

ИМПЛЕМЕНТАЦИЈА СИЈАМСКЕ НЕУРОНСКЕ МРЕЖЕ ЗА ВЕРИФИКАЦИЈУ ЛИЦА ЗАСНОВАНУ НА „ONE-SHOT“ ПРЕДИКЦИЈИ**IMPLEMENTATION OF A SIAMESE NEURAL NETWORK FOR FACE VERIFICATION BASED ON ONE-SHOT PREDICTION**

Петар Сурла, Факултет техничких наука, Нови Сад

Област – МЕХАТРОНИКА

Кратак садржај – У овом раду представљена је имплементација сијамске неуронске мреже за верификацију лица. Описан је процес креирања скупа података, процесирања слика, као и дизајн и обука мреже. Циљ рада је примена „one-shot“ предикције како би се систем обучио на малом броју података. Приказани су експериментални резултати и поређење са другим моделима, као и коначни закључци.

Кључне речи: верификација лица, сијамске неуронске мреже, конволуционе мреже, препознавање лица

Abstract – This paper presents the implementation of a Siamese neural network for face verification. It describes the process of dataset creation, image preprocessing, as well as the design and training of the network. The objective of this work is to apply one-shot prediction to train the system using a small amount of data. Experimental results and comparisons with other models are provided, along with final conclusions.

Keywords: face verification, Siamese neural networks, convolutional networks, face recognition

1. УВОД

Верификација лица је једна од најзначајнијих примена биометријске идентификације и представља кључни елемент у бројним безбедносним системима, од паметних телефона до система за надзор и контролу приступа. Уз све већу потребу за прецизним и брзим методама верификације идентитета, развој система за препознавање лица представља област која се убрзано развија. Разматра се транзиција од традиционалних метода, које се често показују као непоуздане у условима варијације података, до модерних приступа базираних на дубоком учењу. Посебан акценат је стављен на примену „Siamese Neural Network“ модела, који користи концепт „one-shot“ учења за ефикасну верификацију лица уз минималну количину података. Изазови попут обезбеђивања тачности и робусности модела у реалним условима су такође истакнути.

НАПОМЕНА:

Овај рад проистекао је из мастер рада чији је ментор био др Жељко Кановић, редовни професор.

Циљ рада је да се кроз детаљан опис процеса, укључујући креирање скупа података, обуку и тестирање мреже, развије ефикасан и прецизан систем за верификацију лица. Осим тога, рад обухвата поређење са другим моделима како би се проценила његова ефикасност и предложила унапређења за будућа истраживања.

2. АРХИТЕКТУРА МРЕЖЕ

Ово поглавље представља структуру и функционисање неуронских мрежа, са посебним освртом на мреже за препознавање и верификацију лица. Такође се детаљно разматра архитектура мреже која је примењена за решавање проблема верификације лица.

2.1. Препознавање и верификација лица

Препознавање лица је област машинске визије која има за циљ идентификацију или верификацију особе на основу дигиталне слике или видео записа. Ова технологија се користи у безбедносним системима, паметним уређајима и социјалним мрежама, ослањајући се на биометријске обрасце за препознавање јединствених карактеристика лица. Постоје два типа система препознавања лица (Слика 1):

- Идентификација лица: Систем покушава да идентификује појединца из већег скупа, односно утврђује идентитет особе.
- Верификација лица: Систем утврђује да ли одређена слика лица одговара слици у унапред дефинисаној бази података. Овај процес често укључује бинарну одлуку.



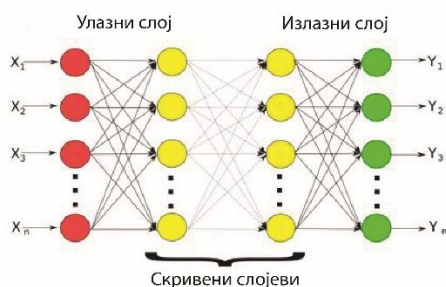
Слика 1. Приказ разлике између идентификације лица и верификације лица

2.2. Вештачке неуронске мреже (ВНМ)

Вештачке неуронске мреже су инспирисане структуром и функцијом биолошких неурона у људском мозгу [1]. Састоје се од улаза, тежина, сумирања, активационе функције и излаза. Основни елементи укључују:

- Улаз: Прима различите облике података (нпр. пикселе слике, нумеричке вредности)
- Тежина: Сваком улазом сигналу придружен је тежински коефицијент који одређује његов значај у процесу обраде.
- Сумирање: Улазни подаци се множе одговарајућим тежинама и сабирају.
- Активационе функције: Функција која одлучује о активацији неурона.
- Излаз: Резултат обраде који се прослеђује даље у мрежи или представља финални излаз

Више повезаних неурона чини слој неуронске мреже, док више повезаних слојева чини неуронску мрежу (Слика 2).



