

Detekcija i klasifikacija da li osobe na slici nose maske

Hajduk Dušan SW40/2018
Fakultet tehničkih nauka, Novi Sad

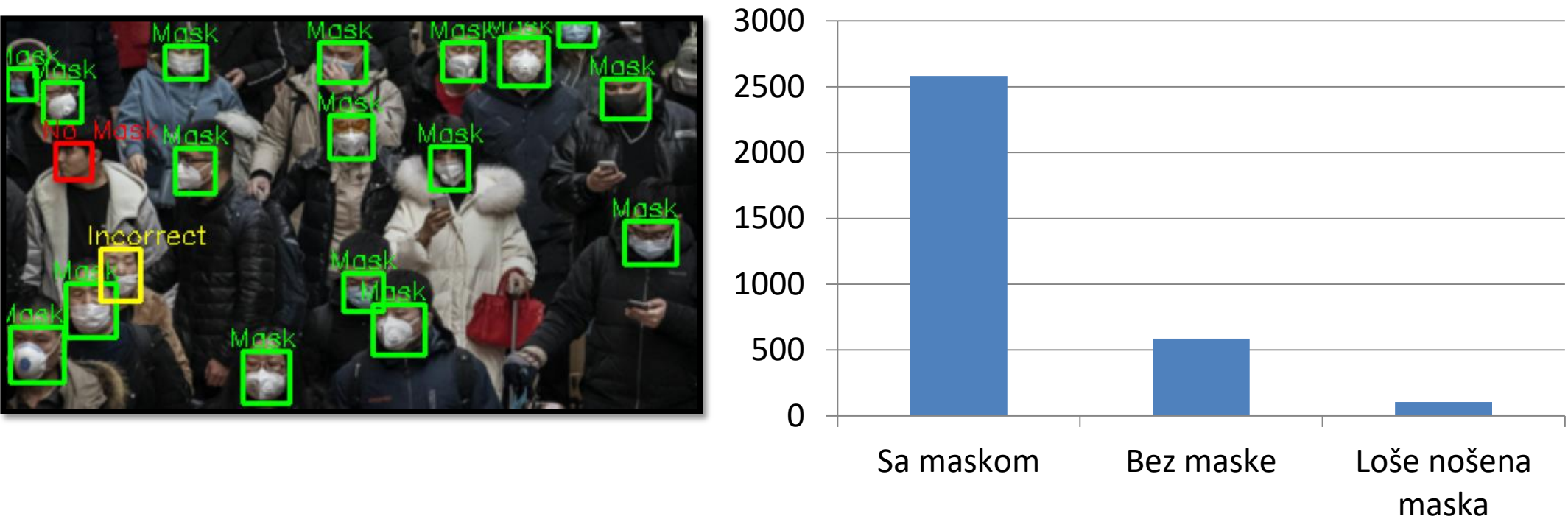


Definicija problema

Potrebno je detektovati sva lica ljudi na slici, i pri tome za svako lice odrediti da li se na njemu nalazi maska.

Skup podataka

Skup podataka koji će se koristiti sadrži slike na kojima se nalazi više ljudi. Uz svaku sliku dolazi i dodatni fajl koji labelira svaku poziciju svakog lica svojim okvirom kao i labelom da li je prisutna, ili dobro nošena maska.



Metodologija

Kao rešenje problema klasifikacije koristićemo tri binarna SVM klasifikatora. Prva dva klasifikatora koriste se za detekciju lica na slikama. Jedan za klasifikaciju lica bez maske i drugi za klasifikaciju lica sa maskom. Detekciju vršimo uz pomoć kombinacije *Sliding Window*, *Image Pyramid* i *Non-Maximum Suppression* tehnika. Treći binarni klasifikator određuje da li se na slici nalazi lice sa maskom ili bez maske, i prema njegovoj sigurnosti određivaćemo konačan rezultat naše predikcije.

Treniranje Klasifikatora

Svaku sliku ćemo podeliti na više malih slika koje sadrže lica. Potom ćemo filtrirati mutnije slike, i sve slike gde je lice manje od 20x20 izbaciti iz skupa podataka.

Posle toga potrebno je skalirati svaku sliku tako da budu iste veličine zbog prirode SVM klasifikatora. Za naš slučaj skaliraćemo slike na 40x40.

