

Олег Жеребух¹, Ігор Фармага²

¹Кафедра систем автоматизованого проектування, Національний університет “Львівська політехніка”, вул. Степана Бандери 12, Львів, Україна, E-mail: oleh.zherebukh.mnknm.2022@lpnu.ua

²Кафедра систем автоматизованого проектування, Національний університет “Львівська політехніка”, вул. Степана Бандери 12, Львів, Україна, E-mail: ihor.v.farmaha@lpnu.ua

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЗНАЧЕННЯ ОБ’ЄКТІВ НА ЗОБРАЖЕННІ

Отримано: березень 12, 2024 / Переглянуто: березень 28, 2024 / Прийнято: квітень 01, 2024

© Жеребух О., Фармага І., 2024

<https://doi.org/>

Анотація. Розроблено модифіковану модель нейронної мережі на базі Yolo V5 та здійснено порівняння метрик якості класифікації об’єктів на відеозображеннях, побудованих на основі базових існуючих відомих архітектур нейронних мереж. Розглянуто застосування згорткових нейронних мереж для обробки зображень з камер відеоспостереження з метою розробки оптимізованого алгоритму для виявлення та класифікації об’єктів на відеозображеннях. Зроблено аналіз існуючих моделей та архітектур нейронних мереж для аналізу зображень і здійснено їх порівняння. Розглянуто можливості оптимізації процесу аналізу зображень за допомогою використання нейронних мереж.

Ключові слова: згорткові нейронні мережі, CNN, виявлення об’єктів, швидкодія обробки відеозображень, властивості та ознаки зображень.

Вступ

Запит на аналіз зображень за допомогою нейронних мереж знаходить віддзеркалення в сучасному світі, де інформація представлена у візуальному вигляді і стає все більш незворотньою частиною нашого життя. Актуальність цього підходу виражається в декількох ключових аспектах.

По-перше, реальність інформаційного суспільства вимагає швидкого та точного аналізу великої кількості візуальних даних. З ростом кількості цих даних нейронні мережі надають змогу автоматизовано розпізнавати об’єкти, аналізувати контент і виділяти суттєву інформацію з великих наборів даних. По-друге, в сфері безпеки і відстеження системи аналізу зображень з нейронними мережами допомагають виявляти небезпечні ситуації, вторгнення і відстежувати об’єкти, забезпечуючи безпеку в об’єктах великого масштабу, таких як аеропорти і міські вулиці. По-третє, в промисловості, автоматизації та виробництві аналіз зображень використовується для підвищення ефективності процесів та оптимізації виробничих ліній. Нейронні мережі допомагають машинам бачити та розуміти своє оточення, що веде до підвищення автоматизації та виробничої продуктивності [1].

Об’єктом дослідження даної роботи є процес відеоспостереження та можливість в реальному часі аналізувати та класифікувати зображення, отримані з відеокамер спостереження.

Предметом дослідження є засоби та методи для оптимізації алгоритму класифікації відеозображень. Це включає в себе оптимізацію алгоритмів обробки відеоданих, адаптацію до різних умов та вирішення проблеми обробки великого обсягу даних в реальному часі.

Метою цієї статті є розробка модифікованої нейронної мережі для вдосконалення ефективності та точності класифікації об’єктів на зображеннях з відеокамер спостереження в режимі реального часу. Ключовою особливістю цієї мережі є вдосконалення здатності систем відеоспостереження в автоматичному режимі виявляти та класифікувати об’єкти, такі як транспортні

засоби, особи чи інші об'єкти інтересу. Покращення цих функцій сприятиме підвищенню рівня безпеки, ефективності та зручності застосування систем відеоспостереження в широкому спектрі відомостей.

Наукова новизна полягає у оптимізації алгоритму відеоспостереження через розробку модифікованої нейронної мережі для класифікації зображень з відеокамер спостереження в режимі реального часу з можливістю подальшої адаптивності.

