

Запобігання можливим пограбуванням за допомогою алгоритму глибокого навчання з обробкою камерою

*Р.Продан, магістр, Д.Шутка, магістр, В.Татарин, к.е.н., доц.
Національний університет «Львівська політехніка», Україна;
e-mail: roman.prodan.mknuo.2022@lpnu.ua*

Анотація

Останнім часом технології глибокого навчання, а саме нейронні мережі [1], привертають все більше уваги з боку бізнесу та наукової спільноти, оскільки вони допомагають оптимізувати процеси та знаходити реальні рішення проблем набагато ефективніше та економніше, ніж багато інших підходів. Зокрема, нейронні мережі добре підходять для ситуацій, коли потрібно виявляти об'єкти або шукати подібні шаблони у відео та зображеннях, що робить їх актуальними в галузі інформаційних та вимірювальних технологій у мехатроніці та робототехніці. Оскільки з кожним роком кількість пограбованих квартир і будинків зростає, вирішення цієї проблеми стало одним із пріоритетних завдань сучасного суспільства. Використовуючи методи глибокого навчання, такі як нейронні мережі, у мехатроніці та робототехніці, можна розробити інноваційні рішення для вдосконалення систем безпеки, що дозволить ефективніше виявляти та запобігати квартирним злочинам.

У нашому дослідженні ми зосередилися на використанні потенціалу YOLO для розробки надійної системи виявлення потенційних злочинців. Щоб досягти цього, ми спочатку зібрали повний набір даних з Інтернету, що містить зображення та відео, що зображують різні злочинні дії та сценарії. Цей набір даних був ретельно підібраний і анотований, щоб гарантувати, що він адекватно охоплює широкий спектр злочинної поведінки та атрибутів.

Потім ми застосували архітектуру YOLO, щоб навчити нашу мережу на цьому наборі даних. Процес передбачав встановлення відповідних параметрів, таких як швидкість навчання, розмір пакету та кількість ітерацій навчання, щоб оптимізувати продуктивність мережі. Етап навчання передбачав передачу в мережу анотованих зображень і надання їй можливості вивчати відповідні особливості та шаблони, пов'язані з потенційною злочинною діяльністю.

У процесі навчання мережа YOLO пройшла кілька етапів експлуатації. Він використовував серію згорткових шарів для вилучення ієрархічних функцій із вхідних зображень, фіксуючи як локальну, так і глобальну контекстну інформацію. Потім ці функції були введені в повністю зв'язані рівні, щоб уможливити просторове мислення та створити точні обмежувальні рамки навколо потенційних злочинних об'єктів або осіб.

Щоб оцінити продуктивність нашої навченої мережі, ми провели масштабні експерименти на окремому тестовому наборі даних, який відрізнявся від даних навчання. Ми ретельно позначили цей набір даних, щоб отримати точні наземні анотації правдивості для порівняння. Вимірюючи показник точності, ми визначили ефективність нашої моделі у виявленні потенційних злочинців.

Наші експерименти показали точність виявлення потенційних злочинців у 97%. Це досягнення демонструє спроможність YOLO та ефективність нашої навченої мережі в точному виявленні злочинної діяльності. Високий рівень точності вказує на те, що наша система може ефективно допомагати в захисті власності, надаючи цінний інструмент для персоналу служби безпеки та правоохоронних органів.

Ключові слова

Злочини, пограбування, нейронні мережі, виявлення об'єктів, машинне навчання.

1. Введення

Сьогодні одним із першочергових завдань, яке необхідно вирішити, є проблема захисту наших осель та майна. Історія систем домашньої безпеки починається з того часу, коли людство почало будувати свої перші будинки,

хоча вони дуже відрізнялися від систем безпеки, які ми маємо сьогодні. Але люди завжди мали інстинкт зберігати свої домівки, сім'ї та майно в безпеці. Можливо, ми розвинули цей інстинкт через історію воєн або через вроджену відповідальність захищати один одного. Ця потреба відкрила двері для розробки як простих, так і складних засобів захисту наших будинків. Системи безпеки приймали різні форми ще до відкриття електрики та її застосування в сучасних технологіях. З розвитком технологій люди почали використовувати ці архітектури та конструкції для створення більш досконалих систем. Але з кожним роком кількість пограбованих квартир тільки збільшується.

