

Detekcija i prepoznavanje brand logoa na slici

Žarko Drageljevic, Milan Desančić, Vladimir Indić

Motivacija i opis problema

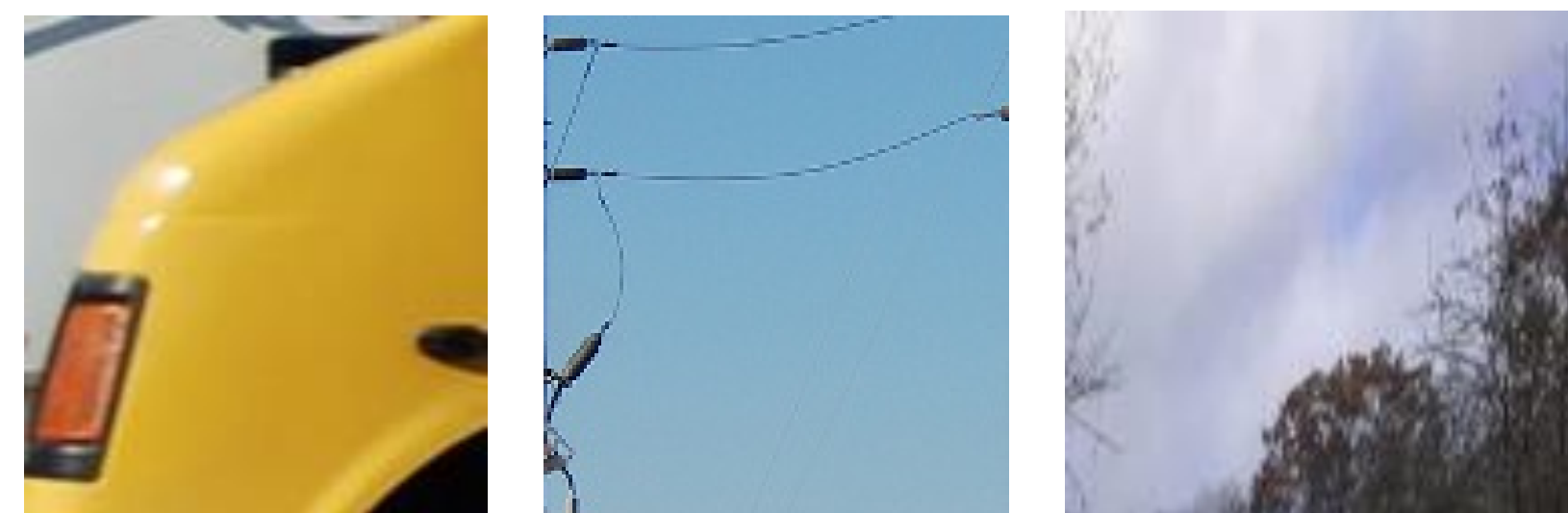
Veliki broj postojećih brendova i njihovih logoa doveo je do toga da ih je jako teško pamtit i raspoznavati. Zar ne bi bilo jednostavno da samo kamerom svog telefona pokažemo na odevni predmet, automobil ili svetlosnu reklamu prodavnice i dobijemo informaciju o kom brendu je reč. Da bi ovako nešto bilo moguće, potrebno je implementirati detektor i klasifikator brend logoa.

Dataset

Skup podataka bazira se na flickr-27 datasetu, koji se sastoji od 810 slika, za svaki od 27 logoa po 30 slika. Pojedine slike su lošeg kvaliteta, te su ručno izdvojene i zamenjene slikama koje su pronadjene na internetu. Primenjen je postupak augmentacije i generisano je po 500 slika dimenzija 128x128 za svaki logo. Pridodato je još oko 6000 slika koje logoe ne sadrže i one predstavljaju posebnu klasu. Ovakav skup podataka sadrži oko 20000 slika i podeljen je na skup za treniranje (70%), validaciju (10%) i testiranje klasifikatora(20%).



Nekoliko slika iz skupa podataka koje sadrže logo brenda



Nekoliko slika koje ne sadrže logo

Metod 1

HOG + NN + Sliding Window + NMS

1. Klasifikacija logoa na slici:

➤ HOG transformacija - uprosćavanje slike dimenzije 128x128 korespondentnim deskriptorom dimenzija 2304



Slika u izvornom obliku i njena uprosćena verzija nastala primenom HOG transformacije

➤ Neuronska mreža - klasifikator koji na ulazu prima HOG deskriptor odgovarajuće slike i određuje njenu klasu pripadnosti. Sastoji se od 4 sloja neurona koji redom imaju 2304, 1024, 256 odnosno 28 neurona. Uspešnost ostvarena prilikom treniranja nakon pete epohe na trening i validacijskom skupu podataka redom iznosi 99.64%, odnosno 96.89%, dok se prilikom testiranja na testnom skupu ostvaruje uspešnost od 96.68%.

