

## СИЈАМСКЕ НЕУРОНСКЕ МРЕЖЕ

### SIAMESE NEURAL NETWORKS

Никола Стојановић, Факултет техничких наука, Нови Сад

Област – ЕЛЕКТРОТЕХНИКА И РАЧУНАРСТВО

**Кратак садржај** – У овом раду је приказан експеримент идентификације знакова са два различита сијамским неуронским мрежама, те истраживање какав ефекат могу имати афине трансформације на тренирање и тестирање неуронских мрежа. Две различите сијамске неуронске мреже су имплементирани и тестирани у различитим сценаријима. Резултати су показали високу прецизност у већини случајева. Дати су и предлози за истраживање у будућности.

**Кључне речи:** Сијамске неуронске мреже, вештачка интелигенција, препознавање слика

**Abstract** – In this paper an experiment in character identification with two different Siamese neural networks is presented, along with the effects of affine transformations on training and testing. Two different Siamese neural networks were implemented and tested in different scenarios. In most scenarios results showed high precision. Guidelines for future research are also presented.

**Keywords:** Siamese neural networks, artificial intelligence, image recognition

#### 1. УВОД

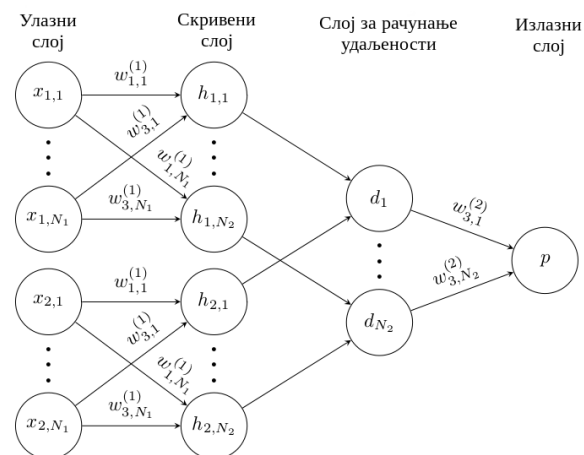
Сијамске неуронске мреже су посебна врста неуронских мрежа које се састоје од најмање две идентичне неуронске мреже (подмреже) спојене на излазу. Под појмом идентичних неуронских мрежа подразумевамо неуронске мреже које имају исту конфигурацију (тј. исте параметре и тежине). Сијамске неуронске мреже се користе како би се поредили улази мапирањем високодимензионалног улаза на вектор карактеристика. На овај начин, мрежа учи меру сличности у подацима, уместо како класификовати улаз у неку од дискретних класа.

Приликом тренирања исправке тежина се одраде на све подмреже. Сијамска неуронска мрежа (слика 1), за разлику од других неуронских мрежа, садржи две неуронске мреже које се састоје од  $N_1$  улазних неурона, са  $N_2$  потпуно повезаних неурона у скривеном слоју. Након скривеног слоја следи слој за рачунање удаљености, који се састоји од  $N_2$  неурона. Коначно, након слоја за рачунање удаљености, следи излазни слој  $p$  који се састоји од једног неурона.

#### НАПОМЕНА:

Овај рад проистекао је из мастер рада чији ментор је био др Стеван Гостојић, ван. проф.

Како су тежине ( $w$ ) унутар двеју подмрежа идентичне, две веома сличне слике се не могу мапирати на две потпуно другачије локације у простору атрибута, тј. ни њихова удаљеност, добијена из излазног слоја  $p$ , не може бити велика.



Слика 1. Структура сијамске неуронске мреже [1]

Циљ овога рада је извршити тренирање неуронских мрежа са и без афиних трансформација, те одредити који ефекат оне имају на прецизност и перформансе неуронских мрежа.

