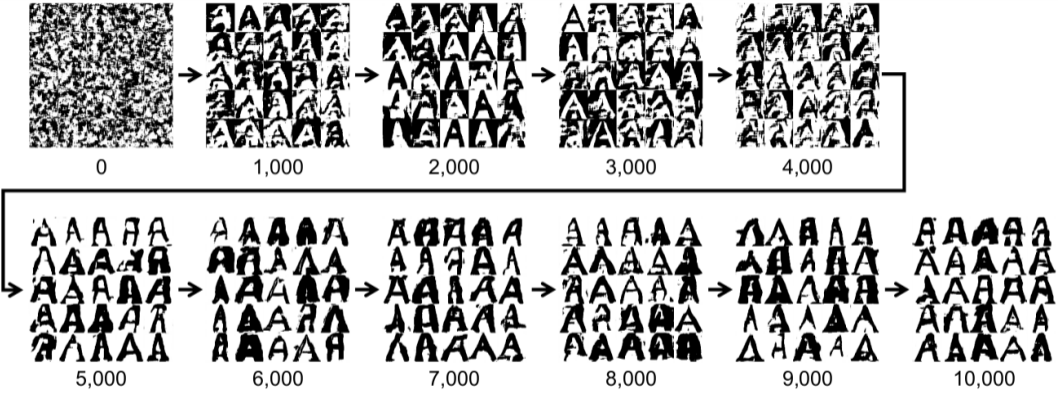


**Slika 19:** Arhitektura za GAN.

## DCGAN (Radford, Metz, and Chintala 2015a)

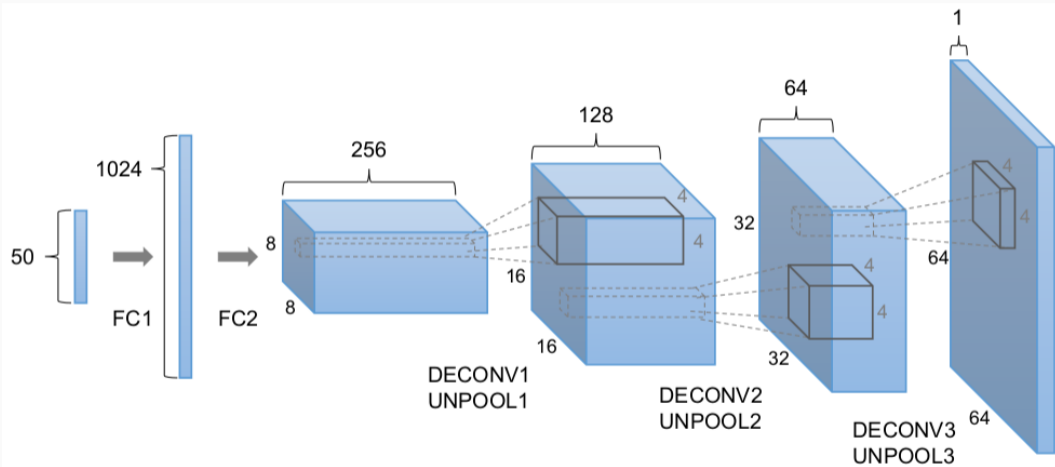
- Koristi konvolutivne GAN-ove
- Proširuje ih trećom mrežom koja predstavlja klasifikator simbola (slova)
- Uspeva da generiše slova zadovoljavajućeg kvaliteta

# DCGAN



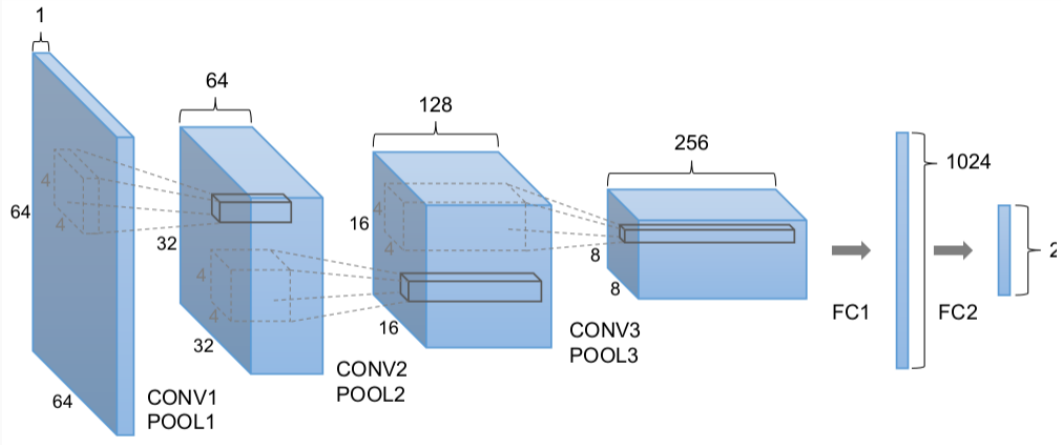
Slika 20: Generisana slova.

