SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Vito Petrinjak

Luka Pošta

VERIFIKACIJA STATIČKOG POTPISA KORIŠTENJEM KLASIČNOG STROJNOG UČENJA

PROJEKTNI ZADATAK

Varaždin, 2026.

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Vito Petrinjak

Matični broj: 0016155011

Studij: Organizacija poslovnih sustava

Luka Pošta

Matični broj: 0016154500

Studij: Organizacija poslovnih sustava

VERIFIKACIJA STATIČKOG POTPISA KORIŠTENJEM KLASIČNOG STROJNOG UČENJA

PROJEKTNI ZADATAK

Mentorica:

Izv. prof. dr. sc. Petra Grd

Varaždin, siječanj 2026.

Vito Petrinjak

Luka Pošta

Izjava o izvornosti

Izjavljujem da je moj završni rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor/Autorica potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi

Sadržaj

1. UVOD

Potpis predstavlja jedan od najdugovječnijih i najraširenijih načina potvrde identiteta pojedinca. Njegova uporaba duboko je ukorijenjena u pravnim, administrativnim i financijskim procesima te se i danas smatra društveno i pravno prihvaćenim oblikom autentikacije. Unatoč razvoju suvremenih digitalnih sigurnosnih mehanizama, poput pametnih kartica, lozinki i različitih biometrijskih senzora, potpis i dalje zadržava važnu ulogu u svakodnevnim postupcima identifikacije i autorizacije.

Razvojem računalnih sustava i sve većom digitalizacijom dokumenata pojavila se potreba za automatiziranom analizom i provjerom autentičnosti potpisa. Tradicionalna, ručna forenzička analiza potpisa zahtijeva visoku razinu stručnosti, vremenski je zahtjevna i podložna subjektivnoj procjeni vještaka. Takav pristup otežava primjenu u suvremenim informacijskim sustavima koji zahtijevaju brzu, skalabilnu i ponovljivu obradu podataka. Zbog toga se sve veća pažnja posvećuje razvoju automatiziranih sustava za verifikaciju potpisa temeljenih na metodama obrade slike i strojnog učenja.

U kontekstu biometrijskih sustava, potpis se klasificira kao bihevioralna biometrijska značajka. Za razliku od fizioloških karakteristika, poput otiska prsta ili šarenice oka, potpis pokazuje izraženu varijabilnost čak i kod iste osobe. Ta varijabilnost proizlazi iz prirode procesa potpisivanja, koji ovisi o motoričkim navikama, psihičkom stanju i vanjskim uvjetima. Iako takva svojstva čine verifikaciju potpisa tehnički zahtjevnim problemom, upravo ta kombinacija društvene prihvaćenosti i složenosti čini potpis zanimljivim i relevantnim istraživačkim područjem unutar strojnog učenja i biometrije.

Ovaj rad usmjeren je na problem verifikacije statičkog, odnosno offline potpisa. Kod statičkog potpisa sustavu je dostupna isključivo vizualna informacija o konačnom izgledu potpisa, bez vremenskih i dinamičkih podataka koji su prisutni kod online potpisa. Takvo ograničenje dodatno povećava složenost problema jer sustav mora razlikovati originalne potpise i krivotvorine na temelju ograničenog skupa informacija. Unatoč tim izazovima, statički potpis ima veliku praktičnu vrijednost zbog svoje široke primjene u stvarnim dokumentima i postojećim arhivama.

Rad objedinjuje teorijsku analizu i praktičnu implementaciju sustava za verifikaciju statičkog potpisa. Teorijski dio rada usmjeren je na sustavno objašnjenje osnovnih pojmova, izazova i metoda koje se koriste u ovom području, uključujući biometrijske sustave, baze potpisa, predprocesiranje slika, ekstrakciju značajki i metode strojnog učenja. Takav teorijski

okvir predstavlja podlogu za praktični dio rada, u kojem se implementira i evaluira konkretan sustav za verifikaciju statičkog potpisa u realističnom scenariju uporabe.

1.1. Motivacija i kontekst problema

Motivacija za istraživanje verifikacije statičkog potpisa proizlazi iz činjenice da potpis i dalje predstavlja jedan od najčešće korištenih oblika potvrde identiteta u svakodnevnim administrativnim, financijskim i pravnim postupcima. Unatoč razvoju naprednih digitalnih metoda autentikacije, potpis je zadržao svoju važnost zbog visoke razine društvene i pravne prihvaćenosti te jednostavnosti primjene (Hafemann, Sabourin i Oliveira, 2017; Bouamra, 2022).

U praksi se provjera potpisa još uvijek često provodi ručno, oslanjajući se na forenzičku analizu i stručnu procjenu vještaka. Takav pristup zahtijeva značajno vrijeme, visoku razinu stručnosti i podložan je subjektivnosti, što može rezultirati nedosljednim ili teško ponovljivim odlukama. S porastom količine digitaliziranih dokumenata i potrebom za brzom obradom podataka, ručna analiza potpisa postaje neprikladna za suvremene informacijske sustave (Singh et al., 2022).

Automatizirani sustavi za verifikaciju potpisa razvijeni su s ciljem rješavanja navedenih ograničenja primjenom metoda obrade slike i strojnog učenja. U takvim sustavima potpis se promatra kao biheviоralna biometrijska značajka, čije su osnovne karakteristike individualnost i varijabilnost. Za razliku od fizioloških biometrijskih obilježja, potpis pokazuje znatne razlike čak i kod iste osobe, što značajno otežava proces verifikacije (Hafemann, Sabourin i Oliveira, 2017).

Poseban izazov u verifikaciji statičkog potpisa predstavlja ograničena količina informacija dostupnih sustavu. Kod offline pristupa raspolaže se isključivo vizualnim prikazom potpisa, bez dinamičkih podataka poput brzine pisanja, pritiska ili redoslijeda poteza. Time je sustav prisiljen donositi odluku na temelju oblika i strukture potpisa, što dodatno povećava osjetljivost na varijabilnost legalnih potpisa i sličnost s vještim krivotvorinama (Singh et al., 2022; Bouamra, 2022).

Dodatni problem predstavlja mali broj dostupnih uzoraka za treniranje modela u realnim scenarijima. U mnogim slučajevima sustav raspolaže s tek nekoliko originalnih potpisa po korisniku, dok su krivotvorine rijetke ili se koriste isključivo u fazi testiranja. Takvi uvjeti zahtijevaju pažljiv odabir pristupa koji su sposobni učinkovito učiti iz ograničenih podataka, pri

čemu se često primjenjuju writer-dependent modeli prilagođeni pojedinom korisniku (Hafemann, Sabourin i Oliveira, 2017; Singh et al., 2022).

U tom kontekstu, verifikacija statičkog potpisa predstavlja složen i aktualan istraživački problem unutar područja biometrije i strojnog učenja. Razvoj pouzdanih sustava zahtijeva kombinaciju prikladnog predprocesiranja, ekstrakcije diskriminativnih značajki i učinkovitih klasifikacijskih algoritama. Razumijevanje navedenih izazova i motivacijskih čimbenika ključno je za daljnju teorijsku analizu i za interpretaciju rezultata praktičnog dijela rada (Bouamra, 2022).

1.2. Biometrijski sustavi i bihevioralne biometrijske karakteristike

Biometrijski sustavi koriste mjerljive karakteristike pojedinca za potrebe identifikacije ili verifikacije identiteta. Takvi sustavi temelje se na pretpostavci da su određena ljudska obilježja dovoljno jedinstvena i stabilna kako bi se mogla koristiti za razlikovanje pojedinaca unutar populacije. U kontekstu sigurnosnih i informacijskih sustava, biometrija predstavlja alternativu tradicionalnim metodama autentikacije temeljenima na znanju ili posjedu, poput lozinki i kartica (Bouamra, 2022).

Biometrijske karakteristike uobičajeno se dijele na fiziološke i bihevioralne. Fiziološke karakteristike odnose se na anatomska obilježja ljudskog tijela, kao što su otisak prsta, šarenica oka ili struktura lica. One se odlikuju visokom postojanošću kroz vrijeme i relativno malom varijabilnošću kod iste osobe. Zbog toga se fiziološka biometrija često smatra pouzdanijom, ali istovremeno zahtijeva specijalizirane senzore i može biti percipirana kao intruzivna (Hafemann, Sabourin i Oliveira, 2017).

Bihevioralne biometrijske karakteristike temelje se na obrascima ponašanja pojedinca tijekom izvođenja određene radnje. U tu skupinu ubrajaju se potpis, rukopis, dinamika tipkanja i način hoda. Za razliku od fizioloških obilježja, bihevioralne karakteristike podložne su većoj varijabilnosti jer ovise o psihičkom i fizičkom stanju osobe, okolinskim uvjetima i kontekstu u kojem se radnja izvodi (Bouamra, 2022). Upravo ta varijabilnost predstavlja jedan od ključnih izazova u razvoju pouzdanih biometrijskih sustava.

Potpis zauzima posebno mjesto među bihevioralnim biometrijskim karakteristikama. Njegova osnovna prednost leži u visokoj razini društvene i pravne prihvaćenosti te

dugogodišnjoj uporabi kao sredstva potvrde identiteta. Istovremeno, potpis posjeduje određenu razinu individualnosti koja proizlazi iz neuromotoričkih procesa uključenih u čin pisanja, što omogućuje njegovu primjenu u biometrijskim sustavima (Hafemann, Sabourin i Oliveira, 2017).

Unatoč navedenim prednostima, bihevioralne biometrijske karakteristike, uključujući potpis, suočavaju se s nizom ograničenja. Visoka intra-class varijabilnost može dovesti do pogrešnog odbacivanja legalnih korisnika, dok sličnost između originalnih potpisa i vještih krivotvorina povećava rizik pogrešnog prihvaćanja neautoriziranih osoba. Zbog toga biometrijski sustavi koji se oslanjaju na bihevioralne karakteristike zahtijevaju pažljivo osmišljene metode obrade podataka i donošenja odluka (Singh et al., 2022).

U kontekstu verifikacije statičkog potpisa, biometrijski sustav mora postići ravnotežu između sigurnosti i upotrebljivosti. Sustav mora biti dovoljno tolerantan na prirodne varijacije potpisa iste osobe, ali istovremeno dovoljno diskriminativan da pouzdano odbaci krivotvorine. Razumijevanje temeljnih obilježja biometrijskih sustava i specifičnosti bihevioralnih karakteristika predstavlja nužnu podlogu za daljnju analizu metoda i pristupa koji se koriste u verifikaciji statičkog potpisa (Hafemann, Sabourin i Oliveira, 2017; Bouamra, 2022).

1.3. Cilj i struktura rada

Cilj ovog rada jest analizirati problem verifikacije statičkog potpisa s teorijskog i praktičnog aspekta te prikazati mogućnosti primjene metoda strojnog učenja u izgradnji pouzdanog biometrijskog sustava za verifikaciju identiteta. Poseban naglasak stavljen je na offline pristup verifikaciji potpisa, u kojem se potpis promatra isključivo kao digitalna slika, bez dostupnih dinamičkih informacija koje su prisutne kod online potpisa.

Teorijski dio rada usmjeren je na sustavno objašnjenje temeljnih koncepata i izazova vezanih uz verifikaciju statičkog potpisa. U tom dijelu obrađuju se osnovna obilježja biometrijskih sustava, potpis kao bihevioralna biometrijska karakteristika, razlike između online i offline pristupa te problematika krivotvorenja potpisa. Nadalje, prikazuju se javno dostupne baze potpisa, metode predprocesiranja slika, postupci ekstrakcije značajki te algoritmi strojnog učenja koji se najčešće koriste u sustavima za verifikaciju potpisa.

Praktični dio rada temelji se na implementaciji cjelovitog sustava za verifikaciju statičkog potpisa u writer-dependent scenariju. Sustav obuhvaća pripremu skupa podataka, predprocesiranje slika potpisa, izlučivanje diskriminativnih značajki iz potpisa, treniranje

klasifikacijskih modela te evaluaciju njihove učinkovitosti primjenom standardnih biometrijskih metrika. Takav pristup odgovara realističnom scenariju uporabe, u kojem je za svakog korisnika dostupan ograničen broj originalnih potpisa, dok se krivotvorine koriste za procjenu performansi sustava.

Struktura rada organizirana je tako da teorijski dio pruža čvrstu podlogu za razumijevanje praktične implementacije. Nakon uvodnog poglavlja, u radu se postupno razrađuju ključni teorijski pojmovi, a zatim se u praktičnom dijelu demonstrira njihova primjena kroz konkretan sustav za verifikaciju potpisa. Na taj se način omogućuje jasno povezivanje teorijskih spoznaja s eksperimentalnim rezultatima i njihova interpretacija u kontekstu postojećih istraživanja.

Na kraju rada daje se evaluacija implementiranog sustava i rasprava o postignutim rezultatima, ograničenjima i mogućim smjerovima budućeg razvoja. Takva struktura omogućuje cjelovit uvid u problem verifikacije statičkog potpisa te naglašava važnost integracije teorijskih znanja i praktičnih rješenja u razvoju biometrijskih sustava.

2. POTPIS KAO BIOMETRIJSKA ZNAČAJKA

Potpis predstavlja specifičnu bihevioralnu biometrijsku značajku koja se temelji na motoričkom ponašanju pojedinca tijekom čina pisanja. Za razliku od fizioloških biometrijskih obilježja, potpis nije statična karakteristika, već rezultat složenih neuromotoričkih procesa koji uključuju koordinaciju mišića, vizualnu percepciju i stečeno iskustvo pisanja. Upravo zbog te složenosti potpis posjeduje određeni stupanj individualnosti koji ga čini pogodnim za biometrijsku verifikaciju.

Jedna od ključnih osobina potpisa jest njegova varijabilnost. Isti potpisnik rijetko kada proizvodi dva potpuno identična potpisa, čak i u sličnim uvjetima. Varijacije mogu nastati zbog promjena u brzini pisanja, emocionalnom stanju, umoru ili vanjskim uvjetima u kojima se potpis daje. Sustavi za verifikaciju potpisa moraju uzeti u obzir takvu prirodnu varijabilnost, a istovremeno razlikovati legitimne potpise od krivotvorina.

Potpis se u biometrijskim sustavima koristi prvenstveno za verifikaciju identiteta, pri čemu sustav provjerava odgovara li dani potpis deklariranom identitetu korisnika. Takav pristup razlikuje se od identifikacijskih sustava koji nastoje odrediti identitet potpisnika među većim brojem mogućih korisnika. Zbog svoje društvene i pravne prihvaćenosti, potpis se i dalje smatra relevantnim oblikom autentikacije, osobito u kombinaciji s automatiziranim metodama analize.

2.1. Povijesni i pravni aspekt potpis

Povijest potpisa kao sredstva potvrde identiteta seže daleko u prošlost, još u razdoblja kada su se dokumenti ručno izrađivali i ovjeravali. Kroz povijest je potpis služio kao dokaz autorstva, suglasnosti i odgovornosti pojedinca za sadržaj dokumenta. Njegova uloga bila je posebno izražena u pravnim i administrativnim kontekstima, gdje je potpis imao snagu osobne izjave volje.

S razvojem pravnih sustava, potpis je postupno dobio formalni status kao pravno obvezujući element dokumenta. U mnogim zakonodavstvima potpis se i danas smatra ključnim dokazom identiteta potpisnika i njegove namjere da prihvati sadržaj dokumenta. Takav status potpisa pridonio je njegovoj dugotrajnoj primjeni i širokoj društvenoj prihvaćenosti, unatoč pojavi novih tehnologija autentikacije (Bouamra, 2022).

U forenzičkom kontekstu, analiza potpisa tradicionalno se provodi ručno, pri čemu se procjenjuju oblik, proporcije, nagib i drugi vizualni elementi potpisa. Ovakav pristup zahtijeva specijalizirano znanje i iskustvo, a rezultati analize mogu varirati ovisno o vještini i subjektivnoj procjeni stručnjaka. Upravo ta ograničenja forenzičke analize predstavljaju jedan od razloga za razvoj automatiziranih sustava za verifikaciju potpisa (Bouamra, 2022).

Pojavom digitalnih dokumenata i elektroničke obrade podataka, tradicionalni koncept potpisa proširen je na elektroničke i digitalne potpise. Iako se takvi oblici potpisa razlikuju u tehničkoj implementaciji, osnovna ideja ostaje ista: potvrda identiteta i izražavanje suglasnosti. U tom kontekstu, automatizirana verifikacija rukom pisanog potpisa predstavlja pokušaj zadržavanja poznatog i pravno prihvaćenog oblika autentikacije uz primjenu suvremenih računalnih metoda.

Razumijevanje povijesne i pravne uloge potpisa važno je za sagledavanje njegove primjene u biometrijskim sustavima. Takav kontekst objašnjava zašto potpis, unatoč svojim ograničenjima i varijabilnosti, i dalje predstavlja relevantan predmet istraživanja u području automatizirane verifikacije identiteta (Bouamra, 2022).

2.2. Karakteristike rukom pisanog potpisa

Rukom pisani potpis predstavlja rezultat složenog neuromotoričkog procesa koji uključuje koordinaciju mišićnih pokreta, vizualnu kontrolu i dugotrajno stečene obrasce pisanja. Tijekom vremena, potpis se oblikuje kao individualni izraz koji odražava osobne navike i motoričke karakteristike potpisnika. Upravo ta individualnost čini potpis pogodnim za primjenu u biometrijskim sustavima (Hafemann, Sabourin i Oliveira, 2017).

Jedna od temeljnih karakteristika rukom pisanog potpisa jest njegova varijabilnost. Čak i kada ista osoba potpisuje dokumente u sličnim uvjetima, potpisi se nikada ne podudaraju u potpunosti. Razlike se mogu očitovati u veličini, nagibu, proporcijama pojedinih dijelova potpisa ili kontinuitetu poteza. Takva intra-class varijabilnost predstavlja ključni izazov za sustave verifikacije, jer sustav mora prepoznati legitimne varijacije potpisa iste osobe bez pogrešnog odbacivanja (Hafemann, Sabourin i Oliveira, 2015).

Na izgled i strukturu potpisa dodatno utječu vanjski i unutarnji čimbenici. Psihičko stanje potpisnika, poput stresa ili umora, može promijeniti brzinu i stabilnost pisanja, dok fizički čimbenici, poput ozljeda ili starosti, mogu utjecati na preciznost i oblik poteza. Također, uvjeti

potpisivanja, uključujući podlogu, sredstvo za pisanje i položaj tijela, mogu uzrokovati dodatne varijacije u izgledu potpisa (Bouamra, 2022).

Osim varijabilnosti, rukom pisani potpis karakterizira i određena razina stabilnosti na globalnoj razini. Iako se pojedini detalji mogu razlikovati, osnovna struktura potpisa, raspored poteza i opći oblik često ostaju relativno konzistentni kroz vrijeme. Sustavi za verifikaciju potpisa nastoje iskoristiti upravo tu ravnotežu između varijabilnosti i stabilnosti, fokusirajući se na značajke koje su dovoljno postojane da razlikuju potpisnike, a istovremeno tolerantne na prirodne promjene potpisa (Hafemann, Sabourin i Oliveira, 2017).

Poseban problem u analizi rukom pisanog potpisa predstavljaju vješte krivotvorine. Takve krivotvorine nastaju kada krivotvoritelj svjesno pokušava oponašati oblik i strukturu originalnog potpisa, često na temelju višestrukih primjera. Vizualna sličnost između originalnog potpisa i vješte krivotvorine može biti vrlo visoka, što dodatno otežava pouzdanu verifikaciju, osobito u offline scenariju gdje nisu dostupne dinamičke informacije o procesu pisanja (Hafemann, Sabourin i Oliveira, 2015; Bouamra, 2022).

Razumijevanje navedenih karakteristika rukom pisanog potpisa ključno je za razvoj učinkovitih sustava verifikacije. Prirodna varijabilnost, utjecaj vanjskih čimbenika i prisutnost krivotvorina zahtijevaju pažljiv odabir metoda predprocesiranja, ekstrakcije značajki i klasifikacije. Ove karakteristike čine temelj za daljnju analizu tipova potpisa i razlika između statičkog i dinamičkog pristupa, koji se razmatraju u sljedećim poglavljima (Hafemann, Sabourin i Oliveira, 2017).

2.3. Vrste krivotvorina potpisa

Krivotvorenje potpisa predstavlja jedan od ključnih problema u sustavima za verifikaciju potpisa jer izravno utječe na sigurnost i pouzdanost sustava. Cilj krivotvoritelja je proizvesti potpis koji će biti dovoljno sličan originalnom potpisu kako bi sustav ili ljudski promatrač pogrešno prihvatio njegovu autentičnost. U literaturi se krivotvorine potpisa najčešće klasificiraju prema razini vještine i namjeri krivotvoritelja (Hafemann, Sabourin i Oliveira, 2017).

Najjednostavniji oblik krivotvorenja su nasumične krivotvorine. U ovom slučaju krivotvoritelj nema nikakvo predznanje o izgledu originalnog potpisa te ispisuje vlastiti potpis ili proizvoljni uzorak, navodeći identitet druge osobe. Takve krivotvorine obično se znatno razlikuju od originalnog potpisa i relativno ih je lako detektirati automatiziranim sustavima, čak i uz jednostavne metode analize (Singh et al., 2022).

Sljedeću skupinu čine jednostavne, odnosno neuvježbane krivotvorine. Kod ovog tipa krivotvoritelj ima uvid u izgled originalnog potpisa, ali ne ulaže značajan napor u njegovo uvježbavanje. Potpis se najčešće pokušava oponašati površno, bez precizne kontrole proporcija, ritma i strukture poteza. Iako su takve krivotvorine sličnije originalu u odnosu na nasumične krivotvorine, i dalje pokazuju znatna odstupanja koja se mogu otkriti analizom oblika i globalnih značajki potpisa (Hafemann, Sabourin i Oliveira, 2017).

Najzahtjevniji oblik krivotvorenja predstavljaju vješte krivotvorine. One nastaju kada krivotvoritelj sustavno vježba oponašanje originalnog potpisa, često na temelju više dostupnih uzoraka. Cilj je reproducirati ne samo opći oblik potpisa, već i fine detalje, poput zakrivljenosti poteza, međusobnih odnosa dijelova potpisa i ukupne strukture. Vizualna sličnost između vješte krivotvorine i originalnog potpisa može biti vrlo visoka, što značajno otežava pouzdanu detekciju, osobito u offline sustavima koji ne raspolažu dinamičkim informacijama o procesu pisanja (Bouamra, 2022).

U kontekstu automatizirane verifikacije statičkog potpisa, upravo vješte krivotvorine predstavljaju najveći izazov. Sustav mora biti dovoljno osjetljiv da prepozna suptilne razlike između originala i krivotvorine, a istovremeno tolerantan na prirodne varijacije legalnih potpisa. Zbog toga se u evaluaciji sustava za verifikaciju potpisa posebna pažnja posvećuje testiranju na vještim krivotvorinama, jer one najrealističnije simuliraju napade u stvarnim uvjetima uporabe (Singh et al., 2022).

Razlikovanje različitih vrsta krivotvorina važno je za pravilno vrednovanje performansi sustava za verifikaciju potpisa. Sustav koji uspješno odbacuje nasumične i neuvježbane krivotvorine, ali ne prepoznaje vješte krivotvorine, ne može se smatrati dovoljno sigurnim za praktičnu primjenu. Stoga razumijevanje tipova krivotvorina predstavlja ključnu podlogu za analizu metoda verifikacije i interpretaciju rezultata dobivenih u praktičnom dijelu rada (Hafemann, Sabourin i Oliveira, 2017).

