

Д. МОРОЗ, аспірант, НУ "Одеська політехніка", Одеса,
В. ЛЮБЧЕНКО, д-р техн. наук, проф., НУ "Одеська політехніка",
Одеса

ІНТЕГРАЦІЯ ДИФУЗІЙНИХ МОДЕЛЕЙ ТА ГЛИБОКОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ І ВІДСТЕЖЕННЯ ОБ'ЄКТІВ В ПІДВОДНИХ ЗОБРАЖЕННЯХ

Складні оптичні властивості води призводять до спотворення кольору, розмиття та низького контрасту, що ускладнює виявлення та відстеження об'єктів. У поєднанні з дефіцитом розмічених даних підводних зображень це серйозно обмежує продуктивність систем виявлення та відстеження об'єктів на основі глибокого навчання. Досліджується застосування дифузійних моделей для попередньої обробки підводних зображень виявлення та відстеження об'єктів за допомогою глибокого навчання. Табл.: 1. Бібліогр.: 22 назв.

Ключові слова: підводні зображення, відновлення зображень, моделі дифузії, глибоке навчання.

Постановка проблеми. Виявлення та відстеження підводних об'єктів мають вирішальне значення в системах різних сфер застосування: моніторинг морського життя, інспекція підводної інфраструктури, навігація автономних підводних транспортних засобів. Підводна інфраструктура, така як трубопроводи, мости та морські платформи, вимагає регулярного огляду, щоб переконатися в її структурній цілісності та запобігти збоям в роботі. Але в процесі розробки відповідних систем доводиться стикатися зі значними проблемами через характерні особливості підводного середовища і його впливом на формування зображень. Погана видимість, спричинена поглинанням і розсіюванням світла, у поєднанні з дефіцитом достатньої кількості розмічених даних підводних зображень серйозно обмежують продуктивність систем на етапі виявлення та відстеження об'єктів при застосуванні глибокого навчання. Складні оптичні властивості води призводять до спотворення кольору, розмиття та низького контрасту, що ускладнює точну ідентифікацію та відстеження об'єктів.

Методи попередньої обробки зображень – препроцесинг, відіграють вирішальну роль у зменшенні впливу цих проблем, покращують якість зображення та доповнюють навчальні набори даних. Традиційні методи часто не можуть ефективно усунути різноманітні складні спотворення підводних зображень. Дифузійні моделі, використовуючи ітераційні процеси усунення шумів, можуть створювати реалістичні підводні зображення, розширюючи набори даних для навчання та підвищуючи надійність моделей глибокого навчання. Крім того, їх можна використовувати для відновлення підводних спотворених, покращуючи їх чіткість і видимість.

Аналіз останніх досліджень. На відміну від звичайних зображень, підводні зображення часто страждають від поганої видимості, спотворення кольорів, низької контрастності та значного шуму [1 – 5]. Ці проблеми виникають через складну взаємодію світла з водою, а також через практичні труднощі, пов'язані з отриманням високоякісних даних у підводних умовах [6, 13]. Погіршення якості зображення перешкоджає виконанню багатьох завдань комп'ютерного зору, зокрема виявлення та відстеження об'єктів за допомогою моделей глибокого навчання. Отже, ефективні методи попередньої обробки мають вирішальне значення для пом'якшення цих проблем і забезпечення більш ефективного аналізу підводних зображень у різних сферах застосування, включаючи морську біологію, підводну робототехніку та археологію, військову справу. Подолання цих проблем може значно підвищити надійність і застосовність систем виявлення та відстеження підводних об'єктів.

Ослаблення та розсіювання світла є основними факторами, що сприяють погіршенню якості підводних зображень. Коли світло проходить через воду, воно одночасно поглинається і розсіюється. Поглинання призводить до зменшення інтенсивності світла, причому різні довжини хвиль поглинаються з різною швидкістю. Червоне світло, наприклад, поглинається швидше, ніж синє світло, що призводить до синьо-зеленого кольору на підводних зображеннях. З іншого боку, розсіювання відбувається, коли світло відхиляється частинками, зваженими у воді. Цей ефект розсіювання зменшує контрастність і різкість зображення, створюючи туманний вигляд. Комбінований ефект поглинання та розсіювання призводить до значного зниження видимості та спотворення

кольорів, що ускладнює точне сприйняття об'єктів у підводному середовищі.