#### 2. СРОДНИ РАДОВИ

##### 2.1 Поновна идентификација објекта

Док се конвенционалне методе поновне идентификације возила фокусирају, или на препознавање броја таблица или на препознавање облика возила, аутори у [2] су одлучили да направе хибридни модел. Тако се слике два возила шаљу једној сијамској неуронској мрежи, док се слике двеју таблица шаљу другој сијамској неуронској мрежи. Ове две мреже су на својим крајевима спојене на слојеве за предикцију, чинећи тиме двоструку сијамску конволутивну неуронску мрежу. Овај модел даје боље резултате када се пореди са једноструким моделима. Чак и са повећаним бројем лажно позитивних примера,  $F$  мера остаје иста. То указује на робустност модела. Кад је реч о поновном препознавању особа, рад [3] представља нову архитектуру која истовремено учи атрибуте и метрику сличности за препознавање основе, за разлику од других решења где се атрибути и метрика сличности уче независно једно од другог. Механизам пажње за бирање најрелевантнијих атрибута је такође предложен. Експерименти су извршени над три скупа података и резултати показују да модел даје резултате као и друга најмодернија решења.

## 2.2 Проналажење слика

Однедавно, у пољу проналажења слика на основу садржаја, појавио се проблем њиховог проналажења без икаквог претходног знања. Радови презентовани овде показују да се сијамске неуронске мреже могу користити на разне начине као помоћ у решавању овог проблема. Рад [4] представља нов начин проналажења слика из слика докумената. Модел се заснива на откривању сваког документа који садржи дати сликовни упит, као и на проналаску шаблона за лоцирање упита на свакој слици.

Сијамска неуронска мрежа је тренирана на *ImageNet* скупу података како би научила разликовати сличне и различите парове слика те је касније коришћена као екстрактор атрибута. Приликом тестирања, предложени метод је прецизнији од метода са подударењем кључних тачака, али не и од метода са филтером базираним на хомографији. Имао је, додуше, бољи одзив од других метода.

У раду [5], показано је да се бинарне хеш вредности за брзи проналазак слика могу генерисати сијамским моделом. Штавише, презентован је нови алгоритам за генерисање позитивних и негативних парова приликом тренирања сијамске неуронске мреже. Овај алгоритам бира такве негативне парове који ће највише да утичу на промену функције грешке. Кад се пореде резултати са најсавременијим хеш методама, модел презентован у раду показује напредак како се број битова повећава, што је последица новог тренинг алгоритма.

## 3. ОПИС ЕКСПЕРИМЕНТА

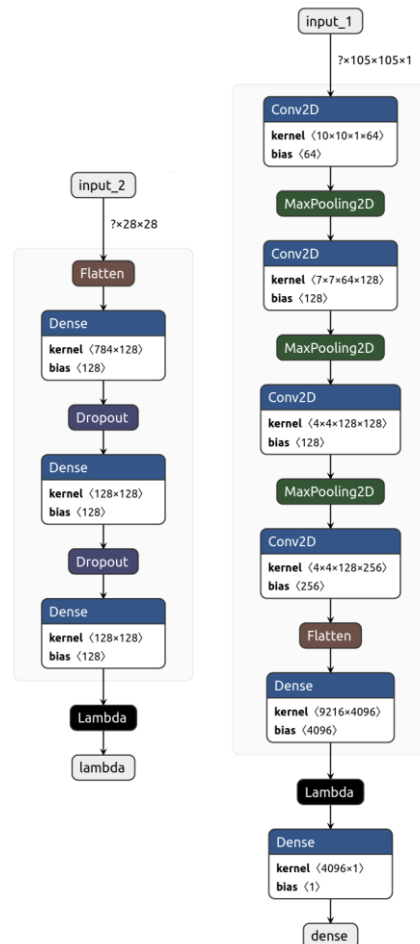
Тестирана су два типа сијамских неуронских мрежа за препознавање слика:

1. Сијамска неуронска мрежа која је развијена као пример за Пајтон (енг. *Python*) библиотеку Керас (енг. *Keras*) [6]. На слици 2 се може видети архитектура једног њеног близанца. Он се састоји од три потпуно повезана слоја између којих се налазе слојеви изостављања. Активациона функција на сваком потпуно повезаном слоју је исправљена линеарна јединица. Последњи слој је ламбда слој који рачуна Еуклидско растојање између два излаза. За оптимизацију приликом тренинга је коришћен модификовани стохастички градијентни спуст тзв. *RMSPProp*. Приликом тренинга за функцију грешке коришћена је контрастивна функција дефинисана у [7]. Ова неуронска мрежа је тренирана и тестирана на *MNIST* скупу података.

