

С. А. Цибуля, В. Ю. Тимчук, Н. С. Цибуля

МОДЕЛЬ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ВИЯВЛЕННЯ ЗАМАСКОВАНИХ ОБ'ЄКТІВ

У статті розглянуто застосування нейронної мережі для виявлення замаскованих об'єктів, прихованих природними масками та засобами маскування. Такі об'єкти зазвичай різноманітні за розмірами, нечіткі та візуально зливаються з місцевістю, що ускладнює їх виявлення системами комп'ютерного зору.

Автори проаналізували алгоритми машинного навчання, які використовуються для сегментації та виявлення об'єктів на зображеннях. Враховуючи отримані результати, для вирішення завдання виявлення замаскованих об'єктів у роботі пропонуємо модель нейронної мережі з архітектурою енкодер-декодер. Її особливостями є: застосування на вході додаткового прошарку, на який подається зображення, оброблене фільтром Собеля, що дозволяє підсилити виявлення країв та текстури об'єктів; використання алгоритму розтягування згортки в блоках енкодера паралельно основній частині визначення ключових ознак зумовлює зниження залежності виявлення від розмірів об'єктів; застосування в блоках декодера механізму зосередження уваги на важливих ділянках зображення підвищує ймовірність правильної класифікації цих ділянок у разі невпевненості моделі щодо їх належності. Проведення експериментів шляхом моделювання з різними гіперпараметрами нейронної мережі дозволило визначити, що бінарна перехресна ентропія найбільше підходить як функція втрат, для розв'язання задачі виявлення об'єктів із сильними фоновими шумами, а вибір для функції активації *Parametric Rectified Linear Unit* дозволяє підвищити точність сегментації об'єктів. Також розглянуто використання різноманітних метрик щодо оцінювання ефективності створеної моделі.

Тестування на наборах даних із реальними замаскованими об'єктами дозволило виявити проблемні питання, що впливають на процес сегментації в цілому та на точність виявлення замаскованих об'єктів зокрема, розв'язання яких дозволяє покращити ефективність роботи нейронних мереж. Результати досліджень пропонуємо використовувати під час створення засобів маскування для визначення їх ефективності на різних природних територіях, а також для пошуку замаскованих об'єктів противника в ході обробки розвідувальних даних.

Ключові слова: комп'ютерний зір; машинне навчання; згорткова нейронна мережа; зображення місцевості; набір даних; маскувальний малюнок; засіб маскування; замаскований об'єкт; сегментація; функція втрат; метрика.

Постановка проблеми в загальному вигляді. Оцінювання ефективності засобів маскування традиційно проводиться шляхом їх випробування в польових умовах, за результатами яких визначають відповідність маскувального малюнка (колірної палітри та текстури) засобу місцевості й оцінюють ефективність приховування військового об'єкта, © С. А. Цибуля, В. Ю. Тимчук, Н. С. Цибуля, 2024

замаскованого цим засобом. Для автоматизації зазначеного процесу та уникання помилок, виключення впливу суб'єктивності в роботі експертів-спостерігачів доцільно використовувати системи комп'ютерного зору. Одним із найскладніших завдань, яке стоїть перед такими системами обробки візуальних даних, є проведення групування та позначення подібних ділянок (сегментів) у зображеннях на піксельному рівні. Цей процес відомий як сегментація, його метою є визначення належності кожної точки зображення (пікселя) певному класу об'єктів. Його результатом є множина сегментів, які разом покривають усе зображення, або множина контурів, виділених на ньому. Усі пікселі в сегменті поєднані за деякою характеристикою чи за визначеною властивістю, наприклад, колір, яскравість або текстура. У ході аналізу зображень на наявність замаскованих об'єктів виконується класифікація пікселів за належністю до класу замаскованого об'єкта або фонові поверхні, на якій його розміщено.

Складність процесу сегментації залежить від характеристик вхідних зображень та їх внутрішнього контексту. Методи сегментації можна поділити на дві групи: на основі виявлення подібності та на основі меж (порогів) [1]. Перші розділяють зображення на сегменти, що мають схожі набори пікселів. Методи другої групи виділяють межі сегментів на основі певної відмінності в характеристиках ділянок зображення. Обидві групи мають складності в роботі із зображеннями, на яких слабо виражені краї ділянок, що можуть бути дуже подібні між собою або на них присутній високий рівень шуму.

Протягом останніх десятиліть було розроблено багато різних алгоритмів сегментації зображень. Проте їх застосування для виявлення замаскованих об'єктів викликає певні труднощі внаслідок варіативності маскувальних малюнків та не відповідає вимогам військової галузі щодо надійності й точності. Тому завдання розроблення алгоритмів пошуку замаскованих об'єктів є актуальною науково-прикладною проблемою.

