

Osnovi računarske inteligencije
-Klasifikacija rasa pasa pomoću dubokog učenja-
Dušan Lečić SV80/2021

1 – Analiza problema

Klasifikacija rasa pasa pomoću dubokog učenja predstavlja složen problem koji zahteva pažljivo razmatranje različitih aspekata kako bi se postigla visoka tačnost i pouzdanost modela. Primarna svrha ovog projekta je automatizacija prepoznavanja rasa pasa sa slike, što može značajno unaprediti veterinarske prakse, rad azila za životinje, kao i doprineti edukaciji javnosti.

2 – Programsko rešenje

Rešenje se zasniva se na korišćenju konvolucijskih neuronskih mreža (CNN) sa prethodno obučenim modelom ResNet50. Plan je da se slike pasa preprocesiraju i da se koristi model koji je obučen na velikim skupovima podataka kao što je ImageNet, kako bi se dobile bolje performanse na specifičnom zadatku klasifikacije rasa pasa.

Preprocesiranje podataka je ključni korak u pripremi slika za obuku modela. Proces preprocesiranja uključuje sledeće korake:

1. **Normalizacija:** Vrednosti piksela slike se skaliraju na opseg $[0, 1]$ ili standardizuju da imaju srednju vrednost 0 i standardnu devijaciju 1. Ovo omogućava stabilniji i brži trening modela, koristeći statistike ImageNet skupa podataka: srednje vrednosti $[0.485, 0.456, 0.406]$ i standardne devijacije $[0.229, 0.224, 0.225]$.
2. **Skaliranje:** Sve slike su skalirane na uniformnu veličinu (224×224 piksela), što je standardna veličina ulaza za mnoge prethodno obučene CNN modele, uključujući ResNet50.
3. **Augmentacija Slika:** Kako bi se povećala generalizacija modela i izbeglo prekomerno prilagođavanje (overfitting), koristiće se tehnike augmentacije kao što su horizontalno prevrtanje, rotacija i zumiranje. Ovo pomaže da model postane robusniji na varijacije u ulaznim slikama. Konkretno, koristiće se:
 - **Horizontalno prevrtanje:** Sa verovatnoćom od 0.3.
 - **Rotacija:** Nasumično rotiranje slike za do 30 stepeni.
 - **Random Crop:** Nasumično isecanje slike na veličinu od 224×224 piksela.
4. **Anotacija i Labeling:** Svaka slika mora biti ispravno označena sa odgovarajućom rasom psa, što je korišćeno kao ciljna oznaka (label) tokom obuke modela.

Trening modela se izvodi korišćenjem PyTorch okruženja i uključuje sledeće korake:

1. Inicijalizacija Modela:

- Model se inicijalizuje korišćenjem prethodno obučениh težina ResNet50 modela. Broj klasa (rasa pasa) se postavlja na osnovu broja različitih rasa u skupu podataka.

2. Optimizacija:

- Optimizacija se vrši korišćenjem SGD (Stochastic Gradient Descent) algoritma sa početnom stopom učenja od 0.001 i momentumom od 0.9.
- Scheduler se koristi za smanjenje stope učenja svakih 7 epoha.

3. Obuka i Evaluacija:

- Model se trenira u dva koraka: treniranje i validacija. Tokom treniranja, model se obučava i ažurira težine. Tokom validacije, model se evaluira na validacionom skupu podataka bez ažuriranja težina.
- Gubitak (loss) i tačnost (accuracy) se izračunavaju za svaku epohu i koriste se za praćenje performansi modela.

4. Čuvanje Najboljeg Modela:

- Najbolje težine modela (one sa najvećom tačnošću na validacionom skupu) se čuvaju i koriste za konačni model.

5. Sačuvavanje Modela:

- Nakon završetka obuke, model se čuva u fajlu `dog_breed_classifier.pth`.

ResNet50

ResNet50 (Residual Networks) je duboka konvolucijska neuronska mreža koja je postala poznata po svojim performansama na zadacima klasifikacije slike. ResNet50 se sastoji od 50 slojeva i koristi inovativnu tehniku nazvanu "residual connections" ili "skip connections". Ova tehnika omogućava da se izbegne problem nestajanja gradijenata, što je uobičajeni problem kod vrlo dubokih mreža.