2. Сијамска неуронска мрежа за препознавање слика на једном примеру коју су развили Кох и остали [1]. На слици 2 је илустрована архитектура једног њеног близанца. То је конволутивна неуронска мрежа која се састоји од четири конволутивна слоја, са слојевима агрегације између сваког конволутивног слоја. Након последњег конволутивног слоја следи слој поравњавања који претвори добијене вредности у један вектор.

На сваки слој се примени филтер одређених димензија, при чему су димензије број дељив са 16. Активациона функција на сваком слоју је исправљена линеарна јединица. Последњи слој у сваком близанцу је потпуно повезани слој са сигмоидном активацио-

ном функцијом која даје предикцију, након чега је спојен са слојем који рачуна удаљеност.



Слика 2. Структура Керасове (лево) и Кохове (десно) сијамске неуронске мреже. Слика је добијена помоћу алата Netron [8]

Затим следи последњи слој за предвиђање који садржи сигмоидну активациону функцију. За иницијализацију тежина и коефицијента склоности неуронске мреже коришћена је униформна дистрибуција, док је за оптимизацију употребљен стохастички градијентни спуст. У сваком конволуционом слоју за регуларизацију је коришћена  $l_2$  регуларизациона казна. Кохова сијамска неуронска мрежа је тренирана и тестирана на *Omniglot* скупу података.

Како би се израчунала прецизност и *F1* мера, припремљени су посебни тест сценарији. Прво су формиран само позитивни, а потом само негативни парови (за сваку класу) који су даље слани неуронској мрежи да утврди да ли су у питању исте или другачије слике. Тестирање је прво извршено само над позитивним, а потом само над негативним паровима. Из ових резултата добијене су стопе истински позитивних, лажно позитивних, истински негативних и лажно негативних примера.

## 4. РЕЗУЛТАТИ

### 4.1 Резултати на *MNIST* скупу

Резултати тестирања по класама (знаменкама) су приказани у табелама 1-4.

Табела 1. Резултати тренирања и тестирања без трансформација

Класа	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
0	435	11	1776	6	0.99	0.98	0.98
1	433	13	1774	8	0.98	0.97	0.98
2	430	16	1777	5	0.99	0.96	0.98
3	429	17	1774	8	0.98	0.96	0.97
4	428	18	1775	7	0.98	0.96	0.97
5	418	28	1776	6	0.99	0.94	0.96
6	432	14	1771	11	0.98	0.97	0.97
7	414	32	1776	6	0.99	0.93	0.96
8	413	33	1773	9	0.98	0.93	0.95
9	417	29	1772	10	0.98	0.93	0.96
Просек	427.4	18.6	1774.9	7.1	0.98	0.95	0.97

Табела 2. Резултати тренирања без трансформација и тестирања са трансформацијама

Класа	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
0	308	138	1665	117	0.72	0.69	0.71
1	321	125	1703	79	0.80	0.72	0.76
2	344	102	1722	60	0.85	0.77	0.81
3	314	132	1653	129	0.71	0.70	0.71
4	318	128	1664	118	0.73	0.71	0.72
5	310	136	1703	79	0.80	0.70	0.74
6	318	128	1683	99	0.76	0.71	0.74
7	280	166	1684	98	0.74	0.63	0.68
8	310	136	1635	147	0.68	0.70	0.69
9	289	157	1648	134	0.68	0.65	0.67
Просек	314.1	131.9	1684.6	97.4	0.75	0.70	0.72

