

РЕЗЮМЕТА НА ПУБЛИКАЦИИТЕ ЗА УЧАСТИЕ В КОНКУРСА

на Галина Димитрова Момчева

кандидат за конкурс за заемане на академичната длъжност „доцент“
по професионално направление 4.6. „Информатика и компютърни науки“ (Архитектури на невронни мрежи)
обявен ИМИ-БАН (ДВ, бр. 69 (80.) / 11.08.2023 г.)

НАУЧНИ СТАТИИ

1. Fuzzy U-Net Neural Network Design for Image Segmentation

Kirichev, M., Slavov, T., Momcheva, G.

Lecture Notes in Networks and Systems, 2022, 374 LNNS, pp. 177–184

Един от проблемите в анализа на биомедицинските изображения е проблемът на сегментация на ядра на клетки. Анотацията на ядра на ръка се доказала с вариращи резултати в зависимост от много фактори. По-скорошните конволюционни невронни мрежи направиха проблема с автоматичното сегментиране на изображението по-лесен, по-бърз и по-надежден. Размитите слоеве наподобяват усещане за несигурност, което се вижда в реалния свят В тази статия се предлага разширение на стандартната U-Net и нова архитектура, която има за цел да подобри качеството на сегментирането на биомедицински изображения. С интегрирането на размити изчисления в стандартната U-Net архитектура постигнахме дори по-добри точности от тези, постигнати от базовата архитектура.

2. Fuzzy U-Net Neural Network Architecture Optimization for Image Segmentation

Kirichev, M.M., Slavov, T.S., Momcheva, G.D.

IOP Conference Series: Materials Science and Engineering, 2021, 1031(1), 012077

В тази статия оптимизирането на модифицирания U-Net модел на невронна мрежа, разширен с размити слоеве, е проучено с използването на Grid search и Keras tuner. Статията е продължение на предишна работа, където моделът е предложен и проучен. От една гледна точка, изследването е фокусирано върху оптимизирането на Fuzzy Layers, вградени в U-Net модела, за да се намери по-добрата архитектура на невронната мрежа за сегментиране на ядра в изследователската работа в екосистемата за научноизследователска и развойна дейност на BioMed Varna за сегментиране на клетъчни ядра. В същото време от глобална гледна точка този експеримент е част от по-големия за търсене на нови техники за проектиране на архитектура на невронни мрежи.

3. Sentiment detection with FedMD: Federated learning via model distillation

Tsankova, P., Momcheva, G.

CEUR Workshop Proceedings, 2020, 2656, pp. 236–247

FedML е техника за разпределено машинно обучение, при която клиентските устройства обучават модели локално, без да споделят никакви данни, с изключение на параметър промени, които се агрегират в централен модел. Този подход за запазване на поверителността има огромен потенциал за съчетаване на необходимостта от големи набори от данни за задълбочено обучение нарастващата чувствителност на собствеността върху данните.

Нашата статия използва FedMD (Federated Learning via Model Distillation) алгоритъм и го прилага за първия време в областта на обработката на естествен език. Резултатите са обещаващи с по отношение на решаването на предизвикателствата, свързани с хетерогенността на данните и персонализирането на модела чрез въвеждане на специфични за клиента модели и съвместно обучение, реализирано чрез моделна дестилация. Получената малка разлика между резултатите на FedMD и не-Внедряването на FedMD се компенсира от по-малкото количество данни за обучение моделите на FedMD и успешното запазване на поверителността за локално достъпни данни.

4. Hyperparameter adjustment in regression neural networks for predicting support case durations

Hristov, H., Momcheva, G.

