УДК 629.7.014-519.05:004.891.032.26 **doi: 10.32620/reks.2018.4.04**

В. В. МОСКАЛЕНКО, А. С. МОСКАЛЕНКО, А. Г. КОРОБОВ, М. О. ЗАРЕЦЬКИЙ, В. А. СЕМАШКО

Сумський державний університет, Україна

МОДЕЛЬ ТА АЛГОРИТМ НАВЧАННЯ СИСТЕМИ ДЕТЕКТУВАННЯ МАЛОРОЗМІРНИХ ОБ'ЄКТІВ ДЛЯ МАЛОГАБАРИТНИХ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

У роботі розроблено ефективні модель та алгоритм навчання бортової системи малогабаритного безпілотного літального апарату для детектування малорозмірних об'єктів, що функціонує за умов обмеженого обсягу обчислювальних ресурсів та навчальних даних. Пропонується 4-х етапна реалізація алгоритму навчання. Перший етап полягає у виборі типу глибокої згорткової мережі, наперед навченої на наборі даних ImageNet, та кількості її низькорівневих шарів, що запозичуються для екстракції інформативних ознак. Другий етап включає навчання без вчителя високорівневих розріджено кодуючих згорткових шарів з використанням модифікованого алгоритму зростаючого нейронного газу, який забезпечує оптимальний розподіл нейронів на навчальних даних та автоматичне визначення необхідної кількості нейронів. Застосування модифікованого алгоритму зростаючого нейронного газу дозволяє ефективно використати нерозмічені навчальні дані і адаптувати верхні шари екстрактора ознак до доменної області застосування. На третьому етапі вихідна карта ознак формується шляхом об'єднання карт ознак з різних шарів глибокої згорткової нейронної мережі. При цьому здійснюється зменшення глибини вихідної карти ознак на основі аналізу головних компонент з наступною побудовою вирішальних правил. Цей підхід дозволяє використати всю наявну контекстну інформацію, що забезпечує ефективне розпізнавання малорозмірних об'єктів на аерофотознімках. Для виконання класифікаційного аналізу результуючої карти ознак запропоновано використовувати інформаційно-екстремальний класифікатор, що навчається за принципом бустінгу. Окрім того для прогнозування координат малорозмірних об'єктів на кадрі зображення пропонується використання регресійного алгоритму ортогональної інкрементальної машини екстремального навчання. Останній етап навчання детектора передбачає тонке налаштування високорівневих шарів глибокої нейронної мережі з використанням метаевристичного алгоритму симуляції відпалу з метою наближення глобального оптимуму комплексного критерії ефективності навчання детектора. В результаті застосування запропонованого підходу вдалося досягнути 96% точності виявлення об'єктів інтересу на тестових зображеннях відкритого набору даних, що свідчить про придатність моделі та методу навчання для практичного використання. При цьому розмір навчальної вибірки, яка використовувалась для побудови моделі складала 500 нерозмічених та 200 розмічених зразків.

Ключові слова: *зростаючий нейронний газ; згорткова нейронна мережа; детектування об'єктів; інформаційний критерій; алгоритм імітації відпалу; екстремальне навчання.*

Вступ

Безпілотна авіація знайшла широке застосування у пошуково-рятувальних роботах, віддаленій інспекції, в задачах охорони територій, тощо. Одним з шляхів підвищення функціональної ефективності безпілотної авіації є впровадження технологій штучного інтелекту для аналізу даних отриманих від бортових сенсорів. Бортові відеокамери є найбільш інформативним сенсором, а функція детектування об'єктів є найбільш затребуваною, що робить розробку моделей, здатних з високою точністю розпізнавати об'єкти інтересу на зображені, перспективним напрямком. Проте обмеження на обсяг обчислювальних та енерго ресурсів, а також габарити і вагу літального апарату унеможливлюють створення моделей аналізу візуальної інформації, що адаптовані до широкого спектру варіації спостережень та модифікацій об'єктів інтересу. Це обумовлює необхідність розробки обчислювально ефективної моделі та методу навчання системи розпізнавання малорозмірних об'єктів інтересу з можливістю адаптації до нових умов функціонування, які характерні в заданій області використання.

У задачах аналізу зображень набули широкого використання численні моделі ієрархічного ознакового подання на основі згорткових нейронних мереж [1]. Найбільш популярними серед них є VGG-16, VGG-19, ResNet-50, GoogleNet, MobileNet, SqueezeNet [2, 3]. Ці мережі відрізняються кількістю

© В. В. Москаленко, А. С. Москаленко, А. Г. Коробов, М. О. Зарецький, В. А. Семашко

шарів, наявністю залишкових зв'язків та різномасштабних фільтрів в кожному з шарів. При цьому в працях [2, 4] було показано, що навчені моделі на наборі даних ImageNet акумулюють в собі важливу інформацію щодо аналізу візуальних образів. Проте чим більша різниця між цільовою область використання мережі від образів представлених в наборі даних ImageNet, тим менше шарів навчених мереж можна повторно використати.

В роботах [4, 5] запропоновано використовувати тонке настроювання детектора об'єктів на основі згорткової мережі VGG-16, шляхом використання алгоритму міні-пакетного зворотного поширення помилки. Однак, це потребувало значних обсягів навчальних даних та використання потужних графічних процесорів, впродовж декількох днів. У праці [6] було запропоновано здійснювати сканування нормалізованої високорівневої карти ознак ковзним вікном, в кожній позиції якого здійснюється класифікаційний аналіз. У праці [7] було запропоновано використовувати в рамках так званої інформаційно-екстремальної технології (IEI-технологія) класифікаційний аналіз високорівневого ознакового подання. Головна ідея такого підходу полягає у адаптивному двійковому кодуванні і побудові оптимальних в інформаційному сенсі радіально-базисних вирішальних правил в двійковому просторі Хеммінга. Цей підхід забезпечує високу обчислювальну ефективність, оскільки використовуються лише прості операції порівняння та логічна XOR. Проте, не було запропоновано вирішення проблеми адаптації високорівневих шарів екстрактора ознак до доменної області використання та методу швидкої оптимізації порогів на ознаки розпізнавання. Індукція ознак на основі випадкового лісу та бустінгу є двома перспективними підходами для прискорення оптимізації бінарного кодування ознак, але є малодослідженими [8].

У працях [3, 9] пропонується навчання згорткових мереж без вчителя на основі автоенкодера або обмеженої машини Больцмана, які потребують великий обсяг навчальних даних і тривалий час навчання для отримання прийнятного результату. У працях [10, 11] пропонується поєднання принципів нейронного газу та розрідженого кодування для навчання згорткових фільтрів за нерозміченими даними. Даний підхід має м'яку конкуруючу схему навчання, що підвищує ймовірність збіжності алгоритму до оптимального розподілу нейронів на навчальних даних. При цьому впровадження алгоритмів розрідженого кодування дозволяє підвищити завадозахищеність і узагальнюючу здатність ознакового подання. Однак вибір кількісті нейронів не формалізований та задається на розсуд розробника, або налаштовується шляхом оптимізації, що призводить до збільшення кількості ітерацій навчання.