Слика 2. Приказ слојева неуронске мреже

Тренирање неуронске мреже укључује иницијализацију тежина, пропагацију унапред („*forward propagation*“), израчунавање грешке (функција губитка), пропагацију уназад („*backpropagation*“) и ажурирање тежина како би се минимизовала грешка.

2.3 Конволуционе неуронске мреже (КНМ)

Конволуционе неуронске мреже су специјализоване за обраду података са мрежастом структуром, као што су слике. Кључни елементи КНМ су:

- Конволуциони слојеви: Користе филтере (кERNEL) који се померају преко слике како би издвојили карактеристике као што су ивице и текстуре (Слика 3).
- Активационе функције: Уводе нелинеарност у мрежу; најчешће се користи „*ReLU*“ функција.
- „*Pooling*“ слојеви: Смањују просторне димензије података (најчешће „*Max Pooling*“), задржавајући најважније информације.
- Потпуно повезани слојеви: На крају мреже, служе за доношење коначне одлуке или класификације.

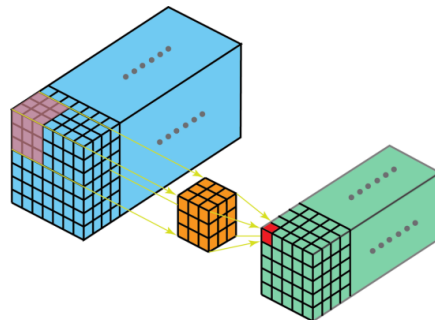
КНМ су веома ефикасне у задацима као што су класификација слика, детекција објеката и препознавање лица због своје способности да аутоматски уче и издвоје релевантне карактеристике из визуелних података. Математички, конволуција се дефинише као:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n) \quad (11)$$

Где су:

- S – решење операције конволуције над матрицом
- I – слика, односно улазни 2Д податак

- K – филтер, односно KERNEL који је 2Д у овом случају
- i, j – индекси елемената за који се рачуна конволуција
- m, n – димензије филтера



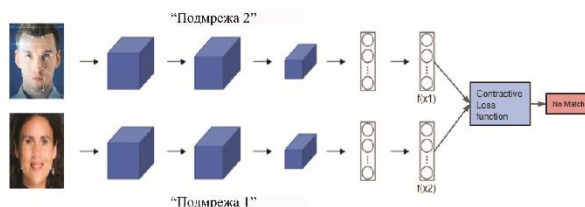
Слика 3. Визуелни приказ конволуције и KERNEL филтера.

3. АЛГОРИТАМ РЕШЕЊА

У овом поглављу представљена је архитектура и начин функционисања алгорита и модела неуронске мреже коришћене за решавање проблема верификације лица. Такође је поједностављено објашњено решење самог проблема.

3.1. Сијамске неуронске мреже (СНМ)

Сијамске неуронске мреже представљају специјални тип неуронских мрежа дизајнираних првенствено за рангирање сличности између два улаза (Слика 4). Мрежа се састоји од две идентичне „подмреже“ које обрађују различите улазе, али деле исте тежине и параметре.



Слика 4. Приказ сијамске неуронске мреже

Циљ сијамских неуронских мрежа је да науче функцију која може да одреди да ли два улаза припадају истој класи или различитим класама. Ово се постиже минимизирањем удаљености између репрезентација сличних улаза и максимизирањем удаљености између репрезентација различитих улаза. Једна од кључних мотивација за сијамске неуронске мреже је њихова примена у учењу на основу једног примера (енг. „*one-shot learning*“) [2]. Овај задатак подразумева да модел мора исправно да класификује нове примере након што је видео само један пример сваке класе. Учење на основу једног примера је важан изазов јер опонаша начин на који људи уче.

