Identifikacija vrsta za potrebe biomonitoringa korišćenjem konvolucionih neuronskih mreža i dubokog učenja

Aleksandar Milosavljević, Đurađ Milošević i Bratislav Predić

Apstrakt—Akvatični i bentonski insekti drugi makrobeskičmenjaci uglavnom se koriste kao bioindikatori ekološkog stanja slatkih voda. Međutim, skup i dugotrajan postupak identifikacije vrsta predstavlja jednu od ključnih prepreka za pouzdan biomonitoring akvatičnih ekosistema. U radu je predložen metod za identifikaciju vrsta zasnovan na dubokom učenju čija je evaluacija obavljena na nekoliko javno dostupnih skupova podatak (FIN-Benthic, STONEFLY9 i EPT29) kao i na sopstvenom CHIRO10 skupu podataka. Predloženi metod se zasniva na tri tehnike dubokog učenja koje se koriste za poboljšanje robusnosti kada se za obučavanje koristi relativno mali skup podataka: preneseno učenje (eng. transfer learning), proširivanje podataka (eng. data augmentation), kao i odbacivanje (eng. dropout). Evaluacija modela je vršena korišćenjem ulaznih slika dimenzija 256x256 piksela gde je 50% slika korišćeno za treniranje, 20% za validaciju, a 30% za testiranje. Dobijeni rezultati pokazuju značajno poboljšanje u odnosu na tradicionalne metode koje su originalno korišćenje i potvrđuju da postoji značajan dobitak kada postoji veći broj slika po uzorku.

Ključne reči—Duboko učenje; konvolucione neuronske mreže; klasifikacija slika; preneseno učenje; proširivanje podataka; biomonitoring; akvatični insekti.

I. Uvod

Raznolikost gena, vrsta i ekosistema opada globalno brže nego ikad pre u ljudskoj istoriji [1]. Akvatični ekosistemi prikazuju među najvećim stopama opadanja sa alarmantnim gubitkom biodiverziteta, te je potreba za isplativim alatima za biomonitoring time i veća.

Tradicionalni pristup morfološkoj identifikaciji u biomonitoringu pretpostavlja korišćenje što šire taksonomske rezolucije [2]. Međutim, identifikacija makrobeskičmenjaka na osnovu morfoloških karakteristika može biti problematična jer broj pogrešno klasifikovanih vrsta raste sa povećanjem taksonomske rezolucije. S druge strane, morfološka identifikacija je vremenski zahtevan i skup proces, te je kao takav neprimenjiv za rutinski biomonitoring [3], [4]. Pobrojani nedostaci tradicionalnog biomonitoringa stvaraju potrebu za razvojem alternativnih pristupa za obradu uzoraka. Nedavni napredak u računarskom vidu u pogledu klasifikacije slika korišćenjem konvolucionih neuronskih mreža (KNM) i dubokog učenja otvorili su put pouzdanoj automatizaciji procesa identifikacije.

Klasifikacija slika u računarskom vidu je problem gde se na osnovu skupa slika određenih kategorija izgrađuje model sposoban da predvidi kategoriju, sa određenom tačnošću, za novo zadatu sliku. Problem nije jednostavan pošto slike mogu da sadrže različite varijacije. Tipičan način za rešavanje ovog problema je pristup zasnovan na podacima [5]. Umesto da pokušavamo da opišemo svaku od klasa koju želimo da identifikujemo, koristi se veliki broj slika za svaku od klasa kako bi se izgradio model (klasifikator) koji je u stanju da ih identifikuje. Tradicionalni pristup klasifikaciji je podrazumevao ručno projektovanje različitih ekstraktora vektora obeležja (eng. feature) na osnovu kojih bi se obučavao klasifikator. Međutim, veliki napredak se javio pojavom dubokih KNM i kompletnog obučavanja (eng. end-to-end learning). Duboke KNM imaju sposobnost da izgrađuju hijerarhiju obeležja kroz različite konvolucione slojeve koje poseduju, te kao takve postaju nezamenjive u ulozi ekstraktora obeležja koji uči iz podataka.

