

УДК [004.8]

СЕРГІЙ МІСОЧЕНКО, КАРИНА СЕЛІВАНОВА

МЕТОДИ ОПТИМІЗАЦІЇ ТА ТРАНСФЕРНЕ НАВЧАННЯ ДЛЯ ПОКРАЩЕННЯ ВИЯВЛЕННЯ ПУХЛИН МОЗКУ НА ОСНОВІ МРТ-ЗОБРАЖЕНЬ

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, Serhii.misochenko@nure.ua, karina.selivanova@nure.ua, hnietniev.konstantin@ill.kpi.ua

Анотація. Це дослідження аналізує методи оптимізації навчання згорткових нейронних мереж, особливості використання малих наборів даних і застосування методів Transfer Learning для підвищення точності детектування пухлин мозку на КТ-зображеннях. Результати навчання цих моделей порівнювалися між собою на тестовому наборі даних з зображеннями пухлин мозку на КТ-знімках.

Ключові слова: охорона здоров'я, медичні інформаційні системи, аугментація, згорткова нейронна мережа, малі набори даних, оптимізація параметрів, розпізнавання образів, трансферне навчання.

Abstract. This research analyzes methods for optimizing the training of convolutional neural networks, the features of using small datasets, and the use of Transfer Learning techniques to improve image detection accuracy. The results of training these models were compared with each other on a test dataset for detecting brain tumor in CT images.

Keywords: healthcare, medical information technologies, augmentation, convolutional neural network, small dataset, optimization parameters, pattern recognition, transfer learning.

DOI: 10.31649/1681-7893-2024-48-2-196-204

I. ВСТУП

Поширеність онкологічних захворювань залишається досить високою, оскільки щорічно діагностуються десятки мільйонів випадків різних видів раку, що призводить до мільйонів смертей у всьому світі. За даними Міжнародного агентства з вивчення раку (IARC) за 2020 рік, тенденції щодо захворюваності на рак та смертності залишаються тривожними в багатьох країнах.

З підвищенням якості та доступності медичної діагностики, а також зі збільшенням попиту на неї, використання нейронних мереж для аналізу діагностичних зображень стає дедалі важливішим. Глибоке навчання (DL) пропонує автоматичне вивчення особливостей медичних зображень за допомогою нейронних мереж. Його застосування охоплює класифікацію, сегментацію та виявлення онкологічного ураження. DL успішно використовується в діагностиці раку, надаючи значні переваги у обробці складних даних та підвищенні точності прогнозів.

З усіма цими перевагами методи глибинного навчання продовжують привертати увагу дослідників у сфері діагностики раку через автоматичний аналіз медичних зображень, розробляючи нові методи та підходи для покращення результатів. [1].

Наразі широке впровадження нейронних мереж стикається з такими викликами: [1]:

- Недостатня кількість високоякісних наборів даних, особливо для рідкісних видів раку.
- Недостатня прозорість моделей, що ускладнює їхнє розуміння для клініцистів.
- Слабка здатність узагальнювати нові набори даних, отримані з іншого обладнання або з іншими параметрами зображення.
- Брак високопродуктивних моделей для мультимодальних зображень.

Таким чином, нейронні мережі стикаються з низкою проблем, які потрібно вирішити для їх повного та успішного впровадження у клінічну практику. Наразі активно пропонуються нові методи модифікацій, оновлень та комбінацій як готових моделей нейронних мереж, так і методів покращення їх наборів даних.