2. Detekcija logoa na slici:

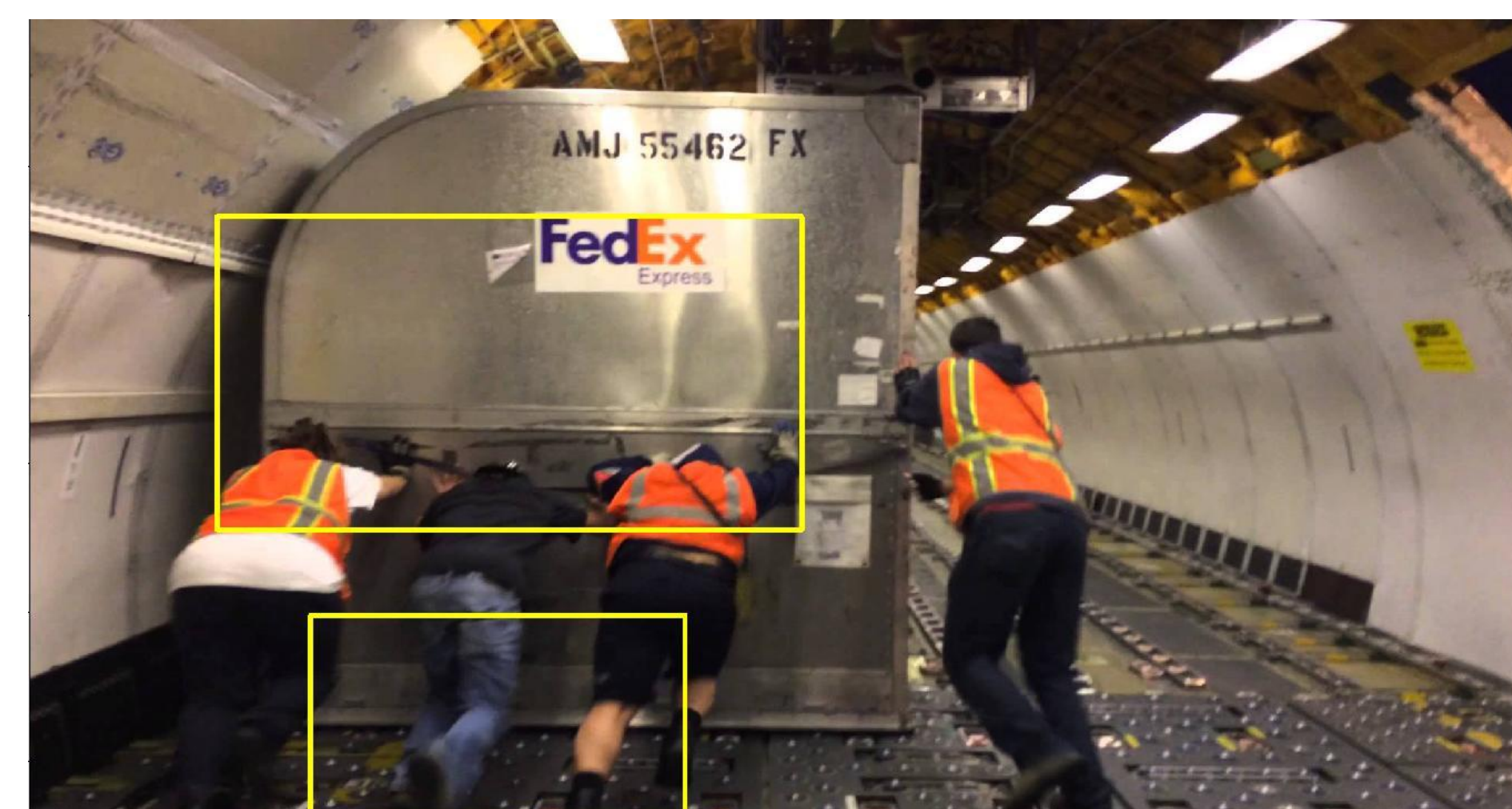
➤ Sliding Window (klizajući prozor) – region pravougaonog oblika dimenzija proporcionalnih polaznoj slici se pomera po istoj. Sa nje izdvaja manje slike na koje se primenjuje HOG transformacija i određeni deskriptor se dovodi na ulaz klasifikatora u cilju utvrđivanja postojanja logoa.

