

# Klasifikacija ugroženih vrsta životinja primenom konvolutivnih neuronskih mreža

Sara Jokić

Fakultet tehničkih nauka  
Univerzitet u Novom Sadu  
Trg Dositeja Obradovića 6  
21000, Novi Sad  
sara.jokic1@gmail.com

Tamara Krgović

Fakultet tehničkih nauka  
Univerzitet u Novom Sadu  
Trg Dositeja Obradovića 6  
21000, Novi Sad  
tasakrgovic@gmail.com

**Sažetak-** Prepoznavanje ugroženih vrsta životinja je od velikog značaja za očuvanje biodiverziteta i održavanje ekološke ravnoteže. Olakšati prepoznavanje ugroženih vrsta doprinelo bi podizanju svesti javnosti i edukaciji o značaju očuvanja prirode. Konvolutivne neuronske mreže sposobne su za automatsko ekstraktovanje značajnih karakteristika iz slika pa shodno tome dobro su se pokazale za rešavanje datog problema. U ovom radu predstavljen je uporedni pregled različitih arhitektura konvolutivnih neuronskih mreža za klasifikaciju ugroženih vrsta životinja u tri klase (panda, polarni medved, beloglavi orao). Skup podataka je prikupljen putem kamera u zoološkim vrtovima i prirodnim staništima životinja, nakon čega su fotografije anotirane koristeći Roboflow alat i eliminisane su neupotrebljive fotografije. Detektovane životinje su sačuvane u novi skup podataka. Eksperimentalni rezultati pokazuju da je InceptionV3 model postigao najbolju tačnost od 99,21%, pri čemu su određeni konvolutivni slojevi, kao i klasifikator, trenirani na našem skupu podataka, uz optimizaciju hiperparametara metodom *random search*. Nasuprot tome, ResNet arhitektura je ostvarila najnižu tačnost od 82,57%.

**Ključne reči:** *ResNet, VGG19, AlexNet, MobileNet, InceptionV3, klasifikacija životinja*

## I. UVOD

Prepoznavanje ugroženih vrsta životinja je od ključnog značaja za očuvanje biodiverziteta i održavanje ekološke ravnoteže. Svaka vrsta igra jedinstvenu ulogu u svom ekosistemu, kada ove vrste nestanu, ravnoteža ekosistema je narušena, a druge vrste mogu biti negativno pogođene. Pored toga, gubitak biodiverziteta ima i druge negativne posledice, može učiniti ekosisteme osetljivijim na stresne faktore iz okoline, kao što su zagađenje, klimatske promene i bolesti[1]. U svetu svake godine izumre između 100 i 10.000 vrsta, od mikroorganizama do velikih biljaka i životinja[2]. Njihov nestanak povezuje se isključivo s ljudskom aktivnošću.

Pravovremena identifikacija ovih vrsta omogućava preduzimanje ciljnih mera zaštite, čime se sprečava njihov potpuni nestanak i čuvaju ključne ekološke zone. Osim toga, prepoznavanje ugroženih vrsta igra ključnu ulogu u podizanju svesti javnosti i edukaciji o značaju očuvanja prirode, što je od presudne važnosti za dugoročno zdravlje naše planete.

Konvolutivne neuronske mreže (CNN) koriste konvolutivne slojeve za automatsko ekstraktovanje značajnih karakteristika iz slika, čime omogućavaju modelima da prepoznaju i razlikuju specifične obrasce i detalje. Ova sposobnost automatskog prepoznavanja karakteristika omogućava precizno klasifikovanje ugroženih vrsta s visokom tačnošću, čak i u lošim uslovima snimanja ili sa slikama lošeg kvaliteta.

U ovom radu želeli smo da prikazemo rezultate dobijene finim podešavanjem (eng. fine-tuning) tri poznate arhitekture konvolutivnih neuronskih mreža za rešavanje problema klasifikacije ugroženih vrsta životinja na tri klase: pandu, polarnog medveda i beloglavog orla. Napravili smo uporedni prikaz performansi ovih modela. Arhitekture koje smo koristili obuhvataju InceptionV3, ResNet i VGGNet. Poseban izazov bio je sakupiti dovoljno veliki i raznovrstan skup podataka nad kojim ćemo trenirati modele. U narednom poglavlju dat je pregled relevantne literature, u trećem poglavlju posvećena je pažnja skupu podataka, dok su u četvrtom poglavlju prikazane metodologije korišćene za fino podešavanje (eng. fine-tuning) modela. U petom poglavlju daćemo pregled naših rezultata, dok je u poslednjem poglavlju dat kratak osvrt na zaključke do kojih smo došli u ovom radu, kao i plan za dalji rad i poboljšanje postojećeg rešenja.

