

Використання трансфертного навчання в задачі класифікації та локалізації крилатих ракет

В роботі розглянуто проблему виявлення та локалізації крилатих ракет. Проведено аналіз архітектур згорткових нейронних мереж для вирішення цього завдання. Представлено приклад застосування архітектури мережі YOLO та трансферного навчання для локалізації крилатих ракет. Показано, що методи машинного навчання для виявлення ракет і їх локалізація є перспективними. Запропонований підхід дає змогу отримати розв'язок задачі виявлення ракети та її напрямку з достатньо високим ступенем точності на модельних даних.

Вступ

У сучасному світі виявлення та локалізація крилатих ракет відіграє ключову роль у національній безпеці. Надзвичайно важливо, перехоплювати будь-які ракети до того, як вони завдадуть удару по об'єктах або цивільним особам.

Для вирішення цього завдання необхідно розглянути існуючі варіанти розпізнавання повітряних цілей. Більшість таких систем не ефективні для розпізнавання ракет та літаків, що їх випускають, це пов'язано з використанням стелс-технологій, які маскують характеристики літальних апаратів у зоні виявлення радарів.

Додавання або навіть заміна радіолокаційних систем виявлення літаків оптичними системами виявлення відкриває нові можливості у визначенні характеру повітряної обстановки.

Обробку відеозображень наземними моніторами повітряного руху можна розділити на ряд завдань. Серед них можна виділити такі завдання:

- виявлення рухомого літаючого об'єкта;
- визначення характеристик об'єкта (дальність, розміри, швидкість, маневреність тощо);
- визначення типу літаючого об'єкта.

Ці завдання можна вирішувати як поетапно, так і одночасно. Можливе виявлення всіх об'єктів на одному кадрі з визначенням їх типу і подальшим їх відстеженням в послідовності відеокадрів або виявлення рухомих об'єктів по відеоряду з подальшою ідентифікацією об'єктів на кожному кадрі. У запропонованому підході передбачається, що рухомий об'єкт або об'єкти спочатку будуть виявлені та локалізовані на відеозображенні, а потім буде вирішено завдання віднесення кожного об'єкта до певного класу.

Постановка проблеми

Класичні методи розпізнавання повітряних цілей, базуються на порівнянні невідомого зображення з бібліотекою відомих. Складність вирішення такого завдання полягає в правильному визначенні невідомого об'єкта,

незалежно від його орієнтації. Для подолання цієї складності використовується підхід з розрахунком числових значень (моментів), інваріантних до повороту зображення на площині.

Однак, революційні досягнення в області аналізу зображень, що були досягнуті після застосування глибоких згорткових нейронних мереж, навчених на тестовому наборі зображень ImageNet[1], дозволяють взяти до уваги їх застосування в завданні виявлення та ідентифікації летючих ракет. Адже, саме модифікації підходів, пов'язані з глибокими нейронними мережами використовувалися для вирішення задачі ідентифікації типів літаків [2].

У задачі класифікації нейронна мережа тим краще, чим точніше вона класифікує об'єкт, який їй подається вперше. Класичні повновзв'язні нейронні мережі не підходять для цього завдання через велику кількість зв'язків між нейронами, тому необхідно вибрати іншу архітектуру. Однією з найефективніших архітектур для вирішення цієї проблеми є згорткові нейронні мережі (CNN).

CNN — спеціалізований тип нейронної мережі для обробки даних, що має мережеву топологію. У цій статті розглядаються наступні архітектури загорнутих нейронних мереж:

- YOLOv7;
- VGG;
- ResNet;
- DenseNet з різною кількістю шарів.

За даними[3] на наборі зображень ImageNet VGG-13 дає похибку 10,75%, VGG-16 – 9,62%, а VGG-19 – 9,12%. З цього можна зробити висновок, що чим глибша мережа, тим вона ефективніша, але на практиці виявилось, що це не так: якість моделі досягла певної межі, а потім почала падати.