Щодо практичного значення, розроблена модифікована нейронна мережа має велику практичну цінність для систем відеоспостереження. Покращення точності класифікації відеозображень сприятиме підвищенню ефективності та достовірності виявлення об'єктів, таких як транспортні засоби чи особи, що є критичним для забезпечення безпеки об'єктів спостереження. Крім того, можливість роботи в режимі реального часу розширює застосування систем відеоспостереження в таких галузях, як безпека на транспорті, міське планування та забезпечення громадської безпеки.

Огляд сучасних літературних джерел за тематикою публікації

Розумні системи відеоспостереження вимагають активного визначення та виявлення подій для уникнення непорозумінь та катастроф. Протягом останнього десятиліття збільшилася кількість природних катастроф і великомасштабних подій під час протестів, що спонукало дослідників аналізувати відеодані від систем відеоспостереження з використанням штучного інтелекту (ШІ). Це також призвело до створення наборів даних для наукових досліджень. Цей розділ надає перегляд досягнень у відеоаналізі та дає уявлення про існуючі системи відеоспостереження та набори даних.

Останні два десятиліття із зростанням міської забудови та населення збільшилася важливість використання відеокамер для громадського спостереження. У більшості традиційних систем відеоспостереження відеозаписи аналізувалися вручну, що призводило до реактивних рішень, а не превентивних. З подальшим розвитком сучасних візуальних сенсорів та алгоритмів штучного інтелекту (ШІ) стало можливим приймати превентивні рішення.

Однією з перших розумних систем відеоспостереження була розроблена IBM для виявлення підозрілої діяльності на парковках, розпізнавання облич, розпізнавання номерних знаків та ідентифікації для контролю доступу [2]. Ця система використовувала Middleware for Large Scale Surveillance (MILS), інтегрований із веб-сервісами для управління даними. Інша система від Fernandez et al. [3] збирала дані від численних візуальних сенсорів Інтернету речей (IoT) для посилення роботи екстреної команди із розподілом відеопотоку та сповіщеннями. Дані від візуальних сенсорів надходили у формі XML і використовувалися для створення семантичного двигуна з онтологією на основі знань для виявлення маршруту руху транспортних засобів та відстеження аномальних траєкторій. В інших роботах [4] вивчалися деталі автомобілів для визначення марки, моделі, кольору та номерного знака за допомогою аналізу низькорівневих ознак з використанням SURF та оптичного розпізнавання символів Tesseract OCR.

Крім того, виробники розумних камер надають рішення для виявлення порушень правил дорожнього руху [5], [6], виявлення пострілів [7], виявлення блукання [5], [6], розпізнавання номерних знаків [5], [6], а також підозрілих людських дій, таких як сутички, біг або падіння [8]. Розгортання розумних систем відеоспостереження для моніторингу дорожнього руху [5], [6], виявлення залишених об'єктів [5], виявлення радіоактивних ізотопів [6] та інтелектуальної маршрутизації [6] також стало актуальним з виникненням концепції розумного міста. Проте жоден з підходів, крім [8] та [9], не може використовуватися для аналізу поведінки толпи. Підхід [8] не враховує складні поведінки при збільшенні густини толпи, тоді як [9] не проводить експериментального аналізу в реальному середовищі.

Протягом останніх років відбулися значні просування у відеоаналізі за допомогою глибокого навчання [10]. Зокрема, багато робіт зосереджувалися на визначенні людської активності (HAR)– завдання розпізнавання конкретних дій людини з послідовності кадрів. HAR привернуло увагу після того, як було показано корисність технік глибокого навчання у відеоаналізі. Tran et al. [11] спростували двовимірні згорткові нейронні мережі (2D CNNs) в тривимірні (3D CNNs), здатні обробляти просторово-часові особливості для аналізу серій кадрів чи відео.

У світі нейронних мереж існує багато типів та архітектур, але два основних типи, які здобули широкий розповсюдження в аналізі зображень, – це згорткові нейронні мережі (Convolutional Neural Networks, CNN) та рекурентні нейронні мережі (Recurrent Neural Networks, RNN). Кожен з цих типів мереж має свої унікальні особливості та застосування [12].

