УДК 004.89

І.О. Складчиков , студент гр. ПК-11мп, доктор філософії, ас. А. С. Момот
КПІ ім. Ігоря Сікорського

АНАЛІЗ ЕФЕКТИВНОСТІ МОДЕЛЕЙ ГЛИБИННОГО НАВЧАННЯ ДЛЯ АВТОМАТИЗОВАНОГО ДЕТЕКТУВАННЯ ОБ'ЄКТІВ НА ІНФРАЧЕРВОНИХ ЗОБРАЖЕННЯХ

Анотація. В даній роботі розглянуто існуючі методи глибинного навчання для автоматизованого виявлення об’єктів на інфрачервоних зображеннях. Проаналізовано актуальність застосування даних моделей на сьогоднішній день та пріоритетність точності розпізнавання об’єктів над швидкістю. Розглянуто мережі R-CNN, YOLO та EfficientDet. Проаналізовано ефективність даних моделей та зроблено висновки щодо можливості їх використання в системах термографічного контролю безпеки.

Ключові слова: глибинне навчання, виявлення об'єктів, інфрачервоні зображення.

ВСТУП

Тепловізійні камери міцно зайняли своє місце серед засобів комп'ютерного бачення. Вони відмінно вирішують поставлені перед ними завдання і дозволяють створювати технічно і економічно ефективні системи. Інфрачервоні камери пройшли довгий шлях від дуже дорогих пристроїв до продукту широкого використання з високою надійністю і доступною ціною, і вони продовжують вдосконалюватися з кожним роком. На сьогоднішній день, спектр застосування тепловізійних систем досить різноманітний: діагностика та ремонт друкованих плат, енергетична галузь, неруйнівний контроль композиційних матеріалів, технічні засоби контролю безпеки, медицина, мисливська та військова промисловості, автоматизоване управління.

Доцільність використання тепловізорів визначається особливостями самих технічних засобів. У системах контролю безпеки інфрачервоні камери – це пристрої спостереження, які більш ефективні для розпізнавання об'єктів, ніж традиційні відеокамери. Якість розпізнавання за допомогою тепловізора не залежить від часу доби. Це дає можливість створювати системи цілодобового моніторингу об'єкта, що охороняється. Системи, призначені для контролю безпеки, застосовуються для захисту периметра аеропортів, залізниць, кордонів, будинків, агрокомплексів та інших об'єктів. Аналіз отриманої інформації може бути автоматизований за допомогою технологій штучного інтелекту [1].

МЕТА РОБОТИ

Тепловізійне обладнання в даний час забезпечує повний і точний аналіз поведінки на основі алгоритмів глибинного навчання. Це включає виявлення таких подій: перетин лінії, вторгнення, вхід до зони, вихід із зони тощо. Інтелектуальне виявлення людини/транспортного засобу може зменшити частоту виникнення помилкових сигналів тривоги, викликаних тваринами, вібрацією, падаючим листям та іншими несуттєвими подіями або об'єктами. Це дозволяє значно підвищити точність роботи системи та алгоритму формування сигналу тривоги [2].

Актуальним і важливим завданням є визначення архітектури глибинної мережі, яка дозволить виявити об'єкт з найбільшою точністю. Оскільки сучасні інфрачервоні камери не мають високої частоти кадрів, швидкість розрахунків

моделі глибинного навчання відходить на другий план. Однак для моніторингу в режимі реального часу частота обробки даних не повинна бути нижче 20-25 Гц. Цей факт необхідно враховувати при виборі відповідної моделі для розпізнавання об'єктів. Крім того, завдання виявлення об'єктів в інфрачервоних зображеннях ускладнюється низькою деталізацією теплових відбитків цих об'єктів, що пов'язано з фізичними особливостями природи інфрачервоного випромінювання [3]. Тому модель глибинного навчання повинна бути в цілому узагальнюючою та мати високу адаптивність.

Для оцінки якості моделей детектування об'єктів використовується метрика mAP (mean Average Precision) – показник середньої коректності розпізнавання різних класів об’єктів. Цей показник обчислює середню достовірність правильних відповідей в діапазоні від 0 до 1 для всіх класів об'єктів, які вміє розпізнавати модель. Для кожного класу середня коректність розпізнавання (Average Precision) визначається як площа області під кривою Precision-Recall. Чим вище значення цього показника, тим менше помилкових розпізнавань модель виконує на тестових даних.

МАТЕРІАЛИ І РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Однією з перших моделей, успішно застосованих в задачах виявлення об'єктів на зображеннях, була мережа R-CNN (Regions with CNNs) [4]. У цій архітектурі на вхід мережі подається не все зображення одночасно, а окремі його регіони. Регіони обираються раніше за допомогою алгоритму вибіркового пошуку, який визначає ділянки зображення що, ймовірно, містять деякі об'єкти. В даний час модифікація архітектури Faster R-CNN забезпечує значення mAP до 35% при швидкості обробки даних близько 15 кадрів в секунду на наборі даних MS COCO [5]. Такі характеристики є відносно поганим показником для використання в термографічних системах контролю безпеки у реальному часі. Існує також модифікація Mask-RCNN, яка дозволяє не тільки виявляти об'єкти, але і вирішувати завдання семантичної сегментації зображення. Однак, швидкість роботи такої архітектури залишається невеликою.

Набагато більш високу продуктивність має мережева архітектура YOLO (You Only Look Once) [6]. Головна її особливість – це розпізнавання об'єктів за один прохід даних. В архітектурі YOLO немає явних циклів, що робить мережу швидкодіючою. YOLO використовує сітку попередньо визначених вікон – областей, в яких класифікуються об'єкти. На наборі даних MS COCO сучасні модифікації YOLO показують до 44% mAP зі швидкістю обробки даних до 70 кадрів в секунду [7]. Мережі цієї архітектури є одними з найшвидших в завданні детектування об'єктів, що робить їх перспективними для використання у складі термографічних систем реального часу. На рисунку 1 представлено приклад застосування мережі YOLOv3 для виявлення людини і собаки на інфрачервоному зображенні, наданому компанією Flir.

