

Hardverska implementacija Viola-Jones algoritma

Risto Pejašinović

Sadržaj

1	Viola-Jones algoritam	3
1.1	Uvod	3
1.2	Integralna slika	3
1.3	AdaBoost i HAAR obeležja	5
1.3.1	HAAR obeležja	5
1.3.2	AdaBoost	6
1.4	Kaskadni klasifikator	7

1 Viola-Jones algoritam

1.1 Uvod

Namena algoritma je detekcija i lokalizacija objekata na slici. Osmišljen od strane Paul Viola i Michael Jones 2001. godine [1].

Dugo godina je zbog brze i pouzdane detekcije bio standardan način detekcije lica na slici. I danas je prisutan u velikom broju mobilnih telefona i digitalnih kamera, ali danas postaje polako zamenjen konvolucionim neuronskim mrežama.

Pouzdanost i brzina su postignuti uvođenjem tri ključna doprinosa:

- **Integralna slika** omogućava brzo izračunavanje obeležja.
- **AdaBoost** algoritam za učenje, odabiranjem obeležja povećava brzinu i pouzdanost detekcije.
- **Kaskadni klasifikator** Realizovanjem algoritma u kaskadama omogućava brzo odbacivanje pozadine slike kako je mala verovatnoća da će se tu naći lice.

1.2 Integralna slika

Kao jedan od ključnih delova algoritma, integralna slika omogućava izračunavanje površine svakog pravouganog obeležja u konstantnom vremenu.

Intenzitet piksela u integralnoj slici na poziciji x,y je zbir svih piksela koji se nalaze gore i levo od pozicije x,y.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

Gde je ii(x,y) integralna slika, a i(x,y) originalna slika.

1	1	1
1	1	1
1	1	1

Ulazna slika

1	2	3
2	4	6
3	6	9

Integralna slika

Slika 1: Primer integralne slike

Piksele integralne slike je moguće računati u paraleli, ili sekvencijalno. Izbor algoritma za računanje integralne slike značajno utiče na performanse i potrebne hardverske resurse.

U paralelnoj implementaciji cena je više pristupa memoriji i više potrebnih sabirača,

dok je kod sekvencijalne implementacije manja brzina.

Osobina koja integralnu sliku čini pogodnu za korišćenje u Viola-Jones algoritmu je da je za računanje bilo koje pravougaone površine unutar integralne slike potrebno 2 oduzimanja i 1 sabiranje.

Originalna					Integralna				
5	2	3	4	1	5	7	10	14	15
1	5	4	2	3	6	13	20	26	30
2	2	1	3	4	8	17	25	34	42
3	5	6	4	5	11	25	39	52	65
4	1	3	2	6	15	30	47	62	81

$5 + 4 + 2 + 2 + 1 + 3 = 17$					$(D) - (B) - (C) + (A) = S$ $34 - 14 - 8 + 5 = 17$				
------------------------------	--	--	--	--	--	--	--	--	--

Slika 2: Primer računanja površine pravougaonika [2]

Na slici (2) je prikazano računanje površine pravougaonika na originalnoj slici i na integralnoj slici. Kao što se može videti za površinu pravougaonika MxN na originalnoj slici nam je potrebno MxN-1 sabiranja. Dok je kod integralne slike broj operacija 2 oduzimanja i 1 sabiranje i ne zavisi od dimenzija pravougaonika.

$$\sum_{(x,y) \in ABCD} i(x,y) = ii(D) + ii(A) - ii(B) - ii(C) \quad [3] \quad (2)$$

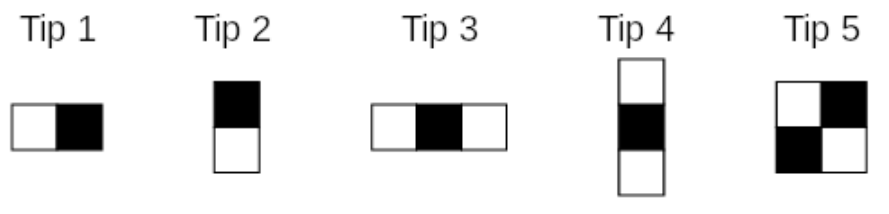
1.3 AdaBoost i HAAR obeležja

1.3.1 HAAR obeležja

Detekcija na slici se vrši na prozorima manjih dimenzija. Na ovim prozorima računaju se HAAR obeležja koja se sastoje od dva ili više pravougaonika, kao na slici (3).