Згорткові нейронні мережі: Згорткові нейронні мережі стали важливим інструментом для аналізу зображень та відео. Вони спеціалізуються на виявленні шаблонів та особливостей візуальної інформації на зображеннях. Основними компонентами CNN є згортка та пулінг, які дозволяють мережі виявляти різні фільтри та області на зображенні та виділяти їх важливі риси. CNN ефективно використовуються для класифікації зображень, детекції об'єктів, сегментації та багатьох інших завдань, пов'язаних з обробкою зображень.

Рекурентні нейронні мережі: Рекурентні нейронні мережі спроектовані для роботи з послідовністю даних, таких як текст, аудіо або відео. Їх головною особливістю є здатність враховувати контекст і історію даних завдяки внутрішнім станам. Це дозволяє RNN здійснювати передбачення на основі попередніх даних в послідовності. RNN використовуються в машинному перекладі, аналізі тексту, генерації мови та інших завданнях, де важлива урахування контексту [12].

Ці два типи нейронних мереж, CNN і RNN, різні за своєю природою та застосуваннями, і часто використовуються в поєднанні для складних завдань аналізу зображень та обробки відображення послідовностей.

Спроектовані та розроблені нейронні мережі мають велике значення в галузі аналізу зображень, і деякі з них стали особливо відомими завдяки своїм вражаючим характеристикам. Нижче наведений огляд кількох таких відомих архітектур [13]:

1. VGG (Visual Geometry Group): VGG – це серія нейронних мереж, розроблена в Університеті Оксфорда, яка відзначається глибокими архітектурами та великою кількістю шарів. Типова VGG мережа має 16-19 шарів і дуже просту структуру: послідовність згорткових і пулінгових шарів, завершених повнозв'язним шаром для класифікації. Хоча VGG вважається застарілою для більш вимогливих завдань, її простота дозволяє легше зрозуміти основи глибокого навчання.

2. ResNet (Residual Network): ResNet відома своєю глибиною та впровадженням концепції "резидуальних блоків". Замість того, щоб намагатися навчати точну функцію, ResNet навчається різниці між вхідним і вихідним значеннями (резидуальні блоки). Це допомагає уникнути проблем градієнта, що зазвичай виникають при навчанні дуже глибоких мереж. ResNet здатна до розрізнення тонких деталей на зображеннях і використовується для багатьох завдань від класифікації до сегментації.

3. Inception (GoogLeNet): Inception (також відома як GoogLeNet) відома своєю архітектурою, що використовує блоки Inception з різними розмірами фільтрів одночасно. Це дозволяє ефективно виявляти різні властивості на зображенні на різних шарах мережі. Inception здатна до зменшення кількості параметрів та при цьому забезпечує високу продуктивність.

4. MobileNet: MobileNet – це архітектура, спеціально розроблена для вбудованих пристроїв та мобільних застосувань. Вона відзначається легкістю та швидкістю, зберігаючи високу точність. MobileNet часто використовується для завдань відстеження об'єктів та класифікації на мобільних пристроях.

5. U-Net: U-Net – це архітектура, спрямована на завдання сегментації зображень, зокрема для медичних зображень. Вона має характерну форму букви "U" та включає в себе згорткові та транспоновані згорткові шари для виявлення об'єктів та відновлення високо деталізованих сегментів.

Нейронні мережі здійснюють вражаючий прорив в аналізі зображень і знайшли застосування в різних галузях. Нижче подаються приклади таких застосувань [14]:

1. Класифікація об'єктів та образів: Нейронні мережі, зокрема згорткові нейронні мережі (CNN), широко використовуються для класифікації об'єктів та образів на зображеннях. Наприклад, вони можуть розпізнавати види тварин, відмічати бренди автомобілів, або навіть діагностувати хвороби на медичних зображеннях.

