

Detekcija lica, prepoznavanje i klasifikacija emocije i generisanje emoji-a na osnovu prepoznatog raspoloženja

Svetozar Vulin

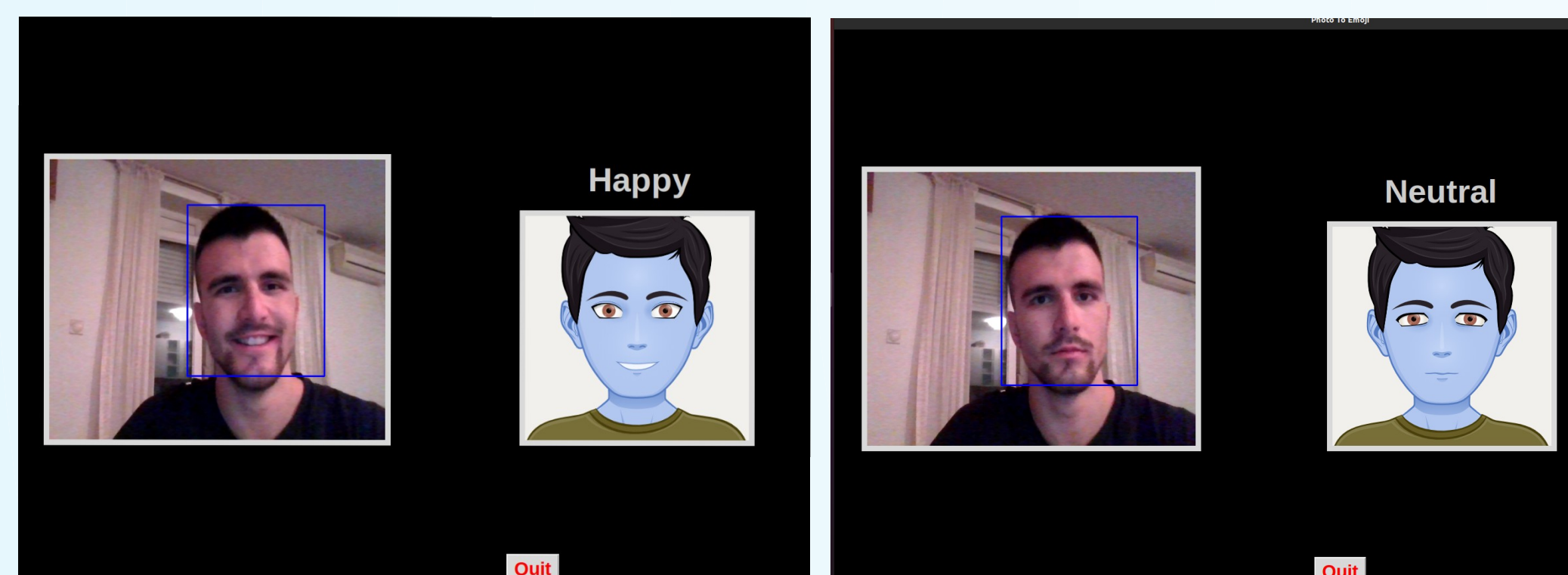
Softversko Inženjerstvo i Informacione Tehnologije, FTN

Uvod

Tema rada predstavlja detektovanje raspoloženja osobe. Raspoloženje osobe se može prepoznati na osnovu facijalne ekspresije kod osobe. Ovaj zadatak se može dekomponovati na manje celine:

- 1) Detekcija lica
- 2) Prepoznavanje facijalne ekspresije osobe
- 3) Klasifikacija emocije na osnovu karakteristika lica

Nakon što su izvršeni navedeni koraci, kao izlaz dobijamo detektovanu emociju, koju dalje pretvaramo u emoji (karikatura osobe koja izgleda kao iz crtaća).



Detekcija lica sa live kamere se vrši korišćenjem openCV biblioteke, koja oko lica nacrtava bounding-box i zatim se frejm sa tim boxom prosleđuje CNN-u (Convolutional Neural Network), koji je istreniran za prepoznavanje facijalnih ekspresija. Zatim se rezultat klasifikacije prosleđuje openCV biblioteci, koja na osnovu prepoznate emocije mapira određenu emociju na emoji i to se dobija kao krajnji output.

Metodologija

U ovom poglavlju će se obraditi koraci prilikom izrade projekta, a to su:

- 1) Analiza i priprema dataset-a
- 2) Kreiranje CNN modela mreže
- 3) Treniranje modela
- 4) Evaluacija

1. Analiza i priprema dataset-a

Dataset korišćen u projektu je FER-2013. Sastoji se od slika osoba sa određenim facijalnim ekspresijama. Slike su sive, veličine 48x48 piksela, gde je faca cropovana i centrirana. Trening set se sastoji od 28000 slika, a test set od 3600. Svaka emocija predstavlja zaseban folder sa slikama koje odgovaraju toj emociji.



2. Kreiranje CNN modela mreže

Prilikom kreiranja modela je korišćena Keras biblioteka. Iz nje su preuzeti slojevi, kao i sam Sequential model. Korišćeni su sledeći slojevi: Konvolucioni, Max-pooling, Dropout, Dense. Aktivacione funkcije u konvolucionim slojevima su relu, a u dense slojevima relu i softmax. Poslednji Dense sloj koristi softmax aktivacionu funkciju za mapiranje 7 output čvorova na 7 mogućih klasa. Max-pooling slojevi su veličine 2x2.

3. Treniranje modela

Pri treniranju je korišćen kerasov Adam optimizator koji je proširenje metode stohastičnog gradijentnog spusta. Broj epoha za treniranje je 50. Loss funkcija je categorical_crossentropy iz razloga što imamo 7 klasa za prepoznavanje.

4. Evaluacija

Nakon završenog treniranja, rezultat preciznosti prepoznatih ekspresija je bio 87%, a loss je iznosio 0.38. Za evaluaciju detekcije je korišćen kerasov Mean IoU (Intersection over Union) koji računa IoU za svaku klasu i onda računa prosek svih klasa. Rezultat Mean IoU je 0.4308.

Testiranje

Prilikom empirijskog testiranja, imamo live kameru koja snima osobu. Oko njene face se iscrtava bounding-box koristeći openCV haar cascade xml. Nakon iscrtanog box-a, uzima se jedan frame tog videa i šalje na obradu. Potrebno je cropovati sliku, tako da se samo deo unutar box-a iskoristi.

Dalje se taj cropovani deo prebacuje u sivu boju, da bi se smanjila kompleksnost slike i lakše izvukle karakteristike sa nje. Tako obrađena slika se šalje do CNN radi dalje obrade, koja zatim prepoznaje karakteristike lica i kao output dobijamo klasifikovano raspoloženje.

Zaključak

Prilikom izrade ovog projekta, primenjene tehnike poput CNN klasifikacije su pokazale koliko segmentacija slika i klasifikacija stvari sa slike napreduje.

Prilikom empirijskih testiranja, primetio sam da je rezultat bio u većini slučajeva tačan, iako je teško samo na osnovu facijalne ekspresije sa slike prepoznati raspoloženje osobe.