



DOI 10.28925/2663-4023.2024.26.698

УДК 004.8:[623.438+623.55.021]

Назаркевич Марія Андріївна

д.т.н., професор, професор кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID ID: 0000-0002-6528-9867
mariia.a.nazarkevych@lpnu.ua

Литвин Василь Володимирович

д.т.н., професор, професор кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID ID: 0000-0002-9676-0180
vasyl.v.lytvyn@lpnu.ua

Висоцька Вікторія Анатоліївна

д.т.н., професор, доцент кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID ID: 0000-0001-6417-3689
victoria.a.vysotska@lpnu.ua

МЕТОД РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ НА ОСНОВІ КЛАСИФІКАЦІЇ КАСКАДІВ ХААРА

Анотація. Розроблено метод розпізнавання рухомих об'єктів у відеопотоці на основі класифікації Хаара. При відстеженні об'єктів у виникає потреба ідентифікувати їх та фіксувати напрям їх руху, швидкість переміщення. Складність розпізнавання полягає не тільки у фіксації об'єкта та слідкування за ним, але й русі самої камери, з якої ведеться відеонагляд. В основі методу Хаара лежать каскадні класифікатори, які швидко виділяють регіони з високою ймовірністю виявлення об'єкта. У каскадах Хаара використовується операція згортки, яка формується на основі пропорційного добутку фур'є-образів функцій. До недоліків каскадів Хаара можна віднести те, що розпізнавання є нестійким при зміні освітлення, нестійке зі зміною масштабу та повороту ключових кадрів. При реалізації цього методу ніхто не здійснює зміну фонів у відеопослідовностях. Цей метод є дуже швидким у реалізації, і відповідно найменш точним, порівняно зі SURF і SIFT. Проте він є доступний до програмування і вільний у використанні. Для застосування Haar Cascades використано класифікатор Adaboost. Це алгоритм вибирає невелику кількість значущих характеристик із більшого набору, щоб забезпечити ефективний результат. Adaboost є методом ансамблевого навчання, який належить до категорії бустингових алгоритмів, що дозволяє комбінувати моделі дерев рішень з малою глибиною, щоб створити сильну модель, здатну забезпечувати високу точність класифікації чи регресії. Окрім розпізнавання об'єкта, реалізовано метод машинного навчання, що базується на методах з вчителем для реалізації прогнозування місцезнаходження об'єкта та ідентифікації цього об'єкта. У навчальній вибірці були «навчені» воєнні машини btr, bmp, танк, автомобіль та гаубиці. Планується використати для ідентифікації об'єкту алгоритми випадкового лісу, SVM, градієнтного бустингу та нейронні мережі. Розглянуто метрики результатів машинного навчання, зокрема оцінки точності, повноти, F1-score, коефіцієнт Каппа та матрицю помилок. Приведено оцінку розроблених моделей. У подальшому планується вдосконалювати початі методи.

Ключові слова: відеопотік; інформаційна система; розпізнавання; ідентифікація.



ВСТУП

Виявлення рухомих об'єктів — це відокремлення рухомих об'єктів від складного фону [1]. Рухомий об'єкт у цьому методі є переднім планом. Багато програм для відеоспостереження та ідентифікації потребують виявлення рухомих об'єктів в полі зору стаціонарної камери. Популярним методом отримання цих силуетів є процес віднімання фону.

В даний час для виявлення рухомих об'єктів використовується в основному кадровий метод віднімання, метод віднімання фону та метод оптичного потоку. Метод оптичного потоку полягає в обчисленні поля оптичного потоку зображення і виконанні кластера обробка відповідно до характеристик розподілу оптичного потоку зображення [2].

Відеоспостереження з трекінгом — це технологія, яка дозволяє автоматично відстежувати рух об'єктів на відеозаписі за допомогою розробленого програмного забезпечення і обладнання. Системи відеоспостереження з трекінгом і машинним навчанням застосовуються в різних сферах, таких як безпека (нагляд за об'єктами, виявлення вторгнень), транспорт (моніторинг дорожнього руху, управління транспортними потоками), промисловість (контроль за виробничими процесами), медицина (моніторинг пацієнтів) тощо.

Постановка проблеми. Дослідження присвячено побудові методу розпізнавання рухомих об'єктів на основі класифікації каскадів Хаара.

Аналіз останніх досліджень і публікацій. Ідентифікація рухомих об'єктів у відеопотоках даних розглянуто у [1]. Там представлено мультимодальний підхід для ідентифікації особи на основі відео в неконтрольованому середовищі, який підкреслює тривимірну форму тіла, позу, і зовнішній вигляд. Представлено два кодери: PSE кодує форму тіла за допомогою двійкових силуетів, рухів скелета та тривимірної форми тіла, а AAE забезпечує два рівні агрегації ознак зовнішнього вигляду: агрегацію ознак на основі уваги та усереднення агрегація. Для агрегування ознак на основі уваги ми використано просторову та часову увагу, щоб зосередитися на ключових областях для розрізнення особи.