Використання нейронних мереж для визначення об'єктів на зображенні

2. Детекція та відстеження об'єктів: Нейронні мережі використовуються для відстеження та детекції об'єктів на відеозаписах та в реальному часі. Вони можуть виявляти пішоходів на дорозі, слідити за рухом автомобілів або виявляти об'єкти на навігаційних зображеннях дронів.

3. Сегментація зображень (рис. 1): Нейронні мережі можуть виділяти окремі області на зображенні, що називається сегментацією. Це застосовується в медицині для виділення органів на знімках, в автономних автомобілях для розрізнення дороги та об'єктів і в астрономії для аналізу зображень зоряних небесних об'єктів.

4. Генерація зображень та мистецтвозедення: З використанням глибоких генеративних моделей, таких як генеративні згорткові нейронні мережі (GAN), можна створювати нові зображення, що не мають аналогів у реальному світі. Це має застосування в мистецтвозеденні, графіці та синтезі зображень.



Рис. 1. Приклад результату сегментації зображення у чорно-білому та кольоровому спектрах.

У аналізі зображень існує безліч алгоритмів та архітектур нейронних мереж, які можна використовувати для різних завдань, таких як класифікація, виявлення об'єктів та сегментація. При порівнянні різних алгоритмів важливо враховувати конкретні вимоги завдання та особливості даних. Ось порівняння для декількох важливих завдань:

Класифікація [14]:

- ResNet (Residual Networks): ResNet добре справляється з проблемою градієнта для дуже глибоких мереж, що дозволяє досягнути високої точності в класифікації. Вона особливо корисна, коли вимагається висока роздільна здатність.
- Inception (GoogLeNet): Inception відзначається низькою кількістю параметрів і високою продуктивністю. Вона добре підходить для завдань, де об'єкти на зображеннях можуть мати різні масштаби та розміри.

Виявлення об'єктів [14]:

- Faster R-CNN: Цей алгоритм відомий своєю швидкістю та точністю в виявленні об'єктів на зображеннях. Він використовує комбінацію CNN для визначення регіонів і RNN для точного визначення об'єктів.
- YOLO (You Only Look Once): YOLO є швидким та легким алгоритмом для виявлення об'єктів в реальному часі. Він робить прогнози на всьому зображенні в одному проході мережі та відзначається високою швидкістю роботи.
- Mask R-CNN: Ця модель, основана на Faster R-CNN, додатково включає сегментацію об'єктів, дозволяючи виділяти області, де знаходиться об'єкт, та сегментувати їх.

Кожен алгоритм та архітектура мають свої переваги та обмеження, і вибір залежить від конкретного завдання та вимог щодо продуктивності. Роблячи вибір, важливо збалансувати точність, швидкість та обсяги ресурсів.

Постановка задачі

Нейронні мережі є однією з ключових інновацій в галузі комп'ютерної науки, які дозволили симулювати роботу мозку та розв'язувати складні завдання. Вони складаються зі штучних нейронів, що функціонують аналогічно нейронам в біологічному мозку. Нейронні мережі можуть мати різну архітектуру, але зазвичай складаються з кількох шарів нейронів, які взаємодіють один з одним.

У контексті аналізу зображень, нейронні мережі стають надзвичайно потужним інструментом. Вони можуть навчатися розпізнавати патерни, особливості та абстракції візуальної інформації на зображеннях. Для цього вони використовують техніку навчання з учителем, де модель навчається на основі великої кількості вхідних прикладів і відповідних вихідних міток.

Процес аналізу зображень за допомогою нейронних мереж передбачає подачу вхідних зображень на вхід штучної нейронної мережі. Нейрони в мережі обробляють цю інформацію та намагаються визначити особливості та характеристики зображення, наприклад, форми об'єктів, кольори, текстури тощо. Нейрони в останньому шарі мережі надають оцінку або класифікацію зображення відповідно до задачі, яку мережа виконує [15].

