

# ZAVRŠNI RAD br. 6987: Semantička segmentacija kolničkih trakova

Jelena Bratulić  
Mentor: prof. dr. sc. Siniša Šegvić



SVEUČILIŠTE U ZAGREBU

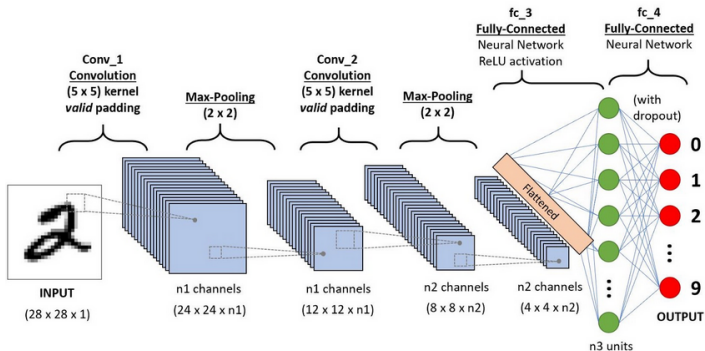
Fakultet  
elektrotehnike i  
računarstva

8. srpnja 2020.

- 1 Duboko učenje
  - Duboki konvolucijski modeli
  - Rezidualni modeli
  - Modeli s ljestvičastim naduzorkovanjem
- 2 Semantička segmentacija
- 3 Arhitektura modela
- 4 Skup podataka LLAMAS
- 5 Eksperimentalni rezultati
  - Eksperiment na skupu podataka CamVid
  - Eksperimenti na skupu podataka LLAMAS
- 6 Zaključak

# Duboki konvolucijski model

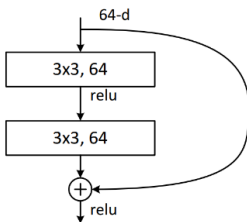
- Konvolucijski sloj, sloj sažimanja i potpuno povezani sloj
- Manji broj parametara i bolja uspješnost od klasičnih modela



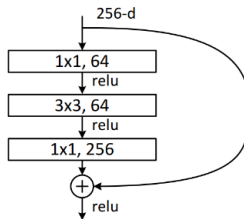
Preuzeto iz [1].

# Rezidualni modeli

- Često korišteni modeli u području računalnog vida
- Ubrzavaju proces učenja i smanjuju grešku prilikom učenja



(a) Osnovna rezidualna jedinica.

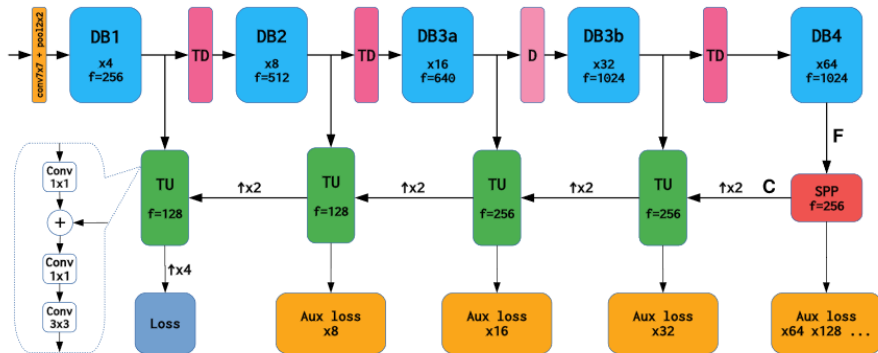


(b) Rezidualna jedinica s uskim grlom.

Preuzeto iz [2].

# Modeli s ljestvičastim naduzorkovanjem

- Rješenje problema guste predikcije



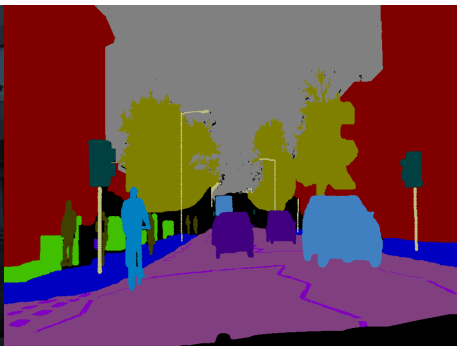
Preuzeto iz [3].

# Semantička segmentacija

- Dodjeljivanje oznake klase (labele) pojedinom pikselu sa slike
- Široka primjena u poljoprivredi, medicini i autonomnoj vožnji



(a) Slika snimljena iz perspektive vozača.

