

Estimacija poze životinja koristeći konvolucijske neuronske mreže

Tomislav Žiger, Klara Krmpotić, Lovro Srebačić,
Sven Skender, Danijel Barišić

1. Uvod i motivacija

Estimacija poza je problem računalnog vida koji na slici identificira lokaciju određenih točaka, takozvane ključne točke. Ključne točke predstavljaju određene točke objekta na slici, primjerice zglobovi ili jedinstvene značajke. Izlaz modela je skup točaka koje predstavljaju ključne točke te ocjena pouzdanosti za svaku od njih.

U projektu je prikazano kako pomoću konvolucijske neuronske mreže možemo procijeniti pozu životinje. Estimacija poza može se primijeniti za praćenje i analizu obrazaca kretanja što može biti od velike koristi u veterinarske svrhe za procjenu zdravlja i dobrobiti životinja, otkrivanje anomalija i ranih prevencija određenih bolesti ili ozljeda. Isto tako i primjena u treniranju životinja.

2. Postojeća rješenja

U nastavku su navedena neka od najviše korištenih rješenja za estimaciju poza životinja.

DeepLabCut je široko korišten okvir za bezmarkernu (engl. markerless) procjenu položaja u analizi ponašanja životinja u istraživanjima iz područja neuroznanosti i etologije. Razvijen kao otvoreni alat, DeepLabCut koristi tehniku dubokog učenja za precizno otkrivanje ključnih točaka. Okvir pruža korisnički prijateljski

alat za označavanje dijelova tijela, podržava prijenosno učenje (engl. transfer learning) i neovisan je o platformi. Svojom svestranošću, DeepLabCut se pokazao vrijednim pri zadatku praćenja poza u stvarnom vremenu [1].

SLEAP (Social LEAP Estimates Animal Pose) je open-source python knjižnica za praćenje udova više životinja u video podacima, temeljena na dubokom učenju. Ima korisničko sučelje za jednostavno označavanje podataka i podržava aktivno učenje za poboljšanje preciznosti rezultata. Primjenjujući duboke neuronske mreže, SLEAP ima za cilj precizno pratiti i procijeniti položaje životinja na slikama i videozapисima. [2] Mogućnosti prilagodbe omogućuju istraživačima prilagodbu okvira različitim eksperimentalnim uvjetima, čineći ga vrijednim alatom u neuroznanosti i analizi ponašanja. Kao open-source rješenje, SLEAP potiče suradnju i doprinose istraživačke zajednice[3].

DeepPoseKit je otvorena Python knjižnica dizajnirana za procjenu položaja, posebno u kontekstu praćenja ključnih točaka u biološkim ili anatomskim studijama. Koristeći okvire dubokog učenja poput TensorFlowa i Kerasa, DeepPoseKit nudi prilagodljivu arhitekturu, alat za označavanje ključnih točaka i podršku za prijenosno učenje. DeepPoseKit je trenutno ograničen na estimaciju individualnih poza zasebno [4].

3. Opis projekta

Za izradu projekta korišten je “YOLOv8n-pose” model za estimaciju poza od Ultralytics-a [5].

YOLOv8 (You Only Look Once) je model najnovije generacije koji nadmašuje uspjehe prethodnih verzija YOLO-a i

uvodi nove značajke i poboljšanja kako bi dodatno povećao performanse i fleksibilnost. Osmišljen je da bude brz, precizan i jednostavan za upotrebu, čineći ga izvrsnim odborom za širok spektar zadataka kao što su prepoznavanje i praćenje objekata, segmentacije instanci, klasifikacije slika te estimacije poza.

YOLO za estimaciju poza ima za cilj identificirati i lokalizirati ključne točke na tijelu, kao što su zglobovi i drugi značajni dijelovi. Umjesto prepoznavanja objekata, model se fokusira na pronalaženje točnih koordinata tih ključnih točaka u okviru slike.

Arhitektura YOLO za estimaciju poza uključuje izlazni sloj koji generira koordinate ključnih točaka, obično u obliku (x, y) za 2D estimaciju ili (x, y, z) za 3D estimaciju. Osim toga, ocjene pouzdanosti za svaku ključnu točku pomažu u procjeni kvalitete estimacije.