Ц.ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ

Лулу Ван у своїй роботі [2] описує основні принципи використання нейронних мереж для діагностики раку легенів, а також розглядає результати застосування моделей, таких як SVM, U-Net, AWEU-Net, DenseNet201 та багатьох інших для цієї мети. Найкращим прикладом є модель Inception_ResNet_V2 для аналізу КТ-знімків, яка досягла максимальної точності 99,7%. Крім того, Жун-Хо Лін і колеги провели дослідження, яке було спрямоване на розробку архітектури моделі для діагностики раку молочної залози, зосереджуючись на підвищенні точності та зменшенні кількості хибних діагнозів. Було проведено тестування AlexNet, ResNet101 та InceptionV3 для аналізу мамограм з метою виявлення раку. Моделі були модифіковані шляхом тонкого налаштування і заміни різних оптимізаторів. Після модифікації моделі показали максимальну точність 81,16%, 85,51% і 91,3% відповідно. Вони також проаналізували мережу зворотного розповсюдження (BPM) і метод опорних векторів (SVM), які показали точність 74,63% і 91,6% відповідно. [3].

Існує багато порівнянь різних моделей для діагностики пухлин. Наприклад, у статті [4] автори порівнюють VGG16, VGG19 та ResNet50. Їх порівнюють на тому ж самому наборі даних, який був обраний для цього дослідження. Було проведено покращення результатів навчання за допомогою методу CLAHE для зміни контрастності зображення, таким чином збільшуючи початковий набір даних новими зображеннями. У цьому дослідженні модель VGG16 показала найкращу та найбільш стабільну продуктивність. Деякі останні статті пропонують інші підходи для підвищення точності моделей. Один із таких підходів — створення алгоритмів, що покращують існуючі моделі, таких як «покращений алгоритм морських хижаків» (IMPA), який використовується для знаходження найкращих значень гіперпараметрів архітектури ResNet50. У результаті створена модель IMPA-ResNet50 показує точність 98,32% у порівнянні з 95,95%, які показала класична ResNet50 [5].

Іншим напрямом покращення продуктивності нейронних мереж є вдосконалення самих наборів даних. Евгін Гочері у своїй статті [6] пропонує ідеї щодо покращення наборів даних, створених з медичних зображень. Зокрема, для медичних МРТ-знімків головного мозку він пише наступне: Найефективнішим методом аугментації МРТ-зображень головного мозку для класифікації випадків гліом високого (HGG) та низького (LGG) ступеня є аугментація шляхом комбінації обертання зі зсувом і трансляцією [6].

Одна з найпоширеніших проблем з біомедичними наборами даних зображень – це недостатня кількість навчальних зразків. Для якісного навчання нейронних мереж потрібні дуже великі та різноманітні набори даних. Одним із найбільш поширених емпіричних правил для машинного навчання в комп'ютерному зорі є «правило 1000» – для ефективного навчання моделі набір даних повинен містити щонайменше 1000 зразків зображень для кожної категорії. Це правило було встановлено Пітом Уорденом, який проаналізував записи з конкурсу класифікації ImageNet, де набір даних містив 1000 категорій, кожна з яких мала трохи менше 1000 зображень у класі. Набір даних був достатнім для навчання перших поколінь класифікаторів зображень, таких як AlexNet, тому автор зробив висновок, що приблизно 1000 зображень є хорошою базовою лінією для алгоритмів комп'ютерного зору. [10].

Оскільки частиною мети є також аналіз найбільш ймовірних проблем, з якими дослідник може стикнутися під час своєї роботи, було вирішено використовувати достатньо малий, але досить популярний набір даних для навчальних цілей: МРТ-зображення мозку для виявлення пухлин мозку. Він містить 253 зображення МРТ-сканування головного мозку, з яких 155 є сканами мозку з пухлинами, а 98 – без них. Для вирішення проблеми нестачі даних були використані методи аугментації даних для генерації відсутніх зразків зображень, необхідних для виконання цього правила. [11].

Для цього навчального завдання всі зображення були стандартизовані до єдиного формату та розміру 150x150x3. Найпростіші методи аугментації зображень включають дзеркальне відображення, обертання та зсув зображень у навчальному наборі [18-27].