Табела 3. Резултати тренирања са трансформацијама и тестирања без трансформација

Класа	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
0	437	9	1777	5	0.99	0.98	0.98
1	435	11	1779	3	0.99	0.98	0.98
2	430	16	1776	6	0.99	0.96	0.98
3	433	13	1776	6	0.99	0.97	0.98
4	428	18	1772	10	0.98	0.96	0.97
5	433	13	1776	6	0.99	0.97	0.98
6	432	14	1777	5	0.99	0.97	0.98
7	435	11	1775	7	0.98	0.98	0.98
8	416	30	1774	8	0.98	0.93	0.96
9	415	31	1769	13	0.97	0.93	0.95
Просек	432.9	13.1	1776.0	6.0	0.98	0.96	0.97

Табела 4. Резултати тренирања и тестирања са трансформацијама

Класа	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
0	423	23	1772	10	0.98	0.95	0.96
1	429	17	1777	5	0.99	0.96	0.98
2	404	42	1761	21	0.95	0.91	0.93
3	399	47	1748	34	0.92	0.89	0.91
4	412	34	1751	31	0.93	0.92	0.93
5	400	46	1751	31	0.93	0.90	0.91
6	419	27	1770	12	0.97	0.94	0.96
7	409	37	1762	20	0.95	0.92	0.93
8	382	64	1756	26	0.94	0.86	0.89
9	388	58	1754	28	0.93	0.87	0.90
Просек	411.9	34.1	1761.5	20.5	0.95	0.91	0.93

#### 4.2 Резултати на Omniglot скупу

Резултати тестирања по класама (алфабетима) су приказани у табелама 5-8.

Када се сагледају сви резултати, види се да су најбољи резултати постигнути на MNIST скупу приликом тренирања са афиним трансформацијама када су резултати прецизности изнад 90%. Увођење трансформација у тестирање доводи до опадања прецизности, при чему она опада 23% и 3% на MNIST скупу, односно 1% и 0% на Omniglot скупу.

Табела 5. Резултати тренирања и тестирања без трансформација

Алфабет	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
Angelic	461	59	455	39	0.92	0.89	0.90
Atemayar Qelisyayer	734	86	687	92	0.89	0.90	0.89
Atlantean	408	92	440	35	0.92	0.82	0.87
Aurek-Besh	412	108	473	21	0.95	0.79	0.86
Avesta	515	5	463	31	0.94	0.99	0.97
Ge'ez	463	57	453	41	0.92	0.89	0.90
Glagolitic	509	91	525	45	0.92	0.85	0.88
Gurmukhi	364	36	371	9	0.98	0.91	0.94
Kannada	709	131	719	79	0.90	0.84	0.87
Keble	815	85	766	89	0.90	0.91	0.90
Malayalam	828	112	806	87	0.90	0.88	0.89
Manipuri	458	62	469	25	0.95	0.88	0.91
Mongolian	497	63	474	58	0.90	0.89	0.89
Old Church Slavonic	745	155	779	76	0.91	0.83	0.87
Oriya	707	213	822	52	0.93	0.77	0.84
Sylheti	474	46	471	23	0.95	0.91	0.93
Syriac Serto	783	117	791	64	0.92	0.87	0.90
Tengwar	418	102	439	55	0.88	0.80	0.84
Tibetan	672	128	669	91	0.88	0.84	0.86
ULOG	399	61	399	38	0.91	0.87	0.89
Средња вредност	568.6	90.5	573.6	52.5	0.92	0.87	0.89

Табела 6. Резултати тренирања без трансформација и тестирања са трансформацијама

Алфабет	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
Angelic	827	213	916	72	0.92	0.80	0.85
Atemayar Qelisyayer	1269	371	1397	161	0.89	0.77	0.83
Atlantean	773	227	865	85	0.90	0.77	0.83
Aurek-Besh	789	251	932	56	0.93	0.76	0.84
Avesta	978	62	904	84	0.92	0.94	0.93
Ge'ez	871	169	904	84	0.91	0.84	0.87
Glagolitic	960	240	1072	68	0.93	0.80	0.86
Gurmukhi	674	126	743	17	0.98	0.84	0.90
Kannada	1244	436	1389	207	0.86	0.74	0.79
Keble	1492	308	1506	204	0.88	0.83	0.85
Malayalam	1452	428	1596	190	0.88	0.77	0.82
Manipuri	815	225	922	66	0.93	0.78	0.85
Mongolian	956	164	971	93	0.91	0.85	0.88
Old Church Slavonic	1351	449	1605	105	0.93	0.75	0.83
Oriya	1260	580	1640	108	0.92	0.68	0.79
Sylheti	865	175	923	65	0.93	0.83	0.88
Syriac Serto	1413	387	1536	174	0.89	0.79	0.83
Tengwar	723	317	895	93	0.89	0.70	0.78
Tibetan	1275	325	1375	145	0.90	0.80	0.84
ULOG	729	191	809	65	0.92	0.79	0.85
Средња вредност	1035.8	282.2	1145.0	107.1	0.91	0.79	0.85

