DOI 10.31891/2307-5732-2024-337-3-32
УДК 004.89.032.26:[004.932.72'1:358.31-021.485

ОНИЩЕНКО СВІТЛАНА
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
https://orcid.org/0000-0002-6173-4361
e-mail: s07onyshchenko@gmail.com
ЛАКТІОНОВ ОЛЕКСАНДР
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
https://orcid.org/0000-0002-5230-524X
e-mail: laktionov.alexander@ukr.net
ГЛУШКО АЛІНА
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
https://orcid.org/0000-0002-4086-1513
e-mail: glushk.alina@gmail.com

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ РОЗПІЗНАВАННЯ ТЕРОРИСТИЧНИХ ТА ВОРОЖИХ ВІЙСЬКОВИХ ОБ'ЄКТІВ

В роботі продемонстровано процес автоматизації розпізнавання терористичних та ворожих військових об'єктів засобами згорткових нейронних мереж. Перевагами запропонованого підходу є універсальність щодо розпізнавання різних терористичних та ворожих військових об'єктів. Запропонована модель може бути використана для перенавчання на нових унікальних зображеннях з метою розширення об'єктів класифікації.

Ключові слова: нейронні мережі, глибоке навчання, класифікація об'єктів, терористичні об'єкти, військова техніка, штучний інтелект.

ONYSHCHENKO SVITLANA, LAKTIONOV OLEKSANDR, HLUSHKO ALINA
National University "Yuri Kondratyuk Poltava Polytechnic"

UTILIZING ARTIFICIAL INTELLIGENCE TO RECOGNIZE TERRORIST AND ENEMY MILITARY TARGETS

The development of artificial intelligence technologies and their application in solving military tasks is the main concept of scientific advancement in Ukraine. Despite the availability of existing solutions for classifying limited types of civilian objects, there is a need to develop new technologies to gain an advantage over the enemy.

Existing artificial intelligence models available in the public domain are oriented towards diagnosing restricted types of civilian objects. This limitation prevents the use of these models for analyzing military enemy equipment.

Automating the process of recognizing enemy military objects by constructing a deep learning classification model.

The process of building a classification model for enemy military equipment using deep neural networks is considered. The Sequential model type is chosen as the basis, which includes batch normalization layers added after each convolutional layer before further processing. Accelerated hyperparameter search is provided by the Random Search tool.

The quality of classification models was measured using classical accuracy metrics. The implementation of the proposed solutions is carried out using the Python programming language and the Python Imaging Library, TensorFlow, Keras Tuner, and Matplotlib libraries.

A convolutional neural network classification model has been constructed capable of classifying enemy military tanks and vehicles. The model demonstrates the best accuracy metrics on training/testing datasets of 0.6/0.55 with errors of 0.658/0.705, respectively.

A comparative analysis of classification models with different types of architectures showed the advantage of the model in terms of accuracy and error rate when using an architecture that includes batch normalization layers. The created classification model is recommended for further training on new data or for solving computer vision tasks.

Keywords: neural networks, deep learning, object classification, military equipment, artificial intelligence.

Постановка проблеми

За час повномасштабного вторгнення зріс попит на використання штучного інтелекту для автоматизації робототехнічних комплексів, дронів й іншої військової техніки. Про це свідчить пошукова система Google, рис. 1.

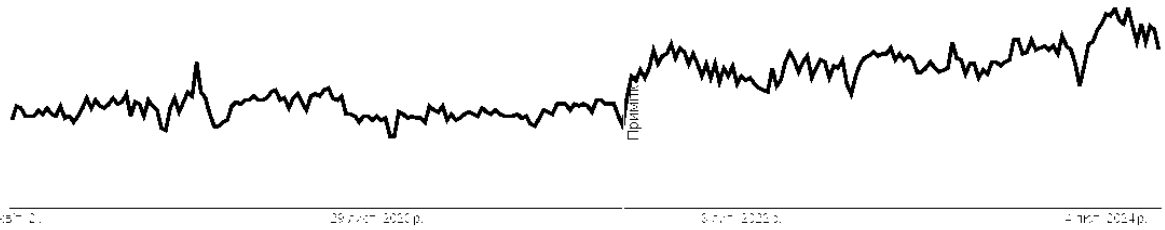

Рис. 1. Викликаний інтерес до пошукового терміну «computer vision» з 28 квітня 2019 року по 01 травня 2024 року у всьому світі, що побудовано за допомогою інструментів Google Trends

Як видно із рис. 1, викликаний інтерес до використання технологій штучного інтелекту за останні п'ять років у світі лише зростав. Збільшення інтересу до вказаних технологій спостерігається з січня 2022 року, де використання штучного інтелекту корелюється з повномасштабним вторгненням в Україну.

