

Kategoriziranje fotografija

Maria Crnjak, Luka Jurić-Grgić i Antonija Pantar

Travanj 2019

Contents

1	Uvodni opis problema	3
2	Cilj i hipoteze istraživanja	3
3	Pregled dosadašnjih istraživanja	3
4	Materijali, metodologija i plan istraživanja	3
5	Očekivani rezultati predloženog projekta	4

1 Uvodni opis problema

U današnje vrijeme društvenih mreža i mobilnih uređaja sve se kreće vrlo brzo, a posebno razvoj mobilne tehnologije. Tako danas u mobilnim telefonima imamo vrlo kvalitetne fotoaparate i fotografija je sve više diljem svijeta. Osim tehnologije, vrlo bitan dio ljudskog života je prehrana. Ako ne znate što je to u Vašem tanjuru, čak i uz ovu današnju tehnologiju, ne biste mogli lako saznati odgovor, osim ako koristite svoj fotoaparat.

Naime, danas postoji mnogo aplikacija koje na temelju fotografije mogu reći nešto o tome što se na fotografiji nalazi - ili ne nalazi. Osim hrane, ovakve aplikacije mogu analizirati i kategorizirati bilo što, što možemo uslikati. Za naš projekt odlučili smo napraviti sličnu aplikaciju.

Na stranici Yelp postoji Dataset challenge koji daje priliku studentima da se okušaju u analizi slika ili teksta. S te stranice skinuli smo dataset u kojem se nalazi 200 000 fotografija hrane, pića, restorana i razni podaci o restoranima, korisnicima tih restorana i samim fotografijama. U ovom Data setu fotografije su podijeljene u 5 kategorija: Inside, Outside, Food, Drink, Menu.

2 Cilj i hipoteze istraživanja

Naš cilj je napraviti program koji što točnije prepozna da li slika koju smo mu dali prikazuje restoran izvana, iznutra, hranu, piće ili meni restorana.

3 Pregled dosadašnjih istraživanja

Ovaj dataset challenge bio je predmet brojnih istraživanja, a neka od njih su:

- 2018. – „Generalized Latent Variable Recovery for Generative Adversarial Networks” (Nicholas Egan, Jeffrey Zhang i Kevin Shen). Generator koji transformira latentne vektore u realistične slike.
- 2017. – „Understanding Hidden Memories of Recurrent Neural Networks” (Hong Kong University of Science and Technology). Modeli učenja koji su vrlo učinkoviti u modeliranju podataka o slijedu, kao što su tekst i govor.
- 2017. – „hotdog app” – Aplikacija koja prepoznaje hot dog-ove.

4 Materijali, metodologija i plan istraživanja

Od materijala nam je dostupan spomenuti dataset sa stranice Yelp. Pravila natjecanja i dataset mogu se naći na [ovom](#) linku.

Kako bismo mogli testirati podatke morat ćemo podijeliti naš Data set na dva dijela: train set (75%) i test set (25%).

Train set ćemo koristiti kako bismo "naučili" program kakve se slike nalaze u kojoj kategoriji, a zatim ćemo testirati "znanje" uz pomoć test seta.

Na kraju ćemo analizirati našu točnost te ju usporediti s rezultatima popularnih aplikacija.

5 Očekivani rezultati predloženog projekta

Očekujemo da ćemo uspjeti napraviti program koji prepoznaje što se od pet kategorija (inside, outside, food, drink, menu) nalazi na danoj fotografiji.

Komercijalne aplikacije imaju točnost čak do 99% dok većina kodova koje smo našli na GitHubu postiže između 70 i 99% točnih rezultata.

Mi pretpostavljamo da ćemo se naći negdje u sredini. Imamo dosta velik train set u odnosu na broj kategorija tako da možemo računati da će nam to povećati točnost.