

Федорова Ю.Є., Отрох С.І., д.т.н.,
Оленєва К.М.

РОЗПІЗНАВАННЯ НЕБЕЗПЕЧНИХ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

Fedorova Y.Y., Otrokh S.I. Olienieva K.M. Detection of dangerous objects in images using neural networks. In recent years, along with the continuous improvement of the computer speed, the application of the computer for the examination and recognition of dangerous goods has become more and more widespread. In order to overcome the shortcomings of the high frequency of false detection in the process of classifying targets using existing feature-based classifiers, this paper proposes an algorithm for detecting dangerous objects using a convolutional neural network based on deep learning. For the image being checked, sliding windows of different scales are used to determine the presence of an object window. To detect objects, a convolutional neural network is trained with a large number of positive and negative samples. For better adaptation to object detection, the topology of the convolutional neural network has been improved. A window of suspected dangerous object is fed into an improved convolutional neural network to detect dangerous objects, and the false detection rate is reduced while maintaining the original detection rate.

Keywords: detection of dangerous objects, differential processing, convolutional neural network, deep learning

Федорова Ю.Є., Отрох С.І., Оленєва К.М. Розпізнавання небезпечних об'єктів на зображеннях за допомогою нейронних мереж. Протягом останніх років, разом з безперервним підвищенням швидкості роботи комп'ютера, все більше і більше розповсюджуються застосування комп'ютера для проведення експертизи та розпізнавання небезпечних вантажів. З метою подолання недоліків високої частоти помилкового виявлення в процесі класифікації цілей за допомогою існуючих класифікаторів на основі ознак, у цій роботі запропоновано алгоритм виявлення небезпечних предметів з використанням згорткової нейронної мережі на основі глибокого навчання. Для зображення, яке перевіряється, використовуються ковзні вікна різних масштабів для визначення наявності вікна об'єкта. Для виявлення об'єктів навчається згорткова нейронна мережа з великою кількістю позитивних і негативних зразків. Для кращого адаптування до виявлення об'єктів поліпшено топологію згорткової нейронної мережі. Вікно підозрілого небезпечного предмета вводиться в покращену згорткову нейронну мережу для виявлення небезпечних об'єктів, і при цьому зменшується частота помилкового виявлення, зберігаючи при цьому вихідну частоту виявлення.

Ключові слова: виявлення небезпечних об'єктів, диференційна обробка, згорткова нейронна мережа, глибоке навчання

Вступ

Приблизно 80 відсотків зовнішньої інформації, з якою мають справу люди, належить до візуальної інформації. Для людей зображення та відео є описом зображення та достовірністю об'єктивних речей, і вони є найважливішим джерелом інформації для людей [1]. Виявлення та відстеження цілей є популярною темою у галузі досліджень комп'ютерного зору. Воно інтегрує обробку зображень, розпізнавання образів, штучний інтелект, автоматичне керування та інші передові технології в багатьох галузях та широко використовується в інтелектуальних системах транспорту, системах інтелектуального моніторингу, промисловому контролі, а також в аерокосмічних та багатьох інших галузях.

Постановка завдання. У зв'язку з великою частотою змін змісту полів сцен, їх важко вивчати за погодженою методикою. На сьогоднішній день головні проблеми виявлення об'єктів включають у себе, як точно та швидко сегментувати ціль, як зменшити вплив складного фону на виявлення цілей та як забезпечити точність виявлення цілей для об'єктів різного розміру та форми. Крім того, у системі виявлення цілей, стійкість та реально-часова продуктивність системи є протиріччям. Стійкість є дуже важливою для системи виявлення цілей, особливо для системи виявлення об'єктів, оскільки система виявлення об'єктів зазвичай

повинна автоматично та постійно працювати і не може бути занадто чутливою до впливу шуму, світла, погоди та інших факторів. Реальний час є вимогою, яку система повинна виконати, що означає, що складність обчислення алгоритму обробки зображень системи не може бути занадто високою.