AIP Conference Proceedings, 2021, 2333, 070015

Регресията е мощна техника за прогнозиране на единична скаларна стойност с висока степен на сигурност. Регресионният модел изисква набор от данни, който се състои само от числови характеристики. Въпреки това наборите от данни често съдържат както числови, така и характеристики тип категории. Тази статия има за цел да проучи няколко различни техники за кодиране на текст, които могат да решат този проблем чрез трансформиране на определена текстова стойност в съответна числена стойност, която е статистически разумна по отношение на набора от данни. Сценарият, който изучаваме, е регресионен модел на невронна мрежа, предвиждащ продължителността на случаите на поддръжка. Най-предизвикателният аспект на кодирането е произтичащата от това промяна на размерността на набора от данни. Всеки

метод на кодиране влияе на размерността в различна степен в зависимост от кардиналността на характеристиката. Този резултат е основното предизвикателство за настройката на хиперпараметрите на невронната мрежа. Трябва да направим такава настройка, която да може стабилно да обработва променения набор от данни. Документът сравнява пет различни подхода за кодиране: еднократно, хеширане, двоично, целево и вграждане на обекти. Настройките на хиперпараметъра за всеки подход са представени чрез използване на общи показатели за ефективност на невронната мрежа и базова настройка на невронна мрежа. Можем да заключим, че умерено увеличената размерност може да подобри предсказващата сила на модела, както се наблюдава в случая на двоичния и хеширащия енкодер.

5. Sustainability of Research-based Ecosystem

Ivanova, A., Momcheva, G., Zhekova, R., Tankova, E., Pavlov, S.

AIP Conference Proceedings, 2022, 2505, 060008

Целта на статията е да се проучи устойчивостта на базираната на изследвания предприемаческа екосистема БиоМед-Варна. BioMed-Varna се фокусира върху интердисциплинарни изследвания в областта на биомедицинския анализ на изображенията, компютърните науки за живота и невронауките. Основните му дейности включват мултидисциплинарни научни изследвания, образователни проекти, развитие на STEM практики и събития/казуси за образователни институции, насърчаване на иновациите и предприемаческата култура и подкрепа за научния напредък на своите членове. Този документ изследва развитието и извършва оценка на устойчивостта на базираната на изследвания предприемаческа екосистема, използвайки данни от реалния свят. Анализът, приложен чрез анализ на социалните мрежи (SNA), разкрива

ролите, дейностите и взаимоотношенията на участниците. Използвайки индикатори за ключови точки (KPI) и мрежови модели, това изследване може да подкрепи решенията за необходими промени в екосистемата. Резултатите могат да бъдат приложени за текуща оценка, както и за прогнозна оценка на устойчивостта. Методологията може да бъде прехвърлена за по-нататъшни изследвания на други регионални екосистеми.

6. Feasibility of Haralick's Texture Features for the Classification of Chromogenic In-situ Hybridization Images

Pavlov, S., Momcheva, G., Burlakova, P., Atanasov, S., Stoyanov, D., Ivanov, M., Tonchev, A.

Proceedings of the International Conference on Biomedical Innovations and Applications, BIA 2020, pp. 65–68, 9244282

Тази статия представя пилотни доказателства за полезността на текстурните характеристики от втори ред за количествения анализ и класификацията на изображения с висока разделителна способност от цели тъканни срези, оцветени с хромогенна *in situ* хибридизация в рамките на образни експерименти с висока производителност. Предизвикателството е, че в момента златният стандарт за класифициране на генната експресия в такива изображения е експертна оценка. Идеята на изследователския екип е да използва различни подходи при анализа на тези изображения, които ще бъдат използвани за структурна сегментация и функционален анализ в генната експресия. Статията представя такава перспективна идея за избор на набор от текстурни характеристики, които ще бъдат използвани за неконтролирана

класификация. В нашия експеримент естественото групиране на образци от изображения (плочки) в зависимост от локалните свойства на текстурата им беше изследвано в процедура за класифициране без надзор. Размерността на характеристиките е сведена до две измерения чрез компонентен анализ, и след това плочките бяха класифицирани чрез групиране на размити с-средни. Общото заключение от този експеримент е, че текстурните характеристики по Haralick са обещаващ избор за класифициране и анализ на изображения от хромогенна *in situ* хибридизация. Редукцията на размерността чрез компонентния анализ генерира малко по-„разбираеми“ от гледна точка на анотатора класове.