Процес побудови моделей комп'ютерного зору щодо виявлення терористичних об'єктів та ворожої техніки потребує створення моделей класифікації. Це може бути або бінарна, або багатокласова

класифікація, яка побудована відомим чином [1]. Створена модель класифікації використовується для вирішення задач комп'ютерного зору, зокрема виявлення скупчення людей тощо [2].

Запропонована систематизація існуючих рішень розглядалася у напрацюванні [3], де не передбачалося вивчення процесу побудови класифікаційних моделей й розглянуто лише теоретичні рекомендації щодо інструментів побудови моделей штучного інтелекту.

Обмеженнями існуючих підходів є використання загально доступних датасетів із вихідними зображеннями, котрі не містять зображень необхідних для виявлення ворожої техніки. Крім того, ефективне планування військових операцій потребує мінімального часу створення моделей штучного інтелекту, котрі надаються розвідкою.

Аналіз останніх джерел

У науковому доробку [4] систематизовано існуючі дослідження щодо комп'ютерного зору безпілотників з 2018 по 2023 роки, що підтверджує виявлену тенденцію зростання попиту на вказані технології. При цьому технології групування об'єктів використовуються у різних напрямах, наприклад для моніторингу асфальтового покриття [5]. Особливостями дослідження [5] є не лише створення моделей групування об'єктів, а й розробка програмного забезпечення для керування дронами щодо моніторингу дорожнього покриття.

Крім моніторингу дорожнього покриття моделі групування об'єктів використовуються для виявлення сміття з високовольтних ліній [6], де передбачається додаткова розробка апаратного забезпечення. Тобто, моделі штучного інтелекту здатні виявляти об'єкти залежно від завдань їх використання. Це може бути сміття, як у роботі [6] або дефекти на виготовленій продукції [7].

Виявлення дефектів продукції [7], опирається на унікальну методологію дослідження, де використовується модель ResNet. Перевагами удосконаленої моделі [7] є врахування оптимізації з використанням модуля ефективної уваги каналу квантування світла. За результатами експериментів модель демонструє точність 97,2%.

Досягнення високих показників точності також забезпечується техніками перетворення й балансування класів, котрі використовуються у поєднанні з пакетом CARET [8]. Аналогічними дослідженнями займалися у роботі [9], де вивчали методи поглибленого аналізу роботи з незбалансованими вибірками. Для цього здійснювався попередній аналіз даних з метою виявлення незбалансованих класів та видалення вказаних ознак для створення класифікатора. Запропонована авторами техніка [9] передбачає поліпшення ефективності моделей й мінімізує похибку. З точки зору дослідження [10] хибно класифіковані класи відносять до помилок, де пропонується шлях вирішення проблеми за допомогою систем числення.

У роботі [11] розглядаються питання оптимізації моделей штучного інтелекту, де доведення адекватності запропонованих рішень здійснюється аж на двадцяти класифікаторах за метриками точності та F1-міри. При цьому методологія дослідження не передбачає врахування шуму як у роботі [12], де виявлено залежність збільшення шуму при збільшенні обсягу вибірки.

Вирішення проблеми мінімізації шуму в напрацюванні [12] здійснено за допомогою агностичних методів, котрі не впливають на складність мережі й працюють з різними типами даних. Поліпшення якості нейронної мережі також залежить від типу досліджуваних вхідних даних, котрі використовуються.

Крім побудови моделей класифікації варто акцентувати увагу на їх практичному використанні. Так у роботі [13] розглянуто концепцію інтернету речей, котрий взаємодіє з різними пристроями підключеними до єдиної мережі у будь-якій точці простору. Саме тому необхідно забезпечити збереження побудованих моделей класифікації у формати залежно від апаратного забезпечення, зокрема Hierarchical Data Format version 5 [14]. Важливим аспектом створення моделей нейронних мереж є також врахування відповідності вхідних даних архітектурі нейронної мережі.

