УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Василије Тодоровић

ВЕРИФИКАЦИЈА ПОТПИСА МЕТОДАМА ДУБОКОГ УЧЕЊА СЛИЧНОСТИ

мастер рад

Београд, 2026.

Ментор:

др Младен Николић, ванредни професор
Универзитет у Београду, Математички факултет

Чланови комисије:

др Јелена Граовац, ванредни професор
Универзитет у Београду, Математички факултет

др Јована Ковачевић, ванредни професор
Универзитет у Београду, Математички факултет

Датум одбране: _____

Баби Лепи и деди Драгомиру из Бора

Наслов мастер рада: Верификација потписа методама дубоког учења сличности

Резиме: Верификација руком писаних потписа представља значајан задатак у области биометрије, посебно изазован у офлајн поставци где је на располагању само статичка слика потписа. Главни изазов представљају вешто израђени фалсификати, намерно конструисани да наликују циљном потпису, услед чега се тешко разликују од аутентичних узорака.

Савремени приступи засновани на дубоком учењу сличности пресликавају потписе у векторски простор у коме су потписи различитих аутора међусобно удаљени, док су потписи истог аутора груписани. Међутим, већина постојећих метода фалсификате третира као произвољне негативне узорке, чиме се занемарује чињеница да је сваки фалсификат намерно везан за одређени идентитет.

У овом раду предлаже се методолошки оквир за дубоко учење сличности који експлицитно моделује однос између оригиналних потписа и њихових фалсификата током обучавања. Стратегије прилагођеног узорковања и избора информативних примера обезбеђују да парови оригинал–фалсификат истог идентитета увек улазе у функцију грешке као тешки негативни узорци. Архитектура користи мрежу ResNet18 са 512-димензионалним векторским репрезентацијама, а упоређене су три функције грешке: триплет функција, функција вишеструке сличности и њихова хибридна комбинација.

При фиксираној стопи прихватања фалсификата од 10% (GAR@FAR=10%), модели свесни фалсификата прихватају 97,0–98,1% аутентичних потписа на валидацији и 85,7–92,5% на потписницима који нису били доступни током обучавања, у односу на 83,9% и 42,4% за контролни модел обучен без података о фалсификатима. Резултати показују да експлицитна свест о фалсификатима током обучавања значајно доприноси робусности система за верификацију потписа.

Кључне речи: машинско учење, рачунарски вид, биометрија, верификација потписа, дубоко учење сличности

Садржај

Глава 1

Увод

Руком писани потпис је један од најстаријих и најчешће коришћених начина за биометријску потврду идентитета појединца. Иако су електронски облици потврде идентитета све заступљенији, потпис на папиру и даље има правну тежину у уговорима, пуномоћјима, банкарским чековима и архивским документима, а често је и једина веза између потписника и документа [5, 12]. Отуда постоји потреба за поузданим аутоматским системима за верификацију потписа, који могу да утврде да ли је дати потпис прави или је реч о покушају имитације.

Верификација потписа се класично дели на два сценарија. У *онлајн* поставци, потпис се снима током самог писања, па су доступни и динамички подаци попут притиска оловке, брзине и редоследа потеза. У *офлајн* поставци, на располагању је само статичка слика већ потписаног документа [10]. Други сценарио је по својој природи тежи, јер се губе сви подаци о току и начину настанка потписа, али је практично значајнији, будући да се већина пословних и правних докумената скенира тек након потписивања. Главни изазов у офлајн верификацији представљају *вешто израђени фалсификати*, потписи које је намерно направио други човек пошто је претходно проучио циљни потпис. За разлику од насумичних покушаја, овакви фалсификати пажљиво опонашају облик и распоред оригинала, па су одступања толико мала да их ни искусни вештаци не могу увек поуздано уочити.

Класични приступи офлајн верификацији ослањали су се на ручно осмишљене особине потписа (геометријске мере, кодове оријентације ивица и текстурне дескрипторе), које су потом прослеђиване класификаторима попут методе потпорних вектора [12]. Осим што су захтевале мукотрпно ручно осми-

шљавање особина, овакве методе су се тешко преносиле на нове потписнике и нове услове прикупљања података. Прелазак на дубоке конволутивне неуронске мреже омогућио је учење репрезентација непосредно из слика потписа [4, 9]. Међу приступима дубоког учења, овом задатку посебно природно одговара *дубоко учење сличности* (енг. *deep metric learning*), које учи пресликавање слика у векторски простор тако да сличним потписима одговарају мала растојања. Оваква организација простора омогућава да се потпис верификује поређењем растојања, чак и када је на располагању само мали број референтних потписа, а исти модел се примењује и на нове потписнике без поновног обучавања.

Иако је дубоко учење сличности постало водећи приступ у верификацији потписа, већина постојећих метода фалсификате током обучавања третира као обичне негативне узорке или их сасвим изоставља [18, 23, 28]. Тако остаје неискоришћена једна важна информација: сваки фалсификат намерно је везан за одређени идентитет и представља управо ону врсту тешких примера на које ће систем наилазити у пракси. Ако се однос између оригинала и његовог фалсификата не моделује изричито, векторски простор се добро уређује по идентитетима, али граница између правих потписа и њихових фалсификата остаје слабо одређена.

Циљ овог рада је развој модела за офлајн верификацију потписа који током обучавања изричито користи везу између оригиналних потписа и њима намењених фалсификата. Уз то, рад обухвата анализу и поређење различитих стратегија дубоког учења сличности, које се међусобно разликују по избору функције грешке и по начину на који користе фалсификате.

Главни доприноси овог рада су следећи. Прво, предлаже се методолошки оквир *свесшан фалсификаша*, у коме прилагођене стратегије узорковања и избора информативних примера обезбеђују да парови оригинал–фалсификат истог идентитета увек уђу у функцију грешке као тешки негативни узорци. Друго, систематски се пореде три функције грешке (триплет [24], вишеструка сличност [30] и њихова хибридна комбинација) на истом архитектонском оквиру (мрежа ResNet18 [11] и 512-димензионални векторски простор) и под истим условима обучавања. Треће, спроводи се аблациона анализа у којој се исти модел обучава без података о фалсификатима, чиме се издваја колики је допринос свести о фалсификатима у односу на остале факторе обучавања.

Експерименти показују значајно побољшање у односу на контролни модел.

При фиксираној стопи прихватања фалсификата од 10% (GAR@FAR=10%), модели свесни фалсификата прихватају 97,0–98,1% аутентичних потписа на валидацији, у односу на 83,9% за контролни модел. На потписницима који нису били доступни током обучавања разлика је још израженија: модели свесни фалсификата прихватају 85,7–92,5% аутентичних потписа, у односу на свега 42,4% за контролни модел. Овако велики јаз показује да модели свесни фалсификата знатно боље генерализују на нове потписнике, што је од пресудног значаја у пракси, где систем готово увек ради са идентитетима који нису били део обучавања.

Остатак рада организован је на следећи начин. Поглавље 2 излаже теоријске основе на којима рад почива: вештачке и конволутивне неуронске мреже, са нагласком на архитектуру ResNet, појам угњеждења и његову улогу у учењу сличности, сијамске архитектуре и основне алате дубоког учења сличности. Поглавље 3 даје преглед области верификације потписа: разлику између онлајн и офлајн поставке, класичне приступе и савремене методе засноване на дубоком учењу сличности. Поглавље 4 формално поставља проблем верификације и излаже предложену методологију: архитектуру модела, испитиване функције грешке, стратегију обучавања свесну фалсификата и целокупан ток система током обучавања и закључивања. Поглавље 5 описује експерименталну поставку: скуп података и његову поделу, имплементационе детаље и протоколе евалуације. Најзад, поглавље 6 доноси експерименталне резултате и њихову анализу, а завршна дискусија изводи кључне закључке за офлајн верификацију потписа.

Глава 2

Теоријске основе

Предложени приступ верификацији потписа заснива се на неколико појмова из дубоког учења. Њихов преглед обухвата вештачке и конволутивне неуронске мреже, векторске репрезентације, сијамске мреже и дубоко учење сличности, на којима почивају методологија и експерименти који следе.

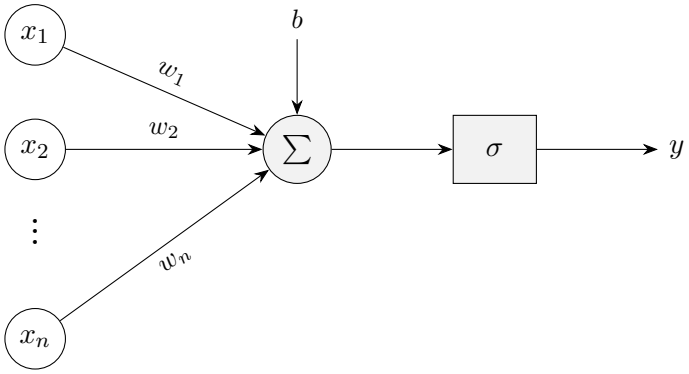
2.1 Вештачке неуронске мреже

За разлику од класичних програма, у којима човек мора унапред да пропише правила решавања, вештачке неуронске мреже уче да реше задатак самостално, на основу великог броја примера. Изграђене су од мноштва једноставних, међусобно повезаних јединица које заједно откривају обрасце у подацима [35]. Основна градивна јединица мреже је *вештачки неурон* — параметризована функција која прима вектор улаза, израчунава линеарну комбинацију његових компоненти и пропушта је кроз нелинеарну активациону функцију. Формално, за улазни вектор $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$, неурон израчунава излаз y као

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right) = \sigma(\mathbf{w}^\top \mathbf{x} + b), \tag{2.1}$$

где су $\mathbf{w} = (w_1, \ldots, w_n)^\top$ тежине, b пристрасност (енг. *bias*), а σ нелинеарна активациона функција. Шематски приказ једног неурона дат је на слици 2.1.

Активациона функција σ игра кључну улогу: она у мрежу уводи нелинеарност и тиме јој омогућава да представи и зависности које нису линеарне. Без нелинеарне активације мрежа би се, ма колико слојева имала, свела на једну линеарну трансформацију улаза, па би јој изражајна моћ била строго огра-



Слика 2.1: Шематски приказ једног вештачког неурона.

ничена. У пракси постоји више уобичајених избора активационе функције, а конкретан избор утиче на својства мреже и на ток обучавања.

Један изоловани неурон може да представи једино линеарно одвојиве функције, али се његова изражајност драматично повећава *слагањем* више неурона у организационе јединице које се називају *слојеви*. Слој се састоји од скупа неурона који паралелно израчунавају своје излазе на основу истог улазног вектора, али уз сопствене тежине, чиме сваки неурон производи једну димензију заједничког излазног вектора слоја. Слојеви се потом међусобно *надовезују* тако што излаз једног слоја постаје улаз наредног, формирајући ланац трансформација.

Стандардна архитектура изграђена на овом принципу, *вишеслојни перцептрон* (енг. *multilayer perceptron*, MLP), састоји се од *улазног слоја*, једног или више *скривених слојева* и *излазног слоја*, при чему је сваки неурон једног слоја повезан са сваким неуроном наредног слоја. Ако са $\mathbf{a}^{(l)}$ означимо излаз слоја l, а са $\mathbf{W}^{(l)}$ и $\mathbf{b}^{(l)}$ његову тежинску матрицу и вектор пристрасности, рекурзивна веза између слојева гласи

$$\mathbf{a}^{(l)} = \sigma\big(\mathbf{W}^{(l)}\,\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}\big)\,, \tag{2.2}$$

при чему је $\mathbf{a}^{(0)}$ улазни вектор, а излаз последњег слоја $\mathbf{a}^{(L)}$ представља коначан одговор мреже. На тај начин, цела мрежа представља композицију оваквих параметризованих трансформација, $\mathbf{a}^{(L)} = f_\theta(\mathbf{a}^{(0)})$, где $\theta = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}_{l=1}^{L}$ обухвата све параметре. Уз умерене претпоставке, већ са једним довољно широким скривеним слојем може се произвољно тачно апроксимирати било која непрекидна функција (теорема о универзалној апроксимацији); у пракси се,

међутим, дубље мреже са више слојева мање ширине показују као ефикасније и њима се лакше уче сложене хијерархијске репрезентације.