# DCGAN - generator $G$



Slika 21: Arhitektura za generator  $G$ .

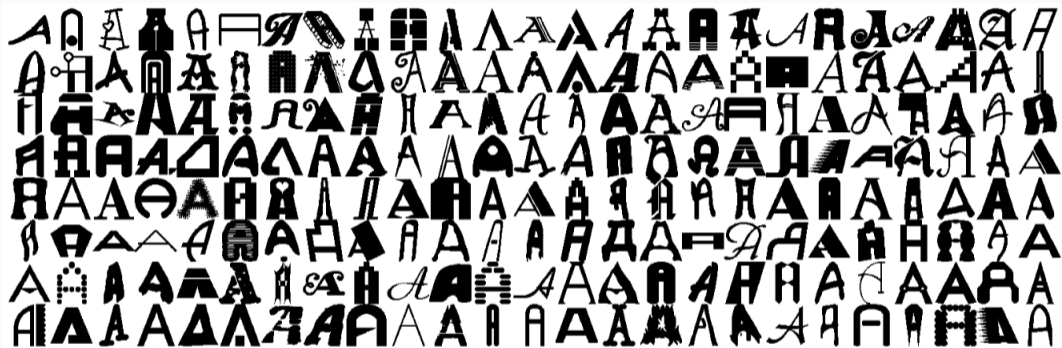
# DCGAN - diskriminator $D$



Slika 22: Arhitektura za diskriminator  $D$ .

- Odabrano 6628 različitih fontova i 26 slova engleskog alfabeta
- Veličina slike postavljena na  $64 \times 64$
- Po uzoru na (Radford, Metz, and Chintala 2015b) korišćenjem autoenkodera eliminisani slični fontovi
- Time je skup fontova smanjen na 200 predstavnika





Slika 23: 200 odabranih fontova.



Slika 24: Generisani simboli koristeći DCGAN

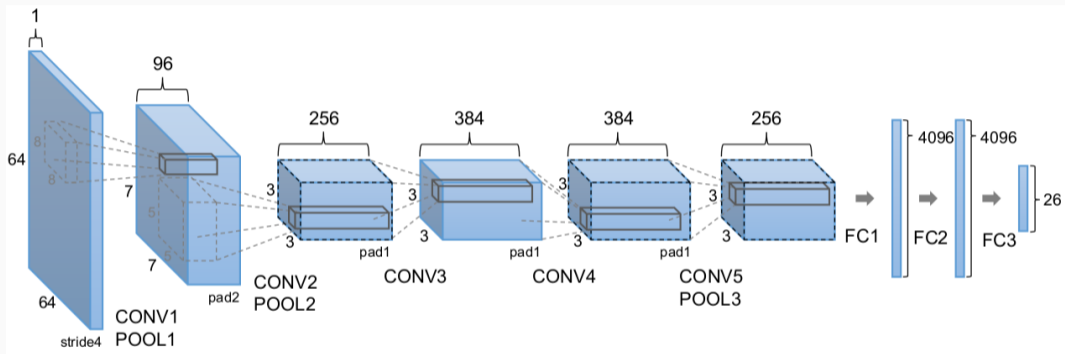
## Uzrok loših rezultata

- Tokom obučavanja ne postoje navođenje mreža da održe strukturu slova
- One se jednostavno takmiče da budu bolje jedna od druge
- Jedan od razloga je i nestabilnost pri obučavanju GAN-ova