3. ONLINE I OFFLINE VERIFIKACIJA POTPISA

Verifikacija potpisa može se provoditi primjenom dvaju temeljnih pristupa, ovisno o vrsti dostupnih podataka o potpisu. Ti pristupi razlikuju se prema načinu prikupljanja potpisa i vrsti informacija koje su dostupne sustavu za analizu. U literaturi se najčešće razlikuju online, odnosno dinamička verifikacija potpisa, i offline, odnosno statička verifikacija potpisa (Hafemann, Sabourin i Oliveira, 2017).

Osnovna razlika između ova dva pristupa leži u dostupnosti vremenskih i dinamičkih informacija. Online verifikacija temelji se na praćenju procesa potpisivanja u stvarnom vremenu, dok se offline verifikacija oslanja isključivo na konačni vizualni prikaz potpisa. Zbog toga se ova dva pristupa razlikuju po složenosti obrade podataka, vrsti značajki koje se mogu izlučiti te razini pouzdanosti sustava.

Unatoč zajedničkom cilju, online i offline sustavi imaju različita područja primjene. Online pristup se češće koristi u kontroliranim okruženjima s namjenskim uređajima za prikupljanje potpisa, dok offline pristup omogućuje analizu potpisa na papirnatim dokumentima i digitaliziranim zapisima. Razumijevanje razlika između ovih pristupa ključno je za pravilno sagledavanje prednosti i ograničenja sustava za verifikaciju potpisa.

3.1. Online (dinamički) potpis

Online, odnosno dinamički potpis, prikuplja se korištenjem uređaja koji omogućuju praćenje procesa potpisivanja u stvarnom vremenu. Takvi uređaji uključuju grafičke tablete, digitalne olovke, pametne telefone ili specijalizirane potpisne plohe. Tijekom potpisivanja sustav bilježi niz vremenski ovisnih podataka koji opisuju dinamiku pokreta potpisnika (Hafemann, Sabourin i Oliveira, 2017).

Najčešće prikupljane dinamičke informacije uključuju prostorne koordinate položaja olovke, brzinu i ubrzanje pokreta, pritisak olovke na podlogu te podatke o nagibu i azimutu olovke. Ove informacije omogućuju detaljan uvid u način na koji je potpis izveden, a ne samo u njegov konačni izgled. Zbog toga online sustavi raspolažu bogatijim skupom podataka u usporedbi s offline pristupom (Hafemann, Sabourin i Oliveira, 2015).

Prednost online verifikacije potpisa očituje se u većoj diskriminativnoj moći sustava. Dinamičke značajke, poput ritma pisanja ili promjena brzine, teško je precizno oponašati, čak

i kod vještih krivotvoritelja. Zbog toga online sustavi često postižu bolje rezultate u razlikovanju originalnih potpisa i krivotvorina u odnosu na offline sustave (Singh et al., 2022).

Unatoč navedenim prednostima, online verifikacija potpisa ima i određena ograničenja. Primjena takvih sustava zahtijeva dostupnost odgovarajuće opreme i kontrolirane uvjete potpisivanja. To može ograničiti njihovu primjenu u stvarnim scenarijima, osobito u situacijama gdje se potpisi već nalaze na papirnatim dokumentima ili su prikupljeni bez namjenskih uređaja. Zbog toga se, unatoč tehničkim prednostima, online pristup ne može uvijek primijeniti u praksi (Bouamra, 2022).

Razumijevanje načina rada online sustava i vrsta podataka kojima raspolažu važno je za usporedbu s offline pristupom. Takva usporedba omogućuje jasnije sagledavanje razloga zbog kojih se u ovom radu fokus stavlja na verifikaciju statičkog potpisa, koja se razmatra u sljedećem potpoglavlju.

3.2. Offline (statički) potpis

Offline, odnosno statički potpis, predstavlja potpis koji je dostupan sustavu isključivo u obliku slike, najčešće dobivene skeniranjem papirnatog dokumenta ili fotografiranjem potpisa. Za razliku od online potpisa, kod offline pristupa ne postoje informacije o procesu potpisivanja, već se analiza temelji isključivo na vizualnom prikazu konačnog rezultata pisanja (Hafemann, Sabourin i Oliveira, 2017).

Zbog nedostatka dinamičkih informacija, offline verifikacija potpisa suočava se s većim izazovima u usporedbi s online pristupom. Sustav mora donositi odluku na temelju oblika, strukture i rasporeda poteza, pri čemu su prirodne varijacije potpisa iste osobe često izraženije. Ovakva ograničenja čine offline verifikaciju osjetljivijom na pogreške, osobito u slučaju vještih krivotvorina koje mogu vizualno snažno nalikovati originalnim potpisima (Singh et al., 2022).

Unatoč navedenim izazovima, offline potpis ima značajnu praktičnu vrijednost. Velik broj stvarnih dokumenata, poput ugovora, obrazaca i arhivskih zapisa, sadrži rukom pisane potpise u papirnatom obliku. Automatizirana analiza takvih potpisa omogućuje primjenu verifikacijskih sustava i u okruženjima gdje online prikupljanje potpisa nije moguće ili nije provedeno u trenutku potpisivanja (Bouamra, 2022).

U offline sustavima posebna se pažnja posvećuje postupcima predprocesiranja slike, čija je svrha smanjiti utjecaj šuma, varijacija u osvjetljenju i razlika u rezoluciji. Također, ključnu

ulogu ima odabir značajki koje su dovoljno diskriminativne da razlikuju potpisnike, ali i tolerantne na prirodne promjene potpisa. Zbog toga se u offline verifikaciji često koriste značajke temeljene na obliku i teksturi potpisa (Hafemann, Sabourin i Oliveira, 2017).

Offline verifikacija potpisa najčešće se primjenjuje u writer-dependent scenariju, u kojem se za svakog korisnika trenira zaseban model na temelju njegovih originalnih potpisa. Takav pristup omogućuje prilagodbu modela specifičnim karakteristikama potpisa pojedinca i često daje bolje rezultate u uvjetima ograničenog broja uzoraka za treniranje (Singh et al., 2022).

Razumijevanje specifičnosti offline potpisa i ograničenja koja proizlaze iz statičke prirode podataka ključno je za pravilnu interpretaciju rezultata sustava za verifikaciju potpisa. U sljedećem potpoglavlju daje se usporedba online i offline pristupa s ciljem jasnijeg isticanja njihovih prednosti i nedostataka.

3.3. Usporedba online i offline pristupa

Online i offline verifikacija potpisa razlikuju se prvenstveno prema vrsti informacija koje su dostupne sustavu za analizu. Kod online pristupa bilježe se dinamički podaci o procesu potpisivanja, dok se kod offline pristupa analiza temelji isključivo na vizualnom prikazu potpisa. Ova razlika ima značajan utjecaj na složenost problema, odabir značajki i ukupnu pouzdanost verifikacijskog sustava (Hafemann, Sabourin i Oliveira, 2017).

Prednost online verifikacije potpisa leži u bogatstvu dostupnih informacija. Dinamičke značajke, poput brzine pisanja, promjena pritiska i ritma poteza, omogućuju precizniju karakterizaciju načina potpisivanja i otežavaju uspješno krivotvorenje. Zbog toga online sustavi u pravilu postižu niže stope pogrešaka u usporedbi s offline sustavima, osobito u scenarijima s vještim krivotvorinama (Singh et al., 2022).

S druge strane, offline verifikacija potpisa nudi veću fleksibilnost u pogledu primjene. Budući da ne zahtijeva posebnu opremu za prikupljanje potpisa, može se koristiti za analizu postojećih papirnatih dokumenata i digitaliziranih arhiva. To offline pristup čini praktičnijim u mnogim realnim situacijama, unatoč njegovim tehničkim ograničenjima i većoj osjetljivosti na pogreške (Bouamra, 2022).

Razlike između ova dva pristupa očituju se i u metodama obrade podataka. Online sustavi često koriste vremenski ovisne značajke i algoritme prilagođene sekvencijalnim

podacima, dok offline sustavi primjenjuju tehnike obrade slike i ekstrakcije značajki temeljenih na obliku i teksturi potpisa. Posljedično, i pristupi klasifikaciji mogu se razlikovati, iako se određeni algoritmi strojnog učenja mogu primijeniti u oba slučaja (Hafemann, Sabourin i Oliveira, 2017).

U praksi, izbor između online i offline pristupa ovisi o specifičnim zahtjevima sustava i dostupnim resursima. Dok online verifikacija nudi veću razinu sigurnosti, offline verifikacija omogućuje širu primjenu i lakšu integraciju u postojeće procese. Zbog toga offline pristup i dalje predstavlja aktivno istraživačko područje, s ciljem smanjenja razlike u pouzdanosti u odnosu na online sustave (Singh et al., 2022).

U kontekstu ovog rada, fokus je stavljen na offline verifikaciju potpisa zbog njezine praktične relevantnosti i mogućnosti primjene na stvarnim dokumentima. Razumijevanje razlika između online i offline pristupa omogućuje jasnije sagledavanje ograničenja i potencijala implementiranog sustava za verifikaciju statičkog potpisa.

4. PROBLEM VERIFIKACIJE STATIČKOG POTPISA

Problem verifikacije statičkog potpisa odnosi se na donošenje odluke o tome pripada li promatrani potpis deklariranom identitetu korisnika. Za razliku od klasičnih problema klasifikacije, u kojima je cilj razvrstati uzorke u unaprijed definirane klase, verifikacija potpisa predstavlja specifičan oblik biometrijskog problema u kojem se provjerava tvrdnja o identitetu potpisnika. Takav problem postavlja posebne zahtjeve na način obrade podataka i evaluacije sustava (Hafemann, Sabourin i Oliveira, 2017).

U kontekstu statičkog potpisa, složenost problema dodatno se povećava zbog ograničenih informacija dostupnih sustavu. Analiza se temelji isključivo na vizualnom prikazu potpisa, bez mogućnosti korištenja dinamičkih značajki koje bi mogle olakšati razlikovanje originalnih potpisa i krivotvorina. Zbog toga sustav mora pronaći ravnotežu između tolerancije na prirodne varijacije potpisa iste osobe i osjetljivosti na razlike koje upućuju na krivotvorenje.

Posebnost problema verifikacije potpisa očituje se i u načinu na koji se sustav trenira i koristi. U praksi se često raspolaže s ograničenim brojem originalnih potpisa po korisniku, dok krivotvorine nisu uvijek dostupne u fazi treniranja. Ovakvi uvjeti zahtijevaju prilagodbu klasičnih metoda strojnog učenja kako bi se postigla zadovoljavajuća razina pouzdanosti sustava (Singh et al., 2022).

4.1. Verifikacija naspram identifikacije

U biometrijskim sustavima razlikuju se dva osnovna zadatka: verifikacija i identifikacija. Verifikacija podrazumijeva provjeru tvrdnje o identitetu korisnika, pri čemu sustav odgovara na pitanje odgovara li dani biometrijski uzorak deklariranom identitetu. Suprotno tome, identifikacija ima za cilj odrediti identitet korisnika među većim brojem mogućih identiteta pohranjenih u sustavu (Hafemann, Sabourin i Oliveira, 2017).

U slučaju verifikacije potpisa, korisnik navodi svoj identitet, a sustav uspoređuje dostavljeni potpis s referentnim potpisima te osobe. Odluka se donosi u obliku binarnog ishoda, pri čemu se potpis prihvaća kao originalan ili odbacuje kao krivotvoren. Ovakav pristup odgovara mnogim stvarnim scenarijima uporabe potpisa, poput potpisivanja ugovora ili autorizacije dokumenata.

Identifikacija potpisa predstavlja složeniji problem jer sustav mora usporediti potpis s potpisima svih korisnika pohranjenih u bazi podataka. Takav pristup zahtijeva veću količinu podataka, veće računalne resurse i složenije algoritme. Zbog izražene varijabilnosti potpisa i sličnosti između potpisa različitih osoba, identifikacija potpisa često je manje pouzdana u odnosu na verifikaciju (Singh et al., 2022).

Zbog navedenih razloga, većina sustava za verifikaciju statičkog potpisa fokusira se upravo na verifikacijski scenarij. Takav pristup omogućuje jednostavniju implementaciju, bolju prilagodbu individualnim karakteristikama potpisa i jasniju interpretaciju rezultata. U kontekstu ovog rada, problem verifikacije odabran je kao primarni istraživački zadatak jer odgovara realnim uvjetima primjene i omogućuje učinkovitiju evaluaciju performansi sustava.

4.2. Writer-dependent i writer-independent sustavi

U sustavima za verifikaciju potpisa razlikuju se dva osnovna pristupa s obzirom na način treniranja modela i odnos prema korisnicima sustava: writer-dependent i writer-independent pristup. Ova podjela temelji se na pretpostavci o tome u kojoj mjeri su modeli prilagođeni pojedinim korisnicima te kako se generaliziraju na nove potpise (Hafemann, Sabourin i Oliveira, 2017).

Writer-dependent pristup podrazumijeva izgradnju zasebnog modela za svakog korisnika sustava. Model se trenira isključivo na potpisima jedne osobe, pri čemu se uči razlikovati originalne potpise tog korisnika od krivotvorina. Takav pristup omogućuje detaljno prilagođavanje modela specifičnim karakteristikama potpisa pojedinca i često postiže bolje rezultate u uvjetima ograničenog broja uzoraka za treniranje (Singh et al., 2022).

Prednost writer-dependent sustava očituje se u njihovoj sposobnosti da se nose s velikom intra-class varijabilnošću potpisa. Budući da je model treniran isključivo na potpisima jedne osobe, lakše prepoznaje legitimne varijacije potpisa i smanjuje vjerojatnost pogrešnog odbacivanja originalnih potpisa. Međutim, ovakav pristup zahtijeva treniranje i održavanje većeg broja modela, što može povećati složenost sustava u scenarijima s velikim brojem korisnika (Zois et al., 2019).

Writer-independent pristup temelji se na izgradnji jedinstvenog modela koji se trenira na potpisima većeg broja korisnika. Cilj takvog modela je naučiti opće obrasce koji razlikuju originalne potpise od krivotvorina, neovisno o identitetu potpisnika. Ovakav pristup omogućuje jednostavniju implementaciju i lakšu skalabilnost sustava, ali često zahtijeva veću količinu

podataka za treniranje kako bi se postigla zadovoljavajuća razina pouzdanosti (Hafemann, Sabourin i Oliveira, 2017).

U praksi, writer-independent sustavi mogu pokazivati slabije performanse u scenarijima s ograničenim brojem uzoraka po korisniku, jer im nedostaje specifična prilagodba individualnim karakteristikama potpisa. Zbog toga se u mnogim istraživanjima i praktičnim primjenama preferira writer-dependent pristup, osobito u offline verifikaciji potpisa gdje su podaci često ograničeni (Singh et al., 2022; Zois et al., 2019).

U kontekstu ovog rada, odabran je writer-dependent pristup jer omogućuje realističnu simulaciju uvjeta u kojima je za svakog korisnika dostupan mali broj originalnih potpisa. Takav pristup pruža jasniji uvid u ponašanje sustava i olakšava interpretaciju rezultata evaluacije u praktičnom dijelu rada.

4.3. Glavni izazovi verifikacije statičkog potpisa

Verifikacija statičkog potpisa predstavlja složen problem zbog kombinacije više međusobno povezanih čimbenika koji utječu na pouzdanost sustava. Za razliku od nekih drugih biometrijskih karakteristika, potpis pokazuje izraženu varijabilnost i osjetljivost na vanjske utjecaje, što otežava precizno razlikovanje originalnih potpisa i krivotvorina (Hafemann, Sabourin i Oliveira, 2017).

Jedan od glavnih izazova je velika intra-class varijabilnost potpisa iste osobe. Potpisi se mogu razlikovati u veličini, nagibu, proporcijama i kontinuitetu poteza, čak i kada ih proizvodi isti potpisnik u kratkom vremenskom razmaku. Sustav za verifikaciju mora biti dovoljno tolerantan da prihvati takve legitimne varijacije, a istovremeno dovoljno osjetljiv da odbaci neautorizirane potpise (Singh et al., 2022).

Drugi značajan izazov predstavlja mali broj dostupnih uzoraka za treniranje. U realnim uvjetima sustav često raspolaže s ograničenim brojem originalnih potpisa po korisniku, dok su krivotvorine rijetko dostupne ili se koriste isključivo u fazi evaluacije. Ovakva ograničenja otežavaju učenje robusnih modela i zahtijevaju pažljiv odabir metoda koje mogu učinkovito raditi s malim skupovima podataka (Zois et al., 2019).

Posebno zahtjevan problem predstavljaju vješte krivotvorine. Takve krivotvorine mogu vizualno vrlo snažno nalikovati originalnim potpisima, jer krivotvoritelj nastoji oponašati osnovni oblik i strukturu potpisa. U offline verifikaciji, gdje nisu dostupne dinamičke informacije o

procesu pisanja, razlikovanje originala i vještih krivotvorina postaje izrazito zahtjevno (Hafemann, Sabourin i Oliveira, 2017).

Dodatni izazovi proizlaze iz kvalitete ulaznih podataka. Razlike u rezoluciji skeniranih dokumenata, prisutnost šuma, neujednačeno osvjetljenje i varijacije u kontrastu mogu negativno utjecati na performanse sustava. Zbog toga je u offline verifikaciji potpisa nužno provoditi odgovarajuće postupke predprocesiranja kako bi se smanjio utjecaj takvih čimbenika (Singh et al., 2022).

Svi navedeni izazovi čine verifikaciju statičkog potpisa zahtjevnim istraživačkim problemom koji zahtijeva pažljivo osmišljene metode obrade slike i strojnog učenja. Razumijevanje tih izazova ključno je za pravilno vrednovanje rezultata sustava i za odabir prikladnih pristupa u praktičnoj implementaciji, što se razmatra u nastavku rada (Hafemann, Sabourin i Oliveira, 2017).

5. BAZE PODATAKA POTPISA

Razvoj i evaluacija sustava za verifikaciju potpisa uvelike ovise o dostupnosti kvalitetnih i reprezentativnih baza podataka. Baze potpisa omogućuju treniranje modela, usporedbu različitih metoda te objektivnu procjenu performansi sustava u kontroliranim uvjetima. Bez standardiziranih skupova podataka bilo bi teško provoditi pouzdanu evaluaciju i uspoređivati rezultate različitih istraživanja (Hafemann, Sabourin i Oliveira, 2017).

Baze podataka potpisa obično sadrže skup originalnih potpisa po korisniku te odgovarajuće krivotvorine različite razine složenosti. Način prikupljanja podataka, broj korisnika, broj potpisa po korisniku i vrsta krivotvorina imaju značajan utjecaj na složenost problema i na dobivene rezultate verifikacije. Zbog toga je razumijevanje strukture i karakteristika korištenih baza ključno za pravilnu interpretaciju performansi sustava (Singh et al., 2022).

U području offline verifikacije potpisa razvijeno je više javno dostupnih baza podataka koje se koriste kao referentni skupovi u znanstvenim istraživanjima. Takve baze omogućuju ponovljivost eksperimenata i usporedbu različitih pristupa pod istim uvjetima. Istovremeno, one predstavljaju pojednostavljenu aproksimaciju stvarnih uvjeta primjene, što treba uzeti u obzir prilikom analize rezultata.

5.1. Uloga baza potpisa u razvoju sustava

Baze podataka potpisa imaju središnju ulogu u razvoju sustava za verifikaciju potpisa jer predstavljaju temelj za treniranje i testiranje modela. Kvaliteta i raznolikost podataka izravno utječu na sposobnost sustava da nauči razlikovati originalne potpise od krivotvorina. Sustav treniran na ograničenom ili nereprezentativnom skupu podataka može pokazivati dobre rezultate u eksperimentalnim uvjetima, ali slabije performanse u stvarnoj primjeni (Hafemann, Sabourin i Oliveira, 2017).

Osim treniranja modela, baze potpisa služe i za objektivnu evaluaciju performansi sustava. Korištenjem standardiziranih baza moguće je uspoređivati rezultate različitih metoda i algoritama na jednakim skupovima podataka. Takva usporedivost ključna je za znanstveni napredak u području verifikacije potpisa, jer omogućuje procjenu stvarnih prednosti i nedostataka pojedinih pristupa (Singh et al., 2022).

Važan aspekt baza potpisa je i struktura podataka, osobito omjer originalnih potpisa i krivotvorina te način njihove podjele na skupove za treniranje i testiranje. U realnim scenarijima često postoji znatna neravnoteža između broja originalnih potpisa i krivotvorina, što dodatno otežava učenje modela. Baze podataka omogućuju simulaciju takvih uvjeta i analizu ponašanja sustava u zahtjevnim situacijama (Zois et al., 2019).

U kontekstu offline verifikacije potpisa, baze podataka također omogućuju proučavanje utjecaja različitih vrsta krivotvorina na performanse sustava. Posebna pažnja posvećuje se vještim krivotvorinama, koje predstavljaju najrealističniji oblik napada. Analiza rezultata na takvim uzorcima daje realniju sliku sigurnosti i pouzdanosti sustava (Singh et al., 2022).

Razumijevanje uloge baza potpisa u razvoju i evaluaciji sustava predstavlja nužan preduvjet za pravilnu interpretaciju eksperimentalnih rezultata. U sljedećem potpoglavlju detaljnije se razmatraju najčešće korištene javne baze statičkih potpisa i njihove osnovne karakteristike.

5.2. Javne baze statičkih potpisa

U području offline verifikacije potpisa razvijeno je više javno dostupnih baza podataka koje se koriste kao standardni referentni skupovi u znanstvenim istraživanjima. Takve baze omogućuju ponovljivost eksperimenata, objektivnu usporedbu različitih metoda te evaluaciju performansi sustava pod jednakim uvjetima. Korištenjem javnih baza istraživači mogu uspoređivati rezultate svojih metoda s postojećim radovima i na taj način procijeniti relativnu uspješnost predloženih pristupa. Među najčešće korištenim bazama statičkih potpisa ubrajaju se CEDAR, GPDS i MCYT baze podataka (Hafemann, Sabourin i Oliveira, 2017).

CEDAR baza potpisa jedna je od najstarijih i najčešće korištenih baza u istraživanjima offline verifikacije potpisa. Sadrži rukom pisane potpise većeg broja korisnika, uključujući originalne potpise i vješte krivotvorine. Posebnost ove baze je dostupnost sustavno prikupljenih krivotvorina, što je čini pogodnom za evaluaciju sustava u zahtjevnim i realističnim uvjetima. Zbog svoje strukture, umjerene veličine i široke primjene u literaturi, CEDAR baza često se koristi kao referentni skup za usporedbu različitih pristupa verifikaciji potpisa te je zbog tih svojstava odabrana i za praktični dio ovog rada (Singh et al., 2022).

GPDS baza potpisa predstavlja jednu od najvećih javno dostupnih baza u ovom području. Karakterizira je velik broj korisnika i značajan broj potpisa po korisniku, što omogućuje detaljnu analizu ponašanja sustava u različitim scenarijima. Zbog svoje veličine i

raznolikosti, GPDS baza često se koristi u istraživanjima koja ispituju skalabilnost sustava, generalizaciju modela i primjenu naprednijih metoda strojnog učenja (Zois et al., 2019).

MCYT baza potpisa kombinira offline i online podatke, što je čini posebno prikladnom za usporedna istraživanja dvaju pristupa verifikaciji potpisa. U offline dijelu baze nalaze se skenirani rukom pisani potpisi prikupljeni u kontroliranim uvjetima, dok online dio uključuje dinamičke informacije o procesu potpisivanja. Ova baza omogućuje analizu utjecaja vrste dostupnih podataka na performanse sustava i često se koristi u istraživanjima koja uspoređuju offline i online metode verifikacije potpisa (Hafemann, Sabourin i Oliveira, 2017).