Разлика у резултатима између Керасове и Кохове неуронске мреже долази због тога што Кохова неуронска мрежа садржи конволутивне слојеве, праћене са слојевима агрегације који неуронској мрежи дају способност инваријантности. Једноставније речено, неуронска мрежа ће бити у стању препознати објекат иако је он померен у односу на центар слике.

Са друге стране, Керасова неуронска мрежа је потпуно повезана неуронска мрежа, без конволутивних слојева, што значи да она нема особину инваријантности.

Табела 7. Резултати тренирања са трансформацијама и тестирања без трансформација

Алфабет	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
Angelic	439	81	454	40	0.92	0.84	0.88
Atemayar Qelisayer	702	118	674	105	0.87	0.86	0.86
Atlantean	423	77	425	50	0.89	0.85	0.87
Aurek-Besh	417	103	463	31	0.93	0.80	0.86
Avesta	513	7	448	46	0.92	0.99	0.95
Ge'ez	450	70	449	45	0.91	0.87	0.89
Glagolitic	513	87	542	28	0.95	0.86	0.90
Gurmukhi	364	36	362	18	0.95	0.91	0.93
Kannada	696	144	711	87	0.89	0.83	0.86
Keble	797	103	740	115	0.87	0.89	0.88
Malayalam	831	109	809	84	0.91	0.88	0.90
Manipuri	460	60	477	17	0.96	0.88	0.92
Mongolian	498	62	464	68	0.88	0.89	0.88
Old Church Slavonic	755	145	783	72	0.91	0.84	0.87
Oriya	792	128	806	68	0.92	0.86	0.89
Sylheti	460	60	472	22	0.95	0.88	0.92
Syriac Serto	789	111	778	77	0.91	0.88	0.89
Tengwar	403	117	435	59	0.87	0.78	0.82
Tibetan	683	117	686	74	0.90	0.85	0.88
ULOG	376	84	406	31	0.92	0.82	0.87
<b>Средња вредност</b>	<b>568.1</b>	<b>91.0</b>	<b>569.2</b>	<b>56.9</b>	<b>0.91</b>	<b>0.86</b>	<b>0.89</b>

Табела 8. Резултати тренирања и тестирања са трансформацијама

Алфабет	Истински позитивни	Лажни негативни	Истински негативни	Лажни позитивни	Прецизност	Одзив	F1 мера
Angelic	811	229	920	68	0.92	0.78	0.85
Atemayar Qelisayer	1302	338	1286	172	0.88	0.79	0.84
Atlantean	740	260	859	91	0.89	0.74	0.81
Aurek-Besh	749	291	916	72	0.91	0.72	0.80
Avesta	973	67	907	81	0.92	0.94	0.93
Ge'ez	799	241	890	98	0.89	0.77	0.82
Glagolitic	831	369	1060	80	0.91	0.69	0.79
Gurmukhi	629	171	734	26	0.96	0.79	0.86
Kannada	1222	458	1441	155	0.89	0.73	0.80
Keble	1481	319	1528	182	0.89	0.82	0.86
Malayalam	1518	362	1609	177	0.90	0.81	0.85
Manipuri	813	227	939	49	0.94	0.78	0.85
Mongolian	901	219	971	93	0.91	0.80	0.85
Old Church Slavonic	1317	483	1592	118	0.92	0.73	0.81
Oriya	1364	476	1581	167	0.89	0.74	0.81
Sylheti	849	191	943	45	0.95	0.82	0.88
Syriac Serto	1385	415	1574	136	0.91	0.77	0.83
Tengwar	706	334	911	77	0.90	0.68	0.77
Tibetan	1223	377	1344	176	0.87	0.76	0.82
ULOG	686	234	826	48	0.93	0.75	0.83
<b>Средња вредност</b>	<b>1015.0</b>	<b>303.1</b>	<b>1141.6</b>	<b>105.6</b>	<b>0.91</b>	<b>0.77</b>	<b>0.83</b>

