

UPUTE ZA PRIKUPLJANJE PODATAKA

SVRHA: Projekt iz predmeta Neuronske Mreže

Opis rješenja: Sustav poduprijet neuronskom mrežom i tehnikama strojnog učenja te računalnog vida, koji u realnom vremenu snima korisnika, te očitava njegove geste dlana kao komande za upravljanje zrakoplovom unutar simulacije/igrice.

Ulaz: Niz slika (video) korisnika i njegovog dlana.

Izlaz: Komande za zrakoplov unutar simulacije, koje čine pozicija u 2d koordinatnom sustavu, rotacija, trenutna brzina.

1. MOTIVACIJA

Prikupljanje podataka na temelju kojih će se u kasnijim fazama projekta graditi model neuronske mreže je temeljni dio svakog ozbiljnog ML projekta. Zbog specifičnosti našeg zadatka, podatke (slike) prikupljamo sami. To je iscrpan proces, tako da će svaki član sudjelovati u tom dijelu.

Kasnije ćemo se dotaknuti svakog od koraka, no sada ćemo ukratko proći što ćete morati učiniti: snimiti par videozapisa, kodom iz njih izvući korisne frameove (slike), sve te slike provesti kroz „LabelStudio“ kako bi na svakoj slici označili područje dlana (bounding box), uploadat takav dataframe na zajednički Google Disk gdje će se svi naši podaci konačno ujediniti u jedan dataset.

Ideja je da svaki član snimi 2 (eventualno 3) videa od otprilike 30 sekundi, ukupno 1 minuta, te time ciljamo dobiti 50 slika po videu, ukupno 100 slika po osobi. Ako netko baš želi, može i tri videa, odnosno 150 slika.

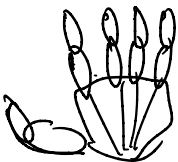
2. SNIMANJE VIDEOA

Dakle, bilo bi dobro snimiti 2 videa od 30ak sekundi, u različitim uvjetima, druga pozicija doma (pozadina) i drugi uvjeti osvjetljenja. Varirajte i udaljenost dlana od kamere, te poze. Pokušajte da u startu, kamera bude u visini dlana.

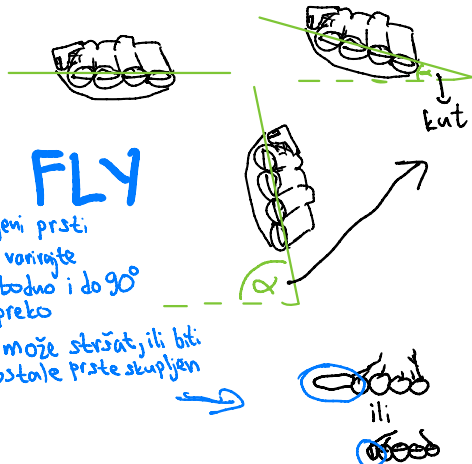
Kvaliteta kamere nije bitna, jer će se svaki frame ionako reducirati u veličini, jedino je bitno da kamera snima u omjeru (ratio) 16/9 ili 3/2 (to su standardni formati pa vjerojatno vaši uređaji to zadovoljavaju).

Što snimati? Zapravo nam je bitno imati 4 ključne poze: otvoreni dlan (PAUSE), dlan prema dolje skupljeni

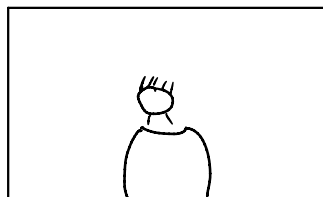
prsti (FLY) [slobodno može i treba biti rotacija], poze dlana između te dvije akcije, neželjene poze (UNDEFINED) i slike bez dlana (NOPALM). Pokušajte varirati udaljenost dlana od kamere, ići s dlanovima i gore/dolje i lijevo/desno, (x,y) koordinate slobodno istražite. Pokušajte odokativno da 20% podataka bude NOPALM, 20% PAUSE i preostalo (najbitnije) FLY.



PAUSE



UNDEFINED



NOPALM

• bez dlana u kadru

3. IZ VIDEOA U PODATKE

Nakon što ste snimili videozapise i prenijeli ih na kompjuter (.mp4, .mov), pripremio sam python kod („l_video_to_dataset.py“), koji će iz vaših videa generirati podatke i spremiti ih u folder. Ima par parametara, pa pogledajte što piše u kodu, ima komentara.

Kod će podijeliti vaš video u tih ciljanih 50ak slika, svaki frame pretvoriti u crno-bijeli, reducirati ga na željenu veličinu, te malo ujednačiti intenzitet svijetlosti. Paralelno, kod će generirati još jedan folder u kojem će se zasebno snimati razlika između 2 susjedna framea (movement). Te slike nakon što se izvuku nećemo više dirati, ali ćete ih uploadati također na kraju. Finalno, trebali bi imati 2 foldera na kompjuteru sa svim slikama i sa slikama movement. Za svaki video je potrebno zasebno pokrenuti kod. (sve će postati jasnije kad pokrenete kod)

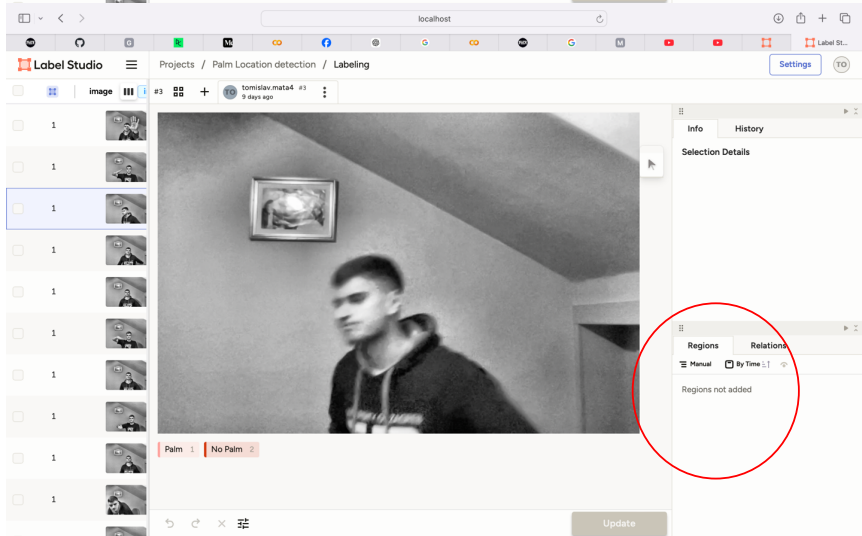
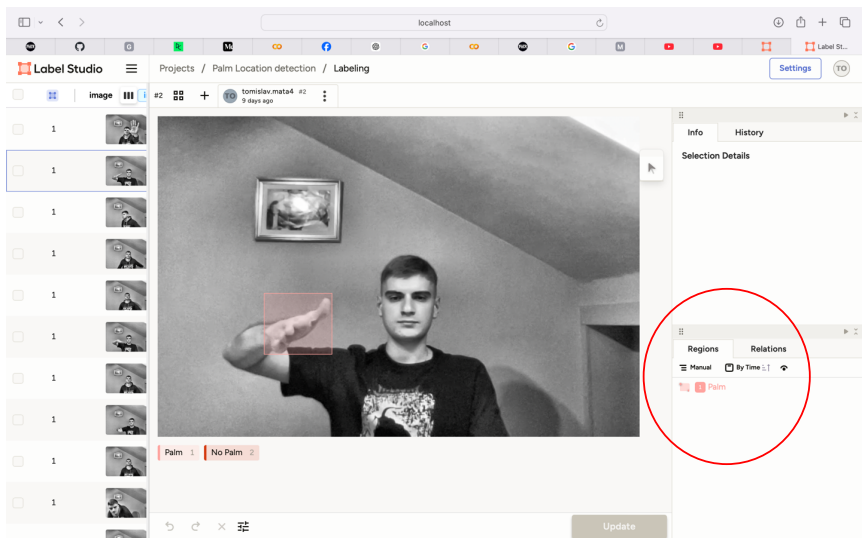
4. OZNAČAVANJE PODATAKA