Ključne karakteristike ResNet50 uključuju:

- **Residual Blocks:** Ovi blokovi sadrže direktne veze koje zaobilaze jedan ili više slojeva, omogućavajući prenos informacija iz ranijih slojeva direktno u kasnije slojeve. Ovo olakšava obuku dubokih mreža.
- **Deep Architecture:** Sa 50 slojeva, ResNet50 može naučiti složenije reprezentacije, što je korisno za preciznu klasifikaciju slika.
- **Transfer Učenje:** ResNet50 se često koristi za transfer učenje. Model se inicijalno obučava na velikom skupu podataka kao što je ImageNet, a zatim se dodatno trenira (fine-tuning) na specifičnom skupu podataka (u ovom slučaju, slike pasa).

Prednosti ResNet50 za ovu primenu:

- **Prethodna Obuka:** Korišćenje modela koji je već obučen na velikom i raznovrsnom skupu podataka smanjuje vreme i resurse potrebne za obuku.
- **Preciznost i Efikasnost:** ResNet50 je pokazao izuzetne rezultate u zadacima klasifikacije slike, što ga čini idealnim izborom za zadatak klasifikacije rasa pasa.

3 – Struktura projekta

Projekat je organizovan kroz sledeće ključne fajlove, od kojih svaki ima specifičnu ulogu u procesu razvoja modela za klasifikaciju rasa pasa:

- **data_processing.py:** Ovaj fajl je odgovoran za pretprocesiranje podataka i podešavanje skupa podataka. Sadrži funkcije za učitavanje slika, normalizaciju, augmentaciju i pripremu skupa podataka za obuku, validaciju i testiranje.
- **model.py:** Ovaj fajl služi za inicijalno podešavanje modela. U njemu se nalazi kod za kreiranje i prilagođavanje ResNet50 modela, uključujući zamenu završnog sloja tako da odgovara broju rasa pasa koje se klasifikuju.
- **train.py:** Ovaj fajl je zadužen za obuku modela. Sadrži funkcije za treniranje modela, praćenje gubitka (loss) i tačnosti (accuracy) tokom epoha, kao i za čuvanje najboljih rezultata modela.
- **evaluate.py:** Ovaj fajl je odgovoran za evaluaciju tačnosti modela. Sadrži kod za testiranje modela na validacionom i test skupu podataka, kao i za izračunavanje relevantnih metričkih performansi.
- **prediction.py:** Pomoćni fajl koji se koristi za testiranje modela na stvarnim fotografijama pasa. Sadrži funkcije za učitavanje pojedinačnih slika.

4 – Analiza rezultata

Izvršavanjem evaluacionog skripta utvrđeno je da model postiže tačnost od 94% na validacionom skupu podataka. Ovaj visok procenat tačnosti ukazuje na to da model uspešno prepoznaje i klasifikuje rase pasa na osnovu dostavljenih slika. Tačnost modela je evaluirana korišćenjem podataka podeljenih na trening, validacioni i test set, što osigurava da rezultati reflektuju realne performanse modela na neviđenim podacima.

5 – Dodatna unapređenja projekta

Iako trenutni dataset obuhvata 120 rasa pasa i pruža solidnu osnovu za obuku i evaluaciju modela, nije sveobuhvatan i ne uključuje sve rase pasa koje postoje. Dodavanje novih rasa u dataset predstavljalo bi značajan korak ka povećanju preciznosti i robusnosti modela.

Pored toga, moguće je unaprediti projekat razvijanjem frontend aplikacije sa drag & drop mehanizmom. Ova funkcionalnost bi korisnicima omogućila jednostavno korišćenje modela nad stvarnim podacima, olakšavajući učitavanje i evaluaciju slika pasa direktno.

6 – Primer izlaza

U sledećem primeru, prosleđena je slika nemačkog ovčara i izvršen je poziv skripte iz terminala. Rezultat je prikazan ispod.



```
dogifier/src on master [!?] via 🐘 v3.12.4 (env) on ☁ (eu-central-1)
> py prediction.py
PATH: ../../german_shephard.jpg | The predicted breed is: German shepherd
```