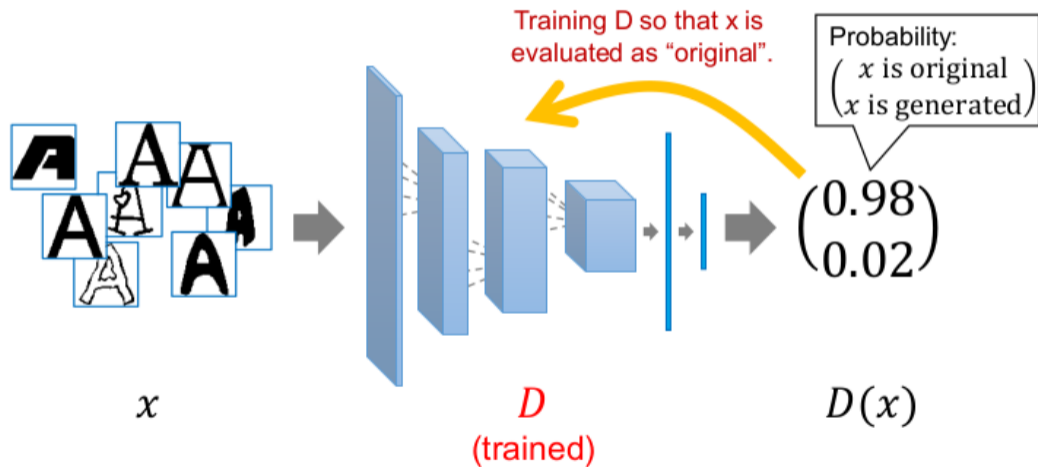
## Rešenje

- Razvija se posebni klasifikator simbola ( $C$ ) koji za sliku pogađa koje je slovo na njoj
- Njegovo obučavanje se vrši pre obučavanja GAN-a
- Odabrana arhitektura dosta liči na AlexNet
- Ostvaruje preciznost od 92.14% na skupu za obučavanje
- Proces obučavanja GAN-a se modifikuje tako da se uzima u obzir i predikcija koju daje  $C$
- Generator se **ažurira** tako što uzima u obzir šta predviđa mreža  $C$

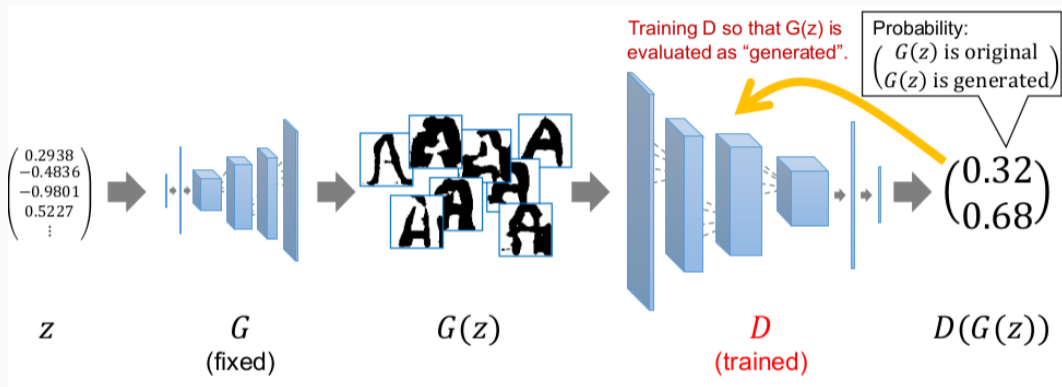
# DCGAN - klasifikator simbola



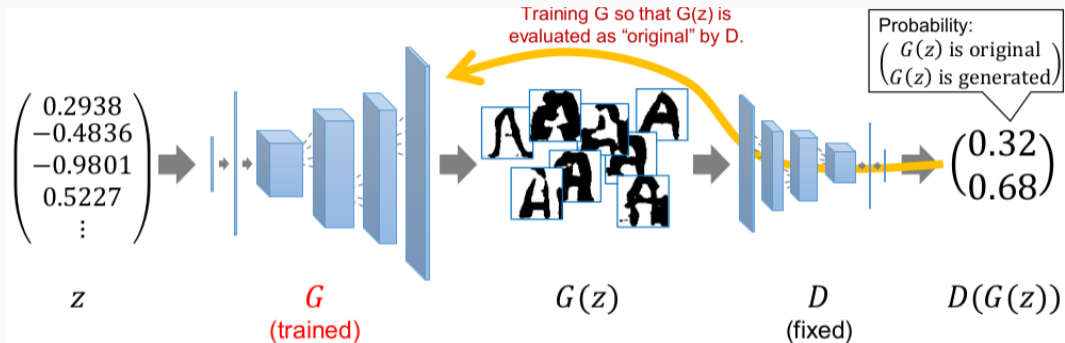
Slika 25: Arhitektura za klasifikator simbola.