Iako navedene baze pružaju vrijednu osnovu za razvoj i evaluaciju sustava za verifikaciju potpisa, važno je naglasiti da one predstavljaju pojednostavljenu aproksimaciju stvarnih uvjeta primjene. Podaci su najčešće prikupljeni u kontroliranim uvjetima, a broj dostupnih uzoraka po korisniku često je veći nego u realnim scenarijima. Zbog toga je prilikom interpretacije rezultata dobivenih na javnim bazama potrebno uzeti u obzir njihova ograničenja, koja se detaljnije razmatraju u sljedećem potpoglavlju (Singh et al., 2022).

5.3. Ograničenja javno dostupnih baza

Iako javno dostupne baze potpisa predstavljaju temelj za razvoj i evaluaciju sustava za verifikaciju potpisa, one imaju određena ograničenja koja je potrebno uzeti u obzir prilikom interpretacije rezultata. Podaci u takvim bazama najčešće su prikupljeni u kontroliranim uvjetima, što može dovesti do razlika u odnosu na stvarne scenarije uporabe, gdje su uvjeti potpisivanja znatno raznolikiji (Singh et al., 2022).

Jedno od glavnih ograničenja odnosi se na način prikupljanja potpisa. U mnogim bazama potpisi su prikupljani tijekom jedne ili nekoliko sesija u relativno kratkom vremenskom razdoblju. Time se smanjuje vremenska varijabilnost potpisa iste osobe, iako je poznato da se potpis može mijenjati kroz dulje razdoblje zbog promjena u navikama pisanja, starosti ili zdravstvenog stanja. Sustavi trenirani na takvim podacima mogu pokazivati smanjenu robusnost pri primjeni na stvarnim, dugoročnim podacima (Singh et al., 2022).

Dodatno ograničenje predstavlja struktura krivotvorina u javnim bazama. Iako mnoge baze sadrže vješte krivotvorine, one su često izrađene u eksperimentalnim uvjetima, uz jasno definirane upute krivotvoriteljima. Takav način prikupljanja ne mora u potpunosti odražavati ponašanje krivotvoritelja u stvarnim napadima, gdje su motivacija, dostupnost uzoraka i razina vještine znatno raznolikiji (Engin et al., 2020).

Javne baze potpisa često se razlikuju i po broju dostupnih potpisa po korisniku. U eksperimentalnim skupovima podataka taj broj može biti veći nego u realnim sustavima, gdje su često dostupna tek dva ili tri originalna potpisa. Takva razlika može dovesti do optimistične procjene performansi sustava i otežati generalizaciju rezultata na stvarne uvjete primjene (Singh et al., 2022).

Konačno, kvaliteta slika potpisa u javnim bazama često je relativno visoka i ujednačena, s minimalnim šumom i dobrim kontrastom. U stvarnim dokumentima potpisi mogu biti degradirani zbog loše kvalitete skeniranja, prisutnosti pečata, linija ili drugih elemenata dokumenta. Ovi faktori mogu značajno utjecati na performanse offline sustava za verifikaciju potpisa, što dodatno naglašava potrebu za oprezom pri interpretaciji rezultata dobivenih na javnim bazama (Engin et al., 2020).

Razumijevanje ograničenja javno dostupnih baza potpisa ključno je za realističnu procjenu mogućnosti sustava za verifikaciju potpisa. Takva analiza omogućuje pravilno sagledavanje dobivenih rezultata i postavlja temelje za daljnju raspravu o metodama predprocesiranja i ekstrakcije značajki, koje se razmatraju u sljedećem poglavlju.

6. PREDPROCESIRANJE SLIKA POTPISA

U sustavima za offline verifikaciju potpisa predprocesiranje slika predstavlja jednu od ključnih faza obrade podataka. Budući da se analiza temelji isključivo na vizualnom prikazu potpisa, kvaliteta ulazne slike ima izravan utjecaj na uspješnost cijelog sustava. Razlike u načinu skeniranja, rezoluciji, osvjetljenju i prisutnosti šuma mogu značajno otežati kasnije faze ekstrakcije značajki i klasifikacije (Hafemann, Sabourin i Oliveira, 2017).

Cilj predprocesiranja nije poboljšati estetsku kvalitetu slike, već pripremiti podatke u oblik koji omogućuje stabilniju i robusniju analizu. Kroz niz standardiziranih postupaka nastoji se smanjiti utjecaj neželjenih varijacija koje nisu povezane s identitetom potpisnika, a istovremeno očuvati one informacije koje nose diskriminativnu vrijednost za verifikaciju potpisa (Singh et al., 2022).

Predprocesiranje se u offline verifikaciji potpisa obično sastoji od više uzastopnih koraka, pri čemu izbor i redoslijed postupaka ovise o karakteristikama podataka i primijenjenim metodama ekstrakcije značajki. Iako ne postoji jedinstveni pristup koji bi bio optimalan za sve sustave, određeni postupci pojavljuju se kao standardni u većini istraživanja.

6.1. Svrha predprocesiranja

Osnovna svrha predprocesiranja slika potpisa jest smanjenje varijabilnosti uzrokovane vanjskim čimbenicima koji nisu povezani s karakteristikama potpisnika. Takvi čimbenici uključuju razlike u osvjetljenju, kontrastu, rezoluciji slike i kvaliteti skeniranja. Uklanjanjem ili ublažavanjem njihovog utjecaja omogućuje se konzistentnija analiza potpisa u kasnijim fazama sustava (Hafemann, Sabourin i Oliveira, 2017).

Predprocesiranje ima važnu ulogu i u normalizaciji podataka. Potpisi mogu biti različitih veličina, položaja i orijentacije unutar slike, što otežava izravnu usporedbu između uzoraka. Postupci normalizacije omogućuju dovođenje potpisa u zajednički referentni okvir, čime se olakšava ekstrakcija značajki i poboljšava stabilnost klasifikacijskih modela (Singh et al., 2022).

Još jedan važan cilj predprocesiranja jest uklanjanje šuma i nepotrebnih elemenata slike. U stvarnim dokumentima potpisi se često pojavljuju zajedno s linijama, tekstom, pečatima ili drugim grafičkim elementima. Ako se takvi elementi ne uklone ili ne ublaže, mogu

negativno utjecati na izračun značajki i dovesti do pogrešnih odluka sustava. Zbog toga se predprocesiranje koristi za izdvajanje samog potpisa iz šireg konteksta dokumenta (Bouamra, 2022).

Pravilno provedeno predprocesiranje omogućuje bolju generalizaciju sustava i smanjuje osjetljivost modela na promjene koje nisu relevantne za identitet potpisnika. Iako predprocesiranje samo po sebi ne rješava problem verifikacije potpisa, ono predstavlja nužan korak koji osigurava kvalitetne ulazne podatke za ekstrakciju značajki i klasifikaciju. U sljedećim potpoglavljima detaljnije se razmatraju uobičajene tehnike predprocesiranja koje se primjenjuju u sustavima za offline verifikaciju potpisa.

6.2. Uobičajene tehnike predprocesiranja

U offline verifikaciji potpisa predprocesiranje se provodi primjenom niza standardnih tehnika čiji je cilj pripremiti sliku potpisa za pouzdanu ekstrakciju značajki. Iako konkretna implementacija može varirati ovisno o sustavu, određene tehnike pojavljuju se kao uobičajene u većini istraživanja i praktičnih rješenja (Hafemann, Sabourin i Oliveira, 2017).

Jedan od najčešćih koraka predprocesiranja je binarizacija slike potpisa. Binarizacijom se slika pretvara u dvovrijedni prikaz, pri čemu se pikseli klasificiraju kao dio potpisa ili pozadine. Time se pojednostavljuje daljnja obrada i smanjuje osjetljivost sustava na varijacije u osvjetljenju i boji podloge. U praksi se često koriste globalne ili adaptivne metode binarizacije, ovisno o kvaliteti ulazne slike (Singh et al., 2022).

Uklanjanje šuma predstavlja sljedeći važan korak predprocesiranja. Šum može nastati zbog kvalitete skeniranja, kompresije slike ili prisutnosti sitnih artefakata na dokumentu. Takvi neželjeni elementi mogu negativno utjecati na izračun značajki, osobito onih temeljenih na lokalnoj strukturi slike. Primjenom filtriranja i morfoloških operacija nastoji se ukloniti šum uz očuvanje bitnih dijelova potpisa (Hafemann, Sabourin i Oliveira, 2017).

Normalizacija veličine i položaja potpisa koristi se kako bi se smanjile razlike među uzorcima koje nisu povezane s identitetom potpisnika. Potpisi se mogu pojaviti na različitim mjestima unutar slike i u različitim razmjerima, što otežava njihovu izravnu usporedbu. Normalizacijom se potpis dovodi u standardizirani okvir, čime se omogućuje konzistentnija ekstrakcija značajki i stabilnije ponašanje klasifikatora (Singh et al., 2022).

U nekim sustavima provodi se i poravnanje ili korekcija nagiba potpisa. Nagib može varirati ovisno o načinu potpisivanja ili položaju dokumenta tijekom skeniranja. Iako korekcija nagiba nije uvijek nužna, u određenim slučajevima može doprinijeti smanjenju varijabilnosti i poboljšanju performansi sustava (Bouamra, 2022).

Odabir i kombinacija tehnika predprocesiranja ovise o karakteristikama korištene baze podataka i primijenjenim metodama ekstrakcije značajki. Pretjerano agresivno predprocesiranje može dovesti do gubitka informacija koje nose diskriminativnu vrijednost, dok nedovoljno predprocesiranje može ostaviti previše šuma i varijabilnosti. Zbog toga je važno pronaći ravnotežu između uklanjanja neželjenih utjecaja i očuvanja relevantnih obilježja potpisa.

6.3. Utjecaj predprocesiranja na performanse sustava

Predprocesiranje slika potpisa ima izravan utjecaj na performanse sustava za verifikaciju potpisa, jer određuje kvalitetu i konzistentnost ulaznih podataka na kojima se temelje sve daljnje faze obrade. Neodgovarajuće ili nedosljedno predprocesiranje može dovesti do značajnog pada točnosti sustava, čak i kada se koriste napredne metode ekstrakcije značajki i klasifikacije (Hafemann, Sabourin i Oliveira, 2017).

Kvalitetno predprocesiranje može smanjiti intra-class varijabilnost potpisa uklanjanjem neželjenih utjecaja, poput šuma, promjena u osvjetljenju i razlika u veličini slike. Time se omogućuje stabilnija ekstrakcija značajki i lakše učenje diskriminativnih obrazaca koji razlikuju originalne potpise od krivotvorina. U mnogim istraživanjima pokazano je da pažljivo odabrani koraci predprocesiranja mogu značajno poboljšati rezultate verifikacije, osobito u offline sustavima (Singh et al., 2022).

S druge strane, pretjerano agresivno predprocesiranje može imati negativan učinak na performanse sustava. Uklanjanje prevelikog broja detalja ili prekomjerna normalizacija mogu dovesti do gubitka informacija koje nose identitet potpisnika. Takav gubitak može smanjiti diskriminativnu moć značajki i otežati razlikovanje originalnih potpisa i vještih krivotvorina (Kumar, 2023).

Utjecaj predprocesiranja posebno je izražen u sustavima koji se oslanjaju na ručno dizajnirane značajke. Takve značajke često su osjetljive na kvalitetu ulazne slike i pretpostavljaju određenu razinu konzistentnosti podataka. Neujednačeno predprocesiranje

može dovesti do nestabilnih značajki i smanjene pouzdanosti klasifikatora (Hafemann, Sabourin i Oliveira, 2017).

Zbog navedenih razloga, predprocesiranje se ne smije promatrati kao izolirani korak, već kao sastavni dio cjelokupnog sustava za verifikaciju potpisa. Odabir tehnika predprocesiranja treba biti usklađen s metodama ekstrakcije značajki i klasifikacije koje se koriste u sustavu. Takav integrirani pristup omogućuje postizanje boljih performansi i realističniju procjenu mogućnosti sustava u stvarnim uvjetima primjene (Singh et al., 2022).

7. EKSTRAKCIJA ZNAČAJKI IZ STATIČKOG POTPISA

Ekstrakcija značajki predstavlja središnju fazu sustava za verifikaciju statičkog potpisa jer omogućuje pretvorbu vizualnog prikaza potpisa u numeričku reprezentaciju pogodnu za analizu i klasifikaciju. Budući da se offline verifikacija temelji isključivo na slici potpisa, kvaliteta i prikladnost izlučenih značajki izravno utječu na uspješnost cijelog sustava (Hafemann, Sabourin i Oliveira, 2017).

Cilj ekstrakcije značajki nije sačuvati cjelokupnu informaciju o slici, već izdvojiti one elemente koji nose diskriminativnu vrijednost za razlikovanje potpisnika. Takvi elementi mogu opisivati globalni oblik potpisa, lokalne strukture poteza ili teksturalna svojstva slike. Odabir vrste značajki ovisi o prirodi problema, vrsti dostupnih podataka i primijenjenom klasifikacijskom pristupu (Singh et al., 2022).

U kontekstu statičkog potpisa, ekstrakcija značajki mora se nositi s prirodnom varijabilnošću potpisa iste osobe i istodobno zadržati osjetljivost na razlike koje upućuju na krivotvorenje. Zbog toga se u praksi često kombiniraju različite vrste značajki kako bi se postigla bolja ravnoteža između robusnosti i diskriminativnosti.

7.1. Uloga značajki u verifikaciji potpisa

Značajke imaju ključnu ulogu u sustavima za verifikaciju potpisa jer predstavljaju vezu između sirovih podataka i klasifikacijskog modela. Kvaliteta značajki određuje koliko će sustav biti sposoban razlikovati originalne potpise od krivotvorina, neovisno o odabranom algoritmu strojnog učenja (Hafemann, Sabourin i Oliveira, 2017).

Dobro odabrane značajke trebaju zadovoljiti nekoliko osnovnih kriterija. One moraju biti dovoljno stabilne kako bi tolerirale prirodne varijacije potpisa iste osobe, ali istovremeno dovoljno osjetljive da detektiraju razlike između potpisa različitih osoba ili između originala i krivotvorine. Ako značajke ne ispunjavaju ove zahtjeve, sustav može pokazivati visoke stope pogrešnog odbacivanja ili pogrešnog prihvaćanja (Singh et al., 2022).

U offline verifikaciji potpisa značajke se najčešće temelje na obliku i strukturi potpisa. Budući da nisu dostupne dinamičke informacije, sustav mora iskoristiti prostorne odnose, distribuciju poteza i teksturalna svojstva slike. Zbog toga je ekstrakcija značajki u offline

sustavima posebno zahtjevna i ima veći utjecaj na performanse u usporedbi s online pristupom.

U literaturi je pokazano da čak i relativno jednostavni klasifikatori mogu postići dobre rezultate ako su korištene značajke dovoljno informativne. Suprotno tome, primjena složenih algoritama strojnog učenja ne može nadoknaditi loš odabir značajki. Zbog toga se u razvoju sustava za verifikaciju potpisa posebna pažnja posvećuje dizajnu i evaluaciji značajki prije samog postupka klasifikacije (Hafemann, Sabourin i Oliveira, 2017).

Razumijevanje uloge značajki predstavlja temelj za daljnju klasifikaciju i analizu različitih vrsta značajki koje se koriste u verifikaciji statičkog potpisa. U sljedećem potpoglavlju razmatra se podjela značajki i njihove osnovne karakteristike.

7.2. Klasifikacija značajki

Značajke koje se koriste u sustavima za verifikaciju statičkog potpisa mogu se klasificirati prema različitim kriterijima, ovisno o načinu na koji opisuju potpis i razini informacija koje obuhvaćaju. Najčešće se dijele na globalne i lokalne značajke, kao i na strukturne i teksturalne značajke. Ova klasifikacija omogućuje sustavno sagledavanje prednosti i ograničenja pojedinih pristupa (Hafemann, Sabourin i Oliveira, 2017).

Globalne značajke opisuju potpis kao cjelinu, bez detaljnog razmatranja lokalnih struktura. Takve značajke uključuju mjere povezane s ukupnim oblikom potpisa, poput omjera širine i visine, gustoće poteza ili globalne distribucije piksela. Prednost globalnih značajki je njihova relativna jednostavnost i robusnost na sitne lokalne varijacije, ali im je nedostatak ograničena diskriminativna moć u slučaju vještih krivotvorina (Hafemann, Sabourin i Oliveira, 2015).

Lokalne značajke fokusiraju se na analizu manjih dijelova potpisa ili lokalnih obrazaca unutar slike. One opisuju detalje poput lokalne orijentacije poteza, zakrivljenosti ili promjena teksture u pojedinim regijama potpisa. Zbog svoje detaljnosti, lokalne značajke često imaju veću diskriminativnu snagu, osobito u razlikovanju originalnih potpisa i vještih krivotvorina, ali su istovremeno osjetljivije na šum i varijabilnost podataka (Hafemann, Sabourin i Oliveira, 2017).

Strukturne značajke nastoje opisati geometrijsku strukturu potpisa, uključujući odnose između pojedinih dijelova, raspored poteza i topološke karakteristike. Takve značajke često se

temelje na analizi kontura ili skeleta potpisa i mogu pružiti uvid u organizaciju i stil pisanja potpisnika. Međutim, njihova ekstrakcija može biti složena i osjetljiva na kvalitetu predprocesiranja (Hafemann, Sabourin i Oliveira, 2015).

Teksturalne značajke opisuju raspodjelu i odnose piksela unutar slike potpisa, bez eksplicitnog modeliranja strukture poteza. One se često temelje na statističkim mjerama ili histogramima koji opisuju lokalne promjene intenziteta ili orijentacije. Teksturalne značajke pokazale su se učinkovitima u offline verifikaciji potpisa jer mogu uhvatiti suptilne razlike u izgledu potpisa koje su teško uočljive na globalnoj razini (Hafemann, Sabourin i Oliveira, 2017).

U praksi se često kombiniraju različite vrste značajki kako bi se iskoristile njihove komplementarne prednosti. Kombinacija globalnih i lokalnih, odnosno strukturnih i teksturalnih značajki može rezultirati robusnijim i diskriminativnijim reprezentacijama potpisa. Takav pristup omogućuje sustavu da se učinkovitije nosi s varijabilnošću potpisa i prisutnošću krivotvorina, što je ključno za postizanje pouzdanih rezultata verifikacije (Singh et al., 2022).

7.3. Ručno dizajnirane značajke

Ručno dizajnirane značajke predstavljaju tradicionalni pristup ekstrakciji informacija iz statičkog potpisa. Ovaj pristup temelji se na eksplicitnom definiranju karakteristika koje se smatraju relevantnima za razlikovanje originalnih potpisa i krivotvorina, pri čemu dizajn značajki proizlazi iz domenskog znanja o strukturi i izgledu potpisa. Ručno dizajnirane značajke dugo su bile temelj sustava za offline verifikaciju potpisa i još se uvijek često koriste u praksi (Hafemann, Sabourin i Oliveira, 2017).

Geometrijske značajke opisuju globalna i lokalna svojstva oblika potpisa. U ovu skupinu ubrajaju se mjere poput širine i visine potpisa, omjera dimenzija, površine zauzete potpisom te raspodjele piksela unutar slike. Takve značajke relativno su jednostavne za izračun i robusne na sitne varijacije, ali često nemaju dovoljnu diskriminativnu moć za razlikovanje vještih krivotvorina koje uspješno oponašaju opći oblik potpisa (Hafemann, Sabourin i Oliveira, 2015).

Teksturalne značajke usmjerene su na analizu lokalnih obrazaca i statističkih svojstava slike potpisa. One opisuju raspodjelu intenziteta piksela, promjene u teksturi i lokalne varijacije koje nastaju kao posljedica načina pisanja. Takve značajke mogu uhvatiti suptilne razlike u izgledu poteza koje nisu vidljive na globalnoj razini, što ih čini posebno korisnima u offline verifikaciji potpisa (Singh et al., 2022).

Među najčešće korištenim teksturalnim značajkama u verifikaciji statičkog potpisa ističu se Histogram of Oriented Gradients (HOG) značajke. HOG opisuje lokalnu distribuciju orijentacija gradijenata unutar slike, čime se učinkovito bilježe rubovi i struktura poteza potpisa. Zbog svoje sposobnosti da opiše oblik i teksturu, HOG značajke pokazale su se učinkovitima u razlikovanju originalnih potpisa i krivotvorina u offline scenarijima (Kumar, 2023).

Prednost ručno dizajniranih značajki leži u njihovoj interpretabilnosti i relativno malim zahtjevima za količinom podataka. Takve značajke često se mogu koristiti u kombinaciji s klasičnim algoritmima strojnog učenja i postići zadovoljavajuće rezultate čak i u uvjetima ograničenog broja uzoraka po korisniku. Međutim, njihova učinkovitost uvelike ovisi o kvaliteti predprocesiranja i pravilnom odabiru značajki (Hafemann, Sabourin i Oliveira, 2017).

Unatoč svojim ograničenjima, ručno dizajnirane značajke i dalje imaju važnu ulogu u verifikaciji statičkog potpisa, osobito u sustavima koji se temelje na writer-dependent pristupu. Njihova jednostavnost, stabilnost i mogućnost primjene u realnim uvjetima čine ih relevantnima i u suvremenim sustavima, što se odražava i u praktičnom dijelu ovog rada.

7.4. Napredni pristupi ekstrakciji značajki

Osim klasičnih ručno dizajniranih značajki, u području offline verifikacije potpisa razvijeni su i napredniji pristupi ekstrakciji značajki koji nastoje modelirati strukturu potpisa na apstraktnijoj razini. Takvi pristupi često proizlaze iz potrebe za boljim razlikovanjem originalnih potpisa i vještih krivotvorina u uvjetima ograničenog broja uzoraka i izražene varijabilnosti potpisa (Zois et al., 2019).

Jedan od značajnih naprednih pristupa temelji se na sparse representation metodama. U ovom pristupu potpis se predstavlja kao linearna kombinacija malog broja elemenata iz unaprijed definirane rječničke baze. Ideja je da originalni potpisi iste osobe mogu biti učinkovito rekonstruirani korištenjem rijetke reprezentacije, dok krivotvorine zahtijevaju drugačiju ili manje učinkovitu kombinaciju elemenata. Takva razlika u reprezentaciji može se iskoristiti za donošenje odluke o autentičnosti potpisa (Zois et al., 2019).

Sparse representation pristupi posebno su zanimljivi u offline verifikaciji potpisa jer omogućuju modeliranje globalne strukture potpisa bez potrebe za velikim količinama podataka. Istraživanja su pokazala da ovakvi pristupi mogu postići konkurentne rezultate u writer-dependent scenarijima, gdje je broj dostupnih uzoraka po korisniku ograničen. Međutim,

njihova primjena može biti računalno zahtjevna, osobito kod većih rječnika i baza podataka (Zois et al., 2019).

Drugi napredni smjer istraživanja uključuje grafovske pristupe ekstrakciji značajki. U grafovskim metodama potpis se modelira kao graf, pri čemu čvorovi i bridovi opisuju strukturne elemente potpisa i njihove međusobne odnose. Takav pristup omogućuje eksplicitno modeliranje topologije i organizacije poteza potpisa, što može biti posebno korisno za razlikovanje originalnih potpisa i vještih krivotvorina (Maergner et al., 2019).

Grafovske reprezentacije omogućuju očuvanje strukturnih informacija koje se često gube pri klasičnim teksturalnim ili globalnim pristupima. U kontekstu verifikacije potpisa, grafovi mogu opisivati odnose između segmenata, raskrižja poteza ili ključnih točaka potpisa. Iako grafovski pristupi nude bogatu i informativnu reprezentaciju, njihova primjena često je povezana s većom složenošću implementacije i povećanim računalnim zahtjevima (Maergner et al., 2019).

