

КЛАСИФИКАЦИЯ С КОНВОЛЮЦИОННИ НЕВРОННИ МРЕЖИ

Александър Иванов, Пенка В. Георгиева
Бургаски свободен университет
alex.milagro@abv.bg
penka.georgieva@bfu.bg

CLASSIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS

Alexander Ivanov, Penka V. Georgieva
Burgas Free University

Abstract: *This paper studies convolutional neural networks as a widely used tool for image recognition. Another area of applications of the convolutional neural networks is the process of classification. Test results with a new type of activation function in a convolutional neural network are presented in the paper.*

Keywords: *convolutional neural networks, classification, artificial intelligence*

I. ВЪВЕДЕНИЕ

Изкуственият интелект е направление в компютърната наука. Терминът изкуствен интелект е въведен от Джон Макарти през 1956 като „науката за създаване на интелигентни машини“. Исторически изкуственият интелект възниква от опитите за формализиране на човешкото знание със средствата на математическата логика и първоначално се прилага за доказване на теореми и за моделиране на игри. Основа на изкуствения интелект стават различни науки: философия, математика, алгоритми, логика, психология, информатика, лингвистика, но постепенно това направление се обогатява с други обекти на изследване и други средства. В настоящия момент изкуственият интелект е съвкупност от традиционния изкуствен интелект, обогатен с различни методологии за числено пресмятане, като обект на изкуствения интелект (за разлика от неизкуствения интелект) стават основно онези задачи, за които липсват директни математически или логически алгоритми и могат да бъдат решавани само интуитивно. Към традиционните базисни за изкуствения интелект науки се добавят нови: компютърни науки, неврология, когнитивна наука, онтология, изследване на операциите, икономика, вероятности, оптимизиране. Областите на приложение на средствата на изкуствения интелект са разнообразни: експертни системи, игри, доказателство на теореми, обработка на естествени езици, разпознаване на образи, роботика, навигация, системи за управление, системи за планиране, извличане на данни, логистика и много други. [1]

Основните изчислителни парадигми на съвременния изкуствен интелект са:

- размити системи, включващи размита логика;
- еволюционни пресмятания, включващи генетични алгоритми;
- изкуствени невронни мрежи;
- машинно обучение;
- вероятностни изводи.

Изследванията в областта на невронните мрежи са свързани с идеята за създаване на интелигентна система за изчисления, която да работи по начин, близък до човешкия мозък.

В областта на автоматичното управление невронни мрежи се използват при работване на системи за управление на работи, манипулатори, транспортни средства, при идентификация и управление на сложни обекти, чиито динамични модели се характеризират с неопределеност и неточност. Освен в автоматиката, изкуствените невронни мрежи се използват и в редица други направления – финанси, медицина, военно дело, криминалистика, разпознаване на образи и др.

Невронните мрежи се състоят от елементи за обработка на информация, наречени неврони или възли. Невроните са свързани и теглата на връзките между тях определят силата на съответните връзки. Входната информация за всеки неврон е претеглената сума от сигналите от останалите неврони. Тази информация се акумулира в неврона като изходният му сигнал се определя посредством активационна функция. Най-простата активационна функция е функцията на Хевисайд. Тази функция е прекъснатата, което при обучителни алгоритми, базирани на градиент, налага заместването ѝ с непрекъснати и диференцируеми функции, пр. сигмоидални. Разнообразието от използвани активационни функции не се изчерпва до споменатите, като някои алтернативи са радиално-базисни функции, ReLU, eLU, softmax и други.

Изкуствените невронни мрежи са модели на биологичните невронни мрежи в следните два аспекта:

- информацията се натрупва в мрежата в процес на обучение;
- силата на връзките между отделните неврони се моделира с теглата на съответните връзки.

Настоящата работа представя възможно решение на проблема за класификация на изображения. Целта е да се определи към кой клас принадлежи дадено входно изображение. Начинът за постигане на това е създаването и тренирането на изкуствена невронна мрежа върху изображения, при което невронната мрежа трябва да определи към кой клас принадлежи даденото изображение.

II. ПРИЛОЖЕНИЯ НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В МОРСКОТО ДЕЛО

В последните десетилетия Изкуствения интелект е сред водещите теми на научния фокус и много експерти виждат революция в технологиите и практиката в областта, която е породена както от развитието на хардуера, обезпечаващо големите нужди от изчислителна мощ, така и от иновативни теоретични концепции, базирани на открития от невробиологията, когнитивната наука и статистиката. Морското дело не е изключение от тази тенденция. Техниките на Изкуствения интелект се използват при автоматизирането на навигационни и товарни системи [2], в обработката на големи обеми от информация, свързани с мониторинга на солените водоноси и техните екосистеми, предсказването на стихийни бедствия [3], идентифицирането на ресурси (например полезни изкопаеми), сигурността в корабоплаването [4], управлението и планирането на икономически дейности (например риболов).

