

Detekcija objekata u slikama

Studentski projekt 2018/19

Izv. prof. dr. sc. Zoran Kalafatić

Zoran.Kalafatic@fer.hr

ZEMRIS, 2018/19

Članovi tima

- Zorana Doždor
- Domagoj Jakšić
- Marko Kršić
- Lovro Ludvig
- Mihael Macuka
- Franko Pandžić

Cilj projekta

- ostvariti programski sustav za detekciju osoba u slikama
- steći iskustva rada u timu
- aktivno koristiti odabrani programski jezik
- implementirati zadanu komponentu uz korištenje raspoloživih knjižnica programskog kôda
- povezati svoje rezultate s gotovom komponentom
- upoznati se s osnovama računalnog vida i strojnog učenja

Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



Klizno detekcijsko okno

- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



- mi ćemo se ograničiti na ponašanje detektora na jednoj poziciji

Klizno detekcijsko okno

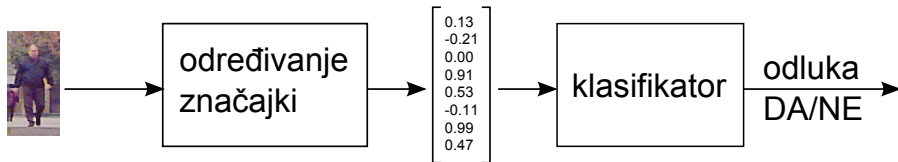
- učitana slika se pregledava pomicanjem detekcijskog okna s nekim korakom, te na različitim *skalama*
- na svakoj poziciji detektor donosi odluku: objekt ili ne



- mi ćemo se ograničiti na ponašanje detektora na jednoj poziciji
- ispitivanje sustava – postavljanjem isječaka s pješacima i bez njih na ulaz

Detektor

- iz sličice izrezane iz detekcijskog okna računaju se *značajke*
- na temelju značajki prethodno naučeni (trenirani) klasifikator donosi odluku



Značajke

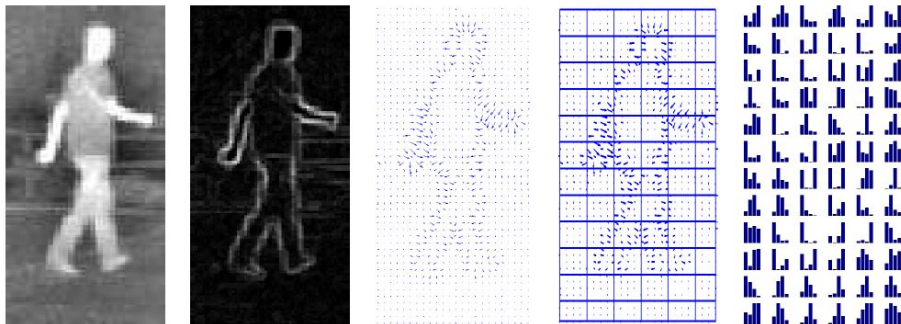
- brojčani podaci izračunati iz slike, obično zapisani u vektor
 - ▶ prosječna svjetlina slike, histogram boje
 - ▶ vrijednosti svih slikovnih elemenata složene u vektor
 - ▶ vrijednosti izračunate iz nekih kombinacija slikovnih elemenata
- u našem slučaju značajke se određuju na temelju *orijentacija gradijenata* u pojedinim područjima slike

Klasifikator

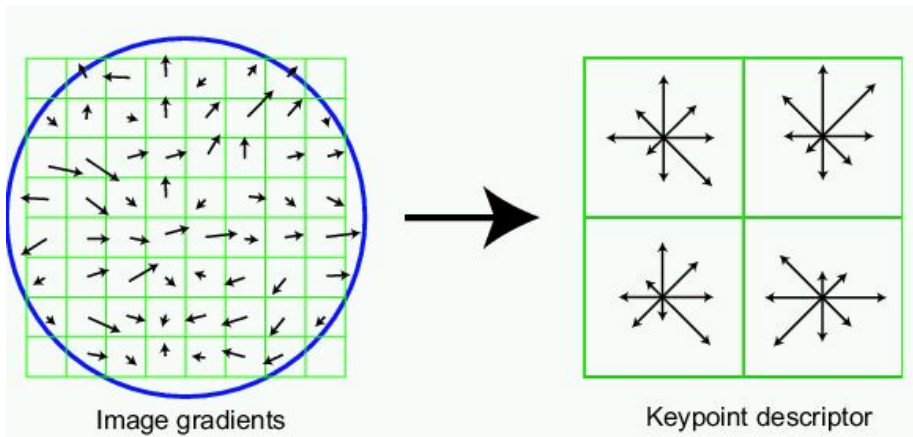
- najprije ga moramo *naučiti* da razlikuje uzorke iz različitih razreda
- na temelju *uzoraka za učenje* klasifikator uči granicu između razreda – *decizijska funkcija*
- u našem slučaju uzorci za učenje su sličice pješaka i sličice bez pješaka
- iz tih sličica računaju se *vektori značajki* – ulaz u klasifikator
- klasifikator iterativno nastoji pronaći/poboljšati decizijsku funkciju kako bi što bolje razdvojio uzorke za učenje u zadane razrede
- umjetne neuronske mreže (ANN)
- stroj s potpornim vektorima (SVM)
- mi ćemo koristiti gotovu komponentu
 - ▶ LIBSVM
 - ▶ SVM u biblioteci OpenCV
 - ▶ bit će potrebno proučiti kako se koriste, odabrati prikladniju
 - ▶ prilagoditi izlaz iz razvijene komponente za računanje značajki
 - ▶ trenirati klasifikator, ispitati sustav na test-primjerima