Ефективне навчання моделей глибокого навчання вимагає великих наборів даних, отримання яких є дорогим і призводить до *дефіциту даних* високоякісних розмічених даних [14]. Для збору підводних даних потрібне спеціальне обладнання, таке як апарати з дистанційним керуванням або автономні підводні апарати, а також кваліфікований персонал для їх керування. Процес ручного анотування підводних зображень займає багато часу та може призвести до помилок через погану видимість і складні сцени. Дослідження з метою збільшення кількості даних і створення синтетичних даних часто не в змозі відтворити складність реального підводного середовища.

Шум і артефакти ще більше погіршують якість даних для навчання. Зважені частинки у воді, такі як осади і планктон, можуть створювати випадковий шум і затемнювати деталі на зображеннях. Перешкоди, пов'язані з обладнанням, наприклад електричний шум від підводних апаратів або датчиків, також можуть сприяти погіршенню зображення. Крім того, нерівномірне освітлення та тіні можуть створювати артефакти, які ускладнюють аналіз зображення (рис. 1).



Рис. 1. Кадри з підводного відео.

Ці шуми та артефакти можуть суттєво вплинути на продуктивність алгоритмів виявлення та відстеження об'єктів, що призведе до помилкових спрацьовувань і пропусках при виявленні. Методи препроцесингу, спрямовані на зменшення шуму та видалення артефактів, є важливими для підвищення точності та надійності систем аналізу підводних зображень [15]. Для вирішення цих проблем і підвищення якості підводних

зображень постійно розробляються передові алгоритми шумозаглушення та методи видалення артефактів.

Мета роботи. Аналіз останніх досягнень в застосуванні методів дифузії, їх переваги та обмеження в контексті підводних зображень та дослідження їх інтеграції з моделями глибокого навчання.

Об'єкт дослідження – задача препроцесорної обробки підводних зображень у контексті виявлення та відстеження об'єктів за допомогою глибокого навчання.

Предмет дослідження – інтеграція дифузійних моделей та глибокого навчання для підвищення точності і надійності систем виявлення і відстеження об'єктів у підводних зображеннях.

Моделі на основі дифузії. Дифузійні моделі стали потужними інструментами в обробці зображень, особливо при їх відновленні та покращенні. Вони значно просунули різні галузі застосування завдяки їх здатності вивчати складні розподіли даних та генерувати нові зображення високої якості [16].

Дифузійні моделі – це імовірнісні генеративні моделі, які вчаться поступово перетворювати розподіл даних (наприклад, зображення) у простий розподіл, наприклад, розподіл Гауса, за допомогою процесу прямої дифузії [14]. Модель включає два етапи: процес прямої дифузії та процес зворотної дифузії.

Прямий процес поступово додає шум до вхідних даних x_0 , протягом T кроків часу, перетворюючи його на чистий шум, x_T :

$$q(x_t | x_{t-1}) = N(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I),$$

де β_t^T – розклад дисперсій, а I – тотожна матриця.

Процес зворотної дифузії, навпаки, вчиться змінювати це перетворення, починаючи з розподілу шуму та поступово зменшуючи його, щоб реконструювати вихідну вибірку даних. Зворотній процес також є марківським і може бути зображений як:

$$p_\theta(x_{t-1} | x_t) = N(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)),$$

де $\mu_\theta(x_t, t)$ і $\Sigma_\theta(x_t, t)$ є вивченими параметрами, часто змодельованими нейронною мережею.

