



UTVRĐIVANJE RASA PASA I MAČAKA UPOTREBOM KONVOLUCIONIH NEURONSKIH MREŽA

DOG AND CAT BREED DETERMINATION USING CONVOLUTIONAL NEURAL NETWORKS

Petar Nikolić, *Fakultet tehničkih nauka, Novi Sad*

Oblast – SOFTVERSKO INŽENJERSTVO I INFORMACIONE TEHNOLOGIJE

Kratak sadržaj – U radu su predložena dva pristupa rešavanju problema utvrđivanja rase pasa i mačaka na slikama. Prvi pristup predstavlja klasifikaciju u jednom koraku u kojem se utvrđuje rasa životinje bez prethodnog utvrđivanja njene vrste. U drugom pristupu, rasa se utvrđuje u dva koraka – u prvom se određuje vrsta životinje, dok se u drugom slika prosleđuje specijalizovanom klasifikatoru koji utvrđuje rasu u okviru prethodno utvrđene vrste. Implementacija oba pristupa zasniva se na upotrebi konvolucionih neuronskih mreža, kao i tehnike augmentacije podataka. Dodatno, kod oba pristupa ispitan je i uticaj primene tehnike aktivnog učenja.

Ključne reči: klasifikacija slika, konvolucione neuronske mreže, augmentacija podataka, aktivno učenje

Abstract – The paper proposes two approaches to solving the problem of determining the breed of dogs and cats in images. The first one is a one-step classification without previously determining the species of the animal. The second one consists of two steps - in the first one, the image is classified based on the species. In the second one, the image is passed to a specialized classifier which determines the breed within the previously determined species. Both solutions are based on the use of convolutional neural networks, as well as the data augmentation technique. Additionally, the impact of the application of the active learning technique was analyzed for both proposed solutions.

Keywords: image classification, convolutional neural networks, data augmentation, active learning

1. UVOD

Kućni ljubimci, pogotovo psi i mačke kao neki od najpopularnijih, imaju bitnu ulogu u životu ljudi, od razonode, preko emocionalne podrške, do pomoći u slučaju zdravstvenih problema. Mačkama i psima se posvećuje dosta vremena na internetu, te tako postoji i veliki broj aplikacija koje su im posvećene, kao što su *Dog Buddy* [1], *MeowTalk* [2] i *Relax My Cat* [3].

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Aleksandar Kovačević, red. profesor.

Neretko se pak dešava da vlasnici nisu dovoljno upoznati sa rasama svojih pasa i mačaka, što može dovesti do nesporazuma među vlasnicima, kako uživo tako i *online*. Stoga bi postojanje načina za pouzdano i automatsko utvrđivanje rase ljubimaca na osnovu slike bilo značajno za njih.

Utvrdjivanje rase pasa i mačaka predstavlja problem višeklasne klasifikacije slika. Rešavanje ovog problema izazovno je iz više razloga. Jedan jesu velike razlike u izgledu koje mogu postojati u okviru jedne rase ali i između rase pasa i mačaka (boja, dužina krzna itd). Drugi problem se tiče kvaliteta samih slika, odnosno visoke varijabilnosti po pitanju rezolucije i dimenzija slike, te osvetljenosti i zamućenosti.

U ovom radu predloženo je rešenje problema utvrđivanja rase pasa i mačaka na primeru 3 rase pasa i 3 rase mačaka. Rešenje se zasniva na višeklasnoj klasifikaciji slika upotrebom konvolucione neuronske mreže sa samostalno razvijenom arhitekturom. Radi rešavanja prvog problema ispitan je utvrđivanje rase sa i bez prethodnog utvrđivanja vrste životinje. U cilju dobijanja robusnih modela kojima bi mogao biti rešen i problem kvaliteta slika, u rešenju je korišćena i tehnika augmentacije podataka primenom različitih transformacija slike. Dodatno, radi dobijanja manjeg trening skupa, a uz očuvanje ili unapređenje performansi rešenja, na modelima korišćenim za klasifikaciju ispitan je uticaj primene tehnike aktivnog učenja.

U drugom poglavlju dat je opis korišćenog skupa podataka. U trećem poglavlju opisane su metode i tehnike korišćene prilikom implementacije rešenja. U četvrtom poglavlju opisano je predloženo rešenje. U petom poglavlju prikazani su ostvareni rezultati primene predloženog rešenja i analizirane njegove greške. U šestom poglavlju nabrojani su najvažniji zaključci i navedeni mogući pravci daljeg razvoja.