## II. SRODNA ISTRAŽIVANJA

U istraživanju “*Real-Time Identification of Animals found in Domestic Areas of Europe*”[3] autori se bave identifikacijom 34 klase životinja na području Evrope, koristeći četiri različite konvolutivne neuronske mreže: VGG-19, InceptionV3, ResNet-50 i MobileNetV2. Razvijen je sistem za klasifikaciju ovih životinja iz slika, kao i u realnom vremenu iz video zapisa. Sistem automatski generiše dva skupa podataka: jedan sa tekstualnim informacijama (naziv klase, datum i vremenski interval prisustva) i jedan sa slikama identifikovanih klasa. Eksperimentalni rezultati pokazuju visoku tačnost za sve četiri arhitekture (90.56% za VGG-19, 93.41% za InceptionV3, 93.49% za ResNet-50, i 94.54% za MobileNetV2). Istraživanje izdvaja MobileNetV2 arhitekturu kao najbolju u rešavanju ovog problema. S obzirom na veliku sličnost korišćenih metodologija sa našim radom, odlučili smo da isprobamo fino podešavanje i MobileNetV2 arhitekture, u nadi da ćemo dobiti još bolje rezultate. Autori su u ovom radu za fino podešavanje ResNet arhitekture koristili SGD optimizator, sa kojim su dobili najbolje rezultate. Po uzoru na ovaj rad, mi smo takođe eksperimentisali is a SGD optimization funkcijom.

U istraživanju koje su sprovedi Alexander[4] i dr. analizirana je efikasnost dubokih konvolutivnih neuronskih mreža za klasifikaciju životinjskih vrsta nad skupovima podataka različitog “kvaliteta”. Korišćen je Snapshot Serengeti (SSE)[5] skup podataka, prikupljen pomoću 225 live kamera u Serengeti Nacionalnom Parku i sadrži 1,2 miliona slika životinja (ručno anotiranih) i obuhvata 48 vrsta od kojih je klasifikacija rađena na 26. Testirani skupovi podataka obuhvataju: neobrađene slike sa neuravnoteženim klasama, neobrađene slike sa izbalansiranim klasama, slike sa životinjama u prvom planu, i ručno segmentisane slike. Sa prvim skupom metode se postigle tačnosti između 35,4% i 60,4%. U drugom skupu tačnost je bila oko 70%, dok je u trećem postizala tačnost od oko 85%. Najbolji rezultati postignuti su na ručno segmentisanim slikama u balansiranom skupu, gde je tačnost dostigla do 98,1%. Autor rada navodi slične probleme u prikupljanju fotografija sa kojima smo se i mi susreli, shodno rezultatima u ovom radu, posebnu pažnju smo posvetili “čišćenju” slika u našem skupu podataka. Takođe, naglašena je važnost balansiranih i raznovrsnih skupova podataka, zbog čega smo eksperimentisali sa tehnikama RandomOversampling i augmentacije slika.

Pored toga, autor ističe da su dublje konvolutivne arhitekture pokazale bolje rezultate, što smo takođe istražili.

U istraživanju Nguyen i dr.[6] dali su okvir za automatsko prepoznavanje životinja u divljini koristeći duboke konvolutivne neuronske mreže. Korišćeni skup podataka preuzet je iz projekta Wildlife Spotter, koji sadrži slike snimljene kamerama-zamkama, koristi se za klasifikaciju tri vrste: ptica, pacov (veliki azijski pacov) i bandicoot. Za augmentaciju slika su korišćene tehnike poput uveličavanja i kropovanja. Autori su izabrali modele AlexNet, VGG-16, i ResNet-50, trenirajući ih od nule i koristeći fine-tuning. Za sve modele tačnost je bila između 89,16% i 90,4%, ali od toga najbolji rezultat je imao ResNet. I ovaj rad ima vrlo sličnu temu i metodologiju kao naše istraživanje pa ćemo shodno tome obratiti posebnu pažnju na ResNet i način na koji je sa njim eksperimentisano u ovom radu, uključujući fino podešavanje i arhitekturu koju su autori kreirali.

## III. SKUP PODATAKA

U ovom istraživanju koristimo dataset koji smo samostalno prikupile pomoću live kamera u zoološkim vrtovima, kao i kamera postavljenih u prirodnim staništima životinja.

### A. Opis skupa podataka

Skup podataka koji koristimo u ovom istraživanju sadrži ukupno 2967 fotografija u jpg formatu kategorizovanih u 3 klase:

- 952 fotografije pande
- 998 fotografija beloglavih orlova
- 978 fotografija belog (polarnog) medveda

Jedan deo fotografija je prikupljen korišćenjem alata VLC, koji je omogućio podešavanje intervala za automatsko snimanje ekrana. Drugi deo fotografija je prikupljen manuelno, snimanjem ekrana sa live streamova dostupnih na YouTube-u u trenucima kada su životinje bile jasno vidljive i u fokusu.