У [2] представляємо SEAS, використовуючи Shape-Aligned Supervision, щоб покращити функцію зовнішнього вигляду для повторної ідентифікації людини. Ми використовуємо неявне представлення форми тіла, щоб контролювати навчання кодера ідентифікації на основі зовнішнього вигляду, із екстрактором форм, щоб перевести карту функцій у форми тіла на рівні пікселів, забезпечуючи піксельний рівень керівництво формою. Для повторної ідентифікації рухомого об'єкту за допомогою відео включено часову узгодженість у появі різних кадрів в одному відео шляхом додавання калібрування на рівні пікселів. Такий підхід забезпечує найсучаснішу продуктивність у оцінках повторної ідентифікації особи як на основі кадрів, так і на основі відео на загальнодоступних наборах даних. Процес детекції сильно залежить від сторонніх факторів, серед яких — технічні можливості камери, та стабільність її положення, яка є ключовою в процесі детекції руху. До них належить стабільне без мерехтіння світло та висока частота кадрів і показники роздільної здатності камери (рис. 1).

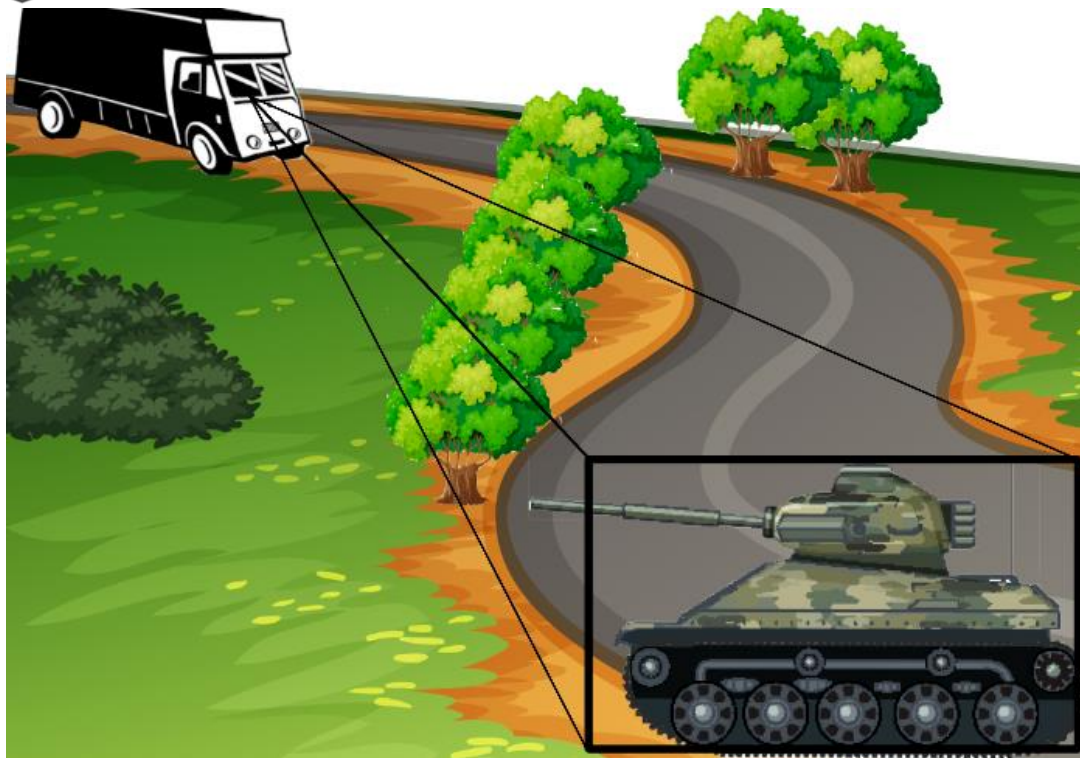


Рис. 1. Функціональна схема відеонагляду

Такий підхід дозволяє значно підвищити ефективність систем відеоспостереження, зменшити витрати на обслуговування і забезпечити швидке реагування на події у реальному часі.

Основні аспекти відеоспостереження з трекінгом включають відстеження руху об'єктів. Система здатна автоматично визначати та відстежувати рух певних об'єктів (наприклад, людей, автомобілів, тварин тощо) на відеозаписі. Це може бути корисно для виявлення небажаних об'єктів або моніторингу безпеки. Для ефективної роботи системи трекінгу використовуються спеціалізовані алгоритми та програмне забезпечення, які аналізують відеопотік і виконують відстеження об'єктів. Відстеження може здійснюватися за допомогою комп'ютерного зору, інфрачервоних датчиків, радіохвильових систем, чи використання кількох вище перелічених технологій.

Відеоспостереження з трекінгом, що використовує машинне навчання, це поєднання комп'ютерного зору, аналізу відеопотоків та використання алгоритмів глибокого навчання для автоматичного виявлення і відстеження об'єктів на відеозаписах.

Система спостереження [3] отримує відеопотік з камери (або декількох камер) і обробляє кожен кадр для виявлення об'єктів, які потрібно трекувати. Застосування алгоритмів комп'ютерного зору і машинного навчання дозволяє системі автоматично виявляти об'єкти на кадрі. Це можуть бути машини, люди, тварини або інші об'єкти, які цікавлять систему. Після виявлення об'єктів система використовує алгоритми трекінгу, які слідкують за рухом об'єктів у часі. Це дозволяє системі не тільки відстежувати моментальне положення об'єктів, але й прогнозувати їхній майбутній шлях на основі історії руху. Використання глибоких нейронних мереж дозволяє системі вдосконалювати відстеження об'єктів шляхом аналізу великої кількості даних і покращення точності трекінгу. Моделі можуть навчатися розпізнавати не тільки форми об'єктів, але й їхні унікальні риси та характеристики руху.