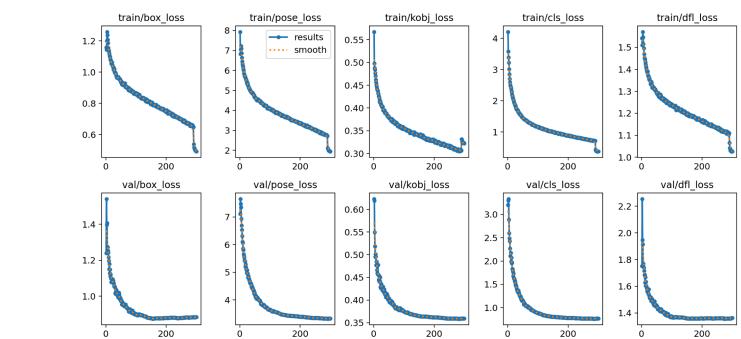
YOLO za estimaciju poza ima široku primjenu u područjima kao što su analiza pokreta, praćenje aktivnosti, biomehanika, virtualna stvarnost i interakcija čovjeka i računala. Također, pruža i alate za bolje razumijevanje i interpretaciju tih istih pokreta.

Skup slika koji smo koristili za treniranje i validiranje modela preuzeli smo sa interneta (<https://github.com/AlexTheBad/AP-10K/tree/main/images>). Skup se sastoji od 10000 slika 54 vrsta životinja raznih dimenzija. Slike su podijeljene na skup za treniranje (70%), skup za validaciju (10%) i skup za testiranje (20%) nakon čega je pokrenuto treniranje modela kroz 300 epoha. Parametri korištene YOLOv8 mreže nalaze se u datoteci weights/best.pt.

4. Opis rezultata

Nakon provedenog učenja i validacije našeg modela, uslijedilo je testiranje na novim slikama kako bi se provjerila njegova učinkovitost. Nakon testiranja, provela se evaluacija rezultata koristeći metrike preciznost, odziv i F1 mjera kako bi se utvrdila točnost detekcije.

Na sljedećoj slici mogu se vidjeti gubici pri treniranju i validaciji modela:



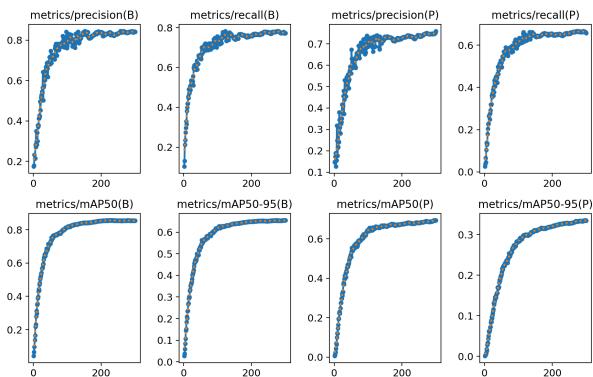
Na ovoj slici gubitaka možemo vidjeti kako su treniranje i validacija modela poprilično uspješne za prepoznavanje životinja na slici (gubitak u iznosu od 0.49 kod treniranja i gubitak u iznosu od 0.88 kod validacije). Kod prepoznavanja poza životinja, treniranje i validacija nisu bili toliko uspješni (gubitak u iznosu od 1.95 kod treniranja i gubitak u iznosu od 3.33 kod validacije).

Ovakav visoki gubitak možemo prepisati loše označenom skupu podataka za treniranje koji možemo vidjeti na sljedećoj slici:

