
OBJECT DETECTION USING VIOLA-JONES ALGORITHM

Dušan Stevanović

College of Applied Professional Studies, Vranje, dusanstev92@gmail.com

Abstract: In this paper it has been described and applied method for detecting face and face parts in images using the Viola-Jones algorithm. The work is based on Computer Vision Systems, artificial intelligence that deals with the recognition of two-dimensional or three-dimensional objects. When Cascade Object Detector script is trained, multimedia content is assigned for recognition. In this work the content will be in the form of an image, where the program will have the task of recognizing the objects in the images, separating the parts of the images in the head area, and on each discovered face, separately mark the area around the eyes, nose and mouth.

Algorithm for detection and recognition is based on scanning and analyzing front part of human head. Common usage of face detection and recognition can be find in biometry, photography, on autofocus option which is implemented in professional photo cameras or on smiling detectors (Keller, 2007). Marketing is also popular field where face detection and recognition can be used. For example, web cameras that are implemented in TVs, can detect every face in near area. Calculating different type of algorithms and parameters, based on sex, age, ethnicity, system can play precisely segmented television commercials and campaigns. Example of that kind of systems is OptimEyes. (Strasburger, 2013)

In other words, every algorithm that has as its main goal to detect and recognize face from image, should give as a feedback information, is there any face and if answer is positive, where is its location on image. In order to achieve acceptable performances, algorithm should minimize false recognitions. These are the cases when the algorithm ignores and does not recognize the real object from the image, and vice versa, when the wrong object is recognized as real. One of the algorithms that is frequently applied in this area of research is the Viola-Jones algorithm. This algorithm is functional in real time, meaning that besides detection, it is also possible to adjust the ability to monitor faces from video material.

In this paper, the problem that will be analyzed is facial image detection. Man can do this task in a very simple way, but to do the same with a computer, it is necessary to have a range of precise and accurate information, formulas, methods and techniques. In order to maximize the precision of recognizing the face of the image using the Viola-Jones algorithm, it is desirable that the objects in the images are completely face-to-face with the image-taking device, which will be shown through experiments.

Keywords: Object detection, Viola-Jones algorithm, Computer Vision Systems, Haar-like features

PREPOZNAVANJE OBJEKATA SA SLIKE KORIŠĆENJEM VAJOLA-DŽONS ALGORITMA**Dušan Stevanović**

Visoka škola primenjenih strukovnih studija, Vranje, dusanstev92@gmail.com

Rezime: U ovom radu opisana je i primenjena metoda detektovanja lica i delova lica sa slika korišćenjem Vajola-Džons (eng. Viola-Jones) algoritma. Rad je baziran na modelu računarskog vizuelnog sistema (eng. Computer Vision Systems), veštačke inteligencije koja se bavi prepoznavanjem dvodimenzionalnih ili trodimenzionalnih predmeta, koristeći logiku pomenutog algoritma. Jednom istreniranoj skripti Kaskadni Klasifikator koja je sastavni deo CVS modela, dodeljuje se multimedijalni sadržaj na prepoznavanje. U ovom radu sadržaj će biti u formi slike, gde će program imati zadatak da prepozna objekte na slikama, izdvaja delove slika u predelu glave, te na svakom otkrivenom licu, odvojeno obeleži predeo oko očiju, nosa i usta.

Algoritmi za detekciju i prepoznavanje lica se usredsređuju na analizi prednje strane ljudske glave. Najčešća primena prepoznavanja lica je u biometriji, fotografiji, prilikom korišćenja autofokus opcije na aparatu ili kod detektora osmeha (Keller, 2007), marketingu, gde veb kamera ugrađena u televizor detektuje svako lice u blizini. Proračunavanjem raznih parametara koji se tiču pola, starosti, etničke pripadnosti lica koje se analizira, sistem na televiziji pušta precizno segmentiranu vrstu reklama i kampanja. Primer ovakvog sistema je OptimAjs (eng. OptimEyes) (Strasburger, 2013).