2. SKUP PODATAKA

U cilju implementacije predloženog rešenja, iskorišćen je *The Oxford-IIIT Pet Dataset* [4] skup podataka sa ukupno 7,349 slika pasa i mačaka. Od tog broja, 4,978 slika za 25 rase pasa i 2,371 slika za 12 rase mačaka. Za svaku od rase postoji po 200 slika. Dodatno, sve slike u skupu poseduju anotacije koje sadrže naziv rase životinje na slici, zatim koordinate pravougaonika koji sadrži glavu životinje, kao i metapodatke o svojstvima same slike, kao što su njene dimenzije i naziv.

Za potrebe razvoja i testiranja predloženog rešenja korišćen je podskup goreopisanog skupa koji se sastoji od slika po 3 rase pasa i mačaka. Konkretno, u pitanju su *Basset*, *Leonberger* i *Samoyed* u slučaju pasa i *Persian*, *Egyptian Mau* i *Siamese* u slučaju mačaka.

3. KORIŠĆENE METODE I TEHNIKE

Od tehnika mašinskog učenja prilikom implementacije predloženog rešenja korišćene su konvolucione neuronske mreže, augmentacija podataka i aktivno učenje. Pomoću konvolucionih neuronskih mreža kreirani su modeli za potrebe klasifikacije, augmentacija podataka je iskorišćena za podizanje kvaliteta trening i test skupova, dok je aktivno učenje upotrebljeno radi smanjenja obima trening skupa uz očuvanje ili unapređenje performansi modela.

3.1. Neuronske mreže

U cilju kreiranja modela koji obavljaju klasifikaciju slika pasa i mačaka, korišćene su konvolucione neuronske mreže sa samostalno razvijenom arhitekturom, koja sadrži 6 konvolucionih slojeva, koji koriste *ReLU* aktivacionu funkciju i ispraćeni su svaki sa po jednim *Max Pooling* i *Dropout* slojevima. Nakon konvolucionih, sledi jedan potpuno povezani sloj, praćen poslednjim slojem sa *Softmax* aktivacionom funkcijom.

3.2. Aktivno učenje

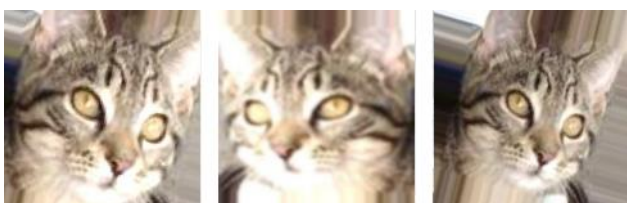
Aktivno učenje predstavlja tehniku polunadgledanog učenja koja se koristi u cilju smanjenja količine podataka potrebne za obučavanje modela mašinskog učenja, a sa očuvanjem performansi ili njihovim unapređenjem [5]. Ovo se postiže iterativnim povećavanjem trening skupa, gde se u svakoj iteraciji utvrđuju najinformativniji podaci, to jest podaci iz kojih bi model mogao najviše naučiti.

Glavni problem primene ove tehnike jeste odabir metrike za informativnost podataka, budući da postoji dosta različitih metrika, počev od nasumičnog odabira kao najjednostavnije, do metrika koje se zasnivaju na različitim teorijskim pretpostavkama.

Prilikom implementacije predloženog rešenja, ispitan je uticaj primene 4 različite metrike – *EP Measure*, *BvSB*, kao i *Bavg* i *Bmax* varijante *Bernoulli entropy* metrike.

Tehnika aktivnog učenja u ovom radu primenjena je na sledeći način:

1. iz celog trening skupa T izdvaja se pseudonasumično odabrani podskup P , koji sadrži 20% slika ukupnog skupa
2. iz podskupa P izdvajaju se trening i validacioni skupovi u odnosu 90-10 i obavlja se treniranje modela u $30 - 5 \times k$ epoha, gde broj k predstavlja trenutnu iteraciju aktivnog učenja



Slika 1. Primeri slika dobijenih augmentacijom podataka

3. primena treniranog modela na skup T i rangiranje dobijenih predikcija primenom metrike informativnosti
4. prebacivanje N najinformativnijih slika iz skupa T u skup P
5. ponavljanje koraka 2-4 u K iteracija

Hiperparametri goreopisanog algoritma su broj iteracija K , broj najinformativnijih podataka N , odnos veličina trening i validacionog skupa, kao i primenjena metrika informativnosti i njihove vrednosti su određene empirijski.

