Izveštaj prvog domaćeg zadatka iz predmeta Duboko učenje

# Image segmentation

## Kratak opis problema

Cilj ovog projekta je razviti modele za segmentaciju slika koje prikazuju različite vrste hrane, kao što su donut, pizza i kolač, a potom njihova kategorizacija. Kroz analizu podataka, implementaciju modela i eksperimente, cilj je postići tačnu klasifikaciju objekata hrane na slikama.

## Informacije o podacima

Podaci korišćeni u projektu dolaze iz COCO (Common Objects in Context) dataset-a. Specifično, identifikatori slika za krofnu, pizzu i kolač su izdvojeni za dalju analizu.

## Analiza podataka

Analiza podataka uključuje učitavanje slika iz URL-ova, promenu dimenzija slika na 224x224 piksela, i dodatne obrade slika kako bi se stvorila matrica odgovarajućeg oblika koja bi bila puštena modelu. Matrica puštena modelu je kategorizovana matrica oblika [[[1,0,0,0],[0,1,0,0],[0,0,1,0],[0,0,0,1]]] gde kada je u pitanju nepoznata kategorija bit je 1 na prvom mestu, kada je donut onda je na drugom, pizza na 3. i cake na 4.

## Zaključci iz analize podataka

Pregledavanjem seta podataka od po 5 slika za svaku kategoriju možemo zaključiti.

Krofne:

Krofne su uglavnom u prvom planu i često ih ima više na slici. Sve krofne su u standardnom obliku krofne ili pregrižene. Na većini slika se jasno vidi da je krfona u pitanju. U slučaju da su pregrižene uglavnom ih drži neka osoba. Uglavnom se od traženih kategorija pojavljuju samo krofne na slikama, tj ne mešaju se.  
  
Pizze:

Na slikama se retko pojavljuje cela pizza. Uglavnom fali jedno ili više pačića. Uglavnom se jasno vidi da je pizza isečena na trouglaste parčiće. Uglavnom su uprvom planu. Korica i centar pice se znatno razlikuju unutar jedne slike ali i izmedju različitih slika. Primećujemo različite tipove pizza. Neretko na slikama imamo osobu ili deo osobe (ruku koja seče picu nožem ili sekačem)

Kolači:

Najčešće su u obliku mafina ili torte. Ukoliko je u pitanju torta najčešće je standardnog kvadratnog oblika. Skoro uvek na nekom tipu tanjira ili podmetača. Slike su najčešće pravljene tako da prikažu i osobu koja jede kolač ili su close-upovi za neki tip reklame ili promociju.

## Opis rešenja

### Modeli:

U rešavanju problema klasifikacije slika, odlučio sam se za implementaciju tri modela koji su poznati po svojoj efikasnosti u radu sa slikama: DenseNet121, MobileNetV2 i EfficientNetB0. Korišćen je pristup transfer learning-a, gde su preuzete težine modela obučene na velikim skupovima podataka, poput ImageNet-a. Dodatno, dodao sam specifične slojeve kako bih prilagodio modele zadatku klasifikacije hrane.

### Opis arhitekture modela:

Modeli su kompajlirani sa Adam optimizatorom, a funkcija gubitka je kategorička unakrsna entropija, koja je prikladna za višeklasnu klasifikaciju.

#### Model 1: DenseNet121

Kao osnovni model koristio sam DenseNet121.

Dodao sam dodatne konvolucione slojeve sa upsampling operacijama jer sam primetio da naglim upsampleovanjem dobijem izuzetno pixelizovanu sliku gde su pixeli ogromni.

#### Model 2: MobileNetV2

Za osnovni model izabrao sam MobileNetV2, poznat po efikasnosti i maloj veličini.

Dodao sam dodatne konvolucione slojeve sa upsampling operacijama prilagođene specifičnostima problema.

#### Model 3: EfficientNetB0

Odabrao sam EfficientNetB0, poznat po postizanju visoke tačnosti sa malim brojem parametara.

Dodao sam dodatne konvolucione slojeve sa upsampling operacijama prilagođene specifičnostima problema.

#### Opis eksperimenata:

Eksperimenti su obuhvatili proces obuke modela na trening skupu podataka i evaluaciju performansi na odvojenom skupu za validaciju. Koristio sam callback-ove, posebno EarlyStopping kako bih praćenjem performansi sprečio preprilagođavanje modela i raniji prekid u tom slučaju. Podatke sam podelio u trening i validacioni skup.

## Zaključci

Sva tri modela su trenirana i evaluirana. Neki od modela su bolji od drugih. Primećujemo i veliku razliku prilikon eksperimenata na validacionom i na trening skupu, gde je trening skup bolje obrađen.