Rad je organizovan na sledeći način: poglavlje 2 sadrži opis srodnih radova koji se bave identifikacijom vrsta za potrebe akvatičnog biomonitoringa, kao i javnih skupova podataka koji su korišćeni u ovom radu. U poglavlju 3 dat je opis predloženog metoda uključujući i detalje konkretne implementacije. Ostvareni rezultati i odgovarajuća diskusija dati su u poglavlju 4. Konačno, u zaključku (poglavlje 5) je dat rezime i naznačeni su pravci daljeg istraživanja.

II. SRODNI RADOVI I SKUPOVI PODATAKA

Problem automatizovane taksonomske identifikacije bentonskih makrobeskičmenjaka zasnovane na klasifikaciji slika je obrađen u radovima [6]–[11]. Lytle i dr. [6] su razvili jedan od prvih sistema ove vrste. Njihov sistem BugID koristi Scale Invariant Feature Transform (SIFT) deskriptore [12] u kombinaciji sa Random Forest (RF) klasifikatorom i na STONEFLY9 skupu podataka [6], [13] postiže tačnu klasifikaciju u 95,5% slučajeva.

Larios i dr. [7] su za klasifikaciju koristili tri različita ekstraktora obeležja: Histogram of Oriented Gradients (HOG), Beam Angle Statistics (BAS) i SIFT specijalizovanih za različite delove prostora obeležja. Testiranje predloženog metoda je vršeno na EPT29 skupu podataka koji sadrži 4722

Aleksandar Milosavljević – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: aleksandar.milosavljevic@elfak.ni.ac.rs).

Đurađ Milošević – Prirodno-matematički fakultet, Univerzitet u Nišu, Višegradska 33, 18000 Niš, Srbija (e-mail: djuradj@pmf.ni.ac.rs).

Bratislav Predić – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: bratislav.predic@elfak.ni.ac.rs).

slike 29 akvatičnih vrsta koje pripadaju redovima *Ephemeroptera*, *Plecoptera* i *Trichoptera* koji se najčešće koriste za procenu stanja akvatičnih ekosistema. Najbolji rezultat od 88,06% je ostvaren korišćenjem Spatial-Pyramid Kernel Support Vector Machines (SVM) klasifikatora u kombinaciji sa stratifikovanom 3-kratnom unakrsnom validacijom.

Kiranyaz i dr. [8] su predložili još jedan klasičan, tj. pre dubokog učenja, pristup morfološkoj identifikaciji makrobeskičmenjaka. Skup podataka korišćen u njihovom istraživanju se sastojao od 1350 slika 8 taksonomskih vrsta. Skup podataka nije javni, te ga nismo koristili u našim eksperimentima. Za ekstrakciju obeležja koristili su ImageJ softver koji generiše 15-to dimenzione vektore obeležja na osnovu kojih su obučavani različiti klasifikatori: SVM, Bayesian Classifiers (BC) i dve neuronske mreže: Multi-Layer Perceptron (MLP) i Radial Basis Function Network (RBFN). Najbolji rezultat i grešku od 3,57% je zabeležio MLP model.

Joutsijoki i dr. [9] su koristili isti skup podataka i metodologiju za ekstrakciju obeležja kako bi ispitali primenjivost veštačkih neuronskih mreža za identifikaciju makrobeskičmenjaka. Eksperimenti su vršeni sa tri arhitekture: MLP, Probabilistic Neural Network (PNN) i RBFN, a MLP se ponovo pokazao najbolje ostvarujući tačnost od 95,3%. Treba napomenuti da je ovde korišćena drugačija metodologija podele skupa, tako da je 80% korišćeno za treniranje, a po 10% za validaciju i testiranje.

