

С. В. Ковбасюк, Р. М. Осадчук, М. П. Романчук, Л. М. Наумчак

АЛГОРИТМ ФОРМУВАННЯ НАБОРУ АПРІОРНИХ ДАНИХ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ОБРОБКИ ЦИФРОВИХ АЕРОЗНІМКІВ

Аналіз застосування безпілотних авіаційних комплексів у зоні проведення бойових дій вказує на постійне підвищення вимог до розвідувальних даних. Одним із шляхів підвищення ефективності застосування цього виду озброєння є використання методів штучного інтелекту. Базисом для побудови надійної моделі нейронної мережі є велика кількість різноманітних даних, що допомагає точніше узагальнити інформацію про поставлене завдання. У статті розглянуто підхід до підвищення точності класифікації зображень об'єктів згортковими нейронними мережами на базі аугментації даних, який відрізняється від відомих адаптацією до факторів знімання та специфіки об'єктів аеророзвідки.

Важливим та найбільш трудомістким етапом побудови точної моделі машинного навчання є пошук та анотація даних, які будуть використовуватися для навчання та тестування точності роботи нейронної мережі. Від кількості зібраних даних залежить стабільність роботи мережі в реальних умовах. На сьогодні добре описано та проаналізовано такі методи аугментації зображень об'єктів, як геометричні перетворення, колірні корекції та просторова фільтрація зображення. Не розглянутими досі залишаються можливості оптимальних комбінацій методів аугментації даних для досягнення потрібного узагальнення слабопомітних інваріантних ознак об'єктів. З урахуванням проведених досліджень було запропоновано алгоритм формування набору апріорних даних нейронної мережі для розпізнавання об'єктів на цифрових аерознімках, що дозволить значно зменшити складність процесу збирання потрібних даних і замінити методами збільшення, які набагато простіші, витрачають менше обчислювального ресурсу та здатні підвищити точність роботи згорткових нейронних мереж. Крім того, у статті проведено моделювання використання запропонованого підходу.

Ключові слова: безпілотний літальний апарат; згорткова нейронна мережа; аугментація даних; автоматизована обробка цифрових аерознімків.

Постановка проблеми в загальному вигляді. Ефективне проведення сучасної військової операції (ведення бойових дій) не можливе без повної та своєчасної розвідувальної інформації, зокрема про стан та положення військ (сил) противника. Досвід ведення широкомасштабної війни росії проти України свідчить про зростання значущості застосування безпілотних авіаційних комплексів (БпАК) для добування розвідувальних даних [1].

Для підвищення ймовірності правильного розпізнавання об'єктів на аерознімках добре себе зарекомендували методи штучного інтелекту, а саме нейронні мережі (нейромережі), які дозволяють проводити обробку великого обсягу даних для

© С. В. Ковбасюк, Р. М. Осадчук, М. П. Романчук, Л. М. Наумчак, 2022

обґрунтування прийняття рішень у режимі реального часу та не витрачають на це значного обчислювального ресурсу. Особливістю такого підходу є застосування алгоритму навчання нейромережі з використанням апріорно розпізнаних фрагментів зображень об'єктів замість програмування процесу їх обробки. Такий підхід використовують, коли складно формалізувати набір ознак розпізнавання і створити алгоритм обробки зображень.

У разі малої кількості зображень об'єктів для такого навчання, незбалансованості їх класів, невідповідності апріорних зображень умовам знімання знижується точність класифікації зображень об'єктів [2–7]. Отже, актуальним є розроблення підходу щодо формування розширеного набору апріорних даних нейромережі для обробки цифрових аерознімків.

Аналіз останніх досліджень і публікацій. Алгоритми глибокого навчання продемонстрували можливість вирішення різноманітних складних інженерних та наукових завдань. Успішне застосування запропонованих алгоритмів неможливе без використання всебічних наборів даних для навчання та тестування [8–10]. Зокрема, алгоритми комп'ютерного зору для розв'язання проблем класифікації, виявлення об'єктів, семантичної сегментації та сегментації екземплярів вимагають великої кількості зображень для забезпечення інваріантності тренуваних моделей [11–12]. Відомі два шляхи покращення характеристик наборів даних, які використовують для навчальних цілей. Перший із них передбачає фізичний збір зразків набору даних, отриманих у різних умовах, для забезпечення їх різноманітності. На теперішній час є безліч великих наборів даних, які були сформовані для вирішення проблем комп'ютерного зору та зазвичай використовуються як еталонні [13–15]. Їх специфікою є те, що вони загальні та придатні для порівняння результатів роботи розроблених алгоритмів. Однак наявні універсальні набори розмічених зображень об'єктів можуть бути непридатними для вирішення конкретних інженерних завдань. Одне з можливих застосувань таких наборів полягає у використанні їх як основи для попередньої підготовки нейромереж (передачі навчання) [16], що дозволяє їх точно налаштувати та адаптувати для конкретної мети. Проте у деяких випадках таке налаштування не забезпечує шуканого результату через те, що деякі події рідко трапляються, і в результаті можна отримати лише декілька зразків зображень [17, 18]. Зрештою це призводить до перенавчання великих моделей. Отже, другий підхід, що ґрунтується на штучних маніпуляціях із початковим набором даних, має перспективу застосування.