За статистикою, минулого року в Україні було зафіксовано понад 40 тисяч випадків квартирних пограбувань, і з кожним роком ця цифра збільшується. Професійні охоронні компанії можуть розмістити різні типи датчиків для контролю вашого будинку. Датчики можуть відстежувати рух у вашому домі, відчинені двері, розбиті вікна, дим, рівень чадного газу та навіть температуру у вашому домі. Компанії, які надають послуги безпеки, повинні постійно стежити за камерами відеоспостереження, щоб зафіксувати незвичну поведінку підозрілих осіб. Велика кількість людей користуються камерами спостереження, які здатні лише фіксувати відео, але не дають конкретної інформації про можливу підготовку до злочину. Системи та підходи, які зараз є на ринку, не забезпечують достатнього рівня безпеки та надійності для сучасних викликів.

У зв'язку зі швидким зростанням популярності використання нейронних мереж для розпізнавання об'єктів у відео зростає потреба в оптимізації аналізу людської поведінки, і нейронні мережі найкраще підходять для цього. Вони можуть обробляти відеопотоки в реальному часі, аналізувати та повідомляти про загрозу пограбування користувача. Також такий підхід дуже корисний для охоронних компаній, які можуть значно оптимізувати свою роботу. Нам потрібно прийняти концепцію комп'ютерного зору в домівках, а не ухилятися від ідеї обміну особистими даними для досягнення нових рівнів захисту, безпеки, комфорту та розваг. Комп'ютерний зір у поєднанні з ML дозволяє комп'ютерам/системам через цифрові зображення чи відео розуміти те, що вони бачать.

2. Недоліки

Обмеження підходів глибокого навчання, такі як недостатня кількість даних і проблеми з якістю даних, створюють значні проблеми для досягнення оптимальної продуктивності та надійності системи. У сфері захисту власності стає обов'язковим отримати найкращі вихідні дані, які охоплюють різні ситуації, пов'язані з потенційним пограбуванням.

Збір і анування даних для навчання моделі глибокого навчання є ручним і трудомістким завданням, яке вимагає кропітких зусиль. Це передбачає пошук відповідних зображень і відео, що зображують потенційні сценарії пограбування, з різних джерел, таких як записи з камер спостереження, онлайн-платформи та реальні випадки. Забезпечення того, щоб зібрані дані адекватно представляли спектр можливих злочинних дій та умов навколишнього середовища, має вирішальне значення для підготовки надійної та точної моделі.

3. Мета

У цій статті висвітлюється застосування глибокого навчання для запобігання квартирним злочинам шляхом використання його здатності аналізувати великі дані та виявляти складні закономірності. Завдяки вдосконаленим алгоритмам глибокого навчання він демонструє здатність обробляти різноманітні джерела даних, у тому числі записи з камер спостереження та дані датчиків, для ідентифікації підозрілих дій і потенційних загроз у реальному часі, що сприяє швидшому часу реагування та превентивним заходам із запобігання злочинам.

4. Комп'ютерний зір для аналізу відео

У 2010-х роках відбувся бум на нейронні мережі. На початку 2010-х років дослідження показали, що для навчання нейронних мереж найкраще використовувати функцію активації ReLU та алгоритм оптимізації Адама[2]. Ці два компоненти допомогли ефективно навчити нейронні мережі.

У 2012 році було зроблено прорив у класифікації зображень, вигравши ImageNet Large Scale Visual Recognition Challenge (ILSVRC) із глибокою архітектурою CNN, відомою як AlexNet[3]. AlexNet використовував кілька згорткових шарів, нелінійні функції активації та такі методи, як регуляризація випадання, щоб покращити узагальнення. Ця перемога значно підвищила популярність і дослідницький інтерес до CNN.