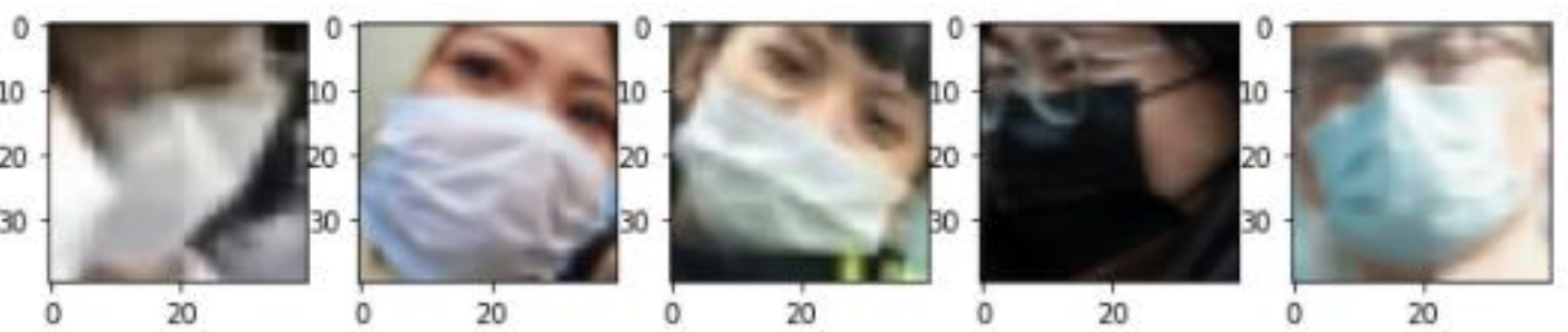
Negativne podatke za treniranje prva dva klasifikatora, tj slike na kojima se ne nalaze lica, generisaćemo tako što prolazimo kroz više slika iz našeg skupa podataka i biramo nasumične okvire koji se ne preklapaju sa labelisanim okvirima na kojima se nalaze lica.

Negativne podatke za treniranje trećeg klasifikatora već imamo, oni predstavljaju lica bez maski.

Broj slika	1 SVM	2 SVM	3 SVM
Trenirani Skup	1972	739	1582
Testni Skup	494	185	396

Preciznost	1 SVM	2 SVM	3 SVM
Trenirani Skup	99.5%	98.7%	97.9%
Testni Skup	96.3%	94.0%	94.9%

Iz testa preciznosti vidimo da su performanse ovih klasifikatora na zadovoljavajućem nivou, pogotovo kada uzmemo u obzir primer slika na kojima su predikcije bile pogrešne.