У сучасному контексті розвитку інформаційних технологій та машинного навчання, завдання автоматичного визначення об'єктів на зображеннях є ключовим елементом, вимагаючим уваги та вдосконалення. Застосування нейронних мереж, зокрема глибоких конволюційних мереж (CNN), виявляється найбільш перспективним напрямком для вирішення даної задачі, однак існують істотні виклики, що обмежують їх ефективність та застосовність.

Однією з ключових проблем є недостатня точність в реальному часі та надійність результатів в умовах високого рівня складності зображень, де можливі фактори, такі як тіні, затемнення, рух об'єктів та перекриття, ускладнюють завдання визначення об'єктів. Методи, які вперто вдосконалюють або адаптуються до змінних умов, залишаються ненадійними в ряді сценаріїв.

Крім того, існуючі підходи часто виявляють обмежену здатність адаптації до об'єктів різних масштабів та геометричних конфігурацій. Це суттєво обмежує їх застосовність у вимогливих візуальних завданнях, де множина об'єктів варіюється за розміром та формою [16].

З іншого боку, суттєвим обмеженням є великі обчислювальні витрати, пов'язані з використанням глибоких нейронних мереж для визначення об'єктів на великій кількості зображень. Це не лише обмежує ефективність алгоритмів у реальному часі, але і вносить високі вимоги до обчислювальної інфраструктури, ускладнюючи їхнє впровадження у різноманітні практичні застосування [17].

Отже, загальною метою є розробка та оптимізація модифікованої нейронної мережі для виявлення об'єктів на зображеннях, які надходять від відеокамер спостереження. Оптимізація цього алгоритму має на меті досягнення кількох конкретних цілей. По-перше, покращення швидкодії алгоритму, зменшення часу обробки зображень та забезпечення його працездатності в режимі реального часу. По-друге, підвищення точності виявлення та класифікації об'єктів, що робить систему більш надійною та ефективною. По-третє, зменшення витрат обчислювальних ресурсів для забезпечення економічної ефективності та оптимального використання обладнання.

Серед ключових проблем, які дослідження має вирішити, варто визначити оптимальні параметри модифікованої нейронної мережі, забезпечення стійкості та адаптивності до різних умов освітлення та перешкод на зображеннях, а також оптимізацію алгоритмів обробки відеоданих для ефективної роботи з великим обсягом даних в реальному часі. Реалізація модифікованої нейронної мережі має виявити значущий внесок у розвиток систем відеоспостереження, підвищуючи їхню продуктивність, точність та реалізуючи вимоги до роботи в реальному часі.

Результати та обговорення

У світі нейронних мереж для аналізу зображень, несумнівно, відбувається бурхливий розвиток. Однак, щоб максимізувати їхній потенціал і вдосконалити результати аналізу зображень, важливо вдосконалювати та оптимізувати процеси. Розглянемо деякі з можливостей оптимізації та покращення результатів з використанням нейронних мереж [18]:

1. Попередня обробка зображень: Попередня обробка зображень є важливим кроком перед їхнім введенням у нейронну мережу. Застосування фільтрів, нормалізація, аугментація та інші техніки можуть покращити якість та інформаційність зображень, зменшити шум та підготувати дані для навчання та тестування.

Використання нейронних мереж для визначення об'єктів на зображенні

2. Підбір архітектури та гіперпараметрів: Вибір правильної архітектури нейронної мережі та гіперпараметрів є вирішальним завданням. Деякі завдання можуть вимагати глибоких мереж, тоді як інші – менше кількості параметрів для досягнення найкращих результатів.

3. Постійне навчання (fine-tuning) та перенос навчання (transfer learning): Використання заздалегідь навчених моделей та їх докладання для вирішення конкретних завдань може значно зекономити час і ресурси. Fine-tuning дозволяє адаптувати модель до нових завдань, використовуючи знання, набуті на попередніх наборах даних.

4. Використання ансамблів моделей: Комбінування декількох моделей, таких як декілька CNN, може покращити результати аналізу зображень. Ансамблі можуть підвищити точність та зменшити ризик перенавчання.