Метод Хаара [4] є одним із найбільш ранніх і ефективних алгоритмів для виявлення об'єктів, у зображеннях. Він був розроблений на основі перетворення Хаара і використовується для створення фільтрів, здатних виявляти текстури та контури об'єктів. На основі методу Хаара лежать каскадні класифікатори, які можуть швидко виділити регіони з високою ймовірністю імовірності об'єкта.

Хоча метод Хаара досить швидкий, він має низькі обмеження: точність його роботи значно знижується при зміні освітлення, курсу, а також при низькій якості зображення. Крім того, він вимагає складного попереднього навчання для створення надійних каскадів для кожного типу об'єкта.

Локальні бінарні шаблони (LBP) [5] Метод локальних бінарних шаблонів (LBP) є популярним підходом для видалення текстурних ознак. LBP перетворює кожен піксель у бінарний шаблон, аналізуючи його значення рядних сусідніх пікселів. Цей підхід є достатньо стійким до змін освітлення, проте його ефективність обмежена, коли йдеться про складні та великий обсяг даних.

LBP є зручним для використання в реальному часі завдяки низькій обчислювальній міцності, однак має обмежену точність при розпізнаванні обличчя з низькою якістю зображення. Традиційно, LBP використовували в системах, де точність не була критичним параметром.

Мащини опорних векторів (SVM) [6] використовує навчання з наглядом для класифікації зображення облич. SVM працює шляхом побудови гіперплощини, яка розділяє зразки двох класів з максимальним зазором між ними. У випадку розпізнавання облич SVM можна використовувати для класифікації ознак, видалених за допомогою методів зниження розмірності.

Недоліком SVM є складність обробки великих наборів даних, більшість збільшення гіперплощини може бути ресурсомістким. Крім того, для досягнення оптимальної точності SVM потрібні додаткові обчислювальні ресурси та налаштування параметрів.

Метод PCA [7] (аналіз основних компонентів). Аналіз головних компонентів (PCA) — це метод зниження розмірності, який часто використовується в системах розпізнавання обличчя. PCA шукає ортогональні компоненти в наборі даних і зберігає ті, що містять найбільшу варіативність. Це дозволяє зменшити обсяг даних без втрати важливих інформаційних ознак.

PCA часто використовується в комбінації з SVM, що покращує точність класифікації. Однак PCA має своє обмеження, оскільки втрати даних можуть негативно вплинути на результат, особливо при складних конфігураціях одягу.

Згорткові нейронні мережі (CNN) [8]. Згорткові нейронні мережі (CNN) є найпопулярнішою архітектурою для розпізнавання зображення, що включає обличчя. CNN складаються із шарів згорток, які виділяють просторові ознаки, і шарів об'єднань, що зменшують розмірність вхідних даних. Завдяки автоматичному виділенню ознаки CNN не потребує складної попередньої зображень, що значно спрощує процес побудови моделі.

CNN є надзвичайно стійкими до змін в освітленні та курсі, що робить їх оптимальними для розпізнавання обличчя. Крім того, CNN здатні працювати з великими наборами даних і ефективно узагальнювати отримані знання, що забезпечують високу точність.

Метою статті є розробка розпізнавання рухомих об'єктів на основі класифікації каскадів Хаара.

Для успішної реалізації поставленої мети були поставлені такі завдання:

- провести аналіз предметної області дослідження, зокрема розкрити теоретичні засади розпізнавання рухомих об'єктів на основі класифікації каскадів Хаара;

- визначити, які існують аналоги інформаційної системи розпізнавання та відслідковування рухомих об'єктів, проаналізувати їх особливості, переваги та недоліки;
- вибрати оптимальний програмний засіб для розробки системи з усіх наявних варіантів;
- розробити інформаційну систему для відслідковування та розпізнавання рухомих об'єктів у відеопослідовностях на основі каскадів Хаара, протестувати її функціонал та продемонструвати результати тестування.

Об'єкт дослідження. є процес відслідковування та розпізнавання рухомих об'єктів у відеопослідовностях.

Предмет дослідження. є засоби і методи, які використовуються для створення інформаційної системи відслідковування та розпізнавання рухомих об'єктів у відеопослідовностях.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Ідентифікація в контексті комп'ютерного зору означає визначення конкретного об'єкта або паттерну на зображенні чи відео. Основні етапи ідентифікації включають виявлення об'єктів, де система знаходить на зображенні або відео-області, що містять об'єкти, які їй потрібно аналізувати. Ознаки і дескриптори, де визначаються ключові ознаки або дескриптори об'єктів, які беруть участь в розпізнаванні. Використання навчених моделей або алгоритмів для встановлення класифікації або ідентифікації, до якого класу або категорії належить об'єкт.

Таким чином, поєднання ідентифікації об'єктів і 3D моделювання може значно покращити здатність системи до аналізу та інтерпретації оточуючого світу, що є важливим для багатьох сучасних застосувань, таких як віртуальна реальність, медична діагностика, автономні системи і багато інших (рис. 2).