Формулювання завдання дослідження. Метою статті є розроблення алгоритму формування необхідного набору апріорних даних цифрових аерознімків нейромережі з наявних геометричних перетворень зображень та їх колірної корекції, збільшення наборів даних (аугментації даних), що адаптований до факторів знімання та специфіки об'єктів, які розпізнаються.

Виклад основного матеріалу. Завдяки формуванню великих наборів зображень, підвищенню обчислювальних потужностей та розробкам на основі згорткових нейромереж (англ. CNN) вже отримані значні результати у вирішенні завдань комп'ютерного зору [19, 20]. Проте більшість великих наборів даних для спеціальних

завдань не є загальнодоступними, а тренування CNN на малих базах призводить до перенавчання, що не забезпечує узагальнення нейромережею невидимих інваріантних ознак.

Для забезпечення інваріантності моделі нейромережі потрібна навчальна вибірка, яка б враховувала зміну характерних ознак класів об'єктів залежно від умов виконання аерознімання із застосуванням безпілотного літального апарата (БпЛА) (освітленість, час доби, пора року, погодні умови тощо). Крім того, для запобігання перенавчанню нейромережі на навчальній вибірці недостатнього обсягу необхідно збільшити кількість апіорних даних. Логічним рішенням є використання методів збільшення набору даних [20], що є схемою регуляризації та штучно розширює наявний набір зображень об'єктів, щоб збільшити кількість інваріантних прикладів. Аугментація даних – це сукупність невитратних обчислювальних методів [2, 4], що складається з геометричних перетворень та колірної корекції зображення [21–23]. Досягнення потрібного узагальнення слабопомітних інваріантних ознак об'єктів забезпечується збільшенням навчальної вибірки шляхом підбору оптимальних комбінацій методів аугментації даних (геометричних перетворень зображень та їх колірної корекції).

Формування набору апіорних даних зображень об'єктів розпочинається з планування контрольованої класифікації (формується конкретне завдання комп'ютерної обробки знімків, визначається список класів об'єктів, що підлягають розпізнаванню).

Згідно з теорією розпізнавання образів для класифікації об'єктів простір ознак доцільно розділити на замкнуті області, кожна з яких містить значення ознак, характерних тільки для одного з класів об'єктів, і відносити піксель знімка до того класу, в область якого потрапив вектор його ознак. Функція призначена для визначення номера й області U_i , до якої належать ознаки даного зображення або його фрагмента. Правило прийняття рішення про належність даного зображення або його фрагмента до одного з класів ґрунтується на основі розбивання простору ознак U на $M_{кл}$ класів U_i , $i = 1, \dots, M_{кл}$, що не перетинаються:

$$\bigcup_{i=1}^{M_{кл}} U_i = U, \quad U_i \cap U_j = \emptyset \text{ при } i \neq j. \quad (1)$$

Для військового дешифрування аерознімків найбільш пристосованою і випробуваною є система класифікації, запропонована Ю. К. Ребріним та І. М. Карповичем [24, 25]. У її основу покладено ієрархічну структуру, яка передбачає поетапний перехід від загального до часткового на основі визначення характерних особливостей об'єкта. Ознаки, які лежать в основі такого поділу, полегшують операцію класифікації об'єктів, оскільки є визначниками для розпізнавання зображень.

Визначення належності зображення об'єкта на аерознімку до певного виду, класу, підкласу, типу можливе завдяки різниці в оптичних щільностях та за яскравістю зображень, розмічених поруч об'єктів або об'єкта й фону. Потрібно оцінити інформативність ознак, що використовуються для забезпечення необхідної ймовірності розпізнавання за мінімального набору ознак та здійснити вибір їх оптимального піднабору.

Отже, під розпізнавальними ознаками будемо розуміти одержані за допомогою технічних засобів розвідки характеристики об'єктів, за якими їх можна розпізнати [25]. До основних прямих розпізнавальних ознак простих об'єктів належать форма, розміри, колір, деталі.