Параметри мреже (тежине и пристрасности свих слојева) уче се на основу тренинг скупа, минимизацијом *функције грешке* $\mathcal{L}$ која мери одступање предвиђања мреже од тачних ознака. Учење се одвија итеративно: у *пролазу унапред* мрежа за дати улаз израчунава излаз и вредност функције грешке, а затим се *пропагацијом грешке уназад* (енг. *backpropagation*) израчунавају градијенти функције грешке по свим параметрима, на основу којих се параметри ажурирају у смеру који смањује грешку. Детаљан опис ових поступака може се наћи у [35].

Иако вишеслојни перцептрон у теорији може да представи произвољну функцију, у пракси се његова примена на сирове слике суочава са озбиљним ограничењима. Ако се свака слика третира као вектор у којем је сваки пиксел независна особина, већ за стандардну улазну слику димензија $224 \times 224 \times 3$ први скривени слој мора имати десетине милиона тежина, а додавање још неколико оваквих слојева чини обучавање рачунски и меморијски неизводљивим. Поред тога, оваква архитектура потпуно занемарује чињеницу да суседни пиксели носе више заједничке информације од удаљених, и да информативне особине објеката треба да буду препознатљиве независно од њиховог положаја унутар слике. Управо ова ограничења мотивишу прелаз на конволутивне неуронске мреже.

2.2 Конволутивне неуронске мреже

Конволутивне неуронске мреже посебно су прилагођене обради слика и других података распоређених у мрежу тачака. У поређењу са потпуно повезаним мрежама, њихове предности потичу од три идеје: сваки неурон гледа само мали део слике (*локална повезаност*), исте тежине користе се на свим положајима (*дељење параметара*), а научене особине препознају се без обзира на то где се у слици налазе (*непроменљивост на померање*) [7].

Уместо да буде повезан са свим пикселима улаза, неурон у конволутивном слоју повезан је само са малим суседством пиксела, које се назива *рецептивно поље*. Тежине које одговарају том пољу чине *језгро* (енг. *kernel*), које се помера преко целе слике, тако да се исте тежине примењују на сваком месту [17]. Захваљујући томе, мрежа препознаје научену особину без обзира на

то где се она у слици појави, а укупан број тежина остаје мали. Ређањем више конволутивних слојева слика се описује на више нивоа: први слојеви уочавају једноставне обрасце попут ивица, средњи их спајају у делове објеката, а дубљи препознају сложеније целине. Између конволутивних слојева обично се додају слојеви за *агрегацију* (енг. *pooling*), који смањују величину слике и чине мрежу отпорнијом на мала померања.

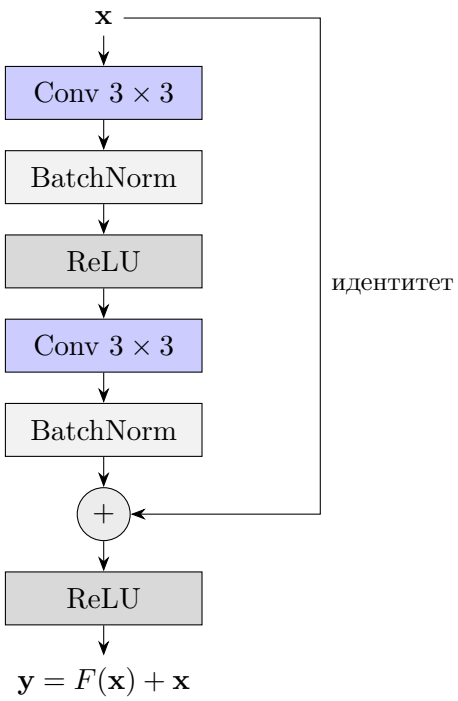
Конволутивне мреже вратиле су се у жижу истраживања мрежом AlexNet [15], која је 2012. године на такмичењу ImageNet постигла дотад невиђен резултат. После ње појављивале су се све дубље мреже. Више слојева начелно значи већу моћ мреже да научи сложеније особине, али се показало да је веома дубоке мреже тешко обучити: градијенти би се током учења толико смањили или разишли да је опадала и тачност на самом тренинг скупу.

Овај проблем решиле су *резидуалне мреже* (енг. *residual network*, ResNet) [11]. Њихова главна идеја је *резидуална веза* (енг. *skip connection*): уместо да блок од неколико конволутивних слојева директно научи жељену трансформацију $H(\mathbf{x})$, мрежа учи само разлику $F(\mathbf{x}) = H(\mathbf{x}) - \mathbf{x}$ (тзв. *резидуал*), док улаз $\mathbf{x}$ споредним путем прескаче блок и сабира се са његовим излазом (слика 2.2). Ово доноси две користи. Прво, ако је најбоља трансформација блиска томе да ништа не мења, мрежи је лако да научи $F(\mathbf{x}) \approx \mathbf{0}$, што у обичним мрежама није лако постићи. Друго, оваква веза пушта градијент да несметано тече уназад, па се избегава његово нестајање и постаје могуће поуздано обучити мреже са педесет, сто и више слојева.

Разне варијанте ResNet мреже разликују се по броју и распореду резидуалних блокова. Међу њима, ResNet18 (варијанта са укупно осамнаест слојева, чији је средишњи део организован у четири секције са по два резидуална блока) добро балансира моћ мреже и брзину рачунања, па се често користи као основа (енг. *backbone*) за задатке у којима скуп података није превелик. Из тих разлога усвојена је и у овом раду.

2.3 Векторске репрезентације

Под *уграђивањем* (енг. *embedding*) подразумева се пресликавање које сировом улазу (слици, тексту, аудио сигналу или било ком другом облику података) придружује вектор фиксне димензије $\mathbf{e} \in \mathbb{R}^d$ у тзв. *простору векторских репрезентација*. Кључна особина овог пресликавања је да семантички сродне



Слика 2.2: Резидуални блок у мрежи ResNet.

инстанце буду пресликане у блиске тачке простора, док семантички различите буду удаљене. На тај начин, простор векторских репрезентација постаје геометријска репрезентација значења.

Једна од најмоћнијих особина векторских репрезентација је њихова *универзалност*: исти концептуални оквир се на природан начин примењује на широк спектар модалитета података. Речи, реченице, документи, слике, лица, аудио сигнали, молекули и многи други типови података могу се представити као вектори у одговарајућем простору векторских репрезентација. Тиме векторске репрезентације постају својеврстан универзални језик репрезентације: уместо да се за сваки задатак развијају *ad hoc* особине, један исти модел учи да сваки податак преслика у вектор, а сви даљи задаци (класификација, груписање, претраживање, верификација) могу се свести на једноставне геометријске операције над тим векторима.

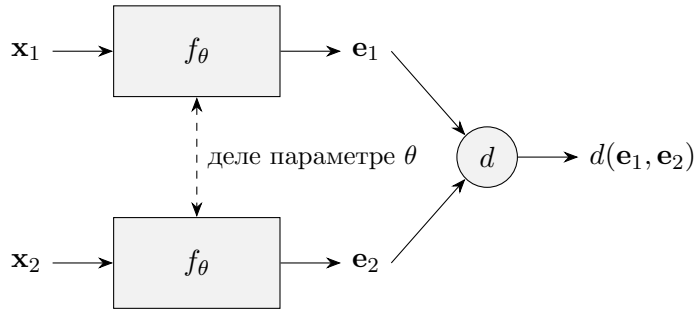
Овакав приступ доноси неколико практичних предности. Прво, једном научене векторске репрезентације често се могу пренети на нове задатке без поновног обучавања, што је основа *преноса знања* (енг. *transfer learning*). Друго, поређење векторских репрезентација је јефтино: за два вектора $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^d$ довољно је израчунати њихово еуклидско растојање $\|\mathbf{e}_1 - \mathbf{e}_2\|_2$ или косинусну

сличност $\cos(\mathbf{e}_1, \mathbf{e}_2) = \frac{\mathbf{e}_1 \cdot \mathbf{e}_2}{\|\mathbf{e}_1\| \, \|\mathbf{e}_2\|}$. Треће, у задацима верификације и идентификације, нови ентитет (на пример идентитет чије потписе модел никада није видео) може се додати у систем простим израчунавањем његове векторске репрезентације, без поновног обучавања.

Главно питање које преостаје је: како обучити модел да производи управо такве векторске репрезентације у којима семантичка сличност одговара геометријској близини? Једна од најранијих и најутицајнијих архитектура управо за тај циљ су *сијамске неуронске мреже*, описане у наставку.

2.4 Сијамске неуронске мреже

Сијамске неуронске мреже представљене су 1993. године у раду Бромлија и сарадника, и њихова прва примена је била управо у домену верификације руком писаних потписа [1]. Назив су добиле по сијамским близанцима: архитектура се састоји од две идентичне гране које *деле параметре*, при чему свака грана прима по један улаз и производи његову векторску репрезентацију. На крају се израчунава функција растојања између две векторске репрезентације, која представља меру различитости два улаза. Шематски приказ архитектуре дат је на слици 2.3.



Слика 2.3: Шематски приказ сијамске неуронске мреже.

Кључни аспект архитектуре је то што су тежине θ *дељене* између обе гране. Ово није само техничка ситница, већ суштинска особина: дељење тежина обезбеђује да поређење улаза буде *симетрично*: ако се на оба улаза пропусти исти узорак, нужно се добијају идентичне векторске репрезентације, а поређење улаза $(\mathbf{x}_1, \mathbf{x}_2)$ даје исти резултат као поређење $(\mathbf{x}_2, \mathbf{x}_1)$. Када би свака грана имала независне параметре, мрежа би у начелу могла да научи различите репрезентације за исте објекте, чиме би поређење изгубило смисао. Поред

тога, дељење параметара значајно смањује укупан број параметара који треба научити и доводи до боље генерализације на ограниченим скуповима података, што је за домен потписа од велике важности будући да је за сваку особу типично доступан само мали број узорака.

Обучавање сијамске мреже одвија се на паровима улаза $(\mathbf{x}_1, \mathbf{x}_2)$ заједно са ознаком $y \in \{0, 1\}$ која говори да ли пар припада истој класи. Циљ је научити мрежу да за позитивне парове ($y = 1$) производи векторске репрезентације малог међусобног растојања, а за негативне парове ($y = 0$) векторске репрезентације великог међусобног растојања. У оригиналном раду [1], ово је постигнуто једноставним функцијама грешке, али је приступ значајно унапређен увођењем *контрастне функције грешке* (енг. *contrastive loss*) [2], у облику са маргином који је усталио рад [8]:

$$\mathcal{L}_{\text{contrastive}}(\mathbf{x}_1, \mathbf{x}_2, y) = y \cdot d^2 + (1 - y) \cdot \max(0, m - d)^2, \tag{2.3}$$

где је $d = \|\mathbf{e}_1 - \mathbf{e}_2\|_2$ растојање векторских репрезентација, а $m > 0$ маргина која дефинише жељени минимални размак између векторских репрезентација негативних парова. Прва компонента, активна за позитивне парове, директно повлачи њихове векторске репрезентације ка истој тачки. Друга компонента, активна за негативне парове, њих одбија све док им растојање не пређе маргину m, након чега је њихов допринос градијенту једнак нули. На тај начин се векторске репрезентације истог идентитета окупљају у компактан кластер, а кластери различитих идентитета бивају међусобно одвојени за бар m.

У фази закључивања, обучена сијамска мрежа користи се на следећи начин. За дати пар $(\mathbf{x}_1, \mathbf{x}_2)$ за који треба утврдити да ли припада истој класи, израчунавају се векторске репрезентације и њихово растојање. Ако је $d(\mathbf{e}_1, \mathbf{e}_2) < \tau$, где је τ унапред одабрани праг, пар се прихвата као позитиван (исти идентитет); иначе се одбацује. Овакав механизам омогућава да се верификација спроводи и над идентитетима који нису постојали у тренинг скупу, што је у домену верификације потписа од суштинског значаја: у пракси сваки нови потписник треба да буде укључен у систем без поновног обучавања модела. Поред директне примене у виду парне верификације, обучена грана сијамске мреже се у реалним системима најчешће користи за издвајање векторских репрезентација, тако што се сваки нови потпис пропусти кроз мрежу и сачува његов вектор у бази идентитета, чиме се верификација своди на једноставно поређење вектора.

Дубоко учење сличности, којем је посвећен наредни одељак, у уопштеном облику почива на истим идејама: учењу из поређења, дељењу параметара и учењу векторских репрезентација на основу односа међу узорцима, а не дискретних ознака класа.

