

# Upravljanje muzičkom platformom Spotify pomoću pokreta ruke

Jovana Jevtić, Andrija Vojnović

Fakultet tehničkih nauka, Novi Sad

## Uvod i motivacija

Sposobnost prepoznavanja oblika i pokreta ruke može biti ključna u poboljšanju korisničkog interfejsa mnogih platformi. Zbog toga, ovaj projekat se bavi prepoznavanjem pet različitih pokreta ruke korisnika i u zavisnosti od prepoznatog pokreta, upravljanjem muzičkom platformom Spotify. Akcije koje su podržane su:

- Prevlačenje rukom u desno – prebacivanje na prethodnu pesmu
- Prevlačenje rukom u levo – prebacivanje na narednu pesmu
- Palac nadole – smanjivanje zvuka za dve jedinice
- Palac nagore – pojačavanje zvuka za dve jedinice
- Stop – pauziranje/nastavljanje pesme

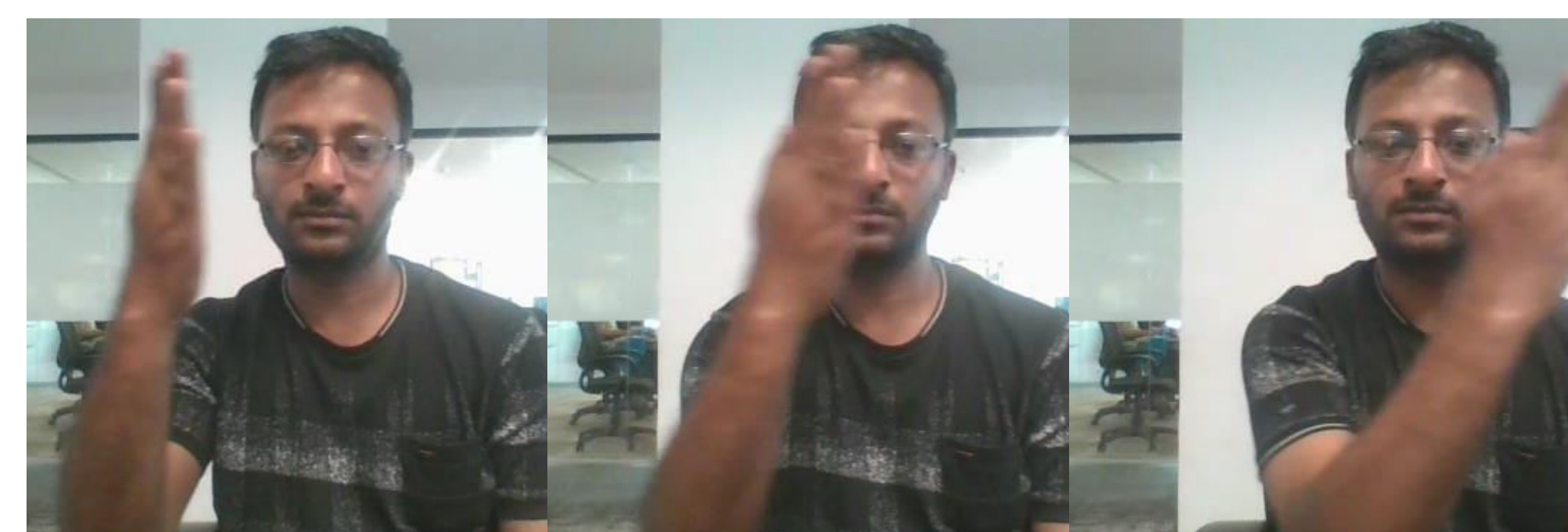
## Skup podataka

Skup podataka je podeljen na trening, validacioni i test skup. Podaci za treniranje i validaciju preuzeti su sa [linka](#) i snimljeni su web kamerom, dok smo podatke za testiranje snimili različitim kamerama. Svaki video sastoji se od 30 frejmaova. Takođe, program će biti testiran u realnom vremenu pomoću web kamere.