5. Покращення архітектури нейронної мережі: Нові інновації в архітектурах, такі як архітектура Transformer для обробки послідовностей, можуть бути адаптовані для аналізу зображень і призвести до покращення результатів.

6. Використання апаратного прискорення: Використання апаратного прискорення, такого як графічні процесори (GPU) та тензорні процесори (TPU), допомагає покращити продуктивність нейронних мереж та дозволяє обробляти великі обсяги даних швидше [19].

Оптимізація та покращення результатів аналізу зображень з використанням нейронних мереж є безперервним процесом, і вона відіграє важливу роль у покращенні продуктивності та точності цих систем. Наявність нових технологій та підходів в цій галузі надає великі можливості для подальших досліджень та розвитку [20].

Розглянемо більш детальний приклад оптимізації процесу виявлення об'єктів на зображенні, побудований на основі вище згаданих можливостей оптимізації та покращення результатів з використанням нейронних мереж. В цьому підході було використано існуючу модель YOLOv5 та введено певні модифікації в методи цієї моделі. Схема роботи такої системи наведена на рис. 2.



Рис. 2. Функціональна схема інформаційної системи для виявлення об'єктів на зображенні за допомогою модифікованих існуючих моделей.

Тут використовується оригінальна архітектура CSPDarknet53 та мережа агрегації шляхів (PANet), як основа й допоміжна частина нашої моделі [21].

На виході додається ще одна деталь для детекції дрібних об'єктів. Щоб додатково покращити ефективність нашої мережі, використовується кілька додаткових етапів. Зокрема, здійснюється популяція даних під час тренування, що сприяє адаптації до різких змін розміру об'єктів на зображеннях. Також додається багатомасштабне тестування та стратегії ансамблю мультимоделей під час виводу для отримання більш переконливих результатів детекції.

Модель реалізується на PyTorch 1.9.0 [22]. На фазі тренування використовується частина попередньо навченої моделі з YOLOv5, оскільки наша модифікація моделі використовується більша частина основи та деяка частина з оригінальної моделі YOLOv5. Використовуючи ці ваги, ми можемо значно скоротити час тренування. Використовується навчальний набір даних VisDrone2021 протягом 70 епох, а перші 3 епохи використовуються для початкового навчання. Також тут присутній оптимізатор Adam для тренування та $3e-5$ як початкова швидкість навчання. Швидкість навчання останньої епохи знижується до 0,11 від початкової швидкості навчання.

Використання нейронних мереж для визначення об'єктів на зображенні

Отже, впроваджені зміни, такі як CSPDarknet53 та PANet архітектури, призвели до значного поліпшення точності та стабільності детекції. Додаткові кроки, такі як популяція даних, багатомасштабне тестування та стратегії ансамблю мультимodelей, сприяли підвищенню ефективності. Реалізація моделі на PyTorch 1.9.0 та її тренування на безкоштовному GPU від Google Collab підтверджують можливість практичного використання. Експерименти та порівняння з іншими моделями свідчать про високий рівень точності та продуктивності запропонованого підходу. Отримані результати надають обґрунтовані підстави для подальших досліджень та розробок в області комп'ютерного зору.

Висновки

У результаті виконаної роботи вдалося досягти покращень в детекції та розпізнаванні об'єктів, що свідчить про успішність використаних методів оптимізації. У ході виконання дослідження була розглянута та реалізована модифікація процесу виявлення об'єктів на зображеннях на основі вдосконалених методів оптимізації та використання нейронних мереж. В рамках даного підходу використовувалася вже існуюча модель YOLOv5, яку було покращено та адаптовано для досягнення більш високих результатів.

Було введено оригінальну архітектуру CSPDarknet53 та мережу агрегації шляхів (PANet) як основу та допоміжну частину модифікованої моделі. Додатково була внесена деталь для ефективної детекції дрібних об'єктів. Для підвищення ефективності було використано популяцію даних під час тренування, що дозволило адаптуватися до різких змін розміру об'єктів на зображеннях. Також були застосовані багатомасштабне тестування та стратегії ансамблю мультимodelей для отримання більш переконливих результатів детекції.