Napredni pristupi ekstrakciji značajki, poput sparse representation i grafovskih metoda, nastoje premostiti ograničenja klasičnih ručno dizajniranih značajki. Iako se rjeđe koriste u praktičnim sustavima zbog složenosti, oni imaju važnu ulogu u istraživačkom kontekstu jer pružaju dublji uvid u strukturna svojstva potpisa i otvaraju mogućnosti za daljnje unapređenje sustava za verifikaciju statičkog potpisa.

8. METODE STROJNOG UČENJA ZA VERIFIKACIJU POTPISA

Metode strojnog učenja predstavljaju ključnu komponentu sustava za verifikaciju statičkog potpisa jer omogućuju donošenje odluke o autentičnosti potpisa na temelju izlučenih značajki. Nakon faza predprocesiranja i ekstrakcije značajki, zadatak klasifikatora je naučiti razliku između originalnih potpisa i krivotvorina te na temelju tog znanja procijeniti nove, nepoznate uzorke (Hafemann, Sabourin i Oliveira, 2017).

U kontekstu verifikacije potpisa, problem se najčešće formulira kao binarni klasifikacijski zadatak, pri čemu se potpisi razvrstavaju u dvije klase: originalni potpisi i krivotvorine. Za razliku od klasičnih problema klasifikacije, ovdje je naglasak na pouzdanosti odluke i kontroli kompromisa između sigurnosti i upotrebljivosti sustava. Zbog toga izbor metode strojnog učenja ima izravan utjecaj na praktičnu primjenjivost sustava (Bouamra, 2022).

Različite metode strojnog učenja nude različite prednosti i ograničenja u kontekstu verifikacije potpisa. Neke metode bolje se nose s malim skupovima podataka, dok druge zahtijevaju veći broj uzoraka za postizanje stabilnih rezultata. Također, određene metode omogućuju jednostavniju interpretaciju rezultata, dok su druge fokusirane isključivo na postizanje što bolje diskriminativne moći. Zbog toga se u literaturi često ispituje više pristupa kako bi se identificirala najprikladnija metoda za određeni scenarij uporabe (Hafemann, Sabourin i Oliveira, 2017).

8.1. Uloga klasifikatora u verifikacijskom sustavu

Klasifikator u sustavu za verifikaciju potpisa ima zadatak donijeti odluku o tome pripada li promatrani potpis deklariranom identitetu korisnika. Ta odluka temelji se na numeričkoj reprezentaciji potpisa dobivenoj postupkom ekstrakcije značajki. Klasifikator uči granicu razdvajanja između originalnih potpisa i krivotvorina te na temelju te granice procjenjuje nove uzorke (Hafemann, Sabourin i Oliveira, 2017).

U verifikacijskom scenariju klasifikator ne dodjeljuje potpis jednoj od više klasa, već procjenjuje sličnost ili pripadnost u odnosu na referentne potpise određene osobe. Rezultat klasifikacije često se izražava u obliku kontinuirane vrijednosti ili skora, koji se zatim

uspoređuje s unaprijed definiranim pragom odluke. Na temelju tog praga sustav donosi konačnu odluku o prihvaćanju ili odbacivanju potpisa (Bouamra, 2022).

Uloga klasifikatora posebno je naglašena u writer-dependent sustavima, gdje se za svakog korisnika trenira zaseban model. U takvim sustavima klasifikator mora učinkovito naučiti granicu između legitimnih varijacija potpisa iste osobe i uzoraka koji predstavljaju krivotvorine. To zahtijeva metode koje su robusne na mali broj uzoraka i sposobne generalizirati na nove potpise (Hafemann, Sabourin i Oliveira, 2017).

Odabir klasifikatora izravno utječe na ponašanje sustava u pogledu sigurnosti i upotrebljivosti. Previše restriktivan klasifikator može dovesti do čestog odbacivanja originalnih potpisa, dok previše tolerantan klasifikator povećava rizik prihvaćanja krivotvorina. Zbog toga se klasifikator i prag odluke moraju promatrati zajedno kao sastavni dio verifikacijskog sustava, što se detaljnije razmatra u kasnijem poglavlju o evaluaciji sustava (Bouamra, 2022).

8.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) predstavlja jednu od najčešće korištenih metoda strojnog učenja u sustavima za offline verifikaciju potpisa. Razlog tome leži u njegovoj sposobnosti učinkovite klasifikacije u uvjetima ograničenog broja uzoraka, što je čest slučaj u writer-dependent scenarijima verifikacije potpisa (Hafemann et al., 2017).

Osnovna ideja SVM-a temelji se na pronalaženju optimalne razdvajajuće granice između dviju klasa u prostoru značajki. Ta granica definira se tako da maksimizira razmak između najbližih uzoraka različitih klasa, koji se nazivaju potporni vektori. Maksimizacijom tog razmaka postiže se bolja generalizacija modela na nove, nepoznate uzorke (Singh et al., 2022).

U kontekstu verifikacije statičkog potpisa, SVM se najčešće koristi kao binarni klasifikator koji razdvaja originalne potpise i krivotvorine. Model se trenira na temelju izlučenih značajki potpisa, pri čemu se u writer-dependent pristupu za svakog korisnika trenira zaseban SVM model. Takav pristup omogućuje prilagodbu granice odlučivanja specifičnim karakteristikama potpisa pojedinca (Zois et al., 2019).

Jedna od ključnih prednosti SVM-a je mogućnost korištenja kernel funkcija, koje omogućuju nelinearno razdvajanje podataka u izvornom prostoru značajki. U offline verifikaciji potpisa često se koristi radijalna bazna funkcija (RBF kernel), jer omogućuje modeliranje

složenih odnosa između značajki potpisa. Time se povećava sposobnost sustava da razlikuje originalne potpise i vješte krivotvorine koje se preklapaju u prostoru značajki (Kumar, 2023).

SVM pokazuje dobru otpornost na prenaučavanje, osobito u scenarijima s malim brojem uzoraka, što ga čini prikladnim za realne uvjete primjene sustava za verifikaciju potpisa. Međutim, performanse SVM-a uvelike ovise o pravilnom odabiru hiperparametara, poput parametra regularizacije i parametara kernel funkcije. Neodgovarajući odabir tih parametara može dovesti do loših rezultata, bez obzira na kvalitetu izlučenih značajki (Singh et al., 2022).

U praksi, SVM se često kombinira s ručno dizajniranim značajkama, poput teksturalnih ili HOG značajki, kako bi se postigla ravnoteža između robusnosti i diskriminativne moći sustava. Takva kombinacija pokazala se učinkovitom u brojnim istraživanjima i predstavlja stabilnu osnovu za implementaciju sustava za offline verifikaciju potpisa, uključujući i sustav razmatran u praktičnom dijelu ovog rada (Kumar, 2023; Zois et al., 2019).

8.3. Ostale metode strojnog učenja

Osim Support Vector Machine metode, u području offline verifikacije potpisa istraživane su i druge metode strojnog učenja koje se mogu primijeniti na problem razlikovanja originalnih potpisa i krivotvorina. Ove metode razlikuju se po načinu modeliranja podataka, zahtjevima u pogledu količine podataka i razini složenosti implementacije (Hafemann et al., 2017).

Jedan od pristupa koji se koristi u verifikaciji potpisa su statistički klasifikatori, poput k-najbližih susjeda (k-NN). Ovakve metode temelje se na mjerenju sličnosti između potpisa u prostoru značajki. Iako su konceptualno jednostavne i lako interpretabilne, njihova učinkovitost uvelike ovisi o odabiru metrike udaljenosti i broju dostupnih uzoraka, zbog čega često pokazuju slabije performanse u usporedbi s naprednijim metodama u offline scenarijima (Hafemann et al., 2017).

U literaturi se također razmatraju metode temeljene na probabilističkim modelima, poput Gaussian Mixture Models. Takvi modeli nastoje opisati distribuciju značajki originalnih potpisa i procijeniti vjerojatnost da promatrani potpis pripada toj distribuciji. Iako mogu pružiti intuitivan probabilistički okvir, njihova primjena u offline verifikaciji potpisa često je ograničena zbog pretpostavki o distribuciji podataka i osjetljivosti na mali broj uzoraka (Maergner et al., 2019).

Ansambl metode, poput Random Forest klasifikatora, također su primjenjivane u sustavima za verifikaciju potpisa. Ove metode kombiniraju više jednostavnijih klasifikatora kako bi se poboljšala stabilnost i točnost odluka. Iako mogu postići dobre rezultate u određenim uvjetima, njihova primjena u writer-dependent sustavima često zahtijeva veći broj uzoraka za treniranje kako bi se izbjeglo prenaučavanje (Hafemann et al., 2017).

U istraživačkom kontekstu razmatrani su i strukturni pristupi koji kombiniraju ekstrakciju značajki s naprednim metodama klasifikacije, poput grafovskih modela. Takvi pristupi mogu pružiti detaljan opis strukture potpisa, ali su često računalno zahtjevni i složeni za implementaciju, što ograničava njihovu primjenu u praktičnim sustavima (Maergner et al., 2019).

Unatoč postojanju različitih metoda strojnog učenja, Support Vector Machine ostaje jedna od najčešće korištenih i najstabilnijih metoda u offline verifikaciji potpisa. Njegova sposobnost rada s ograničenim skupovima podataka i dobra generalizacija čine ga prikladnim izborom za sustave poput onog razmatranog u ovom radu, dok se ostale metode često koriste u usporednim analizama ili istraživačkim eksperimentima.

9. EVALUACIJA SUSTAVA VERIFIKACIJE POTPISA

Evaluacija sustava za verifikaciju potpisa predstavlja ključni korak u razvoju i analizi biometrijskih sustava jer omogućuje objektivnu procjenu njihove pouzdanosti i primjenjivosti. Cilj evaluacije nije samo utvrditi ukupnu točnost sustava, već i analizirati ponašanje sustava u različitim uvjetima i scenarijima uporabe, osobito u prisutnosti krivotvorina (Hafemann et al., 2017).

U kontekstu verifikacije potpisa, evaluacija se provodi usporedbom odluka sustava s poznatim oznakama uzoraka, pri čemu se ispituje sposobnost sustava da ispravno prihvati originalne potpise i odbaci krivotvorine. Posebna se pažnja posvećuje ravnoteži između sigurnosti i upotrebljivosti, jer sustav koji je previše restriktivan ili previše tolerantan ne zadovoljava praktične zahtjeve (Singh et al., 2022).

Za razliku od klasičnih klasifikacijskih problema, evaluacija biometrijskih sustava zahtijeva primjenu specifičnih metrika koje uzimaju u obzir prirodu verifikacijskog zadatka. Te metrike omogućuju detaljniju analizu pogrešaka sustava i jasnije sagledavanje kompromisa između različitih vrsta pogrešnih odluka.

9.1. Evaluacija biometrijskih sustava

Evaluacija biometrijskih sustava razlikuje se od evaluacije standardnih sustava strojnog učenja zbog specifičnih zahtjeva vezanih uz sigurnost i pouzdanost. U biometrijskim sustavima pogrešne odluke nemaju jednaku težinu, jer pogrešno prihvaćanje neautorizirane osobe i pogrešno odbacivanje legitimnog korisnika imaju različite posljedice u praksi (Hafemann et al., 2017).

U evaluaciji biometrijskih sustava koristi se skup testnih uzoraka koji uključuje originalne potpise i krivotvorine. Sustav se ispituje u kontroliranim uvjetima kako bi se procijenilo njegovo ponašanje pri različitim postavkama praga odluke. Na taj se način može analizirati kako promjena praga utječe na sigurnost i upotrebljivost sustava (Singh et al., 2022).

Važan aspekt evaluacije biometrijskih sustava je ponovljivost rezultata. Korištenjem standardiziranih baza podataka i jasno definiranih protokola evaluacije omogućuje se usporedba različitih metoda i pristupa pod jednakim uvjetima. Takva usporedivost ključna je

za znanstvena istraživanja i za procjenu stvarnih prednosti pojedinih rješenja (Hafemann et al., 2017).

Evaluacija biometrijskih sustava ne svodi se isključivo na jednu numeričku mjeru uspješnosti. Umjesto toga, koristi se skup metrika koje zajedno pružaju cjelovit uvid u ponašanje sustava. Te metrike i njihova interpretacija detaljnije se razmatraju u sljedećem potpoglavlju.

9.2. Evaluacijske metrike

Evaluacija sustava za verifikaciju potpisa temelji se na primjeni specifičnih metrika koje opisuju različite vrste pogrešaka u biometrijskim sustavima. Za razliku od opće točnosti, ove metrike omogućuju detaljniju analizu ponašanja sustava i pružaju uvid u kompromis između sigurnosti i upotrebljivosti (Hafemann et al., 2017).

False Acceptance Rate (FAR) predstavlja udio krivotvorenih potpisa koje sustav pogrešno prihvaća kao originalne. Ova metrika izravno je povezana sa sigurnošću sustava, jer visoka vrijednost FAR-a znači povećan rizik od neautoriziranog pristupa. U sustavima za verifikaciju potpisa FAR se najčešće mjeri korištenjem vještih krivotvorina, koje predstavljaju najzahtjevniji oblik napada (Singh et al., 2022).

False Rejection Rate (FRR) označava udio originalnih potpisa koje sustav pogrešno odbacuje. Ova metrika odražava upotrebljivost sustava, jer visoka vrijednost FRR-a dovodi do nezadovoljstva legitimnih korisnika i smanjene praktične primjenjivosti sustava. Sustav s niskim FRR-om tolerantniji je na prirodne varijacije potpisa iste osobe, ali takva tolerancija može utjecati na sigurnost (Hafemann et al., 2017).

Equal Error Rate (EER) predstavlja točku u kojoj su vrijednosti FAR-a i FRR-a jednake. EER se često koristi kao sažeta mjera performansi biometrijskog sustava, jer omogućuje jednostavnu usporedbu različitih metoda i sustava. Niža vrijednost EER-a ukazuje na bolju ukupnu ravnotežu između sigurnosti i upotrebljivosti sustava (Singh et al., 2022).

Iako su FAR, FRR i EER temeljne metrike u evaluaciji sustava za verifikaciju potpisa, važno je naglasiti da niti jedna od njih sama po sebi ne pruža potpunu sliku o ponašanju sustava. Njihova interpretacija mora se promatrati u kontekstu konkretne primjene i u odnosu na odabrani prag odluke. U sljedećem potpoglavlju razmatra se upravo uloga praga odluke i kompromis između sigurnosti i praktične primjene sustava.

9.3. Prag odluke i kompromis sigurnosti

U sustavima za verifikaciju potpisa odluka o prihvaćanju ili odbacivanju potpisa donosi se usporedbom izlazne vrijednosti klasifikatora s unaprijed definiranim pragom odluke. Taj prag predstavlja graničnu vrijednost na temelju koje se odlučuje smatra li se promatrani potpis originalnim ili krivotvorenim. Odabir praga ima ključan utjecaj na ponašanje sustava i izravno određuje ravnotežu između sigurnosti i upotrebljivosti (Hafemann et al., 2017).

Promjenom praga odluke moguće je upravljati stopama pogrešnog prihvaćanja i pogrešnog odbacivanja. Smanjenjem praga sustav postaje tolerantniji, što dovodi do niže stope pogrešnog odbacivanja originalnih potpisa, ali istovremeno povećava rizik pogrešnog prihvaćanja krivotvorina. Suprotno tome, povećanjem praga sustav postaje restriktivniji, čime se smanjuje FAR, ali raste FRR. Ovaj odnos jasno pokazuje postojanje kompromisa između sigurnosti i praktične upotrebljivosti sustava (Kumar, 2023).

U praksi, optimalan prag odluke ovisi o kontekstu primjene sustava. U sigurnosno osjetljivim sustavima, poput financijskih ili pravnih aplikacija, često se preferira niži FAR, čak i po cijenu višeg FRR-a. U takvim slučajevima sustav daje prednost sigurnosti nad udobnošću korisnika. Suprotno tome, u aplikacijama gdje je korisničko iskustvo važnije, prag se može postaviti tako da se smanji broj pogrešnih odbacivanja, uz prihvatljiv porast sigurnosnog rizika (Hafemann et al., 2017).

Analiza kompromisa između FAR-a i FRR-a često se provodi korištenjem krivulja poput ROC ili DET krivulje, koje prikazuju ponašanje sustava pri različitim postavkama praga odluke. Takva analiza omogućuje detaljnije razumijevanje performansi sustava i olakšava donošenje odluke o odgovarajućoj postavci praga u skladu s konkretnim zahtjevima primjene (Hafemann et al., 2017).

Razumijevanje uloge praga odluke ključno je za pravilnu interpretaciju rezultata evaluacije sustava za verifikaciju potpisa. Prag odluke ne predstavlja statičnu vrijednost, već parametar koji se mora prilagoditi kontekstu uporabe i sigurnosnim zahtjevima sustava. Time se završava teorijska rasprava o evaluaciji sustava verifikacije potpisa i postavlja temelj za završni zaključak teorijskog dijela rada.

10. ZAKLJUČAK TEORETSKOG DIJELA

U ovom teorijskom dijelu rada razmotreni su temeljni koncepti, izazovi i pristupi povezani s verifikacijom statičkog potpisa kao bihevioralne biometrijske značajke. Potpis je sagledan u širem povijesnom, pravnom i biometrijskom kontekstu, čime je istaknuta njegova dugotrajna društvena prihvaćenost i praktična relevantnost unatoč razvoju suvremenih metoda autentikacije.

Posebna pažnja posvećena je razlikama između online i offline verifikacije potpisa, pri čemu su istaknuta ograničenja i izazovi statičkog pristupa. Nedostatak dinamičkih informacija, izražena varijabilnost potpisa iste osobe i prisutnost vještih krivotvorina čine offline verifikaciju zahtjevnim problemom koji zahtijeva pažljivo osmišljene metode obrade slike i strojnog učenja.

U radu su analizirani različiti aspekti sustava za verifikaciju potpisa, uključujući ulogu baza podataka, predprocesiranje slika, ekstrakciju značajki i odabir klasifikacijskih metoda. Poseban naglasak stavljen je na ručno dizajnirane i napredne pristupe ekstrakciji značajki, kao i na primjenu Support Vector Machine metode u writer-dependent scenariju, koji se pokazao prikladnim za uvjete ograničenog broja uzoraka po korisniku.

Evaluacija sustava za verifikaciju potpisa razmotrena je kroz prizmu biometrijskih metrika i kompromisa između sigurnosti i upotrebljivosti. Analiza evaluacijskih metrika, poput FAR-a, FRR-a i EER-a, te uloge praga odluke omogućuje dublje razumijevanje ponašanja sustava i njegovih ograničenja u realnim uvjetima primjene.

Teorijski dio rada pruža cjelovitu podlogu za praktični dio, u kojem se navedeni koncepti primjenjuju kroz implementaciju konkretnog sustava za verifikaciju statičkog potpisa. Time se ostvaruje poveznica između teorijskih spoznaja i praktičnih rješenja, a teorijska analiza služi kao okvir za interpretaciju eksperimentalnih rezultata i raspravu o mogućim smjerovima daljnjeg razvoja sustava.

11. IMPLEMENTACIJA SUSTAVA ZA VERIFIKACIJU STATIČKOG POTPISA

Ovo poglavlje opisuje implementaciju sustava za verifikaciju statičkog potpisa temeljenog na metodama obrade slike i strojnog učenja. Implementacija je izrađena u skladu s teorijskim okvirom predstavljenim u prethodnim poglavljima te je usmjerena na praktičnu realizaciju writer-dependent sustava za offline verifikaciju potpisa.

Sustav je osmišljen tako da verificira tvrdnju identiteta na temelju statičke slike potpisa, bez korištenja dinamičkih informacija o procesu potpisivanja. U skladu s time, ulaz sustava čini digitalna slika potpisa i identifikator osobe čiji se identitet provjerava, dok je izlaz binarna odluka o prihvaćanju ili odbacivanju te tvrdnje. Sustav ne provodi identifikaciju potpisnika među više osoba, već isključivo provjerava odgovara li priloženi potpis deklariranom identitetu.

Implementacija je organizirana kao modularni pipeline koji obuhvaća pripremu podataka, predprocesiranje slika potpisa, ekstrakciju značajki, treniranje klasifikacijskih modela, evaluaciju performansi te verifikaciju pojedinačnih potpisa. Poseban naglasak stavljen je na jasno razdvajanje faze treniranja modela i faze verifikacije, kako bi se osigurala ponovljivost eksperimenata i realističan scenarij primjene sustava.

Cjelokupni izvorni kod sustava, uključujući sve skripte korištene u implementacijskom i evaluacijskom dijelu rada, dostupan je u javnom GitHub repozitoriju na adresi:

https://github.com/lukaposta/static-signature-verification

U implementacijskom dijelu rada svi ključni koraci sustava dokumentirani su na način da se izravno povezuju s odgovarajućim teorijskim konceptima, pri čemu se selektivno navode relevantni isječci izvornog koda radi ilustracije konkretne realizacije opisanih postupaka.

11.1. Koncept i arhitektura implementiranog sustava

Implementirani sustav temelji se na writer-dependent pristupu verifikaciji potpisa. U takvom sustavu za svakog korisnika trenira se poseban binarni klasifikator koji razlikuje originalne potpise te osobe od potpisa koji ne pripadaju toj osobi. Negativnu klasu u procesu učenja čine potpisi drugih osoba, dok se krivotvorine koriste isključivo u fazi evaluacije sustava.

Arhitektura sustava organizirana je u dvije jasno odvojene faze: fazu treniranja i fazu verifikacije. U fazi treniranja, nad unaprijed definiranim skupom podataka provodi se predprocesiranje slika, ekstrakcija značajki i učenje klasifikacijskog modela za svaku osobu. Dobiveni modeli spremaju se na disk i u toj se fazi sustav u potpunosti konfigurira. Tijekom faze verifikacije ne provodi se nikakvo dodatno učenje, već se koriste isključivo prethodno istrenirani modeli.

Svaki model pohranjen je kao zasebna datoteka te sadrži istrenirani klasifikator, parametre predprocesiranja, parametre ekstrakcije značajki i prag odlučivanja. Takva organizacija omogućuje da se tijekom verifikacije pojedinačnog potpisa koristi potpuno ista konfiguracija sustava kao u fazi treniranja, čime se osigurava konzistentnost obrade podataka.

Cjelokupni tok obrade potpisa u sustavu može se sažeti u sljedeće korake: učitavanje slike potpisa, predprocesiranje slike, ekstrakcija značajki, primjena odgovarajućeg klasifikatora te donošenje odluke o prihvaćanju ili odbacivanju tvrdnje identiteta. Ovakav dizajn sustava omogućuje jasnu povezanost između teorijskih razmatranja o verifikaciji potpisa i njihove konkretne implementacije u programskom kodu.

11.2. Skup podataka i organizacija dataseta

Za implementaciju sustava za verifikaciju statičkog potpisa korišten je CEDAR Signature Dataset, koji je jedan od standardnih i najčešće korištenih skupova podataka u istraživanjima offline verifikacije potpisa. Odabrani podskup obuhvaća 55 osoba, pri čemu je za svaku osobu dostupno 24 originalna potpisa i 24 vješte krivotvorine. Takva struktura omogućuje izgradnju i evaluaciju sustava u realističnom scenariju koji uključuje i prisutnost kvalitetnih krivotvorina.