Географските информационни системи (ГИС), предложени за пръв път от Р. Томсън през 1968 г., са широко използвани в съвременните пристанищни администрации. Основните източници на данни за тези системи са топографските карти, като също могат да се използват сателитни и въздушни изображения. Информация, извлечена от такива системи, може да се съхранява в растерни или векторни изображения. Популярни файлови формати за растерни изображения са DRG, GeoTIFF, JPEG2000, MrSID, а за векторни – DXF, DLG, GML, Spatialite, Shapefile и др. В ГИС се използват и geodetic datums, които представляват координатни системи с набор от референтни точки, използвани за локализиране на обекти на земната повърхност. World Geodetic System (WGS) е Geodetic datum, съдържаща координатна система, сферична референтна повърхност и гравитационна еквипотенциална повърхност. Други представители на този вид информационни структури са EGM2008, ED50, OSGB36.

ГИС реформират в голяма степен морското дело, като дефинират нов вид бизнес интелигентност – „пространствена“. Целта е да се трансформира векторна информация в друг вид данни, необходими в практиката. Двете основни характеристики на векторната информация са наклон и аспект. Наклонът обозначава тримерните характеристики на релефа, а аспектът – пространствената ориентация спрямо други обекти. Чрез ГИС могат да се изградят геометрични мрежи. Тези мрежи представляват описания на обекти и връзки между тях и са линейни структури, сходни с графите. Някои от приложенията им са в управлението на транспорта, градското планиране, публичните услуги, електрификацията и други.

Разпознаване на кораби на въздушни снимки чрез конволюционни невронни мрежи е предложено в [5]. Използвани са данни от NWPU VHR-10, който е публично достъпен и използван единствено за научни изследвания масив от данни. Масивът съдържа десет класа геопространствена информация – „самолети“, „кораби“, „пристанища“, „тенис кортове“, „мостове“, „складове“, „бейзболни игрища“, „баскетболни игрища“, „писти“ [6].

III. КОНВОЛУЦИОННИ НЕВРОННИ МРЕЖИ

Конволюционните невронни мрежи (КНМ) са клас изкуствени невронни мрежи с праволинейни връзки. КНМ са вдъхновени от биологични процеси в животинския визуален кортекс и са предложени като наследник на неокогнитрона от Ян Лекун в [7]. КНМ имат широко приложение в практически проблеми при обработка на изображения, обработка на естествени езици, в препоръчващи системи, но също и при класификация, например в големи бази от изображения и други.

Класическата архитектура на конволюционните мрежи се състои от 3 вида скрити слоя, изпълняващи различни функции (фиг. 1).



Фигура 1. Архитектура на КНМ

Първият вид от слоевете е конволюционният. Входовете на невронната мрежа са свързани към участъци от входната информация, наречени полета на възприятие и наподобяващи фоточувствителните клетки в ретината на окото. Всяко поле на възприятие има определен фиксиран размер, еднакъв за всички неврони в съответния слой. Тези полета на възприятие обхождат входната матрица от данни (пр. пиксели) през определена стъпка. От математическа гледна точка операцията, извършвана от конволюционните слоеве е крос-корелация и се състои в откриването на сигнал в рамките на друг сигнал. Синаптичните тегла във всички полета на възприятие са еднакви, което води до значително намаление на броя на обучаемите параметри на мрежата и предпазва от проблеми при спускане по градиента.

Вторият вид от слоевете е сливащият. В такъв слой изходните стойности от група неврони се изпраща към единствен неврон от следващия слой. Съществуват различни техники като максимално, средно аритметично (фиг. 2) или сумиращо сливане. Сливането има за цел да открие характеристики на информацията от съседни полета на възприятие и така да открие шаблони на по-високо ниво на абстракция.



Фигура 2. Максимално сливане (вляво) и средно аритметично сливане (вдясно)

Третият вид слоеве е напълно свързан слой, идентичен със скритите слоеве на класическите праволинейни невронни мрежи. Както и при праволинейните мрежи, в напълно свързаните слоеве на КНМ е препоръчително да се използва регуляризация.

Регуляризацията е процес на въвеждане на допълнителни данни или корекция на налични за справяне с некоректно формулиран проблем или превенция на претрениране и се реализира чрез dropout слоеве за коректното възпроизвеждане на некоректни примери, което намалява или напълно блокира възможностите за класификация на нови данни, неналични при обучението.

Хинтън предлага нова архитектура – „Капсулована мрежа“, която надгражда КНМ чрез формирането на „капсули“ от неврони и добавянето на алгоритъм за динамично маршрутизиране между тях [8].

КНМ традиционно се обучават чрез сходни техники като тези за праволинейни невронни мрежи. Обучението е контролирано (с учител), като може да се извършва на няколко етапа – предварително, основно, фина настройка.

Достигнатата минимална грешка при разпознаване на образи с КНМ е 3,57% (табл. 1). [9], [10]

Таблица 1. Постигната точност на класификацията с *КНМ*

Разработка	Година	Грешка
SuperVision	2012	15,3 %
GoogLeNet	2014	6,67 %
ResNet	2015	3,57 %
Експерт		5,1 %

IV. КЛАСИФИКАЦИЯ С КОНВОЛЮЦИОННИ НЕВРОННИ МРЕЖИ

Класификация е задача за разделяне на елементи на множество в отделни класове. В задачата за класификация на изображения се дефинират няколко различни подзадачи – локализиране на обекти, разпознаване на множество обекти и откриване на контури.

