

# Prepoznavna obrazov z AG opisnikom

Domen Lušina<sup>1</sup>, Jasna Maver<sup>2</sup>

<sup>1</sup>Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Tržaška 25, 1000 Ljubljana, Slovenija

<sup>2</sup>Univerza v Ljubljani, Filozofska fakulteta,

Oddelek za bibliotekarstvo, informacijsko znanost in knjigarstvo, Aškerčeva 2, 1000 Ljubljana, Slovenija

E-pošta: jasna.maver@ff.uni-lj.si

**Povzetek.** Za prepoznavanje obrazov obstajajo številne metode, nekatere med njimi uporabljajo lokalne opisnike. V članku raziščemo možnost uporabe novega lokalnega AG opisnika, ki je zasnovan na modelu odvisnih učinkov, za prepoznavanje obrazov. V ta namen smo razvili sistem za prepoznavanje obrazov, ki iz izračunanih AG opisnikov zgradi vrečo besed in uporabi SVM klasifikator. Predlagani sistem smo testirali na javno dostopnih podatkovnih bazah. Testiranje smo opravili tudi z drugimi priljubljenimi opisniki. Sistem za prepoznavo obrazov je pri uporabi AG opisnika dal primerljivo klasifikacijsko točnost kot pri uporabi drugih testiranih opisnikov, v nekaterih testih pa je bila ta večja. AG opisnik se je izkazal kot dobra alternativa drugim opisnikom.

**Ključne besede:** prepoznavanje obrazov, AG opisnik, lokalni opisnik

## Face recognition with an AG descriptor

There are many methods available for face recognition. Some of them use local descriptors. The paper explores the possibility of using a new local AG descriptor, derived from a dependent-effects model, for the face-recognition function. A face recognition system is developed that builds a bag of words from calculated AG descriptors and uses a SVM classifier. The system is tested on publicly-available databases. Some other popular descriptors are also tested. The recognition scores of the AG descriptor are similar to those of other tested descriptors. In some tests, the recognition scores are higher. The AG descriptor proves to be a good alternative to other descriptors.

**Keywords:** face recognition, AG descriptor, local descriptor

## 1 UVOD

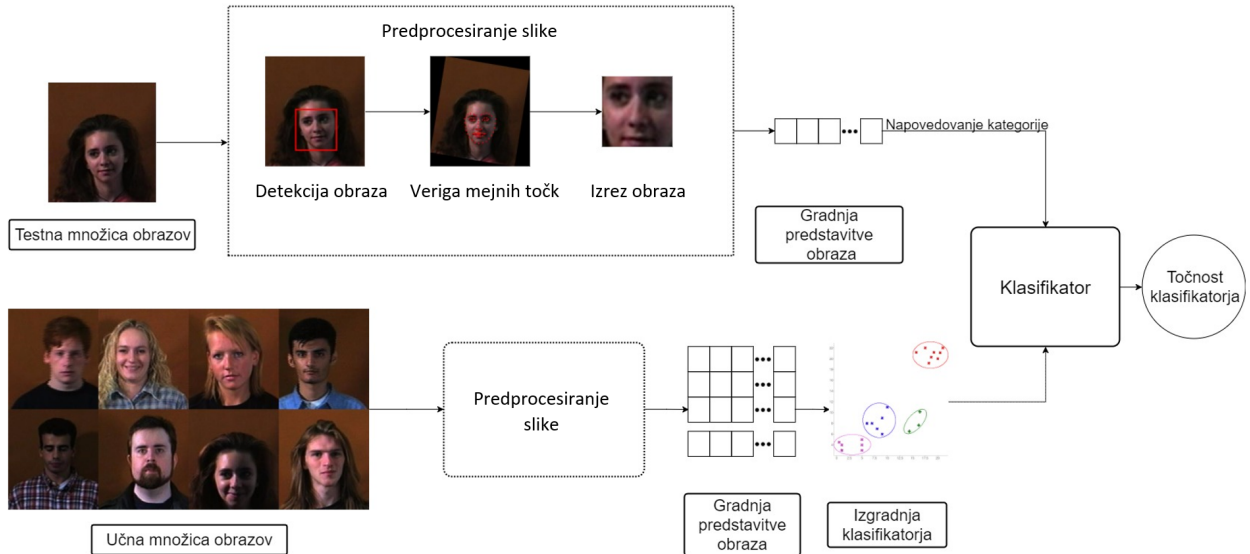
Prepoznavna obrazov je ena osnovnih človeških zmožnosti, ki jo usvojimo v zgodnjem obdobju življenja. Težko si je predstavljati vsakdanje življenje brez zmožnosti, da bi na ulici prepoznali znano osebo, prijatelja ali sorodnika. Čeprav se nam prepoznavanje obrazov zdi samoumevna funkcija, je to zapletena naloga, ki smo jo sposobni izvesti zelo hitro in učinkovito. Kompleksnosti naloge se zavemo, ko želimo zmožnost prepoznavanja obrazov prenesti v računalniški sistem.

Prvi začetki razvoja sistemov za prepoznavanje obrazov segajo v šestdeseta leta prejšnjega stoletja, ko se je začel razvoj prvih polavtomatskih sistemov, ki so zahtevali ročno vnašanje značilnic na obrazu (oči, ušesa, nos in usta) [1]. Skozi leta je prišlo do veliko izboljšav [2], [3], [4].