Dataset je organiziran hijerarhijski na razini datotečnog sustava. Za svaku osobu postoji zaseban direktorij koji sadrži originalne potpise te dodatni direktorij koji sadrži krivotvorine iste osobe. Direktoriji s krivotvorinama označeni su sufiksom `_forg`, dok naziv osnovnog direktorija odgovara identifikatoru osobe. Ovakva organizacija omogućuje jednoznačno mapiranje svakog potpisa na pripadajuću osobu i klasu te jasno razdvajanje originalnih potpisa i krivotvorina već na razini strukture podataka.

U okviru implementacije, struktura direktorija ne koristi se izravno u fazi treniranja i evaluacije, već služi kao ulaz za generiranje centralizirane CSV datoteke koja opisuje cjelokupan skup podataka. Na taj način se logika podjele na skupove za treniranje i testiranje

izdvaja iz samog koda za učenje modela, čime se postiže veća transparentnost i ponovljivost eksperimenata.

Svi daljnji koraci u pipelineu, uključujući predprocesiranje, ekstrakciju značajki, treniranje modela i evaluaciju, oslanjaju se isključivo na informacije sadržane u CSV datoteci. Takav pristup omogućuje strogu kontrolu nad eksperimentalnim uvjetima i sprječava nenamjerno miješanje podataka iz različitih skupova.

Detaljna pravila podjele potpisa na skup za treniranje i skup za testiranje, kao i način generiranja CSV datoteke, opisani su u sljedećem potpoglavlju.

11.3. Generiranje CSV datoteke i podjela na skupove za treniranje i testiranje

Podjela podataka na skup za treniranje i skup za testiranje mora biti strogo definirana kako bi evaluacija sustava bila nepristrana i ponovljiva. U ovom radu podjela se provodi skriptom `make_csv.py`, koja generira CSV datoteku i time postavlja jedini autoritativni izvor informacija o tome koji uzorci pripadaju treningu, a koji testu. U svim kasnijim fazama (predprocesiranje, ekstrakcija značajki, treniranje i evaluacija) koristi se isključivo ova CSV datoteka, bez dodatnog oslanjanja na strukturu direktorija.

Pravila splita su definirana konfiguracijom u skripti. Odabire se 18 originalnih potpisa po osobi za treniranje, preostalih 6 originalnih potpisa ide u test, a svi krivotvoreni potpisi su uvijek u testu. Reproducibilnost je osigurana fiksnim seed-om.

```
# make_csv.py (CONFIG)
DATASET_DIR = Path("dataset55")
OUTPUT_CSV = Path("data.csv")
N_TRAIN_GENUINE = 18
SEED = 42
```

Struktura dataseta mapira se na `person_id` tako da direktoriji s krivotvorinama imaju sufiks `_forg`, ali se za potrebe CSV-a svode na isti identifikator osobe. Time se genuine i forged uzorci povezuju pod istim `person_id`, a klasa se vodi zasebno kroz stupac `label`.

```
# make_csv.py (mapping direktorija)
def is_forg_dir(dir_name: str) -> bool:
    return dir_name.endswith("_forg")

def person_id_from_dir(dir_name: str) -> str:
    # "001" -> "001", "001_forg" -> "001"
    return dir_name.replace("_forg", "")
```

Ključni dio implementacije je način na koji se biraju training genuine potpisi. Skripta koristi `random.Random(SEED)` kako bi odabir bio nasumičan, ali ponovljiv. Zatim se po osobi uzima točno `N_TRAIN_GENUINE` genuine potpisa za trening, ostali genuine idu u test, dok forged uvijek idu u test.

```
# make_csv.py (split logika)
rng = random.Random(SEED)

for pid, data in sorted(persons.items(), key=lambda x: x[0]):
    genuine = sorted(data["genuine"], key=lambda p: p.name)
    forged = sorted(data["forged"], key=lambda p: p.name)

    train_genuine = set(rng.sample(genuine, N_TRAIN_GENUINE))
    test_genuine = [p for p in genuine if p not in train_genuine]

    # Add genuine rows
    for p in genuine:
        split = "train" if p in train_genuine else "test"
        rows.append(
            {"image_path": p.as_posix(), "person_id": pid, "label":
"genuine", "split": split}
        )

    # Add forged rows (all test)
    for p in forged:
        rows.append(
            {"image_path": p.as_posix(), "person_id": pid, "label":
"forged", "split": "test"}
        )
```

Na ovaj način eksplicitno se osigurava da krivotvorine ne sudjeluju u treniranju modela, nego služe isključivo za evaluaciju. Također se uklanja rizik „curenja“ uzoraka između skupova, jer je pripadnost train/test definirana jednom, u CSV-u, i koristi se konzistentno u cijelom pipelineu.

CSV datoteka sadrži stupce `image_path`, `person_id`, `label`, `split`, što omogućuje transparentno filtriranje uzoraka u svim kasnijim skriptama.

12. PREDPROCESIRANJE SLIKA POTPISA

Predprocesiranje je obvezan dio implementiranog pipelinea i primjenjuje se nad svakom slikom potpisa prije ekstrakcije značajki i prije klasifikacije. Cilj predprocesiranja je standardizirati ulazne slike i smanjiti varijabilnost koja nije povezana s identitetom potpisnika, primjerice razlike u osvjetljenju, pozadini, kontrastu, prisutnom šumu i položaju potpisa na skenu. Time se osigurava da kasnije faze sustava, ekstrakcija HOG značajki i SVM klasifikacija, rade nad usporedivim ulazima.

U ovoj implementaciji predprocesiranje je dio jezgre sustava. Primjenjuje se jednako u fazi treniranja i u fazi verifikacije, bez iznimke. Parametri predprocesiranja pohranjuju se uz svaki istrenirani model, kako bi se tijekom verifikacije koristila identična konfiguracija kao tijekom treniranja.

12.1. Uloga predprocesiranja u sustavu

Uloga predprocesiranja u sustavu je osigurati da se iz različitih ulaznih slika dobije standardizirani prikaz potpisa na kojem se značajke mogu pouzdano izračunati. Bez standardizacije, isti potpis može generirati značajno različite značajke zbog pozadine, pomaka, različite debljine poteza ili razlike u intenzitetu piksela, što negativno utječe na stabilnost klasifikatora.

Implementacija predprocesiranja objedinjena je u funkciji `preprocess_signature` u skripti `preprocess.py`. Funkcija prima ulaznu sliku u BGR formatu (OpenCV), provodi definirane korake obrade i vraća predprocesiranu sliku fiksne veličine, uz opcionalne međukorake u `debug` strukturi koji služe za provjeru ispravnosti obrade.

```
# preprocess.py (ulazna točka predprocesiranja)
def preprocess_signature(
    img_bgr: np.ndarray,
    out_size=(220, 155),
    binarize_method="otsu",
    invert_if_needed=True,
    morph_open=True,
    open_ksize=2,
    pad=8,
```

```
):
    """
    Pipeline:
    1) grayscale
    2) binarization (otsu/adaptive)
    3) invert background if needed
    4) small morphology open to remove specks
    5) crop to bounding box of ink pixels
    6) pad and resize to out_size

    Returns:
    - processed grayscale image (out_size)
    - debug: dict with intermediate images (for preview/debug)
    """
```

Ključna dizajnerska odluka je da se predprocesiranje ne tretira kao „pomoćna obrada“, nego kao dio definicije modela. Zbog toga se konfiguracija predprocesiranja sprema uz model i ponovno koristi pri verifikaciji, čime se sprječava razlika između trening i verifikacijskog toka obrade.

Sljedeće potpoglavlje detaljno opisuje svaki korak predprocesiranja i prikazuje ključne isječke koda koji implementiraju binarizaciju, uklanjanje šuma, detekciju bounding boxa, padding i promjenu veličine slike.

12.2. Koraci predprocesiranja

Predprocesiranje u implementiranom sustavu sastoji se od niza uzastopnih koraka koji transformiraju ulaznu sliku potpisa u standardizirani prikaz pogodan za ekstrakciju HOG značajki. Redoslijed koraka je fiksan i primjenjuje se identično u fazi treniranja i u fazi verifikacije, čime se osigurava konzistentnost cijelog pipelinea.

Prvi korak predprocesiranja je konverzija ulazne slike u sivu skalu. Budući da informacija o boji nije relevantna za analizu oblika potpisa, konverzijom u grayscale smanjuje se dimenzionalnost podataka i uklanja nepotrebna varijabilnost.

```
# preprocess.py – konverzija u grayscale
gray = cv2.cvtColor(img_bgr, cv2.COLOR_BGR2GRAY)
```

Nakon konverzije u sivu skalu provodi se binarizacija slike, čime se potpis odvaja od pozadine. U implementaciji je omogućeno korištenje Otsuove metode ili adaptivne binarizacije, pri čemu se u ovom radu koristi Otsuova metoda zbog stabilnog ponašanja na skeniranim dokumentima.

```
# preprocess.py – binarizacija (Otsu)
_, bin_img = cv2.threshold(
    gray, 0, 255, cv2.THRESH_BINARY + cv2.THRESH_OTSU
)
```

Kako bi se osiguralo da su potezi potpisa predstavljeni bijelim pikselima na crnoj pozadini, provodi se provjera omjera bijelih i crnih piksela. Ako je pozadina dominantno bijela, slika se invertira. Ovaj korak je važan za ispravnu detekciju područja potpisa u kasnijim fazama.

```
# preprocess.py – invertiranje pozadine po potrebi
white_ratio = (bin_img == 255).mean()
if invert_if_needed and white_ratio > 0.5:
    bin_img = 255 - bin_img
```

Nakon binarizacije i eventualne inverzije, primjenjuje se morfološka operacija otvaranja s ciljem uklanjanja sitnog šuma i izoliranih piksela koji ne pripadaju potpisu. Time se postiže čišći binarni prikaz bez narušavanja strukture poteza potpisa.

```
# preprocess.py – uklanjanje šuma morfološkim otvaranjem
if morph_open:
    kernel = np.ones((open_ksize, open_ksize), np.uint8)
    bin_img = cv2.morphologyEx(bin_img, cv2.MORPH_OPEN, kernel)
```

Sljedeći korak je pronalaženje pravokutnog okvira koji obuhvaća sve piksele potpisa. Detekcijom koordinata piksela koji pripadaju potpisu izdvaja se minimalni bounding box, čime se uklanja nepotrebna pozadina i normalizira položaj potpisa unutar slike.

```
# preprocess.py – izračun bounding boxa
ys, xs = np.where(bin_img > 0)
y0, y1 = ys.min(), ys.max()
x0, x1 = xs.min(), xs.max()
crop = gray[y0:y1 + 1, x0:x1 + 1]
```

Kako bi se spriječio gubitak informacija na rubovima potpisa, izrezana slika se dodatno proširuje paddingom. Padding omogućuje stabilnije računanje značajki na rubnim dijelovima slike i smanjuje osjetljivost na male pomake.

```
# preprocess.py – padding
crop = cv2.copyMakeBorder(
    crop, pad, pad, pad, pad, cv2.BORDER_CONSTANT, value=0
)
```

U završnom koraku predprocesiranja slika se skalira na fiksnu veličinu od 220×155 piksela. Time se osigurava da sve slike imaju istu dimenziju, što je nužno za konzistentnu ekstrakciju HOG značajki i rad SVM klasifikatora.

```
# preprocess.py – promjena veličine slike
final_img = cv2.resize(crop, out_size, interpolation=cv2.INTER_AREA)
```

Rezultat predprocesiranja je standardizirana grayscale slika potpisa fiksne veličine, koja služi kao ulaz u fazu ekstrakcije značajki. Ovakav niz koraka omogućuje značajno smanjenje varijabilnosti ulaznih podataka, dok se istovremeno zadržavaju ključne strukturne informacije o obliku potpisa.

U sljedećem potpoglavlju detaljno je opisana implementacija predprocesiranja kao cjeline te način na koji se konfiguracija predprocesiranja sprema i ponovno koristi unutar sustava.

12.3. Implementacija predprocesiranja

Implementacija predprocesiranja objedinjena je u jednoj funkciji kako bi se osigurala konzistentna primjena istih koraka u svim fazama sustava. Funkcija `preprocess_signature` predstavlja jedinu ulaznu točku za obradu slike potpisa i koristi se identično tijekom treniranja modela i tijekom verifikacije pojedinačnog potpisa. Time se uklanja mogućnost razlika između trening i inferencijskog toka obrade.

Ključna dizajnerska odluka u implementaciji je razdvajanje konfiguracije predprocesiranja od samog poziva funkcije. Parametri poput metode binarizacije, veličine izlazne slike, paddinga i morfoloških operacija definirani su eksplicitno i spremaju se uz svaki istrenirani model. Tijekom verifikacije, ti se parametri učitavaju iz spremljenog modela i ponovno koriste bez izmjena.

U skripti `preprocess.py` funkcija vraća dvije vrijednosti. Prva je konačna predprocesirana slika fiksne veličine koja se koristi za ekstrakciju značajki. Druga je `debug` struktura koja sadrži međurezultate pojedinih koraka i služi isključivo za vizualnu provjeru tijekom razvoja sustava.

```
# preprocess.py – povratna vrijednost funkcije
debug = {
    "gray": gray,
    "binary": bin_img,
    "cropped": crop,
    "final": final_img,
}
return final_img, debug
```

Važno je naglasiti da se `debug` izlaz ne koristi u daljnjoj obradi niti sudjeluje u treniranju ili verifikaciji. On je korišten isključivo u razvojnoj fazi, primjerice kroz pomoćnu skriptu za pregled predprocesiranih slika, kako bi se vizualno potvrdila ispravnost pojedinih koraka.

U fazi treniranja modela predprocesiranje se primjenjuje nad svim slikama navedenim u CSV datoteci. Konfiguracija predprocesiranja pohranjuje se zajedno s modelom kako bi se tijekom verifikacije koristila identična konfiguracija. To se postiže spremanjem rječnika parametara uz model.

```
# hog_svm_train_eval.py – spremanje konfiguracije predprocesiranja uz model
```

```
model_pack = {
    "person_id": pid,
    "model": clf,
    "prep_cfg": prep_cfg,
    "hog_cfg": hog_cfg,
    "threshold": 0.0,
}
```

Tijekom verifikacije pojedinačnog potpisa, skripta `verify_one.py` učitava spremljeni model i dohvaća konfiguraciju predprocesiranja iz istog paketa. Time se osigurava da se nad ulaznom slikom primijeni potpuno isti niz operacija kao u fazi treniranja.

```
# verify_one.py – korištenje iste konfiguracije predprocesiranja
pack = joblib.load(model_path)
prep_cfg = pack["prep_cfg"]

proc_img, _ = preprocess_signature(img_bgr, **prep_cfg)
```

Ovakav pristup eliminira jedan od čestih izvora pogrešaka u biometrijskim sustavima, gdje se tijekom verifikacije nenamjerno koriste drugačiji parametri obrade nego tijekom treniranja. U implementiranom sustavu predprocesiranje je sastavni dio definicije modela, a ne zasebna, implicitna faza obrade.

Završetkom predprocesiranja dobiva se standardizirana slika potpisa spremna za ekstrakciju značajki. U sljedećem poglavlju opisuje se postupak ekstrakcije HOG značajki iz predprocesiranih slika i njihova uloga u klasifikacijskom dijelu sustava.

13. EKSTRAKCIJA ZNAČAJKI IZ POTPISA

Nakon završetka predprocesiranja, sljedeća faza implementiranog sustava je ekstrakcija značajki iz predprocesirane slike potpisa. Ekstrakcija značajki ima ključnu ulogu jer predstavlja poveznicu između sirove vizualne informacije i klasifikacijskog modela. Umjesto rada izravno nad pikselima slike, sustav koristi vektorsku reprezentaciju koja opisuje strukturne karakteristike potpisa na robusniji i kompaktniji način.

U ovom radu za ekstrakciju značajki korišten je Histogram of Oriented Gradients (HOG), koji se u literaturi često primjenjuje u offline verifikaciji potpisa. HOG omogućuje opis lokalne strukture slike kroz raspodjelu orijentacija gradijenata, čime se naglašavaju rubovi i potezi karakteristični za rukom pisani potpis.

Ekstrakcija značajki provodi se isključivo nad predprocesiranim slikama fiksne veličine, čime se osigurava da svi uzorci imaju istu dimenziju ulaznog vektora. Parametri HOG-a su fiksni i definirani prije treniranja modela, a njihova konfiguracija sprema se uz svaki istrenirani model kako bi se tijekom verifikacije koristila identična postavka.

13.1. Uloga značajki u verifikaciji potpisa

Uloga značajki u sustavu za verifikaciju potpisa je izdvajanje diskriminativnih informacija koje omogućuju razlikovanje originalnih potpisa od krivotvorina. Budući da je potpis bihevioralna biometrijska karakteristika s izraženom intra-class varijabilnošću, odabrane značajke moraju biti dovoljno osjetljive da uhvate individualne obrasce pisanja, ali istovremeno dovoljno robusne da toleriraju prirodne varijacije potpisa iste osobe.

Rad izravno nad pikselima slike potpisa često dovodi do loših rezultata zbog visoke osjetljivosti na šum, pomake i promjene u intenzitetu. Značajke poput HOG-a transformiraju lokalne promjene intenziteta u deskriptore koji bolje opisuju oblik i smjer poteza, što je posebno važno u offline scenariju gdje nisu dostupne dinamičke informacije o procesu potpisivanja.

U implementiranom sustavu značajke predstavljaju ulaz klasifikacijskom modelu i definiraju prostor u kojem se provodi razdvajanje originalnih potpisa i neautoriziranih uzoraka. Kvaliteta i stabilnost značajki izravno utječu na performanse SVM klasifikatora, uključujući kompromis između False Acceptance Rate i False Rejection Rate. Zbog toga je odabir HOG značajki, u kombinaciji s dosljednim predprocesiranjem, ključan preduvjet za pouzdanu verifikaciju potpisa u writer-dependent sustavu.

U sljedećem potpoglavlju detaljno su opisani parametri HOG značajki korišteni u implementaciji te njihov utjecaj na reprezentaciju potpisa.

13.2. Parametri HOG značajki

Kako bi HOG značajke bile konzistentne i usporedive među svim uzorcima, njihovi parametri moraju biti unaprijed definirani i fiksni kroz cijeli sustav. U implementiranom sustavu parametri HOG-a odabrani su tako da omoguće dovoljno detaljan opis lokalne strukture potpisa, uz razumnu dimenzionalnost vektora značajki.

U skripti `hog_svm_train_eval.py` parametri HOG-a definirani su kroz konfiguracijski rječnik koji se koristi tijekom treniranja modela i sprema se uz svaki istrenirani model. Time se osigurava da se tijekom verifikacije koristi identična konfiguracija ekstrakcije značajki.

```
# hog_svm_train_eval.py – konfiguracija HOG značajki
hog_cfg = {
    "orientations": 9,
    "pixels_per_cell": (8, 8),
    "cells_per_block": (2, 2),
    "block_norm": "L2-Hys",
    "transform_sqrt": True,
}
```

Parametar `orientations` definira broj diskretnih orijentacija gradijenata koje se računaju unutar svake ćelije. Vrijednost 9 predstavlja kompromis između osjetljivosti na smjer poteza i robusnosti na male varijacije u pisanju.

Parametar `pixels_per_cell` određuje veličinu osnovne ćelije nad kojom se računa histogram orijentacija. U ovoj implementaciji koristi se veličina 8×8 piksela, što omogućuje hvatanje lokalnih detalja poteza potpisa, a istovremeno ne fragmentira sliku na previše sitne dijelove.

Parametar `cells_per_block` definira broj ćelija koje čine jedan blok za normalizaciju. Korištenje blokova veličine 2×2 omogućuje lokalnu normalizaciju histograma, čime se smanjuje utjecaj promjena kontrasta i debljine poteza.

Normalizacija blokova provodi se metodom `L2-Hys`, koja se u praksi pokazala stabilnom za opisivanje rubnih struktura. Dodatno, parametar `transform_sqrt` omogućuje primjenu kvadratnog korijena nad intenzitetima prije izračuna gradijenata, čime se smanjuje utjecaj jakih intenzitetskih razlika i dodatno povećava robusnost značajki.

Ova konfiguracija HOG parametara odabrana je kako bi se naglasile strukturne karakteristike potpisa, poput smjera i kontinuiteta poteza, uz istovremeno smanjenje osjetljivosti na šum i varijacije koje nisu povezane s identitetom potpisnika. Definirani parametri koriste se dosljedno u cijelom sustavu i čine sastavni dio definicije svakog istreniranog modela.

U sljedećem potpoglavlju opisana je implementacija ekstrakcije HOG značajki i način na koji se dobiveni vektori koriste u procesu treniranja i verifikacije potpisa.

13.3. Implementacija ekstrakcije značajki

Ekstrakcija HOG značajki u implementiranom sustavu provodi se nad predprocesiranom slikom potpisa i predstavlja izravnu vezu između faze obrade slike i faze strojnog učenja. Implementacija je centralizirana u jednoj funkciji kako bi se osigurala dosljedna primjena istih parametara i postupaka tijekom treniranja modela i tijekom verifikacije pojedinačnog potpisa.

U skripti `hog_svm_train_eval.py` definirana je pomoćna funkcija `extract_hog`, koja prima predprocesiranu grayscale sliku i konfiguraciju HOG parametara te vraća jednodimenzionalni vektor značajki. Funkcija koristi implementaciju HOG-a iz biblioteke `scikit-image`, pri čemu se svi parametri eksplicitno prosljeđuju iz konfiguracijskog rječnika.

```
# hog_svm_train_eval.py – ekstrakcija HOG značajki
from skimage.feature import hog

def extract_hog(img_gray: np.ndarray, hog_cfg: dict) -> np.ndarray:
    feats = hog(
        img_gray,
        orientations=hog_cfg["orientations"],
        pixels_per_cell=hog_cfg["pixels_per_cell"],
        cells_per_block=hog_cfg["cells_per_block"],
        block_norm=hog_cfg["block_norm"],
        transform_sqrt=hog_cfg["transform_sqrt"],
        feature_vector=True,
```

```
    )
    return feats
```

Ekstrakcija značajki provodi se nad slikama koje su prethodno skalirane na fiksnu veličinu, čime se osigurava da svi uzorci generiraju vektore iste duljine. Dimenzionalnost prostora značajki time je konstantna i ne ovisi o izvornim dimenzijama slike potpisa.

U fazi treniranja, HOG značajke se računaju za sve uzorke navedene u CSV datoteci i koriste se za izgradnju matrice značajki. Kako bi se ubrzali eksperimenti i omogućilo višekratno treniranje modela nad istim skupom značajki, implementiran je mehanizam spremanja izračunatih HOG značajki u cache datoteku.

```
# hog_svm_train_eval.py – spremanje značajki u cache
np.savez_compressed(
    cache_path,
    X=X,
    image_path=img_paths,
    person_id=person_ids,
    label=labels,
    split=splits,
    prep_cfg=prep_cfg,
    hog_cfg=hog_cfg,
)
```

Spremanjem značajki u cache omogućuje se da se predprocesiranje i ekstrakcija značajki izvrše samo jednom, dok se kasniji eksperimenti s različitim konfiguracijama klasifikatora mogu provoditi bez ponovne obrade slika. Ovakav pristup značajno smanjuje vrijeme izvođenja eksperimenata i povećava preglednost procesa evaluacije.

Tijekom verifikacije pojedinačnog potpisa koristi se identična funkcija za ekstrakciju HOG značajki, pri čemu se konfiguracija `hog_cfg` učitava iz spremljenog modela. Time se osigurava da su značajke korištene u verifikaciji izračunate na potpuno isti način kao i one korištene tijekom treniranja.

```
# verify_one.py – ekstrakcija HOG značajki pri verifikaciji
hog_feats = extract_hog(proc_img, hog_cfg).reshape(1, -1)
```

Dosljedna implementacija ekstrakcije značajki u svim fazama sustava osigurava da razlike u odlukama klasifikatora proizlaze isključivo iz sadržaja potpisa, a ne iz razlika u načinu obrade podataka. Time je stvoren stabilan temelj za fazu treniranja writer-dependent SVM modela, koja je opisana u sljedećem poglavlju.