Slika 1: Pikaz beloglavog orla u skupu podataka



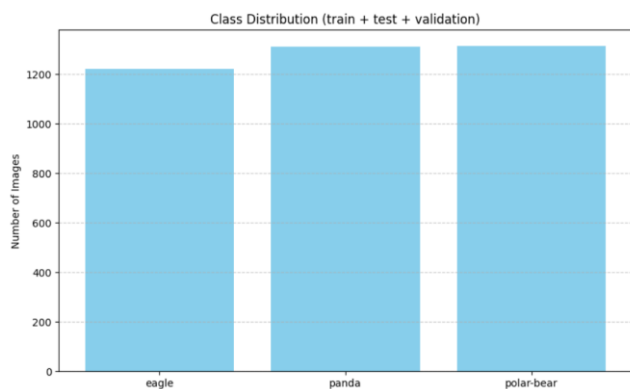
Slika 2: Pikaz pande u skupu podataka



Slika 3: Pikaz polarnog medveda u skupu podataka

Tokom prikupljanja podataka, suočili smo se sa nekoliko izazova. U zoološkim vrtovima, ograničen prostor i ograničeno kretanje životinja unutar kaveza otežali su prikupljanje raznovrsnih slika. Takođe, životinje su se često skrivale van dometa kamere (posebno u toku noći). Sa druge strane, kod live kamere koje su smeštene u prirodnom staništu životinja, pojavljivale su se nejasne slike zbog stalnog kretanja životinja i njihove udaljenosti od kamere. Problem su predstavljale i slike gde bi se više životinja našlo ispred kamere često u položaju neadekvatnom za detekciju, neretko na kameri bi se pojavile i vrste kojima se nismo bavili u ovom projektu. Kako bismo prevazišli ove probleme, značajan deo slika je ručno screenshotovan sa live snimaka.

Za anotaciju životinja koristili smo alat Roboflow, pomoću koga smo ručno anotirali slike. Prazni kadrovi i slike lošeg kvaliteta su eliminisani, dok su preostale slike sa jasno vidljivim životinjama obeležene kao panda, orao ili polarni medved. Nakon toga, slike su se isečene po koordinatama bounding boxova i na taj način je kreiran konačan dataset za obuku modela.



Slika 4: Pikaz distribucije po klasama za ceo skup podataka

S obzirom da se na nekim slikama nalazilo više od jedne životinje, dobili smo blago neizbalansiran konačan skup podataka sa 1314 slika polarnog medveda, 1309 slika pande i 1220 slika beloglavog orla.

## B. Predprocesiranje podataka

Da bismo proširili skup podataka i smanjili rizik od pretreniranja mreže, primenili smo tehnike augmentacije slika, uključujući rotaciju, horizontalno i vertikalno preslikavanje. Eksperimentisali smo i sa drugim tehnikama augmentacije, ali su ove 3 pokazale najbolje rezultate u skladu sa našim skupom podataka. Slike smo prilagodili uniformnom formatu od 224x224 piksela. Skup podataka, koji obuhvata sve slike, nasumično je podeljen u sledećim proporcijama:

- Obuka: 65%
- Validacija: 15%
- Testiranje: 20%

Ova podela omogućila je dovoljno veliki skup podataka za obuku modela, kao i za kontinuirano prilagođavanje na osnovu rezultata nad validacionim skupom podataka.

## IV. METODOLOGIJE

U daljem tekstu dat je pregled odabranih arhitektura konvolutivnih neuronskih mreža, koje su fino podešene (eng. fine-tuning) za detekciju ugroženih životinjskih vrsta.

### A. InceptionV3

InceptionV3 je konvolutivna arhitektura koju je Google predstavio 2015. godine kao naslednika InceptionV1 i InceptionV2. Ova mreža koristi kombinaciju konvolucionih, pooling i inception modula, pri čemu svaki modul primenjuje filtre različitih dimenzija kako bi identifikovao detalje na različitim skalama. Ova struktura omogućava paralelnu obradu informacija kroz više slojeva, čime se postiže šira arhitektura mreže koja poboljšava efikasnost učenja i analize složenih obrazaca u slikama, bez potrebe za povećanjem dubine mreže.[7]

### B. VGG19

VGG19 je arhitektura konvolutivnih neuronskih mreža trenirana na ImageNet[8] skupu podataka. Sastoji se od 19 slojeva s težinskim parametrima, od kojih su 16 konvolutivni slojevi, a preostala tri su potpuno povezana sloja (eng. fully connected). Ulaz u model čine slike dimenzija 224 x 224 piksela. Konvolutivni slojevi koriste male filtere dimenzija 3 x 3, uz primenu ReLU aktivacione funkcije posle svakog sloja. Model sadrži i pet max-pooling slojeva, a završava se s tri potpuno povezana sloja i softmax slojem za konačnu klasifikaciju. [9]

### C. ResNet50

ResNet50 je duboka konvolutivna mreža, trenirana na ImageNet[8] skupu podataka. Ova arhitektura koristi rezidualne veze kako bi prevazišla problem nestajanja gradijenta u vrlo dubokim mrežama. Model se sastoji od 50 slojeva, uključujući konvolute, pooling i potpuno povezane slojeve, omogućavajući efikasno treniranje čak i kod složenih zadataka klasifikacije.[10]