Такі ключові моделі дифузії, як Denoising Diffusion Probabilistic Models (DDPM) [15] і Denoising Diffusion Implicit Models (DDIM) [16, 17] є теоретичною основою для практичної реалізації дифузійних моделей. DDIM, наприклад, вводять немарківське формулювання для прискорення процесу вибірки. Дифузійні моделі добре справляються зі створенням високоякісних зображень шляхом вивчення основного розподілу даних і вибірки з нього. Процес зворотної дифузії, починаючи з випадкового шуму, поступово покращує зображення, додаючи деталі та структуру, доки не буде сформовано реалістичне зображення. Моделі Stable Diffusion, використовують моделі прихованої дифузії для виконання процесу дифузії в просторі нижчої розмірності, значно знижуючи обчислювальні витрати та створюючи зображення з високою роздільною здатністю. Моделі каскадної дифузії додатково покращують якість зображення, використовуючи кілька моделей дифузії [18]. Ці моделі спочатку генерують зображення з низькою роздільною здатністю, а потім поступово вдосконалюють його за допомогою наступних моделей дифузії для досягнення високої точності.

Методи ICDT (Iterative Color Correction and Denoising Transformer) і UW-DiffPhys демонструють потенціал дифузійних моделей для вирішення проблем покращення підводних зображень. ICDT використовує архітектуру на основі перетворення дифузії для ітеративного виправлення спотворень кольору та видалення шуму, що призводить до отримання чіткіших і яскравіших зображень. UW-DiffPhys поєднує в нейронній мережі фізичні методи обчислення світла UIE (Underwater Image Enhancement) з низькою обчислювальною складністю з моделями на основі дифузії U-Net. Це дозволяє усунути обмеження існуючих методів і зменшити обчислювальну складність структури UW-DDPM, зберігаючи високу продуктивність.

Дифузійні моделі пропонують переконливу альтернативу традиційним методам попередньої обробки підводних зображень, особливо в контексті виявлення та відстеження об'єктів на основі глибокого навчання [7, 17, 18, 19]. Хоча дифузійні моделі надають переваги для попередньої обробки підводних зображень, але для забезпечення їх ефективного застосування необхідно вирішити кілька проблем і обмежень:

обчислювальні вимоги, доступність даних, складність реального підводного середовища. Подолання цих перешкод має вирішальне значення для реалізації повного потенціалу попередньої обробки на основі дифузії. Різниця в продуктивності виявлення малих і великих об'єктів є серйозною проблемою, особливо коли набори даних для навчання та тестування суттєво відрізняються за масштабом об'єкта. Невеликі об'єкти часто не мають чітких візуальних ознак, через що їх важко відрізнити від фону.

Застосування дифузійних моделей до попередньої обробки підводних зображень з метою виявлення та відстеження об'єктів є новою областю зі значним потенціалом для майбутніх досліджень. Сучасні дослідження демонструють перспективність дифузійних моделей для покращення підводних зображень. Вони включають дослідження нових дифузійних архітектур [19], адаптованих до характеристик підводного зображення, таких як нерівномірне освітлення та спотворення кольору. Інтеграція мультимодальних даних, таких як акустичні або сонаРНі дані, з візуальним покращенням на основі дифузії також може забезпечити більш комплексний підхід до розуміння підводної сцени [20]. Нарешті, дослідження оцінки невизначеності в рамках дифузійної моделі має вирішальне значення для надання надійної інформації для алгоритмів виявлення та відстеження об'єктів.

Методика апробації моделей та проблеми реалізації. Основною перешкодою для розгортання дифузійних моделей для попередньої обробки підводних зображень є їх обчислювальна вартість. Додатки в реальному часі, особливо на підводних апаратах з обмеженими ресурсами, вимагають ефективної обробки. Дослідження апаратного прискорення за допомогою графічних або спеціалізованих процесорів може значно підвищити швидкість роботи дифузійних моделей для підводних зображень у реальному часі.