Аналіз останніх досліджень і публікацій. Із розвитком згорткових нейронних мереж відкрилися нові перспективи розв'язання задачі комп'ютерного зору щодо сегментації в реальному часі та з достатньою точністю. Першою успішною моделлю була FCN-8 (англ. fully convolutional network), яка стала основою для подальших розробок нейронних мереж для сегментації зображень [2]. Базуючись на цій роботі, більшість мереж використовують архітектуру енкодер-декодер. На даний час популярними моделями нейронних мереж є: U-Net, SegNet, Mask R-CNN, YOLACT/YOLACT, YOLO (англ. you only look once) тощо. Кожна зі згаданих моделей має свої унікальні характеристики та оптимізована під різні види сегментації: семантичну (англ. semantic segmentation), об'єктну (англ. instance segmentation) та паноптичну (англ. panoptic segmentation).

Семантична сегментація – процес призначення всіх пікселів зображення до об'єктів певного класу. Її недоліком є неможливість визначати окремо об'єкти одного класу. Відомі такі моделі нейронних мереж для семантичної сегментації: SegNet, U-Net, DeconvNet і FCN тощо.

Об'єктна сегментація стосується задач комп'ютерного зору, пов'язаних із кількісним аналізом об'єктів. Надає можливість виявити кожен об'єкт або екземпляр класу, присутній на зображенні, і призначити йому власний сегмент або виділити рамкою

з унікальним ідентифікатором. Найбільш відомі такі моделі нейромереж: Mask R-CNN, Faster R-CNN, PANet і YOLACT.

Паноптична сегментація – поєднання підходів семантичної та об'єктної сегментації, коли кожному пікселю зображення присвоюється семантична мітка та унікальний ідентифікатор об'єкта. Більшість моделей паноптичної сегментації базуються на моделі Mask R-CNN: UPSNet, FPSNet, EPSNet і VPSNet.

В останні роки у зв'язку з активним розвитком машинного навчання питанню виявлення об'єктів в інтересах збройних сил різних країн за допомогою комп'ютерних систем приділяється все більше уваги. У статті [3] проаналізовано шляхи розв'язання проблеми сегментації високотекстурованих об'єктів, які мають великий діапазон зміни можливих значень кольору, в інтересах оброблення та автоматизації процесу дешифрування даних аерокосмічного знімання. З'ясовано, що найкращий результат можливо отримати в разі використання методів опису й виміру текстури, що ґрунтуються на обчисленні кількості перепадів яскравості на одиницю площі зображення та опису текстури на основі визначення її фрактальної розмірності.

Дослідження [4] присвячено виявленню людей, одягнених у військову форму з маскувальним малюнком (камуфляж), які знаходяться в природному середовищі. Для його проведення було створено спеціальний набір даних (англ. dataset) із зображеннями 20 зразків найбільш популярної військової форми, наприклад: British Multi-Terrain Pattern, Desert Digital MARPAT тощо. Для виявлення таких осіб автори розробили нейронну мережу на основі згорткової нейронної мережі CNN (англ. convolutional neural network). Для покращення результатів пошуку були застосовані алгоритми створення суперпікселів та обмеження просторової гладкості. Результати експериментів показали перевагу запропонованого підходу над класичними CNN-базованими методами виявлення. Продовження досліджень було висвітлено в роботі [5], основними їх особливостями є: розширення кількості зразків форми до 26; нова модель нейронної мережі; застосування для трансферного навчання моделі VGG-Net; використання розширених згорткових шарів для збільшення рецептивного поля нейронів без збільшення кількості ваг.

У [6] автори розглядають процес виявлення військовослужбовців за допомогою безпілотного літального апарата. Була створена нейронна мережа, особливістю якої є поєднання навчання одночасно на промаркованих (анотованих) і немаркованих даних, а також застосування модуля уваги до країв об'єктів, який становить собою додаткові прошарки, що інтегруються в основну архітектуру мережі та відповідають за модифікації ваг нейронної мережі щодо підвищення чутливості до країв. Це дозволило покращити точність виявлення об'єктів.

Стаття [7] присвячена виявленню замаскованих об'єктів COD (англ. camouflaged object detection), що є складнішим завданням порівняно з виявленням інших об'єктів через високий ступінь їх злиття з фоном. Автори розробили нейронну мережу MAGNet (англ. MAGnifier network), особливістю якої є наявність двох паралельних модулів: ергодичного збільшення зображення та фокусування уваги. MAGNet була протестована на загальнодоступних COD-датасетах та на власному датасеті із замаскованими військовими об'єктами. Ця нейронна мережа показала кращі результати на восьми метриках оцінювання порівняно з іншими моделями виявлення.

Автори дослідження [8] пропонують використовувати для визначення замаскованих об'єктів згорткову нейронну мережу із сімейства мереж YOLO, особливістю роботи якої є розрахунок імовірності правильної класифікації для всіх виявлених об'єктів, на основі яких будується теплова карта ефективності маскувального малюнка на місцевості. З'ясовано, що ймовірність класифікації моделі з високим відсотком сходиться з оцінкою людини-спостерігача, а чим ефективніший маскувальний малюнок, тим нижчу ймовірність класифікації видає модель.