7. Gabor Features for the Classification and Evaluation of Chromogenic In-Situ Hybridization Images

Pavlov, S., Momcheva, G., Burlakova, P., Atanasov, S., Stoyanov, D., Ivanov, M., Tonchev, A.

Lecture Notes in Networks and Systems, 2022, 374 LNNS, pp. 375–383

Резюме. Високопроизводителната хромогенна *in situ* хибридизация (CISH) е микроскопска техника, която разкрива пространственото разпределение на генната експресия в животински клетки и тъкани чрез лесно откриваема оцветена утайка. “Златният стандарт” за класифициране на CISH-оцветени тъкани включва качествено оценяване от експерт в областта. Този метод е предубеден, страда от ниска възпроизводимост и намалява ефективността на експериментите с висока производителност. Малко количествени методи за анализ разрешават тези проблеми, но предложените методи са чувствителни експериментални условия или изискват експертна настройка на множество параметри. Идеята на нашия изследователски екип е да извлече текстурна

информация от CISH-изображения които ще бъдат използвани за генериране на пространство за характеристики за семантично сегментиране и функционално анализ на генната експресия. В настоящата ни работа ние изследваме идеята от неконтролирана класификация въз основа на характеристики, генерирани чрез енергийни филтри на Gabor. Тъканта беше разделена на припокриващи се плочки от 150 μm и обработена с Gabor филтърна банка (5 дължини на вълните, 16 посоки, честотна лента 1.4). Резултатите за 16 посоките на всяка дължина на вълната бяха комбинирани чрез максимална суперпозиция ведно изображение и средната стойност на сивото, стандартното отклонение и ентропията бяха измерени. След подходящо намаляване на размерите, плочките бяха класифицирани по размит алгоритъм за C-means. Четирима експерти без предварителни познания за класификацията резултатите оценяват силата и модела на генна експресия на набор от произволно избрани плочки и независимо всеки клас в оригиналния цял слайд изображения. Използвано е сравнение между оценките по скалата на класа и по скалата на плочката за оценка на полезността на избраните функции.

8. Text and source readability-A step to cognition

Momcheva, G., Spasova, V., Ivanova, A., Zhelyazkov, M.

AIP Conference Proceedings, 2021, 2333, 070013

Статията има за цел да изследва връзката между четимостта на текстове от инструкции (текстове на задания, техническа документация, текстове на закони/наредби) и съответния изходен код (решение на задача) от конкретен образователен ресурс за програмиране по темата с невронни/положителни настроения „документи за самоличност“. Статията разглежда два вида

изследвания: едното е на прилагане на софтуерни показатели и второто е от емпирично изследване на разбирането на текст за целева група от 15-годишни ученици (върху текстове и програмен изходен код). Резултатите са силна основа за продължаване на изследванията, за да се получи обективна методология за четливост за хората на тази възраст. Това има както изследователска, така и бизнес добавена стойност на издателските компании като обективни препоръки за авторите на учебниците, а също и за учителите по компютърни науки и езици за общи преподавателски практики.

9. Application of process mining approach to the developmental process of the roundworm *C. elegans*

Chervenkov, T., Pavlov, S., Marinov, D., Hristov, H., Momcheva, G.

Proceedings of the International Conference on Biomedical Innovations and Applications, BIA 2021, pp. 51–53

Извличането на процеси е аналитичен подход, който произтича от и се сближава с науката за данните и моделирането на процеси. Първоначално създаден за подпомагане на управлението на бизнес процесите, обаче подходът за копаене на процеси е универсален и приложим в други области. Известно е, че извличането на знания от процеси споделят прилики с тези, използвани в биоинформатиката и че нововъзникващата дисциплина за извличане на знания от процеси може да се възползва от прилагането на техники, разработени в компютърната биология. Тук обаче демонстрираме обратното: че извличането на знания от процеси може да се приложи за изучаване на биологични процеси. Тъй като извличането на знания от процеси работи върху регистрационни файлове на събития, за да се анализира определен биологичен процес, е необходимо информацията за

поредица от биологични събития да се трансформира в регистър на събития. За това изследване ние приложихме техники за извличане на знания от процеси към набор от данни за развитието кръглия червей *C. elegans*. Данните за експресията на едноклетъчния темпорален ген бяха трансформирани в регистър на събитията и анализирани с инструменти за извличане на данни от процеси.