Необхідну кількість згорткових фільтрів у високорівневих згорткових шарах наперед оці-нити важко, тому перспективним підходом до навчання згорткових фільтрів є використання принципів зростаючого нейронного газу, який дозволяє автоматично визначити необхідну кількість нейронів [11]. Однак механізм додавання нових нейронів в алгоритмі нейронного газу на основі заданого періоду вставки часто призводить до викривлення утворених структур і не-стабільності процесу навчання. Проте у праці [12] було показано, що забезпечити стабільність навчання можна шляхом задавання «радіусу досяжності» нейронів. Це передбачає використання замість періоду вставки нейронів порогу максимального віддалення нейрону від кожної з віднесених до нього точок навчальної множини. Однак досі не було розглянуто механізми оновлення нейронів та оцінки віддаленості точок вхідного простору до нейронів з метою адаптації процесу навчання до процедури розрідженого кодування спостережень

Задачу детектування об'єктів на основі карти ознак згорткової мережі вирішують на основі архітектур згорткових нейронних мереж таких як: Yolo (You only look once), Faster R-CNN, DetectNet та SSD (Single Shot MultiBox Detector) [13, 14]. Важливою частиною таких архітектур є регресійна модель для прогнозування обмежувальної рамки об'єктів інтересу, яка дозволяє точно локалізувати об'єкт на зображенні. Однак, навчання таких моделей в умовах обмежених обчислювальних ресурсів та обсягу навчальних даних методом стохастичного градієнтного спуску є неефективним. Одним з перспективних шляхів реалізації регресійної моделі обмежувальних рамок є використання екстремальної машини навчання (Extreme Learning Machine, ELM), для якої характерна висока швидкість навчання [15]. Актуальним питанням є дослідження застосування методу інкрементального навчання шляхом послідовного додавання прихованих вузлів для уникнення проблеми перенавчання, яка виникає при великій кількості вузлів прихованого шару. В цьому випадку питання тонкого настроювання параметрів екстрактора ознак може бути вирішене шляхом застосування метаевристичних алгоритмів оптимізації як альтернативи алгоритму градієнтного спуску. Серед таких алгоритмів варто виділити алгоритм симуляції відпалу, який характеризується кращою конвергенцією та меншою ймовірність затрягання в «поганому» локальному оптимумі [16]. Проте, його використання в задачах тонкої настройки згорткових фільтрів все ще залишається не достатньо дослідженим.

Тому, дослідження спрямоване на розробку моделі та ефективного алгоритму навчання системи

детектування об'єктів інтересу на зображенні за умов обмежених обчислювальних ресурсів та об'єму навчальних даних є актуальним.

Постановка задачі

Нехай $D_{train} = \{ I_k, B_k \mid k = \overline{1, K_1} \}$ та $D_{test} = \{ I_k, B_k \mid k = \overline{1, K_2} \}$ є набором навчальних та тестових даних, де I_k – k-те зображення з кольоровою схемою RGB з роздільною здатністю 300x300, B_k – набір правильно заданих обмежувальних рамок, які відповідають об'єктам інтересу на k-му зображенні, K_1, K_2 – обсяги навчального та тестового наборів даних відповідно. Нехай алфавіт класів розпізнавання $\{X_z^o \mid z = \overline{1, Z}\}$ характеризує малорозмірні об'єкти. В даному випадку кількість правильно заданих обмежувальних рамок об'єктів інтересу на клас розпізнавання становить не більше 200 реалізацій.

Дано структурований вектор просторово-часових параметрів функціонування системи детектування об'єктів інтересу, який у загальному випадку має структуру

$$g =< e_1, ..., e_{\xi_1}, ..., e_{\Xi_1}, f_1, ..., f_{\xi_2}, ..., f_{\Xi_2} >, \quad \Xi_1 + \Xi_2 = \Xi, \tag{1}$$

де $e_1, ..., e_{\xi_1}, ..., e_{\Xi_1}$ – генотипні параметри функціонування, які впливають на параметри алгоритмів екстракції ознак розпізнавання;

$f_1, ..., f_{\xi_2}, ..., f_{\Xi_2}$ – фенотипні параметри функціонування, які впливають на геометрію контейнерів класів розпізнавання.

При цьому відомі обмеження на параметри функціонування:

$$R_{\xi_1}(e_1, ..., e_{\xi_1}, ..., e_{\Xi_1}) \le 0, \quad R_{\xi_2}(f_1, ..., f_{\xi_2}, ..., f_{\Xi_2}) \le 0. \tag{2}$$

Необхідно в процесі машинного навчання визначити оптимальні значення координат вектора (1), які забезпечують максимум усередненого за алфавітом класів розпізнавання комплексного критерію

$$J = J_{Cls} \cdot J_{Loc}, \tag{3}$$

$$g^* = \arg\max_G \{J(g)\}, \tag{4}$$

де J_{Cls} – інформаційний критерій функціональної ефективності класифікаційного аналізу;

J_{Loc} – критерій точності локалізації обмежувальної рамки об'єкта інтересу.

При функціонуванні системи розпізнавання безпосередньо в робочому режимі необхідно забезпечити максимальну точність локалізації та класифікації об'єктів інтересу на тестових зображення.

Модель та алгоритм навчання системи детектування малорозмірних об'єктів

Важливою умовою вирішення задачі розробки моделі аналізу даних для малогабаритних безпілотних літальних апаратів, що функціонують в умовах ресурсних та інформаційних обмежень є використання всієї доступної апріорної інформації. Техніка передачі знань є одним з прикладів використання апріорної інформації, яка накопичена в попередньо навченій нейронній мережі [17, 18]. Такий підхід дозволяє запозичувати низькорівневі шари з попередньо навченої глибокої мережі на великому об'ємі даних, а високо рівневі шари адаптувати до поточної доменної області використання.

Однак, коли об'єкти інтересу мають малий розмір, то зона інтересу є недостатньо інформативною, тому важливим є використання додаткової інформації щодо контексту в якому знаходиться об'єкт розпізнавання. На рис. 1 зображено пропоновану архітектуру моделі детектування малорозмірних об'єктів інтересу, на основі комбінації техніки передачі знань та інформації про контекст, яка отримана шляхом об'єднання карт ознак отриманих з різних шарів штучної нейронної мережі.

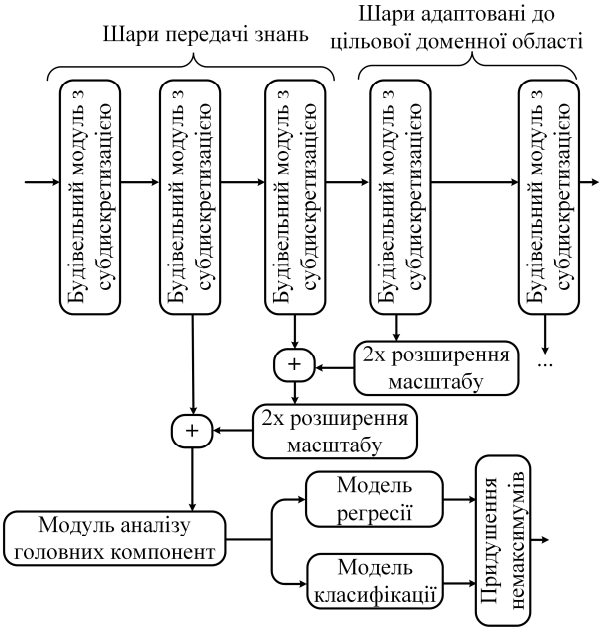


Рис. 1. Загальна архітектура детектора малорозмірних об'єктів на зображенні

Використання шару масштабування забезпечує однаковий розмір кожного каналу карти ознак. Конкатенація та масштабування розглядається як один шар масштабування-конкатенації.