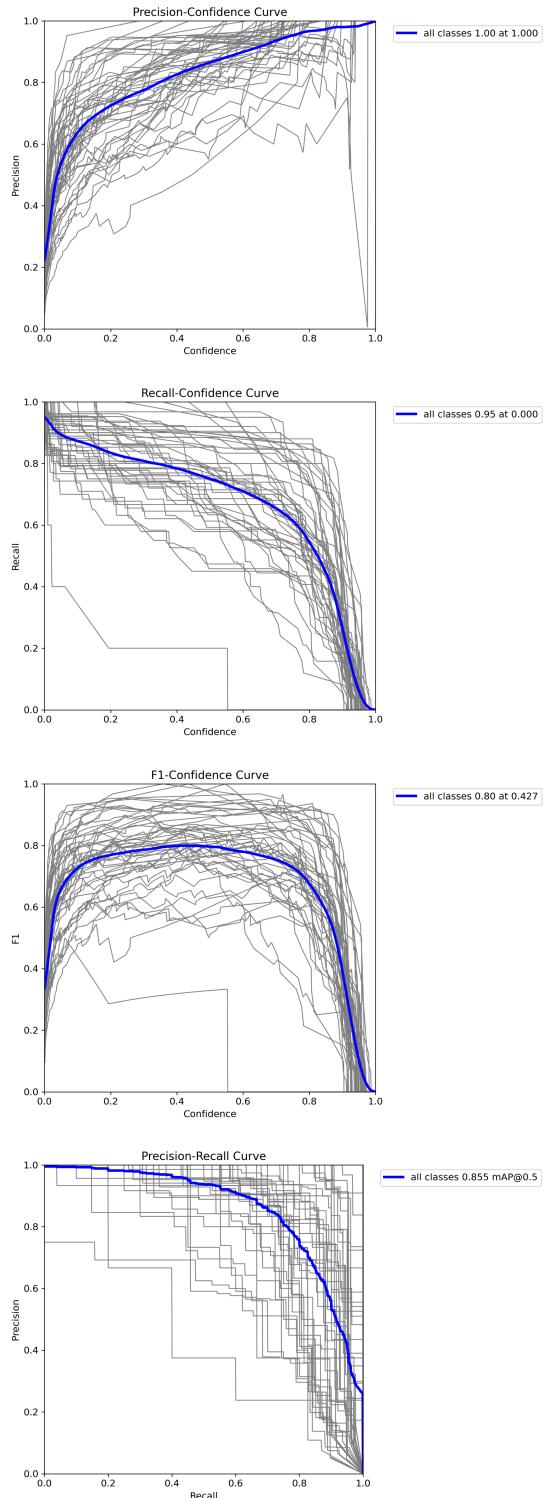


Ne uzimajući u obzir samu estimaciju poze, na temelju zadanih keypoints-a iz dataseta, model je uspješno prepoznao keypoints-e na životinjama sa novih slika, što vidimo iz kobj_loss za treniranje modela koji iznosi 0.32 i za validaciju koji iznosi 0.36.

Sljedeća slika predstavlja metrike koje smo koristili kako bismo ocijenili uspješnost našeg modela:

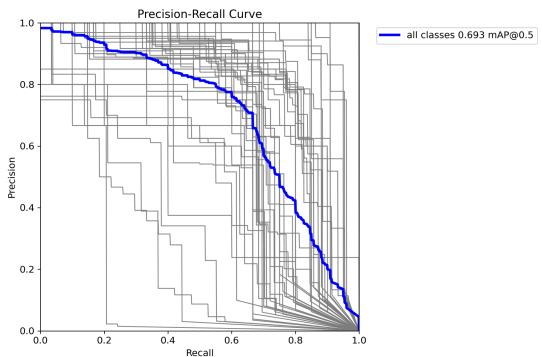
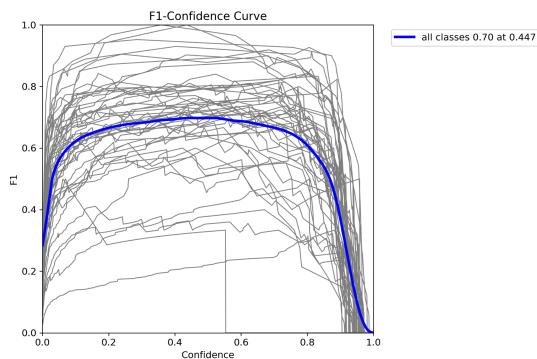
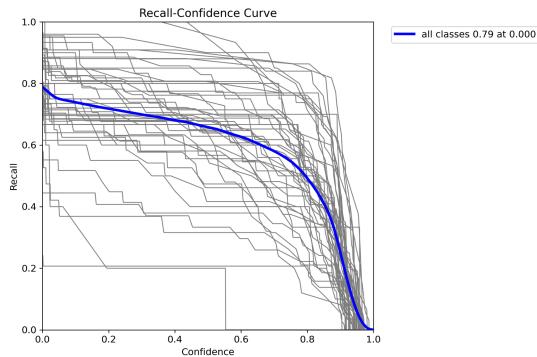
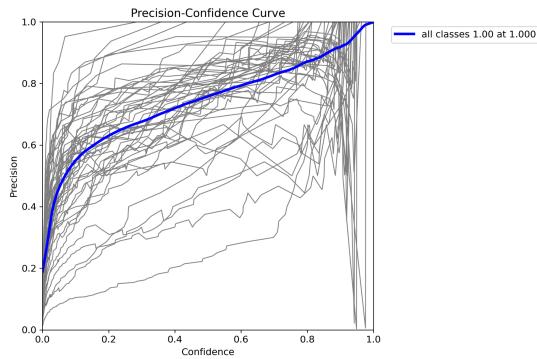