У результаті наповнення класів зображеннями об'єктів потрібно досягти збалансованості (дотримуватися пропорційності зображень, однакових властивостей, відповідності їх кількості). Навчальні вибірки мають бути репрезентативними, однорідними, максимально подібними за розподілом значень яскравості зображень до нормального (гаусівського) закону.

Відповідно до вимог, викладених вище, потрібно емпірично визначити модель нейромережі [9] та набір підготовлених ваг згідно із завданням, здійснивши підбір гіперпараметрів та налаштування мережі. Результатом має бути відібрана модель нейромережі з налаштованими гіперпараметрами, яка дозволяє отримати найбільш ефективний результат, та визначена кількість епох навчання для досягнення глобального мінімуму в ході вирішення завдання оптимізації моделі. На прикладі графіка кривої втрат у ході навчання нейромережі CNN VGG-16 (рис. 1) видно, що різниця втрат на навчальній та валідаційній вибірках починає зменшуватися близько 22-ї епохи. Відповідно, для аналізу ефективності моделі за заданих гіперпараметрів у подальшому достатньо проведення навчання протягом 23 епох для досягнення максимальної ефективності нейромережі.

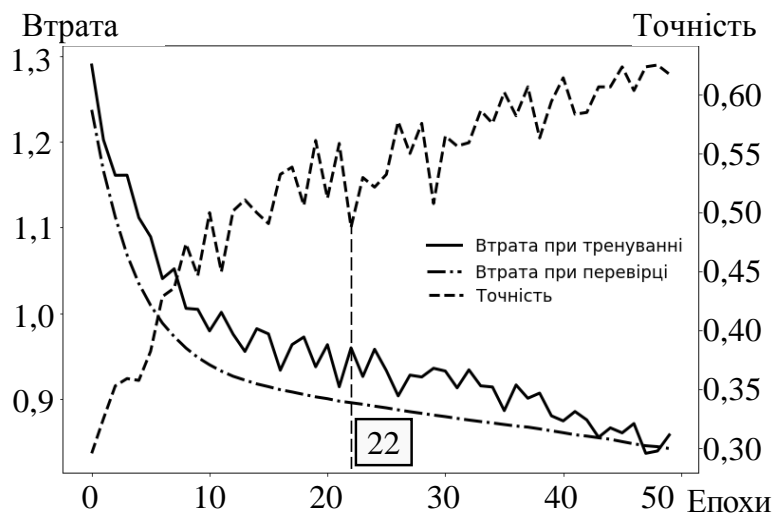


Рис. 1. Графік залежності втрат у ході навчання нейромережі

Покращити попередньо отримані результати дозволяє збільшення кількості навчальних даних зображень об'єктів з використанням інформації лише в навчальних даних. Основні методи належать до категорії деформації зображень об'єктів, що є підходом, спрямованим на безпосереднє збільшення вхідної інформації до моделі об'єкта в просторі ознак. Сказане вище можна подати в такому вигляді:

$$\begin{aligned} \varphi: S &\rightarrow T, \\ S' &= S \cup T, \end{aligned} \quad (2)$$

де S – оригінальний набір зображень об'єктів;

T – відповідні перетворення, визначені функцією φ ;

S' – збільшений набір зображень об'єктів.

Усі методи базуються на комбінації афінних перетворень вхідного зображення для маніпулювання навчальними даними [20]. Для кожного вхідного зображення створюється «дубльований» образ – це зміщення, збільшення / зменшення, повертання, перевертання, відображення, перекладання, обрізання, спотворення або затінення, додавання гаусівського шуму та зміна кольорової палітри. Зображення та дублікат надходять до нейромережі.

Геометричні перетворення зображення змінюють геометрію зображення об'єкта шляхом зміщення окремих значень координат розташування його пікселів. Базова форма об'єкта на зображенні зберігається, але змінюється до якоїсь нової позиції та орієнтації. Так схема обертання повертає зображення навколо його центра шляхом відображення кожного його пікселя (x, y) до (x', y') з таким перетворенням:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (3)$$

Дослідницькі експерименти [4] показали, що зміна θ від -30° до $+30^\circ$ встановлює достатньо великі обертання, щоб генерувати нові зразки. Обрізка є ще однією схемою збільшення набору зображень, популяризованою Кріжевським [20].

Колірна корекція трансформує канали RGB, змінюючи значення кожного пікселя (r, g, b) до нових значень (r', g', b') відповідно до заздалегідь визначеної моделі, або проводиться корекція кольору за допомогою фільтра HSB [20]. Це регулює освітлення та колір і залишає геометрію зображення незмінною.

Просторова фільтрація зображення для пошуку розривності, яка є однією з базових властивостей яскравості сигналу, дозволяє підвищити результати навчання CNN [6, 7], що визначаються текстурою, формою об'єкта.