14. TRENIRANJE WRITER-DEPENDENT MODELA

Nakon predprocesiranja slika i ekstrakcije HOG značajki, sljedeća faza implementiranog sustava je treniranje klasifikacijskih modela za verifikaciju potpisa. U skladu s teorijskim dijelom rada, implementacija se temelji na writer-dependent pristupu, pri čemu se za svaku osobu trenira zaseban binarni klasifikator.

Cilj treniranja nije identifikacija potpisnika među više osoba, već verifikacija tvrdnje identiteta. Za svaki model klasifikacijski problem definiran je kao razdvajanje originalnih potpisa jedne osobe od potpisa koji toj osobi ne pripadaju. Takva formulacija problema omogućuje prilagodbu modela individualnim karakteristikama potpisa svake osobe, što je posebno važno u scenarijima s ograničenim brojem uzoraka po korisniku.

Treniranje se provodi isključivo u offline fazi sustava. Jednom istrenirani modeli spremaju se na disk i koriste se u fazi verifikacije bez dodatnog učenja ili prilagodbe. Time se osigurava stabilnost sustava i jasno razdvajanje faze učenja i faze primjene.

14.1. Writer-dependent pristup i formulacija klasifikacijskog problema

U writer-dependent sustavu za verifikaciju potpisa svaki korisnik ima vlastiti klasifikacijski model. Za osobu s identifikatorom `person_id` trenira se binarni klasifikator koji razlikuje potpise te osobe od potpisa svih ostalih osoba u skupu podataka. Pozitivnu klasu čine originalni potpisi ciljne osobe, dok negativnu klasu čine potpisi drugih osoba.

U implementiranom sustavu negativna klasa formira se isključivo od originalnih potpisa drugih osoba, dok se krivotvorine ne koriste u fazi treniranja. Krivotvoreni potpisi koriste se samo u fazi testiranja i evaluacije, čime se simulira realističan scenarij primjene sustava u kojem krivotvorine nisu dostupne tijekom učenja modela.

Treniranje modela provodi se iterativno za svaku osobu u skupu podataka. U skripti `hog_svm_train_eval.py` implementirana je petlja koja prolazi kroz sve jedinstvene identifikatore osoba i za svaku osobu konstruira odgovarajući trening skup te trenira zaseban klasifikator.

```
# hog_svm_train_eval.py – iteracija po osobama (writer-dependent trening)
for pid in persons:
    # izdvajanje pozitivnih i negativnih uzoraka
    ...
    clf.fit(X_train, y_train)
```

Formulacija klasifikacijskog problema na ovaj način omogućuje da svaki model nauči specifične obrasce potpisa jedne osobe, umjesto da pokušava generalizirati preko svih potpisnika. Time se postiže veća osjetljivost na individualne karakteristike potpisa, ali se istovremeno povećava broj modela koje sustav mora održavati.

Writer-dependent pristup također omogućuje fleksibilnu evaluaciju performansi po osobi i agregaciju rezultata na razini cijelog sustava. U kasnijim poglavljima prikazano je kako se rezultati pojedinačnih modela kombiniraju kako bi se dobila ukupna slika učinkovitosti sustava.

U sljedećem potpoglavlju detaljno je opisan postupak formiranja pozitivnih i negativnih uzoraka za treniranje svakog modela te način na koji se kontrolira omjer klasa u trening skupu.

14.2. Formiranje pozitivnih i negativnih uzoraka

Formiranje trening skupa za svaki writer-dependent model predstavlja jednu od ključnih implementacijskih odluka u sustavu za verifikaciju potpisa. Budući da se za svaku osobu trenira zaseban binarni klasifikator, potrebno je jasno definirati koje uzorke čine pozitivnu, a koje negativnu klasu, kao i način kontrole njihove zastupljenosti u trening skupu.

Pozitivnu klasu čine originalni potpisi ciljne osobe koji su, prema CSV datoteci, označeni kao pripadnici skupa za treniranje. Broj pozitivnih uzoraka po osobi izravno ovisi o konfiguraciji dataseta i u ovom radu varira ovisno o eksperimentalnoj postavci. U konačno odabranoj konfiguraciji korišteno je 18 originalnih potpisa po osobi za treniranje.

Negativna klasa formira se od originalnih potpisa svih ostalih osoba u skupu podataka, koji su također označeni kao pripadnici trening skupa. Krivotvoreni potpisi namjerno su isključeni iz treniranja kako bi se izbjeglo učenje specifičnih obrazaca krivotvorina i zadržao realističan scenarij u kojem sustav nema pristup krivotvorinama tijekom faze učenja.

U skripti `hog_svm_train_eval.py` pozitivni i negativni uzorci izdvajaju se na temelju informacija iz cache-a značajki i pripadajućih oznaka. Za svaku osobu prvo se identificiraju

indeksi pozitivnih uzoraka, a zatim se iz skupa svih ostalih trening uzoraka formira bazen negativnih primjera.

```
# hog_svm_train_eval.py – izdvajanje pozitivnih i negativnih uzoraka
pos_train_idx = train_genuine_idx[train_genuine_pid == pid]

neg_pool_idx = train_genuine_idx[train_genuine_pid != pid]
```

Kako bi se spriječila prevelika neravnoteža između pozitivne i negativne klase, implementiran je mehanizam kontrole broja negativnih uzoraka. Broj negativnih primjera koji se koriste u treniranju definira se kao višekratnik broja pozitivnih uzoraka, putem parametra `neg_ratio`. Na taj način se osigurava da negativna klasa bude dovoljno reprezentativna, ali ne dominira trening skupom.

```
# hog_svm_train_eval.py – kontrola omjera negativnih uzoraka
n_pos = len(pos_train_idx)
n_neg = int(round(neg_ratio * n_pos))

rng = np.random.default_rng(seed)
neg_train_idx = rng.choice(neg_pool_idx, size=n_neg, replace=False)
```

Nakon odabira pozitivnih i negativnih uzoraka formira se konačni trening skup koji se sastoji od odgovarajućih vektora značajki i pripadajućih oznaka klasa. Pozitivni uzorci označeni su vrijednošću 1, dok su negativni uzorci označeni vrijednošću 0.

```
# hog_svm_train_eval.py – formiranje trening skupa
X_train = np.vstack([X[pos_train_idx], X[neg_train_idx]])
y_train = np.hstack([np.ones(len(pos_train_idx)),
np.zeros(len(neg_train_idx))])
```

Ovakav način formiranja trening skupa omogućuje fleksibilno upravljanje složenošću klasifikacijskog problema i prilagodbu modela dostupnoj količini podataka po osobi. Ujedno se osigurava da svaki writer-dependent model uči razlikovati potpise ciljne osobe od reprezentativnog skupa potpisa drugih osoba, bez oslanjanja na specifične obrasce krivotvorina.

U sljedećem potpoglavlju opisan je korišteni klasifikator, njegova integracija u pipeline te odabrani hiperparametri za linearni i RBF SVM.

14.3. SVM klasifikator i pipeline treniranja

Za klasifikaciju potpisa u implementiranom sustavu korišten je Support Vector Machine (SVM), koji se u teorijskom dijelu rada pokazao prikladnim za probleme verifikacije potpisa s ograničenim brojem uzoraka po osobi. U skladu s time, u implementaciji su ispitane dvije varijante SVM klasifikatora, linearni i RBF kernel, kako bi se analizirao utjecaj nelinearnosti modela na performanse sustava.

SVM klasifikator integriran je u cjeloviti pipeline zajedno sa standardizacijom značajki. Budući da HOG značajke mogu imati različite raspone vrijednosti, prije klasifikacije primjenjuje se standardizacija pomoću `StandardScaler`. Time se osigurava da sve značajke imaju usporedivu skalu, što je posebno važno za stabilno ponašanje SVM-a, osobito kod RBF kernela.

U skripti `hog_svm_train_eval.py` pipeline se definira korištenjem `Pipeline` klase iz biblioteke `scikit-learn`, čime se objedinuju svi koraci potrebni za treniranje i kasniju primjenu modela.

```
# hog_svm_train_eval.py – definicija SVM pipelinea
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import StandardScaler
from sklearn.svm import SVC

clf = Pipeline(
    steps=[
        ("scaler", StandardScaler()),
        (
            "svm",
            SVC(
                kernel=kernel,
                C=1.0,
                gamma=gamma,
```

```
                class_weight="balanced",
                probability=False,
            ),
        ),
    ]
)
```

Parametar `kernel` omogućuje odabir između linearnog i RBF kernela, ovisno o eksperimentalnoj konfiguraciji. Linearni kernel koristi se kao stabilan baseline u scenarijima s malim brojem trening uzoraka, dok RBF kernel omogućuje nelinearno razdvajanje u prostoru značajki kada je dostupna veća količina podataka po osobi.

Parametar regularizacije `C` fiksiran je na vrijednost 1.0 u svim eksperimentima kako bi se smanjio broj slobodnih hiperparametara i osigurala usporedivost rezultata između različitih konfiguracija. Za RBF kernel koristi se zadana vrijednost parametra `gamma`, definirana kao `"scale"`, čime se automatski prilagođava raspon značajki.

Zbog moguće neravnoteže između pozitivne i negativne klase, posebno u writer-dependent scenariju, postavljen je parametar `class_weight="balanced"`. Time se pogreške na slabije zastupljenoj klasi dodatno penaliziraju, što doprinosi stabilnijem učenju modela.

Nakon definiranja pipelinea, model se trenira nad prethodno formiranim trening skupom značajki i oznaka klasa.

```
# hog_svm_train_eval.py – treniranje SVM modela
clf.fit(X_train, y_train)
```

Ovako definiran pipeline osigurava da se tijekom treniranja i kasnije tijekom verifikacije nad novim uzorcima primjenjuju identični koraci standardizacije i klasifikacije. Time se dodatno smanjuje mogućnost nekonzistentnosti u obradi podataka i osigurava pouzdano ponašanje sustava.

U sljedećem potpoglavlju opisan je postupak spremanja istreniranih modela i pripadajućih konfiguracija, koji omogućuje njihovu ponovnu upotrebu u fazi verifikacije pojedinačnih potpisa.

14.4. Spremanje modela i konfiguracije

Nakon uspješnog treniranja SVM klasifikatora za pojedinu osobu, istrenirani model i pripadajuća konfiguracija spremaju se na disk kako bi se mogli koristiti u fazi verifikacije bez potrebe za ponovnim učenjem. Ovakav pristup omogućuje jasno razdvajanje faze treniranja i faze primjene sustava te osigurava stabilnost i ponovljivost rezultata.

U implementiranom sustavu za svaku osobu sprema se zasebna datoteka modela u formatu `.joblib`. Naziv datoteke sadrži identifikator osobe, čime se omogućuje jednostavno dohvaćanje odgovarajućeg modela tijekom verifikacije pojedinačnog potpisa. Svaki spremljeni model sadrži ne samo istrenirani klasifikator, već i sve parametre potrebne za konzistentnu obradu ulaznih podataka.

Prije spremanja modela formira se rječnik koji objedinjuje sve relevantne komponente sustava. U taj se rječnik pohranjuju identifikator osobe, istrenirani SVM pipeline, konfiguracija predprocesiranja, konfiguracija HOG značajki te prag odlučivanja. Prag odlučivanja u ovom radu fiksiran je na vrijednost 0.0, čime se odluka temelji izravno na znaku vrijednosti `decision_function`.

```
# hog_svm_train_eval.py – paket za spremanje modela
model_pack = {
    "person_id": pid,
    "model": clf,
    "prep_cfg": prep_cfg,
    "hog_cfg": hog_cfg,
    "threshold": 0.0,
}
```

Nakon formiranja paketa, model se sprema na disk korištenjem biblioteke `joblib`. Time se omogućuje učinkovita serializacija objekata koji sadrže i numeričke podatke i Python objekte, poput SVM klasifikatora i konfiguracijskih rječnika.

```
# hog_svm_train_eval.py – spremanje modela na disk
model_path = models_dir / f"svm_person_{pid}.joblib"
joblib.dump(model_pack, model_path)
```

Tijekom faze verifikacije, skripta `verify_one.py` učitava odgovarajući `.joblib` model na temelju odabranog identifikatora osobe. Iz učitanog paketa dohvaćaju se svi potrebni parametri, čime se osigurava da se predprocesiranje i ekstrakcija značajki provode s istom konfiguracijom kao tijekom treniranja.

Ovakav način spremanja modela omogućuje jednostavnu proširivost sustava, jer se dodavanje nove osobe svodi na treniranje i spremanje novog modela bez potrebe za izmjenom postojećih modela. Istovremeno se osigurava da je svaki writer-dependent model potpuno samostalan i konzistentan s definicijom sustava.

Završetkom ovog potpoglavlja kompletiran je proces treniranja modela. U sljedećem poglavlju opisan je postupak verifikacije pojedinačnog potpisa, koji koristi prethodno spremljene modele za donošenje odluke o prihvaćanju ili odbacivanju tvrdnje identiteta.

15. VERIFIKACIJA POJEDINAČNOG POTPISA

Nakon treniranja i spremanja writer-dependent modela, sustav omogućuje verifikaciju pojedinačnog potpisa za unaprijed odabranu osobu. Verifikacija predstavlja fazu primjene sustava u kojoj se nad novim, neviđenim potpisom donosi odluka o prihvaćanju ili odbacivanju tvrdnje identiteta, bez ikakvog dodatnog učenja ili prilagodbe modela.

Verifikacija je implementirana kao strogo deterministički postupak koji koristi isključivo informacije pohranjene u spremljenom modelu. Ulaz u sustav čini digitalna slika potpisa i identifikator osobe čiji se identitet provjerava, dok je izlaz binarna odluka ACCEPT ili REJECT, uz pripadajuću numeričku vrijednost rezultata klasifikatora. Time se osigurava da se ponašanje sustava tijekom verifikacije u potpunosti podudara s konfiguracijom korištenom tijekom treniranja.

Cjelokupni postupak verifikacije implementiran je u skripti `verify_one.py`, koja objedinjuje učitavanje modela, predprocesiranje ulazne slike, ekstrakciju značajki, izračun rezultata klasifikatora i donošenje odluke. Na taj način verifikacija se provodi kroz isti pipeline kao i treniranje, ali bez faze učenja.

15.1. Postupak verifikacije i učitavanje modela

Postupak verifikacije započinje učitavanjem writer-dependent modela koji odgovara osobi čiji se identitet provjerava. Model se dohvaća na temelju identifikatora osobe, a iz `.joblib` datoteke učitavaju se svi potrebni elementi za obradu potpisa: istrenirani klasifikator, konfiguracija predprocesiranja, konfiguracija HOG značajki i prag odlučivanja.

```
# verify_one.py – učitavanje spremljenog modela
pack = joblib.load(model_path)

clf = pack["model"]
prep_cfg = pack["prep_cfg"]
hog_cfg = pack["hog_cfg"]
threshold = pack.get("threshold", 0.0)
```

Nakon učitavanja modela, ulazna slika potpisa učitava se pomoću OpenCV-a i prosljeđuje funkciji za predprocesiranje. Važno je naglasiti da se tijekom verifikacije koristi ista funkcija `preprocess_signature` i ista konfiguracija predprocesiranja koja je spremljena uz model tijekom treniranja.

```
# verify_one.py – predprocesiranje ulazne slike
img_bgr = cv2.imread(image_path)
proc_img, _ = preprocess_signature(img_bgr, **prep_cfg)
```

Predprocesirana slika zatim se koristi za ekstrakciju HOG značajki, pri čemu se ponovno koristi konfiguracija HOG parametara spremljena uz model. Dobiveni vektor značajki oblikuje se u odgovarajući format za ulaz u SVM klasifikator.

```
# verify_one.py – ekstrakcija značajki
hog_feats = extract_hog(proc_img, hog_cfg).reshape(1, -1)
```

Na ovaj način osigurava se da su svi koraci obrade ulaznog potpisa tijekom verifikacije identični onima korištenima tijekom treniranja. Takva dosljednost ključna je za ispravno funkcioniranje sustava, jer se svaka razlika u obradi mogla negativno odraziti na pouzdanost odluke klasifikatora.

U sljedećem potpoglavlju opisan je postupak donošenja odluke na temelju rezultata SVM klasifikatora, uključujući interpretaciju vrijednosti `decision_function` i primjenu fiksnog praga odlučivanja.

15.2. Donošenje odluke i prag odlučivanja

Nakon što su iz ulazne slike potpisa izračunate HOG značajke, sustav pristupa donošenju odluke o prihvaćanju ili odbacivanju tvrdnje identiteta. Odluka se temelji isključivo na izlazu SVM klasifikatora, bez ikakve dodatne heuristike ili naknadne prilagodbe.

U implementiranom sustavu koristi se metoda `decision_function`, koja vraća realnu vrijednost koja predstavlja udaljenost uzorka od razdjelne hiperravnine klasifikatora. Predznak te vrijednosti određuje klasu, dok njezina apsolutna vrijednost odražava sigurnost

odluke. Pozitivna vrijednost označava pripadnost pozitivnoj klasi, odnosno originalnom potpisu ciljne osobe, dok negativna vrijednost označava nepripadnost toj klasi.

```
# verify_one.py – izračun decision score-a
score = float(clf.decision_function(hog_feats)[0])
```

Prag odlučivanja u sustavu fiksiran je na vrijednost 0.0. Ovakav odabir praga u skladu je s teorijskom definicijom SVM-a, gdje je razdjelna granica upravo definirana nul-tom vrijednošću `decision_function`. Time se izbjegava dodatna optimizacija praga na testnim podacima i osigurava konzistentno ponašanje sustava u svim eksperimentalnim postavkama.

```
# verify_one.py – donošenje binarne odluke
decision = "ACCEPT" if score >= threshold else "REJECT"
```

Fiksiranjem praga odlučivanja pojednostavljuje se interpretacija rezultata i olakšava usporedba različitih konfiguracija sustava. Promjene u performansama, izražene kroz FAR i FRR, tada su izravna posljedica odabira značajki, količine trening podataka i vrste klasifikatora, a ne dodatne prilagodbe praga.

Važno je naglasiti da dobiveni `score` nije vjerojatnost. On ne predstavlja stupanj sigurnosti u probabilističkom smislu, već mjeru udaljenosti od granice razdvajanja. Veća apsolutna vrijednost rezultata upućuje na veću sigurnost odluke, ali se ne može izravno interpretirati kao postotak pouzdanosti.

Ovakav način donošenja odluke omogućuje transparentno i deterministično ponašanje sustava, što je posebno važno u kontekstu biometrijske verifikacije. U sljedećem potpoglavlju detaljnije se razmatra interpretacija decision score-a te način na koji se rezultat prikazuje korisniku u demonstracijskoj web aplikaciji.

15.3. Interpretacija decision score-a i prikaz rezultata verifikacije

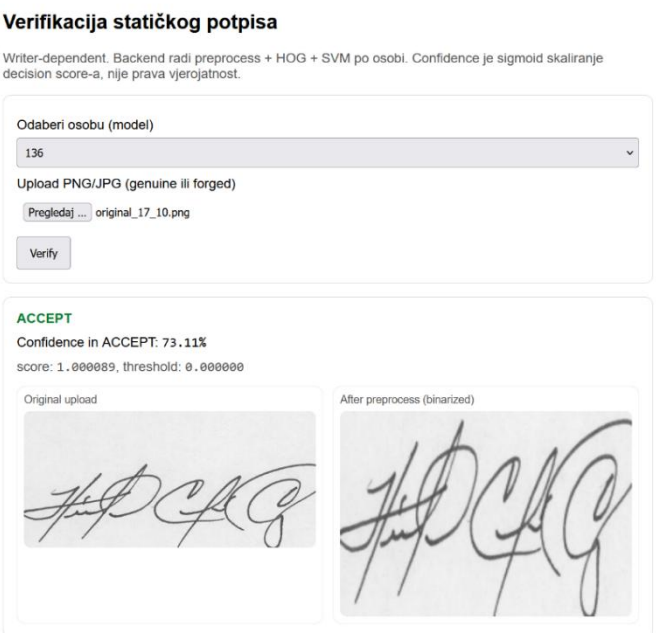
Vrijednost dobivena metodom `decision_function` predstavlja kontinuiranu mjeru udaljenosti uzorka od razdjelne granice SVM klasifikatora. U implementiranom sustavu ta se vrijednost koristi kao osnovni indikator sigurnosti odluke, pri čemu se znak vrijednosti koristi za donošenje binarne odluke, a apsolutna vrijednost za procjenu jačine te odluke.

Pozitivna vrijednost decision score-a označava da se potpis nalazi na strani pozitivne klase, odnosno da ga model smatra originalnim potpisom ciljne osobe. Negativna vrijednost označava da se potpis nalazi na strani negativne klase i da se tvrdnja identiteta odbacuje. Što je apsolutna vrijednost score-a veća, to je potpis udaljeniji od razdjelne granice i odluka klasifikatora sigurnija.

Važno je naglasiti da decision score ne predstavlja vjerojatnost u statističkom smislu. Iako se u demonstracijskoj web aplikaciji koristi jednostavna sigmoidna transformacija kako bi se korisniku prikazala intuitivna mjera „sigurnosti“, takav prikaz služi isključivo informativnoj svrsi i ne predstavlja stvarnu probabilističku procjenu. Odluka sustava uvijek se temelji isključivo na usporedbi decision score-a s fiksnim pragom 0.0.

U demonstracijskoj web aplikaciji rezultat verifikacije prikazuje se korisniku zajedno s ključnim informacijama o odluci: binarnim ishodom (ACCEPT ili REJECT), numeričkom vrijednošću decision score-a, korištenim pragom odlučivanja te vizualnim prikazom ulazne i predprocesirane slike potpisa. Time se omogućuje transparentan uvid u način rada sustava i olakšava interpretacija pojedinačnih slučajeva verifikacije.

Na slici u nastavku prikazan je primjer uspješne verifikacije, u kojem je decision score pozitivan i iznad praga odlučivanja.



Slika 1: Primjer uspješne verifikacije

U ovom primjeru potpis se nalazi dovoljno daleko od razdjelne granice na strani pozitivne klase, što rezultira prihvaćanjem tvrdnje identiteta. Prikaz predprocesirane slike omogućuje dodatnu vizualnu provjeru da je potpis ispravno izdvojen i normaliziran prije ekstrakcije značajki.

Na sljedećoj slici prikazan je primjer odbijanja tvrdnje identiteta, gdje je decision score negativan.

Slika 2: Odbijanje tvrdnje identiteta

U ovom slučaju potpis se nalazi na strani negativne klase, što rezultira odlukom REJECT. Takvi primjeri uključuju i kvalitetne krivotvorine koje, unatoč vizualnoj sličnosti s originalnim potpisom, sustav prepoznaje kao nepripadajuće ciljnoj osobi. Ovakvo ponašanje sustava u skladu je s izmjerenim vrijednostima False Acceptance Rate i predstavlja realno ograničenje biometrijskih verifikacijskih sustava.

Prikaz decision score-a i predprocesirane slike u web aplikaciji ima isključivo demonstracijsku i edukativnu ulogu te ne utječe na samu odluku sustava. Time se osigurava jasna razlika između jezgre verifikacijskog algoritma i korisničkog sučelja, koje služi za ilustraciju rada implementiranog sustava.