### D. AlexNet

AlexNet ima jednostavnu arhitekturu u poređenju sa drugim modelima. Sastoji se od 8 slojeva, uključujući 5 konvolutivnih i 3 potpuno povezana sloja, uz 3 dodatna max-pooling sloja. Konvolutivni slojevi koriste ReLU aktivacionu funkciju. Ulazne slike su dimenzija 224 x 224 piksela, kao i kod ostalih.[12]

### E. MobileNetV2

MobileNetV2 je lagani model neuronske mreže dizajniran za mobilne uređaje i sisteme sa ograničenim resursima. Sastoji se od konvolutivnih slojeva, praćenih depthwise separable convolutions. Takođe uvodi inverzne rezidualne blokove i linearne bottleneck slojeve. Ove komponente zajedno pomažu da se smanji broj parametara i računskih operacija koje model treba da izvede, dok i dalje omogućavaju modelu da prepozna složene karakteristike[13]

### F. Evaluacija modela

Za evaluaciju modela smo koristili tačnost, preciznost, odziv i F1 rezultat.

#### 1) Tačnost (accuracy)

predstavlja odnos između broja tačno klasifikovanih primera i ukupnog broja primera u skupu podataka.[11] Tačnost se izražava kao procenat svih ispravnih klasifikacija u odnosu na ukupan broj klasifikacija i može se izračunati kao:

$$\text{Accuracy} = \frac{\text{correct classifications}}{\text{total classifications}} = \frac{TP + TN}{TP + TN + FP + FN}$$

gde TP predstavlja tačne pozitivne slučajeve, TN tačne negativne slučajeve, FP lažne pozitivne slučajeve, a FN lažne negativne slučajeve.

#### 2) Preciznost (Precision)

Preciznost meri proporciju tačnih pozitivnih predikcija među svim pozitivnim predikcijama koje je model napravio. Računa se kao odnos tačnih pozitivnih predikcija prema zbiru tačnih pozitivnih i lažnih pozitivnih predikcija. Preciznost pokazuje koliko od predviđenih pozitivnih primera su zapravo pozitivni.[11]

$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

#### 3) Odziv (recall, sensitivity)

Odziv meri proporciju tačnih pozitivnih predikcija među svim stvarnim pozitivnim instancama u skupu podataka. Računa se kao odnos tačnih pozitivnih predikcija prema zbiru tačnih pozitivnih i lažnih negativnih predikcija. Osetljivost pokazuje koliko od stvarnih pozitivnih primera model uspeva da pravilno identifikuje. [11]

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

#### 4) F1 score

Predstavlja harmonijsku sredinu preciznosti i odziva. F1- score uzima u obzir i lažno pozitivne i lažno negativne predikcije, što ga čini korisnom metrikom za neuravnotežene skupove podataka gde preciznost i odziv mogu biti iskrivljeni. [11]

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

### F. Treniranje modela

U ovom projektu rađeno je više eksperimenata za svaku navedenu arhitekturu, koji su uključivali



testiranje različitih optimizatora, kao i različite kombinacije hiperparametara (pri čemu su nam radovi iz literature služili kao osnova za početno pogađanje). Takođe smo ispitivali kako primena dropout tehnike i ranog zaustavljanja (eng. early stopping) utiču na modele. S obzirom na blago neizbalansiran skup isprobali smo i random oversampling. Svaki model smo trenirali koristeći dva različita pristupa kako bismo videli koji će se bolje pokazati za naš zadatak.

1. Prvi pristup je obuhvatao korišćenje unapred obučениh modela sa težinama obučениm nad ImageNet skupom podataka kao osnovu, pri čemu su gornji klasifikacioni slojevi isključeni. Osnovni slojevi koji funkcionišu kao ekstraktori osnovnih osobina su zamrznuti, dok smo drugi deo konvolutivnih slojeva trenirali ponovo sa našim podacima, kako bismo bolje prilagodili modele za ekstrakciju specifičnih osobina u našem skupu podataka. Za svaku arhitekturu posebno, eksperimentisali smo i tražili optimalan odnos slojeva koje ćemo zamrznuti i koje ćemo trenirati sa našim podacima. U ovu svrhu koristili smo radove koji su se bavili sličnom temom i eksperimentisali sa brojem slojeva koji su autori navodili kao optimalan za rešavanje ovog problema. Takođe eksperimentisali smo i sa primenom *random search* algoritma za optimizaciju hiperparametara. Na vrhu svake arhitekture dodali smo klasifikacione slojeve koji su uključivali potpuno povezani sloj (Dense) sa ReLU aktivacionom funkcijom, dropout sloj sa 'rate' parametrom posebno podešenim za svaku arhitekturu, kako bismo smanjili pretreniranje mreže (eng. overfitting), i još jedan potpuno povezani sloj sa softmax aktivacionom funkcijom za konačnu klasifikaciju. Takođe, implementirani su mehanizmi ranog zaustavljanja (eng. early stopping) sa individualno podešenim 'patience' parametrom za svaku arhitekturu, kako bi se dodatno poboljšala efikasnost modela i sprečio overfitting.