Удосконалення архітектури YOLO для покращення виявлення замаскованих об'єктів виконано в роботі [9]. Запропонований алгоритм MC-YOLOv5s використовує YOLO версії v5s як базову структуру. Основна ідея полягає в застосуванні двовимірного дискретного косинусного перетворення для аналізу спектральних компонентів зображення. У результаті внесених змін середня точність mAP (англ. mean average precision) алгоритму MC-YOLOv5s досягла 94%, що більше, ніж у базової моделі, на 3,7%.

Заслуговує на увагу також публікація, присвячена виявленню об'єктів флори й фауни, які мають властивості до маскування [10]. Для проведення досліджень було зібрано набір даних, що містить 10000 зображень з об'єктами в різних природних сценах. Ці зображення детально анотовані за 78 категоріями, об'єктами та інстанціями. Була розроблена нейронна мережа SINet (англ. search identification network), яку протестували на створеному датасеті та на інших наборах даних для виявлення замаскованих біологічних об'єктів: CAMO [11], CAMO++, COD10K, UW-RS, NC4K. Під час роботи SINet на першому пошуковому етапі виділяє невеликі просторові зсуви на зображенні, що дозволяє інтегрувати більш дискримінаційні подання ознак в локальному просторі. На другому етапі вона застосовує частинний декодер, який здійснює агрегацію ознак із різних рівнів для генерації кінцевої карти об'єкта, використовуючи механізм уваги для підсилення середньорівневих ознак і кращого виділення об'єктів. Такий підхід дозволяє SINet визначати наявність замаскованого об'єкта на зображенні, точно ідентифікувати його, навіть якщо він добре зливається з навколишнім середовищем.

Формулювання завдання дослідження. Метою статті є розроблення моделі нейронної мережі для виявлення замаскованих об'єктів, прихованих природними масками та засобами маскування.

Виклад основного матеріалу. Особливістю маскувального малюнка засобу маскування є його здатність візуально збігатися та зливатися з місцевістю, у результаті чого він стає нечітким на вигляд. Окрім високої подібності між замаскованими об'єктами та фоном, ці об'єкти зазвичай різноманітні за масштабом (розмірами). На складність їх виявлення впливає також відмінність між процесом детекції особового складу, озброєння та військової техніки, унаслідок різних робочих відстаней їх виявлення. Тому для чіткого розуміння, де знаходяться і що це за об'єкти, доцільно виконувати сегментацію зображень. Виявлення замаскованих об'єктів є складним завданням унаслідок:

високої візуальної схожості засобів маскування з природним фоном, різниця полягає лише у відтінках;

впливу часу доби та погодних умов у процесі знімання на якість цифрових зображень. Недостатня освітленість, наявність крапель роси, рух рослин від вітру можуть призвести

до неякісного зображення;

впливу технічних характеристик апаратури для знімання та форматів зберігання знімків на зміну палітри кольорів у разі стискання вхідного сигналу;

довільних початкових умов формування цифрових зображень об'єктів на місцевості, обумовлених різними ракурсами фотографування, висотами літального апарата під час аерофотографування, положення й місцеперебування об'єктів.

На сам процес сегментації зображень впливають такі фактори:

складність фону на зображенні, де можуть бути ділянки, схожі із цільовими об'єктами за формою, кольором або текстурою;

наявність певної кількості об'єктів, подібних до цільового. Модель може згенерувати кілька сегментів, що їм відповідають. Це особливо актуально для завдань, де об'єкти часто зустрічаються в групах чи скупченнях;

перекриття об'єктів впливає на роботу моделі, яка може інтерпретувати частини одного об'єкта як кілька окремих, особливо якщо перекриття значно змінює його видиму форму;

змінність форми об'єктів, передусім тих, які можуть сильно трансформуватися за формою або розміром. Можлива неправильна ідентифікація частини об'єкта або позначення окремих його ділянок як поодиноких цільових об'єктів;

шум або артефакти на зображенні, спричинені процесом знімання або стискання, можуть призвести до помилкового виявлення ділянок, схожих на об'єкти.

Нейронні мережі мають потенціал до розв'язання більшості зазначених проблем, оскільки здатні навчатися та покращувати якісь свої роботи зі збільшенням кількості наявних прикладів та внесенням змін у їх архітектуру. За критерії для вибору архітектури нейронної мережі може бути використано: забезпечення необхідної точності розпізнавання; мінімально необхідний розмір навчального датасету; мінімальна кількість епох для навчання; достатня швидкість розпізнавання в режимі практичної роботи.

Для визначення замаскованих об'єктів пропонуємо модель нейронної мережі з архітектурою (рис. 1), вхідними даними для якої слугують зображення замаскованого об'єкта.