Аналіз останніх досліджень. Дослідження виявлення небезпечних вантажів головним чином включає виявлення цілей на основі відео та виявлення цілей на основі статичного зображення. У статті розглядається алгоритм виявлення цілей на основі статичного зображення, тобто виявлення та локація набору цілей на статичному зображенні. Складність виявлення цілей на основі статичного зображення полягає в тому, що ціль на зображенні буде змінюватися через зміни в освітленні, перспективі та внутрішньому об'єкті [2]. З огляду на вищезазначені труднощі, дослідження внесли багато спроб. На сьогоднішній день запропоновані методи в основному складаються з алгоритму виявлення цілей на основі контуру форми та методу виявлення на основі аналізу ознак [3].

Метод виявлення цілей на основі форми зазвичай завершує виявлення завдання шляхом вилучення зовнішнього контуру цілі. У виявленні об'єктів цілей форма є очевидною ознакою тіла людини, такою як вертикальний стовбур [4]. Гаврил та інші дослідники побудували бібліотеку шаблонів контурів людини, розрахували схожість між вікном, яке має бути перевірене, та людиною та запропонували ієрархічний алгоритм відповідності на основі контуру. Фархаді та інші покращили алгоритм, поєднуючи характеристики форми об'єкта з інформацією про краї, вказавши, що метод має кращий ефект на об'єкт. Фельценшваль вивчив шаблон контуру об'єкта з великої кількості позитивних зразків та визначив, чи є вікно об'єктом, шляхом вимірювання інформації про краї очікуваного вікна за допомогою відстані Хаусдорфа та визначення шаблону. Переваги трьох вищезазначених методів полягають у тому, що вони прості та не потребують попередньої обробки зображення, і всю інформацію зображення можна зберегти [5]. Однак через різний стан об'єктів алгоритм виявлення об'єктів на основі відповідності шаблонів потребує побудови великої кількості формових шаблонів для досягнення хорошого ефекту відповідності, а класифікаційний ефект не є хорошим, коли є заслонення на цільовому об'єкті.

З огляду на труднощі у точності традиційних алгоритмів виявлення небезпечних вантажів в різних сценаріях, у поєднанні із існуючим станом досліджень та рівнем техніки, стаття поєднує традиційні ознаки об'єкта з методом глибокого навчання, що має важливе теоретичне та практичне значення для технології виявлення небезпечних вантажів [6-8]. Метод, заснований на ознаках та машинному навчанні, реалізує виявлення цілей за допомогою вилучення ознак, навчання ознак та виявлення. Загалом, характеристики тіла людини виражаються приблизно зміною місцевої яскравості статистичного зображення. Загальні ознаки тіла людини включають в себе ознаку хвиль Гаара, ознаку НОG, ознаку Edlgelet та змішану ознаку. Джеффри Хінтон та інші вперше запропонували концепцію глибокого навчання в журналі "Science". Основна ідея полягає в тому, щоб імітувати процес навчання людського мозку впізнавати ціль через глибоку нейронну мережу. Глибоке навчання досягло великого успіху в багатьох галузях і привернуло широку увагу. У 2012 році Хінтон та інші використовували глибоке навчання та виграли чемпіонат з класифікації зображень ImageNet, вони навчали в оригінальному просторі пікселів RGB глибоку згорткову нейронну мережу глибиною, яка містить 60 мільйонів пікселів, 650000 нейронів становлять мережу з п'яти шарів згорткової точності, ніж використання традиційних ручних ознакових проєктів претендентів на 100 відсотків. Джі та інші запропонували багатошарову 3D згорткову нейронну мережу для вивчення просторово-тимчасових характеристик відеоблоку та здійснили навчання ознак усього відео за допомогою операції згортки, замінивши попереднє виявлення просторово-тимчасових точок і вилучення ознак описувачів.

Метою роботи є дослідження методів виявлення та відстеження цілей на основі зображень, зокрема на основі статичного зображення. Основна увага зосереджена на

вирішенні проблем точності, швидкості та стійкості виявлення об'єктів у різних умовах. Також досліджується поєднання традиційних методів виявлення об'єктів на основі форми з новітніми методами глибокого навчання, з метою підвищення ефективності виявлення небезпечних вантажів.