Зміну зображень об'єктів слід виконувати тільки з навчальною вибіркою. Для малих наборів даних варіант збільшення в автономному режимі є кращим порівняно з онлайнним збільшенням, оскільки розмір набору зображень зростає відповідно до здійснених перетворень. Потрібно порівняти отримані результати, програмно збільшивши дані та навчання моделі нейромережі. У результаті аналізу відібрати методи (або їх можливі комбінації) аугментації даних та діапазони зміни їх параметрів, які зумовлюють підвищення ефективності моделі. При цьому слід враховувати розмір утвореного набору зображень та час, витрачений на збільшення.

На основі викладеного матеріалу алгоритм формування набору апріорних даних нейромережі для автоматизованого розпізнавання об'єктів на аерознімках можна подати в такому вигляді:

1. Складання списку класів відповідно до завдання та заданих аерознімків. Правило прийняття рішення про визначення належності того чи іншого зображення об'єктів або його фрагмента одному із класів будується на основі розділення простору ознак U на $M_{кл}$ класів U_i , $i = 1, \dots, M_{кл}$.

2. Формування набору класів, що передбачає створення навчальних збалансованих вибірок для кожного класу об'єктів за їх зображеннями: $\varphi: S \rightarrow T$, $S' = S \cup T$, де S – оригінальний набір зображень об'єктів; T – відповідні перетворення, визначені функцією φ ; S' – збільшений набір зображень об'єктів.

3. Визначення архітектури моделі відповідно до поставленого завдання. Вибирається та архітектура моделі нейромережі, яка відповідає умовам конкретного завдання (1).

4. Налаштування моделі шляхом підбору значень гіперпараметрів $\text{minimize} \sum_{i=1}^m l(h_{\theta}(x^{(i)}), y^{(i)})$, де $x^{(i)} \in X_L$ – множина вхідних зображень об'єктів; $y^{(i)} \in Y_L$ – відповідна множина виходів; $h_{\theta}(x^{(i)})$ – функція гіпотези, яка оптимізується за параметрами моделі θ ; $l(h_{\theta}(x^{(i)}), y^{(i)})$ – функція втрат.

5. Навчання моделі нейромережі, що передбачає визначення достатньої кількості епох m для цього.

6. Аналіз результатів навчання моделі нейромережі за кожним окремим методом аугментації даних та визначення оптимальних діапазонів зміни параметрів відібраних методів збільшення даних.

7. Аналіз результатів застосування комбінацій методів збільшення набору зображень моделі CNN. Якщо отримано шуканий результат, то відбувається перехід до кроку 8, в іншому випадку – перехід до кроку 5.

8. Навчання моделі нейромережі за вибраними комбінаціями методів аугментації даних у межах визначених оптимальних параметрів до повного збігання моделі.

Застосування розробленої методики формування набору оптимальних апріорних даних нейромережі дозволяє підвищити достовірність розпізнавання зображень об'єктів шляхом регуляризації її за рахунок інваріантності до умов проведення аерознімання.

Для аналізу впливу різних методів збільшення набору даних на класифікацію об'єктів використовувалася CNN VGG-16 [4] із вагами, підготовленими на ImageNet, та такими гіперпараметрами: функція активації – *ReLU*; схема ініціалізації ваг – *Xavier*; швидкість навчання – 0,01; алгоритм оптимізації – стохастичний градієнтний спуск; функція оновлення – *Adam*; регуляризація – *L2* градієнтна нормалізація; мінінабори – 16; епохи навчання – 23.

Для навчання нейромережі використано набір даних, сформований для трьох збалансованих класів зразків озброєння та військової техніки, що містить 1658 зображень об'єктів розміром 70 x 70 пікселів. Для оцінювання результатів навчання моделі використовувалася метрика якості класифікації об'єктів точність *Accuracy*:

$$Accuracy = \frac{TP + TN}{N}, \quad (4)$$

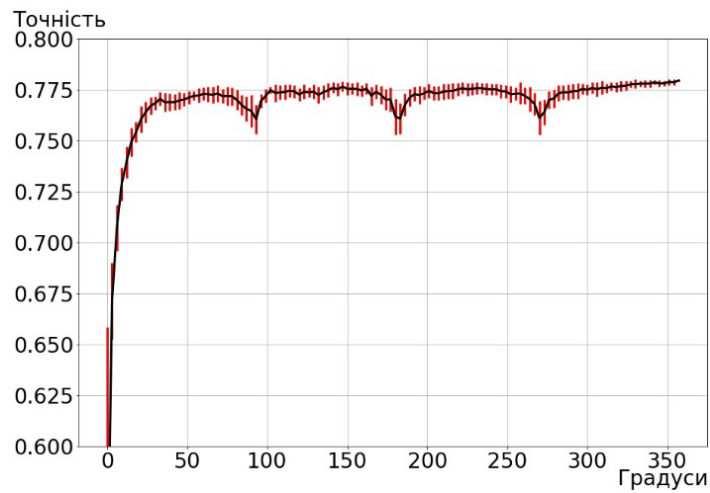
де *Accuracy* – точність класифікації об'єкта;