Inception, Xception, VGG, Fire є найбільш популярними будівельними модулями, які використовуються для побудови глибоких згорткових нейронних мереж. Ці модулі мають різну мікро-архітектуру, що обумовлює різну обчислювальну складність та ефективність навчання цих моделей. Ми пропонуємо адаптувати нижні шари напереднавченої моделі нейронної мережі Squeezenet, яка складається з Fire модулів і характеризуються високою обчислювальною ефективністю. Верхні шари мережі в цьому випаду можуть бути побудовані за допомогою простих модулів архітектури VGG, які дозволяють отримати значну гнучкість при різних підходах навчання.

Першим етапом пропонованого алгоритму навчання, є реалізація навчання без вчителя високорівневих шарів мережі, для максимального використання нерозмічених реалізацій навчальних даних. В цьому випадку, для забезпечення завадозахищенності та інформативності ознакового подання, пропонується розраховувати функції активації кожного пікселя карти ознак за допомогою алгоритму Orthogonal Matching Pursuit (OMP) [8].

Навчання без вчителя високорівневих шарів мережі пропонується виконувати з використанням зростаючого розріджено кодуючого нейронного газу, що оснований на принципах зростаючого нейронного газу та розрідженого кодування.

В цьому випадку, навчальні дані для навчання високорівневих фільтрів згорткової мережі формуються шляхом розбиття вхідних зображень, або карти ознак на патчі. Ці патчі перетворюються в одновимірні вектори, що надходять на вхід алгоритму зростаючого розріджено кодуючого нейронного газу, основні етапи якого наведено в роботі [7].

Об'єднання карт ознак з різних шарів штучної нейронної мережі призводить до проблеми прокляття розмірності ознак. Для усунення даної проблеми пропонується використовувати аналіз головних компонент. Його використання дозволяє знизити розмірність не беручи до розгляду низькорівневі ознаки, які є нечутливими до цільової доменної області застосування.

Окрім того, пропонується реалізація класифікаційного аналізу карти ознак за допомогою так званої інформаційно-екстремальної технології з кодуванням ознак на основі бустінгу. Ця технологія дозволяє виконати класифікаційний аналіз з низькою обчислювальною складністю та відносно високою точністю в умовах обмеженого обсягу навчальних даних [7].

Для класифікаційного аналізу вихідної карти ознак формується навчальний набір даних $D = \{x_j, y_j \mid j = \overline{1,n}\}$, де n – обсяг вибірки, а – мітка j-го пікселя, яка відповідає одному з класів алфавіту розпізнавання $\{X_z^o \mid z = \overline{1,Z}\}$. При цьому інформаційно-екстремальний класифікатор оцінює належність j-го елементу даних (пікселя на карті ознак) за його N_1 ознаками до одного з Z класів. Це здійснюється шляхом кодування ознак на основі дерев рішень та вирішальних правил, побудованих в радіальному базисі бінарного простору Хемінга. Навчання інформаційно-екстремального класифікатора відбувається відповідно до наступних кроків.

1. Ініціалізація вагових коефіцієнтів $w_j = 1/n$.

2. Для $k = 1 \ldots K$ виконуються наступні кроки.

3. Генерація D_k набору навчальних даних з усього набору D з використанням функції розподілу ймовірності $P(X = x_j) = w_j$.

5. Навчання дерева рішень T_k на наборі D_k із застосуванням ентропійного критерію забрудненості вузлів.

6. Виконання бінарного кодування вектора x_j з набору даних D шляхом об'єднання шляхів рішення в деревах T_1, …, T_k. Кожне дерево формує двійковий код, в якому одиниці знаходяться у позиціях, що відповідають номерам вузлів, через які проходить шлях вектора x_j від кореня до термінального вузла. Нарешті, формування результуючого коду вектора x_j відбувається шляхом об'єднання кодів, сформованих кожним із дерев [8]. Результатом роботи цього етапу є бінарна матриця $\{b_{z,s,i} \mid i = \overline{1,N_2}; s = \overline{1,n_z}; z = \overline{1,Z}\}$, де N_2 кількість індукованих бінарних ознак та n_z кількість реалізацій відповідного класу X_z^o. що задовольняє рівність $n = \sum_z n_z$.

7. Побудова інформаційно-екстремальних вирішальних правил в радіальному базисі бінарного простору Хемінга та обчислення інформаційного критерію

$$E_z^* = \max_{\{d\}} E_z(d), \qquad (5)$$

де $\{d\} = \{0,1,...,\left(\sum_i b_{z,i} \oplus b_{c,i} - 1\right)\}$ – набір концентричних радіусів з центром b_z;

b_z – опорний вектор розподілу даних в класі X_z^o, який розраховуються з правилом

$$b_{z,i} = \begin{cases} 1, & \text{if } \frac{1}{n_z}\sum_{s=1}^{n_z} b_{z,s,i} > \frac{1}{Z}\sum_{c=1}^{Z}\frac{1}{n_c}\sum_{s=1}^{n_c} b_{c,s,i}; \\ 0, & \text{otherwise;} \end{cases} \quad (6)$$

де E_z – критерій ефективності побудови вирішальних правил в режимі навчання для класу X_z^o, який розраховується як нормалізована модифікація інформаційної міри Кульбака [7]:

$$E_z = \frac{1-(\alpha_z+\beta_z)}{\log_2(2+\varsigma)-\log_2\varsigma} \cdot \cdot \log_2\left[\frac{2-(\alpha_z+\beta_z)+\varsigma}{(\alpha_z+\beta_z)+\varsigma}\right], \quad (7)$$

де α_z, β_z – частота хибно позитивних та хибно негативних спрацювань вирішальних правил класу X_z^o;

ς – будь-яке мале знакододатнє число, яке вводиться, для уникнення невизначеності при діленні на нуль.

З метою підвищення ефективності навчання, загальноприйнятим є зведення проблеми багатокласової класифікації до серії двокласових класифікацій за принципом «один проти всіх». Для уникнення проблеми незбалансованості класів розпізнавання, обумовленої переважанням у навчальному наборі даних негативних зразків, здійснюється введення синтетичного класу, який є альтернативним для X_z^o. Синтетичний клас представлено n_z векторами з інших класів, які найбільш близькі до опорного вектора b_z, де n_z – обсяг навчального набору даних класу X_z^o.

8. Тестування отриманих інформаційно-екстремальних вирішальних правил на наборі даних D та розрахунок коефіцієнту помилок для кожної реалізації з D. При цьому, в режимі розпізнавання прийняття рішення про належність вектора b до одного з класів розпізнавання з алфавіту $\{X_z^o \mid z = \overline{1,Z}\}$ здійснюється за максимальним значенням функції належності $\mu_z(b)$ згідно виразу $\arg\max_z\{\mu_z(b)\}$. У цьому випадку функція належності $\mu_z(b)$ бінарного подання b вхідного вектора даних *x* до класу X_z^o, оптимальний контейнер якого має опорний вектор b_z^* та радіус d_z^*, розраховується за формулою:

$$\mu_z(b) = \exp\left(-\sum_{i=1}^{N_2} b_i \oplus b_{z,i}^* / d_z^*\right). \quad (8)$$

9. Оновлення коефіцієнтів $\{w_j\}$, пропорційне отриманій помилці для вектора x_j.