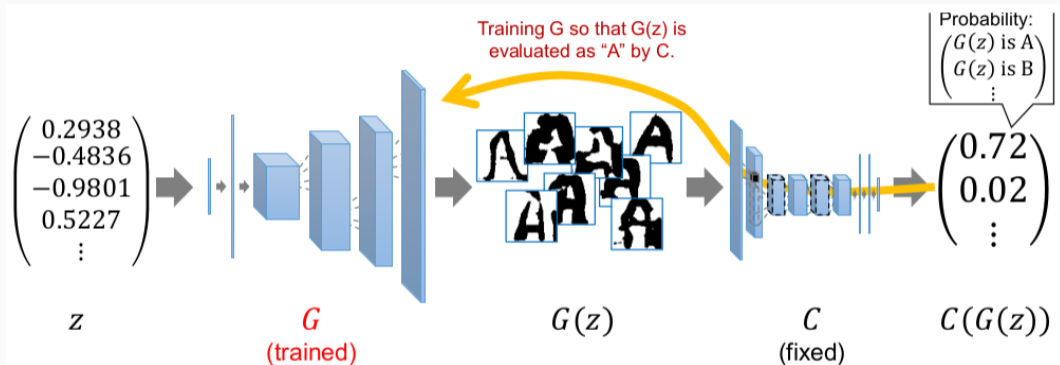
## DCGAN - obučavanje 2



Slika 27: Ažuriranje diskriminatora sa lažnim slikama.



Slika 28: Ažuriranje generatora.



Slika 29: Ažuriranje generatora uz pomoć klasifikatora.



# Učenje tipografskog stila

---

**A** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**B** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**C** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**D** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**E** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**F** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**G** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**H** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**I** ABCDEFGHIJKLMNOPQRSTUVWXYZ

**J** ABCDEFGHIJKLMNOPQRSTUVWXYZ

- Konstruisan skup podataka od 10000 fontova

## Postavka diskriminativnog problema

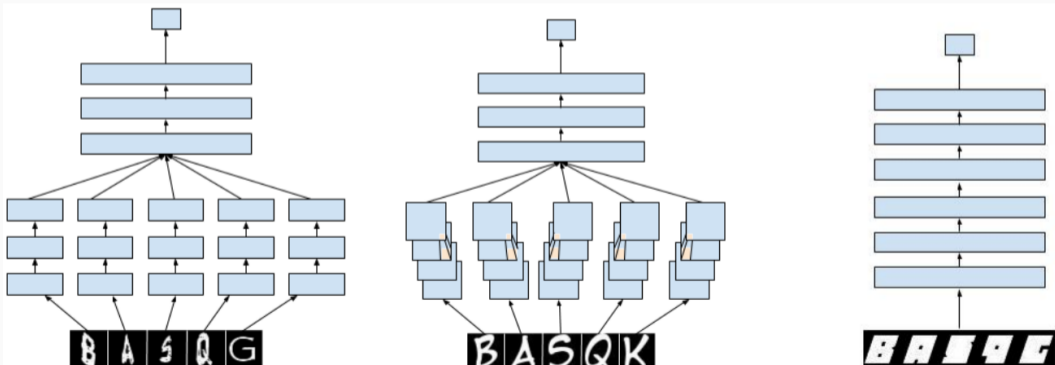
- Data su 4 slova istog fonta  $A$  (na primer B, A, S, Q) i nepoznato slovo  $\Phi$
- Potrebno je odrediti da li  $\Phi$  pripada fontu  $A$



Slika 31: Levo - pozitivni primeri, desno - negativni primeri.