Prepoznavanje obrazov ima velik potencial uporabe. Med drugim bi ga s pridom lahko uporabljali za kontrolo dostopa, nadzor in v zdravstvene namene. V zadnjih letih se je prepoznavna obrazov že začela uporabljati kot biometrična identifikacija npr. za dostop do socialnih omrežij, kot sta Facebook in Google+, ki sta dostopna na pametnih telefonih, kjer vgrajena kamera omogoča sliko obraza.

Prepoznavna obrazov je zelo obsežno raziskovalno področje. Grobo ga lahko razdelimo na dva dela, pri katerih se uporabljata globalni oz. lokalni pristop. Starejši je globalni pristop. Sem uvrstimo metodo glavnih komponent (PCA) [5], linearno diskriminantno analizo (LDA) [6] in analizo neodvisnih komponent (ICA) [7]. Uporaba teh metod ima svoje slabosti, med drugim je znano, da je PCA metoda občutljiva na spremembe v osvetlitvi. Na splošno so te metode, ker opisujejo obraz kot celoto, zelo občutljive na spremembe v videzu obraza, kot je npr. sprememba v izrazih. Pri lokalnem pristopu obraz običajno predstavimo z lokalnimi opisniki, izračunanimi na določenih obraznih točkah, imenovanih značilne lokacije. Lokalni pristopi so zato bolj robustni na spremembe v izrazih na obrazu. Z uporabo ustreznega opisnika lahko minimiziramo vpliv osvetlitve. Med prvimi je bil opisnik SIFT [8], skozi leta pa so bili razviti številni novi opisniki [9], [12], [13], [14], [15]. Prepoznavna obrazov pogosto vključuje uporabo metod s področja umetne inteligence. Za primerno se je izkazal klasifikator SVM, prav tako naključni gozdovi in umetne nevronske mreže. Uporaba metod umetne inteligence je postala skoraj neizogibna, saj je prispevek zelo velik.

Cilj predlaganega dela je preizkusiti lokalni AG opisnik [11] za prepoznavo obrazov. Pripravili smo sistem za prepoznavo obrazov, sestavljen iz modula za detekcijo obraza in modula za prepoznavo. Predstavitev obrazov,



Slika 1: Sistem za prepoznavo obrazov.

ki jo uporablja modul za prepoznavo, bo zasnovana na AG opisnikih. V poglavju 2 predstavimo zgradbo sistema za prepoznavo obrazov oz. osnovne module, ki jih vključuje sistem. V podpoglavju 2.1 opišemo postopek detekcije in izrezave obraza iz slike, v poglavju 2.2 pa predstavimo metodo vreče besed, ki smo jo uporabili za predstavitev obrazov, in dva različna pristopa za pridobivanje značilnih lokacij. Sledi razlaga AG opisnika, s katerim predstavimo značilne lokacije na sliki. V poglavju 3 najprej predstavimo testne podatkovne baze slik obrazov. Sledijo opis eksperimentov ter prikaz in interpretacija rezultatov.

## 2 SISTEM ZA PREPOZNAVNO OBRAZOV

Sistem za prepoznavanje obrazov je sestavljen iz več modulov: predprocesiranja slike, gradnje predstavitev obraza in klasifikatorja, ki je zgrajen s pomočjo učenja. Slika 1 prikazuje osnovne module predlaganega sistema.

### 2.1 Predprocesiranje

Predprocesiranje slike vključuje detekcijo obraza, izračun verige mejnih točk na obraznih delih, poravnavo in izrezovanje obraza ter normalizacijo velikosti obraznega dela.

**2.1.1 Detekcija obraza:** Primer detektiranega obraza prikazuje slika 2. Izmed številnih algoritmov, ki so na voljo, smo izbrali algoritem za detekcijo obrazov, predlagan v delih [17] in [18]. Koda algoritma je na voljo v knjižnici dLib\*. Algoritem uporablja HOG piramido slik, skozi katero pošlje linearni klasifikator. Vsaka raven v piramidi pomeni HOG značilnice, izračunane pri različni resoluciji slike. Za detekcijo obraza uporabimo zaznavalno okno, ki je sestavljeno iz mreže prekrivajočih

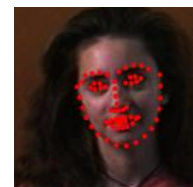
se HOG značilnic. Te združimo v skupni vektor in uporabimo SVM klasifikator, ki ločuje med območjem v sliki, ki predstavlja obraz in območjem, ki to ni.

Detekcijo obrazov smo izboljšali tako, da smo dano sliko  $I$  rotirali za kote  $\alpha \in [-30^\circ, 30^\circ]$  s korakom  $1^\circ$ . Za manjše obraze, katerih drseče okno ni zaznalo, smo sliko povečali in detekcijo ponovili.



Slika 2: Detekcija obraza. Obraz je označen z rdečim okvirjem.

**2.1.2 Izločanje obraza in njegova poravnava:** Obraz izločimo iz slike z določitvijo verige mejnih točk na obraznih delih (slika 3), kot so npr. usta, nos, oči in obrvi. Za to nalogo uporabimo algoritem ansambla regresijskih dreves [19], ki uporablja 68 mejnih točk. Algoritem je implementiran v knjižnici dLib. Prednost tega algoritma je, da z zadovoljivo natančnostjo zazna verigo mejnih točk obraznih delov v milisekundi. Algoritem uspešno zazna položaj mejnih točk tudi v primerih, ko je obraz delno zakrit. Obraz na sliki poravnamo glede



Slika 3: Veriga vključuje 68 mejnih točk.