Рис. 2. Процес ідентифікації об'єктів

Архітектура нейронної мережі

Розробивши потрібну натреновану модель, розглянемо її використання, на прикладі задачі класифікації. Щоби перепрофілювати навчену модель, потрібно замінити вихідний класифікатор на нові класи і донавчити модель на розроблених авторами даних. Існує три варіанти реалізації:

Навчання кількох згорткових шарів передбачає заморожування ваг початкових шарів нейромережі, які є відповідні загальним ознакам зображень, незалежних від завдання, та оновлення вищих рівнів, що стосуються конкретних функцій залежно від завдання.

Навчання класифікатора відповідає збереженню згорткової бази нейронної мережі та використанню її вихідних даних для подачі в новий класифікатор.

Навчання всієї моделі корисне за наявності великого набору даних та обчислювальної потужності. Якщо набір зображень маленький, більшість шарів заморожують, щоб уникнути оверфіту (overfit), коли модель запам'ятовує доступні зображення й погано узагальнюється. Якщо обчислювальної потужності не вистачає, набір даних невеликий та/або попередньо навчена модель розв'язує дуже схоже завдання, підійде класифікатор навчання. У цьому випадку попередньо навчена модель відіграє роль фіксованого механізму отримання ознак.

Метод розпізнавання на основі класифікації каскадів Хаара

Каскад Хаара [9] є набором примітивів, для яких застосовується операція згортки. Згортка $[]$ є математичною операцією двох функцій та дозволяє отримати третю функцію $(f * g)(t)$:

$$(f * g)(t) = \int f(t - \tau)g(\tau)d\tau \quad (1)$$

Основною властивістю згортки є те, що фур'є-образ згортки пропорційний добутку фур'є-образів функцій.

При використанні каскадів Хаара використовуються прості примітиви, що складаються з прямокутників і мають рівні, +1 і -1. При цьому кожен прямокутник використовується кілька разів різного розміру (див. рис. 3)

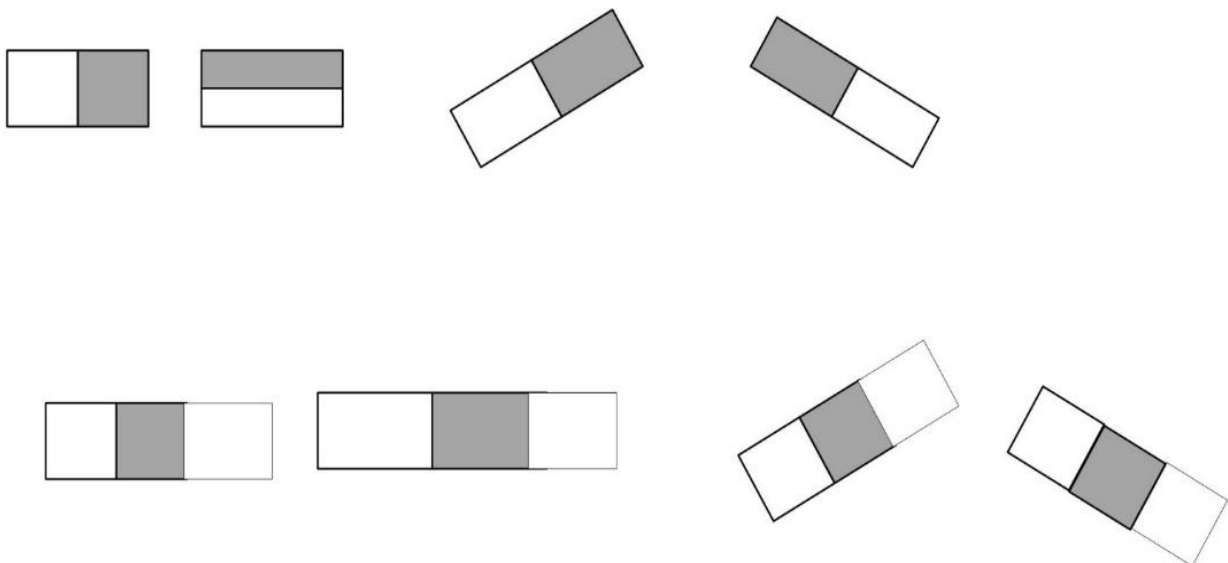


Рис. 3. Варіанти простих примітивів каскадів Хаара



У більшості випадків для простого порівняння двох достатньо схожих фрагментів зображення, його реалізують через їх коваріацію $cov(X, Y)$

$$cov(X, Y) = E([X - E(X)][Y - E(Y)]) \quad (2)$$

Каскади Хаара є базовими для побудови системи виділення складних об'єктів, таких як обличчя, руки чи інші об'єкти. Також часто разом з каскадами Хаара застосовується алгоритм навчання AdaBoosta. Каскад Хаара — це набір примітивів, для яких вважається їх згортка з зображенням. Використовуються самі прості примітиви, що складаються з прямокутників, які мають два рівні, +1 і -1. При цьому кожен прямокутник використовується кілька разів різного розміру. Під згорткою розуміємо $s = X \cdot Y$, де Y — сума елементів зображення в тінях, а X — сума елементів зображення в світловій області. Можна обчислювати X/Y , тоді алгоритм буде стійким при зміні масштабу.

При реалізації каскадів Хаара ми отримуємо структурну інформацію про об'єкт: наприклад, центр об'єкту, що рухається у відеопослідовності, якщо ідентифікуємо танк, завжди будуть темні примітиви. Чим більше використовуються різні примітиви, тим точніше можна потім класифікувати об'єкт. При цьому, якщо точна класифікація не потрібна — можна використовувати меншу кількість примітивів.