Raitoharju i dr. [10] su kreirali javno dostupan skup FINmetodologija Benthic sa ciljem testiranja različitih klasifikacije vizuelno sličnih vrsta akvatičnih makrobeskičmenjaka. FIN-Benthic sadrži 3 podskupa sa 64, 29 i 9 vrsta. Broj slika po klasi varira između 7 i 577. U radu je predložena i metodologija za slikanje uzoraka iz različitih uglova korišćenjem dva kamere. Ovo je značajno za proces identifikacije jer omogućava kombinovanje dve nezavisne predikcije pri određivanju vrste uzorka. Još jedna bitna karakteristika FIN-Benthic skupa je da sadrži 10 eksplicitnih podela na trening (50%), validacione (20%) i test (30%) podskupove. Eksperimenti koje su autori sproveli su prvi put iskoristili KNM (AlexNet [14]), kako za ekstrakciju obeležja u kombinaciji sa SVM klasifikatorom, tako i za kompletno obučavanje koje je dalo i nešto bolje rezultate. Najbolje tačnosti po tri definisana podskupa su 75,74% (podskup 1), 81,04% (podskup 2) i 90,14% (podskup 3).

Konačno, rad [11] predstavlja naš doprinos u oblasti automatske identifikacije vrsta larvi hironomida (*Diptera: Chironomidae*). U radu je predstavljen kreirani skup podataka koji sadrži 1846 slika i sastoji se od 10 morfološki vrlo sličnih vrsta iz istog roda ili podfamilije (vidi Sliku 1). U radu je predstavljen metod zasnovan na korišćenju ResNet-50 [15] KNM prethodno obučene na ImageNet [16] skupu koji je dao tačnost od 99,465% na validacionom skupu koji je činio 20% ukupnih podataka.

U Tabeli I prikazani su detalji skupova podataka koji su korišćeni za potrebe istraživanja predstavljenih u ovom radu. Ilustracija CHIRO10 skupa je data na slici 1.

TABELA I Detalji korišćenih skupova podataka

Naziv skupa	Pod- skup	Br. klasa	Br. uzoraka	Br. slika	Br. slika po uzorku	Br. slika po klasi
CHIRO10	1	10	1846	1846	1	79–207 (~186)
	2	5				199-684
	3	2				199–1647
FIN-Benthic	1	64	7705	15074	1-2 (~2)	7-577
	2	29	6038	11832	1–2 (~2)	230–577 (~408)
	3	9	1692	3272	1-2 (~2)	322-395
STONEFLY9	-	9	774	3845	1–5 (~5)	119-532
EPT29	-	28	1608	4794	1–4 (~3)	27–366 (~171)



Sl. 1. Ilustracija CHIRO10 skupa podataka.

III. OPIS METODA

Predloženi metod identifikacije vrsta na osnovu slika se zasniva na dubokom učenju, tj. na obučavanju rezidualne

KNM u ulozi klasifikatora. Da bi se ostvarila robusnost u režimu ograničenog broja trening uzoraka korišćene su sledeće tehnike pri projektovanju i obučavanju klasifikatora:

- 1. Preneseno učenje (eng. transfer learning)
- 2. Odbacivanje (eng. dropout)
- 3. Proširivanje podataka (eng. data augmentation)

Preneseno učenje predstavlja osnovnu tehniku koja omogućava korišćenje dubokih KNM za rešavanje problema sa relativno malim skupom podataka. Zasniva se na korišćenju prethodno utrenirane mreže na nekom velikom skupu podataka, kao što je npr. ImageNet skup. Treniranje na ImageNet skupu obezbeđuje da mreža izgradi hijerarhiju različitih obeležja koje se mogu naći na fotografijama generalno i koje je pogodno iskoristiti za klasifikaciju novih fotografija. Da bi takav transfer bio moguć, potrebno je zameniti vršni deo mreže zadužen za klasifikaciju i utrenirati ga koristeći niz obeležja koje identifikuje duboka KNM. Uobičajeno je da se deo KNM ispred klasifikatora naziva enkoder. U zavisnosti od prirode novog skupa ovakav pristup može da bude i sasvim dovoljan. Međutim, u našem slučaju ulazni podaci ne predstavljaju nešto što se tipično nalazi na fotografijama, zbog toga je bilo neophodno izvršiti dvofazno obučavanje. Nakon obučavanja klasifikatora u prvoj fazi, u drugoj fazi je vršeno fino podešavanje cele mreže. Fino podešavanje nije ništa drugo do obučavanja celokupne mreže, kako klasifikatora tako i enkodera. Termin fino se koristi da naznači korišćenje vrlo malih koeficijenata učenja kako bi se što manje narušile polazne vrednosti parametara.