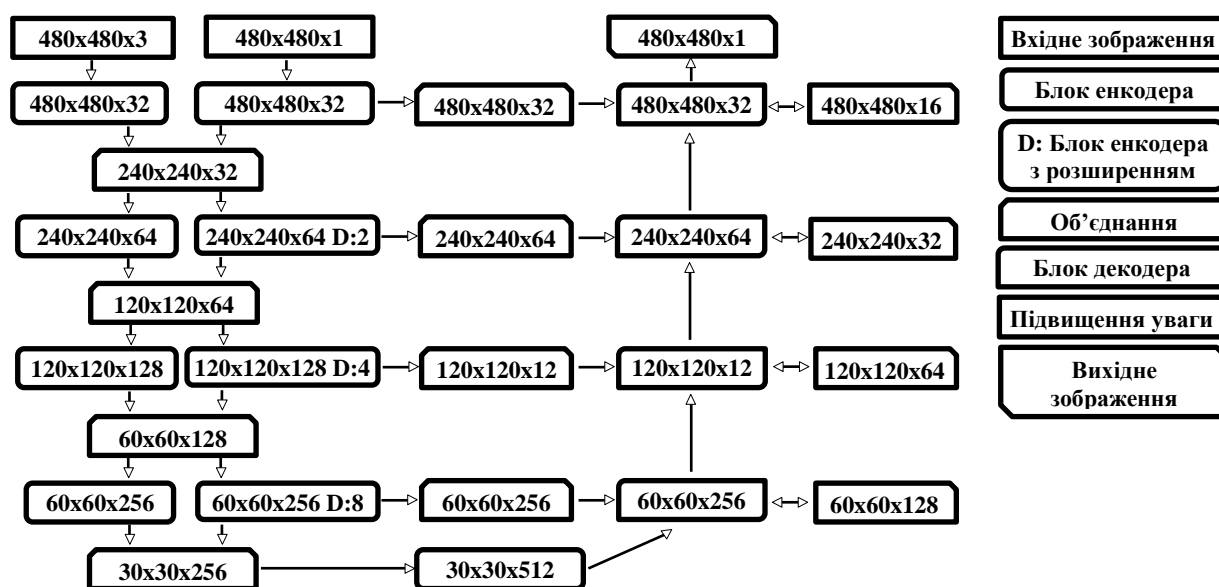


Рис. 1. Архітектура моделі нейронної мережі для визначення замаскованих об'єктів

Для покращення якості роботи моделі застосовується попередня підготовка вхідних даних: вирізання з кожного основного зображення робочої частини розміром 480 x 480 пікселів із розміщенням замаскованого об'єкта посередині та створення додаткового зображення, яке попередньо оброблене фільтром Собеля (рис. 2).

Особливість архітектури розробленої нейронної мережі полягає у:
застосуванні додаткового прошарку для попереднього обробленого фільтром Собеля вхідного зображення;

використанні симетричних пропусків з'єднань (англ. skip connector);

застосуванні розтягування згортки (англ. dilation);

використанні механізму уваги (англ. attention gate);

використання функції активації PReLU (англ. parametric rectified linear unit).

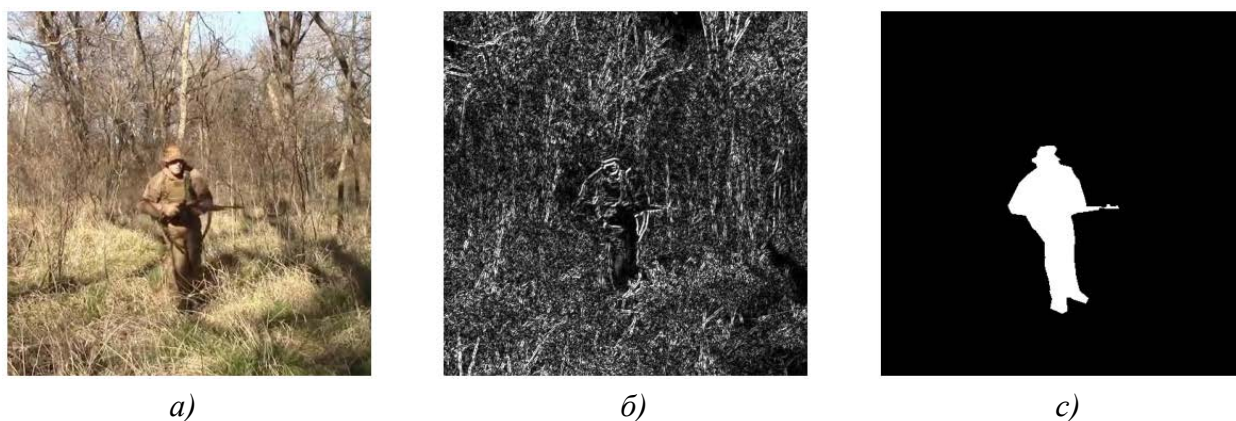


Рис. 2. Зображення, що подаються на вхід нейронної мережі для навчання: а) початкове зображення [4]; б) оброблене фільтром Собеля; в) маска