10. Якщо $|E_k^* - E_{k-1}^*| < \varepsilon$ то вийти з циклу.

Іншою важливою задачею є знаходження точної границі об'єктів за допомогою подвибірки обмежувальних рамок. Пропонується реалізація локалізації об'єктів на зображенні на базі регресійної моделі, яка заснована на нейронній мережі прямого поширення з одним прихованим шаром.

Розглядається n випадкових окремих елементів даних

$$\{(x_j, y_j) \mid x_j \in R^N, y_j \in R^M, 1 \le j \le n\},$$

де $x_j = (x_{j1}, x_{j2}, \ldots, x_{jN})^T$ вхідний вектор, та $y = (y_1, y_2, \ldots, y_n)^T$ набір вихідних цільових векторів, де $y_j = (y_{j1}, y_{j2}, \ldots, y_{jM})^T$ – центри та розміри обмежувальних рамок об'єктів. Одношарова мережа прямого поширення з одним прихованим шаром з додаванням R прихованих вузлів та функцією активації $\varphi(x)$ представлена нижче

$$\sum_{r=1}^{R}\beta_j\varphi(w_r^T x_j + b_r) = o_j, 1 \le j \le n, \quad (9)$$

де $w_r = (w_{r1}, w_{r2}, \ldots, w_{rN})^T$ – вектор вагових коефіцієнтів, який пов'язує вхідний шар з *r*-м прихованим вузлом;

b_r – зміщення r-го прихованого нейрону;

$\beta_r = (\beta_{r1}, \beta_{r2}, ..., \beta_{rM})$ – вектор вагових коефіцієнтів, який пов'язує вихідний шар з r-м прихованим вузлом;

o_j – вихід мережі для вхідного вектора x_j;

$\varphi(x)$ – функція активації.

Мережа з R прихованими вузлами може відтворювати ці N зразків з нульовою помилкою, коли всі параметри можуть вільно коригуватись, тобто існують β_r, w_r та b_r. Вище вказані *n* рівнянь можуть мати компактну форму подання у вигляді матричного рівняння

$$H\beta = Y, \quad (10)$$

де

$$H = \begin{bmatrix} \varphi(w_1^T x_1 + b_1) & \dots & \varphi(w_R^T x_1 + b_R) \\ \dots & \dots & \dots \\ \varphi(w_1^T x_n + b_1) & \dots & \varphi(w_R^T x_n + b_R) \end{bmatrix}_{n \times R}, \quad (11)$$

$$\beta = \begin{bmatrix} \beta_1^M \\ \dots \\ \beta_R^M \end{bmatrix}_{R \times M}, \; Y = \begin{bmatrix} y_1^M \\ \dots \\ y_n^M \end{bmatrix}_{n \times M}. \quad (12)$$

де H – прихований шар вихідної матриці одношарової мережі прямого поширення.

Для вирішення проблем надлишковості прихованих вузлів та повільної збіжності моделі запропоновано використовувати регресійний алгоритм ортогональної інкрементальної машини екстремального навчання. Цей метод уникає надлишковості вузлів та отримує розв'язок рівняння $H\beta = Y$ методом найменших квадратів за рахунок використання методу ортогоналізації Грамма-Шмідта. Детальні теоретичні докази збіжності ортогональної інкрементальної машини екстремального навчання приведено в роботі Li Ying [15].

Метод навчання ортогональної інрементальної ELM відбувається відповідно до наступних кроків.

1. Встановлюється максимальна кількість ітерацій L_{max} та бажана точність навчання E_0.

2. Для $L = 1 \dots L_{max}$ виконуються наступні кроки.

3. Збільшити кількість прихованих вузлів на один : $r = r + 1$.

4. Випадковим чином згенерувати один прихований нейрон та розрахувати його вихідний вектор h_r.

5. Якщо $r = 1$ тоді $v_r = h_r$ інакше

$$v_r = h_r - \frac{\langle v_1, h_r \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, h_r \rangle}{\langle v_2, v_2 \rangle} v_2 - \dots - \frac{\langle v_{r-1}, h_r \rangle}{\langle v_{r-1}, v_{r-1} \rangle} v_{r-1}. \quad (12)$$

6. Якщо $\| v_r \| \geq \varepsilon$, то виконується обчислення вихідної ваги для нового прихованого вузла $\beta_r = v_r^T E / (v_r^T v_r)$ та розраховується нова залишкова помилка $E = E - v_r \beta_r$ інакше $r = r - 1$.

7. Якщо $\| E \| \geq E_0$ відбувається вихід з основного циклу роботи.

Для реалізації навчання класифікатора, регресійної моделі та тонкої настройки екстрактора ознак сформовано навчальні набори в процесі співставлення реальних обмежувальних рамок об'єктів інтересу з рамками за замовчуванням. Обмежувальні рамки за замовчуванням визначаються як карта ознак пікселів спроектована на вхідне зображення. Ми розглядаємо стан аерознімків в якому камера орієнтована вниз під прямим кутом і розташована на великій висоті (висота більша 100 м).

Встановлена за замовчуванням обмежувальна рамка ставиться у відповідність кожній реальній рамці об'єкту, якщо перетин з нею за метрикою Жаккарда більший 0,4. Регресійна модель навчається тільки на позитивних зразках (співставлені обмежувальні рамки за замовчуванням).

Комплексний критерій ефективності навчання J (3) детектора об'єктів враховує як ефективність класифікаційного аналізу J_{Cls}, так і ефективність прогнозування координат обмежувальної рамки об'єкта інтересу J_{Loc}.

Пропонується розраховувати критерій ефективності класифікації об'єктів інтересу за формулою

$$J_{Cls} = \frac{1}{Z} \sum_{z=1}^{Z} E_z, \quad (13)$$

де E_z – інформаційний критерій ефективності машинного навчання класифікатора розпізнавати z-й клас.

Ефективність передбачення границь об'єктів інтересу пропонується розраховувати за формулою

$$J_{Loc} = \frac{1}{\hat{n}} \sum_{i=1}^{\hat{n}} IoU_i, \quad (14)$$

де IoU_i – міра перетину (коефіцієнт Жаккарда) реальної рамки об'єкту інтересу з відповідною і-ю прогнозованою рамкою об'єкту;

$\hat{n}$ – кількість співставлених передбачених обмежувальних рамок.

Останнім етапом алгоритму навчання є виконання тонкого налаштування високорівневих фільтрів екстрактора ознак, отриманих після навчання без вчителя, для врахування суттєвої незбалансованості між об'єктами інтересу та патчами фонових ділянок зображень. Пропонується використання алгоритму симуляції відпалу, який є метаевристичним оптимізаційним алгоритмом. Ефективність алгоритму симуляції відпалу залежить від реалізації процедури пошуку сусідів create_neighbor_solution для формування нового рішення S_i на і-ій ітерації алгоритму. На рис. 2 показано псевдо код алгоритму симуляції відпалу, на кожній ітерації якого виконується обчислення функції критерію f(), яка обчис-

люється шляхом пропускання розміченого навчального набору даних через модель детектора та розрахунку комплексного критерію (3) [16, 20].