Značajke temeljene na deskriptoru HOG

- na ulaznoj slici računa se gradijent (veličina i smjer)
- slika se dijeli na ćelije
- u svakoj ćeliji računa se histogram orijentacija gradijenta
- vrijednosti pretinaca histograma svih ćelija ulančavaju se u vektor značajki



- na ulaznoj slici računa se gradijent (veličina i smjer)
- slika se dijeli na ćelije
- u svakoj ćeliji računa se histogram orijentacija gradijenta



Gradijent

- vektor koji pokazuje smjer i veličinu najveće promjene funkcije
- u našem slučaju funkcija predstavlja vrijednosti slikovnih elemenata u sivoj slici
- u svakoj poziciji slike računa se gradijent u x i u y smjeru

$$\nabla_x f(x, y) = \frac{1}{2\Delta x} (f(x + \Delta x, y) - f(x - \Delta x, y))$$

$$\nabla_y f(x, y) = \frac{1}{2\Delta y} (f(x, y + \Delta y) - f(x, y - \Delta y))$$

- primjena jednostavnih konvolucijskih maski

$$\begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

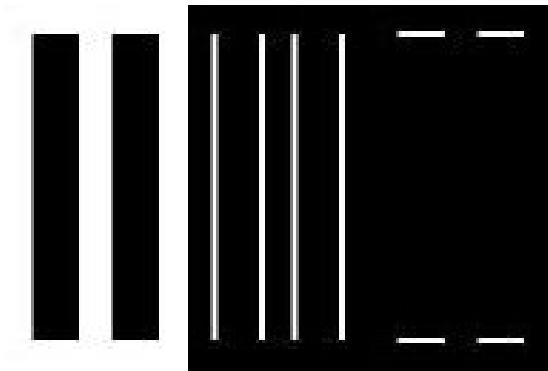
Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



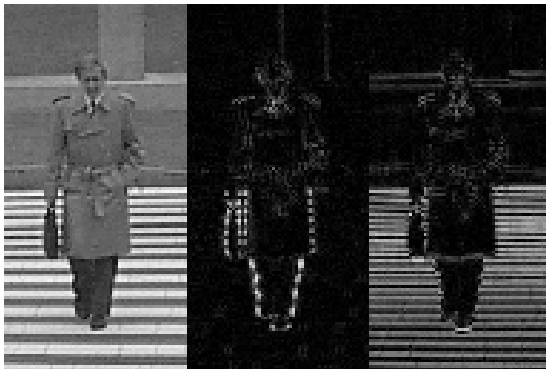
Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



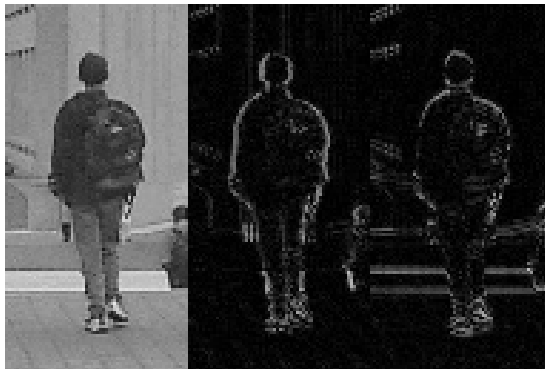
Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



Računanje gradijenta

- gradijenti u smjeru x i y (horizontalni i vertikalni gradijent)
- na svakoj poziciji u slici definiran je vektor



Ocjena uspješnosti postupaka za detekciju objekata

- detektor daje listu pravokutnika u slici za koje “vjeruje” da sadrže primjerak traženog objekta (osoba, prometni znak, nogometaš)
- da bismo automatizirano mogli prebrojati uspješne odnosno neuspješne detekcije, trebaju nam referentne oznake (eng. *ground truth*)
- ispitni skup slika s pripadnim popisom položaja traženih objekata
- broje se:
 - TP – točne detekcije (*true positives*)
 - FP – lažne detekcije (*false positives*)
 - FN – ispuštene detekcije (*false negatives*)
 - TN – pravi negativni (*true negatives*)
- kako uopće brojati pogotke/promašaje? → omjer površine presjeka i unije (*intersection-over-union, Jaccardov indeks*)