2. Drugi pristup je podrazumevao takođe korišćenje unapred obučениh modela sa težinama obučениm na ImageNet skupu, ali sa kompletno zamrznutim konvolutivnim delom arhitekture (težine nisu ažurirane tokom

obuke). Na ovakvu osnovu dodali smo klasifikator koji odgovara opisu iz prethodnog pristupa koji smo na isti način trenirali. Takođe smo eksperimentisali sa early stopping i dropout mehanizmima za prevenciju overfitting-a.

U eksperimentima korišćeni su *Adam* i *SGD* optimizatori, kao i *Categorical Cross-entropy* funkcija gubitka.

## V. REZULTATI

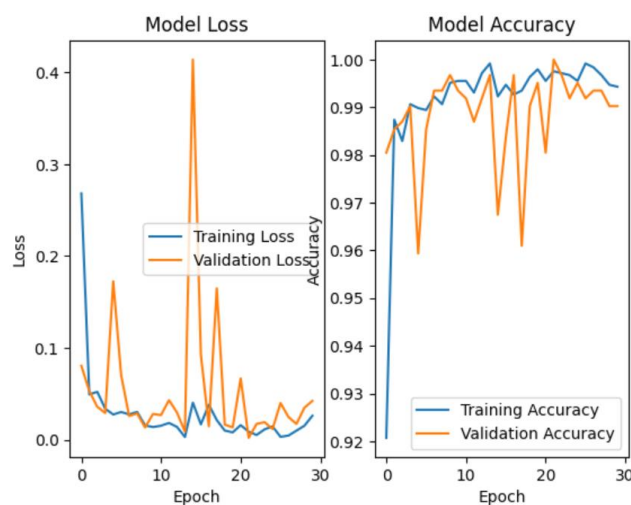
U narednom poglavlju dat je upoređan pregled rezultata dobijenih finim podešavanjem najpoznatijih arhitektura konvolutivnih neuronskih mreža. Prikazan je svaki eksperiment koji smo radili, gde je za svaku arhitekturu eksperiment koji se najbolje pokazao detaljnije prikazan matricom konfuzije i graphicima.

### A. InceptionV3

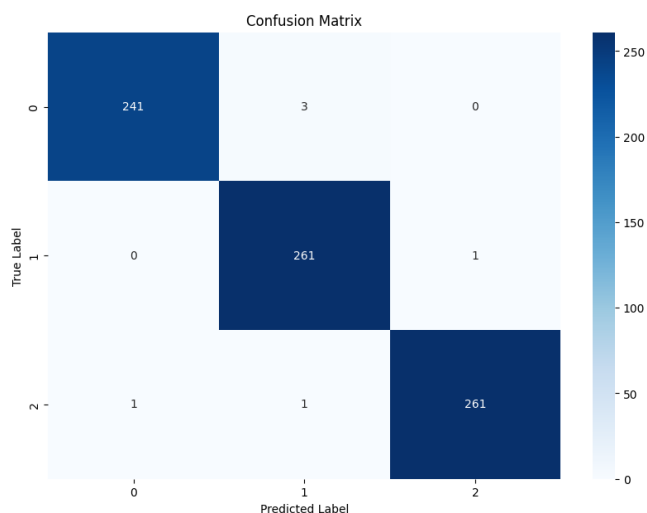
U tabeli 1 prikazani su rezultati eksperimenta 1 za InceptionV3 arhitekturu gde je primenjen pristup 1(navedenom sekciji *treniranje modela*), korišćen je random search algoritam za optimizaciju hiperparametara.

	preciznost	odziv	f1
beloglavi orao	1.00	0.99	0.99
panda	1.00	0.99	0.99
polarni medved	0.98	1.00	0.99
tačnost		99,21%	

Tabela 1: Prikaz rezultata eksperimentalnog za InceptionV3



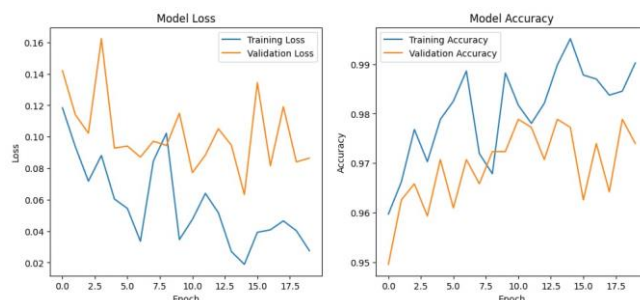
Slika 4: Grafički prikaz rezultata



Slika 5: Matrica konfuzije

	preciznost	odziv	f1
beloglavi orao	1.00	0.96	0.98
panda	0.97	0.98	0.98
polarni medved	0.97	0.99	0.98
tačnost	97,65%		

