

Strojno učenje

Detekcija kaskadom boostanih Haarovih klasifikatora

Siniša Šegvić

Zavod za elektroniku, mikroelektroniku,
računalne i inteligentne sustave
Fakultet elektrotehnike i računarstva
Sveučilište u Zagrebu

Kombiniranje strojno naučenih algoritama

1. Pregled značajnijih pristupa:

- ☐ nezavisno naučeni eksperti (varijante glasanja)
- ☐ višerazinsko učenje (npr: boostanje, kaskadiranje)
- ☐ Alpaydin, poglavlje 15

2. **Postupak Viole i Jonesa:**

- ☐ o detekciji kaskadom ojačanih (boostanih) Haarovih klasifikatora
- ☐ P. Viola and M. Jones, Robust Real-Time Face Detection, IJCV04.

UVOD: KONTEKST

Promotrit ćemo jednu uspješnu primjenu boostanja u računalnom vidu

Računalni vid: izvlačenje simboličkih podataka iz slika

1. 3D rekonstrukcija
2. praćenje objekata kroz slijed slika
3. prepoznavanje objekata (koji tip automobila je ispred rampe?)
4. **detekcija objekata** (ima li koje lice na slici?)

Detekcija objekata:

1. detekcijom specifičnih geometrijskih struktura:
relativno brzo, **gubitak informacije**
2. binarnom klasifikacijom u pokretnom detekcijskom oknu:
računski iznimno zahtjevno, prikladno za **strojno učenje**

UVOD: DETEKCIJA KLASIFIKACIJOM

Ideja: pomicati detekcijsko okno po slici, u svakom položaju pozvati binarni klasifikator

- u svakoj slici ima oko $2.5e6$ interesantnih lokacija ($720*576*60$)
- detekcija treba raditi u stvarnom vremenu...

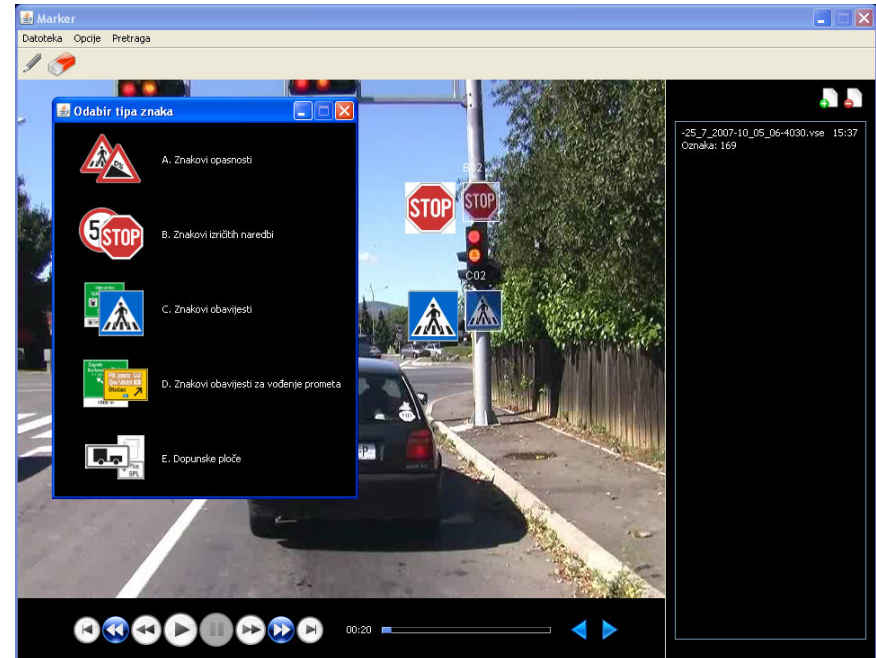


- zbog ogromne računske složenosti, značajke moraju biti jednostavne!
- jednostavne značajke \Rightarrow slabi klasifikatori
- logično rješenje: kaskada boostanih slabih klasifikatora!

Što je s jakim klasifikatorima?

- ne možemo ih izračunati $2.5e6$ puta u 40 ms...