TP – істино-позитивне рішення;

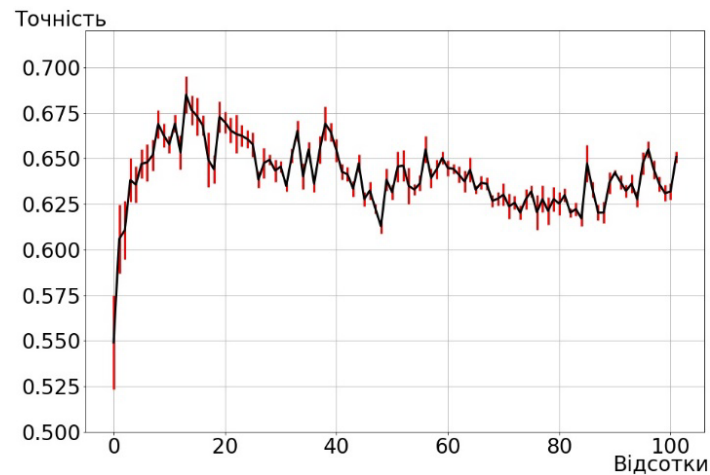
TN – істино-негативне рішення;

N – загальна кількість об'єктів.

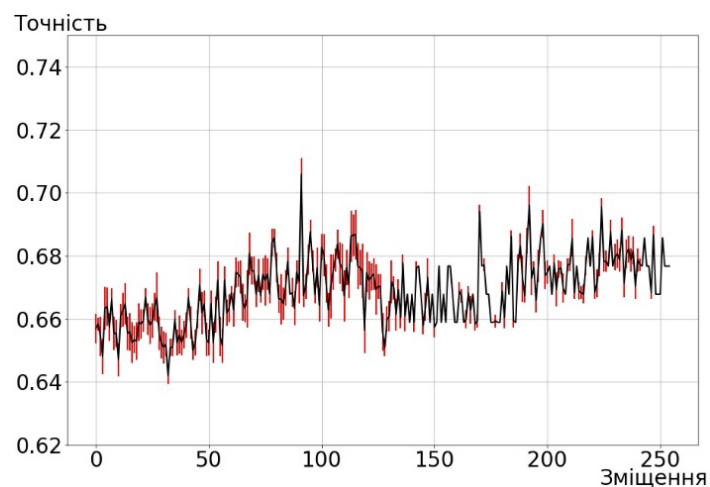
Результати застосування методів аугментації даних для зображень об'єктів, наведені в табл. 1 та зображені на рис. 2, демонструють підвищення точності класифікації для виконання завдання розпізнавання CNN.



а)



б)



в)

Рис. 2. Графіки зміни Ассигасу в разі використання методів обертання від 0° до 360° (а), обрізування та зміщення (б), кольорового зміщення (в) для нейромережі VGG-16

Метод обертання (рис. 2а) у разі повороту на кути, кратні 90° , показав втрату інваріантності дрібнозернистих зображень (зменшення ефективності на 2,5% для завдання класифікації). Метод випадкового обрізування та зміщення (рис. 2б) отримав гірші результати, ніж для крупнозернистих зображень [4], хоча він створив більше зразків зображень об'єктів порівняно з іншими схемами збільшення.

Методи кольорового зміщення (рис. 2в) [6, 10] дали незначне підвищення точності Ассигасу порівняно з геометричними схемами. Отже, CNN дає кращу ефективність виконання завдань у разі навчання на зображеннях, що містять зміни в геометрії, а не редагування освітлення та зміщення кольорів.