Druga tehnika koja je iskorišćena kako bi se povećala robusnost klasifikatora je odbacivanje [17]. Odgovarajući sloj je dodat nakon enkodera, tako da se u svakom koraku obučavanja odbacuje određen procenat (u našem slučaju 50%) obeležja na osnovu kojih se vrši klasifikacija uzoraka. U fazi testiranja i eksploatacije mreže se uzimaju u obzir svi izlazi, ali se vrši skaliranje vrednosti za odgovarajući procenat odbacivanja. Na ovaj način se postiže u proseku isti nivo izlaza koji smo imali kod treniranja. Efekat koji se postiže primenom odbacivanja je da klasifikator mora da se oslanja na više različitih obeležja pri određivanju kategorije. Na ovaj način se izbegava preterano prilagođavanje modela (eng. overfitting), a samim tim i bolji rezultati na validacionom skupu.

Obučavanje neuronske mreže da klasifikuje slike zahteva nalaženje obeležja koje određuju odgovarajuće klase. Taj proces zahteva veliku količinu trening uzoraka kako bi se izolovale ključne karakteristike klasa i dobio klasifikator otporan na različite varijacije koje na slikama mogu da se jave. Kada imamo manju količinu trening podataka, a koristimo model velikog kapaciteta, dešava se model vrlo brzo "zapamti" sve trening uzorke, ali zato daje loše rezultate na validacionom skupu. Tipično korišćena tehnika koja služi da se ovo izbegne je proširivanje podataka. Proširivanje podataka podrazumeva primenu nasumičnih transformacija nad ulaznim slikama tako da se u svakom trening ciklusu mreži predoči nešto što ranije nije "videla". U zavisnosti od prirode skupa podataka, tipične transformacije uključuju obrtanje (eng. flip), rotaciju, translaciju, skaliranje, zakošavanje, promenu osvetljaja i kontrasta, itd.

A. Arhitektura mreže

Vodeći se prethodno obrazloženim principima, za klasifikaciju je usvojena arhitektura zasnovana na ResNet-50 [15] enkoderu prikazana na slici 2.



Sl. 2. Šematski prikaz korišćene arhitekture zasnovane na ResNet-50 mreži.

Na izbor ResNet-50 mreže utrenirane na ImageNet skupu kao enkodera je uticalo nekoliko faktora: dobri rezultati na ImageNet skupu, veličina mreže u pogledu broja parametara, memorijsko zauzeće u toku treniranja, brzina obučavanja, kao i dostupnost modela u korišćenom programskom okruženju. ResNet arhitektura generalno vrlo često predstavlja dobar inicijalni izbor zbog dobrog odnosa preciznosti i brzine obučavanja.

ResNet-50 enkoder poseduje 23.587.712 parametara, dok na izlazu daje 2048 obeležja. Izlazi se dobijaju korišćenjem globalnog usrednjavanja (eng. *global average pooling*) po izlaznoj mapi obeležja koja za korišćeni ulaz dimenzija 512x512 piksela iznosi 16x16x2048.

Na dobijenih 2048 izlaza iz enkodera se primenjuje odbacivanje sa faktorom 50%. Ovo u praksi znači da se u fazi obučavanja polovina, tj. 1024 nasumično izabranih izlaza postavi na nulu. Tako modifikovani izlazi enkodera se dovode na potpuno povezani sloj sa 3 neurona gde svaki izlaz odgovara jednoj klasi. Sloj poseduje 6.147 parametara koji se obučavaju u prvoj fazi. Na izlaze se primenjuje *softmax* aktivaciona funkcija (1) čime ovaj sloj dobija ulogu klasifikatora obeležja koje daje ResNet-50 enkoder. Izlazi softmax funkcije predstavljaju verovatnoće da je tekući uzorak pripadnik neke od klasa. Ovo praktično znači da pojedinačni izlazi imaju vrednosti iz intervala [0, 1] i da je zbir svih izlaza jednak 1.