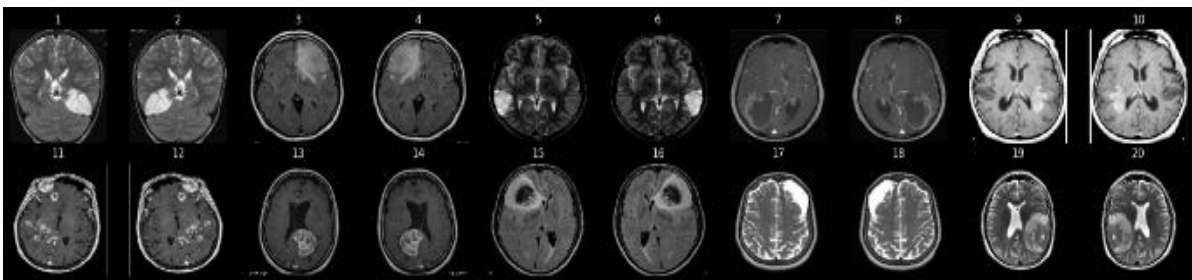


Рисунок 1 – Зображення після зменшення до 150x150 та аугментації (дзеркальне відображення)

III. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ТА ОБГОВОРЕННЯ

Найпопулярнішим типом архітектури нейронних мереж, що використовуються в обробці зображень, є згорткові нейронні мережі (CNN). Зазвичай CNN складається з трьох шарів: згорткового, прямого та повністю з'єднаного. Згортковий шар є основним структурним елементом CNN. Він несе основне навантаження обчислень мережі. Якщо маємо вхідні дані розміру $W \times W \times D$ та D_{out} — кількість ядер зі просторовим розміром F зі кроком S та відстанню P , то розмір вихідного значення можна визначити за наступною формулою:

$$W_{out} = \frac{W - F + 2P}{S} + 1 \quad (1)$$

Модель CNN містить 4 прихованих згорткових шари та 3,223,325 навчених параметрів. Код оригінальної моделі наведено на малюнку нижче:

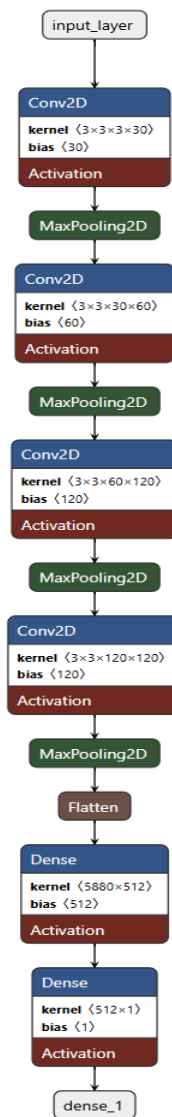


Рисунок 2 – Архітектура моделі

Точність моделі визначається за наступною формулою:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

БІОМЕДИЧНІ ОПТИКО-ЕЛЕКТРОННІ СИСТЕМИ ТА ПРИЛАДИ

Формула бінарної крос-ентропії використовується для оцінки кількості негативних результатів:

$$\frac{1}{N} \sum_{i=1}^N y_i * \log(p(y_i)) + (1 - y_i) * \log(1 - p(y_i)) \quad (3)$$

де y_i представляє фактичний клас;

$\log(p(y_i))$ - це ймовірність цього класу;

$p(y_i)$ - це ймовірність одиниці; $1 - p(y_i)$ - це ймовірність нуля.

Точність цієї моделі на тестовому наборі становить **93.0%**

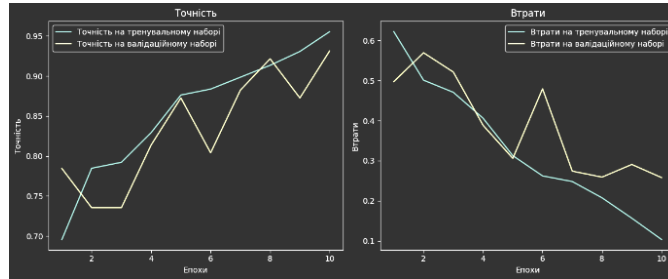


Рисунок 3 – Результат навчання моделі

Після навчання моделі було вирішено протестувати вплив оптимізаторів на якість навчання моделі. Порівнювалися оптимізатори RMSprop та Adamax. Точність моделей з цими оптимізаторами становить 90% та 87% відповідно. На рис.4-5 наведено результати навчання моделі за допомогою цих оптимізаторів [12-16]:

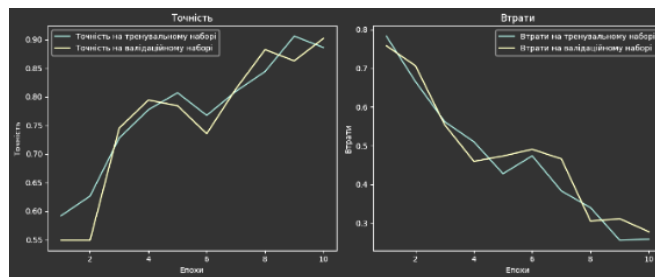


Рисунок 4 – Навчання моделі з використанням RMSprop

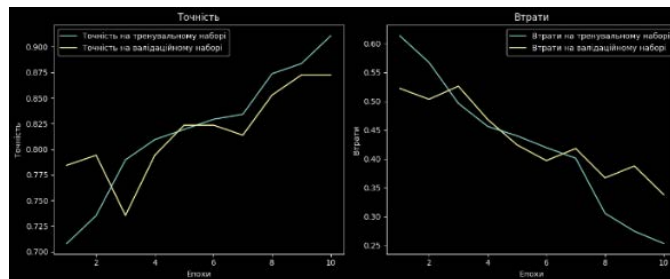


Рисунок 5 – Навчання моделі з використанням Adamax

З отриманих результатів можна побачити, що оптимізатори сімейства Adam надають гарні дані. Додаємо регуляризацію $l_1(0.01)$ та функцію для відключення деяких випадково обраних нейронів (DropOut) зі значенням (0.3):

БІОМЕДИЧНІ ОПТИКО-ЕЛЕКТРОННІ СИСТЕМИ ТА ПРИЛАДИ

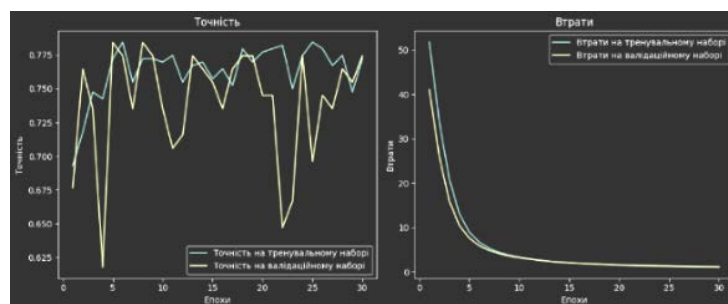


Рисунок 6 – Навчання моделі з регуляризацією 11

Після зміни 11 на 12 з тим самим значенням (0.01). Точність на тестовому наборі становить 91.0%.

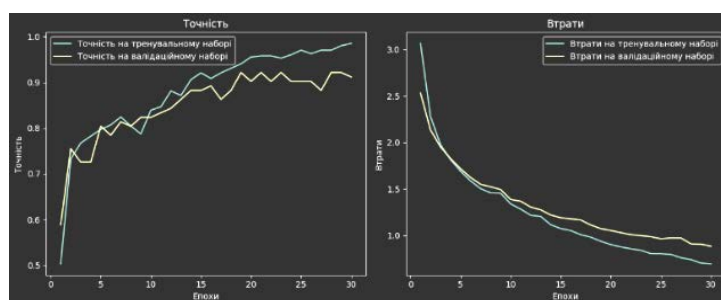


Рисунок 7 – Навчання моделі з регуляризацією 12

Результати використання регуляризації 12 є кращими. Щодо функцій активації, раніше використовувалася функція relu , а тепер необхідно порівняти з tanh та selu . Для tanh точність становить 84%, а для selu – 90%. Результати навчання наведені на рис.8-9:

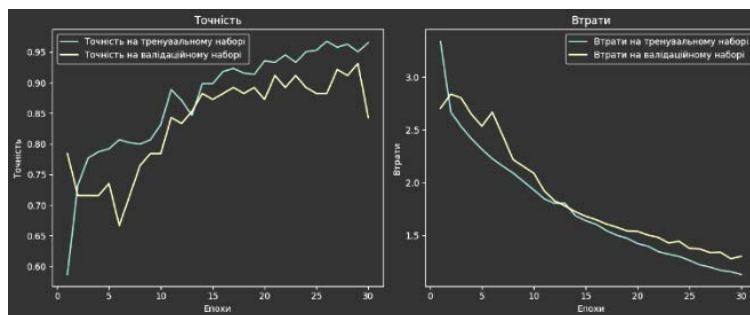


Рисунок 8 – Навчання моделі з функцією активації tanh

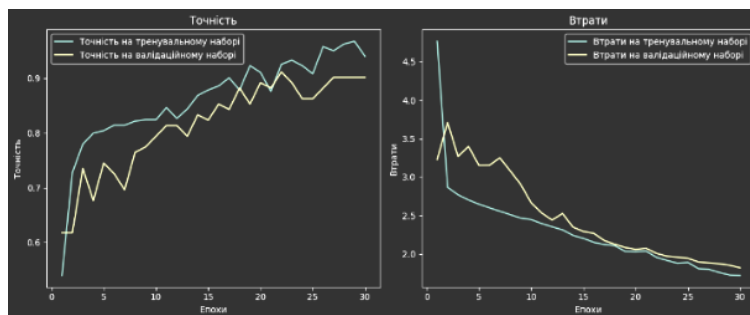


Рисунок 9 – Навчання моделі з функцією активації selu .

БІОМЕДИЧНІ ОПТИКО-ЕЛЕКТРОННІ СИСТЕМИ ТА ПРИЛАДИ

Як видно з отриманих результатів, найкраще зростання точності та мінімізація втрат були досягнуті з функцією активації *relu*. Архітектура остаточної моделі після оптимізації параметрів наведено на рис. 10.

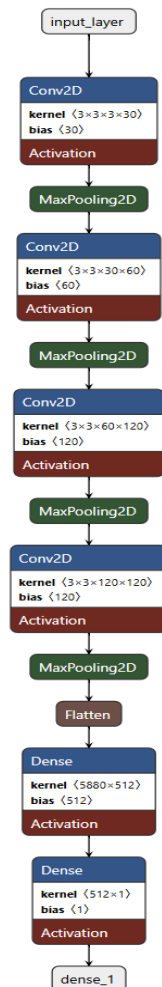


Рисунок 10 – Модифікована архітектура моделі

Таким чином, найоптимальніша конфігурація моделі в цьому завданні:

- оптимізатор - *Adam*
- регуляризація - $l_2(0.001)$
- функція активації – *relu*
- *DropOut* після прихованих шарів зі значенням 0.3.

Як згадувалося раніше, популярна архітектура VGG16 була обрана як попередньо навчена модель. Як модифікацію нейронної мережі було додано Flatten шар з 512 нейронами, за яким слідує вихідний шар з одним нейроном. Ця модель досягає точності 98% [16-19]:

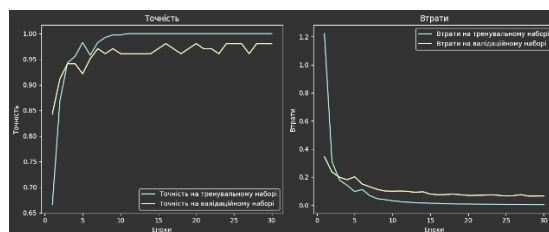


Рисунок 11 – Навчання моделі VGG16

БІОМЕДИЧНІ ОПТИКО-ЕЛЕКТРОННІ СИСТЕМИ ТА ПРИЛАДИ

Результати навчання демонструють точність 98% для набору з 512 зображень. Щоб перевірити результати та вирішити проблему перенавчання, набір даних був подвоєний за рахунок додаткової аугментації зображень..