$s_{current} \leftarrow create_initial_solution()$
$s_{best} \leftarrow s_{current}$
$T \leftarrow T_0$
$C \leftarrow \varepsilon,\ 0 < \varepsilon < 1$
$for (i = 1\ to\ epochs_max)$
  $s_i \leftarrow create_neighbor_solution(s_{current})$
  $if\ f(s_i) \geq f(s_{current})$
    $s_{current} \leftarrow s_i$
    $if\ f(s_i) \geq f(s_{best})$
      $s_{best} \leftarrow s_i$
    $end\ if$
  $elseif\ \exp\left(\frac{f(s_{current}) - f(s_i)}{T}\right) > uniform_random(0,1)$
    $s_{current} \leftarrow s_i$
  $end\ if$
  $T \leftarrow c \times T$
$end\ for$
$return(s_{best})$

Рис. 2. Псевдокод алгоритму симуляції відпалу

Аналіз псевдокоду на рис. 2 показує, що поточне рішення $S_{current,}$, відносно якого відбувається пошук нових кращих рішень S_{best}, оновлюється у випадку знаходження нового рішення, що збільшує критерій (3) або випадковим чином з розподілу Гіббса. При цьому початкове рішення утворюється за допомогою процедури create_initial_solution. Для формування нового рішення пропонується використовувати найпростіший неадаптивний алгоритм, який можна представити у вигляді формули [20]

$$s_{current} = s_{current} + \\ +uniform_random(-1,1) \cdot step_size, \quad (15)$$

де uniform_random – функція генерації випадкових чисел з рівномірного розподілу з заданого діапазону; step_size – це розмір діапазону пошуку нових рішень, сусідніх з $S_{current}$.

Алгоритм придушення немаксимумів використовується для фільтрації непотрібних спрацювань детектора об'єктів на один і той самий об'єкт на зображенні [13, 14].

Таким чином ми пропонуємо модель детектора об'єктів на зображені та алгоритм його навчання, який побудовано на базі різних технік з метою максимізації контекстної інформації отриманої з обмеженого обсягу розмічених навчальних даних та в умовах обмежених обчислювальних ресурсів.

Результати фізичного моделювання

Для навчання детектора об'єктів було використано 200 зображень з розміченого набору даних Inria Aerial Image Labeling Dataset [21]. Кожне зображення має роздільну здатність 5000×5000. При цьому 500 нерозмічених зображень з роздільною здатністю 224x224 було згенеровано з використанням техніки випадкового вирізання з поворотом для навчання без вчителя. Також 200 розмічених зображень з роздільною здатністю 224x224 було згенеровано для навчання з вчителем. Розмічений набір навчальних даних було аугментовано до 1000 зразків за допомогою додавання до зображень шуму, повороту та вирізання, зміни їхнього контрасту.

В наборі навчальних даних Inria Aerial Image Labeling Dataset представлено велику кількість транспортних засобів у міській місцевості. Транспортні засоби були обрані як об'єкти інтересу, при цьому міська зона розглядається як доменна область застосування. В цьому випадку алфавіт класів розпізнавання дорівнює Z=3, де перший клас розпізнавання відповідає автомобілям, другий клас розпізнавання відповідає вантажівкам і третій – фоновим зображенням. Розмір об'єкту інтересу на випадковому зображенні знаходиться в діапазоні [7×7,….10×10] пікселів.

Відповідно до техніки переносу знань перші 7 fire модулів попередньо навченої згорткової нейронної мережі Squeezenet були запозичені. В результаті кожне зображення було перекодовано в карту ознак з розміром 13×13×384.

Наступні шари навчаються без вчителя на нерозміченому наборі навчальних даних з цільової доменної області. При цьому ядра фільтрів рівні 3x3, крок сканування – 1.

Вихідна карта ознак сформована шляхом об'єднання карт ознак, сформованих модулями Fire6 та Fire7, а також карти ознак, сформованої останнім згортковим шаром.

Спочатку пропонується виконувати навчання детектора з попередньо навченим останнім шаром без вчителя з використанням зростаючого розрідженого кодуючого нейронного газу без тонкої настройки. При цьому, гіперпараметр алгоритму навчання без вчителя рівний ν=0,8. В алгоритмі навчання інформаційно-екстремального класифікатора пікселів карти ознак число вузлів в деревах рішень обмежено до 16. При цьому глибина кожного дерева встановлюється на рівні 6.

З метою покращення результатів машинного навчання детектора, інформативність ознакового опису підвищена шляхом використанню тонкої настройки високорівневих згорткових шарів. В цьому випадку поточні параметри алгоритму симуляції відпалу були встановлені таким чином: c=0,98, T0=10, epochs_max=6000, step_size=0,001. На кожному етапі тонкої настройки виконується перенавчання регресора та класифікатора. Для максимізації узагальнюючої здатності моделі та мінімізації

обчислювальної складності, було здійснено послідовну оптимізацію гіперпараметра ν, що відповідає за щільність розподілу нейронів, з кроком оптимізації 0,1.

Попереднє навчання без вчителя верхніх згорткових шарів на нерозмічених даних з цільової доменної області спрямоване на підвищення ефективності машинного навчання на етапі навчання з учителем.

Варто розглянути вплив параметрів алгоритму зростаючого розріджено кодуючого нейронного газу, використаного на етапі навчання без вчителя, на результати навчання з учителем. У таблиці 1 наведено результати машинного навчання та кількість N_c згенерованих згорткових фільтрів (нейронів) як функцій від параметра ν, що характеризує щільність покриття всього набору навчальних даних згортковими фільтрами.

Аналіз табл. 1 показує, що зі зростанням значення гіперпараметра ν (2) зростає кількість нейронів та значення часткових і загального критеріїв оптимізації.

При значенні $\nu \le 0{,}8$ точність моделі за тестовим набором даних зростає зі збільшенням параметра ν, проте подальше збільшення цього параметра призводить до погіршення результатів через перенавчання моделі. В той же час вибір головних компонент для кожного значення ν був виконаний у відповідності з критерієм Кайзера: вибір лише основних компонентів з власними значеннями, що перевищують 1.

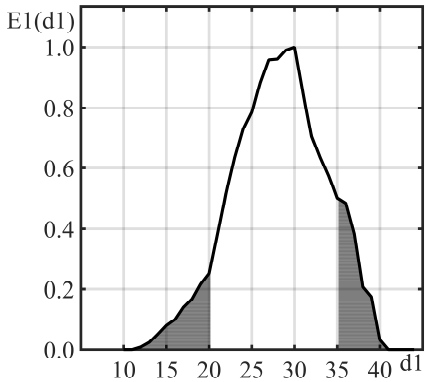
Таблиця 1

Результати навчання без вчителя детектора об'єктів з використанням різних значень параметра ν

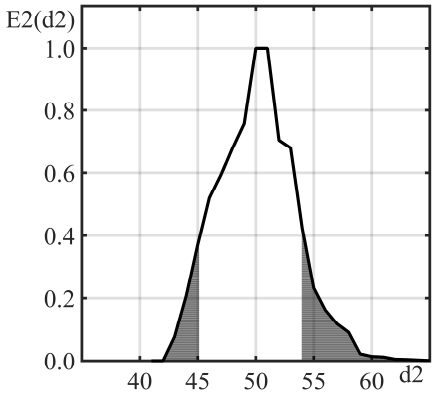
ν	N_c	J_{Cls}	J_{Loc}	J	Відсоток розпізнаних об'єктів на тестовій вибірці, %
0,4	196	0,0912	0,450	0,04104	0,67
0,5	337	0,1502	0,511	0,07675	0,73
0,6	589	0,2508	0,621	0,15574	0,87
0,7	889	0,4203	0,700	0,29421	0,93
0,8	1519	1,0000	0,921	0,92100	0,96
0,9	3058	1,0000	0,923	0,92300	0,95

Рисунок 3 показує залежність інформаційного критерію ефективності навчання (5) від радіусу контейнеру для кожного з класів. За даними графіками може бути оцінена точність та завадостійкість побудованих вирішальних правил.