Стаття присвячена вирішенню питання створення моделей класифікації для їх використання у технологіях комп'ютерного зору щодо вирішення практичних задач розпізнавання терористичних об'єктів й ворожої техніки. Це дозволить ефективніше приймати рішення щодо тактик дій у разі виявлення відповідних об'єктів. Крім того, як зазначено у дослідженні [15], відомі моделі зорієнтовані на використання існуючих банків даних. Вони не містять необхідних взірців ворожої техніки, котрі розглядаються у цьому дослідженні.

Кожен розглянутий підхід [4-15] демонструє унікальні техніки поліпшення якості моделей штучного інтелекту, зокрема шляхом використання зовнішніх інструментів, котрі не впливають на архітектуру нейронної мережі. Тому виявлений науковий факт є не вирішеним у повному обсязі, зокрема для завдань класифікації терористичних об'єктів й ворожої військової техніки.

Метою роботи є автоматизація процесу розпізнавання терористичних й ворожих військових об'єктів шляхом побудови класифікаційної моделі глибокого навчання.

Завдання:

1. Створити класифікаційну модель штучного інтелекту для групування терористичних й ворожих військових об'єктів існуючими засобами глибокого навчання.

2. Провести експериментальну верифікацію запропонованих рішень і довести їх ефективність.

Виклад основного матеріалу

Формальна постановка завдання класифікації терористичних об'єктів та ворожої техніки. Дано датасет зображень, де x_i – зображення досліджуваного об'єкту, y_i – мітка класу до якого належить

досліджуваний об’єкт. Датасет диференційований на тренувальну й тестову вибірки об’єктів дослідження з дотриманням принципу стратифікації. Метою дослідження є створення функції f: X→Y, котра диференціює досліджувані зображення у відповідні класи. Для цього необхідно побудувати оптимальну модель класифікації об’єктів дослідження з точки зору досягнення компромісу між точністю на тренувальній і тестовій вибірках.

Етап дослідження (січень 2024 року – лютий 2024 року) передбачав створення класифікаційної моделі нейронної мережі. При побудові моделей класифікації розглядалися моделі типу Sequential, де процес побудови відбувався за наступними кроками, Крок 1 – Крок 5 [1]. Крок 1. Завантаження зображень із вихідного датасету, де процес передбачає визначення міток labels автоматичним чином за допомогою цілочисельних значень, де batch_size=32.

Крок 2. Створення архітектури нейронної мережі із заданими інтервалами значень гіперпараметрів. Це дозволяє мінімізувати час побудови моделі. Перший згортковий шар (Conv2D) включає фільтри розміром 3x3, функцію активації ReLU і валідну стратегію доповнення. На вхід першого шару подаються зображення визначені у кроці 1.

Крок 2.1. Створення шару нормалізації пакету (BatchNormalization), який додається після кожного згорткового шару для стабілізації і нормалізації вихідних значень шару перед їх подальшою обробкою.

Крок 2.2. Створення шару максимального зведення (MaxPooling2D), який дозволяє зменшити розмірність виходу з попередніх згорткових шарів.

Крок 2.3. Створення розгортки (Flatten).

Крок 2.4. Створення з’єднаного шару (Dense) з функцією активації ReLU.

Крок 2.5. Створення шару регуляризації (Dropout).

Крок 2.6. Створення вихідного шару (Dense) з функцією активації для завдання класифікації.

Крок 3. Компіляція моделі класифікації терористичних об’єктів й ворожої військової техніки з оптимізатором adam та метрикою accuracy для визначення ефективності моделі і функцією втрат для оцінки похибки моделі під час навчання.

Крок 4. Створення сітки пошуку гіперпараметрів та визначення оптимальної моделі, де вказується максимальна кількість ітерацій max_trials.

Крок 5. Виведення результатів побудованої архітектури та графічної інтерпретації результатів дослідження.

Крок 6. Прийняття рішень, людиною котра приймає рішення щодо якості моделей класифікації терористичних об’єктів й ворожої техніки.