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_i}}.$$
 (1)

B. Implementacija i obučavanje mreže

Za implementaciju predložene arhitekture iskorišćen je programski jezik Python i biblioteka Keras [18]. Keras je biblioteka visokog nivoa koja definiše pojednostavljeni interfejs za implementaciju dubokih neuronskih mreža i u našem slučaju se oslanja na TensorFlow [19] biblioteku za realizaciju svih funkcionalnosti.

U okviru *applications* modula Keras poseduje nekoliko dubokih KNM arhitektura istreniranih na ImageNet skupu. Instanciranjem klase *ResNet50*, uz odgovarajuće parametre, dobijamo enkoder za naš model. Na izlaz enkodera se nadovezuje *Dropout* i *Dense* sloj sa *softmax* aktivacijom čime se dobija kompletan model. Kako se u prvoj fazi obučavanja težine ResNet-50 enkodera ne menjaju, potrebno je za sve konvolucione slojeve u enkoderu postaviti atribut *trainable* na *False*.

Kreirani model je kompajliran tako da koristi Adam [20] algoritam (optimizers optimizaciju modul), za sparse categorical crossentropy tip greške (losses modul), dok kao metrika tačnosti koristi se sparse categorical accuracy (metrics modul). Algoritam optimizacije je izabran zbog brze konvergencije, dok su greška za obučavanje i odgovarajuća metrika tačnosti standardni izbor za problem klasifikacije. Varijante ovih funkcija sa prefiksom sparse se koriste kada se klase koje predstavljaju očekivani izlaz zadaju kao celi broj (0, 1 ili 2 u našem slučaju).

Za potrebe proširivanja podataka iskorišćena je Keras ugrađena klasa *ImageDataGenerator* koja se nalazi u *preprocessing* modulu, podmodul *image*. Pri konstruiranju odgovarajućeg generatora slika definišu se opsezi za različite transformacije koje će nasumično primenjivati. U našem slučaju korišćeno je horizontalno i vertikalno obrtanje slike, rotacija do $\pm 90^{\circ}$, translacija do $\pm 15\%$ po oba pravca, promena osvetljaja do $\pm 20\%$, zakošavanje i skaliranje do $\pm 10\%$. Ilustracija je data na slici 3. Proširivanje podataka se vrši samo za trening skup, do za potrebe validacije koriste neizmenjene slike.



Sl. 3. Ilustracija proširivanja podataka. Prva slika (gore-levo) predstavlja ulaz, dok su ostale slike nastale primenom nasumičnih transformacija.

S obzirom da je korišćen generator slika, za obučavanje mreže koristi se metod *fit_generator*. Dodatna kontrola procesa obučavanja je u Kerasu moguća prosleđivanjem liste *callback* objekata. Odgovarajuće klase se nalaze u modulu *callbacks* i u našem slučaju iskorišćene su: *LearningRateScheduler, EarlyStopping, ModelCheckpoint* i *CSVLogger*.

LearningRateScheduler obezbeđuje definisanje proizvoljne funkcije za izmenu koeficijenta brzine obučavanja u zavisnosti od tekuće epohe obučavanja. U našem slučaju, ovaj *callback* je iskorišćen za implementaciju tzv. kosinusnog kaljenja (eng. *cosine annealing*) [21]. Kod kosinusnog kaljenja koeficijent obučavanja se smanjuje po kosinusnoj funkciji od neke inicijalne do neke minimalne vrednosti u toku određenog broja epoha koje su definisane periodom ponavljanja. U prvoj fazi obučavanja je korišćena perioda ponavljanja od 10 epoha, sa inicijalnim koeficijentom obučavanja u rasponu od 10⁻³ do 10⁻⁵, uz smanjivanje 0,7 puta pri svakoj novoj periodi.

Kako ime sugeriše, *EarlyStopping callback* se koristi za ranije zaustavljanje procesa obučavanja ukoliko u određenom broju epoha nema napretka po određenom parametru. U našem slučaju je korišćeno 30 epoha i praćena je tačnost na validacionom skupu.