UČENJE: OZNAČAVANJE

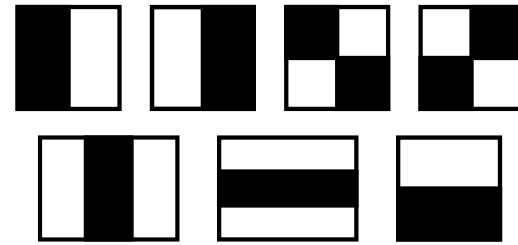


HAAR: ZNAČAJKA

Haarove značajke su pogodne za brzinsku obradu:

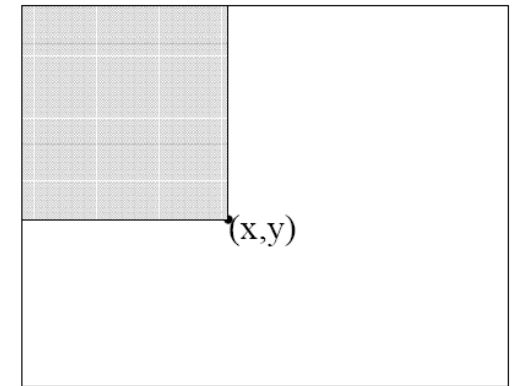
□ $H_i(x, y, sx, sy) = \sum I_{\text{bijeli}} - \sum I_{\text{crni}}$

- mogu se vrlo brzo izračunati
primjenom **integralne slike**



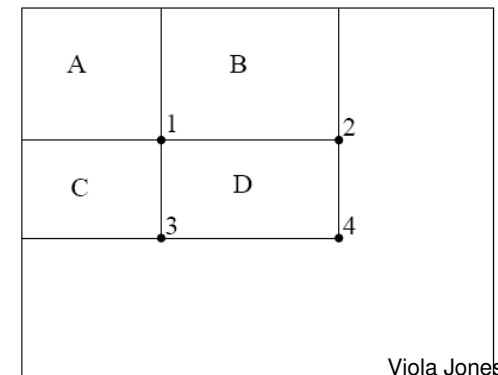
Integralna slika u svakom pikselu sadrži integral po pravokutniku čiji drugi vrh je u ishodištu slike:

$$ii(x, y) = \sum_{x' < x, y' < y} i(x', y')$$



Za izračunavanje pravokutnog integrala potrebna su **samo 4** pristupa integralnoj slici:

$$D = ii_4 - ii_2 - ii_3 + ii_1$$

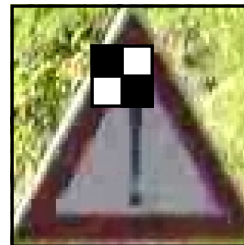


HAAR: KLASIFIKATOR

Kako od značajke $H_i(x, y, sx, sy)$ doći do binarnog klasifikatora?

- fiksirati položaj i mjerilo unutar detekcijskog okna
(svaki tip značajke inducira velik broj klasifikatora)
- testirati vrijednosti značajke parametrima p_j (polaritet) i θ_j (prag)

$$h_j(\mathbf{x}|p_j, \theta_j) = \begin{cases} 1 & \text{ako } p_j f_j(x) < p_j \theta_j, \\ 0 & \text{inače} \end{cases}$$



HAAR: OKNO

Klasifikatore $h_j(\mathbf{x}|p_j, \theta_j)(\mathbf{x})$ evaluiramo u detekcijskom oknu:

- \mathbf{x} ... vektor piksela iz detekcijskog okna
- osnovna veličina okna je $24 \times 24 \Rightarrow$ oko 45000 različitih klasifikatora
- za veća okna **skaliramo klasifikatore**
(**brže** od konstrukcije rezolucijske piramide!)



DETEKCIJA: BOOSTANJE

Slabi Haarovi klasifikatori se kombiniraju u jaki klasifikator **boostanjem**:

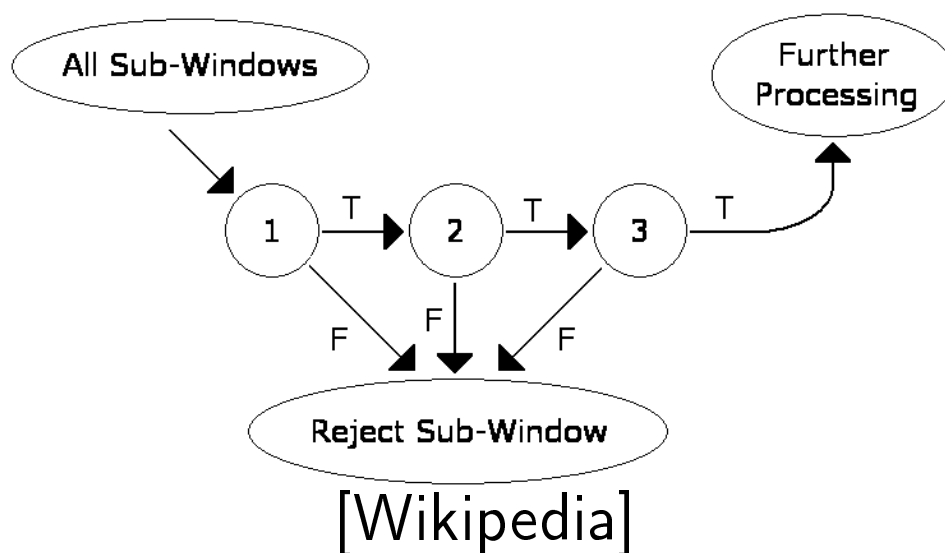
- ulaz: označeni skup za učenje (\mathbf{x}_i, y_i) , te željeni broj T
- inicijaliziraju se težine elemenata skupa za učenje w_i
- for t in $\{1, 2, \dots, T\}$:
 - iscrpnim pretraživanjem pronalazi se klasifikator $h_t(\mathbf{x})$ s najmanjom pogreškom $\epsilon = \sum_i w_i [h_t(\mathbf{x}_i) \neq y_i]$
 - određuju se β_t i $\alpha_t = \log(1/\beta_t)$
 - težine w_i se ažuriraju s obzirom na β_t , te normaliziraju
- konačni jaki klasifikator je:

$$h(\mathbf{x}) = \begin{cases} 1 & \text{ako } \sum_{t=1}^T \alpha_t h_t(\mathbf{x}) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0 & \text{inače} \end{cases}$$

DETEKCIJA: KASKADIRANJE

Problem: za zadovoljavajuće rezultate treba previše slabih klasifikatora

- formiramo binarnu kaskadu: na višim razinama složeniji klasifikatori
- jaki klasifikatori se konfiguriraju tako da:
 - odbacuju minimalan broj pozitivnih primjera,
 - propuštaju poneki negativni primjer
- na većini lokacija prvi jaki klasifikator daje ispravan neg. odgovor



REZULTATI: DETEKCIJA

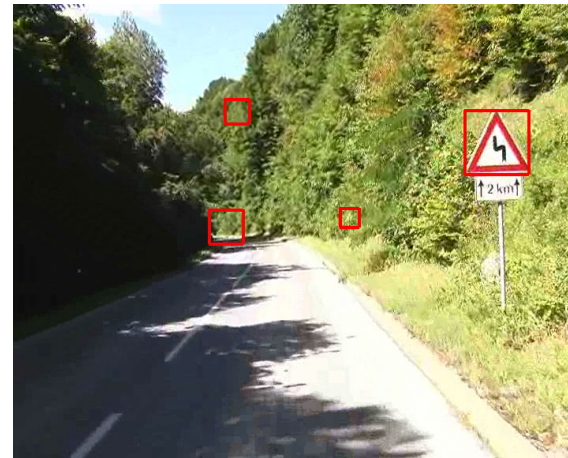
Rezultati testiranja detektora na dva različita ispitna skupa:

Skup	Faktor skaliranja	Znakova	Pogodaka [% ispitni skup]	Promašaja [% ispitni skup]	Lažno pozitivnih [% ispitni skup]
1	1.05	101	96 %	4 %	84 %
1	1.20	101	93 %	7 %	42 %
2	1.05	72	93 %	7 %	163 %
2	1.20	72	90 %	10 %	53 %