Мережа має симетричну структуру та складається з двох частин: енкодера й декодера. У першій її частині вхідне зображення проходить через кілька основних блоків, кожен з яких містить згорткові шари з фільтрами розміром 3 x 3 із застосуванням доповнення (англ. padding), що дозволяє зберігати розміри вхідних даних. Усі згорткові шари супроводжуються функцією активації для введення нелінійності. Після згорток виконується максимальна дискретизація (англ. max pooling), що знижує розмірність даних та збільшує поле зору нейронів. Паралельно до основних блоків енкодера застосовуються блоки з розтягуванням згортки, що забезпечує виявлення об'єктів меншого масштабу, ніж ті, що використовувалися для навчання мережі. Об'єднання результатів роботи основних і додаткових блоків допомагає комбінувати ознаки з різними рівнями контекстуальної інформації, що покращує здатність мережі розрізняти складні структури на зображеннях. Кількість каналів ознак (англ. feature channels) в енкодері подвоюється на кожному рівні зниження дискретизації.

У другій частині мережі застосовуються операції транспонованої згортки (англ. upsampling), що дозволяє збільшувати розміри карт ознак до первісного розміру вхідного зображення. Використання пропусків з'єднань, які передають карти ознак із відповідних шарів енкодера до декодера, допомагає зберегти контекстну інформацію, яка

могла бути втрачена під час стиснення. Це покращує загальне навчання моделі та точність реконструкції сегментованих ділянок.

У моделі реалізовано м'який алгоритм (англ. soft attention) механізму уваги, який допомагає мережі зосередитися на більш важливих ділянках зображення, підвищуючи тим самим якість відтворення та деталізації вихідного зображення. Моделювання уваги здійснюється за допомогою неперервно змінюваних ваг, оптимізація яких відбувається в ході зворотного поширення помилок у процесі навчання мережі.

Ці модифікації роблять модель більш гнучкою та потенційно здатною краще справлятися із задачами сегментації об'єктів зі складними текстурами.

У ході пошуку оптимальної характеристики моделі необхідно проаналізувати вплив кожного гіперпараметра на досягнуту якість сегментації. Процес навчання моделі залежить від початкової ініціалізації її параметрів, зазвичай він не є відтворюваним, оскільки ініціалізація параметрів – це стохастичний процес, однак є підходи щодо використання обмеженої множини значень, з якої розпочинається зчитування випадкової послідовності чисел для ініціалізації ядер. У роботі використовується метод HeNormal, який є вибіркою зі скороченого нормального розподілу, що дозволяє зробити процес навчання дещо відтворюваним.

У ході тренування моделей нейронних мереж для завдань сегментації функція втрат (помилки, англ. loss function) відіграє ключову роль, оскільки вона вимірює розбіжності між передбаченнями моделі та істинними значеннями. Ця функція допомагає керувати процесом навчання моделі, оптимізуючи її здатність точно розрізняти різні семантичні класи на зображеннях. Вона безпосередньо впливає на процес навчання через оптимізатор, який використовує значення втрат для оновлення ваг моделі з метою мінімізації цих втрат. З урахуванням того, що в роботі для виявлення замаскованих об'єктів взято лише один клас, до якого належать усі маскувальні малюнки засобів маскування, то анотацією у датасетів є бінарне зображення – чорно-біла маска (рис. 2). Для сегментації таких зображень, де кожен піксель має бути класифікований як належний до об'єкта (білий) або до фону (чорний), є декілька функцій втрат [12]: бінарна перехресна ентропія (англ. binary cross-entropy (BCE)), binary focal loss with logits, focal loss, dice loss, tversky loss. Функція втрат BCE вимірює розбіжність між істинними мітками та передбаченими ймовірностями кожного класу. Вона поєднує принципи перехресної ентропії з теорії інформації, яка оцінює кількість інформації, необхідної для вираження одного розподілу через інший, та бінарної класифікації, що робить її ефективною для оцінювання моделей, які передбачають належність до одного з двох можливих класів. Решта функцій є розширенням та модифікаціями BCE, вони призначені для розв'язання проблеми дисбалансу класів. Результати тестування різних функцій втрат наведено в табл. 1.

Таблиця 1

Значення метрики середнього IoU залежно від функції втрат

Функція втрат	Binary cross-entropy	Tversky (0,3, 0,7)	Tversky (0,7, 0,3)	Binary focal loss with logits	Dice	Binary cross-entropy + Dice
Середнє IoU	0,914	0,908	0,912	0,907	0,894	0,877

З табл. 1 видно, що найкраще показала себе функція втрат BCE. Для зображення розміром $W \times H$ загальна помилка, що обчислюється функцією BCE, буде дорівнювати сумі втрат на всіх пікселях зображення:

$$L = -\frac{1}{N} \sum_{i=1}^N L_i = -\frac{1}{WH} \sum_{i=1}^{WH} [y_i \times \ln p_i + (1 - y_i) \ln (1 - p_i)], \quad (1)$$

де N – кількість пікселів у зображенні, шт.;

y_i – істинна мітка для i -го елемента (0 – фон або 1 – об'єкт);

p_i – передбачена моделлю ймовірність, що i -й елемент належить до класу об'єкта.

$$N = W \times H, \quad (2)$$

де W, H – висота та ширина зображення відповідно, px.