CSVLogger callback se koristi za snimanje greški i tačnosti nad trening i validacionim skupom u toku procesa obučavanja, a u cilju kasnije vizuelizacija ovog procesa.

Konačno, *ModelCheckpoint* je iskorišćen za snimanje najboljeg rezultata u pogledu tačnosti postignute nad validacionim skupom. Ovako zabeležen model je iskorišćen za narednu fazu obučavanja, tj. fazu finog podešavanja.

Fino podešavanje je vršeno na gotovo identičan način, uz par sitnijih izmena. Nakon učitavanja modela dobijenog iz

prve faze obučavanja, izvršeno je aktiviranje obučavanja za sve slojeve iz ResNet-50 enkodera (*trainable* atribut postavljen na *True*). Da bi se izbegla drastična izmena težina u enkoderu, koeficijent obučavanja se kretao u rasponu od 10⁻⁵ do 10⁻⁷. Konačno, da bi se dobio model koji daje najbolje rezultate na celokupnom skupu podataka, umesto standardnog validacionog, iskorišćen je kompletan neizmenjen skup podataka. Obučavanje je i dalje rađeno sa proširivanjem podataka trening skupa. Odgovarajući rezultati su prezentovani u narednom poglavlju.

IV. REZULTATI I DISKUSIJA

U cilju evaluacije predloženog metoda, za svaki skup podataka, obučili smo 10 modela, po jedan za svaku podelu. Rezultati su dobijeni evaluacijom obučenih modela na odgovarajućem test skupu koji sadrži 30% svih uzoraka iz skupa podataka. Treba naglasiti da slike iz trening skupova nisu ni na koji način korišćene u procesu obučavanja modela, te na taj način predstavljaju realne očekivane performanse modela na nepoznatim podacima.

S obzirom da svi skupovi podataka, sem našeg CHIRO10 skupa, poseduju više slika po uzorku, kao metrika tačnosti je izabrana tačnost na nivou uzorka, a ne slike. Da bi uporedili kako različiti metodi akvizicije više slika po uzorku utiču na rezultat, računali smo i tačnost na nivou slike. Tačnost na nivou uzorka je računata usrednjavanjem predikcija, tj. raspodela verovatnoća, dobijenih za svaku od slika uzorka, te dodelom najverovatnije klase. Na primer, ukoliko imamo 3 klase i 2 slike za uzorak, ukoliko model prediktuje (60%, 10% i 30%) za prvu sliku i (10%, 40% i 50%) za drugu sliku, tada se dobija srednja predikcija od (35%, 25% i 40%), te će uzorak biti pridružen trećoj klasi.

I ABELA II	
SREDNJA TAČNOST I STANDARDNA DEVIJACIJA KLASIFIKACIJE UZOF	AKA

Naziv skupa	Podskup	Srednja tačnost [%]	Standardna devijacija [%]
	1	96,79	0,70
CHIRO10	2	99,17	0,36
	3	99,33	0,37
	1	81,00	0,85
FIN-Benthic	2	85,64	0,70
	3	96,58	0,70
STONEFLY9	-	99,01	0,61
EPT29	_	97,43	0,49

TABELA III Poređenje sa originalnim rezultatima i rezultatima na nivou slike

Naziv skupa	Pod- skup	Origin. rezultati [%]	Naši rezultati [%]	Pobo- ljšanje [%]	Naši rezultati po slici [%]	Pobo- ljšanje po uzorku [%]
FIN- Benthic	1	75,74	81,00	+5,29	76,59	+4,41
	2	81.04	85,64	+4,60	81,19	+4,45
	3	90,14	96,58	+6,44	93,63	+2,95
STONE- FLY9	_	94,50	99,01	+4,55	97,69	+1,32
EPT29	_	88,06	97,43	+9,56	95,37	+2,06

U Tabeli II prikazani su usrednjeni rezultati koji prikazuju tačnost na nivou uzorka kao i odgovarajuća standardna devijacija. U Tabeli III prikazano je poređenje dobijenih rezultata korišćenjem predloženog metoda i originalnih rezultata (FIN-Benthic [10], STONEFLY9 [6] i EPT29 [7]).