3.3. Augmentacija podataka

Radi dobijanja modela koji su otporniji na loš kvalitet slika, te kako bi se nadomestio mali broj slika po rasi u korišćenom skupu podataka, primenjena je tehnika augmentacije podataka u tri koraka:

1. Učitavanje slika, kao i koordinata pravougaonika koji sadrži glavu životinje iz anotacija radi kreiranja isečka originalne slike sa glavom životinje. Isečci su dimenzija 224x224 piksela.
2. Nad dobijenim isečcima primenjuje se jedna ili više transformacija, i to rotacija ulevo i udesno za do 15 stepeni, horizontalno okretanje, isecanje po obe ose do 10%, zumiranje u opsegu 0.95–1.25x, promena osvetljenosti u opsegu od -30% do +40%. Broj transformacija, kao i vrednosti u navedenim opsezima određuju se pseudonasumično za svaki isečak ponaosob.
3. Primena *Gaussian Blur* transformacije sa kernelom dimenzija 7x7 nad oko 30% pseudonasumično odabranih transformisanih isečaka.

Na slici 1 prikazani su primeri slika dobijenih goreopisanom postupkom. Nakon primene augmentacije podataka, dobijen je konačan obučavajući skup sa 6,264 slika, sa 1,044 slika po rasi.

4. OPIS PREDLOŽENOG REŠENJA

Tokom razvoja predloženog rešenja isprobana su dva pristupa klasifikaciji rase pasa i mačaka. Sledi opis oba pristupa.

4.1. Prvi pristup – klasifikacija u jednom koraku

Prvi pristup predstavlja utvrđivanje rase pasa i mačaka u jednom koraku, bez prethodnog utvrđivanja vrste životinje na slici. U tu svrhu korišćena je konvoluciona neuronska mreža sa samostalno razvijenom arhitekturom, koja je obučena da prepozna 3 rase pasa i 3 rase mačaka. Ovaj model će u nastavku nositi oznaku CNN-1.

Prednost ovog pristupa jeste lakoća njegove implementacije, budući da se oslanja na samo jednu neuronsku mrežu. Slabost pristupa leži u tome što dobijeni model može napraviti krupnu grešku ukoliko sliku psa klasifikuje kao sliku mačke i obrnuto.

4.2. Drugi pristup – klasifikacija u dva koraka

Drugi pristup predloženog rešenja (CNN-2S u nastavku) predstavlja utvrđivanje rase u dva koraka:

1. U prvom koraku, slika se klasifikuje po principu pasmačka upotrebom konvolucione neuronske mreže sa samostalno razvijenom arhitekturom (u nastavku CNN-2). Mreža je trenirana na celom trening skupu, podeljenom u klase *dog* i *cat*.
2. U drugom koraku se slika na osnovu rezultata prvog koraka prosleđuje jednoj od dve specijalizovane konvolucione neuronske mreže sa samostalno razvijenom arhitekturom. Jedna od ovih mreža specijalizovana je za klasifikaciju po rasama pasa, a druga po rasama mačaka. Ove mreže su trenirane na odgovarajućim podskupovima celog trening skupa koji sadrže samo slike pasa, odnosno samo slike mačaka. Mreža koja obavlja klasifikaciju po rasama pasa u nastavku će nositi oznaku CNN-D, dok će mreža za klasifikaciju mačaka nositi oznaku CNN-C.

4.3. Ispitivanje metrika informativnosti

Uticao primene tehnike aktivnog učenja ispitano je kod oba pristupa, za sva 4 modela. Broj iteracija primene aktivnog učenja K je za sva četiri modela iznosio 4, dok je vrednost hiperparametra N određena eksperimentalno za svaku iteraciju primene za svaki od četiri korišćena modela ponaosob (tabela 1).

Tabela 1. Vrednost hiperparametra N po modelima i iteracijama

Iteracija	CNN-1	CNN-2	CNN-D	CNN-C
1	1,050	1,050	525	525
2	1,050	1,050	525	525
3	900	900	450	450
4	900	900	450	450

Radi evaluacije primene odabranih metrika, postupak ispitivanja metrika informativnosti je odrađen tri puta i uporedo su analizirane srednje vrednosti tačnosti (*accuracy*) i *loss*-a na validacionom skupu, i to po završetku druge iteracije primene aktivnog učenja. Srednje vrednosti su analizirane budući da se početni trening skup u prvom koraku postupka određuje pseudonasumično.

