

Izdvajanje karakterističnih obeležja lica

Cilj ovog rada je realizacija sistema za izdvajanje karakterističnih obeležja lica na portretu. Za pronalaženje lica na fotografiji korišćene su Haar odlike, za izdvajanje ključnih tačaka korišćeni su histogrami osvetljenja, a za izdvajanje karakterističnih obeležja je korišćen GentleBoost klasifikacioni algoritam. Vektor odlika, na osnovu kojeg se vrši klasifikacija, se sastoji od 49 slojeva, gde se prvih 48 dobija filtriranjem slike različitim Gaborovim filtrima, a poslednji sloj predstavlja luminansu slike. U radu je korišćena baza fotografija Extended Cohn-Kanade, na kojoj je postignuta tačnost od 72% za izdvajanje ključnih tačaka, a tačnost klasifikacije karakterističnih obeležja je bila 88% nakon 10 000 trening iteracija.

Uvod

Izdvajanje karakterističnih obeležja

Izdvajanje karakterističnih obeležja lica ima primenu u raznim sistemima, kao što su: prepoznavanje lica, prepoznavanje emocija lica, praćenje lica i čitanje sa usana. Karakterističnim obeležjima lica se najčešće imenuju izražene tačke kao što su: ivice očiju, obrva i usana, nozdrve, vrh nosa i brade.

Popularnost u rešavanju ovog problema se može primetiti u broju metoda, među kojima su popularni oni koji se zasnivaju na prepoznavanju tekstuira (Reinders *et al.* 1996; Holden i Owens 2002; Feris *et al.* 2002), prepoznavanju teksturom ograničenih modela oblika (Liu *et al.* 2003), prepoznavanju oblika (Hu *et al.* 2003; Yan *et al.* 2003) i klasifikacijom pomoću Haar odlika kombinovanih sa statističkim modelima oblika (Cristinacce i Cootes 2003).

Gaborov filter

Gaborov filter je linearni filter koji se koristi za pronalaženje ivica, kreiran od strane Denisa Gabora. Frekvencijska i orijentaciona reprezentacija kernela Gaborovog filtra je u osnovi slična informacijama po-

*Filip Parag (1999),
Subotica, Eugena
Kumičića 20, učenik 2.
razreda Gimnazije
„Jovan Jovanović
Zmaj” u Novom Sadu*

*MENTOR: Marija
Janković,
Elektrotehnički
fakultet Univerziteta u
Beogradu*

moću kojih vizuelni sistem ljudi i životinja stvara sliku prostora; takođe, ove osobine dolaze do izražaja u prikazu i razlikovanju tekstura.

Gaborov filter se koristi za detekciju ivica. Težinske vrednosti pojedinačnih piksela određene su vrednostima dvodimenzionalne Gausove funkcije pomnožene sinusnim funkcijama. Na ovaj način se dobija „talasast” kernel koji omogućava da se vrednosti piksela periodično ističu (slika 1).



Slika 1.
Primer kernela
Gaborovog filtra

Figure 1.
Gabor kernel example

Boosting algoritmi

Boosting predstavlja vrstu algoritama za mašinsko učenje koji se koriste za smanjenje prezasićenja i varijanse u nadgledanom treniranju slabih klasifikatora. Boosting se zasniva na ideji da se mnogo slabih klasifikatora (eng. weak learners) može izjednačiti sa jednim jakim klasifikatorom (eng. strong learner); statistički gledano, veći broj slabih klasifikatora će zajedno u proseku biti tačniji od jednog jakog klasifikatora. Slabim klasifikatorima se smatraju oni čija je tačnost neznatno bolja od nasumičnosti, dok se jakim klasifikatorima smatraju klasifikatori čija je tačnost, u određenim okolnostima, približna idealnoj. Za slabe klasifikatore se uglavnom biraju stabla odlučivanja (eng. decision trees).

Neka je I broj elemenata u trening skupu, skup $X = \{x_1, \dots, x_I\}$ vektor odlika trening skupa, $Y = \{y_1, \dots, y_I\}$ skup željenih rezultata gde je $y_i \in \{-1, 1\}$, a $\Omega = \{\omega_{1,1}, \dots, \omega_{I,1}\}$ skup početnih težina elemenata vektora odlika postavljenih na vrednost $1/I$. Iteracija algoritma, odnosno broj slabog klasifikatora se označava sa $t \in (1, T)$, gde je T zadat broj iteracija.