Автори ResNet змогли знайти таку топологію, що зі збільшенням глибини мережі її ефективність не знижується. Архітектура мережі ResNet[4] складається з шарів, які, у свою чергу, складаються з таких блоків. Останні два шари — це шар підвибірки та повністю підключений шар.

Щоб покращити інформаційний потік між шарами, автори запропонували дещо відмінну від ResNet схему: вони ввели прямі з'єднання з будь-якого рівня на всіх наступних. Завдяки такій щільності з'єднань в мережі він і отримав свою назву. Також було зменшено кількість карт функцій на перехідних шарах. Ця мережа має ту ж функцію, що й ResNet: збільшення глибини покращує точність моделі. Порівняно з ResNet, DenseNet був більш ефективним.

Однак вказані архітектури є не ефективними в плані часу відпрацювання системи, тому слід розглянути архітектуру YOLO. YOLO — це структура виявлення об'єктів, призначена для промислових додатків, з ефективною конструкцією та високою продуктивністю, орієнтована на апаратне забезпечення[5]. Наприклад, YOLOv6 досягає 35,0 mAP в наборі даних COCO val2017 з 1242 кадрами в секунду на T4 з використанням TensorRT FP6 для виведення bs32, а YOLOv6-s досягає 43,1 mAP в наборі даних COCO val2017 з 520 кадрами в секунду на T4.

Тому, для вибору найкращої архітектури для цього завдання було проведено порівняльні характеристики точності відпрацювання нейронної мережі та декількох стандартних наборах даних. В результаті було отримано наступну порівняльну таблицю 1:

Таблиця 1.

Точність класифікації зображень з тестової вибірки після навчання

Architecture CNN	MNIST	CIFAR-10	CIFAR-100	Fashion MNIST	LabelMe	SDSS17	Vegetable Image Dataset	Caltech 101
YOLOv7	0.9465	0.8391	0.6138	0.8619	0.9013	0.9124	0.9215	0.9425
VGG-16	0.9303	0.8398	0.6253	0.8650	0.9538	0.8237	0.8214	0.9605
VGG-19	0.9204	0.8515	0.6321	0.8632	0.9535	0.9287	0.7893	0.9786
ResNet 50	0.9221	0.8183	0.6181	0.8469	0.9409	0.9632	0.8450	0.9962
ResNet 101	0.9156	0.8031	0.5913	0.8485	0.9379	0.9636	0.8446	0.9837
ResNet 152	0.9169	0.8034	0.5894	0.8503	0.9512	0.9479	0.8427	0.9936
DenseNet 121	0.9379	0.8263	0.6119	0.8513	0.9612	0.9697	0.8106	0.9701
DenseNet 161	0.9458	0.8841	0.6719	0.8734	0.9639	0.9818	0.8803	0.9916

Також слід зазначити, що найефективнішими мережами по точності були ResNet і DenseNet, середня точність яких склала 86,19% і 89,91% відповідно. Проте час відпрацювання вказаних архітектур є не достатнім для завдання класифікації повітряної цілі (крилатої ракети), оскільки потрібна максимальна швидкість ідентифікації для передачі інформації органам влади. В кінцевому варіанті було обрано архітектуру з меншою точністю, проте швидшим часом відпрацювання, а саме YOLOv7.

Таблиця 2.

Час детектування об'єкта в залежності від архітектури нейронної мережі

Архітекту ра	YOLO v7	VGG -16	VGG -19	ResN et 50	ResN et 101	ResN et 152	Dense Net 121	Dense Net 161
Час детектува ння	23.1m s	31.5 ms	30.9 ms	27.8 ms	32.6 ms	35.8 ms	40.2ms	44.8ms

Проте, для навчання нейронної мережі “з нуля” потрібна велика кількість шуканих зображень та багато ресурсів для навчання. Однак, сучасні моделі розпізнавання зображень являють собою глибокі згорткові нейронні мережі з численними шарами. Нижні рівні такої мережі відповідають за розпізнавання елементарних ознак низького рівня, таких як краї, контури та частини зображення. Середні шари забезпечують розпізнавання груп ознак низького рівня, які, у свою чергу, є елементарними ознаками для наступних рівнів. Останній шар визначає категорію зображення.