Виклад основного матеріалу дослідження.

Принцип виявлення небезпечних об'єктів. Алгоритм виявлення небезпечних вантажів належить до виявлення цілей у галузі комп'ютерного зору. Реалізація цього алгоритму потребує професійних знань з обробки зображень, машинного навчання та розпізнавання образів [9]. В даний час найпопулярніший метод виявлення об'єктів здійснюється за допомогою ознак та машинного навчання. Велика кількість досліджень оптимізувала та інновувала характеристики об'єктів та алгоритми машинного навчання. Загалом відмінні алгоритми виявлення об'єктів спрямовані на досягнення швидкості виявлення та точності.

Основна ідея алгоритму виявлення об'єктів на основі шаблону полягає в побудові бібліотеки шаблонів об'єктів шляхом збору різних форм зображень об'єктів заздалегідь. При потребі виявлення екстрагуються ознаки шаблону вхідного зображення, і в бібліотеці шаблонів об'єктів шукають відповідні об'єкти. Залежно від типу використаного шаблону, його можна розділити на алгоритм виявлення об'єктів на основі загального шаблону та на алгоритм виявлення об'єктів на основі локального шаблону. Оскільки зовнішній контур об'єкта є найлегше відрізняється серед багатьох об'єктів, контур об'єкта є найбільш поширеним шаблоном на основі загального [10]. Алгоритм виявлення об'єктів на основі загального шаблону має відносно високу швидкість виявлення, але низький рівень виявлення в разі заслонення об'єкта. Навпаки, алгоритм виявлення на основі локального шаблону має повільну швидкість виявлення через складний процес обчислення, але все одно може досягти хорошого ефекту виявлення в умовах заслонення об'єкта.

Найбільш поширеним класифікатором у сфері розпізнавання візерунків та класифікації регресії є машина опорних векторів (SVM). Ідея полягає в тому, щоб визначити оптимальну площину сегментації в просторі ознак, максимізуючи відстань між площиною сегментації та гіперплощиною, та створити найкраще класифікаційне рішення. Звичайні класифікатори SVM поділяються на лінійні та нелінійні. Лінійний класифікатор має переваги простої структури, швидкості обчислення та може використовуватися разом з численними більш складними нелінійними наборами ознак для підвищення точності класифікатора. Нелінійний класифікатор відображає простір ознак у вищі виміри для реалізації класифікаційного рішення, проте для покращення результатів класифікації потрібні складні математичні обчислення.

Структура мережі з конволюційними шарами (CNN). Існує багато варіацій архітектури CNN, але основна структура все рівно дуже схожа. Основна архітектура CNN зазвичай складається з трьох шарів: конволюційного шару, шару пулінгу та повнозв'язного шару. Конволюційний шар призначений для вивчення характеристичного представлення входу. Як показано на рисунку 1, конволюційний шар складається з декількох карт ознак. Кожен нейрон у карті ознак з'єднаний з його сусідніми нейронами у попередньому шарі. Така сусідня область називається локальним полем сприйняття нейрона у попередньому шарі. Для обчислення нової карти ознак вхідна карта ознак спершу зводиться в зв'язку з вивченим ядром конволюції (також відомим як фільтр, детектор ознак), а потім результат передається до нелінійної активаційної функції. Нову карту ознак отримують за допомогою застосування різних ядер конволюції. Така модель однакових ваг спільного використання має кілька переваг, таких як скорочення складності моделі і полегшення навчання мережі. Функція активації описує нелінійність CNN, що ідеально підходить для багатошарових мереж для виявлення нелінійних характеристик.

Шар пулінгу призначений для досягнення просторової інваріантності за рахунок зменшення роздільної здатності карти ознак. Він зазвичай розташований між двома конволюційними шарами. Карта ознак кожного шару пулінгу з'єднана з відповідною картою ознак попереднього конволюційного шару, тому кількість карт ознак однакова. Типові операції пулінгу - це середнє пулінг та максимальне пулінг. Шляхом накладання кількох конволюційних та пулінгових шарів можна видобути більш абстрактне характеристичне представлення.