Възможен подход за класификация е използване на *КНМ*.

Използваната в тази разработка *КНМ* има следните характеристики:

- в архитектурата има 2 скрити слоя – 1 конволюционен с 4 неврона и 1 сумираещ с 1 неврон;
- сливането е сумиращо;
- активационната функция е дробна логистична [11];
- използван е алгоритъм за обратно разпространение на грешката.

Логистична (сигмоидална) функция се дефинира с формулата:

$$(1) \quad f(x) = \frac{L}{1 + e^{-k(x-x_0)}}$$

Дефиницията на функция на Миттаг-Леффлер е:

$$(2) \quad E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}$$

където Γ е гама функцията параметри α и β , z е комплексен аргумент.

Замествайки експоненциалната функция от (1) с (2), при $L = k = 1$ и $x_0 = 0$ се получава следната дефиниция за дробна логистична функция:

$$(3) \quad f(x) = \frac{1}{1 + E_{\alpha, \beta}(-x)}$$

Подробен анализ на тази функция е направен в [11]. В същата публикация са тествани определени стойности на параметрите α и β във връзка с обучението на праволинейни невронни мрежи.

Като продължение на тези експерименти са направени тестове на обучението на *КНМ* с дробна логистична активационна функция. Част от резултатите са представени в таблица 2. Тестовите са направени върху генерирани данни, които са 4 масива

от числа, интерпретирани като черно-бели изображения с размерност 3×3 . Целта на направените експерименти е да се илюстрира влиянието на стойностите на параметрите върху обучението. За тестовете е използван модифициран код на Python.

Таблица 2. Резултати от проведените експерименти

α	β	Обучителни епохи	М-Л функция*	Грешка преди обучение	Грешка след обучение
1	1	2500	22	0,783147780763	0,221895927387
0,5	-1	2500	22	0,528558520858	0,252199999598
0	-1	2500	22	0,556138985534	0,39264922664
1	-1	2500	22	0,764365559752	0,650841781334
1,7	-1	2500	22	1,22169554113	1,41739375902
1,8	-1	2500	22	1,21679073668	1,41725397982
1,9	-1	2500	22	1,2415135158	1,42479190441
2,1	-1	2500	22	1,03558124386	1,23881851029
2,4	-1	2500	22	227,176300702	0,0224634504472
3	1	2500	22	10,1091000849	1,5706682264
4	1	2500	22	0,352118671916	0,0296797091703

Миттаг-Леффлер функцията се дефинира като безкрайна сума, което налага използването на приближения в практиката. В таблицата **М-Л функция*** е броят изчислявани членове на тази сума за съответния експеримент. Стойности на параметъра α , близки до 2, при фиксирано $\beta=1$, водят до значително редуциране на обучителната грешка.

V. ЗАКЛЮЧЕНИЕ И РАЗВИТИЕ

При проведените експерименти за някои стойности на параметрите на дробната логистична активационна функция се получава по-малка грешка (в най-добрия случай около 5 пъти) в сравнение с логистична функция при еднакъв брой обучителни епохи. Резултатите от експериментите потвърждават някои от откритите тенденции в [11] и демонстрират възможности за използване на дробната логистична функция като активационна функция в конволюционни мрежови архитектури.

Настоящата разработка е обещаваща техника за създаване на по-точни интелигентни системи в морското дело и други области на практиката.

Бъдещи тестове с разширен набор от стойности и при използване на реални данни може да допринесе за по-дълбокото разбиране на поведението на конволюционна невронна мрежа с дробна логистична активационна функция.

Литература:

- [1] П. Георгиева, Изследване на модели на софт компютинг за управление в реално време, Академик Дринов, София , 2013.
- [2] P. Bruno, Maritime Applications for Machine Learning and Data Science. ThoughtCo, 2016.
- [3] W. Cousins, M. Onorato, A. Chabchoub, T. P. Sapsis, Predicting ocean rogue waves from point measurements: an experimental study. 2017 (http://sandlab.mit.edu/Papers/17_SA1.pdf) [4] I. Obradovich, M. Milicevic, K. Zubrinici, Machine learning approaches to maritime anomaly detection, *Nase More* 61 (5):96-101, 2014.
- [5] A. Gallego, A. Pertusa, P. Gil, Automatic Ship Classification from Optical Aerial Images with Convolutional Neural Networks. *Remote Sens.* Vol. 10, 511; 2018.
- [6] <http://www.escience.cn/people/gongcheng/NWPU-VHR-10.html>.
- [7] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. *Proceedings of the the IEEE*, 86(11):2278–2324, 1998.
- [8] G. Hinton, S. Sabour, N. Frosst, Matrix Capsules with EM routing, *ICLR*, 2018.
- [9] G. Cheng, J. Han, A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117: 11-28, 2016.
- [10] G. Cheng, P. Zhou, J. Han, Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7405-7415, 2016.
- [11] A. Ivanov, Fractional activation functions in feed forward artificial neural networks, *XX-th International Symposium on Electrical Apparatus and Technologies*, 2018.