\*Zaznava slik frontalnih obrazov knjižnice dLib: [http://dlib.net/imaging.html#get\\_frontal\\_face\\_detector/](http://dlib.net/imaging.html#get_frontal_face_detector/).

na položaj oči. Središči obeh očes smo izračunali s pomočjo mejnih točk, pridobljenih v prejšnjem koraku. Naj  $\mathbf{p}_{i,s}$  označuje koordinati  $(x_{i,s}, y_{i,s})$   $i$ -te mejne točke, v verigi mejnih točk, ki pripadajo očesu. Naj velja  $i \in \{1, \dots, N\}$ , kjer je  $N$  število mejnih točk, razporejenih okoli očesa, in naj  $s$  označuje, ali gre za levo ali desno oko ( $s \in \{l, r\}$ ). Središče očesa  $C_s$  zapišemo kot:

$$C_s = (x_{c,s}, y_{c,s}) = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_{i,s}. \quad (1)$$

Iz izračunanih središč lahko določimo kot nagiba obraza na sliki

$$\alpha = \tan^{-1} \left( \frac{y_{c,l} - y_{c,r}}{x_{c,l} - x_{c,r}} \right). \quad (2)$$

Sliko nato zarotiramo z rotacijsko matriko  $R$ :

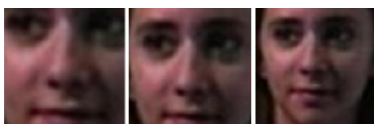
$$R = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

Hkrati ob rotaciji slike pa moramo koordinate mejnih točk v verigi transformirati v koordinate rotirane slike. Poznamo lokacijo zgornjega levega vogala  $(x_0, y_0)$  slike v novem koordinatnem sistemu. Nove koordinate mejnih točk v verigi  $(x', y')$  izračunamo z naslednjo enačbo:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \left( R \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \right). \quad (4)$$

**2.1.3 Izrezovanje obraza:** Z izrezovanjem obraza želimo izločiti vpliv ozadja na prepoznavo obraza. S slike poravnane obraza odstranimo del, ki ne vključuje obraza. Ustvarili smo tri različne tipe izrezovanja, ki jih prikazuje slika 4:

- tip 1 - okvir izrezave zajame oči, usta in nos,
- tip 2 - okvir izrezave zajame oči, usta, nos, lica in obrvi,
- tip 3 - okvir izrezave zajame cel obraz.



Slika 4: Različni tipi izrezave: tip 1, tip 2 in tip 3

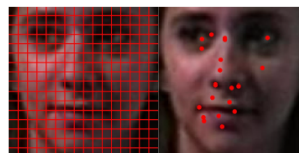
Velikost izrezanega območja smo izračunali iz koordinat mejnih točk v verigi. Za vsak tip izrezave izračunamo minimalni in maksimalni koordinati  $x$  in  $y$ . Izračunane koordinate določajo štiri vogalne točke okvirja izrezave. Območje v sliki zunaj okvirja odstranimo in izrezano sliko normaliziramo na 128 pikslov v višino in širino.

## 2.2 Gradnja predstavitev obraza in klasifikatorja

Za predstavitev obrazov uporabimo metodo *vreče besed* (BOW - *angl. bag of words*) [20]. Značilne lokacije na sliki obraza predstavimo z opisniki, s pomočjo katerih zgradimo besednjak za vrečo besed. Vrečo besed uporabi klasifikator, ki je zmožen ločevanja med različnimi osebami. Glavni koraki predlagane metode so:

- 1) Določanje značilnih lokacij na sliki in gradnja opisnikov. Opisnik opisuje bližnjo okolico značilne lokacije in je predstavljen kot vektor.
- 2) Gradnja besednjaka oz. razvrščanje opisnikov v roje.
- 3) Gradnja histograma opisnikov oz. vreče besed, ki prešteje število opisnikov, dodeljenih k vsakemu roju.
- 4) Uporaba večrazrednega klasifikatorja, kjer vrečo besed, predstavljeno v vektorski obliki, uporabimo za dodeljevanje h kategoriji oz. osebi.

Idealno so ti koraki zasnovani tako, da maksimirajo klasifikacijsko točnost, medtem ko minimizirajo računsko zahtevnost. Za opisnike, pridobljene v prvem koraku, mora veljati, da so robustni na spremembe, ki niso povezane s kategorizacijo (npr. razlike zaradi osvetlitve), hkrati pa morajo vsebovati dovolj informacij, ki omogočajo, ločevanje med kategorijami. Poskrbeti moramo, da je besednjak dovolj velik, da omogoča ločevanje pomembnih sprememb v delih slike in hkrati ne preveč velik, da se izognemo ločevanju nepomembnih podrobnosti, kot je npr. šum.



Slika 5: Izbera značilnih lokacij. Levi del slike prikazuje izbiro značilnih lokacij z mrežo, desni del slike prikazuje značilne lokacije, izbrane z detektorjem. Uporabili smo detektor SURF.

**2.2.1 Značilne lokacije:** Najprej na sliki določimo značilne lokacije, okoli katerih želimo opisati lokalno regijo oz. jo predstaviti z opisnikom. Raziskali smo dva pristopa. Prvi deluje tako, da na sliko postavimo *mrežo* (slika 5 levo). Presečišča mrežnih črt pomenijo značilne lokacije, za katere izračunamo opisnike. Drug pristop uporablja detektorje značilnih točk (slika 5 desno).