4.1 Нейронні мережі згортки

На сьогоднішній день основним методом розпізнавання об'єктів на зображенні є згорткові нейронні мережі та алгоритм зворотного поширення помилок.

Згорткові нейронні мережі використовуються як основа нейронної мережі для виявлення об'єктів. Convolution Neural Network (ConvNet/CNN)[4] - у машинному навчанні це основний інструмент для розпізнавання та класифікації об'єктів, облич, символів тощо на зображеннях і відео. Згорткові нейронні мережі засновані на багатопарових перцептронах і розроблені для використання мінімальної попередньої обробки зображень і відео. Конволюційна нейронна мережа використовує операцію згортки під час навчання, яка водночас дозволяє використовувати ці операції для зменшення обсягу інформації, яка використовується в пам'яті, і кращої роботи із зображеннями чи відео з вищою роздільною здатністю. Операція згортки під час навчання нейронної мережі вчиться витягувати опорні елементи в зображеннях і відео, такі як «краї», «контури» або «грані». На наступному рівні, з цих країв і граней нейрона, мережа намагається розпізнавати повторювані фрагменти текстур, які потім складаються у фрагмент зображення. Кожен шар нейронної мережі має свою операцію згортки. Якщо на перших шарах мережа вчиться розпізнавати візерунки країв і граней, то глибше шари нейрона мережа пробує текстури та частини об'єктів. В результаті такого навчання ми можемо правильно класифікувати картинку або виділити потрібний об'єкт на завершальному етапі.

Основні компоненти згорткових нейронних мереж [5]:

1. Вхідний шар. Як описано вище, згорткові нейронні мережі навчаються на зображенні. CNN розрізняє зображення за трьома кольорами: червоний, зелений, синій - це так зване RGB-подання. Першим входом до будь-якого типу нейронної мережі є вхід.
2. Шар згортки[6][13]. Згортковий шар є найважливішою частиною CNN. Шар згортки працює наступним чином: нехай у нас є зображення $D = 300 \times 150 \times 3$, де 300 — висота, 150 — ширина, а 3 — кількість каналів RGB. Ядро — це вагова матриця, яка використовується для обчислення операції згортання. Як показано на малюнку 2.2, ядро рухається від правого кута із заданим кроком і виконує поелементне множення пікселів, після чого результат підсумовується. Результат матриці залежить від кількох параметрів:
 - Крок
 - ЯдроОсновною метою цієї операції є дослідження низькорівневого шаблону зображення:
 - лінії,
 - кути,
 - бали,
 - контури і обличчя.

У більш глибоких шарах нейрона мережа використовує шаблони, які вона вже вивчила, і намагається зрозуміти високорівневі фрагменти зображення. Результатом операції згортки є отримання нової матриці зображення, яка використовується в наступних рівнях нейронних мереж.

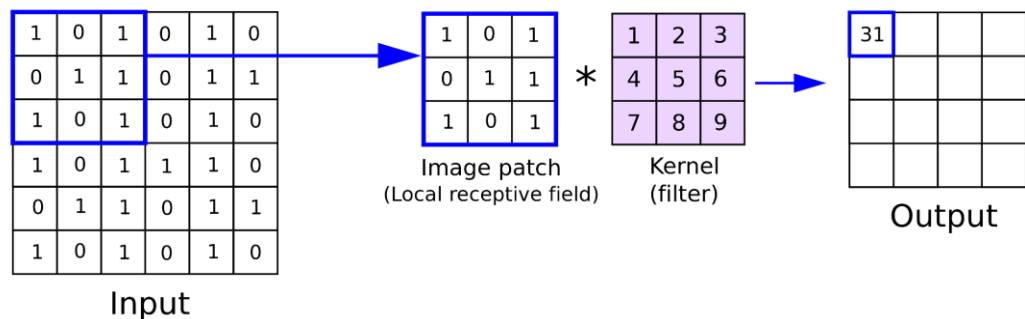
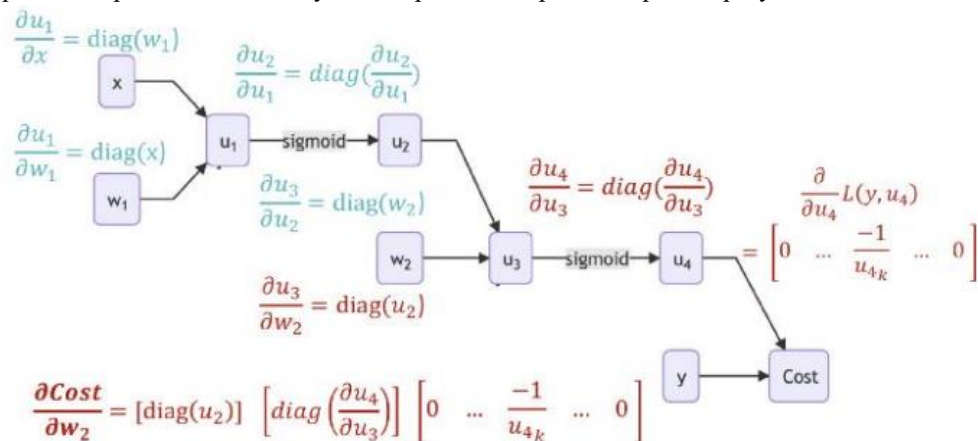