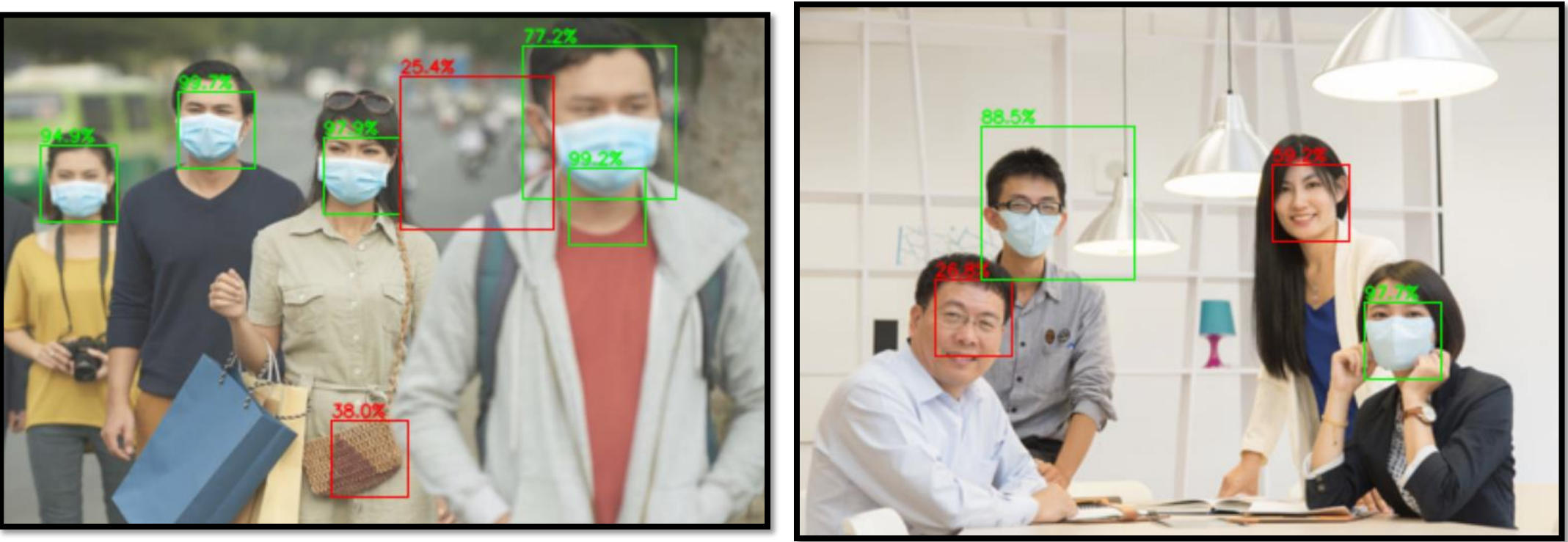
Detekcija

Kao što smo prethodno naveli detekciju lica radimo uz pomoć kombinacije *Sliding Window*, *Image Pyramid* i *Non-Maximum Suppression* tehnika.

Algoritam funkcioniše na sledeći način:

- Prvo prolazimo kroz sliku sa okvirom veličine 40x40, to će biti najmanji okvir sa kojim ćemo raditi. Za svaki okvir koji uzmemo potrebno je računati *HOG*, gde izdvajamo određene karakteristike svakog okvira.
- Potom izračunate vrednosti sprovodimo kroz prva dva binarna klasifikatora, i ukoliko predikcija prevazilazi određen prag, pamtimo taj okvir.
- Kada prođemo celu sliku grupišemo sve izdvojene okvire sa *Non-Maximum Suppression* tehnikom, skaliramo sliku na pola od njene prethodne veličine i krećemo ceo proces ponovog. Ovaj proces se završava kada veličina slike prevaziđe veličinu okvira.
- Potom grupišemo sve okvire koji se preklapaju iz različitih slojeva, tj veličina okvira.
- Na samom kraju za sve prikupljene okvire računamo *HOG* i sprovodimo dobijene vrednosti kroz treći *SVM* klasifikator. Prema sigurnosti klasifikatora i određenog praga koji postavimo određujemo da li se na licu nalazi maska ili ne.

Na slici vidimo rezultat naše detekcije. Procenat iznad svakog okvira predstavlja sigurnost trećeg *SVM* klasifikatora. Iz primera vidimo nekoliko *false positive* slučajeva u detekciji.



Rezultati

Preciznost detekcije računamo kao odnos tačnih predikcija i zbira tačnih i netačnih predikcija. Uzimanjem uzorka od 20 slika iz skupa podataka koji su korišteni za testiranje preciznosti *SVM* klasifikatora dobijamo 22.3% preciznosti.

Na ovaj broj najviše utiče to da se u testnim podacima nalazi velik broj lica koja su manja od 40x40, što ih čini teško uočljivim za klasifikator.

Druga stvar koja najviše utiče na ovaj broj jesu velike razlike kod skaliranja prozora jer se on u svakom prolazu povećava za dva puta.

Predikcija da li lica na detektovanim okvirima nose maske ili ne ima preciznost od 85.2%, što znači da detektovani okvir ne utiče previše mnogo na preciznost *SVM* klasifikatora.

Zaključak

Izvlačenje *HOG* karakteristika iz slika za detekciju lica se nije pokazalo kao idealno rešenje. Sledeći korak bi bio koristiti umesto *HOG*-a koristiti neki algoritam koji bolje prati konture i karakteristike lica.

Loša strana ovog programa jeste što se oslanjamo na dva sloja (prva dva klasifikatora u prvom sloju i treći klasifikator u drugom sloju) koji su međusobno zavisni. Prema tome najbolja preciznost klasifikacije da li osoba nosi masku se dobija množenjem preciznosti tih slojeva. Idealno rešenje bi se oslanjalo na jedan sloj klasifikatora.

Ukoliko bi hteli unaprediti ovaj program za detekciju na videu, morali bi promeniti ceo pristup, jer su vremenske performanse ovog programa poprilično loše, pogotovo kada radimo sa slikama velikih dimenzija. Bilo koja od iteracija *YOLO* algoritama bi značajno poboljšala ovaj problem.