## Metodologije

Rešenje za prethodno opisani problem sastoji se iz dva dela. Prvo je neophodno prepoznati ruku na svakom frejmu videa i to smo postigli pomoću MediaPipe Hands. MediaPipe je rešenje koje se sastoji iz više modela. Prvi model detektuje dlan, dok drugi na osnovu bounding boxa oko detektovanog dlana, izračunava i vraća ključne koordinate ruke. Kada dobijemo koordinate ruke na svakom frejmu, potrebno je prepoznati akciju koja se odigrala u samom videu. To smo postigli pomoću konvolucione neuronske mreže. Pored modela koji smo sami napisali, iskoristili smo ZFNet i VGG arhitekture. Sve tri mreže imaju isti ulaz, koji se sastoji od 30 listi sa x i y koordinatama 21 tačke ruke za svaki video. Izlaz iz mreže je 5 verovatnoća za svaku klasu, čiji je zbir 1.

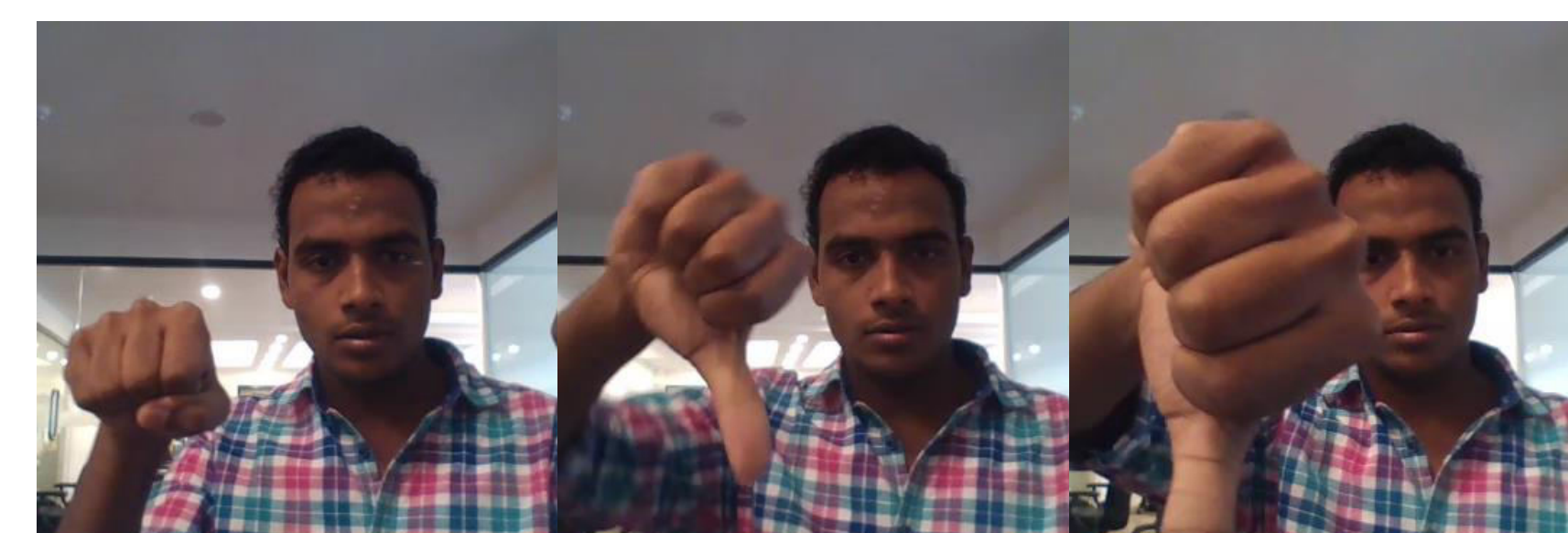
## Primeri pokreta ruku nad kojima su modeli trenirani