Якщо передбачена ймовірність (p_i) дуже близька до істинної мітки (y_i), то компонент втрати для цього пікселя має мале значення. Коли ж передбачена ймовірність (p_i) далека від істинної мітки (y_i), тоді компонент втрати для цього елемента великий.

Виділяють такі способи оцінювання якості роботи алгоритмів семантичної сегментації: об'єктивний та суб'єктивний. Недоліком суб'єктивних способів оцінювання є наявність людського фактора в роботі експертів. Об'єктивні методи використовують кількісні показники та базуються на застосуванні метрик для порівняння вихідного результату алгоритму з еталоном. Тому для навчання моделей алгоритми метрик відіграють важливу роль в оцінюванні їх ефективності. Вони дозволяють аналізувати, наскільки добре модель відрізняє цільові об'єкти від фону. Метрики не впливають на процес оновлення ваг під час навчання моделі, а призначені для оцінювання її продуктивності, надають інтуїтивно зрозумілі показники ефективності. Для бінарної сегментації зображень є кілька ключових метрик: коефіцієнт Дайса (англ. Dice coefficient), точність (англ. precision), чутливість (англ. recall), F1-Score, специфічність (англ. specificity), AUC-ROC (англ. area under the receiver operating characteristic curve), MCC (англ. matthews correlation coefficient). Найбільш простим є індекс Жаккара (англ. Jaccard index), інша назва IoU (англ. intersection over union), який вимірює відношення перетину між передбаченою (A) та істинною маскою (B) і показує, наскільки вони сходяться. IoU визначають у такий спосіб:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \quad (3)$$

де A – передбачена маска;

B – істинна маска;

$(A \cap B)$ – кількість пікселів, які належать як передбаченій, так й істинній масці, тобто їх перетину (англ. intersection);

$(A \cup B)$ – загальна кількість пікселів, що включає всі пікселі, які належать хоча б до однієї з масок (передбаченої або істинної) або їх об'єднанню (англ. union).

Чим ближче значення IoU до одиниці, тим краща точність сегментації моделі.

Функції активації згорткових прошарків нейронної мережі є важливою складовою її архітектури. Значення метрики IoU залежно від функції активації наведено в табл. 2. За результатами моделювання в роботі було вибрано для використання функцію активації PReLU.

Таблиця 2

Значення метрик IoU залежно від функції активації

Функція активації	LeakyReLU	ReLU	PReLU
Середнє IoU	0,927	0,914	0,931

Одним зі способів удосконалення роботи нейронної моделі та підвищення швидкості її навчання є використання попередньо навчених нейронних мереж типу VGG, DenseNet, EfficientNet тощо (англ. transfer learning). За результатами моделювання визначено, що цей спосіб не покращує результати розпізнавання замаскованих об'єктів, унаслідок навчання інших нейромереж на типах даних, які за текстурями та зовнішнім виглядом значно відрізняються від маскувальних малюнків.

Оскільки в завданнях сегментації важливо точно знати межі об'єкта, то додаткове їх виокремлення перед початком навчання має покращити роботу мережі. Ці межі (краї) позначають місця розриву зображення в рівнях сірого, інших кольорів, текстур тощо. Відомі різні алгоритми детекторів країв, як-от: оператори Собеля, Кірша, Превітта тощо. Додавання в нейромережу додаткових шарів із зображеннями, обробленими такими алгоритмами, може бути доцільним, оскільки ці методи виявлення меж підкреслюють важливі особливості зображення, які допомагають нейронній мережі краще розуміти контури та текстурні особливості маскувального малюнка об'єкта.

Результатом роботи нейронної мережі є отримання на виході моделі ймовірності належності пікселя до об'єкта чи до фону. Важливим при цьому є визначення значення оптимального порога передбачення (англ. threshold), за яким вихід моделі розділяється на два класи: позитивний і негативний. Для цього застосовують такі алгоритми: ROC Curve і AUC-ROC, індекс Юдена J тощо.

Завершальним етапом проектування та навчання нейронної мережі є її перевірка щодо можливості визначення об'єктів на зображеннях, які не використовувалися в ході навчання та верифікації. За її результатами приймається рішення про доцільність використання отриманої моделі на практиці.

Висновки. У статті проаналізовано підходи до застосування нейронних мереж для сегментації зображень та перевірено можливість їх використання для визначення об'єктів, прихованих засобами маскування. Тестування на наборах даних із реальними замаскованими об'єктами дозволило визначити архітектуру моделі нейронної мережі. Її особливостями є: використання на вході додаткового прошарку, на який подається зображення, оброблене фільтром Собеля, а також алгоритму розтягування згортки в блоках енкодера паралельно основній частині визначення ключових ознак із вхідних зображень; застосування в блоках декодера механізму зосередження уваги на важливих ділянках зображення.