Обмеженнями дослідження є нормування зображень, де використовувався розмір 229x229. Для зменшення часу побудови моделі використовувалися обмеження гіперпараметрів шарів, де користувалися значеннями у діапазоні [1, 5] з кроком 1. У процесі налаштування пошуку гіперпараметрів, інструментом Random Search, використовувалися наступні налаштування як максимальна кількість спроб max_trials=5, де максимізувалася метрика визначення точності при кількості епох 5.

Програмна реалізація наведених рішень передбачала використання мови програмування Python й ряду бібліотек. Для вказаного етапу досліджень використовувалися бібліотеки tensorflow й Keras Tuner. За допомогою інструментів бібліотек будувалися моделі глибокого навчання й проводився пошук архітектури та визначалися оптимальні значення гіперпараметрів. Попередня обробка зображень здійснювалася засобами Python Imaging Library. Реалізація матриці плутанини здійснювалася бібліотекою metrix, а графічна інтерпретація отриманих результатів дослідження засобами matplotlib. Таким чином здійснюється побудова моделі класифікації терористичних об’єктів й ворожої техніки, де для порівняльного аналізу будувалася модель без використання шарів нормалізації пакету.

Однією з проблем побудови моделей штучного інтелекту є існування вихідних даних, зокрема терористичних та ворожих військових об’єктів. Вихідні зображення можна добути за допомогою відео з YouTube, сфотографувати власноруч або використати існуючі датасети. У якості вихідних даних використано датасет [16], котрий містить двадцять чотири терористичних та ворожих військових об’єктів. Кожен тип об’єкта нараховує різну кількість зображень, наприклад зенітні гармати – 11 зображень, а вантажівки – 2072 зображення. З метою уникнення незбалансованої вибірки для побудови моделі обрано теки із зображеннями наближено однакової кількості. Так, обрано теки з вантажівками й танками, що нараховують 2072 і 1597 зображень відповідно. Тобто вихідний досліджуваний датасет нараховував 3669 зображень, котрі диференційовані на навчальний і тестовий набори зображень. Навчальний датасет нараховував два класи по 959 зображень терористичних та ворожих військових об’єктів, а тестовий – два класи по 638 зображень.

Відповідно до методології дослідження побудова моделі класифікатора зображень терористичних та ворожих військових об’єктів здійснювалася за наведеними кроками з обмеженнями. У процесі налаштування пошуку гіперпараметрів, інструментом Random Search, використовувалися наступні налаштування як максимальна кількість спроб max_trials=5, де максимізувалася метрика визначення точності при кількості епох 5, табл.. 1.

Кожну з п’яти спроб пошуку оптимальної моделі згорткової мережі класифікації терористичних та ворожих військових об’єктів можна охарактеризувати з точки зору балансу між навчальною і тестовою вибірками й з точки зору похибки навчання.

Таблиця 1

Оптимальна архітектура досліджуваної згорткової мережі визначена засобами Keras Tuner на основі досліджуваного датасету та зазначених обмежень

№	Тип шару	Вихідна форма	Параметр
1	conv2d_3 (Conv2D)	(None, 297, 297, 4)	112
2	batch_normalization_3 (BatchNormalization)	(None, 297, 297, 4)	16
3	max_pooling2d_3 (MaxPooling2D)	(None, 148, 148, 4)	0
4	conv2d_4 (Conv2D)	(None, 146, 146, 5)	185
5	batch_normalization_4 (BatchNormalization)	(None, 146, 146, 5)	20
6	max_pooling2d_4 (MaxPooling2D)	(None, 73, 73, 5)	0
7	conv2d_5 (Conv2D)	(None, 71, 71, 2)	92
8	batch_normalization_5	(None, 71, 71, 2)	8
9	max_pooling2d_5 (MaxPooling2D)	(None, 35, 35, 2)	0
10	flatten_1 (Flatten)	(None, 2450)	0
11	dense_3 (Dense)	(None, 2)	4902
12	dropout_2 (Dropout)	(None, 2)	0
13	dense_4 (Dense)	(None, 4)	12
14	dropout_3 (Dropout)	(None, 4)	0
15	dense_5 (Dense)	(None, 1)	5

Під час першої спроби побудови моделі, при кількості епох 5, точність навчальної вибірки змінювалася з 0,54 до 0,6 у той час як тестова вибірка демонструвала діапазон значень 0,5-0,55. Похибки змінювалися у діапазоні від 0,689 до 0,658 для навчальної вибірки і від 0,693 до 0,705 – тестової. Компромісу між навчальною й тестовою підвибірками не спостерігалося, оскільки моделях був характерний процес перенавчання. Результати діагностики інших спроб подано у табл. 2.