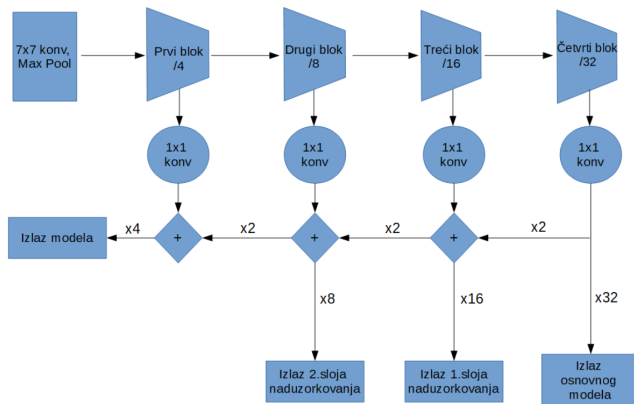


(b) Slika s označenim klasama.

Preuzeto iz [4].

# Arhitektura modela

- Pretrenirani ResNet-18
- Ljestvičasto naduzorkovanje
- Bilinearno naduzorkovanje na izlazu





(a) Primjerak slike iz skupa za učenje.



(b) Ručno označena točna segmentacija.

Preuzeto iz [5].



# Skup podataka LLAMAS

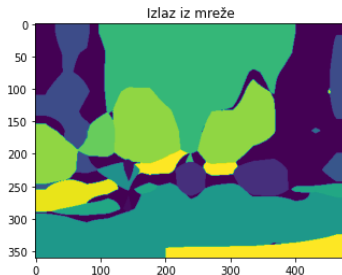
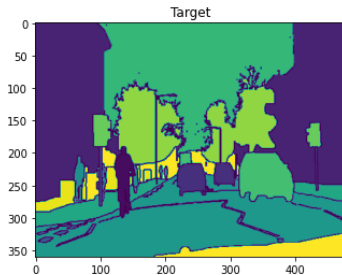
Ukupan broj klasa	5
Broj korištenih klasa	2, 5
Ukupan broj slika	79,113
Skup za učenje	58,269
Podrezani skup za učenje	2,400
Skup za validaciju	10,029
Skup za testiranje	10,815
Izvorna rezolucija	1276 × 717
Korištena rezolucija	638 × 200, 960 × 288

**Tablica:** Osnovni podaci za skup podataka LLAMAS.

- Platforma: Google Colaboratory (GPU NVIDIA TESLA K80)
- Broj epoha: 80
- Veličina grupe za učenje: 12
- Funkcija gubitka: Unakrsna entropija s težinama
- Optimizator: Adam s pretpostavljenim vrijednostima
  - Stopa učenja: 0.001
  - $\beta_1$ : 0.900
  - $\beta_2$ : 0.999

# Eksperiment na skupu podataka CamVid

- Preliminarni eksperiment za provjeru modela
- Skup podataka CamVid - 11 klasa
- Razumijevanje scene iz perspektive vozača
- Uspješnost modela evaluirana udjelom točno klasificiranih piksela u ukupnom broju piksela
- Postignuta točnost od 94% na skupu za učenje i 83 % na skupu za testiranje



# Eksperimenti na skupu podataka LLAMAS

## Prvi eksperiment - binarna segmentacija

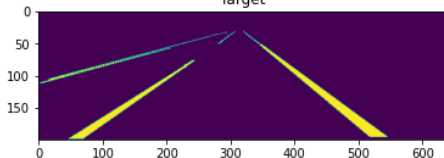
Vrsta modela	Mode	Gubitak	F1	IoU
Osnovni model	učenje	0.102	0.388	0.242
	validiranje	0.104	0.368	0.226
	testiranje	0.133	0.275	0.161
Jedan sloj naduzorkovanja	učenje	0.074	0.521	0.341
	validiranje	0.088	0.519	0.350
	testiranje	0.101	0.398	0.255
Dva sloja naduzorkovanja	učenje	0.054	0.640	0.471
	validiranje	0.079	0.610	0.431
	testiranje	0.090	0.434	0.298
Tri sloja naduzorkovanja	učenje	0.044	0.678	0.514
	validiranje	0.070	0.641	0.468
	testiranje	0.080	0.499	0.341

**Tablica:** Brojčani pokazatelji za skup podataka LLAMAS na skupovima za učenje, validaciju i testiranje po slojevima ljestvičastog naduzorkovanja.