2.5 Дубоко учење сличности

Дубоко учење сличности (енг. *deep metric learning*) представља природно уопштење сијамских мрежа: уместо експлицитне сијамске архитектуре, циљ постаје да се научи једна параметризована функција угњеждења $f_\theta : \mathcal{X} \to \mathbb{R}^d$ таква да растојање у простору векторских репрезентација одражава семантичку сличност. Формално, за било која два узорка $\mathbf{x}_i$ и $\mathbf{x}_j$, желимо да важи

$$d(f_\theta(\mathbf{x}_i),\, f_\theta(\mathbf{x}_j)) = \begin{cases} \text{мало}, & \text{ако } \mathbf{x}_i \sim \mathbf{x}_j, \\ \text{велико}, & \text{ако } \mathbf{x}_i \not\sim \mathbf{x}_j, \end{cases} \tag{2.4}$$

где је d изабрана метрика (најчешће еуклидска или косинусна), а $\sim$ релација семантичке сличности (нпр. припадност истом идентитету). У контексту овог рада, два потписа сматрају се сличним ако припадају истом потписнику, а различитим ако припадају различитим потписницима или ако је један фалсификат другог.

Породица функција грешке у дубоком учењу сличности развијала се постепено, а њени представници разликују се пре свега по томе на колико узорака истовремено делују и каквим оптимизационим механизмом их повезују. *Контрастна функција грешке* (енг. *contrastive loss*) [2], описана у одељку 2.4, ради на паровима узорака и представља најједноставнији облик учења сличности. *Тријлет функција грешке* (енг. *triplet loss*) [24] ради на уређеним тројкама (референтном узорку, позитивном и негативном примеру) и уместо апсолутне маргине намеће *релативну*: позитиван пример мора бити ближи референтном узорку него негативан. Уопштене функције попут *подигнуте структурне* (енг. *lifted structured loss*) [27] и *N-парне* (енг. *N-pair loss*) [26] истовремено разматрају све негативне примере у мини-групи по једном референтном узорку, чиме у једно ажурирање параметара улази информација о свим негативним паровима, а не само о једном. *Функција грешке вишеструке сличности* (енг. *multi-similarity loss*) [30] надограђује ову идеју тако што допринос сваког пара множи коефицијентом који зависи од сличности тог пара: за позитивне

парове коефицијент опада са сличношћу, а за негативне расте. Тако најинформативнији парови (позитивни са малом и негативни са великом сличношћу) улазе у функцију грешке са највећим коефицијентом.

У наставку се детаљније описују *тријплет* функција грешке и функција *вишеструке сличности*, које су у овом раду коришћене као експериментална основа. Триплет функција грешке, популаризована радом на верификацији лица FaceNet [24], дефинише се над уређеним тројкама $(\mathbf{a}, \mathbf{p}, \mathbf{n})$, у којима је $\mathbf{a}$ референтни узорак (енг. *anchor*), $\mathbf{p}$ позитиван пример из исте класе, а $\mathbf{n}$ негативан пример из друге класе:

$$\mathcal{L}_{\text{triplet}}(\mathbf{a}, \mathbf{p}, \mathbf{n}) = \max\big(0,\ \|f(\mathbf{a}) - f(\mathbf{p})\|_2^2 - \|f(\mathbf{a}) - f(\mathbf{n})\|_2^2 + m\big)\,. \tag{2.5}$$

Идеја је да позитиван пример буде ближи референтном узорку него негативан, и то најмање за маргину m. Оваква функција пружа директну, локалну контролу над геометријом у околини сваког референтног узорка. *Функција грешке вишеструке сличности* [30] истовремено разматра *све* позитивне и негативне парове сваког референтног узорка у мини-групи и сваком пару додељује тежину пропорционалну његовој информативности: тешки позитивни парови (са ниском сличношћу) добијају јаче привлачење, а тешки негативни парови (са високом сличношћу) јаче одбијање. На тај начин, једно ажурирање параметара укључује глобалну информацију из читаве мини-групе, што се у пракси показује изузетно ефикасним.

Једно од најважнијих запажања у дубоком учењу сличности [31] јесте да *избор тројки или парова који улазе у функцију грешке* често утиче на коначан учинак подједнако као сама формула функције грешке, или чак и више од ње. Разлог је једноставан: ако се тројке бирају насумично, већина њих биће *лака* (позитиван пример је већ много ближи референтном узорку него негативан), па у таквим тројкама функција грешке даје градијент близу нуле и практично не доприноси учењу. Зато се у савременим системима пажљиво раздвајају две улоге:

- *Узорковање* (енг. *sampling*) одређује *како се формира мини-група*: насумично, балансирано по класама, или по некој другој стратегији која обезбеђује довољно информативних парова за учење.
- *Рударење* (енг. *mining*) одређује *које парове или тројке унутар већ формираних мини-група* прослеђује функцији грешке: типично се бирају

тешки (енг. *hard*) или *полутешки* (енг. *semi-hard*) примери који највише доприносе учењу.

Управо у та два корака, узорковању и рударењу, лежи кључна идеја овог рада. Тиме што узорковање фалсификате распоређује по мини-групама, а рударење их бира као најинформативније примере за функцију грешке, општи оквир дубоког учења сличности претвара се у систем посебно прилагођен верификацији потписа са вешто израђеним фалсификатима. Овај приступ, *свестан фалсификата* (енг. *forgery-aware*), детаљно се разрађује у поглављу 4.

Глава 3

Верификација потписа: преглед области

Теоријски појмови из претходног поглавља добијају пун смисао тек у области у којој се примењују. Верификација руком писаних потписа развија се већ деценијама, са јасном поделом на онлајн и офлајн поставку и богатом историјом метода, од класичних, заснованих на ручно осмишљеним особинама, до савремених приступа заснованих на дубоком учењу сличности. Преглед који следи прати тај развој и издваја ограничења постојећих метода из којих непосредно произлази мотивација овог рада.

3.1 Онлајн и офлајн верификација

Технологија верификације руком писаних потписа дели се на две фундаментално различите гране према начину на који се потпис прикупља и обрађује. У *онлајн* (енг. *online* или *dynamic*) поставци, потпис се снима у самом тренутку настанка, најчешће помоћу таблета осетљивог на додир или дигиталне оловке, па систем има на располагању богат скуп динамичких сигнала: путању оловке у функцији времена, вертикалну компоненту притиска, угао нагиба оловке, брзину и убрзање покрета, као и временске ознаке почетка и завршетка сваког потеза. Ови сигнали носе информацију о томе *како* је потпис настао, а не само о томе *шта* се види у његовом завршеном изгледу. У *офлајн* (енг. *offline* или *static*) поставци, систем има приступ једино финалној статичкој слици потписа (скенираном документу или фотографији), без иједног трага о динамици настанка [5, 13].

Из ове структурне разлике произлазе и различити изазови. Онлајн системи су у начелу прецизнији, јер динамички сигнали значајно помажу разликовању аутентичног потписа од имитације: фалсификатор може да опонаша облик, али тешко може да репродукује идентичан темпо, притисак и редослед потеза који карактеришу конкретног потписника. Међутим, онлајн поставка захтева наменски хардвер у самом тренутку потписивања, што је практично применљиво само у строго контролисаним сценаријима, попут експозитуре банке. У огромном броју случајева који имају правну вредност, потпис се ставља на папир, потом скенира или фотографише, и тек тада се обрађује рачунарски. Управо ту наступа офлајн верификација, која ради са оним што је у пракси најчешће доступно: статичком сликом [10].

Поред поделе по начину прикупљања, постоји и стандардна подела фалсификата по начину настанка. *Насумични* фалсификати (енг. *random forgeries*) су аутентични потписи других особа употребљени уместо потписа циљног потписника, без икаквог покушаја имитације. *Једноставне* фалсификате (енг. *simple forgeries*) производи неко ко зна само име циљног потписника, без увида у изглед оригиналног потписа. *Вешто израђене* фалсификате (енг. *skilled forgeries*) производи особа која је пажљиво проучила оригинални потпис и активно покушава да га опонаша [12]. Прва два типа готово сваки систем може поуздано да детектује, јер је њихова сличност са оригиналом површна или непостојећа. Управо вешто израђени фалсификати чине срж проблема и њима је посвећена највећа пажња у целој модерној литератури. Овај рад је усмерен искључиво на офлајн верификацију са вешто израђеним фалсификатима као главним изазовом.

3.2 Класични приступи

До средине прошле деценије, офлајн верификација потписа била је доминантно заснована на двостепеном приступу: експерт би прво ручно дефинисао низ *особина* (енг. *features*) које се извлаче из слике потписа, а затим би се над њима обучавао стандардни класификатор.

Особине су се грубо груписале у неколико породица: *геометријске* (нпр. однос ширине и висине или дужина контура), *дескрипторе облика* (Зерникеови и Хуови моменти), *текстурне* (LBP, HOG, GLCM) и *усмерене* особине засноване на градијентним операторима [12].

Над извученим вектором особина најчешће се обучавала *метода потпорних вектора* (енг. *SVM*), а користили су се и скривени Марковљеви модели и, у онлајн домену, динамичко поравнање временских серија (енг. *dynamic time warping*), којим се пореде сигнали оловке два потписа. Разликују се приступи *зависни од потписника*, у којима се за сваку особу обучава засебан модел, и *независни од потписника*, у којима један модел учи општу функцију сличности примењиву и на нове особе [29].

Ови приступи донели су значајан напредак, али и границе својствене ручно конструисаним особинама: оне су уско прилагођене унапред дефинисаним обрасцима, па се тешко преносе између различитих скупова података, услова скенирања и стилова писања [5]. Управо то ограничење отворило је простор за дубоке моделе, који особине уче директно из података.

3.3 Приступи засновани на дубоком учењу сличности

Прелазак на дубоке конволутивне мреже потпуно је изменио слику офлајн верификације потписа. Хафеман и сарадници [9, 10] први су показали да мрежа може директно из сирове слике да научи особине које надмашују све раније ручно конструисане скупове особина; њихов хибридни рецепт, дубоке репрезентације праћене класичним верификатором зависним од потписника, донео је велики скок у тачности и остао дуго утицајан.

Паралелно се развио и чисто метрички правац. Сијамска мрежа SigNet [4], обучавана контрастном функцијом грешке, постала је широко прихваћен верификатор независан од потписника и до данас служи као основа за поређење. На истом трагу развили су се приступи који дубоке конволутивне особине прослеђују засебном верификатору независном од потписника [28], као и хибридни модели који спајају одлуке верификатора независног и зависног од потписника [32]. Ранч и сарадници [23] први су на овај задатак применили триплет функцију грешке и тиме га повезали са општим напретком у учењу метричких репрезентација.

Одатле се поље гранало у више смерова: од спајања дубоких мрежа са удаљеношћу уређивања графова [18] и структурно богатијих репрезентација потписа [34], преко вишестепених поступака који дубоке мреже комбинују са класичним детекторима кључних тачака [22] и синтетичких потписа који у

онлајн домену ублажавају хронични недостатак података [16], до повратка једнокласном приступу зависном од потписника са дубоким особинама [29] и најновијих генеративних, дифузионих модела [33]. Упркос разноликости, готово сви деле исти скелет: дубоку мрежу која потпис пресликава у векторски простор у коме се доноси одлука о аутентичности.

Ипак, кроз сву ту литературу провлаче се две празнине које мотивишу овај рад. Прво, већина метода *не моделује експлицитно* однос између аутентичног потписа и његовог фалсификата: фалсификати се или користе тек у евалуацији, или се третирају као обични негативни узорци без везе са идентитетом који имитирају. Тако научени простор добро раздваја *идентитете*, али не и аутентичне потписе од њихових вешто израђених имитација. Друго, не постоји систематско поређење различитих функција грешке под истим, контролисаним условима и уз усклађен избор узорковања и рударења тешких примера. На тај недостатак је, у општем контексту дубоког учења сличности, већ указано [19]. Управо ове две празнине попуњава методологија изложена у наредном поглављу.

Глава 4

Методологија

Методологија почива на три стуба: формалној поставци проблема верификације, избору архитектуре и функција грешке, и стратегији обучавања свесној фалсификата. У наставку се сваки од њих излаже редом, а њихово спајање приказује се јединственим дијаграмом система.