Funkcija $h_t : x \mapsto \{-1, 1\}$ je funkcija i -tog elementa vektora odlika u t -toj iteraciji, koja ukoliko je klasifikacija tačna ima vrednost 1, a u suprotnom -1 . Funkcija $f(x_i)$ predstavlja težinsku predstavu slabog klasifikatora, gde je $f(x_i) = \alpha_t h_t(x_i)$, a α_t težinski koeficijent tog klasifikatora.

Funkcija $E(f(x), y, I)$ je funkcija greške koja kod AdaBoost-a ima vrednost:

$$E_A(f(x), y, I) = \sum_{i=1}^I e^{-y_i f(x_i)}$$

a kod GentleBoost-a ima vrednost:

$$E_G(f(x), y, I) = \sum_{i=1}^I (y_i - f(x_i))^2$$

Iz funkcija greške se uočava da AdaBoost algoritam ističe najveće greške pomoću eksponencijalne funkcije, dok GentleBoost koristi kvadratnu funkciju, što dovodi do regulisanja srednje kvadratne greške. Može se zaključiti da će GentleBoost mnogo manji prioritet dati netačno klasifikovanim elementima u odnosu na AdaBoost.

Težina $\omega_{t+1,i}$, gde je $t + 1$ broj slabih klasifikatora, a i broj elementa vektora odlika se izražava formulom:

$$\omega_{t+1,i} = \omega_{t,i} \exp(-y_i f(x_i))$$

kod AdaBoosta, a kod GentleBoosta formulom:

$$\omega_{t+1,i} = \omega_{t,i} (y_i - f(x_i))^2$$

Cilj svake iteracije algoritma jeste da, promenom parametra α_t , u slučaju AdaBoosta minimizuje izraz:

$$\min \sum_{i=1}^I \omega_{t,i} \exp(-y_i f(x_i))$$

a u slučaju GentleBoosta:

$$\min \sum_{i=1}^I \omega_{t,i} (y_i - f(x_i))^2$$

Cilj istraživanja

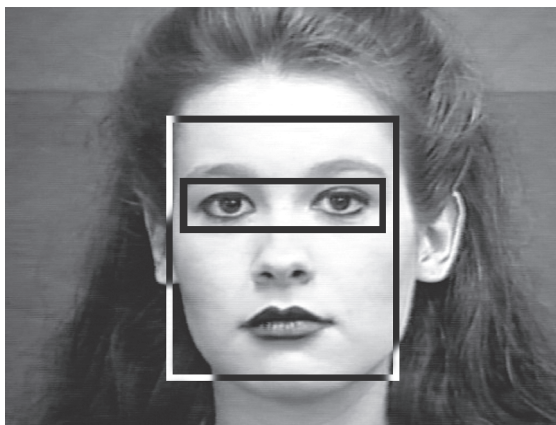
Cilj istraživanja bio je da se ispita sledeće:

- da li povećanje broja slabih klasifikatora pozitivno utiče na tačnost klasifikacije;
- da li povećanje broja primera u skupu za treniranje u odnosu na skup za testiranje pozitivno utiče na tačnost klasifikacije;
- da li je GentleBoost klasifikacija brža od AdaBoost klasifikacije.

Metod

Detekcija lica. Za detekciju lica koristi se Haar Cascade algoritam iz biblioteke OpenCV. Na fotografiji se traži oblast para očiju pomoću Haar odlika. Na osnovu širine pomenute oblasti i empirijski određenih koeficijenata se obeležava granični okvir oko unutrašnjeg lica (slika 2). Ukoliko se pronađe više parova očiju ili pak nijedan, smatra se da je pronalaženje neuspešno.