Аналіз рис. 3 показує, що отримані вирішальні правила є безпомилковими за навчальним набором даних. Оптимальні радіуси контейнерів класів розпізнавання мають такі значення $d_1^* = 30$, $d_2^* = 50$ відповідно. Відстань між центрами контейнерів класів розпізнавання дорівнює $d_{12}^* = 95$.



а



б

Рис. 3. Графіки залежності інформаційного критерію ефективності навчання (5) від радіусу контейнеру для кожного з класів:

а – клас X_1^o; б – клас X_2^o

Таким чином, було отримано високодостовірні вирішальні правила в умовах обмеженого обсягу навчальних даних, які забезпечують отримання компактного розподілу в просторі ознак та значного міжкласового зазору.

Висновки

Наукова новизна одержаних результатів в даній роботі полягає у розробці нових моделі та алгоритму навчання бортової системи детектування малорозмірних об'єктів для малогабаритних безпілотних літальних апаратів, що здатна функціонувати в умовах обмежених обчислювальних ресурсів та навчальних даних. Розроблена модель складається з 7 перших модулів обчислювально ефективної згорткової мережі Squeezenet, одного згортковго розрі-

джено кодуючого шару, шару об'єднання з масштабуванням, перетворювача головних компонентів, інформаційно-екстремального класифікатора та регресійної моделі екстремального навчання. Такий підхід забезпечує використання більшої кількості контекстної інформації для ефективного розпізнавання малорозмірних об'єктів на аерофотознімках.

Розроблений алгоритм навчання складається з таких етапів: навчання без вчителя додаткових високорівневих згорткових розріджено кодуючих шарів, перетворення методом головних компонент вихідної карти ознак, інформаційно-екстремальне навчання класифікатора, інкрементальне навчання регресійної моделі та тонке налаштування верхніх шарів моделі за допомогою метаевристичного алгоритму симуляції відпалу. Було показано, що застосування навчання без вчителя дозволяє зменшити ефект перенавчання та більш ніж у 10 разів прискорити знаходження глобального оптимуму під час тонкої настройки моделі на навчальних даних обмеженого обсягу.

Практична цінність одержаних результатів полягає у зменшені вимог щодо обсягу розмічених навчальних даних та обчислювальних ресурсів. Розроблена модель забезпечує прийнятну для практичного застосування точність детектування об'єктів, розмір яких в 7-14 разів менший за розмір найменшої сторони аерофотознімку. При навчанні пропонованої моделі детектування об'єктів було використано 500 нерозмічених та 200 розмічених зразків навчальних даних з набору навчальних даних Inria Aerial Image Labeling dataset. При цьому точність розпізнавання об'єктів на тестових зображеннях складає 96%.

Публікація містить результати досліджень, проведених за грантом Президента України за конкурсним номером ***№ Ф75/144–2018*** *Державного фонду фундаментальних досліджень.*

Література

1. Subbotin, S. O. The special deep neural network for stationary signal spectra classification [Text] / S. O. Subbotin // Proceedings of the 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET) – Lviv, Slavske, Ukraine, 20-24 Feb, 2018. – IEEE, 2018. – P. 123–128.
DOI:10.1155/2017/3296874.

2. Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles [Text] / A. Carrio, C. Sampedro, A. Rodriguez-Ramos, P. Campoy // Journal of Sensors. – 2017. – Vol. 2017. – P. 1–13.

3. Xu, X. Scaling for edge inference of deep neural networks [Text] / X. Xu, Y. Ding, S. X. Hu // Nature Electronics. – 2018. – Vol. 4, No. 4. – P. 216-222. DOI:10.1038/s41928-018-0059-3.

4. DroNet: Learning to Fly by Driving [Text] / A. Loquercio, A. I. Maqueda, C. R. del-Blanco, D. Scaramuzza // IEEE Robotics and Automation Letters. – 2018. – Vol. 3, No. 2. – P. 1088-1095.
DOI: 10.1145/2733373.2806332.

5. An improved transfer learning approach for intrusion detection [Text] / A. Mathew, J. Mathew, M. Govind, A. Mooppan // Procedia Com-puter Science. – 2017. – Vol. 115. – P. 251-257.
DOI: 10.1016/j.procs.2017.09.132.

6. Radenović, F. Fine-tuning CNN Image Retrieval with No Human Annotation [Electronic resource] / F. Radenović, G. Tolias, O. Chum – Available at: https://ieeexplore.ieee.org/document/8382272. – 2.12.2018. DOI: 10.1109/TPAMI.2018.2846566.

7. Improving the effectiveness of training the onboard object detection system for a compact unmanned aerial vehicle [Text] / V. V. Moskalenko, A. S. Dovbysh, I. V. Naumenko, A. S. Moskalenko, A. G. Korobov // Eastern-European Journal of Enterprise Technologies. – 2018. – Vol. 4/9, No. 94. – P. 19-26.
DOI: 10.15587/1729-4061.2018.139923.

8. Vens, C. Random Forest Based Feature Induction [Text] / C. Vens, F. Costa // Proceedings of the IEEE 11th International Conference on Data Mining – Vancouver, Canada, 11-14 Dec, 2011. – IEEE, 2011. – P. 744–753. DOI: 10.1109/ICDM.2011.121.

9. Feng, Q. Compressed auto-encoder building block for deep learning network [Text] / Q. Feng, C. L. P. Chen, L Chen // Proceedings of the 3rd International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS) – Jinzhou, Liaoning, China, 26-29 Aug, 2016. – P. 131–136.
DOI: 10.1109/ICCSS.2016.7586437.

10. Labusch, K. Sparse coding neural gas: learning of overcomplete data representations [Text] / K. Labusch, E. Barth, T. Martinetz // Neurocomputing. – 2009 – Vol. 72, No. 7–9 – P. 1547–1555.
DOI: 10.1016/j.neucom.2008.11.027.

11. Mrazova, I. Image Classification with Growing Neural Networks [Text] / I. Mrazova, M. Kukacka // International Journal of Computer Theory and Engineering. – 2013. – Vol. 5, No. 3 – P. 422-427. DOI:10.7763/IJCTE.2013.V5.722.

12. Palomo, J. The Growing Hierarchical Neural Gas Self-Organizing Neural Network [Text] / J. Palomo, E. López-Rubio // IEEE Transactions on Neural Networks and Learning Systems. – 2017. – Vol. 28, No. 9. – P. 2000-2009. DOI:10.1109/TNNLS.2016.2570124.