Таблиця 1

Результати Ассигасу для різних видів трансформації нейромережі VGG-16

| № | Метод аугментації даних | Ассигасу, % |
|---|---|-------------|
| 1 | Без збільшення даних | 72 |
| 2 | Обертання 0° – 360° | 77,56 |
| 3 | Випадкове обрізування та зміщення | 74,5 |
| 4 | Обрізування з фіксованими розмірами | 75,3 |
| 5 | Збільшення / зменшення | 74,1 |
| 6 | Зміна зеленого кольору | 72,3 |
| 7 | Випадкове обрізування + обертання 360° | 72,9 |
| 8 | Контраст | 73 |

Отже, збільшення навчальних даних знижує ймовірність перенавчання моделі нейромережі, поліпшує узагальнення і, як результат, підвищує загальну ефективність виконання завдання розпізнавання зображень об'єктів. Для визначення ефективності застосування комбінацій оптимальних методів аугментації даних проводиться навчання моделі нейромережі з оцінюванням отриманих результатів за метрикою точності. З метою уникнення надмірних часових витрат на навчання моделі на значній кількості епох кількість навчальних епох обмежується моментом початку збігання моделі (рис. 1). За вибраними комбінаціями методів аугментації даних у межах визначених оптимальних параметрів проводиться навчання моделі нейромережі до повного її збігання.

Висновки. На основі результату аналізу поширених методів аугментації даних (геометричних та кольорової корекції) розроблено алгоритм формування необхідного набору апріорних даних нейромережі, який забезпечує адаптацію до специфіки об'єктів дешифрування та факторів аерознімання із застосуванням БпАК. Запропонований алгоритм дозволяє оптимально збільшити набір зображень об'єктів, запобігаючи навчанню нейромережі невідповідним шаблонам, та підвищити точність класифікації об'єктів моделлю нейромережі. Використання ефективних методів аугментації даних вхідного набору зображень об'єктів дозволить значно зменшити складність збирання потрібних даних та заощадити обчислювальний ресурс. Даний алгоритм може бути використаний у комплексах БпАК у разі застосування нейромереж для автоматизації обробки зображень за результатами знімання та для формування еталонів зображень військових об'єктів.

Перспективним напрямом подальших досліджень є порівняння впливів комбінації традиційного збільшення наборів даних із результатами роботи генеративних змагальних мереж та формування синтетичних наборів даних для розпізнавання об'єктів.

СПИСОК БІБЛОГРАФІЧНИХ ПОСИЛАНЬ

1. Петрук С. Безпілотні авіаційні комплекси в збройних конфліктах останніх десятиріч. URL: <https://journal.cndiovt.com.ua/article/download/210/222/> (дата звернення: 21.12.2022).
2. Lemley J., Bazrafkan S., Corcoran P. Smart Augmentation Learning an Optimal Data Augmentation Strategy // IEEE Access. 2017. Vol. 5. P. 5858–5869. <https://doi.org/10.1109/ACCESS.2017.2696121>
3. DeVries T., Taylor G. Dataset Augmentation in Feature Space // Workshop track- ICLR 2017. <https://doi.org/10.48550/arXiv.1702.05538>
4. Inoue H. Data Augmentation by Pairing Samples for Images Classification. <https://doi.org/10.48550/arXiv.1801.02929>
5. Learning Data Augmentation Strategies for Object Detection / B. Zoph, E. Cubuk, G. Ghiasi et al. // European Computer Vision Association. <https://doi.org/10.48550/arXiv.1906.11172>
6. Dwibedi D., Misra I., Hebert M. Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection // IEEE International Conference on Computer Vision. 2017. <https://doi.org/10.1109/iccv.2017.146>
7. InstaBoost: Boosting Instance Segmentation via Probability Map Guided Copy-Pasting / Fang H.-Shu, Sun J., Wang R. et al. // IEEE/CVF International Conference on Computer Vision. 2019. <https://doi.org/10.1109/iccv.2019.00077>
8. Kwon O., Sim J. Effects of data set features on the performances of classification algorithms // Expert Systems with Applications. 2013. Iss. 5, Vol. 40. P. 1847–1857. <https://doi.org/10.1016/j.eswa.2012.09.017>
9. Sbai O., Couprie C., Aubry M. Impact of base dataset design on few-shot image classification // European Conference on Computer Vision. 2020. P. 597–613. https://doi.org/10.1007/978-3-030-58517-4_35
10. How good is my test data? Introducing safety analysis for computer vision / Zendel O. et al. // International Journal of Computer Vision. 2017. Iss. 125. P. 95–109. <https://doi.org/10.1007/s11263-017-1020-z>
11. The many faces of robustness: A critical analysis of out-of-distribution generalization / Hendrycks D. et al. // IEEE/CVF International Conference on Computer Vision. <https://doi.org/10.1109/iccv48922.2021.00823>
12. Improving the Robustness of Deep Neural Networks via Stability Training / Zheng S. et al. // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. P. 4480–4488. <https://doi.org/10.48550/arXiv.1604.04326>
13. Imagenet: A large-scale hierarchical image database / Deng J. et al. // IEEE conference on computer vision and pattern recognition. 2009. P. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
14. Microsoft COCO: Common objects in context / Lin T. et al. // In European conference Computer Vision – ECCV 2014. № 8693. P. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