Tabela 4: Pikaz rezultata eksperimenta1 za VGG19

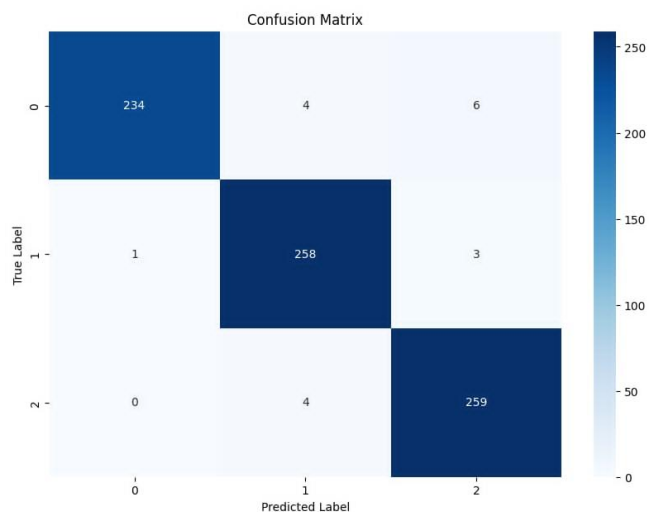


Slika 6: Grafički prikaz rezultata

U tabeli 2 prikazan je rezultat eksperimenta 2, gde je takođe primenjen pristup 1 (navedenom sekciji *treniranje modela*), ali su korišćeni hiperparametri predloženi u pročitanoj literaturi.

	preciznost	odziv	f1
beloglavi orao	1.00	0.99	0.99
panda	0.98	0.99	0.98
polarni medved	0.98	0.98	0.98
tačnost	98,56%		

Tabela 2: Pikaz rezultata eksperimenta2 za InceptionV3



Slika 7: Matrica konfuzije

U tabeli 3 prikazani su rezultati eksperimenta 3 rađenog po pristupu 2 (navedenom sekciji *treniranje modela*), sa primenom oversampling metode nad klasom *eagle*.

	preciznost	odziv	f1
beloglavi orao	1.00	1.00	1.00
panda	0.97	0.99	0.98
polarni medved	1.00	0.98	0.99
tačnost	98,82%		

Tabela 3: Pikaz rezultata eksperimenta3 za InceptionV3

### B.VGG19

U tabeli 4 prikazan je rezultat modela koji je rađen po pristupu 2 (navedenom sekciji *treniranje modela*), gde je za optimizaciju hiperparametara korišćen *random search* i kao optimizator *Adam*.

U tabeli 5 prikazan je rezultat modela koji je rađen po pristupu 1 (navedenom sekciji *treniranje modela*), gde je za optimizaciju hiperparametara takođe korišćen *random search* i kao optimizator *Adam*.

	preciznost	odziv	f1
beloglavi orao	0.97	0.92	0.95
panda	0.92	0.95	0.93
polarni medved	0.94	0.97	0.95
tačnost	94,53%		

Tabela 5: Pikaz rezultata eksperimenta2 za VGG19

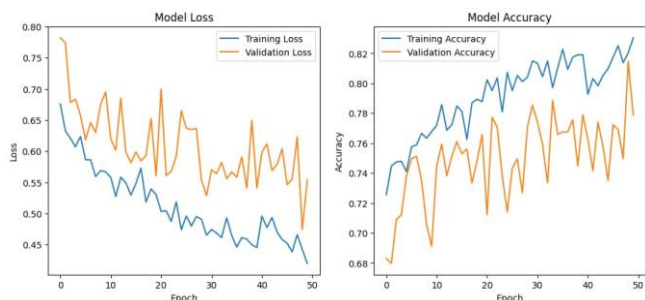
### C.ResNet50

U tabeli 6 je prikazan rezultat modela koji je rađen po pristupu 1 (navedenom sekciji *treniranje modela*), gde

je za optimizaciju hiperparametara takođe korišćen *random search* i kao optimizator *Adam*.

	preciznost	odziv	f1
beloglavi orao	0.88	0.87	0.88
panda	0.70	0.92	0.80
polarni medved	0.92	0.63	0.75
tačnost	82,57%		

Tabela 6: Pikaz rezultata eksperimenta1 za ResNet50



Slika 8: Grafički prikaz rezultata

je za optimizaciju parametara korišćen *random search* i kao optimizator je korišćen *SGD*.

	preciznost	odziv	f1
beloglavi orao	0.36	0.33	0.34
panda	0.52	0.67	0.58
polarni medved	0.80	0.62	0.70
tačnost	56,95%		

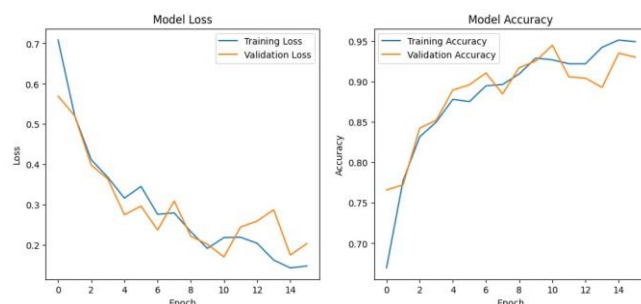
Tabela 8: Pikaz rezultata eksperimenta3 za ResNet50