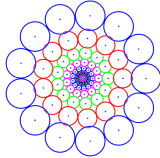
**2.2.2 Gradnja besednjaka:** Pri gradnji sistema za prepoznavo obrazov je pomembno, kako zgradimo besednjak. Besednjak uporablja klasifikator, ki ga je treba učiti. Če bi vsak opisnik iz prejšnjega koraka pomenil besedo v besednjaku, bi bilo učenje časovno zelo zahtevno. Če bi veliko opisnikov dodelili enemu roju oz. besedi, bi izgubili na natančnosti klasifikatorja. Število rojev  $k$  je treba pazljivo izbrati, v našem primeru je bilo rojev

1000. Za združevanje opisnikov v roje oz. besede uporabimo algoritem  $k$ -means [21]. Za vsak obraz preštejemo enake besede in zgradimo histogram oz. vrečo besed, ki je pravzaprav vektor, s katerim predstavimo obraz.

**2.2.3 Razvrščanje v razrede s SVM klasifikatorjem:** Razvrščanje v razrede vključuje učenje klasifikatorja in njegovo testiranje. Predlagan sistem za prepoznavo obrazov uporablja SVM klasifikator. SVM klasifikator najde optimalno hiperravnino, ki razdeli prostor podatkov, v prejšnjem koraku izračunane vreče besed, v dva razreda. Optimalna hiperravnina je določena tako, da je ločitvena meja med razredoma čim širša. SVM se v osnovi uporablja za razvrščanje v dva razreda. Za razširitev na večrazredno razvrščanje moramo opraviti primerjavo vseh mogočih  $N$  razredov med seboj. Končno število primerjav je tako  $N \cdot (N - 1)/2$ .

### 2.3 AG opisnik

Raziskati smo želeli, kako zanesljiva je prepoznavna obrazov, če za opisnik ključnih lokacij uporabimo AG opisnik [11]. Večina popularnih opisnikov, med katerimi je najbolj poznan SIFT, uporablja za predstavitev lokalne regije na sliki informacijo o gradientu oz. o velikosti spremembe intenzivnosti in smeri spremembe. AG opisnik uporabi model odvisnih učinkov (angl. *Dependent Effects Model* [22]). Opisnik vzorči sliko v okolici značilne lokacije na krožnicah z eksponentno naraščajočim polmerom. Zajete vrednosti so Gaussova povprečja, izračunana na okoliških piksljih. Standardni odklon filtrov narašča eksponentno s številko krožnice. Princip vzorčenja prikazuje slika 6. Odčitana Gaussova



Slika 6: Marjetica ponazarja vzorčenje informacije na sliki z AG opisnikom. Modre točke so mesta, kjer vzorčimo z Gaussovimi filtri. Polmeri barvnih krogov ustrezajo standardnemu odklonu Gaussovega filtra. Vzorec je sestavljen iz devetih krožnic. Na vsaki krožnici zajamemo 13 vrednosti.

povprečja  $I_{mn}$ ;  $m = 0, \dots, M - 1$ ,  $n = 0, \dots, N - 1$  lahko ponazorimo s pravokotnikom  $P$  dimenzije  $M \times N$ . Vrednost  $M$  označuje število krožnic, na katerih smo vzorčili, oz. število stolpcev v  $P$ , vrednost  $N$  pa število vzorcev na krožnici oz. število vrstic v  $P$ . V našem primeru velja:  $M = 9$  in  $N = 13$ . Naj  $C_m$  in  $R_n$  označujeta seštevek vrednosti iz  $m$ -tega stolpca in  $n$ -te vrstice pravokotnika  $P$ . Velja  $C_m = \sum_{n=0}^{N-1} I_{mn}$  in  $R_n = \sum_{m=0}^{M-1} I_{mn}$ . Pravokotnik  $P$  lahko predstavimo s tremi tipi učinkov: učinki stolpcev  $\alpha_m$ , učinki vrstic  $\beta_n$ , in učinki preostanka  $\gamma_{mn}$ , za  $m = 0 \dots M - 1$  in  $n = 0 \dots N - 1$ . Vsak  $I_{mn}$  lahko predstavimo kot

linearno kombinacijo treh zgoraj navedenih učinkov

$$I_{mn} = \bar{I} + \alpha_m + \beta_n + \gamma_{mn}. \quad (5)$$

$\bar{I}$  označuje povprečno vrednost, izračunano na vrednostih pravokotnika  $P$ .

$$\bar{I} = \frac{1}{MN} \cdot \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{mn}, \quad (6)$$

$\alpha_m$  in  $\beta_n$  označujeta razliko povprečne vrednosti stolpca  $m$  in vrstice  $n$  od povprečne vrednosti pravokotnika  $\bar{I}$ :

$$\alpha_m = \frac{1}{N} \cdot C_m - \bar{I}, \quad (7)$$

$$\beta_n = \frac{1}{M} \cdot R_n - \bar{I}. \quad (8)$$

Učinke preostanka  $\gamma_{mn}$  lahko izrazimo kot:

$$\gamma_{mn} = I_{mn} - \bar{I} - \alpha_m - \beta_n. \quad (9)$$

Enostavno je pokazati [11], da so učinki  $\beta_n$  enaki učinkom  $-\gamma_{0n}$ , zato opisnik uporabi le učinke  $\alpha$  in  $\gamma$ . Od tod sledi tudi ime AG opisnika kot okrajšava za alfa-gama.