Izdvajanje ključnih tačaka. Gornja polovina detektovanog lica, pronađenog u prethodnom koraku, se vertikalno deli u dva jednaka dela, tako da se u svakom nalazi jedno oko. Analizom horizontalnog i vertikalnog histograma razlike osvetljenja dobijaju se pozicije zenica – vrhovi histograma. Kod histograma razlike osvetljenja se, u odnosu na običan histogram osvetljenja vrednost za svaku kolonu matrice računa sabiranjem

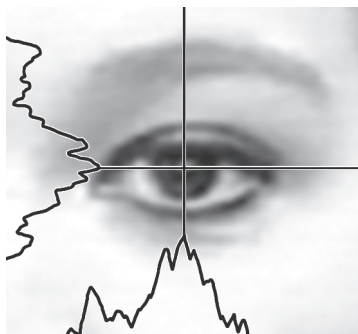


Slika 2.
Izdvajanje unutrašnje
regije lica

Figure 2.
Inner face region
extraction

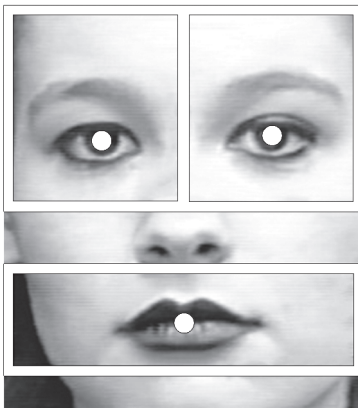
apsolutne razlike intenziteta trenutnog elementa i elementa u prethodnom redu, ukoliko ta apsolutna vrednost prelazi određeni prag (slika 3). Interpupilarna razdaljina (IPR), koja se dobija prethodnom analizom histograma, jeste udaljenost dveju zenica i služi kao osnovna udaljenost u proporcijama lica.

Region u kome se nalaze usne se može ograničiti po vertikalnoj osi sa početkom u 1 IPR i krajem u 1.5 IPR (slika 4). Pronalaženje srednje tačke usana po vertikali je analogno pronalaženju zenica, dok je po horizontali srednja tačka usana jednako udaljena od zenica.



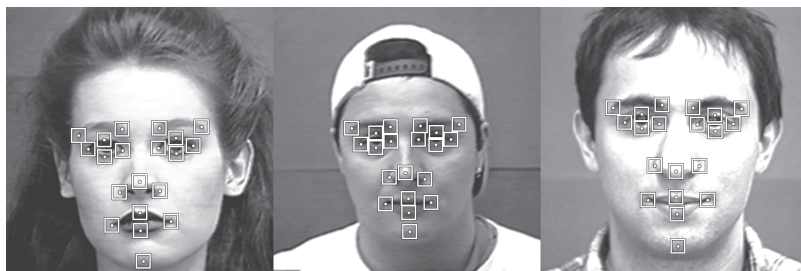
Slika 3.
Pronalaženje središta
oka pomoću
histograma

Figure 3.
Searching for the
center of the eye using
a histogram



Slika 4.
Regioni za
pronalaženje ključnih
tačaka

Figure 4.
Regions for key points
detection



Slika 5.
Primeri sa izdvojenim
regionima interesa

Figure 5.
Examples showing
derived regions of
interest

Izdvajanje regiona interesa. Regioni interesa su pravougaoni delovi unutrašnjeg lica na kojima se sigurno nalaze karakteristična obeležja (slika 5).

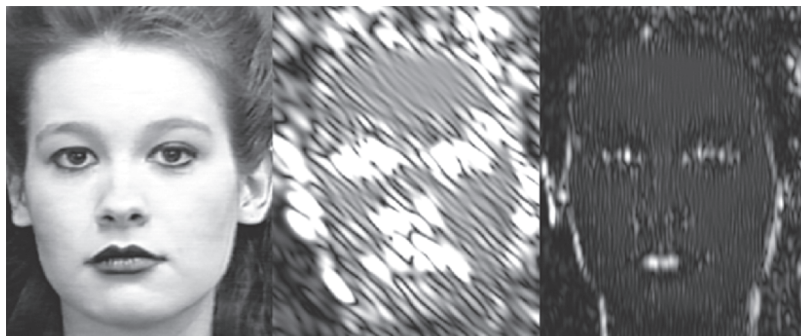
Pozicija svakog regiona se može predstaviti kao zbir pozicije jedne od ključnih tačaka i koeficijentom množene IPR. U predloženom sistemu se izdvaja dvadeset regiona, gde svaki odgovara jednom od karakterističnih obeležja: krajnje leve i desne tačke obrva, krajnje leve, desne, gornje i donje tačke očiju i usana, leva i desna nozdrva, vrh nosa i srednja tačka brade.