У ході проведених досліджень були виокремлені проблемні питання, які впливають на процес сегментації в загальному та на виявлення замаскованих об'єктів зокрема. Для їх

вирішення пропонуємо розглянути такі напрямки:

дослідження можливих шляхів удосконалення архітектури нейронної мережі, що вплине на якість сегментації;

визначення критеріїв та показників точності сегментації щодо належності отриманих сегментів зображення до замаскованих об'єктів;

збір й анотування розширених датасетів із замаскованим особовим складом, а також озброєнням та військовою технікою;

розширення наборів даних за допомогою генерування синтетичних даних;

оцінювання ефективності й доцільності проведення попередньої та завершальної обробки зображень.

У роботі проведено дослідження щодо визначення оптимальної архітектури нейронної мережі для виявлення замаскованих об'єктів на зображеннях, що має значну практичну цінність для використання у системах комп'ютерного зору військового призначення. Результати досліджень пропонуємо використовувати для створення маскувальних малюнків засобів маскування на етапі визначення їх ефективності застосування на різних місцевостях та в природних зонах, а також у ході обробки розвідувальних даних для пошуку прихованих об'єктів противника.

СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ

1. Kaur D., Kaur Y. Various Image Segmentation Techniques: A Review // IJCSMC. 2014. Vol. 3, № 5. P. 809–814.
2. Long J., Shelhamer E., Darrell T. Fully Convolutional Networks for Semantic Segmentation // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 3431–3440. <https://doi.org/10.48550/arXiv.1411.4038>
3. Березіна С. І., Гордієнко Ю. О., Солонець О. І. Аналіз шляхів вирішення проблеми сегментації високотекстурованих об'єктів // Проблеми створення, випробування, застосування та експлуатації складних інформаційних систем : зб. наук. праць. Житомир : ЖВІ, 2019. Вип. 17. С. 27–40. <https://doi.org/10.46972/2076-1546.2019.17.03>
4. Detection of People with Camouflage Pattern via Dense Deconvolution Network / Zheng Y., Zhang X., & Wang F. et al. // IEEE Signal Processing Letters. 2019. Vol. 26, № 1. P. 29–33. <https://doi.org/10.1109/LSP.2018.2825959>
5. Camouflage People Detection via Strong Semantic Dilation Network / Fang Z., Zhang X., & Deng X et al. // Proceedings of the ACM Turing Celebration Conference. China, ACM TURC '19. New York, USA : Association for Computing Machinery. 2019. Article 134. P. 1–7. <https://doi.org/10.1145/3321408.3326662>
6. Liu Y., Wang C.-q., Zhou Y.-j. Camouflaged People Detection Based on a Semi-Supervised Search Identification Network // Defence Technology. 2023. Vol. 21. P. 176–183. <https://doi.org/10.1016/j.dt.2021.09.004>
7. MAGNet: A Camouflaged Object Detection Network Simulating the Observation Effect of a Magnifier / Jiang X., Cai W., & Zhang Z. et al. // Entropy. 2022. Vol. 24, № 12. Article ID 1804. <https://doi.org/10.3390/e24121804>
8. Van der Burg E., Hogervorst M. A., Toet A. Measuring the Dynamics of Camouflage in Natural Scenes Using Convolutional Neural Networks // Target and Background Signatures VIII. 78

2022. Vol. 12270. Article ID 1227003. SPIE. <https://doi.org/10.1117/12.2636107>

9. Zhang W., Zhou Q., Li R., Niu F. Research on Camouflaged Human Target Detection Based on Deep Learning // *Computational Intelligence and Neuroscience*. 2022. Vol. 2022. Article ID 7703444. P. 1–12. <https://doi.org/10.1155/2022/7703444>

10. Fan D.-P., Ji G.-P., Cheng M.-M., Shao L. Concealed Object Detection // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2022. Vol. 44, № 10. P. 6024–6042. <https://doi.org/10.1109/TPAMI.2021.3085766>

11. Anabranh Network for Camouflaged Object Segmentation / Le T.-N., Nguyen T. V., & Nie Z. et al. // *Computer Vision and Image Understanding*. 2019. Vol. 184. P. 45–56. <https://doi.org/10.1016/j.cviu.2019.04.006>

12. Jadon S. A Survey of Loss Functions for Semantic Segmentation // *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. Via del Mar, Chile, 2020. P. 1–7. <https://doi.org/10.1109/CIBCB48159.2020.9277638>

Стаття надійшла до редакції 01.05.2024.