## 5. ЗАКЉУЧАК

У раду су тестиране две различите сијамске неуронске мреже – Керасова неуронска мрежа и неуронска мрежа коју су представили Кох и други.

Обе мреже су прављене да би решавале проблем идентификације руком писаних симбола. Тренирање и тестирање неуронских мрежа је извршено са и без употребе насумичних афиних трансформација. Осим у једном случају, резултати показују висок ниво прецизности (изнад 90%). Кориштење афиних трансформација приликом тестирања доводи до пада прецизности, F1 мере и одзива приликом тестирања. Међутим, тај пад није исти у сваком примеру.

Приликом тренинга Керасове неуронске мреже на MNIST скупу, пад у прецизности је износио 23%, да би се увођењем афиних трансформација у тренинг, смањено на 3%. Насупрот томе, код Кохове неуронске мреже пад у прецизности је износио 1%, да би се увођењем афиних трансформација у тренинг изгубио.

Разлог зашто је Кохова неуронска мрежа показала мањи пад у прецизности јесте у томе што је она конволутивна неуронска мрежа, при чему је она инваријантна на транслације, док Керасова није.

У циљу добијања бољих резултата, потребно је тестирати другачији избор негативних парова приликом тренинга. У овој раду негативни парови су бирани насумично, док би једна опција била користити само оне парове који ће највише допринети смањењу функције грешке приликом тренинга. Аутор сматра да би другачији избор негативних парова коначно могао повећати прецизност на Omniglot скупу изнад људске прецизности.

## 4. ЛИТЕРАТУРА

- [1] Koch, G., Zemel, R. and Salakhutdinov, R., 2015. "Siamese neural networks for one-shot image recognition." In ICML deep learning workshop (Vol. 2).
- [2] de Oliveira, I.O., Fonseca, K.V. and Minetto, R., 2019. "A Two-Stream Siamese Neural Network for Vehicle Re-Identification by Using Non-Overlapping Cameras." arXiv preprint arXiv:1902.01496.
- [3] Wu, L., Wang, Y., Gao, J. and Li, X., 2018. "Where-and-when to look: Deep siamese attention networks for video-based person re-identification." IEEE Transactions on Multimedia, 21(6), pp.1412-1424.
- [4] Wiggers, K.L., Britto Jr, A.S., Heutte, L., Koerich, A.L. and Oliveira, L.S., 2019. "Image Retrieval and Pattern Spotting using Siamese Neural Network." arXiv preprint arXiv:1906.09513.
- [5] Jose, A., Yan, S. and Heisterklaus, I., 2017, September. "Binary hashing using siamese neural networks." In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 2916-2920). IEEE.
- [6] [https://keras.io/examples/mnist\\_siamese/](https://keras.io/examples/mnist_siamese/) (приступљено 13.8.2019.)
- [7] Hadsell, Raia, Sumit Chopra, and Yann LeCun. "Dimensionality reduction by learning an invariant mapping." In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2, pp. 1735-1742. IEEE, 2006.
- [8] <https://lutzroeder.github.io/netron/> (приступљено 9.1.2020.)

## Кратка биографија:



**Никола Стојановић** рођен је 3.10.1995. године у Сомбору. Мастер рад на Факултету техничких наука из области Рачунарства и аутоматике – Интелигентни системи одбранио је 2020. год. Контакт: [stojanovicn129@gmail.com](mailto:stojanovicn129@gmail.com)