15. DOTA: A Large-scale Dataset for Object Detection in Aerial Images / Xia G.-S. et al. // IEEE Conference on Computer Vision and Pattern Recognition. 2018. P. 3974–3983. <https://doi.org/10.1109/cvpr.2018.00418>
16. ActivityNet: A large-scale video benchmark for human activity understanding / F. Heilbron et al. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 961–970. <https://doi.org/10.1109/cvpr.2015.7298698>
17. A survey on deep transfer learning / Tan C. et al. // International conference on artificial neural networks : Springer. 2018 P. 270–279. <https://doi.org/10.48550/arXiv.1808.01974>
18. Vannucci C., Colla V. Classification of Unbalanced Datasets and Detection of Rare Events in Industry: Issues and Solutions // International Conference on Engineering Applications of Neural Networks : Springer, January 01, 2016. P. 337–351. https://doi.org/10.1007/978-3-319-44188-7_26
19. Generalizing from a Few Examples: A Survey on Few-shot Learning / Wang Y. et al. // ACM Computing Surveys. 2022. Vol. 53, № 3. P. 1–34. <https://doi.org/10.1145/3386252>
20. Handwritten Digit Recognition with a Back-Propagation Network / LeCun Y. et al. // Advances in Neural Information Processing Systems. 1996. P. 396–404.
21. The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches / Z. Alom et al. // Preprint arXiv:1803.01164. 2019. <https://doi.org/10.48550/arXiv.1803.01164>
22. Russakovsky O., Deng J., Su H. Imagenet large scale visual recognition challenge // Advances in Neural Information Processing Systems. 2016. P. 3108–3116. <https://doi.org/10.1007/s11263-015-0816-y>
23. Simonyan K., Zisserman A. Two-Stream Convolutional Networks for Action Recognition in Videos // Advances in Neural Information Processing Systems 27 (NIPS 2014). 2014. Vol. 1. P. 568–576.
24. Карпович И. Н. Военное дешифрирование аэроснимков : учебник. Москва : Воениздат, 1990. 544 с.
25. Ребрин Ю. К. Оптико-электронное разведывательное оборудование летательных аппаратов : учебник. Киев : КВВАИУ, 1988. 452 с.

Стаття надійшла до редакції 24.12.2022.

REFERENCES

1. Petruk, S. (n.d.). Bezpilotni aviatsiini komplekxy v zbroinykh konfliktakh ostannykh desiatyrich. Retrieved from <https://journal.endiovt.com.ua/article/download/210/222/> [in Ukrainian].
2. Lemley, J., Bazrafkan, S., & Corcoran, P. (2017). Smart Augmentation Learning an Optimal Data Augmentation Strategy. *IEEE Access*, Vol. 5, 5858–5869. <https://doi.org/10.1109/ACCESS.2017.2696121>
3. DeVries, T., & Taylor, G. (2017). Dataset Augmentation in Feature Space. *Workshop track-ICLR*. <https://doi.org/10.48550/arXiv.1702.05538>
4. Inoue, H. (n.d.). *Data Augmentation by Pairing Samples for Images Classification*. <https://doi.org/10.48550/arXiv.1801.02929>
5. Zoph, B., Cubuk, E., & Ghiasi, G. et al. (n.d.). Learning Data Augmentation Strategies for Object Detection. *European Computer Vision Association*. <https://doi.org/10.48550/arXiv.1906.11172>