Рисунок 1. Шар згортки

3. Шар об'єднання [7]. Основне призначення цього шару – зменшити розмірність зображення. Таким чином, це зменшує кількість параметрів і обсяг обчислень, що виконуються в мережі. Подальші операції виконуються над зменшеним зображенням
4. Функція активації[8]. Це математична функція[7], яка змінює кожне негативне значення зображення на нуль і посилює нелінійні властивості функції. Простіше кажучи, функція активації – це функція, яка

додається до нейронної мережі, щоб допомогти мережі вивчати складні закономірності в даних. Функція активації приймає вихідний сигнал від попереднього нейрона і перетворює його в певну форму, яка може бути прийнята як вхідний для наступного нейрона.

5. Пряме поширення[9]. У нейронних мережах алгоритм прямого розповсюдження має велике значення та виконує дві основні функції під час навчання:
 - Сума добутків. Алгоритм працює в одному напрямку і множить ваги нейронної мережі на вхідний вектор (зображення), а потім підсумовує їх.
 - Застосування функції активації. При проходженні кожного шару нейронної мережі та отриманні вихідної матриці в результаті множення вагових коефіцієнтів на вхідний вектор алгоритм прямого розповсюдження передає суму у функцію активації. А саме, функція активації активує нейрони, які найбільше підходять для навчання даних, і передає ці значення на наступний рівень нейронної мережі.
6. Функція втрат[10]. Функція втрати є важливим етапом нейронної мережі, вона оцінює, наскільки добре та точно працює нейронна мережа. Функція витрат є показником того, наскільки помилкова модель в оцінці співвідношення між прогнозованими та фактичними даними. Функція застосовується після кожного проходу алгоритму прямого поширення. Основна мета полягає в тому, щоб надати інформацію для алгоритму зворотного розповсюдження помилок про те, наскільки ваги потрібно оновити, щоб отримати кращий результат на наступній ітерації мережі.
7. Зворотне розповсюдження [11] стало одним із методів, які зробили революцію в представленні та ефективності нейронних мереж. Зворотне поширення є ітеративним і рекурсивним методом. Після того, як алгоритм прямого розповсюдження запущений і функція вартості обчислена, ми знаємо, наскільки добре наша модель працює на заданій ітерації навчання. І відповідно до значення функції вартості ми мінімізуємо помилку нейронної мережі шляхом диференціювання складної функції та обчислення градієнтів. Іншими словами, алгоритм зворотного поширення повертається через мережу, регулюючи ваги нейронної мережі, щоб на наступній ітерації ми отримали кращий результат.



Малюнок 2. Алгоритм зворотного поширення

8. Виявлення об'єктів на зображеннях. Нейронні мережі для виявлення об'єктів на зображенні складаються з двох частин. Перший — це кодер інформації, який використовує згорточні мережі, а другий — декодер, який передбачає обмежувальні рамки та мітки для кожного об'єкта.