4.1 Формална поставка проблема верификације

Пре описа самог приступа, задатак верификације треба поставити прецизно. Нека је $\mathcal{X}$ скуп свих слика потписа, а $\mathcal{I}$ скуп идентитета (потписника) присутних у систему. Сваком потпису $\mathbf{x} \in \mathcal{X}$ придружени су ознака идентитета $i \in \mathcal{I}$ и индикатор аутентичности $f \in \{0, 1\}$, при чему $f = 0$ означава аутентичан потпис, а $f = 1$ фалсификат.

Верификација потписа формулише се на следећи начин: за пар $(\mathbf{x}_q, i)$, у којем је $\mathbf{x}_q$ потпис који се верификује, а i тврђени идентитет[1], систем треба да донесе бинарну одлуку о томе да ли потпис заиста припада идентитету i. У оквиру дубоког учења сличности, ова одлука се доноси преко обучене функције угњеждења $f_\theta : \mathcal{X} \to \mathbb{R}^d$. За идентитет i из његових расположивих аутентичних узорака израчунава се прототип $\hat{\mathbf{p}}_i \in \mathbb{R}^d$ (у наставку рада биће наведена формула којом се рачуна), а одлука се своди на поређење косинусне сличности вектора потписа који се верификује са прототипом и проверу прага

[1]Тврђени идентитет је особа за коју се тврди да је потпис њен, на пример она наведена на документу. То је део улаза у систем, а не податак који се из слике изводи; особа која покушава да фалсификује потпис тврди да има туђи идентитет.

τ:

$$\cos\big(f_\theta(\mathbf{x}_q),\, \hat{\mathbf{p}}_i\big) \geq \tau \implies \text{потпис се прихвата.} \tag{4.1}$$

Циљ обучавања је да се параметри θ науче тако да аутентичан потпис који се верификује буде близак прототипу свог идентитета, а фалсификован, независно од његове физичке сличности са оригиналом, довољно удаљен.

Управо у овој последњој реченици налази се суштинска тешкоћа разматрана у целом овом раду. Док су потписи различитих особа, било аутентични било насумични фалсификати, у простору пиксела изразито различити, *вешто израђени фалсификати* су намерно конструисани да наликују циљном потпису. Њихова визуелна сличност са оригиналом је врло висока и без посебне интервенције та сличност се пресликава и у простор векторских репрезентација. Стандардни приступи дубоког учења сличности фалсификате третирају као обичне негативне узорке, чиме оптимизују модел да добро разликује *идентитете*, али не нужно и да разликује *аутентичне потписе од њихових циљаних имитација*. На тај начин, граница између два кластера потписа истог идентитета (аутентичних и фалсификованих) остаје у простору векторских репрезентација слабо дефинисана.

Ипак, у тренинг скупу имамо једну специфичну информацију коју претходни приступи нису систематски користили: за сваки фалсификат знамо *ког конкретног потписника* имитира. То значи да је за сваког потписника i могуће формирати парове $(\mathbf{x}_i^{\text{gen}}, \mathbf{x}_i^{\text{fals}})$ његовог аутентичног потписа и његовог фалсификата, који представљају управо ону врсту тешких примера на које систем треба да буде припремљен. Питање је, међутим, како ову информацију унети у обучавање. Због тога је фокус методологије овог рада постављен на два нивоа *пре* саме формуле функције грешке: на *формирање мини-група* тако да оне увек садрже довољно парова оригинал–фалсификат истог идентитета, и на *избор најинформативнијих узорака* унутар мини-групе који ће заиста улазити у функцију грешке. Управо се на тим двема улогама, узорковању и рударењу, гради остатак методологије: функције грешке које их користе и целокупан ток система излажу се у наставку.

4.2 Архитектура модела

Архитектура модела заснована је на конволутивној мрежи ResNet18 [11], претходно обученој на скупу за класификацију ImageNet [3]. Коришћење мре-

же обучаване на општем визуелном задатку као основе модела доноси неколико предности: језгра (филтери) из почетних и средњих слојева су већ научила корисне обрасце (ивице, текстуре, локалне облике) применљиве и на потписе, што значајно убрзава конвергенцију и побољшава генерализацију у поставкама са ограниченом количином података.

Изворна ResNet18 архитектура завршава се потпуно повезаним слојем димензија 512×1000, намењеним класификацији у 1000 класа ImageNet такмичења, чији нам излаз у задатку верификације потписа није користан. Због тога је тај завршни слој замењен **новим линеарним слојем** димензија $512 \times d$, где је $d = 512$ изабрана димензија простора векторских репрезентација. Сврха увођења овог слоја је двојака. Прво, он одваја врх мреже од конкретне поставке ImageNet-а и омогућава да се у фази обучавања учи управо она пројекција која је најкориснија за поређење потписа. Друго, он служи и као *пројекциона глава*: последње особине мреже трансформише у репрезентацију специјализовану за метричко поређење у простору сличности.

Излаз модела је, дакле, вектор $\mathbf{e} = f_\theta(\mathbf{x}) \in \mathbb{R}^d$ за сваки улазни потпис. Овај вектор се пре било каквог поређења *L_2-нормализује*, чиме сви вектори леже на јединичној сфери, а њихова косинусна сличност своди се на скаларни производ:

$$\hat{\mathbf{e}} = \frac{\mathbf{e}}{\|\mathbf{e}\|_2}, \qquad \cos(\mathbf{e}_1, \mathbf{e}_2) = \langle \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2 \rangle. \tag{4.2}$$

Из ове перспективе, обучен модел постаје *екстрактор векторских репрезентација*, а сама верификација своди се на следеће кораке. За сваки познати идентитет i из његових расположивих аутентичних потписа израчунавају се вектори $\{\mathbf{e}_k^{(i)}\}_{k=1}^{K_i}$, и они се чувају у *векторској бази идентитета*. Поред саме листе вектора, у бази се за сваког потписника одржава и *прототип* — нормализовани просек његових аутентичних вектора:

$$\mathbf{p}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \hat{\mathbf{e}}_k^{(i)}, \qquad \hat{\mathbf{p}}_i = \frac{\mathbf{p}_i}{\|\mathbf{p}_i\|_2}. \tag{4.3}$$

Ова листа се ажурира инкрементално: када стигне нова потврђена аутентична слика идентитета i, њен вектор се додаје у листу тог идентитета, а његов прототип се поново израчунава. Када стигне потпис $\mathbf{x}_q$ који се верификује, систем израчунава његов вектор $\mathbf{e}_q = f_\theta(\mathbf{x}_q)$ и упоређује га са прототипом тврђеног идентитета i; одлука се доноси на основу прага:

$$s = \cos(\mathbf{e}_q, \hat{\mathbf{p}}_i) \geq \tau \implies \text{прихвати, иначе одбаци.} \tag{4.4}$$

Праг τ се одређује посебно за сваки модел на основу валидационог скупа, а његов избор је детаљно обрађен у поглављу 5.

4.3 Избор функције грешке

Разматрају се три варијанте функције грешке: *тријлет* (енг. *triplet*), *вишеструке сличности* (енг. *multi-similarity*) и њихова *хибридна* комбинација, предложена у овом раду. Триплет функција и функција вишеструке сличности постижу исти циљ на два различита начина, а хибридна варијанта спаја њихове предности.

Триплет функција грешке. Триплет функција грешке, популаризована радом FaceNet [24], делује над тројкама $(\mathbf{a}, \mathbf{p}, \mathbf{n})$ где је $\mathbf{a}$ референтни узорак, $\mathbf{p}$ позитиван пример из исте класе, а $\mathbf{n}$ негативан пример из друге класе:

$$\mathcal{L}_{\text{triplet}}(\mathbf{a}, \mathbf{p}, \mathbf{n}) = \max\big(0,\ d(\mathbf{a}, \mathbf{p}) - d(\mathbf{a}, \mathbf{n}) + m\big), \tag{4.5}$$

где је d изабрана метрика (овде еуклидско растојање нормализованих вектора), а $m > 0$ маргина. Функцију треба читати овако: њена вредност је позитивна само док је негативни пример ближи референтном узорку од позитивног, или их раздваја мање од маргине m. У том случају корак градијентног спуста смањује $d(\mathbf{a}, \mathbf{p})$, а повећава $d(\mathbf{a}, \mathbf{n})$, све док размак међу њима не пређе m; чим је услов испуњен, вредност постаје нула и тројка престаје да утиче на учење. Захваљујући оператору $\max(0, \cdot)$, у обзир се тако узимају само *тешке* тројке, оне које још нису довољно раздвојене (слика 4.1).

Главна снага триплет функције је *локална, прецизна* контрола: она захтева да у околини сваког референтног узорка позитиван пример буде ближи од негативног за најмање m, па је посебно погодна за раздвајање појединачних тешких случајева у којима су аутентичан потпис и његов фалсификат веома слични у простору слика. Њена мана је, међутим, *сужености погледа*: свака тројка узима у обзир само један позитиван и један негативан пример, па глобалну структуру простора не користи непосредно. Осим тога, веома је осетљива на *избор тројки*: ако се бирају насумично, већина је лака и има градијент једнак нули, због чега учење може бити споро и нестабилно без пажљивог одабира информативних примера.

(а) пре учења (б) после учења

циљ: $d(a,p) + m \leq d(a,n)$

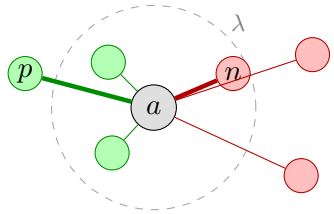
Слика 4.1: Дејство триплет функције грешке на тројку (a, p, n).

Функција вишеструке сличности. Функција вишеструке сличности, представљена у раду [30], истовремено разматра све позитивне и негативне парове сваког референтног узорка у мини-групи и сваком пару додељује тежину пропорционалну његовој информативности. Њен облик је

$$\mathcal{L}_{\text{MS}} = \frac{1}{\alpha}\log\Big[1 + \sum_{k\in\mathcal{P}_i} e^{-\alpha(S_{ik}-\lambda_{\text{MS}})}\Big] + \frac{1}{\beta}\log\Big[1 + \sum_{k\in\mathcal{N}_i} e^{\beta(S_{ik}-\lambda_{\text{MS}})}\Big], \tag{4.6}$$

где је S_{ik} косинусна сличност између референтног узорка i и узорка k, $\mathcal{P}_i$ скуп његових позитивних, а $\mathcal{N}_i$ скуп негативних парова; α, β су параметри скале, а λ_{MS} померај (праг око кога се пар сматра тешким). Функцију чине два члана. Први делује на *све* позитивне парове одједном и привлачи их ка референтном узорку, при чему теже позитиве (оне са ниском сличношћу) наглашава јаче; други делује на *све* негативне парове и одбија их, наглашавајући теже негативе (оне са високом сличношћу). Због облика са логаритмом збира експоненцијала, оба члана понашају се као *мека* верзија максимума, па се пажња аутоматски усмерава на најтеже парове, а сваки пар добија тежину сразмерну својој информативности (слика 4.2).

Главна снага ове функције је *глобална* структура: једно ажурирање параметара користи информације из читаве мини-групе истовремено, тако што се сваки референтни узорак пореди са свим својим позитивима и негативима, а не са само једним паром као код триплет функције. У пракси се показала изузетно ефикасном за организовање читавог простора векторских репрезентација по идентитетима. Њено ограничење је, међутим, то што квалитет учења зависи од тога да ли мини-група садржи довољно информативних (тешких) парова, чиме део одговорности прелази на начин њеног формирања.



Слика 4.2: Функција вишеструке сличности: позитивни (зелено) и негативни (црвено) парови референтног узорка a у мини-групи.

Хибридна функција грешке. Триплет и функција вишеструке сличности дају две различите врсте информације које су у задатку верификације потписа *комплементарне*. Функција вишеструке сличности добро организује глобалну структуру простора, то јест успешно раздваја аутентичне потписе различитих идентитета. Триплет функција пружа локалну прецизност, јер се може прецизно усмерити на парове оригинал–фалсификат и приморати их да и у том најосетљивијем суседству буду довољно удаљени. У овом раду уводи се хибридна функција грешке која експлицитно искоришћава ову поделу одговорности:

$$\mathcal{L}_{\text{hybrid}} = \mathcal{L}_{\text{MS}}^{\text{gen}} + \lambda \cdot \mathcal{L}_{\text{triplet}}^{\text{fals}}, \tag{4.7}$$

где $\mathcal{L}_{\text{MS}}^{\text{gen}}$ означава функцију вишеструке сличности примењену искључиво на парове аутентичних потписа (где негативне парове чине аутентични потписи различитих идентитета), $\mathcal{L}_{\text{triplet}}^{\text{fals}}$ је триплет функција грешке примењена на тројке $(\mathbf{a}, \mathbf{p}, \mathbf{n}_f)$ у којима су референтни узорак и позитиван пример аутентични потписи истог идентитета, а $\mathbf{n}_f$ је *фалсификат истог идентитета*. Параметар $\lambda \geq 0$ балансира два доприноса: за $\lambda = 0$ губи се триплет компонента и хибрид се своди на чисту вишеструку сличност, док се за велике вредности λ приоритет даје одлуци о фалсификатима. На тај начин, хибридна функција омогућава мрежи да истовремено учи и глобалну и локалну структуру простора, при чему свака компонента ради тачно над оним типом парова за који је најкориснија.