➤ Na sve regione za koje klasifikator utvrdi da sadrže logoe, primenjuje se NMS (Non-Maximum Suppression) algoritam. Na osnovu površine preklapanja regiona i sigurnosti predikcije postojanja logoa unutar njih, izdvajaju se najbolji region i oni predstavljaju rezultat detekcije.

Rezultati

Za verifikaciju metoda detekcije i prepoznavanja logoa neophodna je primena na slikama iz stvarnog sveta. Sa javno dostupnih repozitorijuma izdvojeno je 99 slika različitih dimenzija od kojih svaka ima po jedan logo. Zadatak metoda je da ga uspešno detektuje na slici i klasifikuje. Kao ocena uspešnosti posmatraju se tri parametra: broj pokušaja, broj tačnih pogodaka i broj promašaja. Ostvarena je sledeća uspešnost:

- Broj pokušaja: 99
- Broj pogodaka: 21
- Broj promašaja: 322.



Uspešno detektovanje znaka FedEx, ali takođe i detektovanje logoa tamo gde on ne postoji

Metod 2

CNN + Selective Search + NMS

1. Klasifikacija logoa na slici:

➤ Konvoluciona neuronska mreža (CNN) koju čine:

- Konvolucioni sloj (Conv2D): 32 filtera, veličina prozora 5x5
- Pooling sloj (MaxPooling2D): veličina prozora 2x2
- Dropout sloj - zadužen za smanjenje overfittovanja (parametar rate postavljen na 0.2)
- Flatten sloj – matricu prevodi u niz
- Potpuno povezan sloj od 256 neurona čija je aktivaciona funkcija relu
- Izlazni sloj od 28 neurona čija je kativaciona funkcija softmax

Uspešnost ostvarena prilikom treniranja u petoj epohi na trening setu iznosi 99.85%, odnosno 99.79% na validacijskom setu, dok je uspešnost ostvarena prilikom testiranja na testnom setu 99.48%.

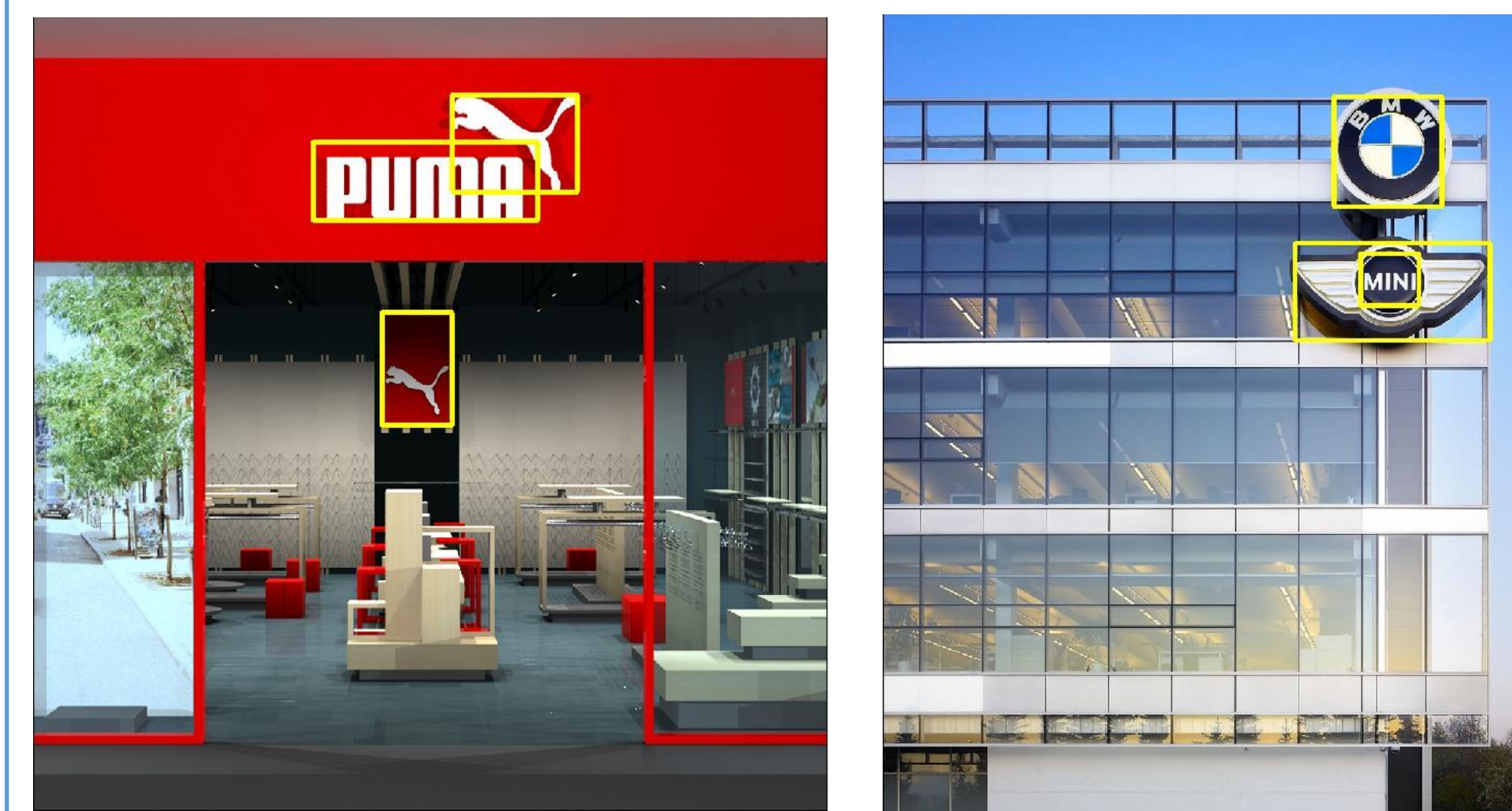
2. Detekcija logoa na slici:

➤ Selective Search algoritam – podržan u Pythonovoj biblioteci OpenCV. Služi za grupisanje i izdvajanje regiona na slici koji su slični po određenim karakteristikama kao što su npr. boja i tekstura. Izdvojeni regioni se puštaju na ulaz konvolucione neuronske mreže koja određuje kojoj klasi i sa kojom sigurnošću joj pripadaju. Ukoliko je izdvojeno više karakterističnih regiona sa logoima, na njih se primenjuje NMS algoritam (objašnjen u prethodnom poglavlju) radi izdvajanje onog regiona koji sa najvećom sigurnošću sadrži logo.

Rezultati

Na istom skupu od 99 slika obavljena je verifikacija ovog metoda. Ostvarena je sledeća uspešnost:

- Broj pokušaja: 99
- Broj uspešnih pogodaka: 88
- Broj promašaja: 7



Uspešno detektovanje svih logoa prikazanih na slici

Zaključak

➤ Upoređujući rezultate oba metoda vidimo da i jedan i drugi dobro rešavaju problem prepoznavanja datog logoa (klasifikacije).

➤ Sa detekcijom logoa na slici ipak nije tako. Prvi metod pokazuje izuzetno loše rezultate i to prvenstveno zbog primene klizajućeg prozora. On predstavlja “brute-force” način za pretragu regiona slike. Jako sporo traje, jer prolazi kroz sve njene delove, kako ne bi nešto propustio i ne vodi računa o tome da li delovi slike koje poklapa imaju međusobnu zavisnost. Osim toga, njegova veličina se mora predefinisati pre upotrebe. Logoi sami po sebi imaju različit oblik i dimenzije. Neki su uski, drugi duguljasti, a treći pak kvadratnog oblika ili okrugli, što dodatno otežava upotrebu jednoličnog klizajućeg prozora.

➤ Zbog kompleksnosti problema mora se koristiti pametan način pretrage slike po regionima koji zavise jedni od drugih. Na taj način se pretraga znatno ubrzava, jer sama dovodi do isključenja pojedinih regiona koji nisu od važnosti. Iako ni drugi opisan metod ne ostvaruje izvanredne rezultate, on je ipak dovoljan da se prikaže kako osnovna pretraga slike po regionima zavisnosti dovodi do znatnog poboljšanja rezultata detekcije. Ovaj metod se može smatrati uprosćenom verzijom današnjih rekurentnih neuronskih mreža (R-CNN, Fast R-CNN i Faster R-CNN) koje predstavljaju “state of the art” oblasti koja se bavi detekcijom objekata.