#### D.AlexNet

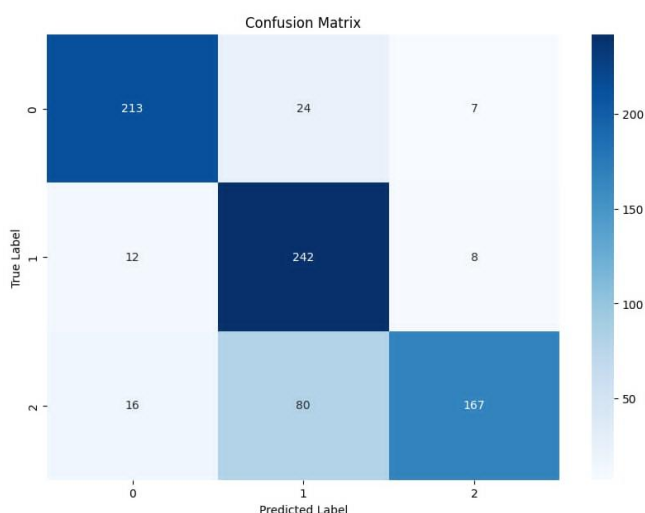
U tabeli 9 prikazani su rezultati našeg pokušaja da konsultujući relevantnu literature implementiramo AlexNet.

	preciznost	odziv	f1
beloglavi orao	0.97	0.91	0.94
panda	0.92	0.97	0.94
polarni medved	0.93	0.94	0.94
tačnost	93,88%		

Tabela 9: Pikaz rezultata eksperimenta1 za AlexNet



Slika 10: Grafički prikaz rezultata



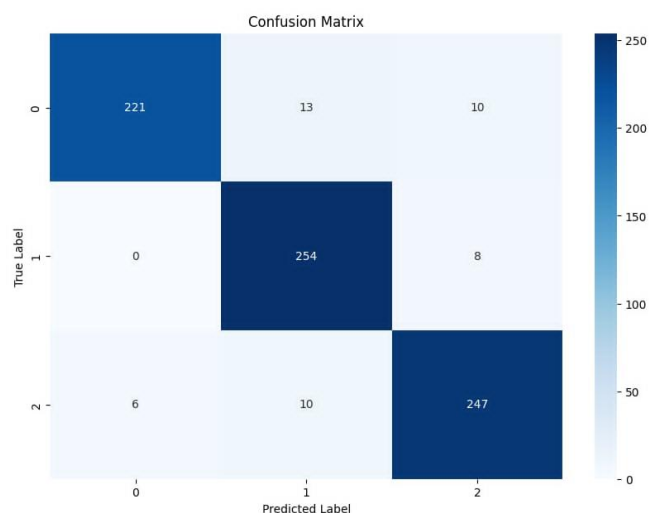
Slika 9: Matrica konfuzije

U tabeli 7 je prikazan rezultat modela koji je rađen po pristupu 2 (navedenom sekciji *treniranje modela*), gde su hiperparametri preuzeti iz literature i kao optimizator je korišćen *SGD*.

	preciznost	odziv	f1
beloglavi orao	0.44	0.95	0.60
panda	0.87	0.32	0.47
polarni medved	0.84	0.43	0.57
tačnost	58,12%		

Tabela 7: Pikaz rezultata eksperimenta2 za ResNet50

U tabeli 8 je prikazan rezultat modela koji je rađen po pristupu 2 (navedenom sekciji *treniranje modela*), gde



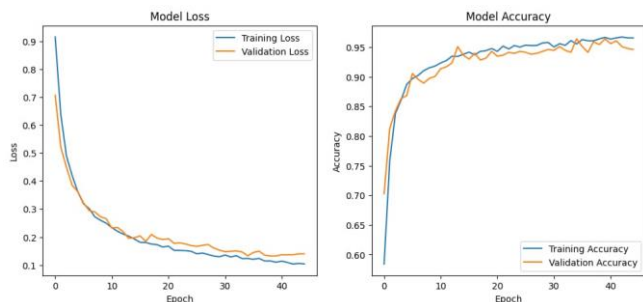
Slika 11: Matrica konfuzije

### E.MobileNetV2

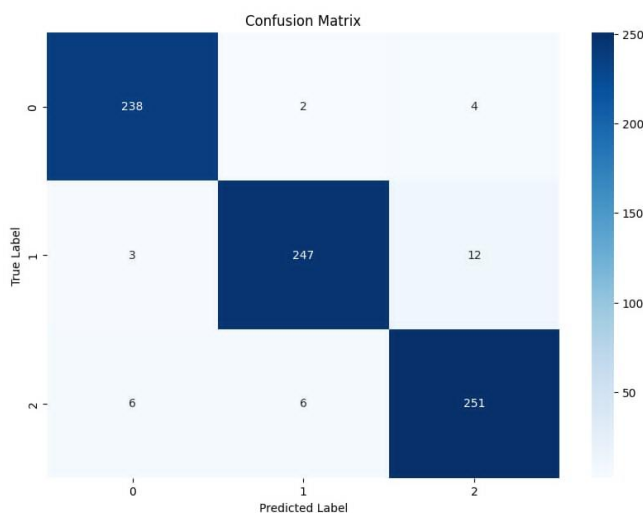
U tabeli 8 je prikazan rezultat gde su svi konvolucionirani slojevi zamrznuti. Kao optimizator je korišćen *Adam*.