## Učenje tipografskog stila - diskriminativni problem

- Pitanje: na koji način mreži proslediti više slova?
- Rad tvrdi da korišćenje razdvojenih kula (eng. *towers*) prikazano kao bolje u njihovim eksperimentima



## Učenje tipografskog stila - odabir klasifikatora

- Isprobano 60 arhitektura, odabrano 7 najboljih koje čine ansambl
- Testirano na fontovima nad kojima model nije obučavan

network	Description of hidden layers in each tower	Fully Connected hidden units in aggregation layers	Total free parameters	performance
1	1 Fully Connected (50 units)	2 (250,200)	437,602	90.3%
2	3 Fully Connected (50,50,50 units)	2 (250,200)	463,102	89.2%
3	3 Fully Connected (100,100,100 units)	2 (250, 200)	925,352	88.2%
4	3 Fully Connected (100,100,100 units)	2 (50, 50)	777,202	89.3%
5	2 Conv Paths, 2 Deep (3x3 → 3x3, 4x4 → 3x3)	2 (50, 50)	216,952	90.7%
6	2 Conv Paths, 2 Deep (3x3 → 3x3, 4x4 → 3x3)	4 (50, 50, 50, 50)	222,052	91.1%
7	2 Conv Paths, 2 Deep (3x3 → 3x3, 4x4 → 3x3)	4 (10, 10, 10, 10)	52,612	90.0%
<b>Voting Ensemble</b>				<b>92.1%</b>

**Slika 33:** 7 odabranih arhitektura koje čine ansambl.

# Učenje tipografskog stila - odabir klasifikatora

Correctly Classified

		1
		2
		3
		4
		5
		6
		7
		8
		9
		10

Incorrectly Classified


## Izdvajaju se dva pristupa

- Simultano generisanje svih slova
- Generisanje jednog slova

## Simultano generisanje svih slova

- Može biti potencijalno bolje jer dozvoljava deljenje težina
  - stil serifa, debljina linije. . .



# Učenje tipografskog stila - generativni problem



Slika 35: Generisanje jednog ili više slova.

## Učenje tipografskog stila - generisanje jednog slova

	Inputs	Gen.	SSE	shape	serif	accept?	comments by rater	Actual
1			1574	Yes	Yes	Yes		
2			1527	Yes	Yes	No	Missing divider	
3			1358	Yes	Yes	Yes		
4			1996	No	Yes	Yes		
5			1728	Yes	Yes	Yes		
6			1955	No	Yes	Yes	Top of R should not be closed	
7			1399	Yes	Yes	Yes		
8			2258	Yes	No	No	Serifs missing	

Slika 36: Komentari ljudi na generisana slova.

## Učenje tipografskog stila - generisanje jednog slova

	Inputs	Gen.	SSE	shape	serif	accept?!	comments by rater	Actual
9			3290	Yes	Yes	No	Similar, but not quite	
10			1355	Yes	Yes	Yes		
11			1359	Yes	Yes	Yes		
12			2593	Yes	Yes	Yes		
13			2428	Yes	Yes	Yes		
14			2144	No	No	No	Not Same	
15			1934	Yes	Yes	Yes		
16			1512	Yes	Yes	Yes	Shape very close, maybe tilted?	

Slika 37: Komentari ljudi na generisana slova.

## Učenje tipografskog stila - generisanje više slova

- Rad tvrdi da se kvalitet generisanih simbola nije popravio
- Ipak, prijavljuju da je smanjena greška za nekih 5-6%
- Usled toga, biraju da dalje koriste arhitekturu koja generiše više slova odjednom

## Učenje tipografskog stila - generisani primeri



## Učenje tipografskog stila - generisani primeri



# Alat Fontool

---

## Pristup

- Svođenje problema na problem generisanja slika
- Korišćenje variacionih konvolutivnih autoenkodera
- Obučavanje sprovedeno na grafičkoj karti NVIDIA GEFORCE 1060 GTX 6GB
- Korišćene biblioteke Keras i TensorFlow

## Razlozi

- Stabilnost pri obučavanju
- Latentni prostor
- Čistija matematička formulacija



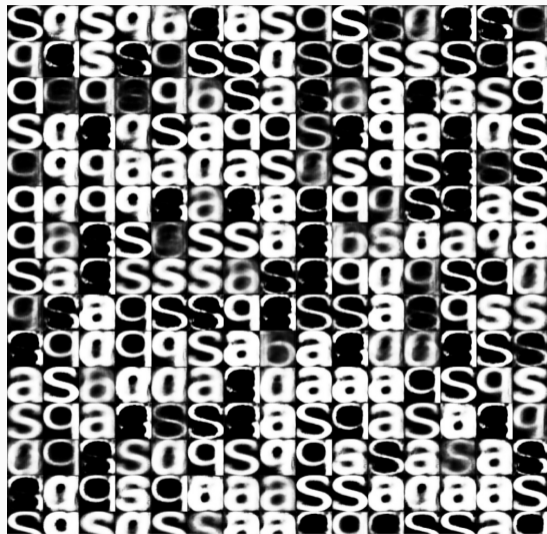
## Predloženi pristup problemu - skup podataka