13. Nakahara, H. An object detector based on multiscale sliding window search using a fully pipelined binarized CNN on an FPGA [Text] / H. Nakahara, H. Yonekawa, S. Sato // International Conference on Field Programmable Technology (ICFPT) – Melbourne, VIC, 11-13 Dec, 2018 – IEEE Publisher, 2018. – P. 168–175. DOI: 10.1109/FPT.2017.8280135.

14. Aircraft Detection by Deep Convolutional Neural Networks [Text] / X. Chen, S. Xiang, C. L. Liu, C. H. Pan // IPSJ Transactions on Computer Vision and Applications. – 2015. – Vol. 7. – P. 10-17. DOI: 10.2197/ipsjtcva.7.10.

15. Zou, W. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incremental Extreme Learning Machine Based on Driving Amount [Text] / W. Zou, Y. Xia, H .Li // IEEE Transactions on Cybernetics. – 2018. – Vol. 48, No. 12. – P. 3403–3410. DOI: 10.1109/TCYB.2018.2830338.

16. Rere, R. L. M. Metaheuristic Algorithms for Convolution Neural Network [Text] / R. L. M. Rere, M. I. Fanany, A. M. Arymurthy // Computational Intelligence and Neuroscience. – 2017. – Vol. 2016 – P. 1–13. DOI:10.1155/2016/1537325.

17. Patricia, N. Learning to Learn, from Transfer Learning to Domain Adaptation: A Unifying Perspective [Text] / N. Patricia, B. Caputo // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Columbus, Ohio, 23-28 June, 2014. – Conference Publishing Services, 2014. – P. 1442–1449. DOI: 10.1109/CVPR.2014.187.

18. Nguyen, A. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images [Text] / A. Nguyen, J. Yosinski, J Clune // IEEE Conference on Computer Vision and Pattern Recognition, (CVPR). – Boston, MA, 7-12 June, 2015 – IEEE Publisher 2015. – P. 427–436. DOI: 10.1109/CVPR.2015.7298640.

19. Optimization of convolutional neural network using microcanonical annealing algorithm [Text] / V. Ayumi, L. M. R. Rere, M. I. Fanany, A. M. Arymurthy // International Conference on Advanced Computer Science and Information Systems (ICACSIS). – Malang, Indonesia, 15-16 Oct., 2016: – IEEE Publisher 2016. – P. 506–511. DOI: 10.1109/ICACSIS.2016.7872787.

20. High-Resolution Aerial Image Labeling With Convolutional Neural Networks [Text] / E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez // IEEE Transactions on Geoscience and Remote Sensing. – 2017. – Vol. 55, No. 12. – P. 7092–7103. DOI: 10.1109/TCYB.2018.2830338.

References

1. Subbotin, S. O. The special deep neural network for stationary signal spectra classification. *14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 20-24 Feb, 2018, pp. 123–128. DOI:10.1155/2017/3296874.

2. Carrio, A., Sampedro, C., Rodriguez-Ramos, A., Campoy, P. A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles. *Journal of Sensors*, 2017, vol. 2017, pp. 1–13. DOI: 10.1155/2017/3296874.

3. Xu, X., Ding, Y., Hu, S. X. Scaling for edge inference of deep neural networks. *Nature Electronics*, 2018, vol. 1, no. 4, pp. 216-222. DOI:10.1038/s41928-018-0059-3.

4. Loquercio, A., Maqueda, A. I., Del-Blanco, C. R., Scaramuzza, D. DroNet: Learning to Fly by Driving. *IEEE Robotics and Automation Letters*, 2018, vol. 3, no. 2, pp. 1088–1095. DOI: 10.1145/2733373.2806332.

5. Mathew, A., Mathew, J., Govind, M., Mooppan, A. An Improved Transfer learning Approach for Intrusion Detection. *Procedia Computer Science*, 2017, vol. 115, pp. 251–257. doi:10.1016/j.procs.2017.09.132.

6. Radenović, F., Tolias, G., Chum, O. Fine-tuning CNN Image Retrieval with No Human Annotation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018. Available at: https://ieeexplore.ieee.org/document/8382272. (Accessed 03.12.2018). DOI: 10.1109/TPAMI.2018.2846566.

7. Moskalenko, V. V., Dovbysh, A. S., Naumenko, I. V., Moskalenko, A. S., Korobov, A. G. Improving the effectiveness of training the on-board object detection system for a compact unmanned aerial vehicle. *Eastern-European Journal of Enterprise Technologies*, 2018, vol. 4/9 no. 94, pp. 19–26. DOI: 10.15587/1729-4061.2018.139923.

8. Vens, C., Costa, F., Random Forest Based Feature Induction. *IEEE 11th International Conference on Data Mining*, Vancouver, Canada, 11-14 Dec, 2011, pp. 744–753. DOI: 10.1109/ICDM.2011.121.

9. Feng, Q., Chen, C. L. P., Chen, L. Compressed auto-encoder building block for deep learning network. *3rd International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS)*, Jinzhou, Liaoning, China, 26-29 Aug, 2016, pp. 131–136. DOI: 10.1109/ICCSS.2016.7586437.

10. Labusch, K., Barth, E., Martinetz, T. Sparse coding neural gas: learning of overcomplete data representations. *Neurocomputing*, 2009, vol. 72, no. 7–9, pp. 1547–1555. DOI:10.1016/j.neucom.2008.11.027.

11. Mrazova, I., Kukacka, M. Image Classification with Growing Neural Networks. *International Journal of Computer Theory and Engineering*, 2013, vol. 5, no. 3, pp. 422–427. DOI:10.7763/IJCTE.2013.V5.722.

12. Palomo, J., López-Rubio, E. The Growing Hierarchical Neural Gas Self-Organizing Neural Network. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, vol. 28, no. 9, pp. 2000–2009. DOI: 10.1109/TNNLS.2016.2570124.

13. Nakahara, H., Yonekawa, H., Sato, S. An object detector based on multiscale sliding window search using a fully pipelined binarized CNN on an FPGA. *International Conference on Field Programmable Technology (ICFPT)*, Melbourne, 11-13 December, 2018, pp. 168–175. DOI: 10.1109/FPT.2017.8280135.

14. Chen, X., Xiang, S., Liu, C.-L., Pan, C.-H. Aircraft Detection by Deep Convolutional Neural Networks. *IPSJ Transactions on Computer Vision and Applications*, 2015, vol. 7, pp. 10–17. DOI: 10.2197/ipsjtcva.7.10.

15. Zou, W., Xia, Y., Li, H. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incre-

mental Extreme Learning Machine Based on Driving Amount. *IEEE Transactions on Cybernetics*, 2018, vol. 48, no. 12, pp. 1-8. DOI: 10.1109/TCYB.2018.2830338.

16. Rere, R. L. M., Fanany, M. I., Arymurthy, A. M. Metaheuristic Algorithms for Convolution Neural Network. *Computational Intelligence and Neuroscience*, 2017, vol. 2016, pp. 1–13. DOI:10.1155/2016/1537325.

17. Patricia, N., Caputo, B. Learning to Learn, from Transfer Learning to Domain Adaptation: A Unifying Perspective. IEEE *Conference on Computer Vision and Pattern Recognition (CVPR),* Columbus, Ohio, 23-28 June, 2014, pp. 1442–1449. DOI: 10.1109/CVPR.2014.187.