Після кількох конволюційних та пулінгових шарів, зазвичай є один або кілька повнозв'язних шарів. Вони з'єднують всі нейрони у попередньому шарі з кожним нейроном у поточному шарі і не зберігають просторову інформацію у повністю з'єднаному шарі.

Розмір вхідного зображення конволюційної нейронної мережі, навченої в цій роботі, становить 62×62 . І для масштабування та налаштування зображення використовується метод бікубічної інтерполяції, який зберігає кращу якість деталей, ніж домінуючий алгоритм білінійного фільтрування. Процес вилучення ознак схожий на процес прямого розповсюдження під час навчання. Тому він має загалом 12 карт ознак розміром 3×3 . У цій роботі ці карти ознак розтягуються у вектор з 300-ма вимірами, який є результатом виявлення ознак вихідного зображення.

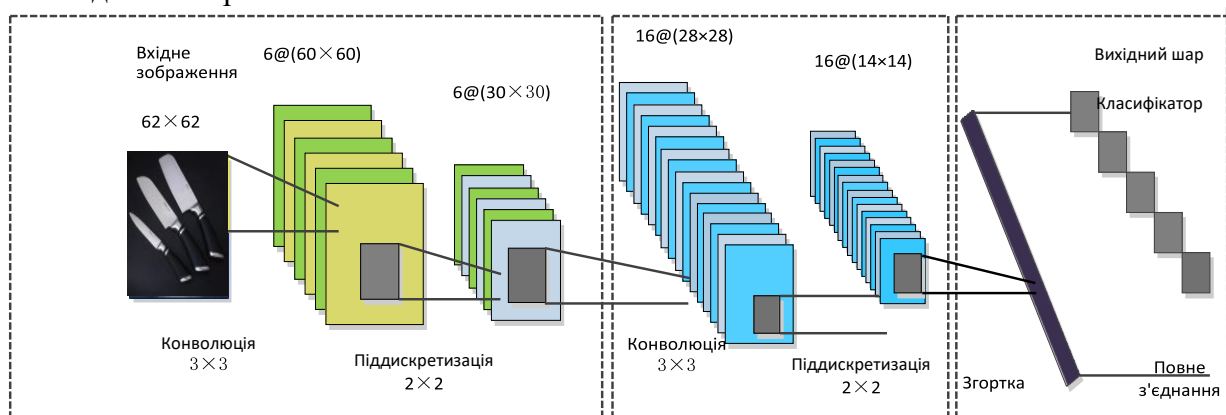


Рис. 1. Базова архітектура згорткової нейронної мережі (CNN)

Перед виявленням об'єкта та класифікацією за допомогою CNN, необхідна диференційна обробка. Різниця - це виявлення змін на рівні пікселів, яке може бути розділене на неперервну міжкадрову різницю та різницю фону. Спочатку використовується неперервна міжкадрова обробка, щоб порівняти два послідовні кадри та витягнути інформацію про рухомий об'єкт. Після витягнення відеосеквенції образу рухомого об'єкта різницю картини буде бінаризовано для виявлення фону та рухомої точки об'єкта піксель за пікселем, що базується на виявленні на рівні пікселів. Традиційний бінарний метод - це встановлення фіксованого порогового значення. Якщо значення різниці менше порогового значення T , воно буде визнано як піксель фону; в іншому випадку це буде піксель об'єкта. Якщо різниця значень пікселя в картині різниці більше заданого порогового значення T , то піксель є пікселем об'єкта, тобто піксель може бути точкою на об'єкті, в іншому випадку він вважається пікселем фонового графіку.

Розмір порога визначається емпіричним значенням, в різних умовах середовища поріг не є фіксованим, тому простий метод визначення значення не бажаний. Люди вводять метод сегментації в бінарну обробку картини різниці в надії, що алгоритм сегментації може автоматично вибрати потрібне мінімальне значення, щоб розділити картину різниці на об'єкт і фон. Наприклад, застосували адаптивний метод сегментації порогу. Спочатку вважається, що гістограма картини різниці є сумішшю трьох зважених одноденних гаусівських моделей з одиничними вершинами, потім змішується ймовірність густоти різницевої картини.