Završetkom ovog poglavlja opisan je kompletan postupak verifikacije pojedinačnog potpisa. U sljedećem poglavlju razmatra se evaluacija sustava na testnom skupu, uključujući korištene metrike i analizu dobivenih rezultata.

16. EVALUACIJA SUSTAVA VERIFIKACIJE

Nakon implementacije cjelokupnog pipelinea za treniranje i verifikaciju potpisa, provedena je sustavna evaluacija performansi sustava. Cilj evaluacije je kvantitativno procijeniti sposobnost sustava da ispravno prihvati originalne potpise i odbaci neautorizirane uzorke, kao i analizirati kompromis između sigurnosti i pouzdanosti verifikacije.

Evaluacija se provodi isključivo nad testnim skupom definiranim u CSV datoteci. Tijekom evaluacije ne koristi se nijedan uzorak koji je sudjelovao u treniranju modela, čime se osigurava nepristrana procjena performansi. Posebna pažnja posvećena je činjenici da se u writer-dependent sustavu evaluacija mora promatrati na razini pojedinačnih osoba, ali i agregirano na razini cijelog sustava.

U implementiranom sustavu evaluacija obuhvaća izračun standardnih biometrijskih metrika, analizu distribucije decision score-ova, generiranje ROC i Precision–Recall krivulja te izradu matrice zabune za odabrane konfiguracije. Ovakav skup analiza omogućuje detaljan uvid u ponašanje sustava u različitim scenarijima i pri različitim postavkama modela.

16.1. Metodologija evaluacije

Metodologija evaluacije u ovom radu temelji se na writer-dependent pristupu i provodi se po osobi, a zatim agregirano nad svim osobama u testnom skupu. Za svaku osobu evaluira se odgovarajući binarni klasifikator korištenjem testnih uzoraka definirane u CSV datoteci.

Testni skup za pojedinu osobu sastoji se od dva tipa uzoraka: originalnih potpisa te osobe koji nisu korišteni u treniranju i svih krivotvorenih potpisa koji pripadaju toj osobi. Originalni potpisi u testnom skupu služe za procjenu stope pogrešnog odbacivanja (False Rejection Rate), dok krivotvorine služe za procjenu stope pogrešnog prihvaćanja (False Acceptance Rate).

Za svaki testni uzorak provodi se postupak verifikacije identičan onome opisanom u prethodnom poglavlju. Izračunava se decision score korištenjem metode `decision_function`, primjenjuje se fiksni prag odlučivanja od 0.0 te se bilježi ishod verifikacije. Dobiveni ishodi uspoređuju se s stvarnim oznakama uzoraka kako bi se formirala matrica zabune za svaku osobu.

U implementaciji evaluacije rezultati se bilježe po osobi, uključujući broj točnih i netočnih odluka, te se zatim agregiraju kako bi se dobila ukupna procjena performansi sustava. Agregacija se provodi na razini uzoraka, čime se osigurava da svaki testni potpis ima jednak utjecaj na ukupne metrike, neovisno o pripadnosti pojedinoj osobi.

Ovakva metodologija evaluacije omogućuje analizu performansi sustava u realističnom scenariju verifikacije potpisa, gdje se odluke donose neovisno za svaku tvrdnju identiteta, a ukupna učinkovitost sustava proizlazi iz ponašanja svih writer-dependent modela zajedno.

U sljedećem potpoglavlju opisane su korištene evaluacijske metrike te način njihovog izračuna u implementiranom sustavu.

16.2. Evaluacijske metrike

Za kvantitativnu procjenu performansi sustava korišten je skup standardnih biometrijskih metrika koje omogućuju analizu ponašanja sustava iz perspektive sigurnosti i pouzdanosti. U implementiranom sustavu metrike se računaju na temelju matrice zabune dobivene nad testnim skupom, pri čemu se odluke donose korištenjem fiksnog praga odlučivanja od 0.0.

Osnovu evaluacije čini matrica zabune s četiri vrijednosti: True Positive (TP), True Negative (TN), False Positive (FP) i False Negative (FN). U kontekstu verifikacije potpisa, TP predstavlja ispravno prihvaćen originalni potpis, TN ispravno odbačenu krivotvorinu, FP pogrešno prihvaćenu krivotvorinu, a FN pogrešno odbačeni originalni potpis.

U skripti `hog_svm_train_eval.py` matrica zabune računa se na temelju stvarnih oznaka testnih uzoraka i predviđenih odluka klasifikatora.

```
# hog_svm_train_eval.py – matrica zabune
from sklearn.metrics import confusion_matrix

tn, fp, fn, tp = confusion_matrix(y_true, y_pred).ravel()
```

Na temelju vrijednosti matrice zabune izračunavaju se ključne evaluacijske metrike. Stopa pogrešnog prihvaćanja (False Acceptance Rate, FAR) definira se kao omjer pogrešno prihvaćenih krivotvorina i ukupnog broja krivotvorina u testnom skupu. Ova metrika izravno odražava sigurnost sustava.

```
# hog_svm_train_eval.py – FAR
far = fp / (fp + tn) if (fp + tn) > 0 else 0.0
```

Stopa pogrešnog odbacivanja (False Rejection Rate, FRR) mjeri udio originalnih potpisa koji su pogrešno odbačeni. FRR predstavlja mjeru pouzdanosti sustava iz perspektive legitimnog korisnika.

```
# hog_svm_train_eval.py – FRR
frr = fn / (fn + tp) if (fn + tp) > 0 else 0.0
```

Uz FAR i FRR, računa se i ukupna točnost sustava (Accuracy), koja predstavlja omjer ispravnih odluka i ukupnog broja testnih uzoraka. Iako se točnost često koristi kao opća metrika uspješnosti, u kontekstu biometrijske verifikacije ona se interpretira s oprezom jer ne razlikuje vrste pogrešaka.

```
# hog_svm_train_eval.py – Accuracy
accuracy = (tp + tn) / (tp + tn + fp + fn)
```

Metrike se u implementaciji najprije računaju za svaki writer-dependent model pojedinačno, a zatim se agregiraju nad svim osobama kako bi se dobila ukupna procjena performansi sustava. Takav pristup omogućuje uvid u ponašanje sustava na razini pojedinačnih korisnika, ali i cjelokupnog sustava.

Korištenjem kombinacije FAR i FRR moguće je analizirati kompromis između sigurnosti i upotrebljivosti sustava, dok ukupna točnost služi kao dodatni indikator opće uspješnosti. U sljedećem potpoglavlju proširuje se evaluacija analizom ROC i Precision–Recall krivulja, koje omogućuju detaljniji uvid u ponašanje sustava u odnosu na decision score.

16.3. ROC i Precision–Recall analiza

Uz osnovne evaluacijske metrike, ponašanje sustava dodatno je analizirano pomoću ROC i Precision–Recall krivulja. Ove krivulje omogućuju dublji uvid u način na koji se sustav ponaša pri različitim vrijednostima decision score-a, neovisno o fiksnom pragu odlučivanja korištenom u konačnoj konfiguraciji.

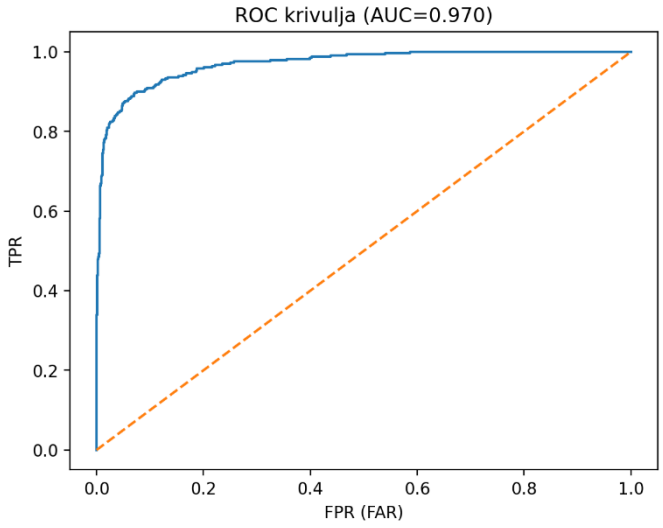
ROC krivulja (Receiver Operating Characteristic) prikazuje odnos između True Positive Rate (TPR) i False Positive Rate (FPR) pri promjeni praga odlučivanja. U kontekstu verifikacije potpisa, TPR odgovara udjelu ispravno prihvaćenih originalnih potpisa, dok FPR odgovara udjelu pogrešno prihvaćenih krivotvorina. Površina ispod ROC krivulje (AUC) služi kao agregirana mjera razdvojivosti klasa u prostoru značajki.

U implementiranom sustavu ROC krivulja računa se na temelju decision score-ova dobivenih nad svim testnim uzorcima, agregirano preko svih osoba. Izračun i iscrtavanje krivulje implementirani su u skripti `plot_roc_pr.py`.

```
# plot_roc_pr.py – ROC krivulja
from sklearn.metrics import roc_curve, roc_auc_score

fpr, tpr, _ = roc_curve(y_true, scores)
auc = roc_auc_score(y_true, scores)
```

Dobivena ROC krivulja pokazuje izrazito dobru razdvojivost između originalnih potpisa i krivotvorina, s vrijednošću AUC približno 0.97. Takav rezultat upućuje na to da sustav u velikoj mjeri uspijeva rangirati originalne potpise iznad krivotvorina, čak i prije primjene fiksnog praga odlučivanja.



Slika 3: ROC krivulja

Precision–Recall (PR) krivulja pruža dodatni uvid u ponašanje sustava, osobito u uvjetima neravnoteže klasa, što je čest slučaj u biometrijskim sustavima. Precision predstavlja udio ispravno prihvaćenih originalnih potpisa među svim prihvaćenim uzorcima, dok Recall

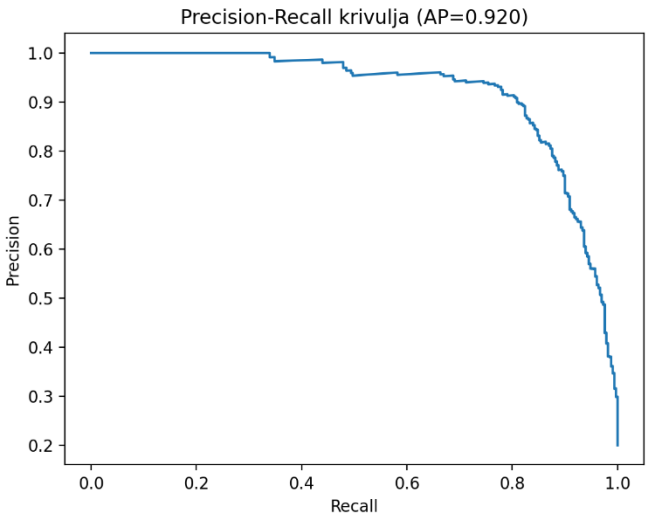
odgovara udjelu ispravno prihvaćenih originalnih potpisa među svim originalnim potpisima u testnom skupu.

PR krivulja u implementiranom sustavu također se računa na temelju decision score-ova, a kao agregirana mjera koristi se Average Precision (AP).

```
# plot_roc_pr.py – Precision–Recall krivulja
from sklearn.metrics import precision_recall_curve, average_precision_score

precision, recall, _ = precision_recall_curve(y_true, scores)
ap = average_precision_score(y_true, scores)
```

Dobivena vrijednost Average Precision iznosi približno 0.92, što ukazuje na visoku preciznost sustava u širokom rasponu vrijednosti recall-a. Takav rezultat potvrđuje da sustav u velikoj mjeri izbjegava pogrešno prihvaćanje krivotvorina, uz razumnu stopu prihvaćanja originalnih potpisa.



Slika 4: Precision - Recall krivulja

Zajednička analiza ROC i Precision–Recall krivulja potvrđuje da odabrana konfiguracija sustava ostvaruje dobar kompromis između sigurnosti i pouzdanosti. Visoka AUC vrijednost pokazuje snažnu razdvojivost klasa, dok visoka Average Precision vrijednost potvrđuje stabilno ponašanje sustava u uvjetima neravnoteže klasa. Ovi rezultati dodatno opravdavaju odabir RBF SVM-a s dovoljnim brojem originalnih potpisa po osobi kao konačne konfiguracije sustava.

U sljedećem poglavlju daje se sažeta analiza eksperimentalnih konfiguracija i ključnih zaključaka implementacijskog dijela rada.

16.4. Matrica zabune i analiza pogrešaka

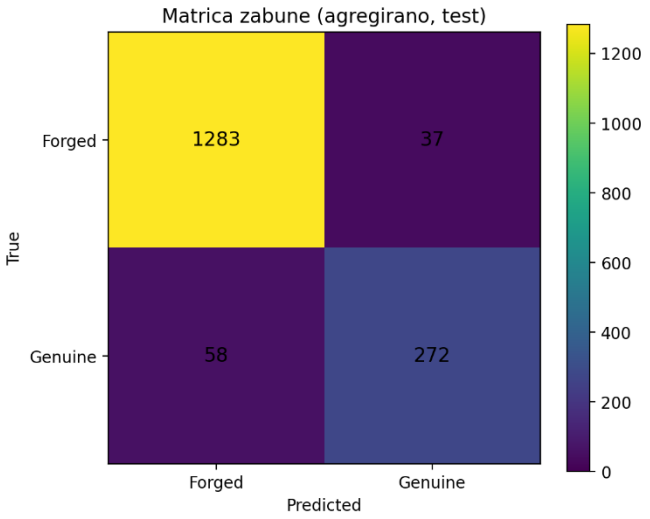
Za dodatnu interpretaciju performansi sustava koristi se matrica zabune (confusion matrix), koja prikazuje broj točnih i netočnih odluka sustava na testnom skupu. Matrica zabune omogućuje izravno sagledavanje dviju ključnih vrsta pogrešaka u biometrijskoj verifikaciji, pogrešno prihvaćanje krivotvorina (FP) i pogrešno odbacivanje originalnih potpisa (FN).

U implementaciji se matrica zabune računa na temelju stvarnih oznaka testnih uzoraka i predikcija dobivenih primjenom fiksnog praga odlučivanja (0.0) nad decision score-ovima. Izračun se provodi pomoću funkcije `confusion_matrix` iz biblioteke `scikit-learn`, pri čemu se matrica u agregiranom obliku dobiva kombiniranjem rezultata svih osoba u testnom skupu.

```
# hog_svm_train_eval.py - matrica zabune (agregirano)
from sklearn.metrics import confusion_matrix

tn, fp, fn, tp = confusion_matrix(y_true, y_pred).ravel()
```

U konačno odabranoj konfiguraciji sustava (CEDAR, 55 osoba, 18 genuine za trening, RBF SVM) dobivena je sljedeća matrica zabune nad agregiranim testnim skupom:



Slika 5: Matrica zabune

Matrica zabune prikazana na slici predstavlja agregirane rezultate verifikacije nad testnim skupom za konačnu konfiguraciju sustava (CEDAR, 55 osoba, 18 genuine potpisa za treniranje, RBF SVM, prag 0.0).

Vrijednosti matrice zabune su sljedeće:

- True Negative (TN) = 1283
- False Positive (FP) = 37
- False Negative (FN) = 58
- True Positive (TP) = 272

Ove vrijednosti omogućuju jasnu i kvantitativnu analizu ponašanja sustava.

Velik broj istinitih negativnih odluka (TN = 1283) pokazuje da sustav u velikoj većini slučajeva ispravno odbacuje krivotvorine. Od ukupno 1320 krivotvorenih potpisa u testnom skupu, samo 37 je pogrešno prihvaćeno, što izravno odgovara izmjerenom FAR-u od približno 2.8 %. To potvrđuje da sustav ima visoku razinu sigurnosti i rijetko donosi sigurnosno kritičnu pogrešku.

Broj lažno prihvaćenih krivotvorina (FP = 37) predstavlja mali, ali ne zanemariv dio testnog skupa. Ti uzorci odgovaraju kvalitetnim, vještim krivotvorinama koje u HOG prostoru značajki leže blizu ili s iste strane razdjelne granice kao originalni potpisi ciljne osobe. Ovakvi slučajevi jasno ilustriraju realna ograničenja offline verifikacije potpisa, gdje vizualna sličnost može dovesti do pogrešnog prihvaćanja.

S druge strane, broj lažno odbijenih originalnih potpisa (FN = 58) veći je od broja FP pogrešaka. Od ukupno 330 originalnih potpisa u testnom skupu, 58 je pogrešno odbijeno, što odgovara FRR-u od približno 17.6 %. Ove pogreške proizlaze iz prirodne intra-class varijabilnosti potpisa, gdje se pojedini originalni potpisi razlikuju po obliku, nagibu, kontinuitetu poteza ili kvaliteti skeniranja u odnosu na uzorke korištene za treniranje.

Broj istinitih pozitivnih odluka (TP = 272) pokazuje da sustav ispravno prihvaća većinu originalnih potpisa, ali ne sve. To potvrđuje da je model naučio reprezentativan opis potpisa ciljne osobe, ali da je zbog naglaska na niskom FAR-u granica razdvajanja postavljena relativno konzervativno, što dovodi do povećanog broja FN pogrešaka.

Ova matrica zabune jasno ilustrira FAR/FRR kompromis u konačnoj konfiguraciji sustava. Sustav je optimiziran prema sigurnosti, uz vrlo nizak broj pogrešno prihvaćenih krivotvorina, po cijenu umjereno povišene stope odbacivanja originalnih potpisa. Takav kompromis je svjestan i u skladu s ciljem verifikacijskog sustava, gdje je pogrešno prihvaćanje neautoriziranog potpisa u pravilu kritičnije od povremenog odbacivanja legitimnog korisnika.

Na temelju ove analize moguće je izravno povezati numeričke metrike (FAR, FRR, Accuracy) s konkretnim ponašanjem sustava, čime matrica zabune služi kao ključni alat za razumijevanje stvarnih performansi implementiranog rješenja.

17. EKSPERIMENTALNA ANALIZA KONFIGURACIJA SUSTAVA

U ovom poglavlju provedena je detaljna eksperimentalna analiza više konfiguracija sustava za verifikaciju statičkog potpisa s ciljem ispitivanja utjecaja ključnih čimbenika na performanse writer-dependent modela. Analiza obuhvaća promjene u broju originalnih potpisa dostupnih za treniranje po osobi, izbor kernela SVM klasifikatora te skaliranje sustava na veći broj korisnika.

Sve konfiguracije koriste identičan implementacijski pipeline: istu podjelu na train i test skup temeljem CSV datoteke, isto predprocesiranje slika, iste HOG parametre, isti način treniranja modela po osobi te fiksni prag odlučivanja od 0.0. Na taj način osigurano je da su razlike u rezultatima isključivo posljedica promjene eksperimentalnih uvjeta, a ne implementacijskih razlika.

Rezultati su prikazani kroz makro i mikro prosječne metrike FAR, FRR i točnosti, kao i kroz agregirane matrice zabune. Poseban naglasak stavljen je na kompromis između sigurnosti sustava, izražene kroz FAR, i pouzdanosti prihvaćanja legitimnih korisnika, izražene kroz FRR.

17.1. Custom dataset (119 osoba, 5 genuine potpisa po osobi)

Prva eksperimentalna konfiguracija temelji se na prilagođenom datasetu koji obuhvaća 119 osoba, pri čemu je za svaku osobu korišteno samo 5 originalnih potpisa za treniranje writer-dependent klasifikatora. Ovakav scenarij predstavlja izrazito nepovoljne uvjete zbog vrlo malog broja pozitivnih uzoraka po osobi te služi kao test stabilnosti različitih SVM pristupa.

17.1.1. Linearni SVM

U konfiguraciji s linearnim SVM klasifikatorom postignute su sljedeće performanse:

- Macro prosjek: FAR ≈ 11.6 %, FRR ≈ 24.3 %, Accuracy ≈ 83.0 %
- Micro prosjek: FAR ≈ 11.7 %, FRR ≈ 21.3 %, Accuracy ≈ 84.1 %

Agregirana matrica zabune pokazuje:

- TN = 1839
- FP = 243
- FN = 343
- TP = 1269

Rezultati pokazuju da linearni SVM u uvjetima vrlo ograničenog broja trening uzoraka po osobi postiže relativno stabilne performanse. Iako je FAR umjereno povišen, sustav uspijeva zadržati prihvatljivu ukupnu točnost. FRR je također značajan, ali ne ekstreman, što ukazuje da linearni model pokazuje određenu toleranciju prema varijabilnosti originalnih potpisa čak i uz mali broj uzoraka.

17.1.2. RBF SVM

U istoj konfiguraciji, ali uz korištenje RBF kernela, dobiveni su sljedeći rezultati:

- Macro prosjek: FAR ≈ 0.14 %, FRR ≈ 76.6 %, Accuracy ≈ 67.4 %
- Micro prosjek: FAR ≈ 0.19 %, FRR ≈ 70.0 %, Accuracy ≈ 69.3 %

Agregirana matrica zabune:

- TN = 2078
- FP = 4
- FN = 1129
- TP = 483

RBF SVM u ovom scenariju pokazuje izrazito nisku stopu lažnog prihvaćanja, praktički eliminirajući pogrešno prihvaćene krivotvorine. Međutim, cijena takve sigurnosti je vrlo visok FRR, pri čemu se velik dio originalnih potpisa pogrešno odbacuje. Model postaje pretjerano restriktivan i nije sposoban generalizirati obrazac potpisa na temelju samo pet uzoraka po osobi.

Usporedba ova dva pristupa jasno pokazuje da u uvjetima izrazito oskudnih podataka linearni SVM predstavlja stabilniji baseline, dok RBF kernel zahtijeva znatno bogatiji skup trening uzoraka kako bi njegova veća izražajna moć došla do izražaja. Ovi rezultati služe kao polazišna točka za analizu konfiguracija s većim brojem originalnih potpisa po osobi u sljedećim eksperimentima.

17.2. CEDAR dataset (55 osoba) – utjecaj broja genuine potpisa po osobi

U nastavku eksperimentalne analize korišten je CEDAR dataset koji obuhvaća 55 osoba, pri čemu je postupno povećavan broj originalnih potpisa korištenih za treniranje writer-dependent modela. Cilj ove analize bio je ispitati kako količina pozitivnih uzoraka po osobi utječe na stabilnost modela, ravnotežu FAR i FRR te ukupnu pouzdanost sustava.

U svim konfiguracijama korišten je isti postupak podjele podataka, isto predprocesiranje, iste HOG značajke i identični kriterij odlučivanja. Time je omogućena izravna usporedba rezultata između različitih postavki.

17.2.1. 5 genuine potpisa po osobi, RBF SVM

Kod korištenja samo 5 originalnih potpisa po osobi i RBF kernela dobiveni su sljedeći rezultati:

- Macro i micro prosjek: FAR = 0.0 %, FRR ≈ 88.1 %, Accuracy ≈ 61.1 %

Agregirana matrica zabune:

- TN = 1320
- FP = 0
- FN = 921
- TP = 124

Ovi rezultati jasno pokazuju da model u potpunosti odbacuje sve krivotvorine, ali pritom odbacuje i gotovo sve originalne potpise. Sustav u ovoj konfiguraciji postaje izrazito konzervativan i praktično neupotrebljiv za realnu verifikaciju identiteta. Ovakvo ponašanje potvrđuje da RBF SVM nije prikladan za scenarije s izrazito malim brojem uzoraka po osobi.

17.2.2. 12 genuine potpisa po osobi

Linearni SVM

- Macro i micro prosjek: FAR ≈ 28.5 %, FRR ≈ 9.7 %, Accuracy ≈ 77.8 %

Agregirana matrica zabune:

- TN = 944
- FP = 376

- FN = 64
- TP = 596

Linearni model u ovoj konfiguraciji pokazuje nisku stopu odbacivanja originalnih potpisa, ali uz vrlo visoku stopu lažnog prihvaćanja krivotvorina. Takav profil pogodan je za sustave u kojima je prioritet prihvat legitimnih korisnika, ali predstavlja ozbiljan sigurnosni rizik.