Таблиця 2

Спроби пошуку оптимальної моделі класифікації терористичних та ворожих військових об'єктів засобами глибокого навчання

Спроба №2			
Навчальний датасет		Тестувальний датасет	
Точність	Похибка	Точність	Похибка
0,51	0,792	0,5	0,693
0,48	0,693	0,5	0,692
0,49	0,693	0,5	0,692
0,53	0,685	0,52	0,692
0,55	0,68	0,52	0,691
Спроба №3			
Навчальний датасет		Тестувальний датасет	
Точність	Похибка	Точність	Похибка
0,54	0,732	0,52	0,694
0,56	0,666	0,53	0,699
0,6	0,642	0,53	0,707
0,62	0,616	0,52	0,714
0,63	0,61	0,53	0,74
Спроба №4			
Навчальний датасет		Тестувальний датасет	
Точність	Похибка	Точність	Похибка
0,52	0,698	0,5	0,693
0,55	0,673	0,5	0,693
0,57	0,669	0,5	0,699
0,57	0,657	0,5	0,704
0,59	0,643	0,5	0,717
Спроба №5			
Навчальний датасет		Тестувальний датасет	
Точність	Похибка	Точність	Похибка
0,47	0,703	0,5	0,693
0,48	0,692	0,5	0,693
0,48	0,692	0,5	0,693
0,47	0,7	0,5	0,692
0,49	0,71	0,49	0,695

Результати точності і похибок, які продемонстрували моделі вказують на наступне. Показники точності навчального і тестового наборів протягом спроб 2, 3 зростали, де не досягався баланс між оцінками точності. Оцінки похибки корелювалися з показниками точності, де при збільшенні точності похибка зменшувалася. За результатами спроб 4, 5 точність навчального набору збільшувалася і зменшувалася, а тестового лишалася стабільною. Це протиріччя спостерігається і у масивах похибок.

У кінцевому результаті оптимальні показники точності демонструє модель, побудована під час спроби №1, де точність на навчальному тестовому наборах 0,6/0,55 при похибці 0,658/0,705. Отриманий результат точності пояснюється використанням неякісних зображень, котрі подаються на вхід мережі. Результати точності моделей можуть бути ліпшими за умови використання якісніших зображень та використання рекомендацій з роботи [8].

Для додаткової перевірки адекватності отриманих результатів проведено додаткове дослідження щодо побудови моделі класифікації, яке показало результати точності оптимальної моделі на рівні 0,48. Це пояснюється використанням архітектури мережі без шарів batch_normalization.

Перевагами запропонованої моделі є час витрачений на її навчання, що становить понад 1 годину 12 хвилин. За вказаний відрізок часу проведено 5 спроб по 5 епох щодо обрання оптимальної моделі серед 25 можливих. При використанні гіперпараметрів з іншим кроком, наприклад у діапазоні від 64 до 256 з кроком 64 час однієї епохи у спробі сягає понад 4 години.

З практичної точки зору створену модель можна перенавчити на нових зображеннях та дослідити показники її якості. Модель рекомендовано використовувати для вивчення процесу створення комп'ютерного зору засобами open cv.

Висновки

1. Завдання розробки моделі штучного інтелекту для групування терористичних та ворожих військових об'єктів вирішується шляхом використання згорткових нейронних мереж типу Sequential, що використовують шари нормалізації пакету для стабілізації і нормалізації вихідних значень зображень шару перед їх подальшою обробкою.

2. За результатами експериментальної верифікації запропонованих рішень побудовано модель класифікації терористичних та ворожих військових об'єктів точність якої, на навчальному/тестовому наборах 0,6/0,55, при похибці 0,658/0,705, у той час як точність альтернативної моделі сягає лише 0,48 на тестовому наборі даних.