4.4 Обучавање свесно фалсификата

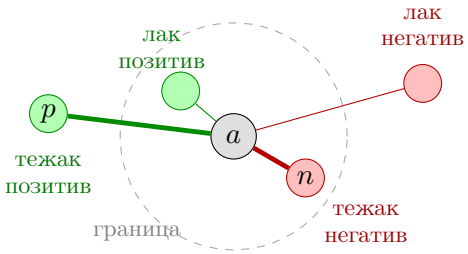
Кључни допринос овог рада не лежи у самим формулама функција грешке, већ у томе *како се мини-групе формирају и који примери из њих улазе у функцију грешке*. За то су задужене две компоненте: *узорковање*, које саставља мини-групе, и *рударење*, које из њих издваја најинформативније парове. Обе се овде уводе у варијантама *свесним фалсификата*, прилагођеним да фалсификате ставе у средиште учења.

Узорковање. Једноставно насумично узимање узорака из тренинг скупа доводи до два проблема. Прво, само мали део идентитета у скупу има придружене фалсификате, па уз насумично формирање мини-група велики део њих ефективно нема ниједан пар оригинал–фалсификат, те у функцију грешке не улази ниједан члан који се односи на фалсификате. Друго, чак и када су у мини-групи присутни оригинал и фалсификат, њихова појава је случајна и неуравнотежена.

Због тога је развијено *узорковање по потписницима* специјално прилагођено задатку. У сваком кораку, узорковање прво бира одређен скуп потписника који ће чинити мини-групу, при чему је избор пристрасан у корист потписника који имају бар један фалсификат, али тако да и потписници без фалсификата редовно долазе на ред ротацијом редова чекања, чиме се не гуши учење међуидентитетске структуре. Затим се за сваког изабраног потписника узима неколико његових аутентичних потписа и, ако их има, неколико његових фалсификата. На тај начин, скоро свака мини-група садржи валидне парове оригинал–фалсификат истог идентитета, па се у сваком кораку истовремено оптимизују и међуидентитетско раздвајање и раздвајање оригинала од фалсификата.

Кључни параметри овог узорковања су: *број потписника по мини-групи*, *број аутентичних потписа по потписнику*, *број фалсификата по потписнику* (где они постоје), и *вероватноћа пристрасности* која контролише колико чешће ће у мини-групу ући потписници са фалсификатима у поређењу са онима без њих. Конкретне вредности свих ових параметара коришћене у експериментима наведене су у поглављу 5. Уз ово, контролише се и *ефективна дужина епохе*, то јест број мини-група обрађених између две евалуације на валидационом скупу.

Рударење. Други део поступка тиче се избора *најинформативнијих* парова или тројки унутар већ формираних мини-група. Разлог је интуитиван: ако у функцију грешке убацимо све могуће парове, велики део њих биће *лак* (позитиван пар је већ много ближи референтном узорку него било који негативан) и за тај пар функција грешке производи градијент близу нуле. Са једне стране, такав пар *не доприноси* учењу; са друге стране, његов нулти члан улази у просек по свим изабраним паровима и сразмерно умањује норму укупног градијента. Зато се користи рударење: оно унутар мини-групе задржава само оне парове или тројке у којима је гранични услов функције грешке нарушен или је близу нарушавања, то јест оне у којима још увек постоји простор за учење (слика 4.3).



Слика 4.3: Лаки и тешки примери из угла рударења, за референтни узорак a.

Будући да триплет и функција вишеструке сличности раде над различитим структурама (тројкама, односно паровима), за потребе овог рада развијена су **два одвојена поступка рударења свесна фалсификата**, по један за сваку од њих. Хибридна поставка користи оба истовремено: рударење за функцију вишеструке сличности обезбеђује парове за компоненту $\mathcal{L}_{\text{MS}}^{\text{gen}}$, а рударење за триплет функцију грешке обезбеђује тројке за компоненту $\mathcal{L}_{\text{triplet}}^{\text{fals}}$. И поред разлике у излазу, оба поступка деле иста два принципа дизајна свесног фалсификата. Прво, *референтни узорак увек мора бити аутентичан потпис*, јер модел никада не учи да фалсификат буде близак неком прототипу, већ само супротно. Друго, скуп негатива сваког референтног узорка дели се на *међуидентитетске негативе* (аутентични потписи других потписника) и *унутаридентитетске негативе* (фалсификати истог потписника), а избор тешких негатива врши се *унутар сваког подскупа засебно* пре него што се споје. Ова два подскупа притом не служе истом циљу, већ одговарају *две-*

ма суштински различитим врстама грешке које систем при верификацији потписа може направити:

- *Међуидентитетски негативи* (аутентични потписи *других* потписника) уче модел да разликује идентитете, то јест да потпис једног потписника не помеша са потписом другог. Они обликују *глобалну* структуру простора репрезентација, у којем свака особа заузима засебну област.
- *Унутаридентитетски негативи* (фалсификати *истог* потписника) уче модел да, унутар већ издвојене области једног идентитета, раздвоји његове аутентичне потписе од фалсификата. Они су у средишту самог задатка откривања фалсификата и чине његов најтежи и најосетљивији део.

Заједно, ова два подскупа покривају оба начина на која верификација може да омане: прихватање туђег аутентичног потписа (међуидентитетска грешка) и прихватање фалсификата (унутаридентитетска грешка). Разлог за њихово *засебно* рударење је практичан: аутентичних потписа других потписника по правилу има много више него фалсификата истог потписника, па би, када би сви негативи чинили један јединствен скуп, бројнији међуидентитетски негативи чинили највећи део чланова функције грешке и тиме потиснули допринос ретких, али критичних фалсификата. Бирањем тешких негатива унутар сваког подскупа засебно, уз горњу границу на број изабраних примера по подскупу, обезбеђује се да оба типа грешке доследно доприносе учењу.

Рударење за функцију вишеструке сличности ради на нивоу појединачних парова. За сваки референтни узорак, оно независно у сваком од два негативна подскупа задржава оне парове чија сличност прелази *праг тежине*, а потом за позитивне парове задржава оне чија сличност пада испод прага одређеног на основу спојеног скупа негатива. Тиме у функцију грешке улази скуп свих тешких парова (референтни узорак, негатив) и (референтни узорак, позитив) из мини-групе.

Рударење за тројлет функцију грешке ради на нивоу уређених тројки $(\mathbf{a}, \mathbf{p}, \mathbf{n})$. За сваки пар референтног узорка и позитива, у сваком од два негативна подскупа бира тешке негативе и формира тројке. Између могућих тројки преферирају се полутешке (енг. *semi-hard*) — оне у којима је негативан пример већ удаљенији од референтног узорка него позитиван, али их раздваја мање од маргине, па тројка и даље нарушава гранични услов и производи

ненулти градијент. Оваквим избором се свесно заобилазе најтеже тројке, у којима је негатив ближи референтном узорку од позитива, будући да оне у раној фази обучавања воде ка лошим локалним минимумима и урушавању модела [24].

Параметри рударења свесног фалсификата за функцију вишеструке сличности укључују *праг тежине*, који одређује ширину појаса сличности унутар којег се пар сматра тешким, и *горњу границу броја негатива по сваком подскупу*, која спречава да многобројнији подскуп доминира у укупној вредности функције грешке. Параметри рударења за триплет функцију грешке укључују *маргину* која одређује очекивано раздвајање позитивних и негативних парова, *режим избора тешких тројки* (нпр. полутешки, тешки или насумични), и *горње границе на број изабраних негатива по подскупу*, засебно за подскуп међуидентитетских и подскуп унутаридентитетских (фалсификатних) негатива. Конкретне вредности свих ових параметара дате су у поглављу 5.

Ток обучавања. Уз све претходно описане компоненте, један корак обучавања тече на следећи начин:

1. Узорковање формира мини-групу индекса узорака.
2. Сваком индексу одговара слика потписа. На њу се у фази обучавања примењује скуп стохастичких *аугментација* који служи као регуларизатор.
3. Аугментоване слике пропуштају се кроз мрежу и за сваку се добија векторска репрезентација $\mathbf{e}$, која се потом L_2-нормализује.
4. Рударење на основу мини-групе векторских репрезентација и припадајућих ознака идентитета и фалсификата формира скуп информативних парова или тројки.
5. На том скупу израчунава се изабрана функција грешке (триплет, вишеструке сличности или хибридна).
6. Алгоритмом пропагације грешке уназад израчунавају се градијенти, а параметри мреже ажурирају изабраним оптимизационим алгоритмом.

Конкретан скуп аугментација, избор оптимизатора и његови хиперпараметри, заједно са распоредом корака учења и критеријумом по којем се прати квалитет обучавања, биће детаљно описани у поглављу 5. Целокупни ток за-

кључивања, заједно са управљањем векторском базом идентитета, графички је приказан у наредном одељку.

4.5 Ток система

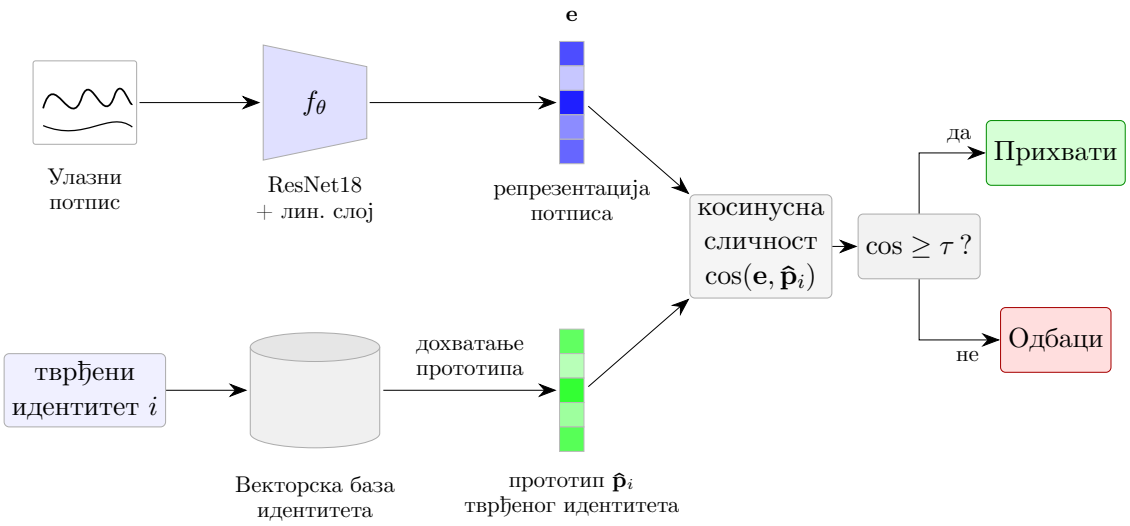
Слика 4.4 приказује цео ток *закључивања* (енг. *inference*), то јест пут којим систем у радном режиму пролази од сирове скениране слике до коначне одлуке да потпис прихвати или одбаци. У овој фази мрежа је већ обучена и њени параметри θ су фиксирани. Кораци који су потребни само током обучавања, попут аугментације и узорковања или рударења и пропагације грешке уназад, у радном режиму се не изводе, па нису ни приказани на дијаграму.

Припрема улаза. Претпоставља се да је потпис већ детектован и исечен са документа и да је познат *тврђени идентитет*, то јест особа за коју се тврди да је потпис њен. Улаз у систем чине, дакле, слика самог потписа и тврђени идентитет. Слика потписа пропушта се кроз обучену мрежу f_θ (ResNet18 са додатим линеарним слојем), чији се излаз L_2-нормализује, чиме настаје *векторска репрезентација* $\mathbf{e}$, тачка на јединичној хиперсфери која изглед потписа сажето представља помоћу $d = 512$ бројева.