Z normalizacijo opisnika dosežemo robustnost opisnika na velikost spremembe v kontrastu. Vrednosti učinkov normaliziramo s standardnimi odkloni, izračunanimi za vektor  $\alpha$  učinkov in stolpce pravokotnika  $\gamma$  učinkov. Velja:

$$\hat{\alpha} = \left( \frac{\alpha_0}{\sigma_\alpha}, \frac{\alpha_1}{\sigma_\alpha}, \dots, \frac{\alpha_{M-1}}{\sigma_\alpha} \right) \quad (10)$$

in

$$\hat{\gamma} = \begin{pmatrix} \frac{\gamma_{00}}{\sigma_\gamma} & \frac{\gamma_{10}}{\sigma_\gamma} & \dots & \frac{\gamma_{(M-1),0}}{\sigma_\gamma} \\ \frac{\gamma_{01}}{\sigma_\gamma} & \frac{\gamma_{11}}{\sigma_\gamma} & \dots & \frac{\gamma_{(M-1),1}}{\sigma_\gamma} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma_{0,(N-1)}}{\sigma_\gamma} & \frac{\gamma_{1,(N-1)}}{\sigma_\gamma} & \dots & \frac{\gamma_{(M-1),(N-1)}}{\sigma_\gamma} \end{pmatrix}, \quad (11)$$

kjer so  $\sigma_\alpha$  in  $\sigma_\gamma$ ;  $m = 0 \dots M - 1$  ocenjeni standardni odkloni, izračunani z naslednjima enačbama:

$$\sigma_\alpha = \sqrt{\frac{\sum_{m=0}^{M-1} \alpha_m^2}{M-1}}, \quad \sigma_\gamma = \sqrt{\frac{\sum_{n=0}^{N-1} \gamma_{mn}^2}{N-1}}. \quad (12)$$



### 3 EKSPERIMENTIRANJE

#### 3.1 Testne baze slik

Predlagani sistem smo testirali na treh bazah, *faces94*, *faces95* in *faces96* zbirke slik obrazov Univerze v Essexu\*. Podatkovne baze se po težavnosti razlikujejo.



Slika 7: Primeri slik iz zbirke slik obrazov Univerze v Essexu. Prva vrstica predstavlja slike iz *faces94*, druga iz *faces95* in tretja vrstica iz podatkovne baze *faces96*.

Podatkovna baza *faces94* vsebuje 3059 slik 153 oseb, med katerimi je 20 žensk in 133 moških. Ozadje na slikah je zeleno platno. Osvetlitev obrazov je enaka, velikosti obraza na slikah so enake. Obrazi so lahko rahlo nagnjeni, prisotna pa je tudi sprememba v izrazih na obrazu.

Podatkovna baza *faces95* vsebuje 1440 slik 72 oseb. Ozadje je rdeča zavesa, ki vsebuje sence, padajoče s slikane osebe. S premikanjem osebe naprej se zgodijo velike spremembe v osvetlitvi obraza zaradi umetnega osvetljevanja. Prisotne so manjše razlike v velikosti obraza, nagibu glave in izrazih na obrazu.

Podatkovna baza *faces96* vsebuje 3016 slik 152 oseb. Iz podatkovne baze so bile izločene tri slike, na katerih nismo zaznali obraza. Naš detektor obrazov zaznava frontalni del obraza in ne stranskega profila obraza, ki je prisoten na teh slikah. Ozadje je kompleksno, saj vsebuje bleščeče posterje. Prisotne so razlike v velikosti, izrazih in nagibu obraza. Osvetlitev obraza se spremeni, ko se oseba pomakne naprej.

#### 3.2 Rezultati

**3.2.1 Izbira tipa izrezave obraza:** Ugotoviti smo želeli, kateri tip izrezave obraza da najvišjo klasifikacijsko točnost. Klasifikacijska točnost je delež pravilno prepoznanih obrazov glede na vse testirane obraze. Prepoznavna obraza je pravilna, če pravilno prepoznamo osebo, ki ji obraz pripada. Podatkovno bazo smo naključno razdelili na učno in testno množico. Učno množico predstavlja 80 % vseh slik, testno množico pa

preostalih 20 % slik. Parameter  $k$ , ki pomeni število rojev za algoritem *k-means*, smo nastavili na 1000 in izvedli 10 ponovitev. Testiranje smo ponovili za vse tri tipe izrezav. Značilne lokacije smo najprej določili z mrežo z velikostjo celic  $8 \times 8$  pikselov. Testirali smo naslednje opisnike lokalne regije: SURF [10], SIFT [8], BRISK [9], ORB [15], KAZE [14] in AG [11]. Rezultate za vse tri testne baze in tri različne tipe izrezav prikazuje tabela 1. Kot je razvidno iz posamičnih rezultatov in povprečja rezultatov, izrezava tipa 3 dosega najvišjo klasifikacijsko točnost. Rezultat se zdi smiseln, saj izrezava tipa 3 vključuje največji del obraza, torej največ informacije. Opazimo lahko, da se klasifikacijska točnost zmanjšuje s težavnostjo podatkovne baze. Klasifikacijska točnost je najvišja za bazo *faces94* in najnižja za bazo *faces96*. Rezultati, ki so jih dosegli različni opisniki, so si podobni, gre za majhne razlike, manjše od 1 %. Za izrezavo tipa 3 dosega opisnik AG na bazi *faces96* najboljši rezultat, na bazi *faces95* pa drugi najboljši rezultat. Najslabše rezultate je dosegel opisnik BRISK.