Predloženi metod je pokazao značajno poboljšanje u identifikaciji vrsta na svim javnim skupovima podataka koje se kreće u opsegu od 4,55 do 9,56%. Jedini lošiji rezultat je zabeležen na našem CHIRO10 skupu, ali razlog za to je drugačija strategija evaluacije rezultata koja koristi dvostruku unakrsnu evaluaciju, tj. odvojeni validacioni i test skup. Takođe razlog za lošiji rezultat leži u činjenici da su postojeći eksperimenti vršeni sa slikama 256x256 piksela, dok je originalni rezultat postignut sa slikama 512x512 piksela.

Ukoliko pogledamo matricu konfuzije prikazanu na slici 4, uočava se da je najlošiji rezultat od $82,7\pm8,1\%$ ostvaren kod klasifikacije *Polypedilum laetum* jedinki. Razlog za to treba tražiti u činjenici da sve ostale klase poseduje ~200 slika po klasi, dok za ovu klasu imamo svega 79 slika. Klasa *Tvetenia calvescens* se nalazi na drugom mestu sa $87,2\pm5,4\%$, međutim ovaj put je većina pogrešno klasifikovanih uzoraka (11,3±4,4%) otišla na klasu *Tvetenia discoloripes* koja je vrlo slična (pripada istom rodu).

Uticaj više slika po uzorku na tačnost klasifikacije uzoraka je pokazalo poboljšanje od 1,32 do 4,45%. Pokazuje se da metod akvizicije slika korišćen kod FIN-Benthic skupa, gde dve kamere slikaju uzorak iz dva ugla, daje bolje rezultate od odgovarajućih metoda korišćenih kod STONEFLY9 i EPT29 skupova podataka gde je broj slika po uzorku 5 i 3, ali se koristi jedna kamera.



Sl. 4. Matrica konfuzije za CHIRO10 skupa podataka (podskup 1).

Konačno, kvalitet slika takođe igra vrlo bitnu ulogu. Ukoliko uporedimo FIN-Benthic podskup 2 (29 klasa, 6038 uzoraka i 11832 slika) sa EPT29 skupom (28 klasa, 1608 uzoraka i 4794 slika) svi parametri govore u korist FIN-Benthic skupa. Međutim, tačnost FIN-Benthic podskup 2 skupa iznosi 85,64%, dok je tačnost kod EPT29 skupa 97,43%. Razlog za ovo je, po našem mišljenju, kvalitet slika koji se ogleda u nivou detalja. Kod FIN-Benthic skupa je taj kvalitet mnogo lošiji nego kod ostalih skupova što se verovatno ogleda i na tačnost klasifikacije.

V. ZAKLJUČAK

U radu je predložen i evaluiran pristup za identifikaciju vrsta za potrebe akvatičnog biomonitoringa. Predloženi metod se oslanja na tri tehnike dubokog učenja koje imaju za cilj poboljšanje robusnosti kada se obučavanje vrši na relativnom malim skupovima podataka: preneseno učenje, proširivanje podataka i odbacivanje. Preneseno učenje je primenjeno korišćenjem ResNet-50 KNM prethodno obučene na ImageNet skupu podataka. Za evaluaciju je iskorišćen naš CHIRO10 skup podataka, kao i nekoliko javnih skupova (FIN-Benthic, STONEFLY9 i EPT29).

Da bi mogli da poredimo rezultate na različitim skupovima podataka, izvršili smo unifikaciju trening procesa korišćenjem slika veličine 256x256 piksela i 10 podela podataka kako bi izmerili srednju tačnost i standardnu devijaciju. Rezultati su pokazali značajna poboljšanja u odnosu na originalne radove, potvrdili značajan uticaj korišćenja više slika po uzorku, ali i pokazali da se broj uzoraka po klasi mora dobro izvagati.