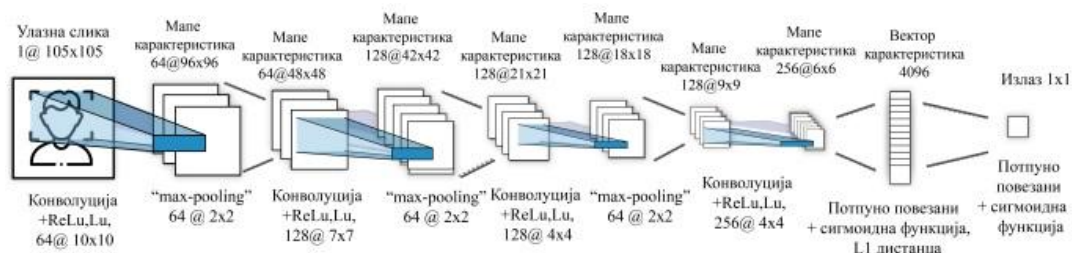
3.2. Архитектура сијамских неуронских мрежа

Подмреже обрађују улазне слике и извлаче карактеристике кроз низ слојева. Након обраде, резултујући вектори карактеристика се пореде коришћењем метричке дистанце (обично „*L1*“

дистанца). Ако је удаљеност мања од одређеног прага, уноси се сматрају сличним. Уместо типичног задатка класификације, сијамске неуронске мреже се тренирају да верификују да ли два уноса припадају истој класи [3].

3.3. Креирање неуронске мреже

Архитектура мреже (Слика 5) је сијамска неуронска мрежа која у основи користи конволуциону неуронску мрежу за извлачење карактеристика улазних слика. Циљ мреже је да утврди да ли улазне слике припадају истој особи, односно врши поређење улазне слике са сликом за валидацију.



Слика 5. Архитектура мреже

4. ПРОГРАМСКО РЕШЊЕ

Програмско решење проблема верификације лица, које је реализовано у програмском језику „Python“ због његове широке примене и подршке за рад са неуронским мрежама. За решавање проблема верификације лица (у овом случају верификације аутора рада), коришћена је сијамска неуронска мрежа („SNN“) која у себи садржи конволуционе слојеве ради побољшања перформанси.

4.1. Скуп података

Коришћени скупови података: Креирана су три скупа података: „anchor“ (референтни скуп), „positive“ (позитиван скуп) и „negative“ (негативан скуп).

- „Anchor“ скуп: Садржи референтне слике особе коју треба верификовати.
- „Positive“ скуп: Укључује додатне слике исте особе за формирање позитивних парова.
- „Negative“ скуп: Обухвата слике различитих особа преузете из скупа „Labeled Faces in the Wild“ (LFW) како би се обезбедили негативни примери [4].

Да би се повећала робустност модела, технике аугментације података су примењене на слике из „anchor“ и „positive“ скупова. То је укључивало насумичне промене осветљења, контраста, окретања, квалитета и засићења, чиме је сваки скуп проширен са 300 на 3.000 слика. Све слике су промењене на величину од 100x100 пиксела и нормализоване како би се обезбедио униформан улаз за неуронску мрежу.

4.2. Неуронска мрежа

Модел обрађује улазне слике кроз вишеструке конволуционе слојеве са „MaxPooling“ и густо повезаним слојевима како би извукао значајне карактеристике лица у векторску репрезентацију.

Слој удаљености „L1“ израчунава апсолутну разлику између кодованих векторских репрезентација („embedding-a“) две слике како би проценила сличност. Конструкција сијамске мреже и слој удаљености су комбиновани да формирају сијамску мрежу, која даје резултат сличности указујући да ли две слике представљају исту особу.

4.3. Тренирање неуронске мреже

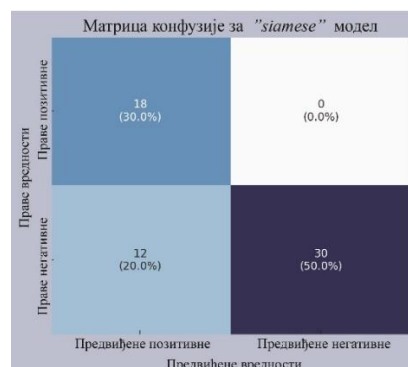
Бинарна унакрсна ентропија је коришћена као функција губитка, а Адам оптимизатор са коефицијентом учења од 0,0001 је примењен за тренинг [5].