- Razvijen skup podataka  $f\tau S$
- Odabrano 500 raznovrsnih fontova
- Za svaki font generisana reprezentacija slova b, a, s, q
- Slova odabrana po uzoru na rad (Baluja 2016)
- Dobijeno 2000 instanci

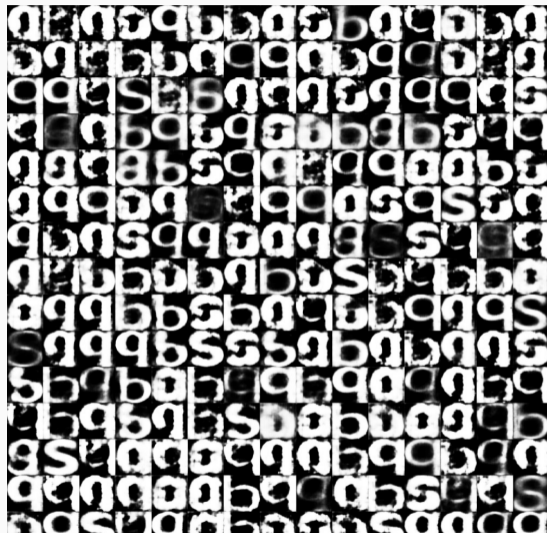
## Predloženi pristup problemu - uзорak iz skupa podataka



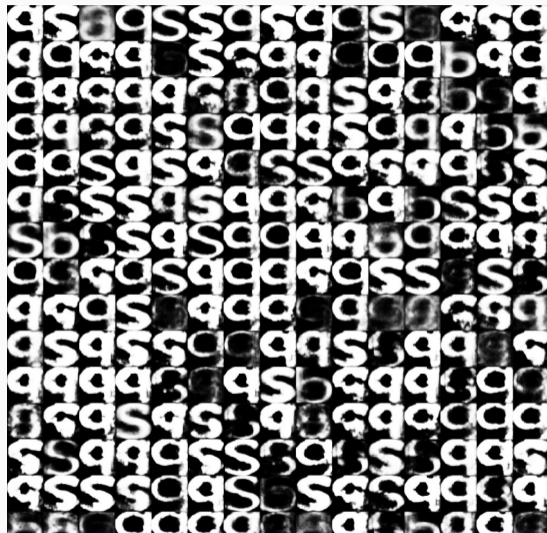
## Predloženi pristup problemu - rezultati (dimenzija 2)



## Predloženi pristup problemu - rezultati (dimenzija 3)



## Predloženi pristup problemu - rezultati (dimenzija 5)



## Predloženi pristup problemu - rezultati (dimenzija 30)



## Font generator

In the trained model the dimension of the latent space is 4. In order to generate a image please provide values for vector  $z$ .

To generate image, simply adjust the values of vector  $z$  and follow the preview bellow.

Image generation is performed using convolutional variational autoencoder.