REFERENCES

1. Kaur, D., & Kaur, Y. (2014). Various Image Segmentation Techniques: A Review. *IJCSMC*, Vol. 3 (5), 809-814.

2. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431–3440). <https://doi.org/10.48550/arXiv.1411.4038>

3. Berezina, S. I., Hordiienko, Yu. O., & Solonets, O. I. (2019). Analiz shliakhiv vyrishennia problemy sehmentatsii vysokoteksturovanykh ob'ektiv [Analysis of Ways to Solve the Problem of Segmentation of Highly Textured Objects]. *Problemy stvorennia, vyprobuvannia, zastosuvannia ta ekspluatatsii skladnykh informatsiinykh system : zb. nauk. prats [Problems of Construction, Testing, Application and Operation of Complex Information Systems. Scientific journal of Korolov Zhytomyr Military Institute]*, Iss. 17, 27–40. <https://doi.org/10.46972/2076-1546.2019.17.03>

4. Zheng, Y., Zhang, X., Wang, F., Cao, T., Sun, M., & Wang, X. (2019). Detection of people with camouflage pattern via dense deconvolution network. *IEEE Signal Processing Letters*, 26 (1) (pp. 29-33). <https://doi.org/10.1109/LSP.2018.2825959>

5. Fang, Z., Zhang, X., Deng, X., Cao, T., & Zheng, C. (2019). Camouflage People Detection via Strong Semantic Dilation Network. In *Proceedings of the ACM Turing Celebration Conference (ACM TURC '19, China)*, Article 134 (pp. 1–7). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/3321408.3326662>

6. Liu, Y., Wang, C.-q., & Zhou, Y.-j. (2023). Camouflaged People Detection Based on a Semi-Supervised Search Identification Network. *Defence Technology*, Vol. 21, 176–183. <https://doi.org/10.1016/j.dt.2021.09.004>

7. Jiang, X., Cai, W., Zhang, Z., Jiang, B., Yang, Z., & Wang, X. (2022). MAGNet: A Camouflaged Object Detection Network Simulating the Observation Effect of a Magnifier. *Entropy*, 24 (12), Article ID 1804. <https://doi.org/10.3390/e24121804>

8. Van der Burg, E., Hogervorst, M. A., & Toet, A. (2022). Measuring the Dynamics of Camouflage in Natural Scenes Using Convolutional Neural Networks. *Target and Background Signatures VIII*, Vol. 12270, Article ID 1227003. SPIE. <https://doi.org/10.1117/12.2636107>

9. Zhang, W., Zhou, Q., Li, R., & Niu, F. (2022). Research on Camouflaged Human Target Detection Based on Deep Learning. *Computational Intelligence and Neuroscience*, Article ID 7703444, 1–12. <https://doi.org/10.1155/2022/7703444>
10. Fan, D.-P., Ji, G.-P., Cheng, M.-M., & Shao, L. (2022). Concealed Object Detection. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44 (10), 6024–6042. <https://doi.org/10.1109/TPAMI.2021.3085766>
11. Le, T.-N., Nguyen, T. V., Nie, Z., Tran, M.-T., & Sugimoto, A. (2019). Anabranh Network for Camouflaged Object Segmentation. *Computer Vision and Image Understanding*, Vol. 184, 45–56. <https://doi.org/10.1016/j.cviu.2019.04.006>
12. Jadon, S. (2020). A Survey of Loss Functions for Semantic Segmentation. In *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. (pp. 1–7). Via del Mar, Chile. <https://doi.org/10.1109/CIBCB48159.2020.9277638>

S. A. Tsybulia, V. Yu. Tymchuk, N. S. Tsybulia

A NEURAL NETWORK MODEL FOR DETECTING CAMOUFLAGED OBJECTS

This article discusses the use of computer vision to detect camouflaged objects that are hidden by natural masks and camouflage devices. Such objects are usually varied in size, fuzzy, and visually merge with the terrain, which makes them difficult to detect.

The authors analyze the models of machine learning algorithms used to segment and detect objects in images. Taking into account the analysis, to solve the problem of detecting camouflaged objects, the paper proposes a neural network model with an encoder-decoder architecture. Its features are: the use of an additional layer at the input, which is fed with an image processed by a Sobel filter, which allows to enhance the detection of object edges; the use of the convolutional stretching algorithm in the encoder blocks, in parallel with the main part of the key features determination, leads to a decrease in the dependence of detection on the size of objects; the use of a mechanism in the decoder blocks to focus on important parts of the image increases the probability of correct classification of image areas in cases of uncertainty of the model regarding their. Experiments by modeling, with different hyperparameters of the neural network, allowed us to determine that binary cross-entropy is most suitable as a loss function for solving the problem of detecting objects with strong background noise, and the choice of Parametric Rectified Linear Unit as an activation function allows to improve the quality of object segmentation. We also consider the use of various metrics to evaluate the effectiveness of the created model.

Testing on datasets with real cloaked objects allowed us to identify problematic issues affecting the segmentation process in general and the accuracy of detecting cloaked objects in particular, the solution of which can improve the efficiency of neural networks in object detection. The results of the research are proposed to be used in the creation of camouflage means to determine their effectiveness, as well as to search for camouflaged enemy objects in the course of intelligence processing.

Keywords: *computer vision; machine learning; convolutional neural network; terrain image; dataset; camouflage pattern; camouflage means; camouflaged object; masking; concealment; object detection; segmentation; loss function; metric.*