Після того, як побудовані набори ознак за тестовим вибором, алгоритми навчання, зокрема AdaBoost, SVM визначають каскад об'єктів. При розпізнаванні об'єкта на зображенні проводиться його порівняння з тестовим зображенням.

Каскади Хаара [10] можна широко використовувати через цілісне представлення зображень і обчислюється формулою

$$I(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j)$$

де $I(i, j)$ — поточний піксель.

МЕТОДИКА ДОСЛІДЖЕННЯ

Тестування моделі проводилося за допомогою метрики, яка відображає якість класифікації для завдання бінарної класифікації. Основними метриками, використовуваними для оцінки моделі, є **точність** (точність), **повнота** (запам'ятовування), **точність класифікації** (точність) та **F1-score** [11]. Крім того, для глибшого аналізу моделі було розраховано коефіцієнт Каппа (Cohen's Kappa), який показує узгодженість між передбаченими та реальними класами.

До метрик продуктивності належать оцінки точності, повноти, F1-score.

Точність (Precision) визначає, наскільки часто модель правильно приймає позитивний клас. Висока точність означає, що модель більшого правильно класифікує рухомих об'єктів.

$$Precision = TP / (TP + FP) \quad (3)$$

Повнота (Recall) показує, скільки істинно позитивних прикладів модель змогла правильно класифікувати. Висока повнота плями про те, що модель не пропускає помітне зображення обличчя.

$$Recall = TP / (TP + FN) \quad (4)$$

F1-score є середнім значенням точності та повноти і використовується як узагальнююча метрика для оцінки балансу між ними. F1-score особливо корисний для оцінки продуктивності моделі у випадку, коли класи можуть бути незбалансованими.

$$F1 - score = 2 * (Precision * Recall) / (Precision + Recall) \quad (5)$$

де TP — істинно позитивні передбачення, FP — хибно позитивні, а FN — хибно негативні.



Коефіцієнт Каппа: Коефіцієнт Каппа вказує на високу узгодженість між передбаченими моделями та реальними класами.

ROC-крива та AUC (Receiver Operating Characteristic) є графічним відображенням здатності моделі розрізняти класи при різних порогах. Площа під кривою (AUC — Area Under Curve) вказує на загальну здатність моделі відрізняти класи. Високе значення AUC (близько до 1) означає, що модель добре справляється з розрізненням класів, тоді як значення найближче до 0,5 вказує на випадкову класифікацію.

Матриця помилок є інструментом для оцінки продуктивності моделі, після чого вона показує кількість правильних та неправильних класифікацій для кожного класу. У випадку нашої матриці помилок можна побачити проаналізувати, як модель справляється з класифікацією обличчя (позначається як клас 1) та не обличчя (клас 0). Нижче наведено приклад типового вигляду матриці помилок:

$$\begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix} \quad (6)$$

де TP — істинно позитивні, FN — хибно негативні, FP — хибно позитивні, TN — істинно негативні класифікації. У разі нашої більшості помилок припадають на FP та FN, які можна вказати на обов'язкову додаткову настройку порогового значення для підвищення точності.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Беремо зразок зображення, а саме фрагмент зображення і переміщаємо по всій площині зображення згідно X і Y координат. На початку роботи алгоритму встановлюємо початкові точки пошуку, і обчислюємо коваріацію відмінності зразка (J) від оригінала зображення (I). Цей метод працює швидко.

До недоліків каскадів Хаара можна віднести те, що розпізнавання є нестійке при зміні освітлення, нестійке зі зміною масштабу та повороту зображення. Він є нестійким, якщо на частині зображення змінюється фон. Якщо потрібно знайти область pхp на зображенні mхm, то кількість операцій буде пропорційно $n^2 \times (m-n)^2$.

Щоб метод працював ефективно, здійснюють операції освітлення та нейтралізують операцією нормування або переходом до бінаризації області. Зміна масштабу та невеликі повороти нейтралізуються зміною роздільної здатності при кореляції. При реалізації цього методу з фоном ніхто не бореться, тобто фон не модифікує.

Якщо камери фіксують кадри з малою роздільною здатністю, то здійснюється оптимізація шляху пошуку рухомого об'єкту з великим кроком.

У ситуаціях, коли результати кореляції є незадовільні, переходять до більш складних методів, як порівняння особливих точок, границь або безпосереднього виділення об'єктів. Ці алгоритми у більшості випадків досить повільні, їх складно написати, є обмеження на структуру зображення.

Стойкість до змінного освітлення, навіть якщо це локальна змінна освітлення, стійкість до шумів (примітиви являють собою простіший смуговий фільтр).

Якщо примітиви були не дуже маленькі, то сильно стійка кореляція при зміні масштабу (розмір примітивів при цьому не буде впливати на точність, якщо обхід з маленьким кроком).



Якщо визнання на великому зображенні розчитати заздалегідь і при виведенні вікна пошуку брати вже перевірені й актуальні для нього — пошук буде значно швидше кореляції (потрібно порівняти меншу кількість елементів).

При цьому видно, що можна провести ряд досить простих оптимізацій, які прискорять і уточняють цей підхід.