Дифузійні моделі можна ефективно інтегрувати з іншими методами обробки зображень для створення більш надійних і універсальних конвеєрів покращення підводного зображення. Крім того, інтеграція моделей дифузії з моделями виявлення об'єктів на основі глибокого навчання може приводити до систем, які можна навчати, як з метою оптимізації для роботи в реальному часі, так і для ефективного виявлення

об'єктів. Аналіз моделей виявлення об'єктів на основі глибокого навчання показав, що на основі критеріїв оцінки перевагу мають моделі сімейства YOLO (рис. 2). Тому дані моделі були вибрані для інтеграції їх з дифузійною моделлю DDPM.

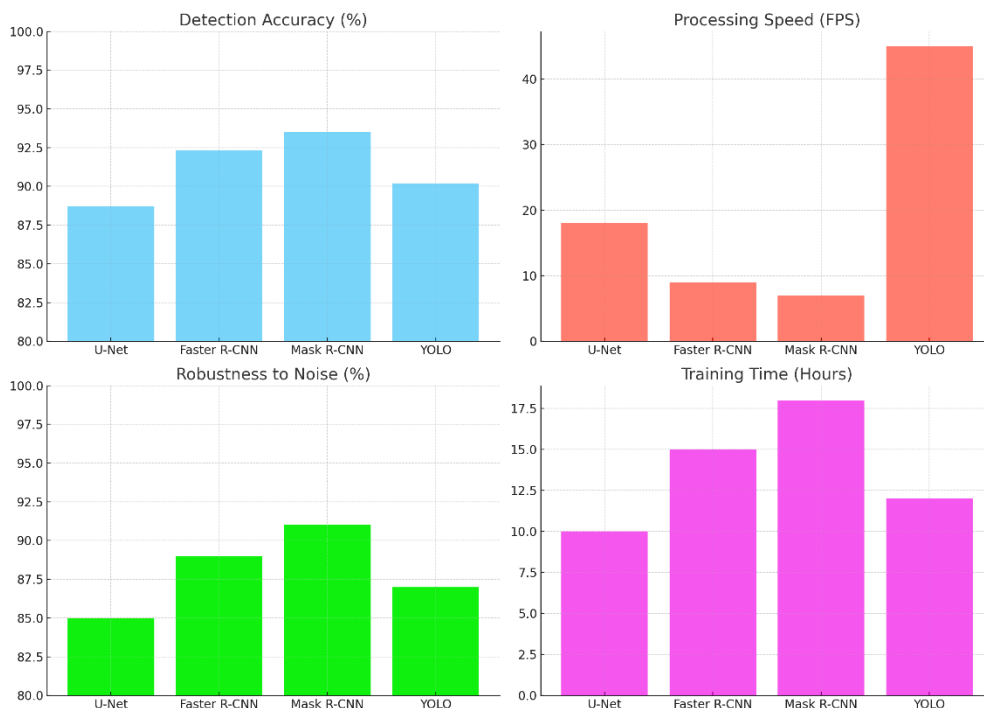


Рис. 2. Результати аналізу моделей виявлення об'єктів на основі глибокого навчання

Підготовка даних, налаштування середовища та процесу навчання.

Аналіз наявних датасетів показав, що WebUOT-1M <https://github.com/983632847/Awesome-Multimodal-Object-Tracking/tree/main/WebUOT-1M> є найбільшим набором даних для відстеження підводних об'єктів, що охоплює різні категорії цілей та сценарії. Набір даних охоплює різні підводні середовища, такі як море, річка, озеро. розділено на навчальний та тестовий набори. Навчальний набір містить 1020 відео (693 тис. кадрів, 6,64 години), тоді як тестовий набір містить 480 відео (407 тис. кадрів, 3,86 години). На відміну від більшості попередніх наборів даних UOT [22], WebUOT-1M включає неживі об'єкти

(наприклад, безпілотний підводний апарат та дрони типу амфібія), що призводить до більш повного та універсального аналізу.