$z_0$

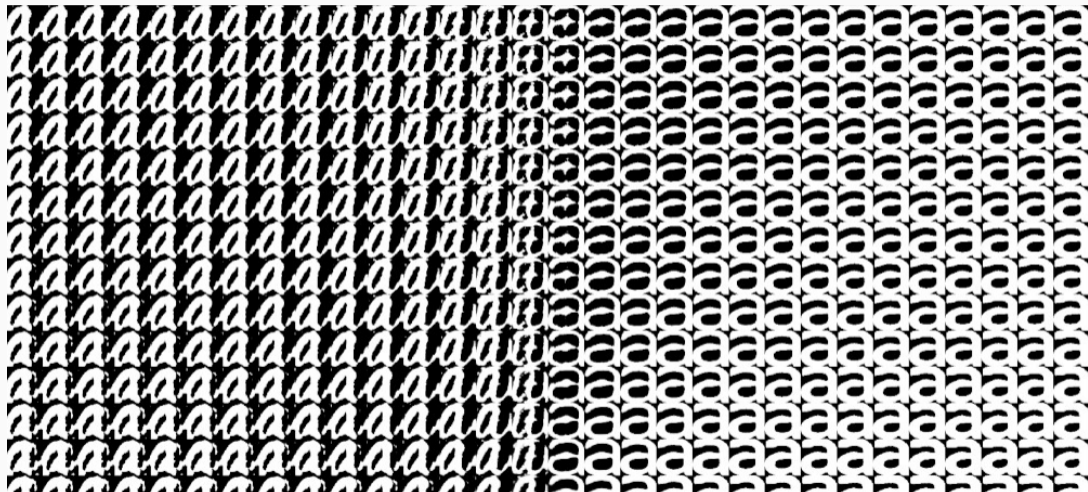
$z_1$

$z_2$

$z_3$



## Predloženi pristup problemu - interpolacija





- Korišćenje većeg broja podataka
- Korišćenje uslovnih variacionih autoenkodera
- Dodavanje dodatnih ulaza u autoenkoder
  - Stil koji je potrebno generisati
  - Širina slova
  - Visina slova
  - Odnos između najtanje i najdeblje linije

# Vizuelizacija fontova

---

- Problem u kojem je potrebno izvršiti grafički prikaz fontova tako da su fontovi grupisani po svojoj sličnosti
- Olakšava se organizacija fontova



**Helvetica**



**Arial**



**Grotesque**

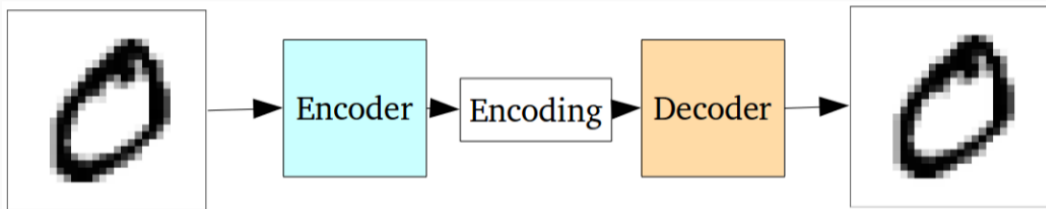
**Slika 40:** Nekoliko sličnih varijanti slova a

### Rad (Solli and Lenz 2009)

- Baziran na smanjivanju dimenzionalnosti
- Metoda glavnih pravaca (eng. *principal component analysis - PCA*)
- Metoda Isomap (Tenenbaum, Silva, and Langford 2000)

- Instance (slike dimenzija  $64 \times 64$ ) se pomoću enkodera preslikavaju u latentni prostor
- Autoenkoder prirodno grupiše slične instance blizu jedne drugih u okviru latentnog prostora
- U takvom prostoru može se podesiti dimenzija podataka
- Radi vizuelizacije, odabrana je dvodimenziona reprezentacija

## Vizuelizacija fontova - autoenkoder







- Dalji razvoj razvijenog modela
- Isprobavanje različitih arhitekutra
- Definisanje metrike sličnosti između fontova
- Korišćenje uslovnih variacionih autoenkodera
- Obučavanje na većem skupu podataka

# Veb aplikacija

---

- Radi potrebe demonstracije razvijena i prpratna veb aplikacija
- Omogućava interaktivno korišćenje i testiranje razvijenih sistema

## Flask

- Radni okvir za serverski deo veb aplikacija

## Bootstrap

- Radni okvir za kreiranje responzivnih veb aplikacija

## JQuery

- JavaScript biblioteka za klijentsko programiranje

## List.js

- JavaScript biblioteka za dinamički prikaz, filtriranje i sortiranje podataka

fontool Home Font explorer Classify font Generate fonts

## Fonttool app

Fonttool is a tool developed for font classification and generation.

It was created as my master thesis during studies at Faculty of Mathematics at University of Belgrade.

You can find the github repository [here](#).



## Fonts

Search font by name

Per page

Set per page

Sort

Appetite

The quick brown fox jumps over the lazy dog

Arilon Expanded

The quick brown fox jumps over the lazy dog

1 2 3 ...

fontool [Home](#) [Font explorer](#) [Classify font](#) [Generate fonts](#)

## Font classifier



Choose File linearstd\_stolica.png

Submit

## Classification results

User image:



Top 10 fonts:

- BallantinesSerial-Black: 0.95
- LinearStd-Heavy: 0.03
- CHE-DUST: 0.02
- inkswipe: 0.00
- Quandary BRK: 0.00
- Freeze Normal: 0.00
- DiamanteShadow-Italic: 0.00
- COMMON6: 0.00
- Blitzkrieg NF: 0.00
- RCMP2: 0.00



## Font generator

In the trained model the dimension of the latent space is 4. In order to generate a image please provide values for vector  $z$ .