Класифікатори використовуються для визначення рухомих об'єктів; це алгоритми, які визначають, чи або ні область зображення містить об'єкт (1) або не містить об'єкт (0) [12], [13]. Щоб досягти вищого рівня з точністю класифікатори навчили розпізнавати рухомий об'єкт шляхом аналізу від тисяч до мільйонів зображень. LBP (локальний бінарний шаблон) і каскади Хаара — це два класи класифікаторів, які використовуються OpenCV [14], [15]. Остання система класифікації буде використана в запропонованому підході. Вейвлет Хаара підхід є основою каскаду Хаара, який ділить пікселі зображення на квадрати на основі їх функції [16]. Це дозволяє використовувати кілька підходів машинного навчання для отримання високого ступеня. Виявлені «особливості» обчислюються за допомогою поняття «цілісного образу». Метод навчання Adaboost використовується Naar Cascades [17], [18]. Це алгоритм вибирає невелику кількість значущих характеристик із більшого набору, щоб забезпечити ефективний результат класифікатора.

Система приймає як вхідні дані цифрове зображення, зібране кадр за кадром із відеопотіку, знятий камерою. Для методу виділення ознак, подібного до Хаара, зображення має спочатку буде перетворено з колірному простору RGB (червоний-зелений-синій) на зображення у градаціях сірого. Розпізнавання обличчя

Процедура виконується із зображенням у градаціях сірого як початковою точкою. Функція, подібна до Хаара, використовується як основу для цього процесу. Якщо об'єкт не знайдено, система продовжить використовувати наступний кадр як вхідні дані. Крім того, якщо на зображенні знайдено об'єкт, наступним кроком є обрізання зображення таким чином, щоб воно генерувало зображення меншого розміру, яке містить лише компоненти об'єкта.

При трекінгу рухомих об'єктів в умовах різної видимості слід врахувати наявність системи GPS та навігації. Використання відеокамер для визначення місцезнаходження машин є основою для будь-якої системи трекінгу. Однак в умовах обмеженої видимості, таких як тунелі або висока забудова, сигнал з відеокамери може бути слабким або відсутнім взагалі. Тому можна передбачати додаткові системи навігації у якості інерціальних давачів або радарів для допомоги визначення місцезнаходження. Необхідно забезпечити надійний зв'язок та точну інформацію про місцезнаходження та стан машин в реальному часі. Тому можемо проектувати для вирішення цих завдань мережі мобільного зв'язку, супутникову комунікацію або мережі LoRaWAN для віддаленого моніторингу у важкодоступних місцях. Програмне забезпечення має містити такі алгоритми обробки даних: Врахування умов різної видимості вимагає розробки алгоритмів, які можуть адаптуватися до різних умов та забезпечувати надійний трекінг машин. Це може включати в себе фільтрацію шуму, корекцію помилок відеокамер, прогнозування місцезнаходження та інші методи. На наступному етапі відбувається візуалізація даних щодо місцезнаходження машин на мапі для зручного відстеження (рис. 4). Розробка інтуїтивного інтерфейсу користувача допоможе операторам ефективно контролювати рух машин в реальному часі. Для забезпечення безпеки даних Оскільки трекінгова інформація може містити конфіденційні дані про маршрути та розміщення машин, важливо забезпечити безпеку цих даних шляхом шифрування та інших методів захисту.



Рис. 4. Розпізнавання на основі каскадів Хаара

Таблиця 1

Отримані результати навчання за допомогою технік машинного навчання

Optimizer	Precision	Recall	map@0.5	map@0.95	F1 score
512×512					
Adam W	0,340	0,873	0,352	0,248	0,488
Adam	0,835	0,842	0,811	0,489	0,839
Adaboost	0,965	0,950	0,977	0,200	0,950
640×640					
Adam W	0,930	0,810	0,878	0,681	0,865
Adam	0,821	0,765	0,788	0,433	0,792
Adaboost	0,969	0,960	0,982	0,833	0,964
1024×1024					
Adam W	0,921	0,722	0,808	0,622	0,810
Adam	0,820	0,764	0,787	0,432	0,791
Adaboost	0,963	0,967	0,972	0,825	0,965

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Розроблено метод розпізнавання рухомих об'єктів у відеопотоці на основі класифікації Хаара.

Для забезпечення об'єктивної оцінки моделі розпізнавання рухомих об'єктів було використано датасет, який складався з зображень, що не використовувалися в процесі навчання. Це дозволяє оцінити загальну здатність моделі та її продуктивність у реальних умовах.

У навчальній вибірці були «навчені» воєнні машини btr, bmp, танк, автомобіль та гаубиці.

У навчальній вибірці були «навчені» воєнні машини btr, bmp, танк, автомобіль та гаубиці. Планується використати для ідентифікації об'єкту алгоритми випадкового лісу, SVM, градієнтного бустингу та нейронні мережі. Розглянуто метрики результатів машинного навчання, зокрема оцінки точності, повноти, F1-score, коефіцієнт Каппа та матрицю помилок.

При розробленні моделі з Haar Cascades використано класифікатори Adam W, Adam, Adaboost.

Аргументація даних мала позитивний вплив на загальну продуктивність моделі. Додавання таких методів, як повороти, масштабування та зміни яскравості, підвищили загальну здатність моделі, дозволивши їй краще розпізнати рухомі об'єкти у змінених умовах. Завдяки цій моделі досягнуто високої точності навіть при тестуванні на раніше невідомих даних.

