

Prepoznavna hoje

Jan Krivec, 63170013

Fakulteta za računalništvo in informatiko
Interaktivnost in oblikovanje informacij

Abstract. V seminarski naogi je predstavljena izdelava programa za klasifikacijo oseb po hoji. V delu najprej raziščemo predhodna dela, ki so povezana s to tematiko in nato predstavimo implementacijo našega algoritma. Za predhodnimi deli sledi opis metodologije, kjer opišemo vse uporabljene metode. V tem delu predstavimo uporabljen način segmentacije, generiranja GEI slik in klasifikacije z analizo glavnih komponent in logistično regresijo. Sledijo rezultati, v katerih predstavimo način testiranja klasifikacije na podlagi GEI slik in rezultate, ki so za vse zorne kote snemanja posnetkov večji od 96%. Nato predstavimo delovanje programa in v zaključku povzamemo vse ugotovitve.

1 Uvod

V današnjih časih se umetna inteligenca in strojno učenje uporabljata za identifikacijo ljudi. Posameznika lahko na te načine prepoznamo po njegovih prstnih odtisih, sliki očesa in obraza. V tej seminarski nalogi pa smo se posvetili prepoznavi osebe po hoji. Čeprav so druge oblike identifikacije natančnejše, za njih potrebujemo predhodne podatke za osebo, kot na primer prstne odtise. V primeru identifikacije po hoji potrebujemo le posnetek osebe, na katerem ta hodi. Identifikacija ljudi po hoji je danes že uporabljena na Kitajskem, kjer prebivalstvo nadzirajo preko kamer.

2 Predhodna dela

Identifikacija ljudi po hoji ima dva pristopa. Prvi pristop temelji na ekstrakciji silhete oz. obrisa postave. Tako se modeli učijo na osnovi vhodnih slik, ki predstavljajo bitne mape silhuet. Obstaja več programov za identifikacijo ljudi po hoji, ki uporabljajo silhete. Med bolj pomembne pa spada GaitSet [2], kjer avtorji hojo predstavijo kot globoko množico, kar pomeni, da je množica slik hoje integrirana z globalno-lokalno združenim globokim omrežjem. Na ta način je identifikacija neodvisna od odstopanj pri posameznih sličicah, saj lahko integrira slike iz ostalih zornih kotov in scenarijev, kot na primer druga oblačila ali pa celo nošenje ročne prtljage [2].

Med poznane modele sodi tudi model CSTL [3], kontekstno občutljivo omrežje za učenje časovnih značilnosti (*angl. context-sensitive temporal feature learning network*). Model združuje učenje v treh nivojih, da pridobi predstavitev

gibanja glede na časovne kontekstualne informacije. Natančneje, CSTL uvaja modeliranje odnosov med značilnostmi več meril za ocenjevanje pomembnosti značilnosti, na podlagi česar omrežje prilagodljivo krepi pomembnejše mere in zavira manj pomembne mere [3].

Drugi pritop temelji na ekstrakciji človeškega okostja iz slike. Slike silhuet lahko zaradi motenj izgubijo drobne prostorske informacije. Trpijo tudi zaradi (samo) okluzije in jih je v resničnem svetu težko pridobiti. Poleg tega te silhuete vsebujejo tudi druge vizualne namige, ki niso dejanske značilnosti hoje in jih je mogoče uporabiti za identifikacijo, pa tudi za zavajanje sistema. Model Gait-Graph [7] tako temelji na grafovskih konvolucijskih omrežjih, ki združuje vhode višjega reda in preostala omrežja v učinkovito arhitekturo za prepoznavanje hoje.

Na sliki 1 lahko vidimo primerjavo med vhodnimi podatki za modele, ki delajo klasifikacijo nad silhuetami in vhodnimi podatki za modele, ki za klasifikacijo uporabijo človeški skelet.

