

# Detekcija i klasifikacija ljudskih akcija

Srdan Stjepanović SV16/2020

## Uvod:

Detekcija i prepoznavanje ljudskih pokreta sa slike/videoa je danas veoma popularan izazov. Detekcija položaja tijela čovjeka kao i klasifikacija njegove akcije može imati razne primjene, kao što su praćene nadzornih kamera, praćenje pravilnosti izvođenja vježbi, virtualna realnost i mnoge druge. Cilj ovog projekta je da se detektuje položaj tijela čovjeka i na osnovu ključnih tačaka njegovog tijela da klasifikujemo njegovu aktivnost.

## Podaci:

Korišten je [HAR-dataset](#) koji je sačinjen od 12000 slika, gdje je svaka slika anotirana jednom od 15 ljudskih akcija. Neke klase su predstavljene veoma sličnom pozom, s toga je empirijski izdvojeno 5 klasa koje će se klasifikovati, kako bi se poboljšala preciznost klasifikatora. Skup podatak je podijeljen na trening i test dio u odnosu 80:20. Svaka slika je pretprocesirana sa dva modela OpenPose i MediaPipe pose. Prije samog pretprocesiranja slike su skalirane na 240x240. Izlaz iz OpenPose modela predstavljaju koordinate 18 ključnih tačaka čovjekovog tijela, dok MediaPipe pose kreira 33 ključne tačke.

## Metodologija:

Za detekciju čovjeka i ključnih tačaka njegovog tijela korištena su dva zasebna modela, a to su OpenPose i MediaPipe pose model. Zatim se rezultati pretprocesiranja u vidu koordinata ključnih tačaka prosleđuju klasifikatoru koji klasifikuje ulazne koordinate na jednu od ljudskih akcija. Iako je prvobitno planirano da se koristi samo OpenPose model, zbog svojih ograničenja u brzini, dodat je i MediaPipe model koji je korišten u detekciji čovjeka u realnom vremenu. Korištena su 3 modela za klasifikaciju i poređeni su njihovi rezultati, a to su: SVM, Random forest i neuronska mreža. Arhitektura neuronske mreže koja je korištena se može vidjeti na slici.



Model: "sequential"

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 66)	0
dense (Dense)	(None, 64)	4288
dense_1 (Dense)	(None, 128)	8320
dense_2 (Dense)	(None, 512)	66048
dropout (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 128)	65664
dense_4 (Dense)	(None, 128)	16512
dense_5 (Dense)	(None, 3)	387

## Rezultati:

Evaluacija je vršena pomoću sledećih metrika: accuracy, precision, f1-score i recall. Klasifikatori postižu sl. accuracy, dok ostale metrike mogu da se vide u doc/results folderu.

	OpenPose	MediaPipe
Neuronska mreža	61%	59%
SVM	62%	60%
Random forest	66%	60%

## Zaključak:

Postignuti rezultati nisu idelani, ali postoji mnogo prostora za unapređenje modela. Neuronska mreža pokazuje lošije rezultate od linearnih klasifikatora, ali podešavanjem parametara modela, kao i promjenom arhitekture rezultati bi se mogli poboljšati. Takođe ulaz u klasifikator bi se mogao normalizovati, a skup podataka proširiti. Hiperparametri linearnih klasifikatora bi se mogli optimizovati korištenjem nekog metoda za „hyperparameter tuning“ kao što je Grid Search.