Вибір відповідного обладнання, програмного забезпечення, фреймворків та підготовки навчальних даних відіграє вирішальну роль у забезпеченні ефективності та надійності експериментів. Крім того, для ефективного навчання моделей та отримання висновків важливе значення мають характеристики графічного процесора. Робоча станція на основі Ubuntu 21.04 обладнана двома графічними процесорами GeForce 2080Ti з 11 ГБ VRAM. Для управління процесом навчання та відстеження експерименту використовувалась платформа Weights and Biases, призначена для відстеження та візуалізації експериментів проєктів машинного навчання. Вона дозволила реєструвати та відстежувати параметри навчання, метрики та гіперпараметри, полегшуючи відтворюваність та надаючи комплексну картину про хід дослідження. Крім того, були використані переваги платформи паралельних обчислень NVIDIA CUDA та бібліотеки cuDNN. Ці технології з GPU-прискоренням значно підвищили швидкість навчання та ефективність моделей завдяки використанню можливостей паралельних обчислень графічних процесорів.

Застосування алгоритмів відстеження. В дослідженні, було апробовано два алгоритми відстеження об'єктів: BoT-SORT та ByteTrack. Проте, ці алгоритми не показали якісні результати трекування на тестових відео даних. Особливо, вони мали невисоку здатність аналізувати форму об'єктів, і тому мало ефективно вирішували проблему оклюзій. Крім того, обидва алгоритми не вирішували проблему виходу об'єкта за межі кадру.

Аналіз результатів. Порівняння моделей на основі дифузії та інтеграції з моделями глибокого навчання YOLO з застосуванням метрик <https://docs.ultralytics.com/guides/yolo-performance-metrics/> представлено в табл. 1.

Таблиця 1

Результати тестування моделей YOLO

Model	Anchor	Backbone	mAP 50	mAP @ [.5: .95]
YOLOv5s	Yes	V5CSPDarknet	94.4	60.1
YOLOv6s	No	EfficientRep	94.0	54.5
YOLOv8s	No	V8CSPDarknet	94.7	64.4

Аналіз продемонстрував можливості та обмеження кожної версії моделі. Результати показують, що YOLOv8 з її вдосконаленою архітектурою YOLOv8CSPDarknet демонструє найвищу загальну продуктивність. Вона досягла виняткової середньої точності mAP, що свідчить про точне виявлення у широкому діапазоні порогових значень IoU. Крім того, YOLOv8 продемонструвала конкурентний час виведення, що свідчить про її потенціал для виявлення об'єктів у реальному часі. З іншого боку, YOLOv6, використовуючи як основу anchor-free EfficientRep архітектуру, продемонструвала найшвидший час виведення, але дещо нижчі показники метрик.

Таким чином, YOLOv8 пропонує кілька переваг порівняно зі своїми попередниками та іншими моделями виявлення об'єктів:

- YOLOv8 досягає найвищої точності в різних тестах виявлення об'єктів;
- модель має високу швидкість логічного висновку, що робить її придатною для застосувань у реальному часі, таких як автономні транспортні засоби та робототехніка;
- YOLOv8 є легкою та вимагає менше обчислювальних ресурсів, ніж інші моделі, що робить її ідеальною для розгортання на периферійних пристроях.

В процесі дослідження було апробовано два алгоритми відстеження – BoT-SORT і ByteTrack. При використанні цих алгоритмів на тестових відеоданих прийнятна якість відстеження не була досягнута.

Висновки та подальші дослідження. Інтеграція попередньої обробки на основі дифузії є важливою для використання більшого потенціалу нейронних моделей у реальних підводних сценаріях. Майбутні дослідження повинні бути зосереджені на оптимізації архітектур дифузійних моделей і дослідженні їх інтеграції з різними платформами глибокого навчання [19, 21], такими як YOLO, щоб досягти вищої продуктивності в складних підводних середовищах. Крім того, вивчення використання рекурентних нейронних мереж (RNN) або трансформерів у структурі дифузії може дозволити моделі краще фіксувати тимчасові залежності у відеопослідовностях, що призведе до покращення продуктивності в задачах відстеження підводних об'єктів.