Algoritem smo testirali tudi za primer značilnih lokacij, ki jih dobimo z detektorjem. Testirali smo naslednje detektorje: SURF [10], SIFT [8], BRISK [9], ORB [15] in KAZE [14]. Rezultate prikazuje tabela 2. Tudi tu da izrezava tipa 3 v večini primerov najboljše rezultate. Klasifikacijska točnost se zmanjšuje s težavnostjo podatkovne baze. Če izberemo izrezavo tipa 3, dosežemo z uporabo detektorjev boljšo kvalifikacijsko točnost kot pri izbiri ključnih lokacij z mrežo.

**3.2.2 Rezultati ob prisotnosti zameglitve in zakritij:** Preizkusiti smo želeli tudi robustnost predlaganega sistema ob prisotnosti zameglitve slike in zakritja dela obraza.

Za prvo nalogo smo slike filtrirali z Gaussovimi filtrom s standardno deviacijo  $\sigma = 1,1$  piksel. Eksperiment smo izvedli pri določitvi značilnih lokacij z mrežo in detektorjem. Rezultate za izrezavo tipa 3 prikazujeta tabeli 3 in 4. Rezultati, doseženi pri uporabi detektorja značilnih lokacij, so večinoma boljši na bazah *faces95* in *faces96*, pri bazi *faces94* pa so rezultati v večini primerov malenkostno slabši od rezultatov, dobljenih pri uporabi mreže. Opisnik AG je dosegel najboljše rezultate za detektor BRISK in ORB.

Za drugo nalogo smo četrtino izrezane slike obraza zakrili s kvadratom intenzivnosti nič. Vsako sliko smo zakrili na štiri različne načine s polaganjem kvadrata v drugo oglišče slike. Eksperiment smo izvedli pri določitvi značilnih lokacij z mrežo in detektorjem. Rezultate za izrezavo tipa 3 prikazujeta tabeli 5 in 6. Rezultati so pri uporabi značilnih lokacij, določenih z mrežo malenkostno boljši na testni bazi *faces94* in *faces96* in slabši na bazi *faces95*. V tem testu se je najbolje izkazal opisnik AG. Pri izbiri značilnih lokacij z detektorjem je dosegel najboljše rezultate v kombinaciji z detektorjem KAZE.

\*Zbirka Univerze v Essexu je dosegljiva na: <http://cswwww.essex.ac.uk/mv/allfaces/>.

Opisnik	SURF	SIFT	BRISK	ORB	KAZE	AG	povprečje
KT[%] <i>faces94</i> -tip 1	<b>99,36</b>	<b>99,45</b>	99,18	<b>99,40</b>	99,12	<b>99,45</b>	99,33
KT[%] <i>faces94</i> -tip 2	<b>99,79</b>	99,77	99,64	<b>99,78</b>	<b>99,91</b>	<b>99,78</b>	99,78
KT[%] <i>faces94</i> -tip 3	<b>99,97</b>	<b>99,93</b>	<b>99,90</b>	<b>99,90</b>	99,88	99,86	99,91
KT[%] <i>faces95</i> -tip 1	<b>94,13</b>	93,34	93,87	<b>94,14</b>	<b>94,06</b>	93,67	93,87
KT[%] <i>faces95</i> -tip 2	97,16	<b>97,91</b>	97,50	<b>97,61</b>	97,56	<b>97,68</b>	97,57
KT[%] <i>faces95</i> -tip 3	<b>98,71</b>	98,43	98,53	98,42	<b>98,92</b>	<b>98,88</b>	98,65
KT[%] <i>faces96</i> -tip 1	84,43	<b>84,68</b>	84,56	<b>84,92</b>	<b>85,00</b>	84,54	84,69
KT[%] <i>faces96</i> -tip 2	<b>93,89</b>	<b>93,71</b>	93,35	<b>94,04</b>	93,50	93,58	93,68
KT[%] <i>faces96</i> -tip 3	96,66	96,54	96,45	<b>96,71</b>	<b>96,76</b>	<b>96,78</b>	96,65

Tabela 1: Klasifikacijska točnost (KT), dosežena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z mrežo in pri uporabi različnih tipov izrezave. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

Detektor	SURF	SIFT	BRISK	ORB	KAZE	SURF	SIFT	BRISK	ORB	KAZE	
Opisnik	SURF	SIFT	BRISK	ORB	KAZE	AG	AG	AG	AG	AG	povprečje
KT[%] <i>faces94</i> -tip 1	<b>99,50</b>	99,43	99,36	99,34	99,44	99,28	99,43	99,36	<b>99,62</b>	<b>99,66</b>	99,44
KT[%] <i>faces94</i> -tip 2	<b>99,80</b>	99,76	99,67	99,78	99,70	99,76	99,69	<b>99,92</b>	99,76	<b>99,81</b>	99,76
KT[%] <i>faces94</i> -tip 3	<b>99,85</b>	<b>99,86</b>	99,83	99,80	99,84	99,85	99,76	99,79	<b>99,86</b>	<b>99,88</b>	99,83
KT[%] <i>faces95</i> -tip 1	97,22	<b>97,96</b>	97,53	97,54	<b>97,89</b>	98,06	97,48	97,69	97,53	<b>97,90</b>	97,68
KT[%] <i>faces95</i> -tip 2	<b>98,62</b>	98,44	98,53	98,38	98,47	<b>98,60</b>	98,49	98,29	<b>98,55</b>	<b>98,55</b>	98,49
KT[%] <i>faces95</i> -tip 3	98,75	<b>99,20</b>	98,67	<b>98,97</b>	98,86	98,75	<b>99,31</b>	98,96	98,88	98,74	98,91
KT[%] <i>faces96</i> -tip 1	85,39	85,47	85,52	<b>85,59</b>	84,98	<b>85,63</b>	84,77	<b>85,97</b>	85,56	84,53	85,34
KT[%] <i>faces96</i> -tip 2	92,51	<b>93,53</b>	93,23	92,72	92,72	92,65	92,69	<b>93,39</b>	<b>92,96</b>	92,50	92,89
KT[%] <i>faces96</i> -tip 3	<b>97,78</b>	97,28	97,29	97,26	<b>97,87</b>	97,37	97,20	97,56	<b>97,70</b>	97,65	97,50