6. Dwibedi, D., Misra, I., & Hebert, M. (2017). Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection. In *IEEE International Conference on Computer Vision*. <https://doi.org/10.1109/iccv.2017.146>
7. Fang, H.-Shu, Sun, J., & Wang, R. et al. (2019). InstaBoost: Boosting Instance Segmentation via Probability Map Guided Copy-Pasting. In *IEEE/CVF International Conference on Computer Vision*. <https://doi.org/10.1109/iccv.2019.00077>
8. Kwon, O., & Sim, J. (2013). Effects of data set features on the performances of classification algorithms. *Expert Systems with Applications, Iss. 5, Vol. 40*, 1847–1857. <https://doi.org/10.1016/j.eswa.2012.09.017>
9. Sbai, O., Couprie, C., & Aubry, M. (2020). Impact of base dataset design on few-shot image classification. In *European Conference on Computer Vision*. (pp. 597–613). https://doi.org/10.1007/978-3-030-58517-4_35
10. Zendel, O. et al. (2017). How good is my test data? introducing safety analysis for computer vision. *International Journal of Computer Vision*, 125, 95–109. <https://doi.org/10.1007/s11263-017-1020-z>
11. Hendrycks, D. et al. (n.d.). The many faces of robustness: A critical analysis of out-of-distribution generalization. In *IEEE/CVF International Conference on Computer Vision*. <https://doi.org/10.1109/iccv48922.2021.00823>
12. Zheng, S. et al. (2016). Improving the Robustness of Deep Neural Networks via Stability Training. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. (pp. 4480–4488). <https://doi.org/10.48550/arXiv.1604.04326>
13. Deng, J. et al. (2009). Imagenet: A large-scale hierarchical image database. In *IEEE conference on computer vision and pattern recognition*. (pp. 248–255). <https://doi.org/10.1109/CVPR.2009.5206848>
14. Lin, T. et al. (2014). Microsoft COCO: Common objects in context. In *European conference Computer Vision – ECCV*, 8693, 740–755. https://doi.org/10.1007/978-3-319-10602-1_48
15. Xia, G.-S. et al. (2018). DOTA: A Large-scale Dataset for Object Detection in Aerial Images. In *IEEE Conference on Computer Vision and Pattern Recognition*. (pp. 3974–3983). <https://doi.org/10.1109/cvpr.2018.00418>
16. Heilbron, F. et al. (2015). ActivityNet: A large-scale video benchmark for human activity understanding. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (pp. 961–970). <https://doi.org/10.1109/cvpr.2015.7298698>
17. Tan, C. et al. (2018). A survey on deep transfer learning. In *International conference on artificial neural networks*. (pp. 270–279). <https://doi.org/10.48550/arXiv.1808.01974>
18. Vannucci, C., & Colla, V. (2016). Classification of Unbalanced Datasets and Detection of Rare Events in Industry: Issues and Solutions. In *International Conference on Engineering Applications of Neural Networks*. (pp. 337–351). https://doi.org/10.1007/978-3-319-44188-7_26
19. Wang, Y. et al. (2022). Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Computing Surveys, Vol. 53, № 3*, 1–34. <https://doi.org/10.1145/3386252>
20. LeCun, Y. et al. (1996). Handwritten Digit Recognition with a Back-Propagation Network. *Advances in Neural Information Processing Systems*. P. 396–404.
21. Alom, Z. et al. (2019). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. *Preprint arXiv:1803.01164*. <https://doi.org/10.48550/arXiv.1803.01164>

22. Russakovsky, O., Deng, J., & Su, H. (2016). Imagenet large scale visual recognition challenge. *Advances in Neural Information Processing Systems*. P. 3108–3116. <https://doi.org/10.1007/s11263-015-0816-y>
23. Simonyan, K., & Zisserman, A. (2014). Two-Stream Convolutional Networks for Action Recognition in Videos. *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, Vol. 1, 568–576.
24. Karpovich, I. N. (1990). *Voennoe deshifrirovaniye aerosnimkov [Military interpretation of aerial photographs]*. Moscow [in Russian].
25. Rebrin, Iu. K. (1988). *Optiko-elektronnoe razvedyvatel'noe oborudovaniye letatel'nykh apparatov [Optoelectronic reconnaissance equipment of aircraft]*. Kyiv [in Russian].

S. V. Kovbasyuk, R. M. Osadchuk, M. P. Romanchuk, L. M. Naumchak
AN APPROACH TO FORMING A PRIOR DATASET OF NEURAL NETWORK FOR PROCESSING DIGITAL AERIAL PHOTOS

Analysis of the use of unmanned aerial systems in the combat zone indicates a constant increase in the requirements for intelligence data. One of the ways to increase the effectiveness of UaV application is the use of artificial intelligence methods. The basis for building a reliable neural network model is a large amount of various data, which helps to more accurately summarize information about the given task. The article considers an approach for increasing the accuracy of object image classification by convolutional neural networks based on data augmentation, which differs from existing ones in its adaptation to shooting factors and the specificity of aerial reconnaissance objects.

An important and most time-consuming step in building an accurate machine learning model is finding and annotating the data that will be used to train and test the accuracy of the neural network. The accuracy and stability of the network in real conditions depends on the amount of collected data. To date, such methods of object image augmentation as geometric transformations, color correction, and spatial image filtering have been well described and analyzed. The possibilities of optimal combinations of data augmentation methods to achieve the desired generalization of poorly visible invariant features of objects remain unexplored at the moment. Taking into account the conducted research, an approach was proposed for the formation of a set of a priori data of a neural network for object recognition on digital aerial photographs, which will significantly reduce the complexity of the process of collecting the necessary data and replace it with magnification methods that are much simpler, consume less computing resources, and increase the accuracy of work convolutional neural networks, simulation of the use of the proposed approach was carried out.

Keywords: *unmanned aerial vehicle; convolutional neural network; data augmentation; automated processing of digital aerial photographs.*