

Detekcija i određivanje rase pasa na slikama

Dragana Grbić, Petar Nikolić
Fakultet Tehničkih Nauka, Novi Sad

Opis problema i motivacija

Cilj projekta je detekcija pasa na slikama i određivanje njihove rase. Ulaz u sistem predstavlja slika na kojoj se nalaze jedan ili više pasa. Izlaz iz sistema je slika na kojoj su psi uokvireni okvirima sa labelama koje predstavljaju rase pasa.

Ovakav sistem može biti iskorišćen kao *proof of concept* za veći sistem za određivanje rase pasa i mačaka, vrsti papagaja i drugih kućnih ljubimaca. Takođe, mogao bi se, u unapređenom obliku, koristiti i prilikom lova ili safarija, zatim za potrebe snimanja dokumentarnih filmova, dakle kada je potrebno sa određene distance pronaći životinju koja je korisniku od interesa.

Skup podataka

Za osnov skupa podataka za treniranje i validaciju iskorišćen je *Staford Dog Dataset* skup slika pasa, koji sadrži anotirane slike 121 rase pasa, ravrstanih po rasama. Od toga, mi smo odabrali 5 rasa pasa:

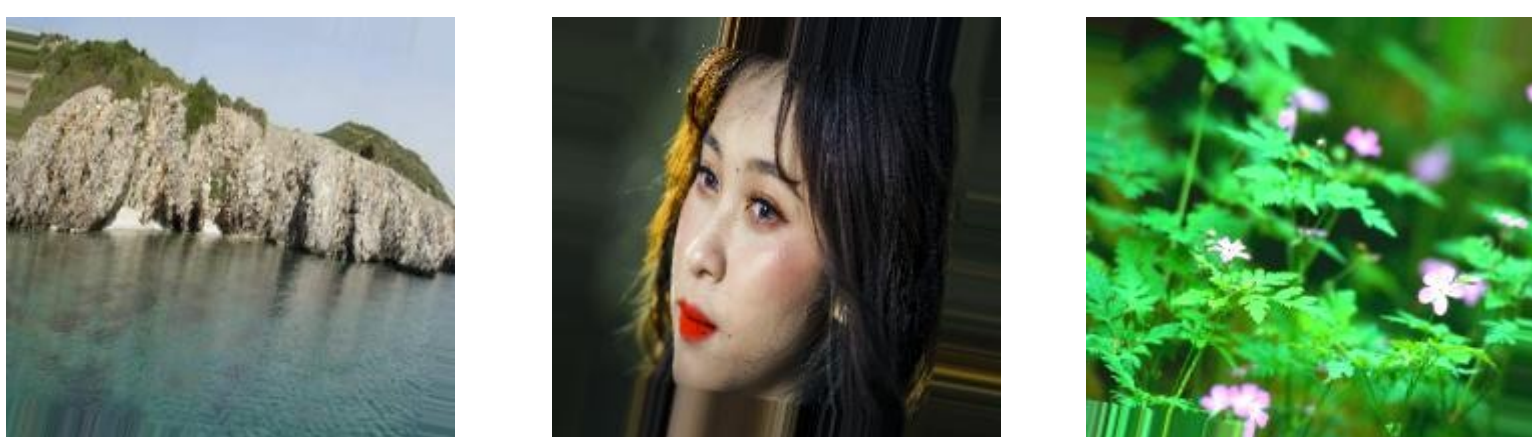
- Kerry blue terrier
- Rottweiler
- Basenji
- Leonberg
- Samoyed

Zbog malog broja slika, korišćena je augmentacija podataka, čime je od polaznih 968 slika dobijeno 13.375 slika pasa, razvrstanih po rasama. Takođe, skup podataka sadrži kao posebnu klasu i 8.990 slika na kojima nisu psi. Od ukupnog broja slika, 10% je iskorišćeno kao validacioni skup.

Kao testni skup, korišćeno je 100 slika pasa, ručno dobavljenih sa interneta. Slike sadrže ukupno 136 pasa.



Slika 1 – primeri slika na kojima su psi



Slika 2 – primeri slika na kojima nisu psi

Metod 1 - Custom CNN

Prvi metod se oslanja na korišćenje konvolucione neuronske mreže (CNN) sa custom-made arhitekturom koja sadrži 6 konvolucionih slojeva.

U zavisnosti od algoritma za generisanje regiona od interesa imamo tri varijante ovog metoda:

- Selective search – fast podešavanje (SSF)
- Selective search – quality podešavanje (SSQ)
- Sliding window – klizajući prozor (SW)

Podešavanja *Selective search* algoritma određuju broj izgenerisanih regiona od interesa, odnosno regiona koji možda sadrže pse. Fast podešavanje daje manje, a quality više regiona. U projektu se koristi implementacija ovog algoritma iz Python OpenCv biblioteke.