Іншими словами, зображення розділяється на два типи пікселів, а саме рухому об'єкту та фону, за допомогою налаштування цього порога. Загалом, коли значення порога

використовується для сегментації бінаризації картини різниці, межа області об'єкта на отриманому бінаризованому зображенні часто не є плавною. Область об'єкта має певні отвори, спричинені неправильним розгляданням, і є деякі шуми в зоні фону, також спричинені неправильним розгляданням. Послідовні операції відкриття та закриття можуть значно покращити ситуацію та досягти потрібного ефекту, але важливо зауважити, що вибір розумного розширення/зменшення або комбінація операцій відкриття та закриття має значний вплив на ефект обробки.

Крім того, оскільки розмір об'єкта на зображенні для виявлення різний, зазвичай використовується багатомасштабне ковзне вікно для виявлення. Пряме використання згорткової нейронної мережі для виявлення в ковзному вікні призводить до низької ефективності виявлення та відсутності реального часу. Тому процес, описаний вище, використовується для швидкого виявлення вікна, яке підозрюється у наявності об'єкта, а потім вікно передається в згорткову нейронну мережу для подальшого виявлення, що має хорошу реальну часову продуктивність та швидкість виявлення.

Результати. Для цього дослідження було виявлено понад 300 зображень небезпечних вантажів, завантажених з Інтернету. На етапі підтвердження підозрюваних вікон небезпечного вантажу, у цій роботі використовується кольоровий канал, амплітуда градієнту та гістограма напрямку градієнту як функції фузії з хорошою здатністю представлення цілей, а детектор, навчений за допомогою цієї функції фузії, має високу частоту виявлення. У цій роботі був використаний метод оцінки на основі ефекту виявлення для оцінки загальної ефективності детектора. Всі існуючі зображення небезпечних вантажів цих зображень були виявлені, і була підрахована частота виявлення вантажів та частота помилкового визначення кожного зображення в середньому.

Те, що підозрюється на виявлення за допомогою функції фузії, є CNN, завершений введенням навченого вікна існуючого небезпечного вантажу, яке використовується для виявлення небезпечних вантажів. Експериментальні результати показані на рис. 3. Порівняно з рис. 2, було виявлено, що після подальшого виявлення за допомогою CNN помилкове виявлення вікна ефективно еліміновано. Аналогічно, цей детектор використовувався для виявлення цих 300 завантажених зображень. Помилкова частота виявлення цього детектора ефективно знизилася, при цьому частота виявлення залишилася високою - 93%.



Рис. 2. Кінцевий результат тестування звичайної CNN

В порівнянні взяли алгоритми, які включають алгоритм виявлення ознак з використанням ознак ACF та алгоритм ConvNet з використанням CNN виключно для виявлення об'єктів. Щоб перевірити частоту виявлення трьох алгоритмів, у цій статті використовуються зображення виявлення небезпечних вантажів як тестовий набір, всі зразки для проведення експериментів, тестовий набір містить понад 300 зображень. Таблиця 1 показує результати трьох алгоритмів на тестовому наборі, цей алгоритм може використовувати частоту виявлення 93%, що на 3% вище, ніж використання алгоритму ACF, але на 4% нижче, ніж алгоритм ConvNet. Однак середня час виявлення кожного зображення за допомогою алгоритму ConvNet становив 1,12 с, що не відповідало вимогам швидкості в

реальних застосуваннях. Алгоритм, запропонований у цій роботі, потребує лише 0,23 с для виявлення зображення в середньому, що може задовольняти вимоги щодо швидкості виявлення, а частота виявлення вища, ніж у детектора, навченого за допомогою алгоритму ACF.