10. Survey of Information Technology Undergraduate Degree Programs in Canada

Marinova, R., Momcheva, G.

2019 IEEE Canadian Conference of Electrical and Computer Engineering, CCECE 2019, 8861715

Проучване на студенти по информационни технологии (ИТ). са представени образователни програми в Канада. ИТ бакалавър степени не са били много популярни в Канада, въпреки че те съществуват от няколко години в други части на света, включително Съединените щати (САЩ) и Европейския съюз (ЕС). Това проучване разглежда текущото състояние на студентите по ИТ дипломни програми и техните учебни планове.

11. Vendor Cybersecurity Risk Assessment in an Autonomous Mobility Ecosystem

Tzoneva, A., Momcheva, G., Stoyanov, B.

10th International Scientific Conference on Computer Science (COMSCI), 2022, DOI: [10.1109/COMSCI55378.2022.9912588](https://doi.org/10.1109/COMSCI55378.2022.9912588)

Оценката на риска за киберсигурността на доставчика е от решаващо значение за инфраструктурата на интелигентния град и устойчивостта на екосистемата за автономна мобилност. Липсата на ангажираност в политиките за киберсигурност и внедряването на процеси от компаниите от ниво, предоставящи хардуер или услуги на OEM производители в рамките на тази екосистема, представлява значителен риск не само за отделните компании, но и за екосистемата като цяло. Предложеният количествен метод за оценка на риска за киберсигурността позволява на доставчиците да имат видимост за финансовия риск, свързан с потенциални заплахи, и следователно да разпределят адекватни ресурси за киберсигурността. Той улеснява по-бързото прилагане на защитни мерки и предоставя полезен инструмент в процеса на избор на доставчик. Документът се фокусира върху оценката на риска за киберсигурността като критична част от цялостната мисия на компанията за създаване на устойчива структура за поддържане на здравето на киберсигурността. Сложният риск за киберсигурността и въздействието върху операциите на компанията като резултати от този количествен анализ представляват уникална възможност за стратегическо планиране и вземане на информирани решения за придобиване на реномирана позиция в устойчива екосистема. Този метод предоставя дървета на атаки и присвоява рисков фактор на всеки доставчик, като по този начин предлага конкурентно предимство и поглед върху картата на риска на веригата за доставки. Това е иновативен начин за разглеждане на позицията на киберсигурността на доставчиците. Чрез подбор на уникални специфични за индустрията параметри и модулен подход, този модел за оценка на риска може да се използва като инструмент за навигация в базата за доставки и предотвратяване на значителни финансови разходи. Той генерира синергии в екосистемата на свързаните превозни средства, водещи до безопасна и устойчива икономика.

12. Voronoi Diagrams and Perlin Noise for Simulation of Irregular Artefacts in Microscope Scans

Alreni, A., Momcheva, G., Pavlov, S.

9th International Conference on Bioimaging, Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - (Volume 2), 2022, DOI: 10.5220/0010833 000003123

Артефактите са често срещано явление в микроскопични изображения и сканирания, използвани в научните изследвания в областта на биологията. Артефактите могат да бъдат редовни и н'

е правилни и да произтичат от различни източници: изкривявания на осветителното поле, чужди частици в осветяването и оптичния път, грешки, нередности по време на фазите на обработка и оцветяване и т.н. Докато няколко изчислителни подхода за справяне с шарени съществуват изкривявания, няма универсален, ефикасен, надежден и лесен метод за отстраняване на неправилни артефакти.