Надання надійних оцінок невизначеності має вирішальне значення для розгортання моделей дифузії в реальних підводних додатках. Оцінки невизначеності можуть допомогти ідентифікувати області зображення, де модель є менш впевненою, дозволяючи алгоритмам виявлення об'єктів і відстеження з нижчим порогом виявлення приймати більш обґрунтовані рішення. Одним із підходів є використання моделей байєсівської дифузії, які забезпечують принципову основу для оцінки невизначеності в прогнозах моделі. Інший підхід полягає у використанні ансамблевих методів, коли кілька моделей дифузії навчаються на різних підмножинах даних, а їхні прогнози об'єднуються для оцінки невизначеності.

Список літератури:

1. Color balance and fusion for underwater image enhancement / C. O. Ancuti et al. IEEE transactions on image processing. 2018. Vol. 27, no. 1. P. 379–393. URL: <https://doi.org/10.1109/tip.2017.2759252>.
2. GUDCP: generalization of underwater dark channel prior for underwater image restoration / Z. Liang et al. IEEE transactions on circuits and systems for video technology. 2022. Vol. 32, no. 7. P. 4879–4884. URL: <https://doi.org/10.1109/tcsvt.2021.3114230>.
3. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior / C.-Y. Li et al. IEEE transactions on image processing. 2016. Vol. 25, no. 12. P. 5664–5677. URL: <https://doi.org/10.1109/tip.2016.2612882>.
4. Peng Y.-T., Cao K., Cosman P. C. Generalization of the dark channel prior for single image restoration. IEEE transactions on image processing. 2018. Vol. 27, no. 6. P. 2856–2868. URL: <https://doi.org/10.1109/tip.2018.2813092>.
5. Uncertainty inspired underwater image enhancement / Z. Fu et al. Computer vision – ECCV 2022. 2022. P. 465–482. URL: https://doi.org/10.1007/978-3-031-19797-0_27.
6. Deep pixel-to-pixel network for underwater image enhancement and restoration / X. Sun et al. IET image processing. 2019. Vol. 13, no. 3. P. 469–474. URL: <https://doi.org/10.1049/iet-ipr.2018.5237>.
7. An underwater image enhancement benchmark dataset and beyond / C. Li et al. 2019. URL: <https://doi.org/10.48550/ARXIV.1901.05495>.
8. Li C., Anwar S., Porikli F. Underwater scene prior inspired deep underwater image and video enhancement. Pattern recognition. 2020. Vol. 98. P. 107038. URL: <https://doi.org/10.1016/j.patcog.2019.107038>.
9. Naik A., Swarnakar A., Mittal K. Shallow-UWnet : compressed model for underwater image enhancement. arXiv, 2021. URL: <https://doi.org/10.48550/ARXIV.2101.02073>.
10. Underwater image enhancement via medium transmission-guided multi-color space embedding / C. Li et al. 2021. URL: <https://doi.org/10.48550/ARXIV.2104.13015>.
11. Fabbri C., Islam M. J., Sattar J. Enhancing underwater imagery using generative adversarial networks. arXiv, 2018. URL: <https://doi.org/10.48550/ARXIV.1801.04011>.
12. Awad, A., Saleem, A., Paheding, S., Lucas, E., Al-Ratrout, S., & Havens, T. C. Evaluating the impact of underwater image enhancement on object detection performance: a comprehensive study. arXiv, 2024. URL: <https://arxiv.org/html/2411.14626v2>.