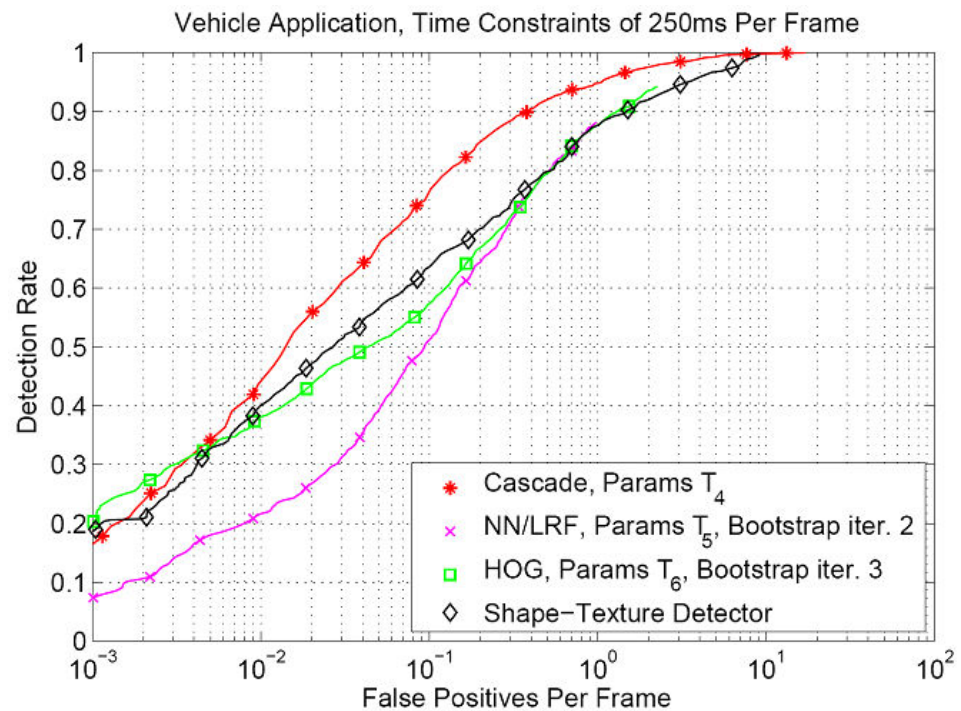
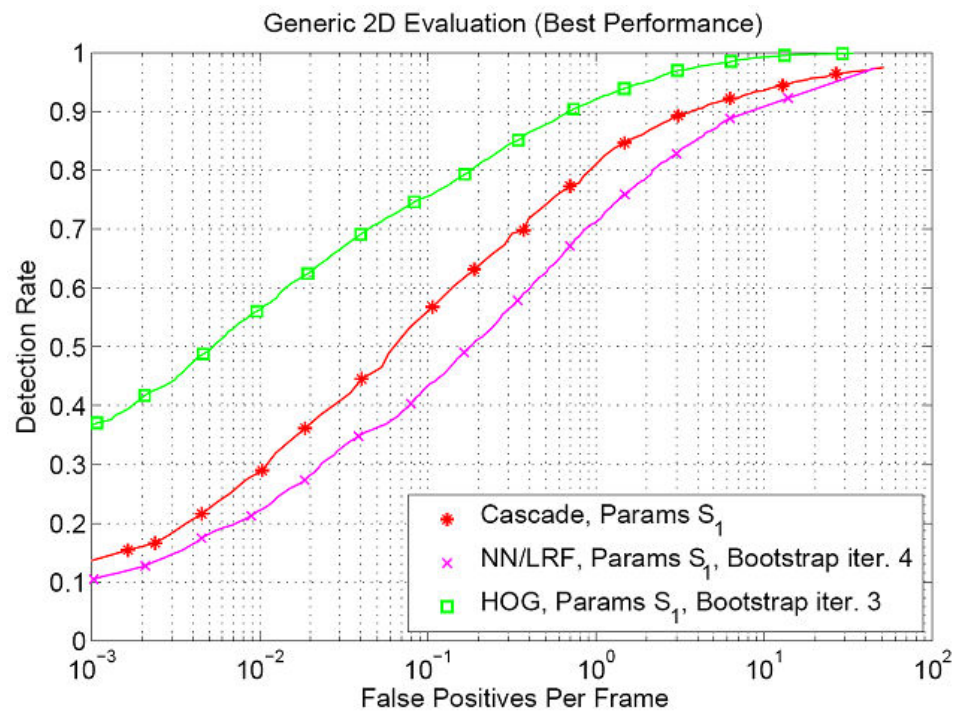
Raniji rezultat, 352 slike za učenje:

- ☐ odziv: 68%
- ☐ lažne pozitivne detekcije: 46%

REZULTATI: PRIMJERI

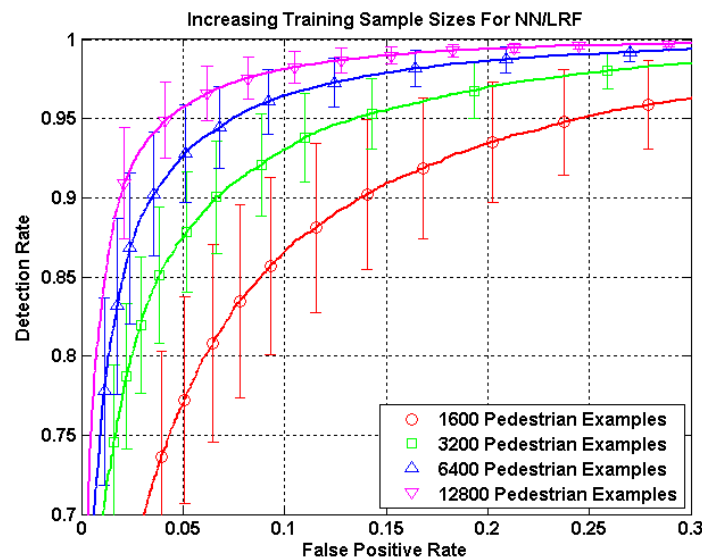


REZULTATI: VJ vs HOG SVM



[MunderGavrillaPami06]

REZULTATI: GROUNDTRUTH



[MunderGavrillaPami06]

Intrigantni rezultati:

- ROC performansa raste s povećanjem skupa za učenje:
- efekti zasićenja nisu zamijećeni ni za $N = 12.800$
- udvostručenje skupa za učenje postiže bolje rezultate od odabira najbolje kombinacije značajka - klasifikator.

REZULTATI: KRAJ