Для наших експериментів ми використовували нейронну мережу з одноразовим детектором [12]. SSD покладається на набір попередньо визначених регіонів. Сітка опорних точок накладається на вхідне зображення, і в кожній опорній точці об'єкти різних форм і розмірів служать регіонами. Для кожної області в кожній точці модель передбачає наявність об'єкта в цій області та змінює розташування та розмір обмежувальної рамки, щоб краще відповідати об'єкту. Оскільки кожна опорна точка може мати кілька регіонів і опорні точки можуть бути близько одна до одної, твердотільні накопичувачі створюють багато потенційних перекриваючих виявлень. Постобробка повинна бути застосована до виходів SSD, щоб видалити більшість цих прогнозів і вибрати найкращий. Найпопулярніший метод постобробки відомий як немаксимальне придушення.

4.2 Реалізація

Для реалізації виявлення грабіжників ми використовували архітектуру нейронної мережі для виявлення об'єктів в реальному часі - YOLO4. Дана модель є четвертим поколінням моделі YOLO. Аббревіатура від You Only Look Once означає, що мережа робить прогноз за один прохід через мережу.

YOLO має три основні переваги перед іншими типами моделей:

- Швидкість: Ця мережа значно підвищила швидкість виявлення об'єктів і дозволяє прогнозувати в реальному часі.
- Висока точність: працює дуже точно з мінімальними фоновими помилками.
- Навчання: архітектура YOLO розроблена для розширення можливостей навчання.

Архітектура YOLO складається з двох частин:

- Згорточна нейронна мережа.
- Одноразові детектори, SSD

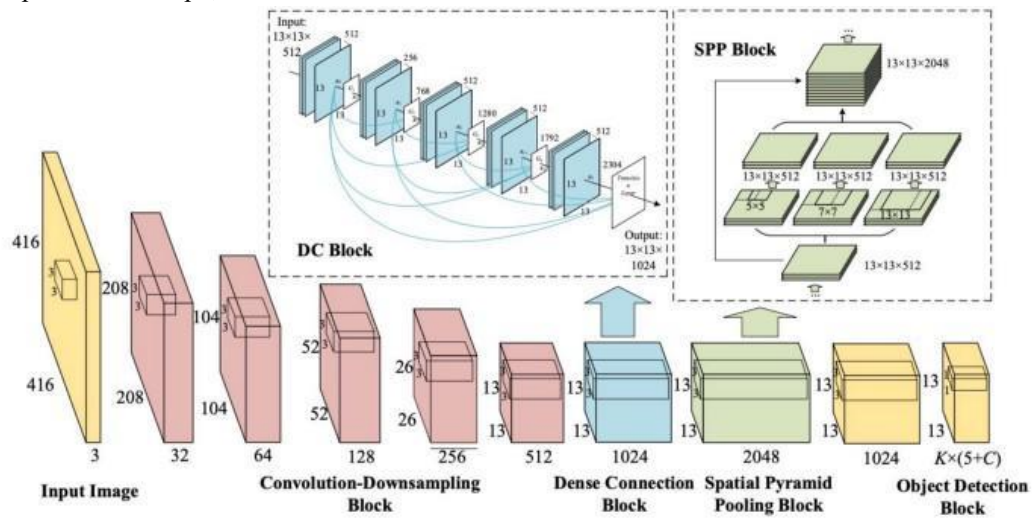


Рисунок 3. Архітектура YOLOv4

Навчальний набір даних був зібраний в Інтернеті. Проаналізувавши кілька десятків відео та виявивши у грабіжників відповідні моделі поведінки. Було сформовано набір даних із 1000 зображень, який включає два класи людської поведінки:

- потенційний грабіжник
- нормальна поведінка



Малюнок 4: Потенційний грабіжник

Після аналізу даних кожне зображення потрібно було анотувати відповідним класом поведінки за допомогою інструменту LabelMe.