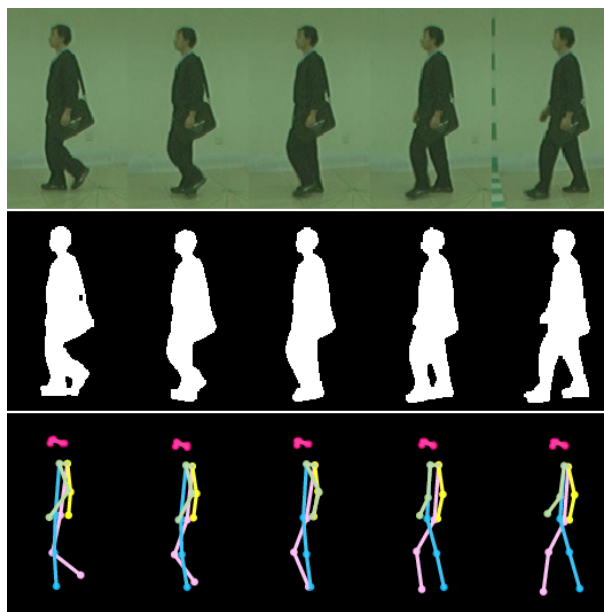


Fig. 1: Primerjava vrst modelov za klasifikacijo hoje.

Vsi opisani modeli prejmejo zaporedje slik in nato klasificirajo podatke. V naši implementaciji modela smo poskusili klasifikacijo osebe izvesti v realnem času. Kot osnovo za naš projekt smo vzeli github repozitorij uporabnika Chun-feng Song [6], saj je bil le-ta edini, ki je zasnovan za klasifikacijo v realnem času. Kot osnovo klasifikacije uporablja posnetke silhuet, ki jih uporabi za generacijo GEI slike.

3 Metodologija

Implementacija algoritma je sestavljena iz treh delov.

3.1 Segmentacija:

Algoritem kot vhod prebere video s pomočjo spletne kamere. Pred naslednjim korakom moramo tako iz trenutne slike razbrati obris osebe. Obstaja več algoritmov za pridobivanje silhuete iz slike. Obstajajo natančni algoritmi, ki iz slike zelo dobro izločijo silhueto človeka. Visoka natančnost teh algoritmov pa je za ceno hitrosti. Tako uporaba teh algoritmov ni mogoča v realnem času. Tako smo v našem projektu uporabili knjižnjico cv2, omogoča ekstrakcijo silhuete v realnem času. Za zaznavanje obrisov potrebujemo sprva sliko, na kateri je le ozadje. Tako najprej najdemo absolutno razliko med trenutno sliko in sliko ozadja, nato pa na dobljeni razliki poiščemo obrise. Med obrisi poiščemo največjega in preverimo, če je njegova višina večja od širine. S tem dodatno preverimo, da gre za obris osebe. Metoda je zelo občutljiva na spremembe v ozadju.

3.2 Generacija GEI slike:

GEI (*angl. Gait Energy Image*) slika se zgenerira iz dobljenih silhuet vsake slike. Iz vsake zajete slike izrežemo njeno silhueto in spremenimo njeno velikost, da lahko nato zgeneriramo GEI sliko. Le-to dobimo tako, da naredimo povprečje zajetih slik. Pred izvedbo programa lahko nastavimo koliko zaporednih slik bomo uporabili za generacijo GEI slike. Na sliki 2 lahko vidimo primer zgenerirane GEI slike.

3.3 Klasifikacija:

Pridobljene slike iz drugega koraka uporabimo za klasifikacijo. Če želimo ljudi zaznavati v realnem času so algoritmi, opisani v poglavju 2 neuporabni, saj so preveč časovno kompleksni, da bi jih lahko uporabili v realnem času. Avtorji so v svoji prvotni implementaciji [6] za klasifikacijo uporabili le primerjavo vrednosti na posameznih pikslih. Ta deluje dovolj hitro, da lahko rezultate klasifikacije dobimo v realnem času, vendar so pa rezultati zaradi tega zelo slabi. Za našo rešitev smo zato vzeli idejo javnega github repozitorija, kjer so avtorji s svojo metodo uspeli doseči 95,3% klasifikacijsko točnost [5].