RBF SVM

- Macro i micro prosjek: FAR ≈ 1.0 %, FRR ≈ 33.9 %, Accuracy ≈ 88.0 %

Agregirana matrica zabune:

- TN = 1307
- FP = 13
- FN = 224
- TP = 436

Usporedbom s linearnim modelom vidljivo je značajno smanjenje FAR-a uz povećanje FRR-a. RBF kernel počinje pokazivati prednost u sigurnosti, ali još uvijek nema dovoljno podataka za pouzdano prihvaćanje svih varijacija originalnih potpisa.

17.2.3. 16 genuine potpisa po osobi, RBF SVM

- Macro i micro prosjek: FAR ≈ 1.7 %, FRR ≈ 25.7 %, Accuracy ≈ 92.3 %

Agregirana matrica zabune:

- TN = 1297
- FP = 23
- FN = 113
- TP = 327

Povećanjem broja genuine potpisa na 16 po osobi dolazi do vidljivog poboljšanja u ravnoteži između sigurnosti i pouzdanosti. FRR se značajno smanjuje u odnosu na konfiguraciju s 12 potpisa, dok FAR ostaje nizak. Ova konfiguracija predstavlja prijelaznu točku prema stabilnijem ponašanju RBF modela.

17.2.4. 18 genuine potpisa po osobi, RBF SVM (odabrana konfiguracija)

- Macro i micro prosjek: FAR ≈ 2.8 %, FRR ≈ 17.6 %, Accuracy ≈ 94.2 %

Agregirana matrica zabune:

- TN = 1283
- FP = 37
- FN = 58
- TP = 272

Ova konfiguracija pokazuje najbolji kompromis između FAR i FRR. Sustav uspješno odbacuje većinu krivotvorina, uz istovremeno visoku stopu prihvaćanja originalnih potpisa. Broj test uzoraka i stabilnost rezultata čine ovu postavku najpouzdanijom za evaluaciju, zbog čega je odabrana kao konačna konfiguracija sustava.

17.2.5. 20 genuine potpisa po osobi, RBF SVM

- Macro i micro prosjek: FAR ≈ 2.5 %, FRR ≈ 18.6 %, Accuracy ≈ 95.2 %

Agregirana matrica zabune:

- TN = 1287
- FP = 33
- FN = 41
- TP = 179

Iako ova konfiguracija postiže nešto višu ukupnu točnost, smanjenje veličine test skupa dovodi do manje pouzdane evaluacije. Zbog toga, unatoč blagom numeričkom poboljšanju, konfiguracija s 18 genuine potpisa po osobi predstavlja uravnoteženiji i metodološki ispravniji izbor.

17.3. Skalabilnost sustava – spojeni dataset (174 osobe)

U završnoj fazi eksperimentalne evaluacije analizirana je skalabilnost predloženog sustava povećanjem broja korisnika. U tu svrhu spojen je custom dataset s 119 osoba i CEDAR dataset s 55 osoba, čime je formiran objedinjeni skup od ukupno 174 osobe. Za svaku

osobu zadržan je writer-dependent pristup s posebnim binarnim klasifikatorom, dok su ostali dijelovi pipelinea ostali nepromijenjeni.

Cilj ove analize bio je ispitati utjecaj povećanja broja osoba na performanse sustava te utvrditi ima li ukupna veličina populacije značajan utjecaj na FAR, FRR i točnost, u odnosu na broj dostupnih uzoraka po osobi.

17.3.1. 5 genuine potpisa po osobi, linearni SVM

- Macro prosjek: FAR ≈ 11.5 %, FRR ≈ 23.2 %, Accuracy ≈ 83.3 %
- Micro prosjek: FAR ≈ 10.4 %, FRR ≈ 23.3 %, Accuracy ≈ 83.9 %

Agregirana matrica zabune:

- TN = 3047
- FP = 355
- FN = 620
- TP = 2037

Rezultati pokazuju ponašanje vrlo slično onome dobivenom na custom datasetu s 119 osoba. Linearni SVM zadržava stabilnost i umjerenu točnost, ali uz povišen FAR. Povećanje broja osoba samo po sebi ne donosi poboljšanje performansi kada je broj genuine potpisa po osobi ograničen.

17.3.2. 11 genuine potpisa po osobi, RBF SVM

- Macro prosjek: FAR ≈ 3.0 %, FRR ≈ 31.6 %, Accuracy ≈ 89.3 %
- Micro prosjek: FAR ≈ 2.5 %, FRR ≈ 28.6 %, Accuracy ≈ 89.1 %

Agregirana matrica zabune:

- TN = 3318
- FP = 84
- FN = 462
- TP = 1151

Usporedbom s konfiguracijom od 5 potpisa vidljivo je značajno smanjenje FAR-a, što potvrđuje da RBF kernel počinje pokazivati prednosti kada je dostupno više podataka po osobi. Ipak, FRR ostaje relativno visok, što upućuje na to da 11 genuine potpisa još uvijek nije dovoljno za pouzdano modeliranje intra-class varijabilnosti kod većeg broja osoba.

17.3.3. Zaključak skalabilnosti

Rezultati dobiveni na spojenom datasetu jasno potvrđuju da ukupni broj osoba nema presudan utjecaj na performanse sustava u writer-dependent scenariju. Ključni čimbenik ostaje broj kvalitetnih genuine potpisa po osobi, dok povećanje populacije bez povećanja količine podataka po korisniku ne poboljšava pouzdanost sustava.

Ova analiza dodatno potvrđuje ranije zaključke da je za primjenu nelinearnih modela, poput RBF SVM-a, nužno osigurati dovoljno bogat skup uzoraka po osobi. U suprotnom, sustav pokazuje povišene stope odbacivanja legitimnih korisnika, bez obzira na veličinu ukupnog skupa podataka.

17.4. Sažetak eksperimentalnih rezultata i odabir konačne konfiguracije

Provedena eksperimentalna analiza omogućila je sustavnu usporedbu različitih konfiguracija sustava za verifikaciju statičkog potpisa u writer-dependent scenariju. Analizirani su utjecaji vrste SVM klasifikatora, broja originalnih potpisa po osobi i veličine populacije korisnika, pri čemu su svi eksperimenti provedeni uz identičan implementacijski pipeline i fiksni prag odlučivanja.

Rezultati jasno pokazuju da linearni SVM predstavlja stabilan baseline u scenarijima s vrlo ograničenim brojem trening uzoraka po osobi. U takvim uvjetima linearni model zadržava umjerene vrijednosti FAR i FRR te relativno stabilnu točnost, ali uz povišenu stopu lažnog prihvaćanja krivotvorina. Zbog toga linearni SVM nije prikladan za primjene u kojima je sigurnost primarni zahtjev.

Nelinearni RBF SVM pokazuje izrazitu osjetljivost na količinu dostupnih podataka po osobi. U konfiguracijama s malim brojem genuine potpisa RBF model postaje pretjerano restriktivan, što rezultira ekstremno visokim FRR-om i neprihvatljivom uporabnom vrijednošću

sustava. Međutim, s postupnim povećanjem broja originalnih potpisa po osobi dolazi do značajnog poboljšanja performansi i uspostave ravnoteže između sigurnosti i pouzdanosti.

Eksperimenti na CEDAR datasetu pokazali su da se optimalna ravnoteža postiže pri korištenju 18 genuine potpisa po osobi u kombinaciji s RBF SVM klasifikatorom. U toj konfiguraciji sustav postiže nisku stopu lažnog prihvaćanja krivotvorina, uz istovremeno prihvatljivu stopu odbacivanja originalnih potpisa i visoku ukupnu točnost. Dodatno povećanje broja trening uzoraka na 20 potpisa donosi tek marginalno poboljšanje točnosti, ali uz smanjenje pouzdanosti evaluacije zbog manjeg testnog skupa.

Analiza skalabilnosti pokazala je da povećanje broja osoba u sustavu nema presudan utjecaj na performanse verifikacije, sve dok je broj genuine potpisa po osobi nedostatan. Time je potvrđeno da je u writer-dependent pristupu kvaliteta i količina podataka po korisniku važniji čimbenik od ukupne veličine populacije.

Na temelju svih provedenih eksperimenata, kao konačna konfiguracija sustava odabran je CEDAR dataset s 55 osoba, 18 genuine potpisa po osobi za treniranje i RBF SVM klasifikator s fiksnim pragom odlučivanja od 0.0. Ova konfiguracija predstavlja najbolji kompromis između sigurnosti, pouzdanosti i metodološke ispravnosti evaluacije te je korištena za detaljnu analizu verifikacije i prikaz rezultata u ovom radu.

Završetkom ovog poglavlja zaokružena je eksperimentalna evaluacija implementiranog sustava. Dobiveni rezultati i odabrana konačna konfiguracija predstavljaju podlogu za demonstraciju rada sustava u praktičnom okruženju, koja je opisana u sljedećem poglavlju.

18. DEMONSTRACIJSKA WEB APLIKACIJA

Uz implementaciju i eksperimentalnu evaluaciju sustava za verifikaciju statičkog potpisa, razvijena je i jednostavna demonstracijska web aplikacija. Svrha aplikacije je omogućiti praktičan prikaz rada implementiranog sustava nad pojedinačnim potpisima te povezati teorijski i implementacijski dio rada s konkretnim primjerom uporabe.

Web aplikacija ne predstavlja sastavni dio verifikacijskog sustava u smislu treniranja ili evaluacije modela. Ona ne uvodi nove algoritme, ne mijenja postojeći pipeline i ne utječe na dobivene eksperimentalne rezultate. Svi modeli korišteni u aplikaciji prethodno su istrenirani u offline fazi i spremljeni na disk, dok se tijekom rada aplikacije provodi isključivo verifikacija pojedinačnih potpisa.

Implementacija web aplikacije omogućuje korisniku odabir osobe, učitavanje slike potpisa i prikaz rezultata verifikacije. Time se na jasan i pregledan način demonstrira način rada writer-dependent sustava za verifikaciju potpisa u realističnom scenariju, bez dodatne složenosti koja bi mogla zamagliti osnovne principe rada sustava.

18.1. Svrha i uloga web aplikacije u radu

Uloga demonstracijske web aplikacije u ovom radu isključivo je ilustrativna. Aplikacija služi za prikaz načina na koji se prethodno implementirani i evaluirani sustav za verifikaciju statičkog potpisa može koristiti u praksi, nad pojedinačnim ulaznim uzorcima.

Web aplikacija nije korištena za prikupljanje rezultata, izračun evaluacijskih metrika niti za usporedbu različitih konfiguracija sustava. Svi rezultati prikazani u evaluacijskom dijelu rada dobiveni su izvođenjem skripti za treniranje i testiranje u offline okruženju. Web aplikacija koristi iste modele i isti verifikacijski postupak, ali isključivo za demonstraciju rada sustava nad odabranim primjerima.

Posebno je važno naglasiti da aplikacija ne provodi treniranje modela, ne koristi krivotvorine za učenje i ne prilagođava prag odlučivanja. Odluka o prihvaćanju ili odbacivanju potpisa temelji se isključivo na izlazu SVM klasifikatora i fiksnom pragu odlučivanja od 0.0, u skladu s implementacijom opisanom u prethodnim poglavljima.

Na taj način web aplikacija predstavlja tanak prezentacijski sloj iznad jezgre verifikacijskog sustava. Njezina svrha je olakšati razumijevanje rada sustava i omogućiti

vizualnu interpretaciju pojedinačnih odluka, bez ikakvog utjecaja na algoritamski dio rješenja ili eksperimentalne zaključke rada.

18.2. Povezanost web aplikacije s verifikacijskim sustavom

Web aplikacija povezana je s verifikacijskim sustavom na način da u potpunosti ponovno koristi postojeću logiku za verifikaciju pojedinačnog potpisa. Backend aplikacije ne implementira vlastitu obradu potpisa, već služi isključivo kao orkestrator koji prikuplja korisnički unos i prosljeđuje ga verifikacijskom pipelineu.

Centralna točka povezivanja je skripta `app.py`, koja prima zahtjev korisnika, sprema učitanu sliku potpisa i poziva funkcionalnost za verifikaciju. U toj fazi web aplikacija ne provodi nikakvu dodatnu obradu slike niti interpretaciju rezultata.

```
# app.py – pozivanje verifikacije potpisa
from verify_one import verify_signature

result = verify_signature(
    image_path=uploaded_image_path,
    person_id=selected_person_id,
    models_dir=MODELS_DIR,
)
```

Funkcija `verify_signature` definirana je u skripti `verify_one.py` i predstavlja jedini ulaz u verifikacijski pipeline. Time se osigurava da se tijekom rada web aplikacije koristi identičan postupak verifikacije kao i u komandno-linijskom načinu rada i tijekom evaluacije sustava.

Unutar skripte `verify_one.py` provode se svi ključni koraci verifikacije: učitavanje writer-dependent modela s diska, predprocesiranje ulazne slike, ekstrakcija HOG značajki te izračun decision score-a pomoću SVM klasifikatora.

```
# verify_one.py – učitavanje modela i izračun score-a
pack = joblib.load(model_path)
clf = pack["model"]
```

```
score = float(clf.decision_function(features)[0])
```

Modeli se učitavaju iz direktorija s prethodno istreniranim `.joblib` datotekama, pri čemu svaka datoteka odgovara jednoj osobi. Web aplikacija ne upravlja treniranjem niti pohranom modela, već pretpostavlja da su modeli već dostupni na disku.

Na ovaj način web aplikacija ostaje strogo odvojena od algoritamskog dijela sustava. Svi rezultati koje aplikacija prikazuje izravno su rezultat verifikacijskog pipelinea opisanog u implementacijskom dijelu rada, bez ikakvih izmjena ili aproksimacija. Time se osigurava da demonstracija vjerno odražava ponašanje sustava analiziranog u prethodnim poglavljima.

U sljedećem potpoglavlju opisan je način prikaza rezultata verifikacije korisniku i interpretacija odluke ACCEPT ili REJECT.

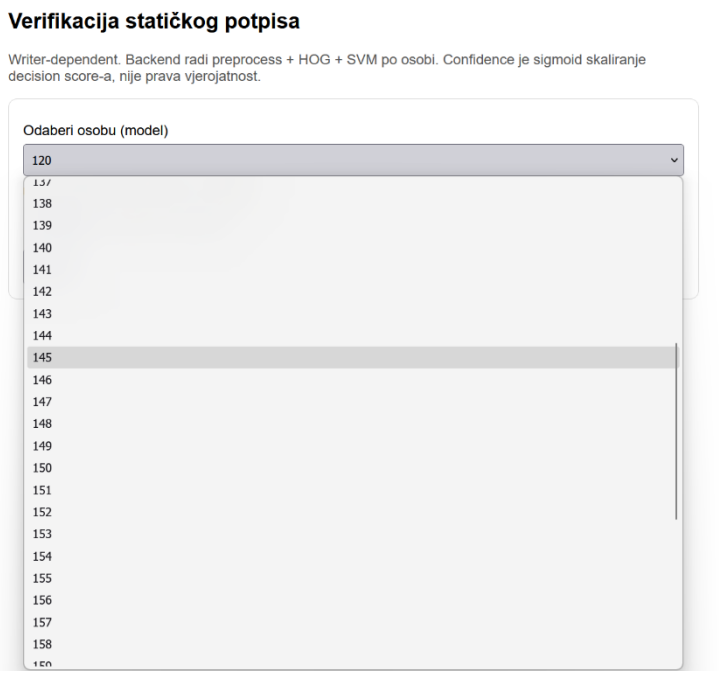
18.3. Prikaz rezultata verifikacije

Nakon provedene verifikacije pojedinačnog potpisa, web aplikacija korisniku prikazuje rezultat u jednostavnom i preglednom obliku. Prikaz rezultata ima isključivo informativnu i demonstracijsku svrhu te ne utječe na odluku sustava.

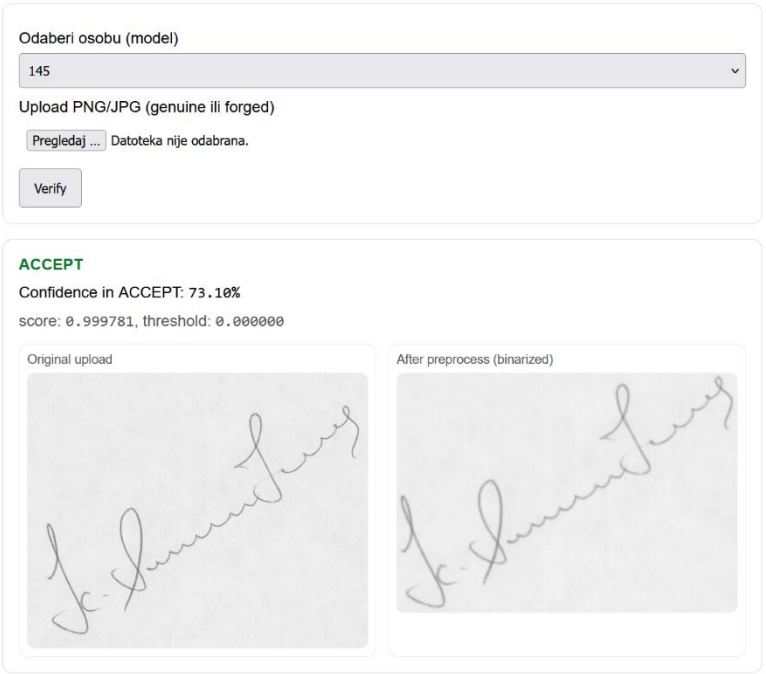
Središnji element prikaza je binarna odluka ACCEPT ili REJECT, koja izravno proizlazi iz usporedbe decision score-a s fiksnim pragom odlučivanja od 0.0. Uz binarnu odluku prikazuje se i numerička vrijednost decision score-a, kako bi se korisniku dao uvid u udaljenost uzorka od razdjelne granice klasifikatora. Veća apsolutna vrijednost score-a upućuje na sigurniju odluku sustava.

Važno je naglasiti da decision score ne predstavlja vjerojatnost. Iako se u korisničkom sučelju može prikazati dodatna, pojednostavljena mjera sigurnosti, takav prikaz služi isključivo radi intuitivnosti i nema formalno statističko značenje. Odluka sustava uvijek se temelji isključivo na znaku decision score-a i fiksnom pragu odlučivanja.

Na slici u nastavku prikazan je primjer web sučelja u kojem je vidljiv odabir identiteta osobe putem padajućeg izbornika, kao i rezultat uspješne verifikacije potpisa. Prikaz odabira identifikatora osobe dodatno ilustrira writer-dependent prirodu sustava, gdje se verifikacija uvijek provodi u odnosu na točno određenu osobu.



Slika 6: Odabir identiteta osobe putem padajućeg izbornika



Slika 7: Uspješna verifikacija potpisa

U prikazanom primjeru sustav donosi pozitivnu odluku ACCEPT, što znači da ulazni potpis pripada odabranoj osobi prema kriterijima naučenim tijekom treniranja. Prikaz decision

score-a omogućuje uvid u jačinu te odluke, dok je sama interpretacija rezultata u skladu s ponašanjem sustava analiziranim u evaluacijskom dijelu rada.

Ovakav način prikaza rezultata omogućuje jasnu i transparentnu demonstraciju rada sustava, bez uvođenja dodatnih elemenata koji bi mogli utjecati na algoritamski dio verifikacije. Web aplikacija time ostaje jednostavan prezentacijski sloj iznad već evaluiranog verifikacijskog pipelinea.

19. ZAKLJUČAK IMPLEMENTACIJSKOG DIJELA

Implementacijski dio rada prikazao je cjelovitu realizaciju sustava za writer-dependent verifikaciju statičkog potpisa, u skladu s teorijskim postavkama i jasno definiranim eksperimentalnim pravilima. Sustav je implementiran kao modularni pipeline koji obuhvaća predprocesiranje slika, ekstrakciju HOG značajki, treniranje SVM klasifikatora po osobi, evaluaciju performansi te verifikaciju pojedinačnih potpisa uz fiksni prag odlučivanja.

Poseban naglasak stavljen je na strogo razdvajanje faze treniranja, evaluacije i verifikacije. Podjela podataka provedena je isključivo putem CSV datoteke kao jedinog izvora istine, čime je osigurana ponovljivost eksperimenata i metodološka ispravnost evaluacije. Writer-dependent pristup omogućio je prilagodbu modela individualnim karakteristikama potpisa svake osobe, što se pokazalo ključnim za postizanje konkurentnih rezultata u offline scenariju.

Eksperimentalna analiza različitih konfiguracija jasno je pokazala da količina dostupnih originalnih potpisa po osobi ima presudan utjecaj na performanse sustava. Linearni SVM pokazao se kao stabilan baseline u uvjetima oskudnih podataka, dok je RBF SVM ostvario značajnu prednost tek uz dovoljan broj trening uzoraka. Konačna konfiguracija temeljena na CEDAR datasetu s 18 genuine potpisa po osobi i RBF SVM-om ostvarila je najbolji kompromis između sigurnosti i pouzdanosti, što je potvrđeno kroz FAR, FRR, ROC i Precision–Recall analize.

Demonstracijska web aplikacija dodatno je potvrdila ispravnost implementacije, omogućujući praktičan prikaz rada sustava bez utjecaja na algoritamski dio ili evaluacijske rezultate. Time je pokazano kako se implementirani sustav može integrirati u aplikacijsko okruženje uz zadržavanje istog verifikacijskog pipelinea.

Zaključno, implementacijski dio rada potvrđuje da je predloženi sustav metodološki konzistentan, tehnički ispravno realiziran i sposoban za pouzdanu offline verifikaciju potpisa u writer-dependent scenariju. Dobiveni rezultati i struktura implementacije pružaju čvrstu osnovu za daljnja poboljšanja, koja se razmatraju u zaključnom dijelu rada.

Popis literature

Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017). *Offline handwritten signature verification: A literature review.* IEEE.

Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2015). *Offline handwritten signature verification: A literature review.* arXiv

Singh, D., Mittal, A., Aggarwal, N., & Kumar, R. (2022). *Offline signature verification: A systematic review.* Research Square.

Kumar, D. S. S. (2023). *Offline signature verification based on ensemble of features using support vector machine. International Journal of Computer Applications,* 184(45), 24-29.

Engin, D., Kantarcı, A., Arslan, S., & Ekenel, H. K. (2020). *Offline signature verification on real-world documents. U Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).* IEEE.

Zois, E. N., Tsourounis, D., Theodorakopoulos, I., Kesidis, A. L., & Economou, G. (2019). *A comprehensive study of sparse representation techniques for offline signature verification. IEEE Transactions on Information Forensics and Security*, 14(10), 2637-2652.

Maergner, P., Howe, N. R., Riesen, K., Ingold, R., & Fischer, A. (2019). *Graph-based offline signature verification. Pattern Recognition.* Preprint.

Bouamra, W. (2022). *Offline handwritten signature verification and forgery detection (Doctoral dissertation).* Larbi Ben M'hidi University & Universidad de Las Palmas de Gran Canaria.

Popis slika

Popis datasetova

Shreelakshmigp. (n.d.). *Cedar dataset* [Data set]. Kaggle. preuzeto 2.1.2026. sa https://www.kaggle.com/datasets/shreelakshmigp/cedardataset

Mallapraveen. (n.d.). *Signature matching* [Data set]. Kaggle. preuzeto 2.1.2026. sa https://www.kaggle.com/datasets/mallapraveen/signature-matching/data