Това поставя учените в областта на живота в тромави затруднения, губи ценно време и може да промени резултатите от анализа. В тази статия авторите очертават систематичен начин за въвеждане на синтетични неправилни артефакти в микроскопични сканирания чрез шум на Перлин и диаграми на Вороной. Причината зад такава задача е да се създадат двойки от „успешни“ и произведени „неуспешни“ двойници на изображения, които да се използват като двойки за обучение в изкуствен невронна мрежа, настроена за премахване на артефакти. В момента описаният метод работи само за изображения в сива скала.

13. Advanced Image Processing Techniques for the Detection and Monitoring of TV Datacenter System

Mohamed Fakhreddine; Teodora Bakardjieva; Galina Momcheva

TW център за данни е основното съоръжение за съхранение на данни. Телевизионните центрове за данни избухнаха в популярност през последното десетилетие. С нарастването на размера и възможностите на центровете нараства и сложността на тяхното управление. Информационната и функционална мрежа са застрашени от взривяване, с потенциално катастрофални последици. В резултат на това са необходими решения, които могат да вървят в крак с нарастването на доставката на данни, както и с размерите за съхранение на центрове за данни. Основната цел зад това изследване е да се предложи цялостна и усъвършенствана система за приспособяване към разширяването на центровете за данни, за подобряване на доставката на данни в центровете за данни и за откриване на грешки и проблеми в телевизионните центрове за данни. Освен това системата ще извършва и усъвършенствано и надеждно наблюдение на телевизионни центрове за данни. Значението на изследването се основава на предложените алгоритми, тъй като тези алгоритми са напълно нови, както и с висока производителност. В допълнение, тези алгоритми включват нови технологии за обработка на изображения. Резултатите, получени с помощта на алгоритмите, показаха получен процент на откриване от 83,33%, което може да се счита за показател за висока производителност на предложената. Освен това, според резултатите, минималното време за замяна е нула секунди, което означава, че замяната е извършена автоматично след откриване на повредата във видеото. Резултатите също така показаха, че за всички видеоклипове откритият брой кадри бяха успешно заменени. Промяната показва 100% ефективност за всяко видео.

14. Multi-Activation Dendritic Neural Network (MA-DNN) Working

Example of Dendritic-Based Artificial Neural Network

Konstantin Tomov, Galina Momcheva

През годините невронните мрежи са базирани на перцептронния модел на изкуствения неврон. Опитите да се отклоните от него са малко или никакви. Перцептронът просто работи и това е обезсърчило изследванията около други невронни модели. Нови открития подчертават значението на дендритите в неврона, но моделът на перцептрон не ги включва. Това ни води до целта на статията, която е да представи и тества различни модели на изкуствени неврони, които използват дендрити за създаване на изкуствен неврон, който по-добре представя биологичния неврон. Авторите предлагат два модела. Едната е направена с цел тестване на идеята за дендритния неврон. Отличителната черта на втория модел е, че той изпълнява активиращи функции след своите дендрити. Резултатите от втория модел показват, че той се представя толкова добре или дори по-добре от модела на перцептрон.

15. Utilizing Autoencoders for Analysis and Classification of Microscopic In Situ Hybridization Images

Aleksandar A. Yanev, Galina D. Momcheva, Stoyan P. Pavlov

Понастоящем анализът на микроскопични изображения от in situ хибридизация (ISH) се извършва ръчно от експерти. Прецизната и обективна оценка и класификация на такива микроскопски изображения може да улесни работата на експертите и да разкрие допълнителни закономерности при този вид данни. В тази работа ние предлагаме работен процес за дълбоко учене за

откриване и класифициране на области от микроскопични изображения със сходни нива на генна експресия. Анализът на данните се извършва чрез използване на тип ANN – Deep Learning Autoencoders – подходящи за неконтролирано обучение. Производителността на модела е оптимизирана чрез балансиране на дължината и сложността на латентните слоеве и фина настройка на хиперпараметрите. Резултатите се валидират чрез адаптиране на показателя за средноквадратична грешка (MSE) и сравнение с оценката на експерта. За обобщаване и визуализиране на резултатите се използва реконструкция и цветно кодиране на класовете върху цялостните микроскопски изображения.