Dakle, svaki algoritam koji za cilj ima da detektuje lice sa slike, treba da kao povratnu informaciju pruži da li se uopšte na slici nalazi lice i ako je odgovor potvrđan, koja je tačno njegova pozicija. Kako bi dostigao prihvatljive

performanse, algoritam bi trebalo da smanji na minimum kako lažno negativne, tako i lažno pozitivne procene. U pitanju su slučajevi kada algoritam ignoriše i ne prepozna pravi objekat sa slike, tj. obrnuto, kada pogrešan objekat prepozna kao pravi. Jedan od algoritama koji se često primenjuje u ovoj oblasti istraživanja jeste Vajola-Džons algoritam (Viola, Jones, 2001). Ovaj algoritam je funkcionalan u realnom vremenu, što znači da se pored detektovanja, može podesiti i mogućnost da prati lica sa video materijala.

U ovom radu, problem koji će se analizirati jeste detekcija lica sa slika. Čovek ovakav zadatak može uraditi na veoma jednostavan način, ali da bi to isto odradio i računar, neophodno je da raspolaže nizom preciznih i tačnih informacija, formula, metoda i tehnika. Kako bi se maksimizirala preciznost prepoznavanja lica sa slika korišćenjem Vajola-Džons algoritma, poželjno je da objekti na slikama budu u potpunosti licem okrenuti ka uređaju koji ih slika, što će se pokazati kroz eksperimente.

Ključne reči: Prepoznavanje objekata, Viola-Jones algoritam, Computer Vision Systems, Haar-like funkcije

1. UVOD

Algoritmi koji se zasnivaju na detektovanju i prepoznavanju određenog paterna, danas imaju široku primenu. Oni nastoje da pruže razumno, a po mogućstvu i tačno rešenje za skup ulaznih parametara, tako što će u datom skupu tražiti najbolje podudaranje očekivanih i datih uzoraka. Takvi algoritmi nalaze široku primenu u današnjem svetu. U medicini, predstavljaju osnovu za računarsku dijagnozu (eng. Computer-aided diagnosis systems, CAD systems), često za otkrivanje raka i tumora (Baker, Rosen, Lo, 2003). Uz pomoć računarske dijagnoze, naučnicima i doktorima se olakšava tumačenje i interpretacija rezultata sa slike. Praktična primena je prisutna i u automatskom prepoznavanju govora, pisanog rukopisa, u prepoznavanju saobraćajnih registarskih tablica, otisaka prstiju ili pak lica. Glavni zadatak je pronaći poziciju i veličinu svih objekata/predmeta na slici koji pripadaju uzorku koji se treba detektovati.

2. TEORIJSKE OSNOVE VAJOLA-DŽONS ALGORITMA

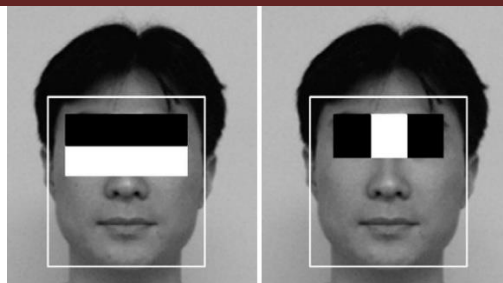
Vajola-Džons je jedan od najranijih algoritama za detekciju objekata sa različitih sadržaja, koji omogućava paralelno otkrivanje objekata u realnom vremenu. Njegovi autori su Pol Vajola i Majkl Džons (eng. Paul Viola, Michael Jones). Iako se algoritam kao takav može trenirati za otkrivanje različitih klasa objekata, najveća motivacija autora za njegovo kreiranje se bazirala na detekciji lica. Osobine koje ovaj algoritam čine dobrim za prepoznavanje lica su visoka tačnost prepoznavanja, tj. mali procenat lažno pozitivnih vrednosti. Vajola-Džons algoritam koristi Haar-like funkcije (Viola, Jones, 2001). Haar-like funkcije su zastupljene u svetu digitalnih slika i koriste se za prepoznavanje objekata. Ime su dobile po sličnosti logike koja se zasniva na Haar talasima u matematici (Viola, Jones, 2001). Često se koriste za prepoznavanje objekata u realnom vremenu. Haar funkcije se zasnivaju na sledećim pretpostavkama, a koja se tiču izgleda i oblika ljudskog lica koje je potrebno prepoznati:

- Predeo oko očiju je tamniji od obraza
- Predeo oko nosa i sam nos je svetliji od predela oko očiju,
- Raspored očiju, nosa i ustiju kod objekata je uvek isti
- Lice se posmatra kao skup lokalnih gradijenata merljivih na nivou piksela

Na osnovu ovih pretpostavki, kreirane su Haar-like funkcije. One razmatraju susedne pravougaone regije, na specifičnim lokacijama u oblasti koja se detektuje. Haar-like funkcija sumira intezitete piksela u svakoj regiji i računa razliku između tih suma. Tako dobijena razlika se zatim koristi za kategorizaciju podsekcije na slici do koje se stiglo prilikom rada algoritma za detekciju. Kako je već rečeno da je predeo oko očiju tamniji od obraza, primer jedne Haar-like funkcije za detekciju lica bi se sastojao od seta dve susedne pravougaone regije, tamnije u predelu očiju i svetlije, u predelu oko obraza i nosa.

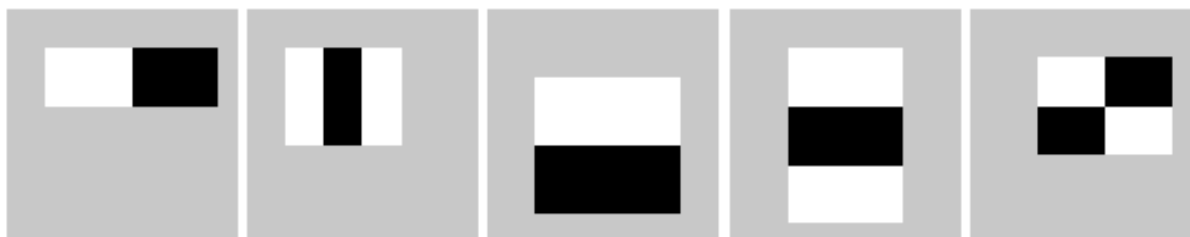
Matematički gledano, neka I predstavlja sliku, a P pravougaonu regiju i neka su obe dimenzija N x N, tada će Haar-like funkcija pridružena pravougaonoj regiji P na slici I biti definisana formulom:

$$\sum_{1 \leq i \leq N} \sum_{1 \leq j \leq N} I(i, j) 1_{P(i, j)} \text{ beta regija} - \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq N} I(i, j) 1_{P(i, j)} \text{ crna regija} \quad (1)$$



Slika 1. Primer pravougaonih regija

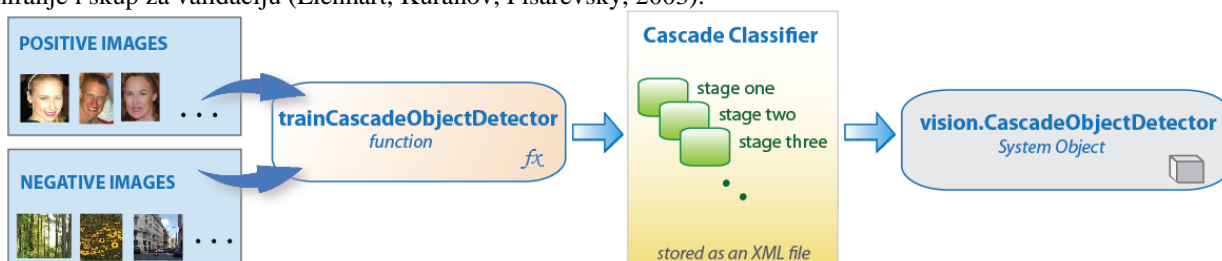
Na slici 2 prikazano je još nekoliko dodatnih pravougaonih regija, a koje se često koriste (Papageorgiou, 1998).