Поређење са базом. Систем проверава тврђени идентитет тако што из *векторске базе идентитета* преузима *прототип* $\hat{\mathbf{p}}_i$ тог идентитета i и рачуна косинусну сличност $\cos(\mathbf{e}, \hat{\mathbf{p}}_i)$ између репрезентације улазног потписа и тог прототипа. Ако је сличност изнад унапред подешеног прага τ, потпис се прихвата као аутентичан; у супротном се одбацује. Праг τ на тај начин представља једини параметар којим се подешава строгост система.

Улога базе. Векторска база идентитета има двојаку улогу. Пре свега, она је *меморија система*: за сваког познатог потписника чува листу његових до сада прихваћених аутентичних вектора и из ње одржава нормализовани просек $\hat{\mathbf{p}}_i$, који служи као прототип тог идентитета. Када у систем стигне нова, потврђено аутентична слика, њен вектор се додаје у листу одговарајућег потписника, а прототип се поново израчунава, па систем остаје ажуран без поновног обучавања мреже, чак и када се дода потпуно нов идентитет. Поред тога, база је и *референца за поређење*: управо она, на захтев, враћа прототип тврђеног идентитета који улази у поређење описано у претходном кораку.

Из овог тока јасно се виде улоге појединих компоненти. Мрежа са линеарним слојем учествује и у обучавању и у закључивању и представља заједнички



Слика 4.4: Ток закључивања.

део обе фазе. Векторска база се, међутим, укључује тек у закључивању и захваљујући њој систем ради без поновног обучавања, па се нови идентитети могу додавати без икакве измене мреже. Узорковање и рударење, супротно томе, припадају само обучавању: узорковање одређује шта улази у мини-групу пре него што слике прођу кроз мрежу, а рударење шта од тога улази у функцију грешке после мреже.

Глава 5

Експериментална поставка

Да би се предложена методологија могла проверити, потребно ју је сместити у прецизно дефинисану експерименталну поставку. Ту поставку чине три целине: скуп података о потписима са својом поделом, имплементациони детаљи обучавања и протокол евалуације по којем се сви испитивани модели пореде под једнаким условима. Наредни одељци излажу сваку од њих редом.

5.1 Скуп података и подела

За потребе овог рада коришћен је приватни скуп података са руком писаним потписима, који није јавно доступан, прикупљен у контролисаним условима. Сваки узорак у скупу представљен је тројком: *слика потписа*, *ознака идентитета* (јединствени број који одговара потписнику) и *бинарни индикатор фалсификата* који разликује аутентичан потпис од вешто израђеног фалсификата. Све фалсификате у скупу ручно су израдили људи који су претходно проучили циљни потпис, тако да припадају најтежој варијанти описаној у одељку 3.1.

Скуп података најпре је подељен на два дисјунктна подскупа *на нивоу идентитета*: *скуп виђених* (енг. *seen*) идентитета и *скуп невиђених* (енг. *unseen*) идентитета. Скуп виђених садржи 924 идентитета и сви они учествују у некој фази обучавања или евалуације модела, док скуп невиђених садржи 495 идентитета који се током обучавања *никада* не појављују, па служе искључиво за процену колико добро модел генерализује на потпуно нове потписнике. Овакав начин поделе је значајан управо због тога што у реалним системима верификације модел редовно среће потписнике који нису постојали у тренинг

скупу, и његова употребна вредност умногоме зависи од перформанси на тако новим идентитетима. Што се фалсификата тиче, у скупу виђених је 105 идентитета са придруженим фалсификатима, а у скупу невиђених 54.

Виђени идентитети. Унутар скупа виђених идентитета спроведена је трострука подела на нивоу појединачних узорака. *Тренинг скуп* служи за директно ажурирање параметара мреже током обучавања: из њега узорковање формира мини-групе и на њима се рачуна функција грешке. *Валидациони скуп* служи за избор најбоље конфигурације модела: на њему се после сваке епохе процењује квалитет научених репрезентација и бира контролна тачка обучавања са најбољим показатељима, а на њему се и калибрише праг τ. *Тест скуп* држи се у потпуности по страни током обучавања и калибрације и користи се само за финалну, непристрасну евалуацију одабраног модела. Будући да сва три скупа садрже потписе иста 924 идентитета, али различите узорке тих идентитета, резултати на овим скуповима показују колико добро модел ради на познатим потписницима.

Невиђени идентитети. У скупу невиђених идентитета поставка је битно другачија. За сваког од 495 потписника, његови узорци подељени су у две улоге: *галерију* (енг. *gallery*) и *упит* (енг. *query*). Галерија се састоји од малог броја аутентичних потписа сваког идентитета и њена улога је да попуни векторску базу: на основу галеријских вектора рачуна се прототип $\hat{\mathbf{p}}_i$ сваког невиђеног идентитета, на исти начин као што је описано у одељку 4.2. Упитни скуп садржи преостале потписе истих идентитета, и аутентичне и фалсификате, и они представљају улазе које систем у евалуацији треба да верификује поређењем са одговарајућим галеријским прототиповима.

Овакво раздвајање на галерију и упит постоји управо због чињенице да за невиђене потписнике не постоји тренинг скуп из којег би се прототипи могли преузети: систем мора *у тренутку увођења* новог потписника да сам формира прототип од његових расположивих аутентичних узорака, што галерија и симулира. Резултати на овом скупу стога директно мере оно што се од модела очекује и у радном систему: способност да на основу неколико изабраних аутентичних узорака новог идентитета поуздано одлучује о аутентичности нових упита.

5.2 Имплементациони детаљи

Окружење и улазни формат. Реализација је заснована на библиотеци PyTorch [21] уз допуну библиотеке PyTorch Metric Learning [20]. Окосница модела је ResNet18 са ImageNet тежинама као почетним; завршни линеарни слој пројектује на 512-димензионалан простор векторских репрезентација. Све слике се пре улаза у мрежу своде на величину 224×224 пиксела уз очување односа страница (преостали простор се испуњава црном бојом), претварају у троканалну сиву слику и нормализују средњим вредностима и стандардним девијацијама ImageNet скупа.

Аугментације. У фази обучавања на улазне слике примењује се скуп стохастичких аугментација који служи као регуларизатор:

- насумична ротација у опсегу $\pm 10°$;
- насумична промена боје (енг. *color jitter*) са умереним променама светлине и контраста ($\pm 0{,}4$), мањом променом засићења ($\pm 0{,}2$) и веома малом променом нијансе ($\pm 0{,}02$);
- стохастичко Гаусово замућење језгром величине 3×3 са $\sigma \in [0{,}1,\ 0{,}5]$, примењено са вероватноћом 0,3.

У фази валидације и тестирања не примењује се ниједна од ових аугментација: слике пролазе само кроз исту основну предобраду (промена величине, претварање у сиве тонове и нормализација), без икаквих случајних измена.

Оптимизатор и распоред корака учења. За оптимизацију параметара мреже коришћен је алгоритам Adam [14]. Корак учења контролисан је *цикличним распоредом* [25] у режиму *triangular2*, који периодично помера корак између доње и горње границе $[\eta_{\min}, \eta_{\max}] = [5 \times 10^{-5},\ 5 \times 10^{-4}]$, при чему је пола циклуса дугачко 20 епоха (изражено у корацима, $20 \cdot \lceil N_{\text{tr}}/B \rceil$, где је N_{tr} величина тренинг скупа а B величина мини-групе). Граничне вредности корака учења идентификоване су унапред помоћу *теста распона корака учења* (енг. *learning rate range test*, LRRT) [25], у којем се корак експоненцијално повећава од 10^{-8} до 10^{-2}, а функција грешке се прати у зависности од корака. Као доња граница изабрана је вредност при којој грешка стабилно почиње да

опада, а као горња изабрана је вредност близу врха непосредно пре дивергенције. Поред тога, у хибридној поставци примењује се и *одсецање градијента* (енг. *gradient clipping*) на L_2 норму од 5, у циљу стабилизације обучавања. Сваки модел обучаван је кроз 1000 епоха.

Параметри узорковања. За све три стратегије (триплет, вишеструке сличности, хибридна) користи се узорковање свесно фалсификата описано у одељку 4.4. Његови параметри по стратегији дати су у Табели 5.1. Хибридна стратегија има посебност. Пошто две њене компоненте раде над различитим типовима парова, она користи *два паралелна поступка узорковања*: стандардно узорковање са фиксним бројем узорака по класи (4 узорка) за компоненту вишеструке сличности (која ради само над аутентичним потписима), и узорковање свесно фалсификата за триплет компоненту, са параметрима наведеним у последњој колони табеле. У табели је такође наведена и резултујућа величина мини-групе B као производ броја потписника по мини-групи и укупног броја узорака по потписнику. Вредности у Табели 5.1 нису унапред задате, већ су резултат подешавања: за сваки модел обучавање је више пута понављано са различитим комбинацијама параметара, а приказане су оне које су дале најбоље резултате на валидационом скупу.

Табела 5.1: Параметри узорковања по стратегији обучавања.

Параметар	Вишеструка сл.	Триплет	Хибридна (триплет)
Потписника по мини-групи	12	16	6
Аутентичних по потписнику (m_g)	8	8	2
Фалсификата по потписнику (m_f)	4	2	3
Вероватноћа потписника са фалс. (p_f)	0,5	0,5	1,0
Резултујућа величина мини-групе (B)	144	160	30

Параметри рударења. Рударење за функцију вишеструке сличности користи праг тежине $\varepsilon = 0{,}1$ и горњу границу од 10 негатива по сваком од два подскупа (међуидентитетски и унутаридентитетски). Рударење за триплет функцију грешке користи маргину $m = 0{,}15$, режим избора полутешких тројки, и горње границе од 3 негатива по подскупу, уз ограничење од 100 тројки по референтном узорку. Параметри саме функције вишеструке сличности (једначина 4.6 у одељку 4.3) су $\alpha = 2{,}0$, $\beta = 30$ за самосталну стратегију и $\beta = 15$ у хибридној поставци, и померај $\lambda_{\text{MS}} = 0{,}5$.

Хибридна функција грешке. У хибридној стратегији две компоненте функције грешке нормализују се пре сабирања: појединачне вредности грешке деле се текућом медијаном (евалуираном у почетном прозору од 50 корака) пре него што се пондеришу, како ниједна компонента не би доминирала рачуном градијента услед разлике у скали. Тежина $\lambda = 0{,}5$ примењује се на триплет компоненту, а компоненти вишеструке сличности додељује се коефицијент 1.

Избор најбоље конфигурације. После сваке евалуационе епохе модел се процењује на валидационом скупу, а као мера квалитета користи се *изједначена стопа грешака* (енг. *equal error rate*, EER) на верификационом задатку, то јест стопа грешке на оном прагу сличности на коме је стопа лажног прихватања једнака стопи лажног одбацивања. Ова метрика има предност што не зависи од избора конкретног прага и стога даје непристрасну меру квалитета научених репрезентација. Контролна тачка са најнижом вредношћу EER-а на валидацији преузима се као финални модел и користи у завршној евалуацији на тест и невиђеним скуповима. Дефиниције ове и осталих евалуационих метрика дате су у наредном одељку.

Доступност кода. Комплетан изворни код имплементације — припрема података, обучавање модела, евалуација и генерисање приказаних резултата — доступан је од аутора на захтев.

Употреба алата вештачке интелигенције. Током израде рада коришћени су алати засновани на великим језичким моделима, и то за језичко обликовање текста и за генерисање илустрација. Спроведени експерименти, добијени резултати и изведени закључци су ауторски и за њих аутор сноси пуну одговорност.

5.3 Евалуација

Испитивани модели. У евалуацији се пореде четири модела. Три су конфигурације *свесне фалсификата* које се међусобно разликују само по функцији грешке: *тријлет*, *вишеструке сличности* и *хибридна*. Четврти је *контролни* модел, обучен функцијом вишеструке сличности искључиво на ау-

тентичним потписима, без иједног податка о фалсификатима. Он служи као контрола која изолује допринос свести о фалсификатима: пошто са моделима свесним фалсификата дели све остале елементе поставке (исту архитектуру и димензију простора векторских репрезентација, исти тренинг скуп и исти протокол обучавања, евалуације и калибрације прага), свака разлика у његовим резултатима може се приписати једино одсуству података о фалсификатима у обучавању.