Svako obeležje se računa sabiranjem piksela crnih pravougaonika potom oduzimanjem zbira piksela belih pravougaonika.

Reprezentacija prozora pomoću integralne slike omogućava da se obeležja izračunavaju u konstantnom vremenu.



Slika 3: HAAR obeležja [4]

Dimenzije prozora se razlikuju od modela klasifikatora. Jedan često korišćeni model koristi prozor dimezije 24x24.



Slika 4: Primer obeležja za detekciju lica [1]

Oblik odabranih obeležja će zavistiti od namene detektora, na slici(4) se mogu videti dva tipična obeležja koja su od interesa za detekciju lica.

Prvo obeležje namenjeno je merenju razlike intenziteta regiona čela i očiju. Obeležje koristi činjenicu da je oblast čela svetlija od očiju.

Dok drugo obeležje poredi intenzitet regiona mosta nosa sa očima.

Kako su obeležja od interesa koja se koriste u modelima na slici(3). Za datu dimenziju prozora kombinacije svih mogućih varijacija oblika i pozicija datih obeležja čini skup od 160.000 različitih obeležja. Kako je većina ovih obeležja slična i davaće

slične rezultate, ovaj broj se može drastično smanjiti korišćenjem algoritma za učenje AdaBoost.

1.3.2 AdaBoost

Kako je već rečeno može se dobiti oko 160.000 obeležja za prozor dimenzije 24x24. Od ovog broja samo neka obeležja mogu dati dobre rezultate prilikom detekcije, kao na slici(4) u primeru detektora lica.

Kako bi se odabrao skup korisnih obeležja može se koristiti neki od algoritama mašinskog učenja. Viola i Jones predlažu modifikovani AdaBoost algoritam.

Ideja AdaBoost-a je kombinovanje više *weak learner*-a kako bi se dobila pouzdana detekcija.

Weak learner je kalsifikator koji ima pouzdanost pogađanja malo bolju nego nasumičnu. Odnosno pouzdanost *weak learner*-a mora biti bar malo iznad 50%.

Kombinacijom ovako dobijenih *weak learner*-a može se dobiti *strong classifier*

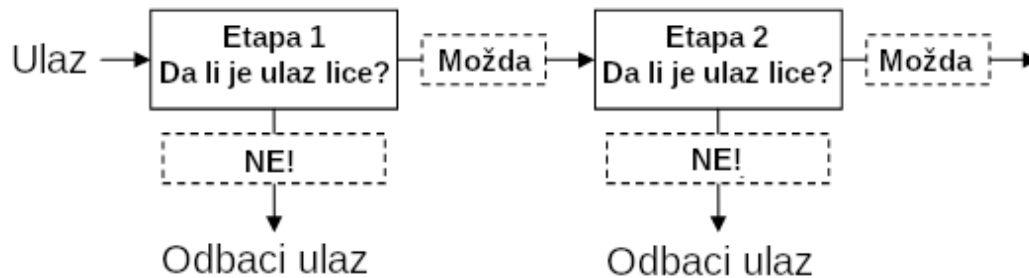
Kao rezultat AdaBoost algoritma dobićemo skup obeležja, u ovom radu se koristi model sa skupom od 2913 obeležja.

1.4 Kaskadni klasifikator

Osnovni princip Viola-Jones algoritma je da se na osnovu svih obeležja u modelu dobije informacija da li se na trenutnom položaju prozora nalazi traženi objekat (npr. lice).

Kako na slici većina skeniranih regiona ne sadrži lice, računanje svih obeležja na svakoj poziciji bi bilo suvišno. Tako da je korišćenje jednog jakog klasifikatora neefikasno.