Slika 2. Pravougaone regije koje se koriste prilikom primene Vajola-Džons algoritma

3. IMPLEMENTACIJA ALGORITMA

U Matlabu, skripta koja koristi Vajola-Džons algoritam je Kaskadni klasifikator (eng. *Cascade Object Detector*). Kako bi se omogućila potpuna funkcionalnost i obezbedila što veća preciznost, neophodno je prvo istrenirati skriptu. Za treniranje je obezbeđen skup za treniranje (eng. *dataset*) koji se sastoji od različitih slika sa i bez ljudskih lica na njima, koje možemo imenovati pozitivnim i negativnim slikama. Slike su zatim podeljene u dva skupa: skup za treniranje i skup za validaciju (Lienhart, Kuranov, Pisarevsky, 2003).



Slika 3. Prikaz implementacije algoritma u Matlabu

Funkcionalnost Kaskadnog klasifikatora se realizuje kroz nekoliko etapa. Svaka etapa se trenira koristeći *boosting* metodu. Funkcioniše tako što se na svakoj etapi beleži region koji je definisan trenutnom lokacijom pravougaone regije. Vrednost te regije će se označiti bilo kao pozitivna bilo kao negativna. Pozitivna vrednost znači da je objekat pronađen, dok negativna označava da objekat nije pronađen. Ukoliko je vrednost negativna, klasifikacija na ovom nivou za taj uzorak se završava. Ukoliko je vrednost pozitivna, klasifikator provlači taj region u sledeću etapu. Detektor će proglasiti da je objekat pronađen tek onda kada, u poslednjoj etapi, klasifikator prepozna regiju kao pozitivnu. Prilikom rada klasifikatora, može doći do sledećih situacija – stvarno pozitivna (eng. *true positive*), lažno pozitivna (eng. *false positive*) i lažno negativna situacija (eng. *false negative*). Kako bi se dobilo najbolje rešenje, neophodno je da svaka etapa u lancu, ima što manju lažno negativnu vrednost, jer ukoliko se u uzorku, objekat pogrešno obeleži kao negativan, klasifikacija nad tim uzorkom staje (Ojala, Pietikainen, Maenpaa, 2002).

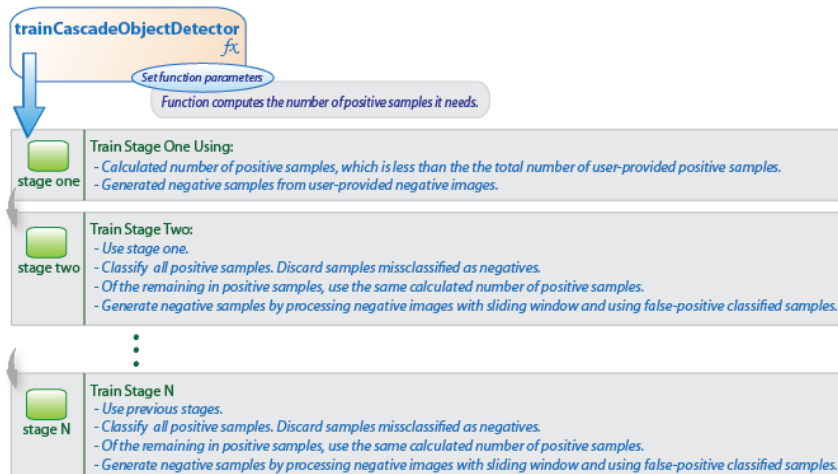
Ukupna lažno pozitivna vrednost Kaskadnog klasifikatora je funkcija:

$$V = f^s, \quad (2)$$

gde f predstavlja lažno pozitivnu vrednost u etapi, dok je s broj etapa. Analogno tome, možemo reći da je ukupna stvarno pozitivna vrednost funkcija:

$$T = t^s, \quad (3)$$

gde je t stvarno pozitivna vrednost u etapi, a s broj etapa. Dodavanje više etapa, smanjuje sveukupnu stvarno pozitivnu vrednost. Trening Kaskadnog klasifikatora je, kao što je već rečeno, obavljen nad skupom pozitivnih i negativnih slika.