Тому слід звернути увагу на трансферне навчання, яке у випадку класифікації зображень крилатої ракети, дозволяє перенавчити останній рівень такої мережі власним набором зображень за прийнятний час без зміни ваг інших шарів і досягнення необхідної точності. Це одна перевага трансферного навчання полягає в тому, що додаткове навчання останнього рівня згорткової мережі не вимагає стільки даних, скільки навчання з нуля, оскільки потрібно налаштувати менше ваг.

Вирішення проблеми

Рішенням описаних проблем визначення крилатої ракети та напрямку її польоту є використання нейронної мережі YOLOv7 та трансферного навчання, що дозволить використовувати попередньо навчену (навіть для інших класів об'єктів) мережу для побудови класифікаторів зображень повітряних летючих об'єктів.

У процесі навчання були «заморожені» ваги всіх нейронів моделі, попередньо навчених на зображеннях ImageNet, за винятком нейронів останнього шару. 1000 епох навчання дозволили отримати рішення задачі класифікації. Процес оцінки якості класифікації та помилок у ході навчання показано на рис. 1. Результати роботи на реальних зображеннях показано на рис. 2.

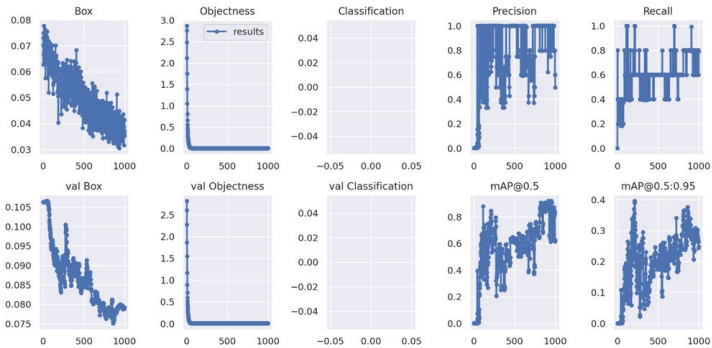


Рис. 1. Графік оцінки якості класифікації та помилок у процесі навчання



Рис. 2. Результати роботи класифікатора

Висновок

Методи машинного навчання для виявлення ракет і їх локалізація є перспективною. Запропонований підхід дає змогу отримати розв'язок задачі виявлення ракети та її напрямку з достатньо високим ступенем точності на модельних даних.

Для вирішення задачі в реальних умовах цей метод можна запропонувати як один із методів, який разом з іншими методами (наприклад, визначення динамічних характеристик польоту з відеоряду зображень) забезпечить коректне розв'язання задачі.

Список літератури

1. Krizhevsky, Alex & Sutskever, Ilya & E. Hinton, Geoffrey. (2012).” ImageNet Classification with Deep Convolutional Neural Networks.”/ Neural Information Processing Systems. 25. 10.1145/3065386.
2. Mash, Robert, Becherer, Nicholas, Woolley, Brian, Pecarina, John ” Toward aircraft recognition with convolutional neural networks”/ [IEEE 2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS) - Dayton, OH, USA (2016.7.25- 2016.7.29)]
3. “How to Retrain an Image Classifier for New Categories” 2018 [Online]. URL: https://www.tensorflow.org/hub/tutorials/image_retraining
4. Szegedy, Christian, Vanhoucke, Vincent, Ioffe, Sergey, Shlens, Jon, Wojna, Zbigniew ”Rethinking the Inception Architecture for Computer Vision”/ [IEEE 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Las Vegas, NV, USA (2016.6.27-2016.6.30)] 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
5. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv.org. URL: <https://arxiv.org/abs/2207.02696>