Дефинисане су прилагођене функције за тренинг које извршавају тренинг кроз више епоха, укључујући обраду у серијама („batch processing“) и метрике перформанси као што су прецизност и одзив. Обучени модел је сачуван у .h5 формату за будућу употребу без потребе за поновним тренингом.

5. РЕЗУЛТАТИ

Сваки модел је тестиран на 30 позитивних и 30 негативних примера, а матрице конфузије су креиране за сваки од њих (слике 6, 7 и 8). Анализирана су три модела: сијамска неуронска мрежа, „CNN“ модел и „Simple Stroke“ модел

Матрица конфузије сијамске неуронске мреже:



Слика 6: Матрица конфузије модела сијамске неуронске мреже

- Прецизност: 100%

- Одзив: 60%
- Тачност: 80%

Модел сијамске неуронске мреже постиже изузетно високу прецизност, што указује да успешно избегава лажне позитивне. Међутим, одзив од 60% сугерише да пропушта значајан број стварних позитивних примера. Ово значи да модел боље препознаје негативне него позитивне примере, што може бити проблем у апликацијама где је критично идентификовати све позитивне случајеве.

Матрица конфузије за „Simple Stroke“ модел:

Матрица конфузије за „simple stroke“ модел		Предвиђене вредности	
Праве вредности	Праве позитивне	23 (38.3%)	20 (33.3%)
	Праве негативне	7 (11.7%)	10 (16.7%)
		Предвиђене позитивне	Предвиђене негативне

Слика 7: Матрица конфузије „Simple stroke“ моде

- Прецизност: 53,5%
- Одзив: 76,7%
- Тачност: 55%

„Simple stroke“ модел, који се ослања на препознавање структурних карактеристика (контуре, облике) лица, има најнижу прецизност и тачност међу свим анализираним моделима. Иако одзив од 76,7% показује да успева да идентификује већину стварних позитивних примера, ниска прецизност ограничава његову укупну ефикасност.

Матрица конфузије за „CNN“ модел:

Матрица конфузије за „CNN“ модел		Предвиђене вредности	
Праве вредности	Праве позитивне	27 (45.0%)	14 (23.3%)
	Праве негативне	3 (5.0%)	16 (26.7%)
		Предвиђене позитивне	Предвиђене негативне

Слика 8: Матрица конфузије „CNN“ модел

- Прецизност: 65,9%
- Одзив: 90%
- Тачност: 71,7%

„CNN“ модел показује добар баланс између прецизности и одзива. Висок одзив указује на

способност модела да ефикасно идентификује већину позитивних примера. Ипак, умерена прецизност значи да постоји одређен број лажних позитивних, што може утицати на примену модела у пракси где је прецизност критична.

6. ЗАКЉУЧАК

Утврђено је да сијамска неуронска мрежа пружа веома поуздане резултате за овај тип задатка. Модел сијамске неуронске мреже показао је изузетно високу прецизност од 100%, чиме је елиминисан број лажних позитивних предикција. Са тачношћу од 80%, модел се показао стабилним у разликовању позитивних и негативних примера. Међутим, одзив модела је био ограничен на 60%, што указује на то да пропушта одређени број позитивних примера. Ово је важан аспект који би требало побољшати у будућим истраживањима.

Предности сијамске неуронске мреже у верификацији лица су вишеструке. Модел, заснован на принципу „one-shot“ учења, показао се ефикасним и при раду са мањим скуповима података. Висок број тачних негативних предикција указује на способност модела да успешно препознаје примере који не припадају циљној класи, односно особи за коју се врши верификација.

7. ЛИТЕРАТУРА

- [1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
- [2] Hesaraki, S. (2023, Novembar 15). One-shot learning. Preuzeto sa: <https://medium.com/@saba99/one-shot-learning-312e385993bc>
- [3] Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition.
- [4] Labeled Faces in the Wild (LFW) Dataset. (2018). Preuzeto sa: <https://vis-www.cs.umass.edu/lfw/>
- [5] Raschka, S., Liu, Y., & Mirjalili, V. (2022). Machine Learning with PyTorch and Scikit Learn. Packt Publishing.

Кратка биографија:



Петар Сурла рођен је 2001. године у Новом Саду. Факултет техничких наука је уписао 2019. године на смеру Мехатроника. Дипломски рад под називом „Примена неуронске мреже за детекцију и класификацију аудио догађаја“ на Факултету техничких наука одбранио је 2023. године.