Izdvajanje karakterističnih obeležja. Karakterističnim obeležjem se smatra jedan piksel, ali pošto isti piksel ne sadrži dovoljno informacija, koristi se isečak veličine 13×13 piksela sa sredinom u pikselu karakterističnog obeležja.

Za potrebe treniranja, na svakom regionu interesa se izdvaja 25 isečaka, od kojih se 9 smatraju tačnim, a ostalih 16 netačnim; od tačnih je jedan isečak sa centrom u ručno obeleženom karakterističnom obeležju, 8 isečaka sa središtem udaljenim 1 piksel od karakterističnog obeležja; od netačnih su 8 isečaka sa središtem udaljenim između 3 i 5 piksela od karakterističnog obeležja, a ostalih 8 su na udaljenosti većoj od 5 piksela. Pozicije isečaka smatranih netačnim su nasumično određene, tako da im udaljenost zadovoljava kriterijum.

Pri testiranju se svaki isečak u regionu interesa klasifikuje pomoću istreniranog Boosting klasifikatora, čime se pronalazi pozicija karakterističnog obeležja.

Svaki isečak se pretvara u vektor odlika za algoritam klasifikacije i sastoji se od 49 slojeva. Prvih 48 slojeva predstavljaju isečak filtriran različitim Gaborovim kernelima, gde se za parametre karakteristike kernela



Slika 6.
Primer primene
Gaborovog filtra

Figure 6.
Application of a
Gabor filter

bira 8 orijentacija između 0° i 180° i 6 frekvencija između 2 i 12 piksela sa ujednačenim uvećanjem. Poslednji sloj vektora odlika je forma luminanse – sivog prostora boje (eng. grayscale) respektivnog isečka. Na slici 6 se mogu videti primeri fotografije filtrirane Gaborovim filtrom sa različitim parametrima.

Klasifikacija karakterističnih obeležja. Treniranje Boosting algoritma se vrši na vektorima odlika isečaka iz trening skupa. Algoritmi koji se koriste su AdaBoost i GentleBoost, gde se oba algoritma treniraju u 1 000 iteracija sa slabim klasifikatorima na bazi 500 stabala odlučivanja. Algoritam GentleBoost se naknadno trenira i sa 10 000 slabih klasifikatora kako bi se našla tačka nakon koje tačnost konvergira.

Nakon izvršenog treniranja, Boosting algoritam se testira na vektorima odlika isečaka iz test skupa. Za tačnost se uzima procenat tačno klasifikovanih obeležja.

Rezultati i diskusija

Trening skup

Baza fotografija korišćena u ovom radu je Extended Cohn-Kanade (CK+), u kojoj se nalazi 593 video snimka 113 osoba kako iz neutralne ekspresije razvijaju emocije (Lucey *et al.* 2013). Sve osobe na fotografijama su između 18 i 50 godina starosti, fotografisane sprema u dobro osvetljenoj prostoriji. Fotografije korišćene za trening i test skup su početni frejmovi (neutralna ekspresija) iz svih snimaka, gde je trening skup 70% nasumično odabranih fotografija, a test skup ostalih 30%. Na slici 7 se može videti 6 uzoraka iz baze.



Slika 7.
Primeri iz baze
fotografija

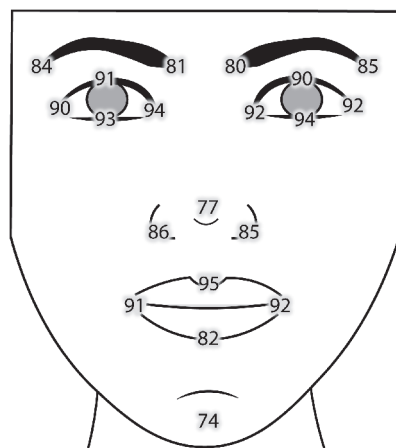
Figure 7.
Examples from the
photo database

Eksperimentalni rezultati

Detekcija lica. Algoritam za detekciju parova očiju na bazi Haar odlika uspešno izdvaja parove očiju na 85% fotografija, odnosno u 504 od ukupno 593 primerka. Fotografije na kojima je detekcija lica neuspešna se izuzimaju iz skupa u narednom koraku.