Реалізація моделі використовувала PyTorch 1.9.0, частину попередньо навченої моделі YOLOv5 для скорочення часу тренування та навчальний набір даних VisDrone2021. Оптимізація ваг та використання оптимізатора Adam сприяли підвищенню швидкості навчання та ефективності моделі.

Розглянутий підхід може слугувати основою для подальших досліджень та розвитку в галузі виявлення об'єктів на зображеннях з використанням нейронних мереж.

Список використаних джерел

- [1] Farmaha I., Salo Y. Medical object detection using computer vision tools and methods only // САІР у проектуванні машин. Питання впровадження та навчання : матеріали XXX Міжнародної польсько- української науково-технічної конференції (Львів, Україна, 1–2 грудня 2022 р.). – 2022. – С. 18.
- [2] Y.-L. Tian, L. Brown, A. Hampapur, M. Lu, A. Senior and C.-F. Shu, "IBM smart surveillance system (S3): Event based video surveillance system with an open and extensible framework", *Mach. Vis. Appl.*, vol. 19, no. 5, pp. 315-327, Oct. 2008. <https://doi.org/10.1007/s00138-008-0153-z>
- [3] J. Fernández, L. Calavia, C. Baladrón, J. Aguiar, B. Carro, A. Sánchez-Esguevillas, et al., "An intelligent surveillance platform for large metropolitan areas with dense sensor deployment", *Sensors*, vol. 13, no. 6, pp. 7414-7442, Jun. 2013. <https://doi.org/10.3390/s130607414>
- [4] R. Baran, T. Rusc and P. Fornalski, "A smart camera for the surveillance of vehicles in intelligent transportation systems", *Multimedia Tools Appl.*, vol. 75, no. 17, pp. 10471-10493, Sep. 2016. <https://doi.org/10.1007/s11042-015-3151-y>
- [5] D. Eigenraam and L. J. M. Rothkrantz, "A smart surveillance system of distributed smart multi cameras modelled as agents", *Proc. Smart Cities Symp. Prague (SCSP)*, pp. 1-6, May 2016. <https://doi.org/10.1109/SCSP.2016.7501018>
- [6] Bosch Intelligent Video Analysis, May 2023, [Електронний ресурс] // Режим доступу: <https://www.boschsecurity.com/xc/en/>.
- [7] Bhubaneswar's Smart Safety City Surveillance Project Powered By Honeywell Technologies, May 2023, [Електронний ресурс] // Режим доступу: https://buildings.honeywell.com/content/dam/hbtbt/en/documents/downloads/Bhubaneswar-CS_0420_V2.pdf.
- [8] Hitachi: Data Integration Helps Smart Cities Fight Crime Iot-hitachi-smart Communities-solution, May 2023, [online] Available: <https://www.intel.com/content/dam/www/public/emea/xc/en/documents/>.
- [9] Iomniscient, May 2023, [Електронний ресурс] // Режим доступу: <https://iomni.ai/oursolutions/>.
- [10] E. B. Varghese and S. M. Thampi, "A cognitive IoT smart surveillance framework for crowd behavior analysis", *Proc. Int. Conf. Commun. Syst. Netw. (COMSNETS)*, pp. 360-362, Jan. 2021. <https://doi.org/10.1109/COMSNETS51098.2021.9352910>