Tabela 2: Klasifikacijska točnost (KT), dosežena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z detektorjem in pri uporabi različnih tipov izrezave. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

	SURF	SIFT	BRISK	ORB	KAZE	AG
KT[%] <i>faces94-zameglitev</i> -tip 3	99,86	99,82	99,91	<b>99,93</b>	<b>99,88</b>	<b>99,86</b>
KT[%] <i>faces95-zameglitev</i> -tip 3	<b>98,80</b>	<b>98,66</b>	98,37	98,32	<b>98,56</b>	98,35
KT[%] <i>faces96-zameglitev</i> -tip 3	<b>96,85</b>	<b>96,83</b>	96,72	<b>97,06</b>	96,50	96,48

Tabela 3: Klasifikacijska točnost (KT), dosežena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z mrežo pri uporabi izrezave tipa 3 in zameglitvi slike z Gaussovimi filtrom. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

Detektor	SURF	SIFT	BRISK	ORB	KAZE	SURF	SIFT	BRISK	ORB	KAZE
Opisnik	SURF	SIFT	BRISK	ORB	KAZE	AG	AG	AG	AG	AG
KT[%] <i>faces94-zameglitev</i> -tip 3	99,80	99,81	<b>99,85</b>	99,76	<b>99,93</b>	99,76	99,77	<b>99,86</b>	99,79	99,81
KT[%] <i>faces95-zameglitev</i> -tip 3	99,14	99,08	99,17	<b>99,23</b>	98,76	99,02	99,00	<b>99,35</b>	<b>99,32</b>	98,84
KT[%] <i>faces96-zameglitev</i> -tip 3	<b>97,78</b>	97,28	97,29	97,26	<b>97,87</b>	97,37	97,20	97,56	<b>97,70</b>	97,65

Tabela 4: Klasifikacijska točnost (KT), dosežena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z detektorjem pri uporabi izrezave tipa 3 in zameglitvi slike z Gaussovimi filtrom. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

	SURF	SIFT	BRISK	ORB	KAZE	AG
KT[%] <i>faces94-zakritja</i> -tip 3	<b>99,90</b>	99,87	<b>99,89</b>	<b>99,90</b>	99,86	<b>99,91</b>
KT[%] <i>faces95-zakritja</i> -tip 3	98,02	<b>98,15</b>	98,01	97,99	<b>98,24</b>	<b>98,25</b>
KT[%] <i>faces96-zakritja</i> -tip 3	96,94	<b>97,10</b>	96,97	<b>97,05</b>	96,63	<b>97,02</b>

Tabela 5: Klasifikacijska točnost (KT), dosežena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z mrežo pri uporabi izrezave tipa 3 in prisotnosti zakritij na slikah. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

Detektor	SURF	SIFT	BRISK	ORB	KAZE	SURF	SIFT	BRISK	ORB	KAZE
Opisnik	SURF	SIFT	BRISK	ORB	KAZE	AG	AG	AG	AG	AG
KT[%] <i>faces94-zakritja</i> -tip 3	99,78	<b>99,80</b>	99,75	99,75	99,79	99,79	99,79	99,78	<b>99,81</b>	<b>99,82</b>
KT[%] <i>faces95-zakritja</i> -tip 3	98,37	98,49	98,46	98,47	98,43	<b>98,51</b>	98,46	98,38	<b>98,52</b>	<b>98,55</b>
KT[%] <i>faces96-zakritja</i> -tip 3	<b>95,86</b>	95,51	95,72	95,56	<b>95,74</b>	95,58	95,64	95,67	95,65	<b>95,91</b>

Tabela 6: Klasifikacijska točnost (KT), pridobljena na bazah *faces94*, *faces95* in *faces96* z izbiro značilnih lokacij z detektorjem pri uporabi izrezave tipa 3 ter prisotnosti zakritij na sliki obraza. **Prvi, drugi** in **tretji** rezultat v vrstici so obarvani in poudarjeni.

Na testiranih podatkovnih bazah sta se metodi pridobivanja značilnih lokacij s pomočjo mreže in detektorjem izkazali za uspešni. Klasifikacijsko točnost bi lahko izboljšali s povečanjem števila rojev  $k$ , metodo z mrežo bi lahko še dodatno izboljšali, če bi uporabili gostejšo mrežo. Razlog, da tega nismo storili, je v omejitvi količine virov, ki so nam bili na voljo. Vsi uporabljeni lokalni opisniki in detektorji ključnih točk, z izjemo opisnika AG so implementirani v knjižnici OpenCV. Te implementacije smo tudi uporabili pri testiranju. Koda za opisnik AG je dostopna na [16].