Slika 4. Treniranje Kaskadnog klasifikatora kroz etape

Prilikom treniranja Kaskadnog klasifikatora, predlog je držati se sledećih normi: veliki skup za treniranje će povećati broj etapa i tako će lažno pozitivna vrednost po etapi biti viša kako bi se smanjila mogućnost da se propusti detektovanje objekta, potrebno je povećati stvarno pozitivnu vrednost

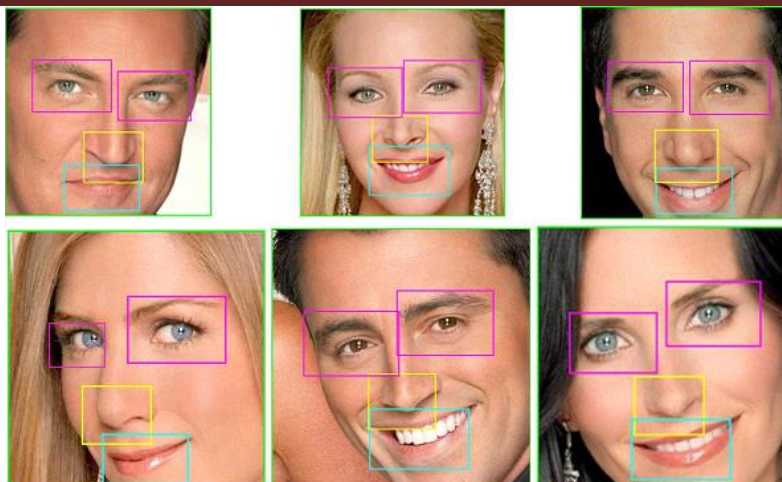
Međutim, treba biti oprezan u ovom slučaju, jer visoka stvarno pozitivna vrednost može da spreči ili oteža pogrešno detektovanje koje je takođe bitno u fazi treniranja (Dalal, Triggs, 2005).

4. PRIKAZ REZULTATA

Za realizaciju algoritma, korišćen je softverski paket Matlab R2016a. Funkcionalnost programa bazirana je na korelaciji između skripti, napisanih u Matlab-u, gde svaka ima svoju glavnu funkciju. Neke od njih, između ostalog su: kreiranjeDetektora.m i DetektovanjeDelovaLica.m. Predefinisana MATLAB-ova funkcija *imread(filename)*, učitava crno-belu ili sliku u boji iz ulaza koji je dodeljen. Povratna vrednost biće vrednost A , koja predstavlja niz, koji sadrži informacije o slici. Ukoliko je slika crno-bela, vrednost A biće niz $m \times n$, a sa druge strane, ukoliko je slika u boji, vrednost A biće dimenzija $m \times n \times 3$ (Castrillón, Déniz, Guerra, Hernández, 2007). Nakon unesene slike, neophodno je kreirati detektore koji će nad unetom slikom primeniti Kaskadni klasifikator, kako bi se nad unetom slikom, primenila metoda za detektovanje lica i njegovih delova. Na konačnom rezultatu će se pored lica, prepoznati i izdvojiti par očiju, usta i nos. U ovom slučaju, kao opcioni, ali ipak bitan parametar, naveli smo i minimalne dimenzije pravougaonih okvira koje uzimamo u obzir, prilikom kreiranja detektora. To znači da će se Vajola-Džons algoritam primeniti nad slikom unetom kao parametar, kreirajući okvire koji će prolaziti po slici, a dimenzija ne manjih od zadate minimalne vrednosti.