To generate image, simply adjust the values of vector  $z$  and follow the preview bellow.

Image generation is performed using convolutional variational autoencoder.

$z_0$

$z_1$

$z_2$

$z_3$



## Zaključak

---

- Metode mašinskog učenja su u navedenim problemima postigle zadovoljavajuće rezultate
- Duboko učenje uspeva da pronađe vrlo široku oblast primene
- Razvijeni sistemi ipak nisu dovoljno robusni za primenu u praksi
- Ipak, predstavljaju dobru osnovu za dalji razvoj i istraživanje
- Ukoliko ste profesionalni tipograf, ne morate strahovati za svoj posao. . . još uvek! :)

**Pitanja?**

---

**Hvala na pažnji!**

---

Avilés-Cruz, Carlos, Risto Rangel-Kuoppa, Mario Reyes-Ayala, Edgar Andrade González, and Rafael Escarela-Perez. 2005. "High-Order Statistical Texture Analysis - Font Recognition Applied" 26 (January): 135–45.

Baluja, Shumeet. 2016. "Learning Typographic Style." *CoRR* abs/1603.04000.  
<http://arxiv.org/abs/1603.04000>.

Campbell, Neill D. F., and Jan Kautz. 2014. "Learning a Manifold of Fonts." *ACM Trans. Graph.* 33 (4). New York, NY, USA: ACM: 91:1–91:11.  
<https://doi.org/10.1145/2601097.2601212>.

Chen, Guang, Jianchao Yang, Hailin Jin, Jonathan Brandt, Eli Shechtman, Aseem Agarwala, and Tony X. Han. 2014. “Large-Scale Visual Font Recognition.” In *CVPR*, 3598–3605. IEEE Computer Society.

Glorot, Xavier, Antoine Bordes, and Yoshua Bengio. 2011. “Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach.” In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, 513–20. ICML’11. USA: Omnipress. <http://dl.acm.org/citation.cfm?id=3104482.3104547>.

Jung, Min-Chul, Yong-Chul Shin, and Sargur N. Srihari. 1999. “Multifont Classification Using Typographical Attributes.” In *ICDAR*, 353–56. IEEE Computer Society.

Ma, Huanfeng, and David Doermann. 2003. "Gabor Filter Based Multi-Class Classifier for Scanned Document Images." In *Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2*, 968. ICDAR '03. Washington, DC, USA: IEEE Computer Society. <http://dl.acm.org/citation.cfm?id=938980.939479>.

Radford, Alec, Luke Metz, and Soumith Chintala. 2015a. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." *CoRR* abs/1511.06434. <http://arxiv.org/abs/1511.06434>.

———. 2015b. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." *CoRR* abs/1511.06434. <http://arxiv.org/abs/1511.06434>.



Ramanathan, R., K. P. Soman, L. Thaneshwaran, V. Viknesh, T. Arunkumar, and P. Yuvaraj. 2009. "A Novel Technique for English Font Recognition Using Support Vector Machines." In *2009 International Conference on Advances in Recent Technologies in Communication and Computing*, 766–69. <https://doi.org/10.1109/ARTCom.2009.89>.

Solli, Martin, and Reiner Lenz. 2009. "Visualization of Large Font Databases," January, 10.

Tenenbaum, Joshua B., Vin de Silva, and John C. Langford. 2000. "A Global Geometric Framework for Nonlinear Dimensionality Reduction." *Science* 290 (5500). American Association for the Advancement of Science: 2319–23. <https://doi.org/10.1126/science.290.5500.2319>.

Wang, Zhangyang, Jianchao Yang, Hailin Jin, Jonathan Brandt, Eli Shechtman, Aseem Agarwala, Zhaowen Wang, et al. 2015. "DeepFont: A System for Font Recognition and Similarity." In *Proceedings of the 23rd Acm International Conference on Multimedia*, 813–14. MM '15. New York, NY, USA: ACM.

<https://doi.org/10.1145/2733373.2807988>.

Zeiler, Matthew D., and Rob Fergus. 2013. "Visualizing and Understanding Convolutional Networks." *CoRR* abs/1311.2901. <http://arxiv.org/abs/1311.2901>.