Druga iteracija je odabrana zbog činjenice da su sva četiri modela ostvarila jako dobre performanse već na toj iteraciji, a uz korišćenje samo oko 53% celog trening skupa, kao i da značajnijeg povećanja performansi u sledećim iteracijama nije bilo.

Tabela 2 sadrži prikaz najboljih metrika informativnosti po modelima, kao i ostvarene srednje vrednosti tačnosti i *loss*-a.

Tabela 2. Rezultati ispitivanja metrika informativnosti

Model	Metrika	Tačnost (%)	Gubitak
CNN-1	Bmax	99.6	0.0145
CNN-2	Bmax	99.2	0.0248
CNN-D	EP Measure	100	0.0029
CNN-C	BvSB	99.8	0.0136

5. REZULTATI I DISKUSIJA

Oba predložena pristupa evaluirana su pomoću $F1$ mere nad test skupom. U nastavku je dat opis korišćenog test

skupa i postupka evaluacije. Dodatno, navedena je i analiza grešaka korišćenih modela.

5.1. Test skup

Za potrebe testiranja performansi predloženog rešenja, korišćen je test skup od ukupno 420 slika, sa 60 slika po rasi. Skup je formiran odabirom po 7 slika za svaku rasu, čime je dobijen inicijalni skup sa 42 slike. Nad ovim skupom je zatim primenjen postupak augmentacije podataka identičan onom koji je primenjen za kreiranje konačnog trening skupa, a koji je opisan u trećem poglavlju.

Kreiranje ovakvog test skupa je bilo potrebno kako bi se mogle proveriti performanse dobijenih modela u slučaju potrebe za klasifikacijom slika niskog kvaliteta i različitih karakteristika, koje se mogu očekivati u realnim uslovima korišćenja.

5.2. Postupak testiranja i ostvareni rezultati

Obavljeno je uporedo testiranje oba pristupa, i to sa i bez primene aktivnog učenja. U slučaju kada aktivnog učenje nije primenjeno modeli su trenirani na celokupnim trening skupovima, dok su u drugom slučaju obučavani primenom dve iteracije aktivnog učenja, uz korišćenje one metrike informativnosti koja se za taj model pokazala najboljom.

U tabeli 3 prikazane su ostvarene vrednosti $F1$ mere za svaki od modela. Druga kolona sadrži vrednosti u slučaju kada ova tehnika nije primenjena (Klasično učenje), a treća u slučaju kada jeste (Aktivno učenje). Poslednji red u tabeli sadrži ukupnu vrednost ove mere za drugi pristup, budući da se on zasniva na upotrebi tri različita modela.

Tabela 3. Rezultati testiranja predloženog rešenja

Model/Pristup	Klasično učenje	Aktivno učenje
CNN-1	0.9361	0.9786
CNN-2	0.9952	0.9881
CNN-D	0.9762	0.9521
CNN-C	0.9845	0.9667
CNN-2S	0.9738	0.9476

Dobijeni rezultati pokazuju da se drugi pristup pokazao osetno boljim u slučaju klasičnog učenja, ali da se u slučaju aktivnog učenja pokazao nešto gorim od prvog pristupa.

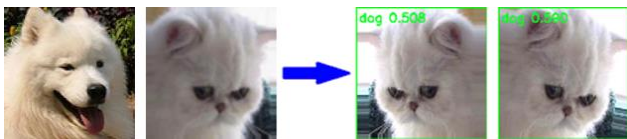
Takođe, može se zaključiti da su se po pitanju performansi modeli dobijeni primenom aktivnog učenja pokazali kao uporedivi modelima dobijenih primenom klasičnog učenja, a uz korišćenje tek samo oko 53% ukupnog trening skupa. Izuzetak predstavlja model CNN-1, koji je pokazao osetno bolje performanse prilikom primene aktivnog učenja.

5.3. Analiza grešaka modela

Svi korišćeni modeli pokazali su odlične performanse uz mali broj grešaka. Razlozi za nastale greške se u većini slučajeva mogu naći u niskom kvalitetu slika ili jakim deformacijama nastalim primenom transformacija tokom augmentacije test skupa.