18. Nguyen, A., Yosinski, J., Clune, J. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. *IEEE Conference on Computer Vision and Pattern Recognition, (CVPR),* Boston, MA, 7-12 June, 2015, pp. 427–436. DOI: 10.1109/CVPR.2015.7298640.

19. Ayumi, V., Rere, L. M. R., Fanany, M. I., Arymurthy, A. M. Optimization of convolutional neural network using microcanonical annealing algorithm. *International Conference on Advanced Computer Science and Information Systems (ICACSIS),* Malang, Indonesia, 15-16 Oct., 2016, pp. 506–511. DOI: 10.1109/ICACSIS.2016.7872787.

20. Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P. High-Resolution Aerial Image Labeling With Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, vol. 55, no. 12, pp. 7092–7103. DOI:10.1109/TGRS.2017.2740362.

Надійшла до редколегії 18.11.2018 розглянута на редколегії 12.12.2018

МЕТОД И АЛГОРИТМ ОБУЧЕНИЯ СИСТЕМЫ ДЕТЕКТИРОВАНИЯ МАЛОРАЗМЕРНЫХ ОБЪЕКТОВ ДЛЯ МАЛОГАБАРИТНЫХ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

В. В. Москаленко, А. С. Москаленко, А. Г. Коробов, М. О. Зарецкий, В. А. Семашко

В работе разработаны эффективные модель и алгоритм обучения бортовой системы малогабаритного беспилотного летательного аппарата для детектирования малоразмерных объектов, функционирующий в условиях ограниченного объема вычислительных ресурсов и обучающих данных. Предлагается 4-х этапная реализация алгоритма обучения. Первый этап заключается в выборе типа глубокой свёрточной сети, заранее обученной на наборе данных ImageNet, и количества ее низкоуровневых слоев, которые заимствуются для экстракции информативных признаков. Второй этап включает обучение без учителя высокоуровневых разреженно кодирующих свёрточных слоёв с использованием модифицированного алгоритма растущего нейронного газа, обеспечивающего оптимальное распределение нейронов по учебных данных и автоматическое определение необходимого количества нейронов. Применение модифицированного алгоритма растущего нейронного газа позволяет эффективно использовать неразмеченные обучающие данные и адаптировать верхние слои экстрактора признаков к доменной области применения. На третьем этапе исходная карта признаков формируется путем объединения карт признаков из разных слоев глубокой свёрточной нейронной сети. При этом осуществляется уменьшение глубины исходной карты признаков на основе анализа главных компонент с последующим построением решающих правил. Этот подход позволяет использовать всю имеющуюся контекстную информацию, обеспечивающую эффективное распознавания малоразмерных объектов на аэрофотоснимках. Для выполнения классификационного анализа результирующей карты признаков предложено использовать информационно-экстремальный классификатор, обучающийся по принципу бустинга. Кроме того, для прогнозирования координат малоразмерных объектов в кадре изображения предлагается использование регрессионного алгоритма ортогональной инкрементальной машины экстремального обучения. Последний этап обучения детектора предусматривает тонкую настройку высокоуровневых слоев глубокой нейронной сети с использованием метаэвристического алгоритма симуляции отжига с целью приближения глобального оптимума комплексного критерии эффективности обучения детектора. В результате применения предложенного подхода удалось достичь 96% точности обнаружения объектов интереса на тестовых изображениях открытого набора данных, что свидетельствует о пригодности модели и метода обучения для практического использования. При этом размер обучающей выборки, которая использовалась для построения модели составляет 500 неразмеченных и 200 размеченных образцов.

Ключевые слова: растущий нейронный газ; свёрточная нейронная сеть; детектирование объектов; информационный критерий; алгоритм симуляции отжига; экстремальное обучение.

A MODEL AND LEARNING ALGORITHM OF SMALL-SIZED OBJECT DETECTION SYSTEM FOR COMPACT DRONES

V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, M. O. Zaretsky, V. A. Semashko

The efficient model and learning algorithm of the small object detection system for compact aerial vehicle under conditions of restricted computing resources and the limited volume of the labeled learning set are developed.

The four-stage learning algorithm of the object detector is proposed. At the first stage, selecting the type of deep convolutional neural network and the number of low-level layers that is pretrained on the ImageNet dataset for reusing takes place. The second stage involves unsupervised learning of high-level convolutional sparse coding layers using the modification of growing neural gas to automatically determine the required number of neurons and provide optimal distributions of the neurons over the data. Its application makes it possible to utilize the unlabeled learning datasets for the adaptation of the high-level feature description to the domain application area. At the third stage, the output feature map is formed by concatenation of feature maps from the different level of the deep convolutional neural network. At that, there is a reduction of output feature map using principal component analysis and followed by the building of decision rules. In order to perform the classification analysis of output, feature map is proposed to use information-extreme classifier learning on principles of boosting. Besides that, the orthogonal incremental extreme learning machine is used to build the regression model for the predict bounding box of the detected small object. The last stage involves fine-tuning of high-level layers of deep network using simulated annealing metaheuristic algorithm in order to approximate the global optimum of the complex criterion of learning efficiency of detection model. As a result of the use of proposed approach has been achieved 96% correctly detection of objects on the images of the open test dataset which indicates the suitability of the model and learning algorithm for practical use. In this case, the size of the learning dataset that has been used to construct the model was 500 unlabeled and 200 labeled learning samples.

Keywords: growing neural gas; convolutional neural network; objects detector; information criterion; simulated annealing algorithm; extreme learning.

Москаленко В'ячеслав Васильович – канд. техн. наук, доцент каф. комп'ютерних наук, Сумський державний університет.

Москаленко Альона Сергіївна – асистент каф. комп'ютерних наук Сумського державного університету.

Коробов Артем Геннадійович – аспірант каф. комп'ютерних наук, Сумський державний університет.

Зарецький Микола Олександрович – аспірант каф. комп'ютерних наук, Сумський державний університет.

Семашко Віктор Анатолійович – аспірант каф. комп'ютерних наук, Сумський державний університет.

Moskalenko Viacheslav Vasilyovich – PhD, associate professor of Computer Sciences Chair of Sumy State University, e-mail: v.moskalenko@cs.sumdu.edu.ua, ORCID Auther ID: 0000-0001-6275-9803.

Moskalenko Alyona Sergiyvna – PhD, teaching assistant of Computer Sciences Chair of Sumy State University, e-mail: a.moskalenko@cs.sumdu.edu.ua, alenarizhova@gmail.com, ORCID Auther ID: 0000-0003-3443-3990.

Korobov Artem Gennadiyovich – PhD student of Computer Sciences Chair of Sumy State University, e-mail: a.korobov@cs.sumdu.edu.ua, ORCID Auther ID: 0000-0003-3239-1977.

Zaretsky Mikolay Alexandrovich – PhD student of Computer Sciences Chair of Sumy State University, e-mail: n.zaretskij@gmail.com, ORCID Auther ID: 0000-0001-9117-5604.

Semashko Victor Anatolyevich – PhD student of Computer Sciences Chair of Sumy State University, e-mail: viktor.s.5994@gmail.com, ORCID Auther ID: 0000-0002-9765-876X.