Model tako vhodnim slikam najprej zmanjša dimenzionalnost z uporabo analize glavnih komponent PCA. Pred uporabo metode PCA vse slike pretvorimo v enodimenzionalne sezname. Klasifikacija na dobljenih rezultatih analize glavnih komponent poteka s pomočjo logistične regresije.

Analiza glavnih komponent Analiza glavnih komponent (*ang. Principal Component Analysis, PCA*) je statistična metoda, ki temelji na ortogonalni transformaciji primarnih spremenljivk. Metoda nam poišče med seboj pravokotne smeri,



Fig. 2: Primer GEI slike, zgenerirane na osebkcu iz baze CASIA-B [1].

v katerih je varianca podatkov največja. Nove smeri imenujemo glavne komponente, njihovo število pa je enako številu prvotnih spremenljivk. Naslednji korak, ki ga omogoča PCA je redukcija dimenzij. Ker so komponente dobljene na način, da vsaka naslednja predstavlja manjši delež skupne variance kot prejšnja, lahko določeno število komponent odstranimo iz nadaljne analize in s tem izgubimo zanemarljivo malo podatkov [4]. Ravno zaradi tega, je PCA zdelo uporabna pri velikih podatkovnih nožicah, katerih dimenzionalnost je velika.

Logistična regresija Logistična regresija se ogosto uporablja za klasifikacijo in napovedno analitoko. Ocenjuje verjetnost izida na podlagi podanega nabora neodvisnih spremenljivk. Mi smo uporabili multinomsko logistično regresijo, ki je posplošitev normalne logistične regresije, omogoča pa klasifikacijo v več kot dva razreda.

4 Rezultati

Uporabljene metode smo testirali s pomočjo podatkovne baze CASIA-B [1], ki vsebuje posnetke 124 subjektov, posnetih iz 11 različnih zornih kotov. Podatki, ki so prosto dostopni na spletni strani vsebujejo slike silhuet za vsako sliko posnetka. Za testiranje smo obdržali posnetke v normalnih pogojih (brez dodatnih oblačil in prtljage). Tako nam je za vsak subjekt ostalo 6 različnih primerov normalne hoje, posnetih z 11 različnih kotov.

Ker prosto dostopna podatkovna baza CASIA-B vsebuje le slike silhuet, smo testirali le klasifikacijo oseb na podlagi GEI slik. Segmentacije nismo testirali, saj nismo imeli pravih podatkov za testiranje, hkrati pa je na lastnih posnetkih težko določiti natančnost segmentacije. Z uporabo že segmentiranih slik za učenje klasifikatorja smo zagotovili, da so bili rezultati odvisni le od kvalitete modela, ne pa tudi od kvalitete segmentacije.

Iz vsake slike silhete smo izluščili le obris in ga centriral na sredino nove slike velikosti 128*88 pikslov. Isto velikost slik uporabljamo tudi v našem programu. Iz obrezanih slik smo nato naredili njihovo povprečje, ki predstavlja GEI sliko.

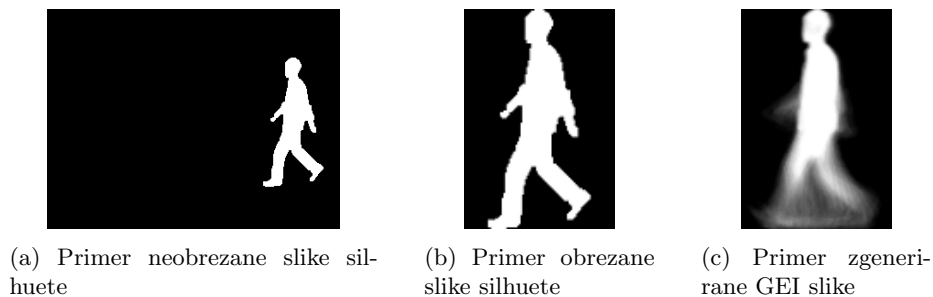


Fig. 3: Proces priprave podatkov baze CASIA-B za testiranje klasifikacije.