Тепер у наборі даних 1012 зображень, але модель все ще показує 96.0% за наступною схемою навчання:

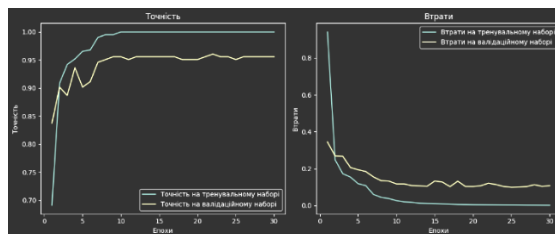


Рисунок 12 – Навчання моделі VGG16 після аугментації

Згідно з результатами експерименту, трансферне навчання виявилось потужним інструментом, здатним значно покращити точність розпізнавання зображень і прискорити навчання моделі, оскільки потрібно повністю навчити лише шари, додані до попередньо навченої моделі. У цьому випадку модель VGG16 показала значне збільшення точності як при подвоєнні розміру набору даних, так і при його потроєнні. Тому в цьому випадку немає потреби розглядати тонке налаштування цієї моделі або використовувати глибші мережі, такі як ResNet-50 або EfficientNetB3+, оскільки збільшення точності не буде значним (~2-3%), як показують нещодавні наукові дослідження, що вивчають використання трансферного навчання для класифікації зображень. [12-17].

ВИСНОВКИ

Під час цього дослідження було проведено аналітичний огляд спеціалізованих літературних джерел, пов'язаних з діагностикою пухлин за допомогою нейронних мереж для аналізу біомедичних зображень. Був обраний набір даних МРТ-зображень мозку для виявлення пухлин мозку, що складається з 253 МРТ-зображень мозку, з яких 155 є сканами з пухлинами, а 98 - без них. Аугментація була виконана для збільшення розміру набору даних і його нормалізації.

Було проаналізовано вплив використання різних методів оптимізації гіперпараметрів на навчання моделі для оцінки їхнього ефекту на покращення точності моделі. Крім того, у дослідженні було продемонстровано доцільність використання трансферного навчання для навчання на малих наборах даних, коли немає специфічних вимог щодо розміру моделі. У результаті модель мала оптимізовані параметри, досягнувши точності 91%, що вважається хорошим показником для моделі з такою кількістю ознак.

Результати цієї моделі були порівняні з отриманими за допомогою трансферного навчання, зокрема з попередньо навченою моделлю VGG16, яка показала точність 98% на неаугментованому наборі даних і 96% на аугментованому наборі. Зниження точності було пов'язано з проблемою перенавчання на малому та однорідному наборі даних.

Як показали експериментальні результати, у випадках недостатньої кількості навчальних зразків, найраціональнішим підходом для покращення точності розпізнавання готової моделі є використання трансферного навчання з попередньо навченими моделями, коли немає вимог щодо розміру моделі. Однак цей підхід може призвести до проблеми перенавчання моделі. Дослідження також чітко продемонструвало доцільність аугментації даних у будь-якому сценарії навчання моделі, оскільки цей метод значно покращує точність для моделей і частково вирішує проблему перенавчання, особливо для великих наборів даних.

REFERENCES

1. Deep learning for medical image-based cancer diagnosis / X. Jiang et al. *Cancers*. 2023. Т. 15, № 14. С. 3608. URL: <https://doi.org/10.3390/cancers15143608>
2. Wang L. Deep learning techniques to diagnose lung cancer. *Cancers*. 2022. Т. 14, № 22. С. 5569. URL: <https://doi.org/10.3390/cancers14225569>