Для навчання нейронної мережі YOLO ми використовували фреймворк Darknet, у якому вже є інструменти для запуску навчання та налагодження процесу. Щоб почати навчання, вам потрібно створити два файли .TXT, де набір даних буде розділено на навчальні та тестові зразки. Набір даних було розділено на 80% даних для навчання та 20% для тестування. Процес навчання запускається за допомогою вбудованих методів Darknet.

5. Результати

У наших експериментах ми використовували Mean Average Precision (mAP) [14] для оцінки мережі. mAP базується на таких показниках:

- матриця плутанини;
- Точність

mAP - розраховується як середня точність для кожного класу, а потім усереднюється по всіх класах.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

Після реалізації всіх компонентів було проведено ряд експериментів з навчанням нейронної мережі Yolo та отримано результати метрик для оцінки нейронної мережі.

Під час навчання функція вартості як основний показник стабільного навчання нейронної мережі завжди виглядала у вигляді спадаючого графіка, що вказує на те, що нейронна мережа добре навчена.

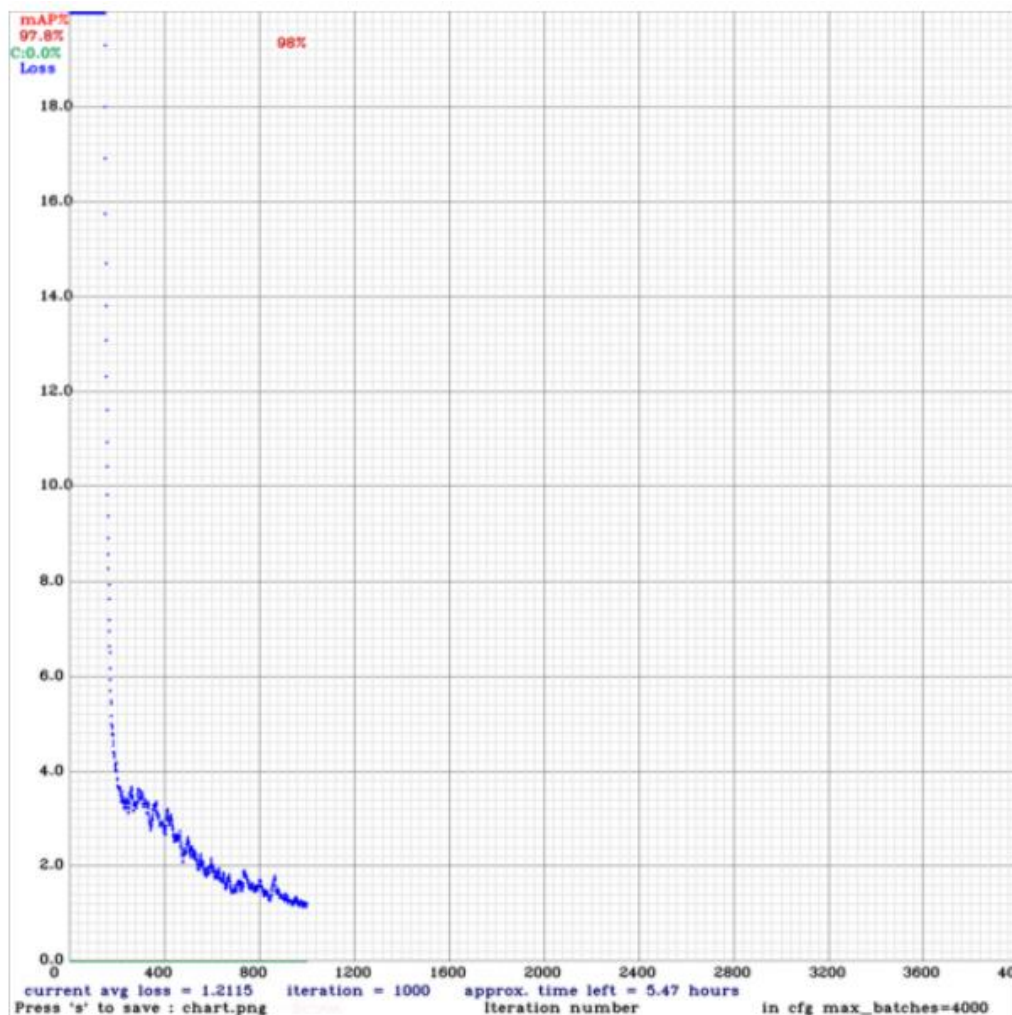


Рисунок 5: Результати функції витрат

Після навчання функція витрат зупинилася на значенні 0,93. Крім того, метрика Mean Average Precision (mAP) показала хороший результат для оцінки того, наскільки добре нейронна мережа розрізняє різні поведінки людей і може точно їх передбачити.