Tako smo za vse subjekte pod vsemi zornimi koti generirali GEI slike iz sekvenc hoje. Klasifikacijo smo naredili za vak zorni kot posebaj. Pri vsakem subjektu smo za vsak zorni kot 4 primere porabili za učenje modela, 2 primera pa za testiranje. Tako smo za testiranje vsakega zornega kota imeli na voljo približno 496 posnetkov (124 osebkov * 4 posnetki) za učenje in 248 posnetkov (124 osebkov * 2 posnetka) za testiranje. Pri številu GEI slik lahko pride do odstopanja zaradi nepopolnih podatkov baze CASIA-B.

Kot lahko vidimo iz rezultatov v tabeli 1, so le-ti zelo dobri za vse zorne kote. Hkrati pa se reduciranje dimenzionalnosti s PCA, treniranje modela za klasifikacijo in nato še testiranje izvede zelo hitro. Tako smo na primer vse podatke za zorni kot 18° (, kjer so vsi podatki popolni) uspeli procesirati in z njimi naučiti in testirati model v 1,63 sekunde za 496 učnih in 248 testnih primerov.

Zaradi dobrih rezultatov in hitrosti smo se odločili, da ta model uporabimo za klasifikacijo oseb v našem programu.

Zorni kot	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°
Klas. točnost (%)	100,0	99,60	99,60	97,98	98,79	97,58	97,58	97,98	98,79	98,79	99,60

Table 1: Rezultati klasifikacije GEI slik z logistično regresijo za vsak zorni kot.

5 Program

K algoritmu je dodan tudi programski vmesnik, ki uporabniku omogoča dodajanje identifikacijo oseb in dodajanje novih oseb v bazo. Za segmentacijo uporabljamo knjižnjico cv2, ki za delovanje potrebuje sprva le sliko ozadja. Tako uporabniku omogočimo tudi, da z gumbom ponastavi sliko ozadja. S tem se izognemo težavam, ki se pojavijo, če je ob zagonu programa v vidnem polju že oseba, ki je nato zaznana kot ozadje, kar pokvari vse naslednje meritve. Program ima dva okna, ki prikazujeta trenutno stanje segmentacije in klasifikacije. Eno okno prikazuje posnetek kamere, kjer je zaznana oseba obdana s kvadratom. Če je

bila le-ta oseba tudi klasificirana, se na obrobi kvadrata izpiše tudi ime indetificirane osebe. Drugo okno prikazuje rezultate segmentacije. V tem oknu lahko tudi vidimo, če je bilo ozadje nepravilno zgenerirano.