Izdvajanje ključnih tačaka. Pri detekciji ključnih tačaka, uspešnost je 72%, odnosno 363 tačno označena od 504 primerka. Uslov da se izdvojeni primerci smatraju tačnim jeste da je greška lokacije ispod 10 piksela udaljenosti od ručno obeležene tačke kod sve tri tačke. Greška se može najviše pripisati nejasnom prelazu između obrve i zenice, koji doprinosi da se na ranije navedenom histogramu najveća vrednost ne podudara sa lokacijom zenice. Nejasan prelaz je najverovatnije uzrokovan senkom, to jest lošim osvetljenjem, ili izrazito malom udaljenošću oka od obrve. Fotografije na kojima je izdvajanje ključnih tačaka neuspešno se izuzimaju iz skupa u narednim koracima.

Klasifikacija karakterističnih obeležja. Treniranje GentleBoost klasifikatorom je izvršeno u 10 000 iteracija, to jest sa 10 000 slabih klasifikatora. Dobijena prosečna tačnost je 88%, a pojedinačna tačnost za svaki region se može videti na dijagramu lica (slika 8).

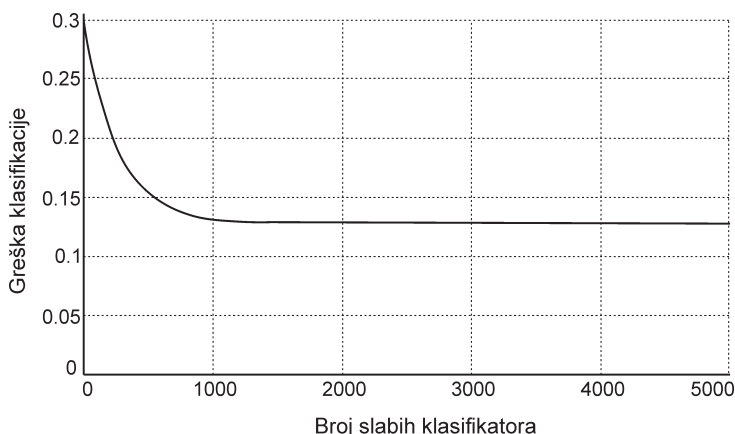


Slika 8.
Grafički prikaz
tačnosti izdvajanja
obeležja u procentima

Figure 8.
Graphical
representation of
landmark detection in
percentages

Na grafiku (slika 9) je prikazana zavisnost prosečne greške od broja iteracija, to jest broja slabih klasifikatora. Može se primetiti da greška u proseku konvergira nakon 1 500 iteracija, što znači da je dalje treniranje suvišno.

Tačnost pronalaženja obeležja sredina brade i sredina vrha nosa je u poređenju sa drugim obeležjima manja, jer je prostor koji okružuje ta obeležja jednoličnog izgleda; sa druge strane, krajnje tačke očiju i usta se nalaze na prelazima koji se karakterišu značajnim kontrastom i izraženim ivicama. Ove odlike povoljno utiču na tačnost pronalaženja tačaka, jer čine lako uočljive obrasce.



Slika 9.
Zavisnost prosečne greške klasifikacije GentleBoosta od broja slabih klasifikatora

Figure 9.
GentleBoost classification error as a function of the number of weak learners

Dok je tačnost pronalaženja krajnjih tačaka očiju međusobno slična, kod usana se primećuje da donja tačka ima znatno manju tačnost u odnosu na druge. Razlog tome je senka stvorena od strane donje usne, koja se može primetiti na slici 2.

Zbog raznovrsnosti u obliku i poziciji obrva, njihove krajnje tačke nemaju ujednačene obrasce i time uzrokuju tačnost manju nego kod usana i očiju. Ipak, očekuje se veća tačnost u odnosu na sredinu nosa i sredinu brade zbog kontrasta i ivica obrva.

Merenje vremena treniranja i testiranja oba algoritma je izvršeno po 10 puta sa 1000 slabih klasifikatora i dobijeno je da je GentleBoost u proseku 2.6 puta brži, odnosno da je prosečno vreme izvršavanja GentleBoosta 2700 sekundi, a AdaBoosta 7200 sekundi. Pretpostavlja se da razlika u računanju težina klasifikatora i njihovih grešaka značajno utiče na vreme izvršavanja. Kod GentleBoosta je greška kvadratnog oblika, dok je kod AdaBoosta eksponencijalna, gde je potrebno izvršiti mnogo više aritmetičkih operacija.