Na ovoj slici možemo vidjeti kako su preciznost i odziv za prepoznavanje životinja visoki, preciznost u iznosu 0.84 i odziv u iznosu 0.77. Kod prepoznavanja poze životinja preciznost iznosi 0.76, a odziv 0.66 što opet možemo prepisati loše označenom skupu za treniranje. Ovdje još možemo vidjeti da mAP pri pragu IoU-a od 0.5 na kraju treniranja za prepoznavanje životinja iznosi 0.85, a za prepoznavanje poze 0.69. Pri IoU pragovima od 0.5 do 0.95 mAP za prepoznavanje životinje iznosi 0.65, a za prepoznavanje poze iznosi 0.33, što je očekivano, jer su pragovi viši.



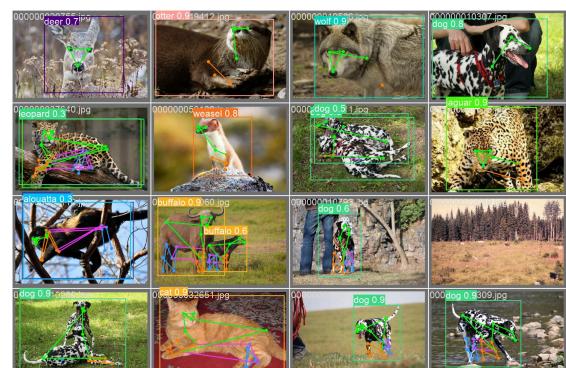
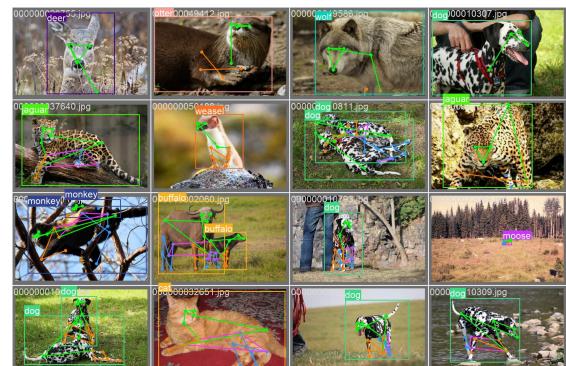
Ove slike prikazuju odnose preciznosti, odziva, pouzdanosti i F1 mjere za prepoznavanje životinje na slici. Sa grafa koji prikazuje krivulju preciznosti u ovisnosti o pouzdanosti možemo vidjeti da se preciznost povećava s povećanjem pouzdanosti. Sa sljedećeg grafa,

odziv-pouzdanost, vidimo da se odziv smanjuje s povećanjem pouzdanosti, što je i očekivano. F1 mjera harmonijska je sredina preciznosti i odziva, a sa grafa koji prikazuje odnos F1 mjere i pouzdanosti vidimo da je mjera pouzdanosti koja optimizira preciznost i odziv jednaka 0.427, jer je za nju F1 mjera maksimalna i ona iznosi 0.8.



Ove slike prikazuju odnose preciznosti, odziva, pouzdanosti i F1 mjere za prepoznavanje poza životinja na slici. Ovdje također možemo vidjeti kako preciznost raste, a odziv pada s povećanjem pouzdanosti. F1 mjera za prepoznavanje poza životinja postiže maksimum od 0.447 za mjeru pouzdanosti od 0.7.

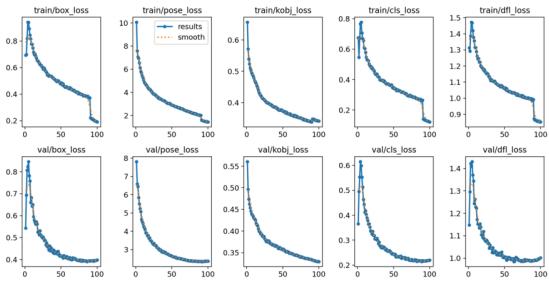
Sljedeća dva skupa slika prikazuju skup označenih slika za validaciju i predict skup slika našeg modela.