#### 4 SKLEP

V delu smo predlagali sistem za prepoznavo obrazov. Sistem najprej predprocesira sliko. Ta korak vključuje detekcijo obraza, iskanje verige mejnih točk in izrezavo slike obraza. Raziskali smo tri izrezave slike, za najboljšo se je izkazala izrezava, ki vključuje cel obraz. Za prepoznavo uporabimo metodo vreče besed in klasifikator SVM. Vreča besed uporablja opisnike AG značilnih lokacij. Predlagali smo dve metodi za določanje značilnih lokacij; prva določi lokacije s pomočjo mreže, druga uporabi detektorje značilnih točk. Testirali smo več detektorjev, rezultati prepoznave so bili podobni. Celoten sistem smo testirali na treh javno dostopnih bazah, ki se po težavnosti razlikujejo. Sistem smo testirali tudi za druge popularne opisnike lokalnih regij. Opisnik AG se je izkazal kot primeren, saj je dosegel povsem konkurenčne rezultate k drugim testiranim opisnikom, pri delno zakritih obrazih pa je dosegel najvišjo klasifikacijsko točnost.

#### ZAHVALA

Raziskavo je omogočilo Ministrstvo za izobraževanje, znanost in šport Republike Slovenije v okviru programa P2-0214-Računalniški vid.

#### REFERENCES

- [1] P. Melin and W. Pedrycz. *Soft Computing for Recognition Based on Biometrics*, volume 312, Springer, 2010.
- [2] A. J. Goldstein, L. D. Harmon and A. B. Lesk. Identification of human faces. *Proceedings of the IEEE*, 59(5), pp. 748–760, 1971.
- [3] L. Sirovich and J. Kirby. Low-dimensional procedure for the characterization of human faces. *Josa A*, 4(3), pp. 519–524, 1987.
- [4] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. *Proceedings of the CVPR*, pp. 586–591, 1991.
- [5] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), pp. 103–108, 1990.
- [6] K. Etamad and R. Chellappa. Discriminant analysis for recognition of human face images. *JOSA A*, 14(8), pp. 1724–1733, 1997.
- [7] M. S. Bartlett, J. R. Movellan and T. J. Sejnowski. Face recognition by independent component analysis. *IEEE Transactions on neural networks*, 13(6), pp. 1450–1464, 2002.
- [8] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), pp. 91–110, 2004.

- [9] S. Leutenegger, M. Chli and R. Y. Siegwart. Brisk: Binary robust invariant scalable keypoints. *ICCV*, pp. 2548–2555, 2011.
- [10] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3), pp. 346–359, 2008.
- [11] R. Mandeljc and J. Maver AGs: local descriptors derived from the dependent effects model. *Journal of Visual Communication and Image Representation* 58, pp. 503–514, 2019.
- [12] M. Calonder, V. Lepetit, C. Strecha and P. Fua. Brief: Binary robust independent elementary features. *ECCV*, pp. 778–792, 2010.
- [13] A. Alahi, R. Ortiz and P. Vandergheynst. Freak: Fast retina keypoint. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 510–517, 2012.
- [14] P. F. Alcantarilla, A. Bartoli and A. J. Davison. Kaze features. *In European Conference on Computer Vision*, pp. 214–227, 2012.
- [15] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, ORB: An efficient alternative to SIFT or SURF, *In Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2564–2571, 2011.
- [16] R. Mandeljc, AlphaGamma descriptor, <https://github.com/rokn/alphagamma-descriptor>, 2017.
- [17] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pp. 886–893, 2005.
- [18] P. F. Felzenszwalb, R. B. Girshick, D. McAllester and D. Ramanan. Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9), pp. 1627–1645, 2010.
- [19] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1867–1874, 2014.
- [20] G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray. Visual categorization with bags of keypoints. *In Workshop on Statistical Learning in Computer Vision, ECCV*, pp. 1–22, 2004.
- [21] D. G. Stork, O. Duda, P. E. Hart. *Pattern classification*. John Wiley & Sons, 2000.
- [22] J. Maver. Self-similarity and points of interest. *IEEE transactions on pattern analysis and machine intelligence*, 32 (7), pp. 1211–1226, 2010.

**Domen Lušina** je leta 2017 postal diplomirani inženir računalništva in informatike (UN) na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Trenutno nadaljuje študij na magistrskem programu Računalništvo in matematika na Fakulteti za matematiko in fiziko Univerze v Ljubljani. Zanima se predvsem za področje umetne inteligence in rudarjenja podatkov.

**Jasna Maver** je leta 1995 doktorirala s področja računalništva na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. V letih 1990 in 1991 je bila raziskovalka v laboratoriju GRASP na Univerzi v Pensilvaniji, kjer je raziskovala na področju aktivnega robotskega vida. Je izredna profesorica za računalništvo in informatiko na Univerzi v Ljubljani. Zaposlena je na Oddelku za bibliotekarstvo, informacijsko znanost in knjigarstvo na Filozofski fakulteti, kjer predava računalniške predmete. Na raziskovalnem področju sodeluje z Laboratorijem za računalniški vid in Laboratorijem za umetne vizualne spoznavne sisteme na Fakulteti za računalništvo in informatiko. Njeno trenutno področje raziskovanja je gradnja nizkonivojskih opisnikov lokalnih regij v slikah, ki jih lahko uporabimo za različne naloge računalniškega vida, med katere lahko uvrstimo tudi prepoznavo obrazov.