Sliding window tehnika je implementirana sa 12 različitih dimenzija za klizajući prozor.

Za otklanjanje preklapajućih regiona od interesa koji sadrže pse, koristi se *Non-Maxima Suppression* (NMS) algoritam, koji je ručno implementiran tako da favorizuje regione većih dimenzija, čime se izbegava obeležavanje jednog psa sa više manjih regiona, a takođe se dobija i efekat grupisanja više pasa iste rase u iste regione, što povećava preglednost rezultata.

Metod 1 - Rezultati

CNN je trenirana u 30 epoha uz korišćenje early-stopper mehanizma. Najbolje performanse ostvarene su u 21. epohi:

- Validation loss – 0.1242
- Validation accuracy – 96.58%

U tabeli ispod, prikazani su rezultati testiranja po varijantama:

Varijanta	Detekcija (%)	Klasifikacija (%)	False-positives
SSF	100	93,38	4
SSQ	99,26	94,85	9
SW	100	94,12	3

Uzimajući u obzir procenat detektovanih i tačno klasifikovanih pasa, ali i broj false-positive-a, najbolje se pokazala treća varijanta ovog metoda.



Slika 3 – primer izlaza bez (ceo) i sa false-positive-om (isečak)

Metod 2 - Custom CNN + ResNet

Drugi metod predstavlja modifikaciju prvog metoda koji u fazi detekcije pasa na slici koristi pretreniranu *ResNet50* neuronsku mrežu iz Python *Tensorflow* biblioteke, dok se određivanje rase, odnosno klasifikacija obavlja našom CNN iz prvog metoda.

Razlog za korišćenje *ResNet50* mreže jeste pokušaj umanjenja broja false-positive-a, koji su uočeni tokom testiranja prvog metoda.

ResNet50 je trenirana na *ImageNet* skupu podataka sa preko 14 miliona slika podeljenih preko 20000 klasa, od čega 118 klasa predstavlja rase pasa.

Ovaj metod takođe ima tri varijante, identične onima iz metoda 1.

Metod 2 - Rezultati

U tabeli ispod, prikazani su rezultati testiranja po varijantama:

Varijanta	Detekcija (%)	Klasifikacija (%)	False-positives
SSF	100	88,23	0
SSQ	100	86,03	0
SW	99,26	80,88	0

Rezultati pokazuju da je korišćenje *ResNet50* mreže ispunilo očekivanja i dovelo do nepostojanja false-positive-a.

Međutim, korišćenje ove mreže je dovelo i do nižeg stepena tačnosti prilikom klasifikacije pasa. Ovo bi se moglo objasniti činjenicom da regioni od interesa koji su dovoljno dobri da ResNet na njima prepozna pse, ne moraju nužno biti i dovoljno dobri da bi naša CNN mogla tačno klasifikovati pse u njima.

Ipak, prva varijanta ovog metoda ostvarila je visok stepen tačnosti, te je njena upotreba i dalje unapređivanje opravdano.



Slika 4 – primer izlaza i primer odsustva false-positive-a

Zaključak i uočeni nedostaci

Na osnovu dobijenih rezultata, može se zaključiti da su oba metoda, pogotovo u trećoj, odnosno prvoj varijanti, pokazali visok stepen efikasnosti u detekciji, ali i određivanju rasa pasa.

Uočeno je i postizanje boljih rezultata korišćenjem *Sliding Window* algoritma u prvom metodu, dok je u drugom metodu bolje rezultate dalo korišćenje *Selective search* algoritma.

Ipak, korišćenje konvolucione neuronske mreže sa custom-made arhitekturom, kao i skup podataka skromne veličine ima za posledicu pojavu false positive-a u prvom, odnosno neusklađenosti naše konvolucione neuronske mreže sa naprednijom *ResNet50* neurosnskom mrežom.

Iz istog razloga se ne očekuju odlične performanse naše CNN u slučaju primene na slike niskog kvaliteta, lošeg osvetljenja, slika koje sadrže pse malih dimenzija, kao i u slučaju varijacija u izgledu u okviru iste rase.

Pravci daljeg razvoja

Iako je naša CNN pokazala visok stepen efikasnosti, za dalji razvoj sistema bi bilo najbolje korisiti neku od state-of-the-art neuronskih mreža, poput *YOLO* i *ResNet* mreža, a efikasnost bi se mogla unaprediti i proširivanjem skupa podataka.

U smislu funkcionalnosti, sistem bi se mogao proširiti tako da omogućava detekciju i klasifikaciju više vrsti životinja.

Takođe, sistem bi se mogao učiniti sposobnim i za real-time obradu video snimaka.