- [11] V. Sharma, M. Gupta, A. Kumar and D. Mishra, "Video processing using deep learning techniques: A systematic literature review", *IEEE Access*, vol. 9, pp. 139489-139507, 2021. <https://doi.org/10.1109/ACCESS.2021.3118541>
- [12] *New trends in production engineering : колективна монографія.* – Warszawa, Poland: Sciendo, 2019. Farmaha I. Wound image segmentation using clustering based algorithms / I. Farmaha, M. Banaś, V. Savchyn, B. Lukashchuk, T. Farmaha. – с.217–225.
- [13]. Jaworski Nazariy, Farmaha Ihor, Farmaha Taras, Savchyn VasyI, Marikutsa Uliana. Implementation features of wounds visual comparison subsystem // Перспективні технології і методи проектування MEMC : матеріали XIV Міжнародної науково-технічної конференції, 18–22 квітня, 2018 р., Поляна, Україна. – 2018. – P. 114–117. (Google Scholar, SciVerse SCOPUS, Web of Science). <https://doi.org/10.1109/MEMSTECH.2018.8365714>
- [14] C. Dhiman and D. K. Vishwakarma, "A review of state-of-the-art techniques for abnormal human activity recognition", *Eng. Appl. Artif. Intell.*, vol. 77, pp. 21-45, Jan. 2019. <https://doi.org/10.1016/j.engappai.2018.08.014>
- [15] Yang, R., Yu, J., Yin, J., Liu, K., & Xu, S. (2022). A dense r-CNN multi-target instance segmentation model and its application in medical image processing. *IET image processing*(9), 16.
- [16] Szajna, A., Kostrzewski, M., Ciebiera, K., Stryjski, R., & Sciubba, E. (2021). Application of the deep cnn-based method in industrial system for wire marking identification. *Energies*(12). <https://doi.org/10.3390/en14123659>
- [17] Took, C. C., & Mandic, D. (2022). Weight sharing for lms algorithms: convolutional neural networks inspired multichannel adaptive filtering. *Digital Signal Processing*.
- [18] Weiller, C., Reiser, M., Glauche, V., Musso, M., & Rijntjes, M. (2022). The dual-loop model for combining external and internal worlds in our brain. *NeuroImage*, 263, 119583. <https://doi.org/10.1016/j.neuroimage.2022.119583>
- [19] Tremeau A., Borel N. A. region growing and merging algorithm to color segmentation [J]. *Pattern Recognition*, 1997, 30(7):1191-1203.R. [https://doi.org/10.1016/S0031-3203\(96\)00147-1](https://doi.org/10.1016/S0031-3203(96)00147-1)
- [20] Levinshtein A., Stere A., Kutulakos K. N., et al. TurboPixels: Fast superpixels using geometric ows [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2009,31(12):2290-2297.D. <https://doi.org/10.1109/TPAMI.2009.96>
- [21] Bazgir O, Zhang R, Dhruva S R, et al. Representation of features as image with neighborhood dependencies for compatibility with convolutional neural networks [J]. *Nature communications*, 2020, 11(1): 4391. <https://doi.org/10.1038/s41467-020-18197-y>
- [22] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2018,40(4):834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>

Oleh Zhrebukh¹, Ihor Farmaha²

¹Computer Aided Design Departmen, Lviv Polytechnic National University, Stepana Bandery st., 12
Lviv, Ukraine, E-mail: oleh.zhrebukh.mnknm.2022@lpnu.ua

²Computer Aided Design Departmen, Lviv Polytechnic National University, Stepana Bandery st., 12
Lviv, Ukraine, E-mail: ihor.v.farmaha@lpnu.ua

USING NEURAL NETWORKS TO IDENTIFY OBJECTS IN AN IMAGE

Recieved: March 12, 2024 / Revised: March 28, 2024 / Accepted: April 01, 2024

© Zhrebukh O., Farmaha I., 2024

Abstract. A modified neural network model based on Yolo V5 was developed and the quality metrics of object classification on video images built on the basis of existing known basic neural network architectures were compared. The application of convolutional neural networks for processing images from video surveillance cameras is considered in order to develop an optimized algorithm for detecting and classifying objects on video images. The existing models and architectures of neural networks for image analysis were analyzed and compared. The possibilities of optimizing the process of image analysis using neural networks are considered.

Keywords: convolutional neural networks, CNN, object detection, speed of video image processing, image properties and features.