3. Application of deep learning to construct breast cancer diagnosis model / R.-H. Lin et al. *Applied sciences*. 2022. T. 12, № 4. C. 1957. URL: <https://doi.org/10.3390/app12041957>
4. Kamal Kamal, Hamid EZ-ZAHRAOUY. A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images, 27 April 2023, PREPRINT (Version 1) available at Research Square URL: <https://doi.org/10.21203/rs.3.rs-2863523/v1>
5. Houssein, E.H., Emam, M.M. & Ali, A.A. An optimised deep learning architecture for breast cancer diagnosis based on improved marine predators algorithm. *Neural Comput & Applic* 34, 18015-18033 (2022). URL: <https://doi.org/10.1007/s00521-022-07445-5>
6. Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. *Artificial intelligence review*. 2023. URL: <https://doi.org/10.1007/s10462-023-10453-z>
7. Pete Wander How many images do you need to train a neural network?, 2017, URL: <https://petewarden.com/2017/12/14/how-many-images-do-you-need-to-train-a-neural-network/>
8. Houda Bichri, Adil Chergui, Mustapha Hain, Image Classification with Transfer Learning Using a Custom Dataset: A Comparative Study, *Procedia Computer Science*, Volume 220, 2023, Pages 48-54, ISSN 1877-0509, URL: <https://doi.org/10.1016/j.procs.2023.03.009>.
9. Rafiq M, Rafiq G, Agyeman R, Choi GS, Jin S-I. Scene Classification for Sports Video Summarisation Using Transfer Learning. *Sensors*. 2020; 20(6):1702. URL: <https://doi.org/10.3390/s20061702>
10. Tymkovich, M. et al. (2021). Application of SOFA Framework for Physics-Based Simulation of Deformable Human Anatomy of Nasal Cavity. In: Jarm, T., Cvetkoska, A., Mahnič-Kalamiza, S., Miklavcic, D. (eds) 8th European Medical and Biological Engineering Conference. EMBEC 2020. IFMBE Proceedings, vol 80. Springer, Cham. https://doi.org/10.1007/978-3-030-64610-3_14
11. Selivanova, K. G., Avrunin, O. G., Tymkovich, M. Y., , T. V., et al., "3D visualization of human body internal structures surface during stereo-endoscopic operations using computer vision techniques," *Przegląd Elektrotechniczny*, 9 30 –33 (2021) DOI [10.15199/48.2021.09.06](https://doi.org/10.15199/48.2021.09.06)
12. Avrunin, O., Tymkovich, M., Semenets, V. and Piatyokop, V. "Computed tomography dataset analysis for stereotaxic neurosurgery navigation," *Proceedings of the International Conference on Advanced Optoelectronics and Lasers, CAOL*, 606-609 (2019).
13. Sokol, Y., Avrunin, O., Kolisnyk, K., & Zamiatin, P. (2020). Using medical imaging in disaster medicine. Paper presented at the 2020 IEEE 4th International Conference on Intelligent Energy and Power Systems, IEPS 2020 - Proceedings, 287-290. doi:10.1109/IEPS51250.2020.9263175
14. Avrunin, O., Kolisnyk, K., Nosova, Y., Tomashevskiy, R., & Shushliapina, N. (2020). Improving the methods for visualization of middle ear pathologies based on telemedicine services in remote treatment. Paper presented at the 2020 IEEE KhPI Week on Advanced Technology, KhPI Week 2020 - Conference Proceedings, 347-350. doi:10.1109/KhPIWeek51551.2020.9250090
15. K. Selivanova and O. Avrunin, "Method of Hand Movement Disorders Determination based on the Surgeon's Laparoscopic Video Recording," *2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2022, pp. 1-4, doi: 10.1109/KhPIWeek57572.2022.9916457.
16. Avrunin, O. G., Tymkovich, M. Y., Farouk Ismail Saed, H., Loburets, A. V., Krivoruchko, I. A., Smolarz, A., & Kalimoldayeva, S. (2019). Application of 3D printing technologies in building patient-specific training systems for computing planning in rhinology. *Information Technology in Medical Diagnostics II*, 1–8. <https://doi.org/10.1201/9780429057618-1>
17. Avrunin, O.G., Alkhorayef, M., Saied, H.F.I., and Tymkovich, M.Y., "The surgical navigation system with optical position determination technology and sources of errors," *Journal of Medical Imaging and Health Informatics*, 5(4), 689-696 (2015).
18. Wójcik Waldemar, Smolarz Andrzej (2017). *Information Technology in Medical Diagnostics*, July 11, 2017 by CRC Press, 210 Pages.
19. *Highly linear Microelectronic Sensors Signal Converters Based on Push-Pull Amplifier Circuits* / edited by Waldemar Wojcik and Sergii Pavlov, Monograph, (2022) NR 181, Lublin, Comitet Inzynierii Srodowiska PAN, 283 Pages. ISBN 978-83-63714-80-2
20. Pavlov Sergii, Avrunin Oleg, Hrushko Oleksandr, and etc. (2021). *System of three-dimensional human face images formation for plastic and reconstructive medicine // Teaching and subjects on bio-medical engineering Approaches and experiences from the BIOART-project* Peter Arras and David Luengo (Eds.), Corresponding authors, Peter Arras and David Luengo. Printed by Acco cv, Leuven (Belgium). - 22 P. ISBN: 978-94-641-4245-7.
21. Pavlov S.V., Avrunin O.G., etc. (2019). *Intellectual technologies in medical diagnosis, treatment and rehabilitation: monograph* / [S. In edited by S. Pavlov, O. Avrunin. - Vinnytsia: PP "TD "Edelweiss and K", 260 p. ISBN 978-617-7237-59-3.