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.978022, or 97.80 %
Total Detection Time: 1 Seconds

Рисунок 6: Результат точності середнього значення

Досягнутий показник середньої точності (MAP) 0,97 демонструє високий рівень продуктивності.

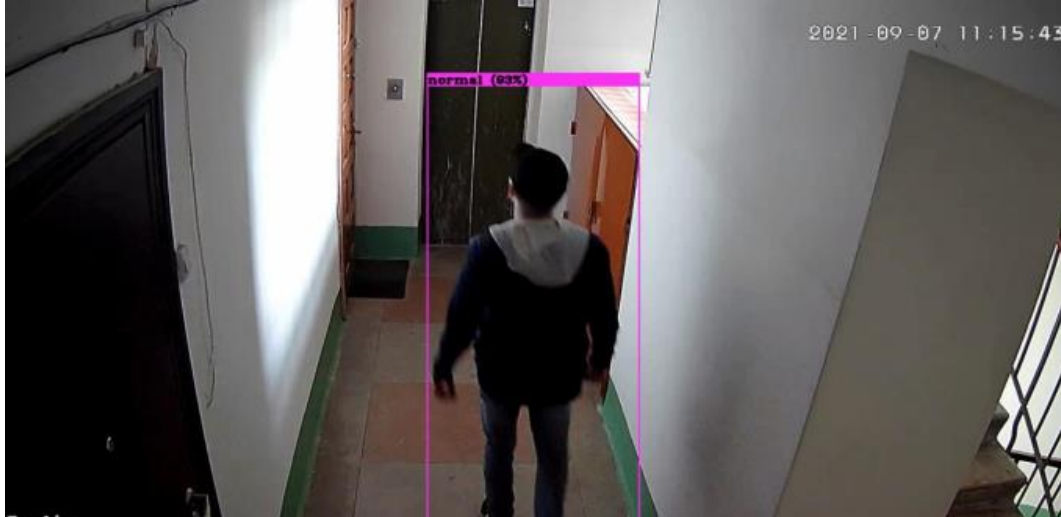


Рисунок 7: Виявлення нормальної поведінки.



Рисунок 8. Виявлення поведінки потенційного грабіжника.

Описаний нами метод виявлення потенційного пограбування на основі нейронної мережі YOLO працює досить точно. Нейронна мережа розпізнає потенційну поведінку злочинців із середньою середньою точністю понад 95%.

5. Висновки

Ця дослідницька робота демонструє потенціал використання нейронних мереж для виявлення та запобігання злочинній поведінці, підкреслюючи її переваги перед існуючими рішеннями. Дослідження також передбачає поглиблений аналіз моделей поведінки потенційних злочинців, які є ключовими міркуваннями під час процесу навчання мережі. Використовуючи структуру Darknet у поєднанні з архітектурою нейронної мережі YOLO, було досягнуто значних успіхів у точному виявленні різноманітної поведінки людей. Це дослідження охоплює презентацію результатів навчання та практичних оцінок, проведених у сценаріях реального світу. Однак важливо зазначити, що для досягнення оптимальної продуктивності нейронної мережі потрібен значний набір даних, щоб підвищити її здатність узагальнювати різноманітну поведінку людей. Експерименти, проведені в рамках цього дослідження, сприяють оцінці потенційної застосовності цього підходу в сфері розпізнавання злочинної поведінки, закладаючи основу для майбутніх експериментів і впровадження. Крім того, цей підхід має значні

перспективи у військовому секторі, де подібні методи розпізнавання поведінки можуть знайти цінні застосування.