13. UnDIVE: generalized underwater video enhancement using generative priors / S. Srinath et al. arXiv, 2024. URL: <https://doi.org/10.48550/ARXIV.2411.05886>.
14. Underwater image enhancement by diffusion model with customized clip-classifier / S. Liu et al. 2025. URL: <https://doi.org/10.2139/ssrn.5165323>.
15. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models. arXiv, 2020. URL: <https://doi.org/10.48550/ARXIV.2006.11239>.
16. Underwater image enhancement method based on denoising diffusion probabilistic model / S. Lu et al. Journal of visual communication and image representation. 2023. Vol. 96. P. 103926. URL: <https://doi.org/10.1016/j.jvcir.2023.103926>.
17. Underwater image enhancement with physical-based denoising diffusion implicit models / N. G. Bach et al. arXiv, 2024. URL: <https://doi.org/10.48550/ARXIV.2409.18476>.
18. Image-Conditional Diffusion Transformer for Underwater Image Enhancement / Nie X. et al. arXiv, 2024. URL: <https://doi.org/10.48550/arXiv.2407.0538>.
19. Bayesian neural networks for one-to-many mapping in image enhancement / G. Huang et al. arXiv, 2025. URL: <https://doi.org/10.48550/ARXIV.2501.14265>.
20. Small object detection: a comprehensive survey on challenges, techniques and real-world applications / M. Nikouei et al. arXiv, 2025. URL: <https://doi.org/10.48550/ARXIV.2503.20516>.
21. Moorton Z., Kurt Z., Woo W. L. State of the art applications of deep learning within tracking and detecting marine debris: a survey. arXiv, 2024. URL: <https://doi.org/10.48550/ARXIV.2403.18067>.
22. Comprehensive underwater object tracking benchmark dataset and underwater image enhancement with gan / Karen Panetta et al. IEEE Journal of Oceanic Engineering. 2021. Vol. 47(1). P. 59–75, <https://neurips.cc/virtual/2024/poster/97605>.

References:

1. Ancuti, C.O., Ancuti, C., De Vleeschouwer, C., & Bekaert, P. (2018). Color Balance and Fusion for Underwater Image Enhancement. IEEE Transactions on Image Processing, 27, 379–393. <https://doi.org/10.1109/TIP.2017.2759252>.
2. Zheng Liang, Xueyan Ding, Yafei Wang, Xiaohong Yan, and Xianping Fu. (2022). GUDCP: Generalization of Underwater Dark Channel Prior for Underwater Image Restoration. IEEE Trans. Cir. and Sys. for Video Technol. 32, 7 (July 2022), 4879–4884. <https://doi.org/10.1109/TCSVT.2021.3114230>.
3. Chong-Yi Li, Ji-Chang Guo, Run-Min Cong, Yan-Wei Pang, and Bo Wang. (2016). Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. Trans. Img. Proc. 25, 12 (December 2016), 5664–5677. <https://doi.org/10.1109/TIP.2016.2612882>.
4. Y.-T. Peng, K. Cao, and P. C. Cosman. (2018). Generalization of the dark channel prior for single image restoration, IEEE Trans. Image Process., vol. 27, no. 6, pp. 2856–2868. <https://ieeexplore.ieee.org/document/8307410>.
5. Zhenqi Fu, Wu Wang, Yue Huang, Xinghao Ding, and Kai-Kuang Ma. (2022). Uncertainty Inspired Underwater Image Enhancement. In Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVIII. Springer-Verlag, Berlin, Heidelberg, 465–482. https://doi.org/10.1007/978-3-031-19797-0_27.
6. X. Sun, L. Liu, Q. Li, J. Dong, E. Lima, and R. Yin. (2019). Deep pixel-to-pixel network for underwater image enhancement and restoration, IET Image Process., vol. 13, no. 3, pp. 469–474. https://llpspark.github.io/Download/Publications/IET_processing_2018.pdf