**ПОДЯКА**

Дослідження виконано за підтримки Міністерства освіти і науки України «Методи та засоби ідентифікації бойових машин на основі технологій глибинного навчання для автоматизованого керування цілерозподілом» за проектом № 0124U000925.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Elhajian, S. Y., El-Sayed, K. M., & Ahmed, S. H. (2008). Moving object detection in spatial domain using background removal techniques-state-of-art. *Recent patents on computer science*, 1(1), 32–54.
2. Chien, S. Y., Ma, S. Y., & Chen, L. G. (2002). Efficient moving object segmentation algorithm using background registration technique. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(7), 577–586.
3. Valera, M., & Velastin, S. A. (2005). Intelligent distributed surveillance systems: a review. *IEE Proceedings-Vision, Image and Signal Processing*, 152(2), 192–204.
4. Awati, V. B., Goravar, A., & Kumar, M. (2024). Spectral and Haar wavelet collocation method for the solution of heat generation and viscous dissipation in micro-polar nanofluid for MHD stagnation point flow. *Mathematics and Computers in Simulation*, 215, 158–183.
5. Chang, M., Ji, L., & Zhu, J. (2024). Multi-scale LBP fusion with the contours from deep CellNNs for texture classification. *Expert Systems with Applications*, 238, 122100.
6. Chen, P. H., Fan, R. E., & Lin, C. J. (2024). A study on SMO-type decomposition methods for support vector machines. *IEEE transactions on neural networks*, 17(4), 893–908.
7. Dorabiala, O., Aravkin, A., & Kutz, J. N. (2024). Ensemble Principal Component Analysis. *IEEE Access*.
8. Arsalan, M., Mubeen, M., Bilal, M., & Abbasi, S. F. (2024, August). 1D-CNN-IDS: 1D CNN-based Intrusion Detection System for IIoT. 29th International Conference on Automation and Computing (ICAC), 1–4.
9. Polvivaara, A., Korpela, I., Vastaranta, M., & Junttila, S. (2024). Detecting tree mortality using waveform features of airborne LiDAR. *Remote Sensing of Environment*, 303, 114019.
10. Vijaya, J., Singh, A. P., Ekka, M., Navya, P., & Otti, S. A. (2024, September). Face Recognition System Using Haar Cascade Algorithm. *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)*, 1–5.
11. Meng, B., Sun, J., & Shi, B. (2024). A novel URP-CNN model for bond credit risk evaluation of Chinese listed companies. *Expert Systems with Applications*, 255, 124861.
12. Saad, R. S. M., Moussa, M. M., Abdel-Kader, N. S. et al. (2024). Deep video-based person re-identification (Deep Vid-ReID): comprehensive survey. *EURASIP J. Adv. Signal Process.* 63. <https://doi.org/10.1186/s13634-024-01139-x>
13. Dong, N., Yan, S., Tang, H., Tang, J., & Zhang, L. (2024). Multi-view information integration and propagation for occluded person re-identification. *Information Fusion*, 104, 102201.
14. Nguyen, V. D., Mirza, S., Zakeri, A., Gupta, A., Khaldi, K., Aloui, R., & Merchant, F. (2024). Tackling Domain Shifts in Person Re-Identification: A Survey and Analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4149–4159.
15. Nazarkevych, M., Petrov, A., Onopriyuk, O., Oleksiv, N., & Kis, Y. (2022). Development of a Fingerprint Pattern Matching Method Using K-Means. *Electronics and information technologies*, 19, 58–65. <http://dx.doi.org/10.30970/eli.19.5>
16. Nazarkevych, M., Voznyi, Y., & Nazarkevych, H. (2021). Development of Machine Learning Method with Biometric Protection with New Filtration Methods. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique"*, 3(11), 16–30. <https://doi.org/10.28925/2663-4023.2021.11.1630>
17. Hrytsyk, V., Nazarkevych, M., & Dyshko, A. (2020). Comparative Analysis of Image Recognition Methods Obtained From Sensors of the Visible Spectrum. *Electronic Professional Scientific Journal «Cybersecurity: Education, Science, Technique»*, 4(8), 149–164. <https://doi.org/10.28925/2663-4023.2020.8.149164>
18. Gamil, S., Zeng, F., Alrifay, M., Asim, M., & Ahmad, N. (2024). An Efficient AdaBoost Algorithm for Enhancing Skin Cancer Detection and Classification. *Algorithms*, 17(8), 353.

**Mariia Nazarkevych**

Doctor of Technical Sciences, Professor,
Professor of the Department of Information Systems and Networks
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-6528-9867
mariia.a.nazarkevych@lpnu.ua

Vasyl Lytvyn

Doctor of Technical Sciences, Professor,
Professor of the Department of Information Systems and Networks
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-9676-0180
vasyl.v.lytvyn@lpnu.ua

Victoria Vysotska

Doctor of Technical Sciences, Professor,
Professor of the Department of Information Systems and Networks
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0001-6417-3689
victoria.a.vysotska@lpnu.ua