Список літератури

- [1] Розенблат Ф. (1958). «Перцептрон: імовірна модель для зберігання та організації інформації в мозку». Доступно - <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.3398&rep=rep1&type=pdf>
- [2] Дідерік П. Кінгма, Джиммі Ба. (2010). «Адам: метод стохастичної оптимізації». Доступно – <https://arxiv.org/abs/1412.6980>
- [3] Алекс Крижевський, Ілля Суцкевер, Джеффри Е. Хінтон. (2012). «Класифікація ImageNet із глибокими згортковими нейронними мережами». Доступно - “ <https://paperswithcode.com/method/alexnet> ”
- [4] Bengio, Y. & Lecun, Yann. (1997). «Згорткові мережі для зображень, мови та часових рядів». Доступно – https://www.researchgate.net/publication/216792820_Convolutional_Networks_for_Images_Speech_and_Time-Series.
- [5] Ян Гудфеллоу, Йошуа Бенгіо та Аарон Курвіль (2016). «Книга глибокого навчання». Доступно - <https://www.deeplearningbook.org/>
- [6] Кейрон О'Ші, Райан Неш. (2015). «Вступ до згорткових нейронних мереж». Доступно – <https://arxiv.org/abs/1511.08458>
- [7] Хосейн Голамалінежад1, Хосейн Хосраві. «Методи об'єднання в глибокі нейронні мережі» (2009). Доступно – <https://arxiv.org/pdf/2009.07485.pdf>
- [8] Шів Рам Дабі, Сатіш Кумар Сінгх, Бідют Баран Чаудхурі. (2021). «Функції активації в глибокому навчанні: всебічний огляд і контроль». Доступно – <https://arxiv.org/abs/2109.14545>
- [9] Джеффри Хінтон (2022). «Алгоритм вперед-наперед: деякі попередні дослідження». Доступно – <https://arxiv.org/abs/2212.13345>
- [10] Катажина Яноха, Войцех Маріан Чарнецький. (2017). «Про функції втрат для глибоких нейронних мереж з класифікації». Доступно – <https://arxiv.org/abs/1702.05659>
- [11] Девід Е. Румелхарт, Джеффри Е. Хінтон і Рональд Дж. Вільямс. (1986). «Навчання представлень шляхом зворотного поширення помилок». Доступно – <https://www.nature.com/articles/323533a0>
- [12] Вей Лю, Драгомір Ангелов, Думітру Ерхан, Крістіан Сегеді, Скотт Рід, Ченг-Ян Фу, Олександр С. Берг (2015) «SSD: Single Shot MultiBox Detector». Доступно – <https://arxiv.org/abs/1512.02325>.
- [13] Бернацкі, Маріуш; Влодарчик, Пшемислав (2004). «Принципи навчання багаторівневої нейронної мережі з використанням зворотного поширення». Доступно - http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html
- [14] Рафаель Паділья, Серхіо Л. Нетто, Едуардо А. Б. да Сілва (2022). «Опитування показників продуктивності алгоритмів виявлення об'єктів». Доступно - <https://ieeexplore.ieee.org/document/9145130>

Дані про авторів

Роман Продан

Магістр

Кафедра комп'ютеризованих систем автоматизації

НУ «Львівська політехніка».

Бандери, 12, Львів, Україна, 79013

Електронна адреса: roman.prodan.mknuo.2022@lpnu.ua

ORCID: 0009-0000-9725-8582

Денис Шутка

Магістр

Кафедра комп'ютеризованих систем автоматизації

НУ «Львівська політехніка».

Бандери, 12, Львів, Україна, 79013

Електронна адреса: denys.shutka.mknuo.2022@lpnu.ua

ORCID: 0009-0008-9759-279X

Василь Татарин

Кандидат технічних наук, доцент

Кафедра комп'ютеризованих систем автоматизації

НУ «Львівська політехніка».

Бандери, 12, Львів, Україна, 79013

E-mail: vasyl.y.tataryn@lpnu.ua

ORCID: 0000-0001-9740-1924