5. Diskusija rezultata

Rezultate na koje ćemo se referencirati odnose se na rad Kunal Dawna: Animal Pose Estimation: Fine-tunning YOLOv8 Pose Models [6]. Rad se odnosi na estimaciju poza pasa na 100 epoha.

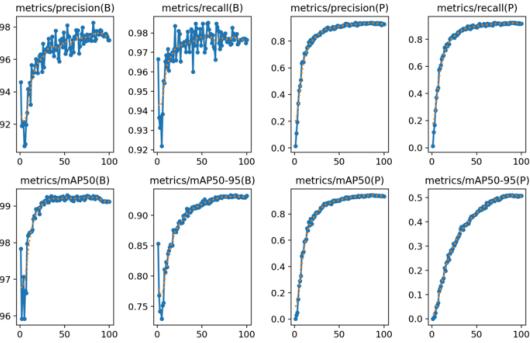
Na sljedećoj slici mogu se vidjeti gubici pri treniranju i validaciji modela u navedenom radu:



Vidimo da što se tiče prepoznavanja životinja i pasa da je gubitak na treniranju ispod 0.2 gdje je u našem radu gubitak 0.49, a kod validacije ispod 0.4 u odnosu na naših 0.88. Što se tiče prepoznavanja poza imamo gubitak ispod 2.0 kod treniranja u odnosu na naših 1.95 i gubitak ispod 2.5 kod validacije gdje je kod nas 3.33. Rezultati gubitaka su poprilično slični.

Što se tiče uspješnog prepoznavanja keypoints-a na životinjama/psima sa novih slika vidimo minimalnu razliku gdje kobj_loss iznosi 0.35 za njihov rad i 0.32 za naš rad i za validaciju u kojoj kobj_loss iznosi 0.33 za njihov rad i 0.36 za naš rad.

Sljedeća slika predstavlja metrike koje su se također koristile u radu kako bi se ocijenila uspješnost modela:



Na ovoj slici možemo vidjeti da su preciznost i odziv u iznosu od 0.97 za razliku od naše preciznosti i odziva čije su vrijednosti 0.84 i 0.77. Kod prepoznavanja poze pasa preciznost iznosi 0.92 u odnosu na naših 0.76, a odziv 0.92 gdje je naš odziv 0.66. Možemo uočiti da su njihove metrike nešto uspješnije od naših.

Ovdje još možemo vidjeti da mAP pri pragu IoU-a od 0.5 na kraju treniranja za prepoznavanje životinja/pasa iznosi 0.99, a za naš rad je to 0.85, dok za prepoznavanje poze 0.94 gdje je kod nas manji, točnije 0.69. Pri IoU pragovima od 0.5 do 0.95 mAP za prepoznavanje životinje/psa iznosi 0.92 i kod nas 0.65, a kod prepoznavanja poze iznosi 0.50, dok kod nas iznosi 0.37. Naši lošiji rezultati u usporedbi s navedenim radom su očekivani zbog korištenja različitih vrsta životinja kod treniranja.

6. Zaključak

Proveden je projekt estimacije poza životinja koristeći YOLOv8 model koji služi za estimaciju poza kroz identifikaciju ključnih točaka na tijelu. Kroz uporabu dubokog učenja i analizu ovih točaka, model je uspješno prepoznao životinje na novim slikama. Gubici pri treniranju i validaciji sugeriraju potrebu za boljim i kvalitetnijim označavanjem podataka. Ovi rezultati pružaju temelj za

daljnje istraživanje i razvoj u području estimacije poza životinja.

7. Literatura

[1] *DeepLabCut*, GitHub.

<https://github.com/DeepLabCut/DeepLabCut>

[2] *Social LEAP Estimates Animal Poses (SLEAP)*.

<https://sleap.ai/>; pristupljeno 14.1.2024.

[3] Pereira, Talmo D., et al. "Fast animal pose estimation using deep neural networks." *Nature methods* 16.1 (2019): 117-125.

[4] *DeepPoseKit*, GitHub.

<https://github.com/jgraving/DeepPoseKit>;

[5] *Pose Estimation*, Ultralytics YOLOv8 docs.

<https://docs.ultralytics.com/tasks/pose/>;

[6] Kunal Dawn, *Animal Pose Estimation: Fine-tuning YOLOv8 Pose Models*, LearnOpenCV (19.9.2023).

<https://learnopencv.com/animal-pose-estimation/>;