Prevlačenje rukom u desno



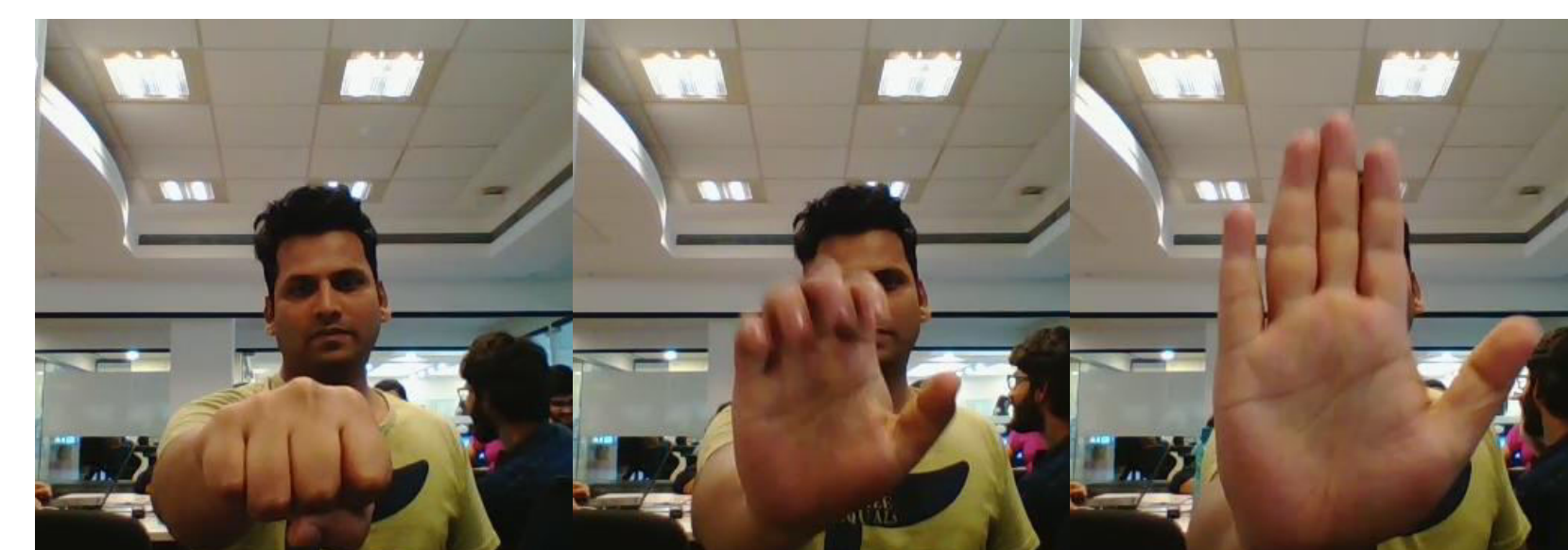
Prevlačenje rukom u levo



Palac nadole



Palac nagore



Stop

## Rezultati

Prilikom testiranja sa testnim podacima, VGG model je imao 58% uspešnosti, ZFNet 82%, dok je naš model imao 84%. Izvršeno je i testiranje preko web camere u realnom vremenu i tom prilikom smo za svaku akciju uradili po 8 ponavljanja za svaki model. ZFNet je bio 0.81% uspešan, naš model 0.91%, dok VGG model nije radio dobro.

## Zaključak

Mislamo da se VGG model nije pokazao dobro kod rešavanja ovog problema jer nemamo dovoljno ulaznih podataka, odnosno da naši podaci ne zahtevaju kompleksnost VGG mreže. Stoga sledi da ukoliko želimo poboljšanja, potrebno je proširiti skup podataka.

Projekat je pokazao da se sa raznim modelima može doći do efikasnog rešenja za upravljanje softverima pomoću pokreta ruke.