	preciznost	odziv	f1
beloglavi orao	0.96	0.98	0.97
panda	0.97	0.94	0.96
polarni medved	0.94	0.95	0.95
tačnost	95,70%		

Tabela 9: Prikaz rezultata eksperimenta 1 za AlexNet



Slika 12: Grafički prikaz rezultata



Slika 13: Matrica konfuzije

### ZAKLJUČAK I BUDUĆI RAD

U ovom radu dat je uporedni pregled različitih arhitektura konvolutivnih neuronskih mreža za klasifikovanje ugroženih vrsta životinja na 3 klase (panda, polarni medved, beloglavi orao) nad skupom podataka koji je sakupljen koristeći live kamere iz zooloških vrtova kao i kamere postavljene u prirodnom staništu životinja. Rezultati našeg rada pokazuju da se

nad našim skupom podataka kao optimalan pokazao InceptionV3 model, sa tačnošću od 99,21%. Ovaj rezultat dobijen je u eksperimentu gde je deo konvolutivnih slojeva treniran sa našim podacima, klasifikator je takođe treniran, a hiperparametri su optimizovani random search algoritmom. Najlošije se pokazala arhitektura ResNet sa 82,57%.

U istraživanju "Real-Time Identification of Animals found in Domestic Areas of Europe" [3], MobileNet je naveden kao najbolje performantan model, u našem radu takođe je pokazao visok procenat tačnosti 95.70%, što se pokazalo kao bolji rezultat na našem skupu podataka. Sa druge strane po uzoru na ovaj rad isprobali smo ResNet sa optimizatorom SGD, sa potpuno zamrznutim konvolutivnim slojevima kao i sa dodatnim treniranjem konvolutivnih slojeva i dobili smo značajno lošije rezultate 58,12% i 56,95%.

U istraživanju koje su sprovedili Alexander[4] i dr., ističe se značaj dobrog skupa podataka, pa smo po uzoru na ovaj rad, potrošili mnogo vremena da skup podataka maksimalno poboljšamo. Takođe koristili smo random oversampling da bismo izbalansirali klase. Random oversampling nije dao revolucionarne promene u rezultatima.

U istraživanju Nguyen i dr.[6] autori posebnu pažnju posvećuju ResNet arhitekturi, po uzoru na ovaj rad pokušali smo dobiti barem približne rezultate, ali u tome nismo uspeli.

Pravac za dalji rad svakako uključuje povećanje broja klasa i generalno povećanje broja slika u kupu podataka, gde bi sa većim i raznovrsnijim skupom podataka imalo smisla proširiti arhitekture koje smo trenirali i možda trenirati na hadrverski sposobnijim računarima gde bismo očekivali bolji rezultat.

### LITERATURA

- [1] <https://www.gvi.ie/blog/smb-how-endangered-species-affect-the-environment/#:~:text=When%20these%20species%20disappear%2C%20the.%2C%20climate%20change%2C%20and%20disease.>
- [2] <https://www.amnh.org/explore/ology/earth/ask-a-scientist-about-our-environment/why-are-so-many-animals-endangered>
- [3] Sorin Liviu Jurj, Flavius Opritoiu, Mircea Vladutiu, "Real-time identification of animals found in domestic areas of Europe," Proc. SPIE 11433, Twelfth International Conference on Machine Vision (ICMV 2019), 1143313 (31 January 2020); <https://doi.org/10.1117/12.2556376>
- [4] Gomez-Villa, Alex & Salazar, Augusto & Vargas-Bonilla, J.. (2016). Towards Automatic Wild Animal Monitoring: Identification of Animal



Species in Camera-trap Images using Very Deep Convolutional Neural Networks. Ecological Informatics. 41. 10.1016/j.ecoinf.2017.07.004.

- [5] Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., & Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific data*, 2(1), 1-14.
- [6] Nguyen, Hung & Maclagan, Sarah & Nguyen, Tu & Nguyen, Thin & Flemons, Paul & Andrews, Kylie & Ritchie, Euan & Phung, Dinh. (2017). Animal Recognition and Identification with Deep Convolutional
- [7] <https://iq.opengenus.org/inception-v3-model-architecture/>
- [8] <https://www.image-net.org/>
- [9] <https://www.geeksforgeeks.org/vgg-net-architecture-explained/>
- [10] <https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f>
- [11] <https://medium.com/@nirajan.acharya777/understanding-precision-recall-f1-score-and-support-in-machine-learning-evaluation-7ec935e8512e>
- [12] <https://medium.com/@siddheshb008/alexnet-architecture-explained-b6240c528bd5>
- [13] <https://www.analyticsvidhya.com/blog/2023/12/what-is-mobilenetv2/>