Tačnost klasifikacije sa 1000 slabih klasifikatora sa raspodelom primeraka iz baze gde je 70% bilo u trening skupu, a 30% u test skupu, je bila 85%, a kod klasifikacije gde je 80% primeraka bilo u trening skupu, a 20% u test skupu, je bila 86%. Može se primetiti minimalna razlika, koja je i očekivana, jer povećanjem broja primeraka u trening skupu se Boosting klasifikator pri treniranju susreće sa većom raznovrsnošću odlika karakterističnih obeležja.

Zaključak

U radu je postignuta tačnost od 72% za izdvajanje ključnih tačaka, a prosečna tačnost klasifikacije karakterističnih obeležja je bila 88%, gde je iznad proseka bila klasifikacija obeležja na očima i usnama, dok je na vrhu nosa i vrhu brade bila znatno ispod proseka. U referentnom radu je postignuta tačnost celog sistema od 92%. Istraživanje je pokazalo da se povećanjem broja slabih klasifikatora dobija veća tačnost klasifikacije do 1500 iteracija, nakon čega je ustanovljeno da tačnost konvergira. Empirijski je dokazano da je GentleBoost algoritam znatno brži od AdaBoosta.

U daljem istraživanju bi trebalo dodati i Bootstrap aggregating algoritam i porediti njegovu tačnost i brzinu izvršavanja sa dobijenim rezultatima u ovom radu. Takođe, trebalo bi istražiti da li postoji vektor odlika koji bi bolje predstavljao karakteristična obeležja lica.

Literatura

- Cristinacce D., Cootes T. 2003. Facial feature detection using AdaBoost with shape constraints. U *British Machine Vision Conference (BMVC 2003)*, str. 10.
- Feris R., Gemmel J., Toyama K. 2002. Hierarchical wavelet networks for facial feature localization. *Proc. IEEE Int'l Conf. Face and Gesture Recognition*, str. 118-123.
- Freund Y., Schapire R. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, **55**: 119.
- Holden E., Owens R. 2002. Automatic facial point detection. *Asian Conference on Computer Vision*, 2: 731.
- Hu C., Feris R., Turk M. 2003. Real-time view-based face alignment using active wavelet networks. U *Proc. IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures*, str. 215-221.
- Liu C., Yan S., Li S. Z., Zhang H., Shum H., Cheng Q. 2003. Face alignment using texture-constrained active shape models. *Image and Vision Computing*, **21** (1): 69.
- Lucey P., Cohn J., Kanade T., Saragih J., Ambadar Z. 2013. *The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression*. Pittsburgh: Carnegie Mellon University
- Reinders M., Koch R. W. C., Gerbrands J. J. 1996. Locating facial features in image sequences using neural networks. *Proc. IEEE Int'l Conf. Face and Gesture Recognition*, str. 230-235.
- Viola P., Jones M. 2001. Robust real-time object detection. U *Second international workshop on statistical and computational theories of vision – modeling, learning, computing, and sampling*. Vancouver, str. 3-30.
- Vukadinović D., Pantić M. 2005. Fully automatic facial feature point detection using Gabor feature based boosted classifiers. U *Man and Cybernetics (SMC'05)*. Waikoloa, str. 1692-1698.
- Yan S., Hou X., Li S. Z., Zhang H., Cheng Q. 2003. Face alignment using view-based direct appearance models. *International Journal of Imaging Systems and Technology*, **13** (1): 106.

Filip Parag

Face Landmark Detection

This paper shows an implementation of a system for face landmark detection on a portrait. Haar-like features detect certain patterns using illumination difference in particular adjacent rectangles and they are used for face detection. Key points on the detected face are located using illumination delta histograms, the global peaks of which mark pupils and lips (Figure 3). Facial landmarks mark 20 specific facial points which can be seen on Figure 8. They are classified using the GentleBoost classification algorithm, which is an ensemble method that uses vast number of weaker classification algorithms, such as decision trees, to improve their common performance. Feature vectors for the GentleBoost algorithm consist of 49 layers. The first 48 layers represent the image filtered using different Gabor filters, and the last layer is luminance. The system is tested on the Extended Cohn-Kanade database, and it achieves 72% accuracy for key point detection and 88% accuracy for face landmark classification after 10 000 iterations of training.