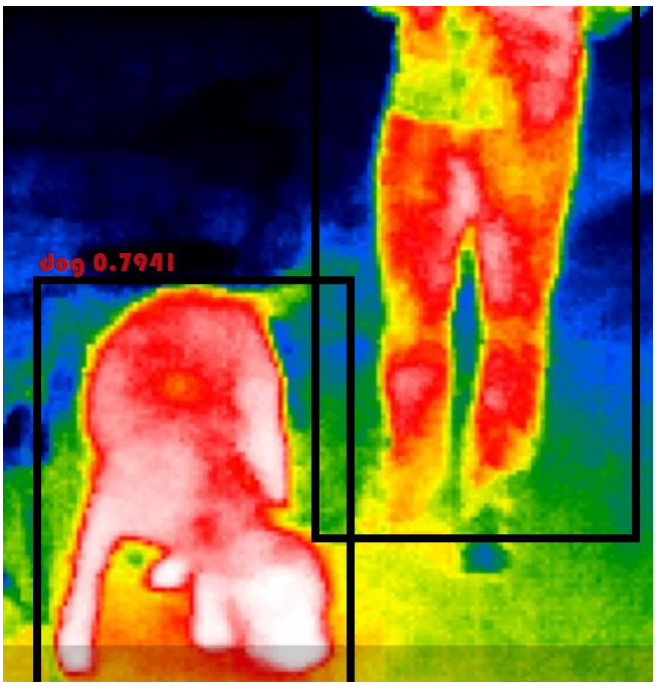

Рис.1. Виявлення людини та собаки на інфрачервоному зображенні за допомогою моделі YOLOv3

Ще одним перспективним класом нових моделей є EfficientNet, що виникли в результаті вивчення алгоритмів масштабування і балансування глибини і ширини (кількості каналів) згорткових глибинних мереж, а також впливу зміни роздільної здатності зображень на показники ефективності моделі. Автори роботи [8] пропонують новий складний метод масштабування, який рівномірно масштабує глибину/ширину/роздільну здатність з фіксованими пропорціями між ними. З існуючого методу під назвою «Пошук нейронної архітектури» для автоматичного генерації нових мереж і власного методу масштабування автори виводять новий клас моделей під назвою EfficientNets. Остання модифікація, названа EfficientDet, призначена для виявлення об’єктів на зображеннях і складається з EfficientNet як бази, до якої додається функціональний пірамідальний блок виявлення ознак BiFPN, а потім і «стандартний» класифікатор кадру. Ця мережа показує mAP до 46% при швидкості обробки даних близько 35 кадрів в секунду на наборі MS COCO [7]. Такі характеристики є оптимальним поєднанням швидкодії та точності детектування і розпізнавання об’єктів. Окрім того, мережі EfficientDet мають меншу кількість внутрішніх параметрів, ніж попередньо розглянуті архітектури, а отже і навчаються швидше.

ВИСНОВКИ

З огляду на швидкий розвиток глибокого навчання, описані моделі є актуальними тільки на момент публікації даної статті. Однак розглянуті класи моделей глибинних мереж вже давно є лідерами в області виявлення об'єктів. Усі ці моделі рекомендуються до розгляду в якості алгоритмів обробки інфрачервоних зображень у складі термографічних систем контролю безпеки. Оскільки в більшості областей застосування таких систем точність розпізнавання є більш пріоритетною, ніж швидкість, моделі EfficientDet представляють особливий інтерес. Слід зазначити, що порівняння ефективності

цих мереж здійснювалось на наборі зображень у видимому спектрі світла. Тому основним завданням для подальших досліджень є порівняння згаданих архітектур на наборі інфрачервоних зображень для отримання кількісних оцінок якості їх роботи.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Galagan R. M. The use of ART-2 neural network for processing information signals of non-destructive testing / R. M. Galagan, A. S. Momot. // Proceedings of the IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). – 2017. – С. 981–985.

[2] Aggregating Deep Convolutional Feature Maps for Insulator Detection in Infrared Images / Z. Zhao, X. Fan, G. Xu та ін. // IEEE Access – 2017. – №5.- С. 21831–31839.

[3] Momot A. Deep learning automated data analysis of security infrared cameras / A. Momot, I. Skladchykov // Slovak international scientific journal. – 2021. – №52. – pp. 13–16.

[4] Girshick R. Rich feature hierarchies for accurate object detection and semantic segmentation / R. Girshick, J. Donahue, J. Malik. // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 580–587.

[5] Derakhshani M. Assisted excitation of activations: A learning technique to improve object detectors / M. Derakhshani, S. Masoudnia, A. Shaker. // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – С. 9201–9210.

[6] Du J. Understanding of Object Detection Based on CNN Family and YOLO / Du. // Journal of Physics: Conference Series,. – 2018. – №1004. – С. 12–29.

[7] Bochkovskiy A. Yolov4: Optimal speed and accuracy of object detection / A. Bochkovskiy, W. Chien-Yao, M. Hong-Yuan. // arXiv preprint arXiv:2004.10934. – 2020.

[8] Mingxing T. Efficientnet: Rethinking model scaling for convolutional neural networks / T. Mingxing, Q. Le. // International Conference on Machine Learning. PMLR. – 2019. – С. 6105–6114.

Наук. керівник – доктор філософії, ас. Момот А.С.