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} > \theta$$

Referentne oznake i izlaz detektora



Mjere uspješnosti

- odziv (*recall, hit rate, sensitivity, true positive rate*)

$$R = \frac{TP}{TP + FN}$$

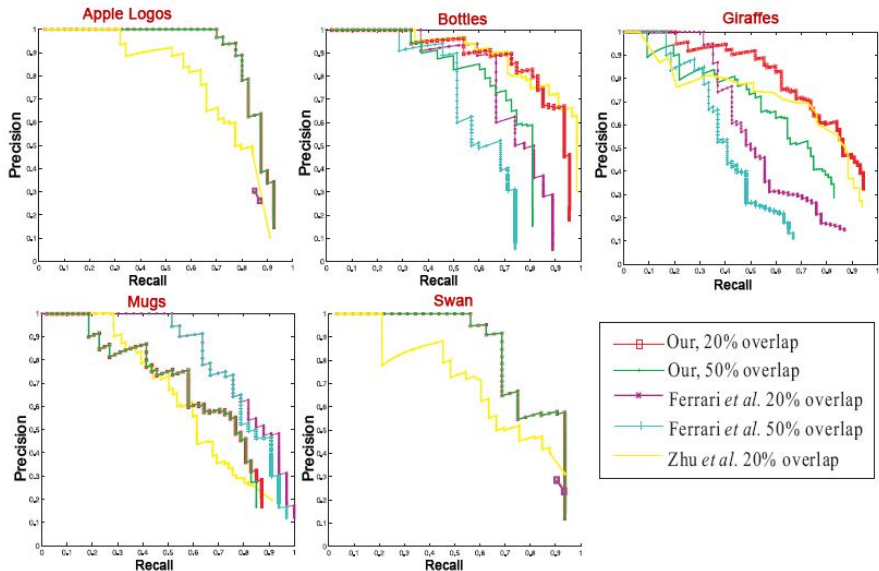
- preciznost (*precision*)

$$P = \frac{TP}{TP + FP}$$

- F1 mjera – harmonijska sredina preciznosti i odziva

$$F = 2 \cdot \frac{R \cdot P}{R + P}$$

Krivulja preciznost-odziv



Datoteka s oznakama (primjer)

894, 3, 1, 246, 340, 11, 22

894, 0, 2, 387, 410, 9, 25

894, 0, 3, 439, 360, 8, 20

...

894, 0, 11, 767, 306, 9, 14

894, 2, 12, 908, 345, 8, 17

894, 1, 13, 413, 367, 9, 18

894, 1, 14, 440, 352, 9, 17

...

894, 1, 22, 653, 299, 8, 16

- redak:

frameNo, oznakaTima (bijeli, plavi, golman1, golman2),
playerID, x i y koordinate lijevog gornjeg vrha opisanog
pravokutnika, širina i visina pravokutnika

Poslovi

- izbor platforme
 - ▶ C/C++ i OpenCV
 - ▶ Python i OpenCV; relativno jednostavna migracija na C/C++
 - ▶ nešto drugo?
- uhodavanje odabrane platforme
 - ▶ instalacija, pokretanje početnih primjera, učitavanje, prikaz i spremanje slike
- upoznavanje s implementacijskim detaljima Dalala i Triggsa [1]
 - ▶ programska implementacija jednostavne verzije HOG deskriptora
- prikupljanje uzoraka za učenje i testiranje
 - ▶ javno raspoložive baze slika osoba (MIT, Dalal & Triggs)
 - ▶ slike bez osoba!
- upoznavanje s bibliotekama *OpenCV* i *LIBSVM*
 - ▶ obratiti pažnju na jednostavnost korištenja i prikladnost za spajanje s vašim podsustavom za računanje značajki
- probati pokrenuti neki od postojećih detektora objekata (npr. detektor osoba iz biblioteke OpenCV [1]), usporediti rezultate

Rokovi

- prva verzija plana projekta – 6.11-12.11.2017.
 - ▶ predati kroz FERweb
 - ▶ odabrati voditelja tima
- primopredaja projekata – 15.1.-26.1.2018.
 - ▶ plan rada (ažurirana, dopunjena, realizacija)
 - ▶ tehnička dokumentacija
 - ▶ prezentacija
- dodjela bodova
 - ▶ nastavnik – do 31.1.2018.
 - ▶ studenti – do 2.2.2018.

Komunikacija

- mail
- sastanci (s mentorom – rjeđi)
- sastanci (bez mentora – češći)
- Google group
(pozivnica na: `zoran.kalafatic@gmail.com`)
- slack?

Literatura



N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893. [Online]. Available: <http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>