

Prepoznavanje emocija na osnovu slike lica

Zoran Jankov – SW8/2018
Dejan Todorović – SW17/2018

UVOD

Rad se bavi temom prepoznavanja emocija na osnovu slike lica. Prepoznavanje emocija je urađeno korišćenjem metode potpornih vektora (eng. *Support Vector Machine*, **SVM**) i konvolucijske neuronske mreže (eng. *Convolutional Neural Network*, **CNN**). Treniranje je vršeno nad **CK+** i **FER2013** skupovima podataka i u oba slučaja je CNN postigla bolje rezultate od SVM-a.

SKUPOVI PODATAKA

- **CK+**
 - Nastao u kontrolisanim uslovima
 - Lice je okrenuto prema kameri
 - Lice je u potpunosti vidljivo
 - Sreća, Tuga, Bes, Strah, Gađenje, Iznenadenje, Prezir
- **FER2013**
 - Slike su prikupljene iz raznih izvora
 - Lice može biti delimično pokriveno
 - Sreća, Tuga, Bes, Strah, Gađenje, Iznenadenje, Neutralno
 - Sadrži preko 30000 slika

REZULTATI

Kako bi SVM mogao da klasifikuje lica na osnovu emocije, potrebno je bilo izvršiti izdvajanje karakterističnih tačaka lica. Za ovo je korišćena *dlib* biblioteka pomoću koje su izdvojene karakteristične tačke obrva, očiju, nosa, usana i vilice. Karakterističnim tačkama je pridodat i histogram orijentisanih gradijenata (eng. *Histogram of oriented gradient*, HOG).

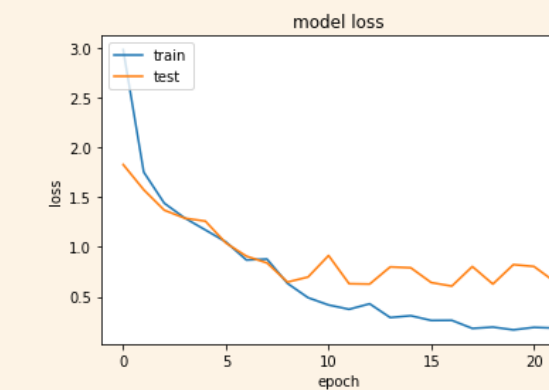
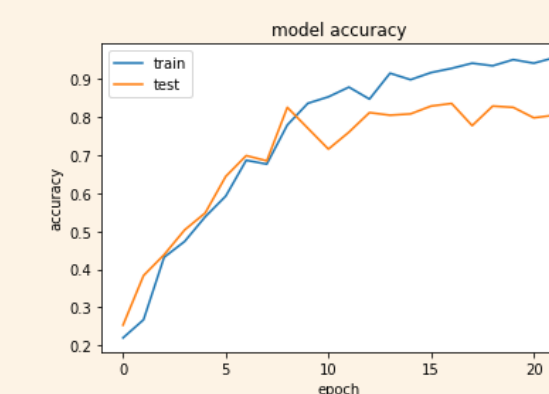
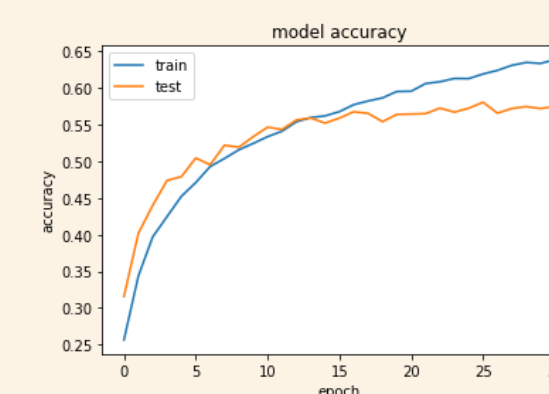
Kako se pokazalo, SVM metoda je dala prilično dobre rezultate nad CK+ skupom podataka, sa tačnošću od 87.7% Ovo je donekle i očekivano jer sva lica tog skupa gledaju ka kameri, što čini izdvajanje karakterističnih tačaka preciznijim.

S druge strane, SVM nad FER2013 skupom ne pokazuje zavidne rezultate s tačnošću od 49.8% Ovo proističe iz prirode slika iz skupa. Zbog neuniformnog načina prikupljanja slika lica, karakteristične tačke lica neće biti dovoljne za kvalitetnu klasifikaciju emocija.

Konvolucijska neuronska mreža sama izdvaja relevantne delove slike kroz *backpropagation* mehanizam. Model koji je korišćen se sastoji od 5 konvolucijskih slojeva, koji su aktivirani uz pomoć *ReLU* funkcije. Nakon svakog konvolucijskog sloja sledi *Max pooling* sloj koji smanjuje dimenzije dva puta. Posle svakog *Max Pooling* sloja se nalazi *Dropout* sloj uz pomoć kojeg se deaktivira određeni procenat neurona. Empirijiskim putem je zaključeno da je potrebno isključiti 20% neurona. *Dropout* slojevi su uvedeni zbog neuniformnosti FER2013 skupa i radi izbegavanja *overfitting*-a.

Ulazne dimenzije slike su 48x48x1, a broj filtera konvolucijskih slojeva je redom 64, 64, 128, 256, 256. Nakon poslednjeg konvolucijskog sloja podaci se svode na sloj dimenzija 1x256 nakon kojeg slede dva sasvim povezana sloja dimenzija 1x256. Nakon njih se nalaze *dropout* slojevi. Konačno, izlaz mreže je sloj dimenzija 1x7 koji je aktiviran uz pomoć *softmax* algoritma.

Mreža se dobro pokazala pri CK+ skupu gde je postigla rezultat od 94.47% tačnosti, dok je nad FER2013 skupom pokazala značajno lošiju tačnost od 58.31%.



ZAKLJUČAK

Rad pokazuje da CNN metoda predstavlja značajno bolje rešenje pri rešavanju problema prepoznavanja emocija na osnovu slike lica. Rezultati SVM metode veoma zavise od okruženja u kojim je slika nastala. Problemi poput lošeg osvetljenja ili nakrivljene slike predstavljaju velike prepreke za ovaj metod. CNN metoda prevazilazi neke od tih problema i pokazuje ubedljivo bolje rezultate, ali i dalje ima svoje nedostatke. Rezultati nad FER2013 skupom nisu zadovoljavajući, jer CNN ne može da se izbori sa velikim razlikama između slika u datom skupu. Kao potencijalni način rešavanja ovog problema moguće je uvesti neku vrstu normalizacije podataka ili nadograditi mrežu uz pomoć prostornog transformera, Ovo bi omogućilo neuronskoj mreži da se fokusira ne bitne delove slike i zanemari obstrukcije poput zaštitnih znakova slike ili delimično prekrivenih delova lica.

	CK+	FER2013
CNN	94.47%	58.31%
SVM	87.7%	49.8%