METHOD OF RECOGNITION OF MOVING OBJECTS BASED ON THE CLASSIFICATION OF HAAR CASCADES

Abstract. A method of recognition of moving objects in a video stream based on the Haar classification has been developed. When tracking objects, there is a need to identify them and record their direction of movement, speed of movement. The complexity of recognition lies not only in fixing the object and following it, but also in the movement of the camera itself, from which video surveillance is conducted. The Haar method is based on cascade classifiers that quickly highlight regions with a high probability of detecting an object. Haar cascades use a convolution operation, which is formed on the basis of the proportional product of Fourier images of functions. The disadvantages of Haar cascades include the fact that recognition is unstable when lighting changes, unstable with changes in scale and rotation of key frames. When implementing this method, no one changes the backgrounds in the video sequences. This method is very fast to implement, and accordingly the least accurate, compared to SURF and SIFT. However, it is accessible to programming and free to use. The Adaboost classifier was used to apply Haar Cascades. This algorithm selects a small number of significant features from a larger set to provide an effective result. Adaboost is an ensemble learning method that belongs to the category of boosting algorithms, which allows combining decision tree models with a small depth to create a strong model capable of providing high accuracy of classification or regression. In addition to object recognition, a machine learning method based on supervised methods was implemented to implement object location prediction and object identification. The training sample included military vehicles btr, bmp, tank, car and howitzer. It is planned to use random forest, SVM, gradient boosting and neural networks algorithms for object identification. The metrics of machine learning results are considered, in particular, the accuracy, completeness, F1-score, Kappa coefficient, and error matrix. The developed models are evaluated. In the future, it is planned to improve the methods that have been started.

Keywords: video stream; information system; recognition; identification.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Elhabian, S. Y., El-Sayed, K. M., & Ahmed, S. H. (2008). Moving object detection in spatial domain using background removal techniques-state-of-art. *Recent patents on computer science*, 1(1), 32–54.
2. Chien, S. Y., Ma, S. Y., & Chen, L. G. (2002). Efficient moving object segmentation algorithm using background registration technique. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(7), 577–586.



3. Valera, M., & Velastin, S. A. (2005). Intelligent distributed surveillance systems: a review. *IEE Proceedings-Vision, Image and Signal Processing*, 152(2), 192–204.
4. Awati, V. B., Goravar, A., & Kumar, M. (2024). Spectral and Haar wavelet collocation method for the solution of heat generation and viscous dissipation in micro-polar nanofluid for MHD stagnation point flow. *Mathematics and Computers in Simulation*, 215, 158–183.
5. Chang, M., Ji, L., & Zhu, J. (2024). Multi-scale LBP fusion with the contours from deep CellNNs for texture classification. *Expert Systems with Applications*, 238, 122100.
6. Chen, P. H., Fan, R. E., & Lin, C. J. (2024). A study on SMO-type decomposition methods for support vector machines. *IEEE transactions on neural networks*, 17(4), 893–908.
7. Dorabiala, O., Aravkin, A., & Kutz, J. N. (2024). Ensemble Principal Component Analysis. *IEEE Access*.
8. Arsalan, M., Mubeen, M., Bilal, M., & Abbasi, S. F. (2024, August). 1D-CNN-IDS: 1D CNN-based Intrusion Detection System for IIoT. 29th International Conference on Automation and Computing (ICAC), 1–4.
9. Polvivaara, A., Korpela, I., Vastaranta, M., & Junttila, S. (2024). Detecting tree mortality using waveform features of airborne LiDAR. *Remote Sensing of Environment*, 303, 114019.
10. Vijaya, J., Singh, A. P., Ekka, M., Navya, P., & Otti, S. A. (2024, September). Face Recognition System Using Haar Cascade Algorithm. *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)*, 1–5.
11. Meng, B., Sun, J., & Shi, B. (2024). A novel URP-CNN model for bond credit risk evaluation of Chinese listed companies. *Expert Systems with Applications*, 255, 124861.
12. Saad, R. S. M., Moussa, M. M., Abdel-Kader, N. S. et al. (2024). Deep video-based person re-identification (Deep Vid-ReID): comprehensive survey. *EURASIP J. Adv. Signal Process.* 63. <https://doi.org/10.1186/s13634-024-01139-x>
13. Dong, N., Yan, S., Tang, H., Tang, J., & Zhang, L. (2024). Multi-view information integration and propagation for occluded person re-identification. *Information Fusion*, 104, 102201.
14. Nguyen, V. D., Mirza, S., Zakeri, A., Gupta, A., Khaldi, K., Aloui, R., & Merchant, F. (2024). Tackling Domain Shifts in Person Re-Identification: A Survey and Analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4149–4159.
15. Nazarkevych, M., Petrov, A., Onoprychuk, O., Oleksiv, N., & Kis, Y. (2022). Development of a Fingerprint Pattern Matching Method Using K-Means. *Electronics and information technologies*, 19, 58–65. <http://dx.doi.org/10.30970/eli.19.5>
16. Nazarkevych, M., Voznyi, Y., & Nazarkevych, H. (2021). Development of Machine Learning Method with Biometric Protection with New Filtration Methods. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique"*, 3(11), 16–30. <https://doi.org/10.28925/2663-4023.2021.11.1630>
17. Hrytsyk, V., Nazarkevych, M., & Dyshko, A. (2020). Comparative Analysis of Image Recognition Methods Obtained From Sensors of the Visible Spectrum. *Electronic Professional Scientific Journal «Cybersecurity: Education, Science, Technique»*, 4(8), 149–164. <https://doi.org/10.28925/2663-4023.2020.8.149164>
18. Gamil, S., Zeng, F., Alrifaey, M., Asim, M., & Ahmad, N. (2024). An Efficient AdaBoost Algorithm for Enhancing Skin Cancer Detection and Classification. *Algorithms*, 17(8), 353.