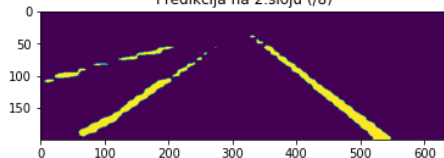
# Eksperimenti na skupu podataka LLAMAS

## Prvi eksperiment - binarna segmentacija

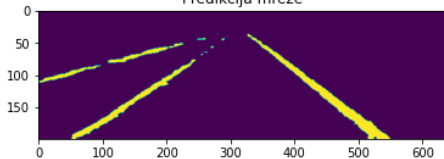
Target



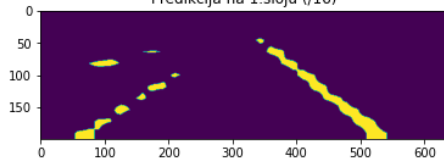
Predikcija na 2.sloju (/8)



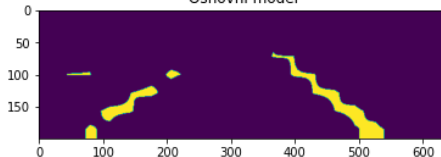
Predikcija mreže



Predikcija na 1.sloju (/16)



Osnovni model



# Eksperimenti na skupu podataka LLAMAS

Drugi eksperiment - utjecaj aux gubitka

Izlaz iz modela	Mode	Gubitak	F1	IoU
Osnovni model	učenje	0.098	0.385	0.239
	validiranje	0.103	0.385	0.238
	testiranje	0.113	0.280	0.164
Jedan sloj naduzorkovanja	učenje	0.086	0.534	0.364
	validiranje	0.093	0.502	0.336
	testiranje	0.110	0.369	0.227
Dva sloja naduzorkovanja	učenje	0.072	0.659	0.493
	validiranje	0.086	0.603	0.432
	testiranje	0.105	0.467	0.305
Tri sloja naduzorkovanja	učenje	0.068	0.706	0.546
	validiranje	0.079	0.639	0.471
	testiranje	0.098	0.506	0.335

**Tablica:** Brojčani pokazatelji za nezavisni model koji ima izlaze na različitim slojevima naduzorkovanja.

# Eksperimenti na skupu podataka LLAMAS

Treći eksperiment - ovisnost veličine skupa za učenje

Veličina skupa za učenje	Vrsta skupa	Gubitak	F1	IoU
2,400	Učenje	0.044	0.678	0.514
	Validacija	0.070	0.641	0.468
	Testiranje	0.080	0.499	0.341
58,269	Učenje	0.045	0.656	0.489
	Validacija	0.058	0.671	0.504
	Testiranje	0.056	0.563	0.393

**Tablica:** Usporedba uspješnosti modela ovisno o različitoj veličini skupa za učenje.

# Eksperimenti na skupu podataka LLAMAS

## Četvrti eksperiment - višeklasna segmentacija

Model mreže	mAP
ERFNet-IntRA-KD	0.598
ResNet-101	0.607
ResNet-50	0.579
UNet-ResNet-34	0.592
ResNet-18 (moj)	0.389

**Tablica:** Usporedba uspješnosti modela na skupu podataka LLAMAS.



# Vizualizacija rezultata

- Razvijeni model uspijeva pronaći oznake na kolniku
- Trakovi za pretjecanje (isprekidane linije) iz udaljenijih dijelova scene se i dalje ne prepoznaju dobro
- Dodatna poboljšanja su moguća korištenjem dubljih modela i drukčije kombinacije aktivacijske funkcije, funkcije gubitka i hiperparametara

- [1] Sumit Saha. A comprehensive guide to convolutional neural networks — the eli5 way, 2018. URL <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way>
- [2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, i Jian Sun. Identity Mappings in Deep Residual Networks. ECCV 2016.
- [3] Ivan Krešo, Josip Krapac, i Siniša Šegvić. Efficient Ladder-style DenseNets for Semantic Segmentation of Large Images. arXiv e-prints: 1905.05661.
- [4] Gabriel J. Brostow, Julien Fauqueur, i Roberto Cipolla. Semantic object classes in video: A high-definition ground truth database. Pattern Recognition Letters, 2008.
- [5] Karsten Behrendt i Ryan Soussan. Unsupervised labeled lane marker dataset generation using maps. U Proceedings of the IEEE International Conference on Computer Vision, 2019.