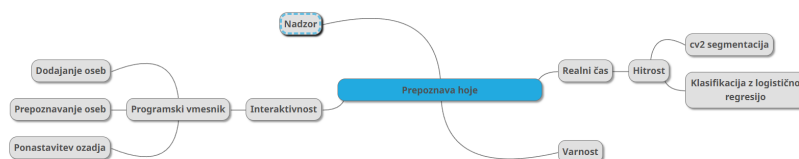


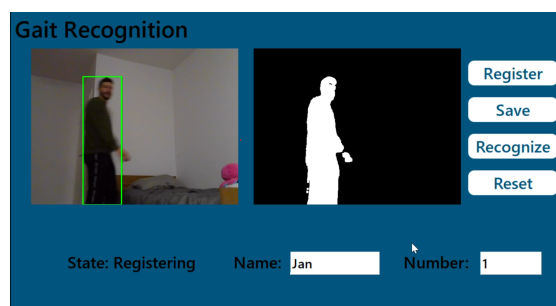
Fig. 4: Miselni vzorec programa

Primer delovanja programa lahko vidimo na sliki 5. Na podsliki a) je predstavljeno zaznavanje osebe, kjer program določeno število slik silhuet (v našem primeru 20 slik) pretvori v GEI sliko in jo shrani. S klikom na gumb "Register" se začne zaznavanje nove osebe. Po posnetku hoje lahko nato zgenerirano sliko shranimo v bazo s klikom na gumb save.

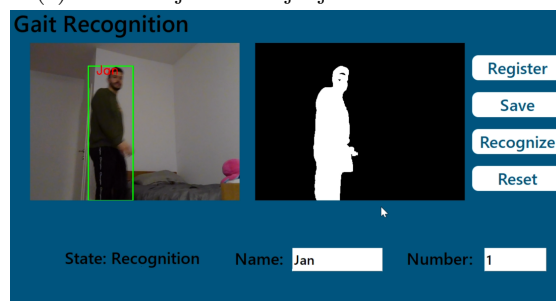
S klikom na gumb "Recognize" nato sprožimo prepoznavanje oseb. Ko je na posnetku zaznana se tudi v levem oknu doda ime prepoznane osebe, kot lahko to vidimo pa podsliki b).

6 Zaključek

V seminarski nalogi smo uspešno implementirali program za razpoznavo oseb po hoji. Učinkovitost programa je odvisna od kvalitete segmentacije in kvalitete klasifikacije. Kot smo videli v poglavju 4, je naša metoda za klasifikacijo glede na GEI sliko zelo učinkovita. Sama metoda za segmentacijo pa prinaša slabše rezultate. Vsaka sprememba v ozadju (npr. dodaten objekt na sliki, sprememba svetlobe, sence, ...) je lahko zaznana kot obris, kar povzroči, da je zgenerirana GEI slika popačena. To posledično tudi poslabša rezultate segmentacije. Vendar pa je sama zaznava človeškega obrisa v realnem času težek problem in tako tudi še ne obstajajo modeli, ki bi lahko z zelo veliko natančnostjo izločili človeško podobo iz ozadja v realnem času. Naš program bi lahko tako izboljšali predvsem z uporabo boljšega modela za segmentacijo. S tem bi bile zgenerirane GEI slike veliko natančnejše, kar bi prineslo boljše rezultate pri klasifikaciji.



(a) Zaznavanje in dodajanje nove osebe v bazo



(b) Prepoznavna osebe

Fig. 5: Primer delovanja programa - dodajanje nove osebe in nato klasifikacija

References

1. CASIA-B dataset. url=<http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>. Dostopano 14.01.2023.
2. Hanqing Chao, Kun Wang, Yiwei He, Junping Zhang, and Jianfeng Feng. Gaitset: Cross-view gait recognition through utilizing gait as a deep set. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2021.
3. Xiaohu Huang, Duowang Zhu, Hao Wang, Xinggang Wang, Bo Yang, Botao He, Wenyu Liu, and Bin Feng. Context-sensitive temporal feature learning for gait recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12909–12918, 2021.
4. Martin Korytowski. *Analiza glavnih komponent: delo diplomskega seminarja*. PhD thesis, University of Ljubljana, 2017.
5. nikitaomare. Gait-recognition. url=<https://github.com/nikitaomare/Gait-Recognition>, 2019. Dostopano 12.01.2023.
6. Chunfeng Song. Gait demo for tutorial v1.0. url=<https://github.com/developfeng/GaitRecognition>, 2021. Dostopano 14.01.2023.
7. Torben Teepe, Johannes Gilg, Fabian Herzog, Stefan Hörmann, and Gerhard Rigoll. Towards a deeper understanding of skeleton-based gait recognition, 2022.