Література

1. François Chollet. Deep Learning with Python. Shelter Island, NY : Manning Publications Co., 2018. 361 p.

2. Density-based clustering with fully-convolutional networks for crowd flow detection from drones / Giovanna Castellano [et al.]. Neurocomputing. 2023. https://doi.org/10.1016/j.neucom.2023.01.059

3. Laktionov O., Boryak B., Pedchenko N., Mykhailichenko O.Overview of computer vision algorithms for detecting hazardous objects by drones. Control, Navigation and Communication System. 2023. Vol. 3, Issue 73. P. 120–122. https://doi.org/10.26906/sunz.2023.3.120

4. Ejaz N., Salimur Choudhury. Computer vision in drone imagery for infrastructure management. Automation in Construction. 2024. Vol. 163. P. 105418. https://doi.org/10.1016/j.autcon.2024.105418

5. Addressing practical challenge of using autopilot drone for asphalt surface monitoring: Road detection, segmentation, and following / Hormazd Ranjbar [et al.]. Results in Engineering. 2023. Vol. 18. P. 101130. https://doi.org/10.1016/j.rineng.2023.101130

6. Semi-Autonomous Mobile Robot Coupled to a Drone for Debris Removal from High-Voltage Power Lines / Rogério S. Gonçalves [et al.]. Robotics and Autonomous Systems. 2024. P. 104697. https://doi.org/10.1016/j.robot.2024.104697

7. Image Defect Classification of Surface Mount Technology Welding Based on the Improved ResNet Model [Electronic resource] / Qiang Zhang [et al.] // Journal of Engineering Research. – 2024. – Mode of access: https://doi.org/10.1016/j.jer.2024.02.007

8. The Automation of the Development of Classification Models and Improvement of Model Quality using Feature Engineering Techniques / Sjoerd Boeschoten [et al.]. Expert Systems with Applications. 2022. P. 118912. https://doi.org/10.1016/j.eswa.2022.118912

9. Review of resampling techniques for the treatment of imbalanced industrial data classification in equipment condition monitoring / Yage Yuan [et al.]. Engineering Applications of Artificial Intelligence. 2023. Vol. 126. P. 106911. https://doi.org/10.1016/j.engappai.2023.106911

10. Onyshchenko S., Alina Yanko, Alina Hlushko. Improving the efficiency of diagnosing errors in computer devices for processing economic data functioning in the class of residuals. Eastern-European Journal of Enterprise Technologies. 2023. Vol. 5, no. 4 (125). P. 63–73. https://doi.org/10.15587/1729-4061.2023.289185

11. Hossain S. M., Md Ahsan Ayub Parameter Optimization of Classification Techniques for PDF based Malware Detection. 2020 23rd International Conference on Computer and Information Technology (ICCIT), DHAKA, Bangladesh, 19–21 December 2020. [S. l.], 2020. https://doi.org/10.1109/iccit51783.2020.9392685

12. Kshitij Tayal, Rahul Ghosh, Vipin Kumar Model-agnostic Methods for Text Classification with Inherent Noise. https://aclanthology.org/2020.coling-industry.19.pdf.

13. Mozhaiev Oleksandr. Study of the internet of things network construction tasks. Control, Navigation and Communication Systems. 2024. Vol. 1, no. 75. P. 137–140. https://doi.org/10.26906/sunz.2024.1.137

14. Hierarchical Data Format 5 – Python 101 0.1.0 documentation. Python 101 – Python 101 0.1.0 documentation. https://scientific-python-101.readthedocs.io/file_formats/hdf5.html

15. Betti Sorbelli F., Lorenzo Palazzetti, Cristina M. Pinotti YOLO-based detection of Halyomorpha halys in orchards using RGB cameras and drones. Computers and Electronics in Agriculture. 2023. Vol. 213. P. 108228. https://doi.org/10.1016/j.compag.2023.108228

16. 2022 Russia Ukraine War, Losses, Oryx + Images. Kaggle: Your Machine Learning and Data Science Community. https://www.kaggle.com/datasets/piterfm/2022-ukraine-russia-war-equipment-losses-oryx/data