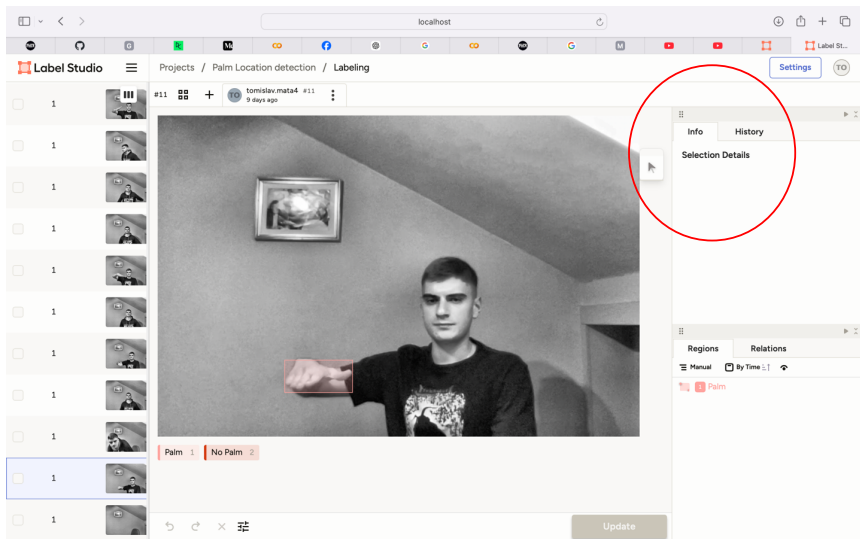
Ovo je malo zeznutiji dio. Kako bi trenirali NN da zna prepoznati dlan na slici, potrebno je pripremiti podatke gdje na slici već ima označeno područje dlana (nacrtan je boundary box oko njega). Za to ćemo koristiti „Label Studio“. Potrebno je imati neko virtualno python okruženje lokalno na računalu. Te kada smo u virtualnom okruženju, u njemu se instalira i pokreće Label Studio iz terminala.

Video s yt kako skinuti i pokrenuti „Label Studio“: <https://youtu.be/hK3poCKIQjM?si=QuTat6lWxB1tDc8g>

Unutar njega stvorite novi projekt i uploadajte podatke (iz foldera sa slikama, ne movement folder), ukupno cca 100 slika. Kao Labeling Setup koristite „Object detection with Bounding boxes“. Tu je još potrebno definirati klase, važno je da svi koristimo iste nazive klasa: PAUSE, FLY, UNDEFINED, NOPALM. Dakle točno te nazive klasa, velika slova. Ostale postavke neka ostanu i gotovi smo.

Sada možemo označavati podatke, postoji plavi gumb, „Label All Tasks“. Važno je biti što precizniji da odaberete točnu klasu i usko i precizno označite dlan. Za NOPALM klasu nije potrebno označavati ništa, samo preskočite. Evo par primjera:





Nakon što ste prošli sve podatke, možemo ih eksportirati. Postoji gumb „Export“ te ćemo odabrati opciju „COCO“. Kao rezultat dobijemo folder sa slikama i result.json koji sadrži informacije o klasi i o lokaciji boundary box-a za svaku sliku. U imena tih datoteka dodajte svoje inicijale npr. (images_TM) (result_TM.json) i to je spremno za staviti na google disk, od tamo će se oni koristiti za treniranje neuronske mreže. Hvala za trud i nadam se da je bilo zanimljivo te da se naučilo ponešto.

5. UPLOAD NA GOOGLE DISK

Na google disk uploadamo 3 stvari:

folder slika (images_TM)[cca 100 slika], folder s movement slikama iz 1. koraka (movement_TM)[cca 100 slika], LabelStudio anotacije (result_TM.json)... umjesto TM vaši inicijali naravno. Molim vas da poštujuete te nazive, kako bi ih bilo lakše spojiti kasnije.

Link na moj google disk, tu uploadate:

https://drive.google.com/drive/folders/1qeVQCfz6e9u_w8X545t7MXT89ROz7PtF?usp=share_link

Ako će biti poteškoća, možemo se dogovoriti i na drugi način.

ROK ZA PREDAJU: 31.12.2024. do kraja dana, što ranije, to bolje J

6. Što dalje...

Ne morate ovo čitati ako Vas ne zanima... uglavnom, ako svatko od nas generira 100 primjeraka, ukupno ćemo imati 600 označenih slika. To je solidan broj, ali plan za dalje je augmentacija podataka, čime možemo slike koje imamo dodatno zrcaliti, rotirati, umanjiti/uvećati itd., kako bi smo umjetno povećali dataset 3-4 puta. To bi bio dosta solidan dataset, s kojim bi trebali moći izgraditi uspješan model.

Za izgradnju neuronske mreže, mislim da bi zgodno bilo koristiti Google Collab, lako se poveže s Google Diskom za dohvaćanje podataka, te nudi besplatno GPU, što iznimno ubrzava treniranje NN, pogotovo kad se radi sa slikama.

Ja sam već malo eksperimentirao oko izgradnje neuronske mreže u TensorFlow-u, sa svojim skromnim datasetom od 100 slika i dobio sam obećavajuće rezultate, ali ne dovoljno dobre zasad. Možete malo pogledati kako se koriste bounding boxes koje ste označavali.