ZAHVALNICA

Prikazani rezultati dobijeni su u okviru istraživanja na projektima III-43007 i III-47003 koje finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

- IPBES, "Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondizio E.S., H. T. Ngo, M. Guèze, J. Aga," Bonn, 2019.
- [2] P. F. M. Verdonschot, "Evaluation of the use of Water Framework Directive typology descriptors, reference sites and spatial scale in macroinvertebrate stream typology," *Hydrobiologia*, vol. 566, no. 1, pp. 39–58, Aug. 2006.
- [3] G. W. Hopkins and R. P. Freckleton, "Declines in the numbers of amateur and professional taxonomists: Implications for conservation," *Anim. Conserv.*, vol. 5, no. 3, pp. 245–249, Aug. 2002.
- [4] F. C. Jones, "Taxonomic sufficiency: The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates," *Environ. Rev.*, vol. 16, no. NA, pp. 45–69, Dec. 2008.
- [5] "CS231n Convolutional Neural Networks for Visual Recognition: Image Classification." [Online]. Available: http://cs231n.github.io/classification/. [Accessed: 23-Dec-2019].
- [6] D. A. Lytle *et al.*, "Automated processing and identification of benthic invertebrate samples," *J. North Am. Benthol. Soc.*, vol. 29, no. 3, pp. 867–874, 2010.
- [7] N. Larios et al., "Stacked spatial-pyramid kernel: An object-class recognition method to combine scores from random trees," in 2011 IEEE Workshop on Applications of Computer Vision, WACV 2011, 2011, pp. 329–335.
- [8] S. Kiranyaz et al., "Classification and retrieval on macroinvertebrate

image databases," Comput. Biol. Med., vol. 41, no. 7, pp. 463–472, 2011.

- [9] H. Joutsijoki *et al.*, "Evaluating the performance of artificial neural networks for the classification of freshwater benthic macroinvertebrates," *Ecol. Inform.*, vol. 20, pp. 1–12, Mar. 2014.
- [10] J. Raitoharju *et al.*, "Benchmark database for fine-grained image classification of benthic macroinvertebrates," *Image Vis. Comput.*, vol. 78, pp. 73–83, Oct. 2018.
- [11] D. Milošević et al., "Application of deep learning in aquatic bioassessment: Towards automated identification of non-biting midges," Sci. Total Environ., vol. 711, p. 135160, Apr. 2020.
- [12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
 [13] G. Martinez-Munoz *et al.*, "Dictionary-free categorization of very
- similar objects via stacked evidence trees," 2010, pp. 549–556.
- [14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, vol. 2, pp. 1097–1105.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference* on Computer Vision and Pattern Recognition, 2016, vol. 2016-Decem, pp. 770–778.
- [16] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [17] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing coadaptation of feature detectors," *arxiv.org*, 2012.
- [18] "Keras: The Python Deep Learning library." [Online]. Available: https://keras.io. [Accessed: 24-Jan-2020].
- [19] "TensorFlow: An end-to-end open source machine learning platform." [Online]. Available: https://www.tensorflow.org/. [Accessed: 24-Jan-2020].
- [20] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in 3rd International Conference for Learning Representations, 2015.
- [21] J. Jordan, "Setting the learning rate of your neural network," 2018.
 [Online]. Available: https://www.jeremyjordan.me/nn-learning-rate/.
 [Accessed: 24-Jan-2020].

ABSTRACT

Aquatic insects and other benthic macroinvertebrates are mostly used as bioindicators of the ecological status of freshwaters. However, an expensive and time-consuming process of species identification represents one of the key obstacles for reliable biomonitoring of aquatic ecosystems. In this paper, we proposed a deep learning-based method for species identification that we evaluated on several available public datasets (FIN-Benthic, STONEFLY9, and EPT29) along with our CHIRO10 dataset. The proposed method relies on three deep learning techniques used to improve robustness when training is done on a relatively small dataset: transfer learning, data augmentation, and feature dropout. The results for all datasets were obtained using 256×256 images and averaging on 10 data splits in training (50%), validation (20%), and test (30%) sets. The results show significant improvement compared to original contributions and confirms that there is a considerable gain when there are multiple images per specimen.

Species identification for aquatic biomonitoring using convolutional neural networks and deep learning

Aleksandar Milosavljević, Đurađ Milošević and Bratislav Predić