Поступак верификације. Евалуација се на сва три скупа (валидационом, тест и невиђеном) спроводи по истом поступку. За сваки потпис $\mathbf{x}_q$ који се верификује, са тврђеним идентитетом i, систем израчунава његову векторску репрезентацију $\mathbf{e}_q = f_\theta(\mathbf{x}_q)$, L_2-нормализује је, и рачуна косинусну сличност са нормализованим прототипом $\hat{\mathbf{p}}_i$ тврђеног идентитета. Одлука се доноси на основу подешеног прага τ: ако је $\cos(\mathbf{e}_q, \hat{\mathbf{p}}_i) \geq \tau$, потпис се прихвата као аутентичан, иначе се одбацује. Прототипи се за виђене скупове (валидациони и тест) рачунају као нормализовани просек векторских репрезентација аутентичних потписа из тренинг скупа истог идентитета, а за невиђени скуп из галерије датог идентитета (одељак 5.1).

Сценарији евалуације. У сваком скупу посматрају се два дисјунктна евалуациона сценарија:

- *Аутентичан против насумичног фалсификата* (енг. *genuine vs. random forgery*) — потпис који се верификује је аутентичан, али припада *другом* идентитету, представљен као кандидат за тврђени идентитет. Овај сценарио мери способност модела да разликује потписнике међусобно.
- *Аутентичан против вешто израђеног фалсификата* (енг. *genuine vs. skilled forgery*) — потпис који се верификује је фалсификат тврђеног идентитета, конструисан да наликује оригиналу. Овај сценарио мери способност модела да разликује аутентичан потпис од његове намерне имитације и представља суштински проблем рада.

У оба сценарија скуп позитивних узорака чине аутентични потписи правог идентитета, а скуп негативних се разликује: насумични аутентични потписи других идентитета у првом, и вешто израђени фалсификати у другом сценарију.

Метрике. На основу скорова сличности у сваком сценарију рачунају се следеће метрике:

- *Изједначена стопа грешака* (енг. *equal error rate*, EER) — вредност на радном прагу на коме је стопа лажног прихватања (енг. *false accept rate*, FAR) једнака стопи лажног одбацивања (енг. *false reject rate*, FRR). EER је сажета мера квалитета која не зависи од конкретно изабраног радног прага.

- *Површина испод ROC криве* (енг. *area under curve*, AUC) — мера укупне способности модела да издвоји позитивне од негативних узорака, узимајући у обзир све могуће прагове [6]. Вредност 1 одговара савршеном раздвајању, а 0,5 насумичном погађању.

- *Стопа прихватања аутентичних при фиксној стопи лажног прихватања* (енг. *genuine accept rate at fixed false accept rate*, GAR@FAR) — говори колико аутентичних потписа систем прихвати када се праг подеси тако да пропусти унапред задат проценат негативних узорака. У овом раду се извештава GAR при FAR = 10%: праг се подеси тако да 10% негатива буде погрешно прихваћено, па се онда мери колики проценат аутентичних потписа систем при том прагу прихвата.

GAR@FAR = 10% изабран је као главна метрика за поређење зато што директно одговара на практично најрелевантније питање: ако се захтева унапред подешен ниво заштите од фалсификата, која конфигурација модела дозвољава да прође највише легитимних потписника? Поред њега, EER пружа алтернативну меру, независну од прага, а AUC сажима квалитет на читавом опсегу радних тачака.

Калибрација прага. Праг τ је граница на косинусној сличности између репрезентације потписа и прототипа тврђеног идентитета. Ако је сличност већа или једнака τ, потпис се прихвата као аутентичан. У супротном се одбацује. Праг је уједно и једина величина којом се подешава строгост система. Виши праг одбацује више фалсификата, али и погрешно одбија више аутентичних потписа, док нижи праг чини обратно. Његова вредност стога није иста за све моделе. Свака научена функција угњеждења производи различиту расподелу сличности и захтева сопствену калибрацију. За сваку конфигурацију праг

се одређује на *валидационом скупу* тако да се у сценарију аутентичан против вешто израђеног фалсификата постигне FAR = 10%. Тако одабран праг се потом примењује *непромењен* на тест и невиђени скуп. Овакав протокол симулира реалистичну поставку у којој се систем једном калибрише, а потом се у пракси користи без накнадног прилагођавања новим подацима.

Поређење модела. Поређење ова четири модела на тест и невиђеном скупу, при једнаком радном FAR = 10%, директно одговара на основно питање рада: колики је допринос експлицитног моделовања фалсификата током обучавања, и која од испитиваних функција грешке у тој поставци даје најбоље резултате.

Глава 6

Резултати и дискусија

У овом поглављу се излажу резултати експеримената спроведених по поставци из претходног поглавља. Излагање има три дела. Прво се сви испитивани модели пореде по основним метрикама верификације на сва три скупа података. Затим се анализирају расподеле сличности код модела са свешћу о фалсификатима и код модела без ње, како би се видело одакле разлике међу њима потичу. На крају, у дискусији се сумирају главни налази и њихове импликације за офлајн верификацију потписа.

6.1 Верификациони резултати по скуповима

У овом одељку упоређују се сви испитивани модели по три основне верификационе метрике, изведене из величина дефинисаних у одељку 5.3: *прихватање аутентичних* (енг. *genuine acceptance*, Gen. Acc.), *одбацивање фалсификата* (енг. *forgery rejection*, Forg. Rej.) и *одбацивање насумичних узорака* (енг. *random rejection*, Rand. Rej.). Резултати на сва три скупа приказани су у Табели 6.1. За сваки модел, праг τ калибрисан је на валидационом скупу у режиму FAR = 10% у сценарију аутентичан против вешто израђеног фалсификата, и потом примењен непромењен на тест и невиђени скуп.

Валидациони скуп. На валидационом скупу сва три модела свесна фалсификата постижу високо прихватање аутентичних, у опсегу 97,0–98,1%, при подешеном FAR = 10%. Међу њима, модел вишеструке сличности остварује највишу вредност (98,1%), а следе га хибридни модел (97,6%) и триплет (97,0%). Контролни модел, обучаван без података о фалсификатима, на истом

Табела 6.1: Верификациони резултати на три скупа при FAR = 10%. За сваки модел, праг τ је калибрисан на валидационом скупу и примењен непромењен на тест и невиђени скуп. Приказане су три метрике у процентима: прихватање аутентичних потписа, одбацивање фалсификата и одбацивање насумичних узорака.

Партиција	Модел	τ	Gen. Acc.	Forg. Rej.	Rand. Rej.
Валидациона	Вишеструка сл.	0,57	**98**,**1**	90,2	99,2
	Хибридна	0,57	97,6	89,5	99,7
	Триплет	0,71	97,0	90,2	99,9
	Контролни	0,88	83,9	90,2	**100**,**0**
Тест	Вишеструка сл.	0,57	**97**,**8**	91,9	99,1
	Хибридна	0,57	97,5	84,0	99,8
	Триплет	0,71	97,6	**98**,**6**	99,8
	Контролни	0,88	83,9	**99**,**8**	**100**,**0**
Невиђена	Вишеструка сл.	0,57	**92**,**5**	90,7	98,7
	Хибридна	0,57	90,1	93,7	99,4
	Триплет	0,71	85,7	94,4	99,5
	Контролни	0,88	42,4	**99**,**3**	**100**,**0**

нивоу заштите од фалсификата прихвата свега 83,9% аутентичних потписа, преко 14 процената мање од модела вишеструке сличности. Стопа одбацивања насумичних узорака је код свих модела врло висока ($\geq$ 99,2%), што показује да међуидентитетско раздвајање није тешко: права тешкоћа је унутаридентитетско, фалсификатно раздвајање.

Тест скуп. На тест скупу, који садржи нове узорке истих идентитета као тренинг, модели свесни фалсификата задржавају висок ниво прихватања аутентичних (97,5–97,8%), уз приметне разлике у одбацивању фалсификата. Триплет функција на тест скупу постиже највишу стопу одбацивања фалсификата међу моделима свесним фалсификата (98,6%), значајно изнад нивоа од око 90% који одговара FAR-у од 10% калибрисаном на валидацији, што сугерише да је њен одабрани праг конзервативнији и да модел ту даје приоритет одбацивању фалсификата. Хибридни модел, насупрот томе, на тесту има мању стопу одбацивања (84,0% у поређењу са 89,5% на валидацији), што указује да његов праг $\tau = 0{,}57$ за тест скуп делује донекле *аг̄ресивно* у корист прихватања. Контролни модел на тесту наставља да даје 83,9% прихватања, готово исто као на валидацији, али и даље знатно ниже од модела свесних

фалсификата.

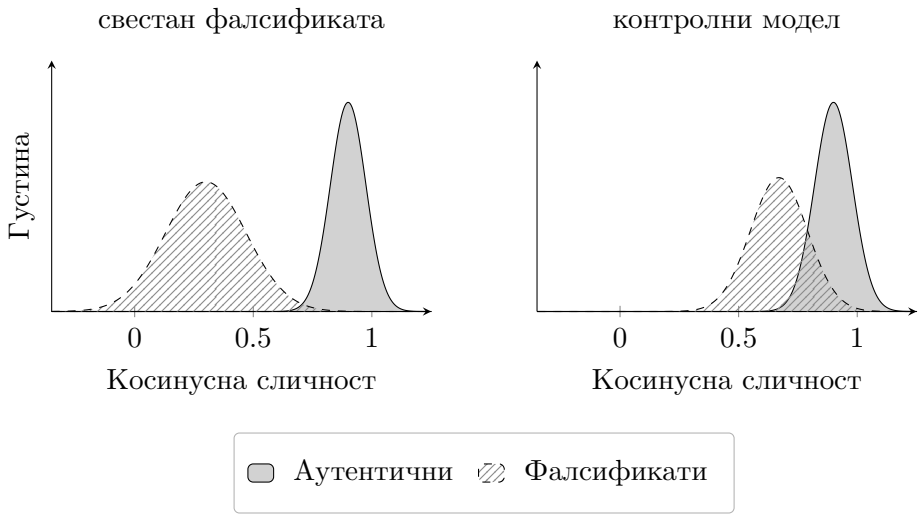
Невиђени скуп. Највећа разлика између модела открива се на скупу потписника који нису били доступни током обучавања. Код свих модела бележи се пад прихватања аутентичних, што је природно, јер се прототипи сада рачунају из мале галерије уместо из обимног тренинг скупа. Ипак, величина пада драматично зависи од присуства фалсификата у обучавању. Модел вишеструке сличности задржава 92,5% прихватања, хибрид 90,1%, а триплет 85,7%. Контролни модел доживљава потпуни колапс: његово прихватање аутентичних пада на свега 42,4%, што значи да више од половине аутентичних потписа нових потписника бива одбачено. Разлика од 50 процената између модела вишеструке сличности и контролног модела је најупечатљивији појединачни налаз овог рада.

Стопа одбацивања насумичних узорака остаје веома висока за све моделе на свим скуповима ($\geq$ 98,7%), што потврђује раније запажање: задатак раздвајања различитих потписника, и при класичном и при савременом моделовању, није суштинска тешкоћа верификације потписа; она лежи искључиво у прецизном раздвајању аутентичних потписа од њихових циљаних имитација.

6.2 Расподеле сличности

Да би се узрок драматичне разлике у перформансама између модела свесних фалсификата и контролног модела учинио експлицитним, у овом одељку се пореде *расподеле косинусних сличности* које две парадигме обучавања производе на истим узорцима. Конкретно, упоређују се два модела који деле све остале елементе поставке (архитектуру, димензију простора векторских репрезентација, тренинг скуп и функцију грешке вишеструке сличности), а разликују се само у томе да ли је коришћена стратегија узорковања и избора тешких примера свесна фалсификата. Слика 6.1 приказује њихове расподеле сличности на валидационом скупу, посебно за аутентичне парове и посебно за фалсификатне.

Лева страна слике приказује модел обучаван стратегијом свесном фалсификата. Његова расподела аутентичних парова концентрисана је око високих вредности косинусне сличности (близу 1,0), док је расподела фалсификатних парова померена ка значајно нижим вредностима (центрирана око 0,3).



Слика 6.1: Расподеле косинусних сличности на валидационом скупу за два модела са функцијом вишеструке сличности.

Између њих постоји широка *гранична* област у којој се готово не налазе ни аутентични ни фалсификатни узорци, па умерен праг (нпр. $\tau = 0{,}57$) даје чисто раздвајање, при коме већина аутентичних бива прихваћена, а већина фалсификата одбачена.