Ideja obrazovanja kaskadnog klasifikatora je da se prozori na kojima se očigledno ne nalazi lice odbace brzo, nakon samo nekoliko izračunatih obeležja.



Slika 5: Kaskadni klasifikator [4]

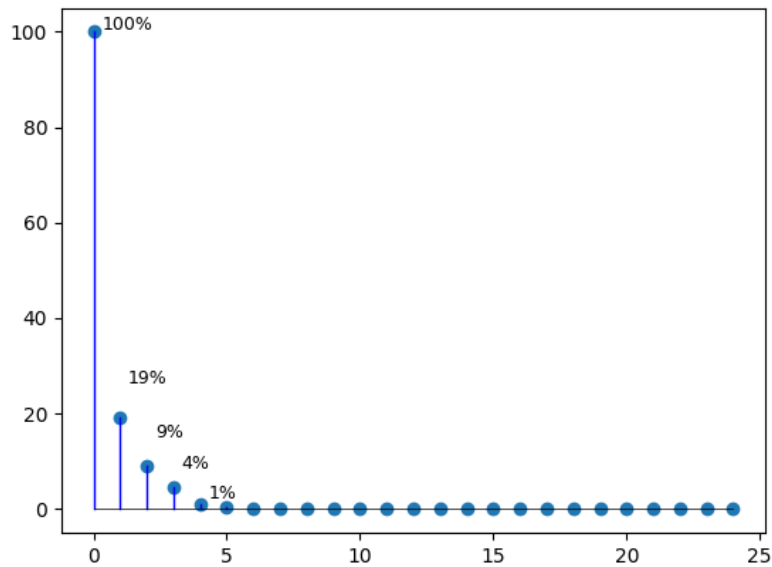
Kako bi se prozori bez lica brzo odbacili predlog je da se jaki klasifikatori grupišu u etape (eng. *stage*). Svaka etapa treba da bude dobra u odlučivanju da li se na analiziranom prozoru definitivno ne nalazi lice. Ukoliko je to slučaj taj prozor će se brzo odbaciti.

Ukoliko rezultat etape ukazuje na to da se na prozoru možda nalazi lice, preći će se na izvršavanje sledeće etape.

Konačno ukoliko sve etape u klasifikatoru na analiziranom prozoru daju rezultat da se možda nalazi lice, može se zaključiti da se na toj poziciji zaista nalazi lice. Zahvaljujući ovome postiže se veoma pouzdan klasifikator sa malim procentnom pogrešno negativnih (eng. *false negative*) rezultata na krajnjim etapama.

Kao primer u ovom radu će se koristiti model kaskadnog klasifikatora za prepoznavanje lica, sa 25 etapa i 2913 obeležja raspoređenih po etapama.

U prvoj etapi se nalazi samo 9 obeležja, dok taj broj raste do 211 u kasnijim etapama.



Slika 6: Procentat prolaska etapa

Na slici(6) je prikazana statistika izvršavanja etapa na *Caltech Dataset-u*[5], koji sadrži 450 slika od 27 različitih ljudi pod različitim osvetljenjima, izrazima i pozadinama.

Vrednosti različitih tačaka na grafu predstavlja procenat izvršavanja date etape na svih 450 slika. Prva etapa će se naravno uvek izvršiti, dok će se druga etapa izvršiti samo u 19% analiziranih prozora, druga 9% itd...

Vidimo da je posle pete etape procenat izvršavanja manji od 1%.

Literatura

- [1] P. A. Viola and M. J. Jones, “Rapid object detection using a boosted cascade of simple features,” in *CVPR*, 2001.
- [2] A. Jain, “Computer vision – face detection,” 2016. [Online]. Available: <https://vinsol.com/blog/2016/06/28/computer-vision-face-detection/>
- [3] K. Cen, “Study of viola-jones real time face detector,” 2016.
- [4] O. Jensen, “Implementing the viola-jones face detection algorithm,” 2008.
- [5] M. Weber, “Frontal face dataset,” 1999. [Online]. Available: www.vision.caltech.edu/Image_Datasets/faces/faces.tar