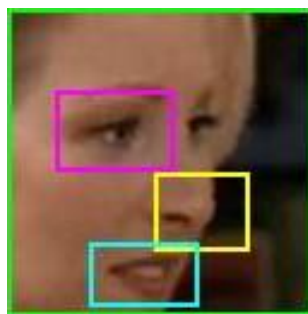
Naravno, ovaj rad je potpuno moguće primeniti za otkrivanje bilo kakvog objekta sa slike, iako se pokazalo da Vajola-Džons algoritam najbolje funkcioniše prilikom prepoznavanja lica. Potrebno je samo obezbediti odgovarajuće skupove slika prilikom treniranja Kaskadnog klasifikatora. Detektor je sposoban da otkrije široki opseg objekata, uključujući i gornji deo tela kod testnih osoba. Tip i vrsta objekta koji će se otkrivati, zavisi i kontroliše se od strane svojstva pod nazivom - Klasifikaci model (eng. *Classification Model Option*). Ovo svojstvo omogućuje da se za više puta istreniran Kaskadni klasifikator, izabere vrsta objekta koja je u samom skupu uspešno istreniranih svojstava, a koja se nalaze u ovom modelu (Matlab R2016a Documentation).

Na slici ispod je prikazan rezultat implementacije algoritma u Matlab-u.



Slika 5. Primena Vajola-Džons algoritma na slici sa objektima okrenutim licem ka fotoaparatu

Analogno tome, na slici ispod je primer rada algoritma kada je osoba uslikana iz profila i kada je kvalitet slike slabiji.



Slika 6. Na objektu koji je uslikan iz profila nije prepoznato levo oko

5. ZAKLJUČAK

Kao što se može videti sa slika u prilogu, koristeći Vajola-Džons algoritam, najbolji rezultati se dobijaju od objekata sa slika kod kojih je vidljivo celo lice. Ovaj algoritam za detektovanje i prepoznavanje lica karakteriše činjenica da se za razliku od drugih algoritama gde se vrši skaliranje slike, ovde sva logika preusmerava na funkcije i pravougaone okvire koji će prolaziti kroz sliku, te vršiti detekciju objekata. Algoritam predstavlja dobru metodu za dalji nastavak implementacije, te detektovanje i prepoznavanje objekata u pokretu na video sadržajima. S druge strane, nedostatak Vajola-Džons algoritma može se naći u kvalitetu materijala koji mu se dodeljuju za prepoznavanje. Slike koje su slabijeg kvaliteta, loše osvetljenosti ili gde objekti nisu okrenuti u potpunosti ka fotoaparatu, povećavaju mogućnost da algoritam ne prepozna objekat uopšte ili ga prepozna delimično.

ЛИТЕРАТУРА

- Baker, J. A.; Rosen, E. L.; Lo, J. Y.; et al. (2003). "Computer-Aided Detection (CAD) in Screening Mammography: Sensitivity of Commercial CAD Systems for Detecting Architectural Distortion". *American Journal of Roentgenology*. 181: 1083–1088. doi:10.2214/ajr.181.4.1811083
- C. Papageorgiou, M. Oren and T. Poggio. A General Framework for Object Detection. *International Conference on Computer Vision*, 1998
- Castrillón Marco, Déniz Oscar, Guerra Cayetano, and Hernández Mario, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams". In *Journal of Visual Communication and Image Representation*, 2007

- Dalal, N., and B. Triggs, "Histograms of Oriented Gradients for Human Detection". IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 1, (2005)
- Keller Jeff, "DCRP Review: Canon PowerShot S5 IS". Dcresource.com. Retrieved 2007-07-11
- Lienhart R., Kuranov A., and V. Pisarevsky "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection.", Proceedings of the 25th DAGM Symposium on Pattern Recognition. Magdeburg, Germany, 2003
- Matlab Documentation, https://www.mathworks.com/help/vision/ref/vision.cascadeobjectdetector-class.html#bs_hhwg-4
- Ojala, T., M. Pietikainen, and T. Maenpaa. "Multiresolution Gray-scale and Rotation Invariant Texture Classification With Local Binary Patterns." IEEE Transactions on Pattern Analysis and Machine Intelligence. Volume 24, No. 7 July 2002
- Paul Strasburger, Tesco face detection sparks needless surveillance panic, Facebook fails with teens, doubts over Google+ Technology, theguardian.com
- Viola, Jones: Robust Real-time Object Detection, IJCV 2001 See pages 1,3
- Viola, Jones, "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition, 2001