22. Romanyuk, O., Zavalniuk, Y., Pavlov, S., etc. (2023). New surface reflectance model with the combination of two cubic functions usage, *Informatyka, Automatyka, Pomiarы w Gospodarce i Ochronie Srodowiska*, 13(3), pp. 101–10
23. Kukharchuk, Vasyl V., Sergii V. Pavlov, Volodymyr S. Holodiuk, Valery E. Kryvonosov, Krzysztof Skorupski, Assel Mussabekova, and Gaini Karnakova. (2022). "Information Conversion in Measuring Channels with Optoelectronic Sensors" *Sensors* 22, no. 1: 271. <https://doi.org/10.3390/s22010271>.
24. Vasyl V. Kukharchuk, Sergii V. Pavlov, Samoil Sh. Katsyv, and etc. (2021). Transient analysis in 1st order electrical circuits in violation of commutation laws", *Przegląd elektrotechniczny*, ISSN 0033-2097, R. 97 NR 9/2021, p. 26-29, doi:10.15199/48.2021.09.05.
25. Pavlov S.V, Petruk V.G., Kolesnik P.F. (2007). Photoplethysmographic technologies of the cardiovascular control: monography, Vinnitsa: Universum-Vinnitsa, 254 p.
26. Wójcik W, Mezhiievska I, Pavlov SV, Lewandowski T, Vlasenko OV, Maslovskiy V, Volosovych O, Kobylanska I, Moskovchuk O, Ovcharuk V, et al. (2023). Medical Fuzzy-Expert System for Assessment of the Degree of Anatomical Lesion of Coronary Arteries. *International Journal*.
27. Pavlov SV, Kozhemiako VP, Petruk VG, Kolesnik PF. (2007). Photoplethysmographic technologies of the cardiovascular control, Vinnitsa: Universum-Vinnitsa, 254 p.

Надійшла до редакції 25.09.2024 р.

МІСОЧЕНКО СЕРГІЙ – аспірант, кафедра біомедичної інженерії, Харківський національний університет радіоелектроніки, Україна, ***e-mail: serhii.misochenko@nure.ua***

СЕЛІВАНОВА КАРІНА – к.т.н., доцент кафедри біомедичної інженерії, Харківський національний університет радіоелектроніки, Україна, ***e-mail: karina.selivanova@nure.ua***

SERHII MISOCHENKO, KARINA SELIVANOVA

METHODS OF OPTIMIZATION AND TRANSFER LEARNING TO IMPROVE THE DETECTION OF BRAIN TUMORS BASED ON MRT-IMAGES

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine