



Univerza v Mariboru

Fakulteta za elektrotehniko,
računalništvo in informatiko

Končno poročilo za projekt “Upravljanje z gestami”

pri predmetu Signale in slike

Člani skupine:

Nik Terglav

Luka Lamprečnik

Simona Zhirova

Marija Jovanova

UVOD

Pri nalogi sem si zadal, da bom na podlagi zavzetih podatkov iz meritev naučil model, da bo prepoznal geste rok iz posnetkov in pravilno razbral funkcijo geste. Ampak ker je pri predmetu umetna inteligenca prišlo do sprememb pri določevanju modelov. Sem pri učenju modela uporabil drugo zbirko podatkov, namreč je moj model prevzel drugo funkcionalnost tj. določevanje okoliških spremenljivk (npr. dan/noč, dež/sneg/jasno). Zato pretekli že zajeti podatki niso ustrezali. Ker pa sem že zajel lastne podatke pri meritvah 1 in 2, sem se odločil za augmentacijo nad temi podatki.

V okviru projekta želim dokončati moj model nevronske mreže, da bo čim natančneje prepoznal vreme oz. čas v dnevu. Za učenje uporabim slike, katere prav tako augmentiram z funkcijami iz knjižnic. Končen izhod bo določitev izvedbe posamezne funkcionalnosti pri vožnji. Npr. če je megleno; vklop meglenk, deževno; vklop brisalcev, noč; vklop ambientne osvetljave.

JEDRO

Zajem podatkov:

Za vsake meritve sem si pripravil 'checklist' z vsemi gestami in število posnetkov, ki jih želim posneti za posamezno gesto. Snemal sem z mobilnim telefonom in uporabil 'wide angle', da sem zajel celotno gesto v okvir. Telefon sem namestil nad volan in ga usmeril proti vozniku, ko sem imel en dolg posnetek vseh gest sem ga razrezal in dodal geste v Mega direktorij. Pri meritvah 2 pa smo se odločili za unifikacijo formata podatkov zato sem začel snemati geste s pomočjo 'media pipe' skripte, ki shrani posneto gesto v formatu '.npy' polja. Vse različne iteracije posameznih gest sem v obliki polj prav tako shranil v skupinski direktorij na platformi mega.

Označevanje/Anotacija podatkov:

Začetne testne podatke sem označeval z CVAT in LabelStudio, da smo preizkusili orodja. Po diskusiji smo se odločili za uporabo 'python' skripte, ki uporablja 'mediapipe' knjižnico za shranjevanje 'keypointov' roke v določenem časovnem trenutku. Preko kamere prenosnika sem zaigral gesto, po koncu vpisal ime in skripta je to gesto pretvorila v določen 'ime_geste.npy array'. Skripta zajeme 'skeleton' premikajoče se roke.

Razširjanje/Augmentacija podatkov:

1. Dodajanje šuma (add_noise):

- Simulira manjše napake zaznave ali gibanja.
- Gaussian šum z nizko intenziteto (noise_level=0.01).

2. Zamik okvirjev (frame_shift):

- Premik časovnega poteka signala v levo ali desno.
- Posnema zamike v začetku ali koncu izvedbe geste.

3. Glajenje (smoothing):

- Uporablja Hannovo okno za glajenje gibanja skozi čas.
- Odstrani nenadne skoke in izenači gibanje.

Statistika:

Število snemanj:	Skupno 4 : 2 x z telefonom, 2 x s skripto
Količina zajetih podatkov:	40 posnetkov, 32 polj z skripto
Zaseden prostor podatkov:	Posnetki: 146.9MB Polja .npy: 800KB
Št. Podatkov v učni množici:	1000+ z uporabo augmentacije

Za učenje nevronske mreže sem uporabil **1.14GB** veliko zbirko podatkov, ki vsebuje več kot 6000 slik vožnje avtomobila ob različnem času v dnevu in med različnimi vremenskimi stanji.

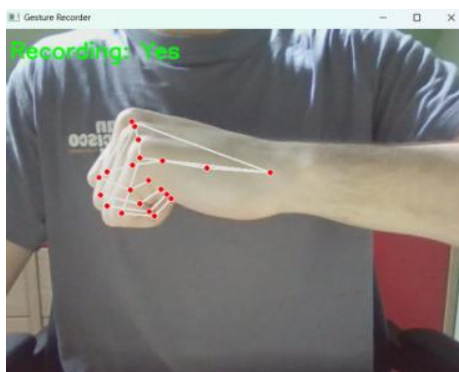
Primeri zajetih podatkov:



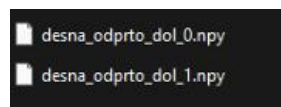
Slika 1: primer zajete geste



Slika 2: primer anotirane geste



Slika 4: označevanje geste s MediaPipe



Slika 3: primer izhoda skripte

```
Shape: (77, 63)
Data type: float64
First few values:
[ 5.73372602e-01  5.16497076e-01  7.71649624e-08  4.90127087e-01
  4.77494806e-01  4.41237018e-02  4.16147768e-01  4.64987844e-01
  5.72532564e-02  3.63016129e-01  4.69781458e-01  6.00338615e-02
  3.20583344e-01  4.77452934e-01  6.02337420e-02  3.70402247e-01
  4.40414906e-01  2.53058504e-02  2.60668427e-01  4.45493907e-01
  1.44262267e-02  1.95509389e-01  4.52451080e-01  9.73674748e-03
  1.44899622e-01  4.57667559e-01  8.13125726e-03  3.69702727e-01
  4.51006800e-01  -6.62557594e-03  2.41439551e-01  4.61825162e-01
 -1.83898918e-02  1.69219092e-01  4.72495586e-01  -2.29422096e-02
  1.13851190e-01  4.77385789e-01  -2.50694510e-02  3.72120261e-01
  4.83825952e-01  -3.60245965e-02  2.49205217e-01  4.94225293e-01
 -4.19408894e-02  1.80424929e-01  5.01029193e-01  -4.48842160e-02
  1.27366751e-01  4.99080539e-01  -4.76316027e-02  3.76863867e-01
  5.35141468e-01  -6.18770681e-02  2.85249352e-01  5.40226698e-01
 -6.38231188e-02  2.31857732e-01  5.39556265e-01  -6.18700683e-02
  1.86785057e-01  5.35250545e-01  -6.10905029e-02]
```



Slika 5: primer deževnega dne



Slika 6: primer noči

ZAKLJUČEK

Projekt »Upravljanje z gestami« je kljub spremembam v začetnem načrtu uspešno prinesel rezultate. Prvotna zamisel o prepoznavanju gest rok se je zaradi sprememb pri modelih umetne inteligence preusmerila v razvijanje sistema za določanje okoliških spremenljivk, kot so vremenski pogoji in čas v dnevu. Pri učenju mreže sem podatke prav tako augmentiral s pomočjo metod iz uvoženih knjižnic.

Čeprav je projekt zahteval prilagoditve, je končni rezultat funkcionalen in predstavlja dobro osnovo za nadaljnje izboljšave, kot so večja natančnost modela ali razširitev na dodatne scenarije uporabe. Sodelovanje v skupini je bilo ključno za uspešno izvedbo projekta.

```
(model-okolje-env) C:\Users\tergl\Desktop\Git repos\Upravljanje_z_gestami\UI\Model 3 - Razvoj modela za a
nalizo okolje-specifičnih podatkov>python Model3.py
Oznake/Labels: ['day', 'night']
Število slik v učnem sklopu: 6438
Oblika vhodnih slik: torch.Size([8, 3, 224, 224])
Oznake: tensor([1, 0, 1, 1, 0, 1, 0, 0])
Oblika izhoda modela: torch.Size([8, 2])
Epoch 1/5: 100%|████████████████████████████████████████████████████████████████████████████████| 804/804 [03:12<00:00, 4.18it/s, Loss=0.104]
Epoch 1/5 - Training Loss: 0.1042, Validation Loss: 0.0796, Validation Accuracy: 97.68%
Epoch 2/5: 100%|████████████████████████████████████████████████████████████████████████████████| 804/804 [03:11<00:00, 4.19it/s, Loss=0.0601]
Epoch 2/5 - Training Loss: 0.0601, Validation Loss: 0.0829, Validation Accuracy: 98.40%
Epoch 3/5: 100%|████████████████████████████████████████████████████████████████████████████████| 804/804 [03:10<00:00, 4.21it/s, Loss=0.0441]
Epoch 3/5 - Training Loss: 0.0441, Validation Loss: 0.0674, Validation Accuracy: 98.55%
Epoch 4/5: 100%|████████████████████████████████████████████████████████████████████████████████| 804/804 [03:11<00:00, 4.21it/s, Loss=0.0203]
Epoch 4/5 - Training Loss: 0.0203, Validation Loss: 0.1428, Validation Accuracy: 96.66%
Epoch 5/5: 100%|████████████████████████████████████████████████████████████████████████████████| 804/804 [03:11<00:00, 4.20it/s, Loss=0.0109]
Epoch 5/5 - Training Loss: 0.0109, Validation Loss: 0.1305, Validation Accuracy: 98.48%
```

Slika 7: primer učenja modela nad novo zbirko