Određen broj grešaka može se objasniti i sličnošću u izgledu između nekih rasa pasa i mačaka (npr. boja krzna).

Model CNN-2 dobijenom primenom aktivnog učenja je tako pogrešno klasifikovao sliku sa mačke kao sliku psa (slika 2). Ova greška se lako objašnjava velikom vizuelnom sličnošću u izgledu mačke rase *Persian*, koja može imati krzno bele boje i psa rase *Samoyed*, koji po pravilu ima krzno bele boje.



Slika 2. Greške modela (desno) usled sličnosti sa drugom rasom (levo)

Takođe, model CNN-2 dobijen bez primene aktivnog učenja je pogrešno klasifikovao psa rase *Basset* kao mačku (slika 3). U ovom slučaju je kombinacija loših karakteristika originalne slike, to jest sličnosti u boji glave psa i okoline i delimične vidljivosti njegovog lica, i primenjenih transformacija prilikom augmentacije (rotacija i zamućenje) dovela do greške u klasifikaciji.



Slika 3. Greška modela (desno) usled deformisanosti finalne i loših karakteristika polazne slike

6. ZAKLJUČAK

Razmotren je problem utvrđivanja rasa pasa i mačaka, kao i značaj za njegovo rešavanje. Predložena su dva različita pristupa za njegovo rešavanje, koja se zasnivaju na upotrebi konvolucionih neuronskih mreža i tehnike augmentacije podataka. Dat je opis korišćenog skupa podataka, kao i primenjenog načina njegove augmentacije, zatim koncepta aktivnog učenja, kao i arhitekture korišćene neuronske mreže. Predloženo rešenje je testirano i rezultati verifikovani.

Na osnovu dobijenih rezultata, pokazalo se da su oba pristupa dala odlične rezultate. Dodatno, kod oba pristupa obavljeno je poređenje rezultata dobijenih sa i bez primene tehnike aktivnog učenja, koje je pokazalo da je moguće smanjenje trening skupa od oko 47%, a uz očuvanje ili unapređenje performansi treniranih modela.

Iako su oba pristupa dala uporedive rezultate, prednost treba dati drugom pristupu zbog njegove modularnosti. Naime, dodavanje podrške za novu vrstu životinja bi se svodilo na treniranje modela koji obavlja klasifikaciju rasa u okviru te vrste i dodatno treniranje modela koji se koristi u prvom koraku, dok narušavanja performansi modela korišćenih u drugom koraku za postojeće vrste ne bi bilo.

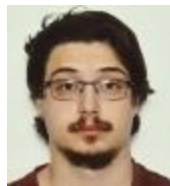
Na osnovu svega navedenog, klasifikacija u dva koraka, gde se u prvom određuje vrsta, a u drugom rasa psa ili mačke, u kombinaciji sa primenom tehnika aktivnog učenja i augmentacije podataka, predstavlja dobro rešenje problema utvrđivanja rasa pasa i mačaka na slikama.

Moguća proširenja predloženog rešenja jesu implementacija podrške za detekciju pasa i mačaka na slikama i klasifikaciju više životinja na slici, kao i povećanje broja podržanih rasa. Dobijeni rezultati bi se mogli unaprediti primenom naprednijih arhitektura konvolucionih neuronskih mreža, drugih transformacija slike u sklopu augmentacije podataka i drugih metrika informativnosti. Dodatno, u cilju bolje analize grešaka modela, mogle bi se primeniti i neke od tehnika vizuelizacije modela mašinskog učenja.

7. LITERATURA

- [1] <https://www.dogbuddy.com/> (pristupljeno u oktobru 2022.)
- [2] <https://www.meowtalk.app/> (pristupljeno u oktobru 2022.)
- [3] <https://relaxmycat.com/> (pristupljeno u oktobru 2022.)
- [4] Zisserman, Andrea Vedaldi Andrew, and C. V. Jawahar. "Cats and dogs." IEEE conference on computer vision and pattern recognition. 2012.
- [5] Settles, Burr. "Active Learning Literature Survey." (2009).

Kratka biografija:



Petar Nikolić rođen je u Pirotu 1998. godine. Diplomirao je na Fakultetu tehničkih nauka iz oblasti Softversko inženjerstvo i informacione tehnologije 2021. godine. Master rad iz iste oblasti odbranio je 2022. godine.