Рис. 3. Кінцевий результат тестування

Таблиця 1

Порівняння трьох алгоритмів виявлення

	ACF	ConvNet	Експериментальний метод
Вірне виявлення	1013	1092	1047
Помилкове виявлення	280	276	284
Пропущене виявлення	113	34	79
Показник виявлення (%)	90	97	93
Середній час (с)	0.17	1.12	0.23

Висновки

Зі стрімким розвитком комп'ютерів застосування комп'ютерного інтелекту стає все більш розповсюдженим. Виявлення небезпечних матеріалів є важливою частиною інтелектуального застосування комп'ютера, яке має широку цінність в інтелектуальних системах транспорту, навчанні роботів, системах інтелектуального моніторингу та інших сферах. З метою досягнення ефекту у цій роботі акцентується на аналізі алгоритму виявлення небезпечних вантажів, поєднуючи глибоке навчання з виявленням небезпечних матеріалів. Через експерименти було доведено, що точність виявлення небезпечних вантажів значно покращується порівняно з традиційним алгоритмом.

Список використаної літератури:

1. Мохамед, А. А. С., Берзой, А., Мохамед, О. А., "Дизайн та апаратна реалізація керування FL-MPPT сонячними електростанціями на основі ГА та аналізу малих сигналів," IEEE Transactions on Sustainable Energy, 8(1), 279-290 (2017).
2. Берхон, Д., Куевас, С., Моран, Ф., Гарсія, Н., "Імплементация на основі GPU оптимізованої непараметричної моделі фону для реального виявлення рухомих об'єктів," IEEE Transactions on Consumer Electronics, 59(2), 361-369 (2013).

3. Вандана, П., Аріндам, Б., Гаррік, О., "Алгоритм фільтрації шуму для сенсорів для виявлення асинхронних змін подій на основі truenorth та його імплементація на truenorth," *Frontiers in Neuroscience*, 12, 118-125 (2018).
4. Маджумдер, Н., Порія, С., Гельбух, А., Камбрія, Е., "Моделювання документів на основі глибокого навчання для виявлення особистостей з тексту," *IEEE Intelligent Systems*, 32(2), 74-79 (2017).
5. Карасулу, Б., Корукоглу, С., "Виявлення та відстеження рухомих об'єктів за допомогою методу анізотропного віднімання фону відео: оптимізація продуктивності," *Expert Systems with Applications*, 39(1), 33-43 (2012).
6. Усама, Й., Лу, С., Імам, Х., Сен, С., Кар, Н. С., "Дизайн та імплементація модуля виявлення та ідентифікації дефектів затишного струму на основі аналізу хвильового пакету для застосування на лініях передач," *IET Generation, Transmission and Distribution*, 8(3), 431-441 (2014).
7. Тан, Й., Дінг, К., "Огляд реалізації на основі GPU алгоритмів інтелектуальних роїв ORCID="0000-0001-8243-4731", " *IEEE Transactions on Cybernetics*, 46, (9), 1-14 (2015).
8. Цю, Ц., У, К., Бішоп, М., Піно, Р.Е., Ліндерман, Р.В., "Паралельна нейроморфна система розпізнавання тексту та її імплементація на гетерогенному високопродуктивному обчислювальному кластері," *IEEE Transactions on Computers*, 62(5), 886-899 (2013).
9. Лі, С., Гуо, Р., Чен, С., "Надійне відстеження та розпізнавання пішоходів з FLIR-відео: уніфікований підхід за допомогою розрідженої кодування," *Sensors*, 14(6), 11245-11259 (2014).
10. Росас-Ромеро, Р., "Віддалене виявлення лісових пожеж за допомогою відеосигналів з класифікаторами на основі навчених словників K-SVD," *Engineering Applications of Artificial Intelligence*, 33, 2014.

Автори статті

Федорова Юлія - студентка, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського». Київ, Україна.

Отрох Сергій - доктор технічних наук, професор, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського». Київ, Україна.

Оленєва Ксенія - асистент кафедри, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського». Київ, Україна.

Authors of the article

Fedorova Yuliia - student, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute». Kyiv, Ukraine.

Otrokh Serhii - Doctor of Science (technic), Professor, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute». Kyiv, Ukraine.

Olenieva Kseniia - department assistant, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute». Kyiv, Ukraine.