Десна страна слике приказује контролни модел обучаван без података о фалсификатима. Његова расподела аутентичних парова је слично концентрисана око високих вредности, али је расподела фалсификата ту померена ка знатно вишим вредностима и сада се у великој мери *преклапа* са расподелом аутентичних. Узрок је директан: модел током обучавања никада није учио да за исти идентитет тежи разликовању аутентичних и фалсификатних узорака, па их у простору векторских репрезентација пресликава у блиске тачке. Резултат је да се ниједан праг не може поставити тако да истовремено постигне високо прихватање аутентичних и високо одбацивање фалсификата: свако побољшање једног од та два нужно иде на штету другог.

Ова визуелна разлика објашњава све нумеричке резултате из претходног одељка. Контролни модел, да би уопште достигао FAR = 10%, мора захтевати веома висок праг ($\tau = 0{,}88$), а тада знатан реп расподеле аутентичних остаје *испод* те границе и бива одбачен. Тиме се објашњава пад прихватања аутентичних на 83,9% на виђеним скуповима и његов колапс на 42,4% на невиђеном скупу, где је расподела аутентичних додатно проширена услед мање обимне галерије по идентитету.

6.3 Дискусија

Резултати приказани у одељцима 6.1 и 6.2 омогућавају више суштинских закључака о методологији и о домену у целини.

Прво и најважније, експерименти потврђују да је *експлицитно моделовање фалсификата током обучавања* суштинско за робусну офлајн верификацију потписа. Поређење конфигурација свесних фалсификата и контролне варијанте показује јаз од 14 процената прихватања аутентичних на виђеним скуповима и драматичних 50 процената на невиђеним потписницима, при истом нивоу заштите од фалсификата. Расподеле сличности на слици 6.1 визуелно показују и зашто: без сигнала о фалсификатима, модел нема разлога да их раздваја од аутентичних потписа истог идентитета у простору векторских репрезентација.

Друго, међу испитиваним функцијама грешке, *вишеструка сличност* се уопштено показала као најуспешнија, нарочито у генерализацији на нове потписнике. Њено фокусирање на глобалну структуру простора, то јест организовање свих идентитета истовремено, очигледно се добро уклапа са захтевима верификације потписа, где је кључно да модел научи општу геометрију сличности применљиву и на потписнике које до тада није видео. Триплет функција, са својом локалном природом, пружа боље одбацивање фалсификата на тест скупу (праг $\tau = 0{,}71$ је конзервативнији), али по цену генерализације: на невиђеном скупу њено прихватање аутентичних пада на 85,7%, што је испод осталих модела свесних фалсификата.

Треће, *хибридна* функција грешке, која експлицитно комбинује две доминантне варијанте, у овим експериментима *не надмашује* вишеструку сличност саму по себи. Њено прихватање аутентичних на невиђеном скупу (90,1%) лежи између триплета (85,7%) и вишеструке сличности (92,5%). Једно тумачење овог налаза је да рударење свесно фалсификата, када се користи са функцијом вишеструке сличности, већ обезбеђује довољно јак сигнал о фалсификатима: фалсификати истог идентитета ионако у њу улазе као најтежи негативни парови и добијају највећи коефицијент, па додатна триплет компонента не доноси нову информацију, већ само још једном пондерише парове које прва компонента већ узима у обзир. Уз то, две компоненте делују над различитим структурама (паровима, односно тројкама), па њихово сабирање уводи и додатни хиперпараметар λ чија лоша вредност може да поништи очекивану

добит. Друга могућност, коју овај рад не може да искључи, јесте да хибридна формулација захтева пажљивију калибрацију тежине λ и параметара рударења од оне коришћене у експериментима.

Глава 7

Закључак

У овом раду разматрана је офлајн верификација руком писаних потписа методама дубоког учења сличности, са фокусом на вешто израђене фалсификате, који представљају најтежу варијанту овог задатка. Предложена методологија почива на обучавању *свесном фалсификата*: узорковање мини-група и рударење парова осмишљени су тако да у функцију грешке редовно улазе парови оригинал–фалсификат истог идентитета, а не само насумично изабрани негативни примери. Три функције грешке (триплет, вишеструке сличности и њихова хибридна комбинација) испитане су под истим условима и упоређене са контролним моделом обученим искључиво на аутентичним потписима, при чему подела скупа података раздваја виђене од невиђених идентитета.

Главни закључак рада јесте да је *експлицитно моделовање фалсификата током обучавања* пресудно за квалитет система. При истом нивоу заштите од фалсификата, модели свесни фалсификата прихватају око 14 процената више аутентичних потписа на виђеним идентитетима и око 50 процената више на невиђенима у односу на контролни модел.

Међу испитаним функцијама грешке најбоље се показала функција *вишеструке сличности*, нарочито у генерализацији на нове потписнике, јер организује све идентитете истовремено и тако учи општу геометрију сличности. Триплет функција даје боље одбацивање фалсификата на тест скупу, али слабије генерализује, док хибридна комбинација у овим експериментима не надмашује вишеструку сличност саму по себи, што сугерише да рударење свесно фалсификата уз вишеструку сличност већ обезбеђује довољно јак сигнал о фалсификатима.

Резултати такође показују да *одбацивање насумичних узорака* код свих

испитиваних модела остаје готово савршено ($\geq$ 98,7% на свим скуповима). Ово показује да *међуидентитетско* раздвајање није тешка страна задатка, јер сви испитивани модели без тешкоћа разликују потписе различитих особа. Права тешкоћа је *унутаридентитетско раздвајање*: треба раздвојити аутентичан потпис од његовог намерног фалсификата. Због тога квалитет оваквих система треба мерити пре свега поређењем аутентичних потписа са фалсификатима истог идентитета, а не са потписима насумично изабраних особа.

Главно ограничење које резултати показују и на које будући радови треба да се фокусирају јесте *генерализација на потпуно нове потписнике*. Иако модел вишеструке сличности у овом сегменту значајно надмашује остале моделе, његова стопа прихватања аутентичних на невиђенима (92,5%) је и даље 5,6 процената нижа од оне на виђенима (98,1%). Овај јаз представља плафон тренутног приступа и природан правац за даља истраживања: дубље основне мреже, богатије стратегије аугментације прилагођене потписима, као и обучавање у коме се већ током учења симулирају услови рада система, то јест ситуација у којој је по идентитету доступно свега неколико потписа.

Практично гледано, рад показује да се употребљив систем за верификацију потписа може добити и са стандардном основном мрежом и умереним ресурсима, под условом да се фалсификати угради у само обучавање, и да проширење галерије новог потписника не захтева поновно обучавање, већ само додавање његових вектора у базу идентитета.

Литература

[1] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a "Siamese" time delay neural network. In *Advances in Neural Information Processing Systems*, volume 6, pages 737–744, 1993.

[2] Sumit Chopra, Raia Hadsell, and Yann LeCun. Learning a similarity metric discriminatively, with application to face verification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 539–546. IEEE, 2005.

[3] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 248–255, 2009.

[4] Sounak Dey, Anjan Dutta, J. Ignacio Toledo, Suman K Ghosh, Josep Lladós, and Umapada Pal. SigNet: Convolutional siamese network for writer independent offline signature verification. *arXiv preprint arXiv:1707.02131*, 2017.

[5] Moises Diaz, Miguel A Ferrer, Donato Impedovo, Muhammad Imran Malik, Giuseppe Pirlo, and Réjean Plamondon. A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 51(6):1–39, 2019.

[6] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[7] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016.

[8] Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 1735–1742, 2006.

[9] Luiz G Hafemann, Robert Sabourin, and Luiz S Oliveira. Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70:163–176, 2017.

[10] Luiz G Hafemann, Robert Sabourin, and Luiz S Oliveira. Offline handwritten signature verification—literature review. In *Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–8. IEEE, 2017.

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[12] Donato Impedovo and Giuseppe Pirlo. Automatic signature verification: The state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(5):609–635, 2008.

[13] Anil K Jain, Arun Ross, and Salil Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105, 2012.

[16] Songxuan Lai, Lianwen Jin, Luojun Lin, Yecheng Zhu, and Huiyun Mao. SynSig2Vec: Learning representations from synthetic dynamic signatures for real-world verification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 735–742, 2020.

[17] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[18] Paul Maergner, Vinaychandran Pondenkandath, Michele Alberti, Marcus Liwicki, Kaspar Riesen, Rolf Ingold, and Andreas Fischer. Offline signature verification by combining graph edit distance and triplet networks. In *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition (S+SSPR)*, pages 470–480. Springer, 2018.

[19] Kevin Musgrave, Serge Belongie, and Ser-Nam Lim. A metric learning reality check. In *European Conference on Computer Vision (ECCV)*, pages 681–699. Springer, 2020.

[20] Kevin Musgrave, Serge Belongie, and Ser-Nam Lim. PyTorch metric learning. *arXiv preprint arXiv:2008.09164*, 2020.

[21] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32, pages 8024–8035, 2019.

[22] Jivesh Poddar, Vinanti Parikh, and Santosh Kumar Bharti. Offline signature recognition and forgery detection using deep learning. *Procedia Computer Science*, 170:610–617, 2020.

[23] Hannes Rantzsch, Haojin Yang, and Christoph Meinel. Signature embedding: Writer independent offline signature verification with deep metric learning. In *International Symposium on Visual Computing (ISVC)*, pages 616–625. Springer, 2016.

[24] Florian Schroff, Dmitry Kalenichenko, and James Philbin. FaceNet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 815–823, 2015.

[25] Leslie N Smith. Cyclical learning rates for training neural networks. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 464–472, 2017.

[26] Kihyuk Sohn. Improved deep metric learning with multi-class N-pair loss objective. In *Advances in Neural Information Processing Systems*, volume 29, pages 1849–1857, 2016.

[27] Hyun Oh Song, Yu Xiang, Stefanie Jegelka, and Silvio Savarese. Deep metric learning via lifted structured feature embedding. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4004–4012, 2016.

[28] Victor L. F. Souza, Adriano L. I. Oliveira, and Robert Sabourin. A writer-independent approach for offline signature verification using deep convolutional neural networks features. In *7th Brazilian Conference on Intelligent Systems (BRACIS)*, pages 212–217. IEEE, 2018.

[29] V. V. Starovoitov and U. Yu. Akhundjanov. A writer-dependent approach to offline signature verification based on one-class support vector machine. *Pattern Recognition and Image Analysis*, 34:340–351, 2024.

[30] Xun Wang, Xintong Han, Weilin Huang, Dengke Dong, and Matthew R. Scott. Multi-similarity loss with general pair weighting for deep metric learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5017–5025, 2019.

[31] Chao-Yuan Wu, R Manmatha, Alexander J Smola, and Philipp Krähenbühl. Sampling matters in deep embedding learning. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2859–2867, 2017.

[32] Mustafa Berkay Yilmaz and Kagan Ozturk. Hybrid user-independent and user-dependent offline signature verification with a two-channel CNN. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 526–534, 2018.

[33] Mingjian Zhang, Lidong Zheng, Jiaen Chen, Yichi Zhang, and Yuchen Zheng. SigLDiff: A signature based latent diffusion model for forged-free offline signature verification. In *International Conference on Document Analysis and Recognition (ICDAR)*, pages 416–432. Springer, 2025.

[34] Elias N Zois, Evangelos Zervas, Dimitrios Tsourounis, and George Economou. Sequential motif profiles and topological plots for offline signature verification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13245–13255, 2020.

[35] Предраг Јаничић and Младен Николић. *Вештачка интелигенција*. Математички факултет, Универзитет у Београду, Београд, 5 edition, 2025. Пето, електронско издање.

Биографија аутора

Василије Тодоровић – рођен је 4. априла 2000. године у Београду. У Младеновцу је завршио општи смер гимназије као носилац Вукове дипломе.

Математички факултет Универзитета у Београду, смер Информатика, уписао је 2019. године. Основне студије завршио је 2023. године са просечном оценом 9,28, након чега је на истом факултету уписао мастер студије. Током студија био је сарадник у настави на Катедри за рачунарство и информатику, где је држао вежбе из предмета Истраживање података 1 и 2, Релационе базе података, Пројектовање база података и Програмирање 2.

Током студија највише га је заинтересовала област машинског учења. У том правцу обавио је стручну праксу на којој је радио са визуелно-језичким моделима, а паралелно са мастер студијама ради као инжењер машинског учења.