7. C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao. (2020). An underwater image enhancement benchmark dataset and beyond, *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389. <https://ieeexplore.ieee.org/document/8917818>.
8. C. Li, S. Anwar, and F. Porikli. (2020). Underwater scene prior inspired deep underwater image and video enhancement, *Pattern Recognit.*, vol. 98, pp. 107038–107049. <https://www.sciencedirect.com/science/article/abs/pii/S0031320319303401>.
9. A. Naik, A. Swarnakar, and K. Mittal. (2021). Shallow-UWnet: Compressed model for underwater image enhancement, in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 15853–15854. <https://doi.org/10.48550/arXiv.2101.02073>.
10. C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren. (2021). Underwater image enhancement via medium transmission-guided multi-color space embedding, *IEEE Trans. Image Process.*, vol. 30, pp. 4985–5000. <https://ieeexplore.ieee.org/document/9426457>.
11. C. Fabbri, M. J. Islam, and J. Sattar. (2018). Enhancing underwater imagery using generative adversarial networks, in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 7159–7165. <https://doi.org/10.48550/arXiv.1801.04011>.
12. Awad, A., Saleem, A., Paheding, S., Lucas, E., Al-Ratrout, S., & Havens, T. C. (2024). Evaluating the Impact of Underwater Image Enhancement on Object Detection Performance: A Comprehensive Study. <https://arxiv.org/html/2411.14626v2>.
13. Srinath, S., Chandrasekar, A., Jamadagni, H., Soundararajan, R., & Prathosh A. (2024). UnDIVE: Generalized Underwater Video Enhancement Using Generative Priors. <https://doi.org/10.48550/arXiv.2411.05886>.
14. Liu, S., Li, K., Ding, Y., & Qi, Q. (2024). Underwater Image Enhancement by Diffusion Model with Customized Clip-Classfier. <https://dx.doi.org/10.2139/ssrn.5165323>.
15. Jonathan Ho, Ajay Jain, and Pieter Abbeel. (2020). Denoising diffusion probabilistic models. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20)*. Curran Associates Inc., Red Hook, NY, USA, Article 574, 6840–6851. <https://dl.acm.org/doi/abs/10.5555/3495724.3496298>.
16. Lu, S., Guan, F., Zhang, H., & Lai, H. (2023). Underwater image enhancement method based on denoising diffusion probabilistic model. *J. Vis. Commun. Image Represent.*, 96, 103926. <https://doi.org/10.1016/j.jvcir.2023.103926>.
17. Bach, N. G., Tran, C. M., Kamioka, E., & Tan, P. X. (2024). Underwater Image Enhancement with Physical-based Denoising Diffusion Implicit Models. *arXiv preprint arXiv:2409.18476*.
18. Nie, X., Pan, S., Zhai, X., Tao, S., Qu, F., Wang, B., Ge, H., & Xiao, G. (2024). Image-Conditional Diffusion Transformer for Underwater Image Enhancement. <https://doi.org/10.48550/arXiv.2501.14265>.
19. Huang, G., Anantrasirichai, N., Ye, F., Qi, Z., Lin, R. R., Yang, Q., & Bull, D. (2025). Bayesian Neural Networks for One-to-Many Mapping in Image Enhancement. *arXiv preprint arXiv:2501.14265*.
20. Nikouei, M., Baroutian, B., Nabavi, S., Taraghi, F., Aghaei, A., Sajedi, A., & Moghaddam, M. E. (2025). Small Object Detection: A Comprehensive Survey on Challenges, Techniques and Real-World Applications. *arXiv preprint arXiv:2503.20516*.
21. Moorton, Z., Kurt, Z., & Woo, W. L. (2024). State of the art applications of deep learning within tracking and detecting marine debris: A survey. *arXiv preprint arXiv:2403.18067*.
22. Karen Panetta, Landry Kezebou, Victor Oludare, and Sos Agaian. (2021). Comprehensive underwater object tracking benchmark dataset and underwater image

enhancement with gan. IEEE Journal of Oceanic Engineering, 47(1):59–75, <https://neurips.cc/virtual/2024/poster/97605>.

*Статтю представив д.т.н., проф. Одеського імені І.І. Мечникова
Національного Університету Ю.О. Гунченко*

Надійшла (received) 15.05.2025

Moroz Dmytro, PhD student
National University "Odesa Polytechnic"
Shevtchenko Ave., 1, Odesa, Ukraine, 65044
Tel.: +380-93-234-78-27, email: mordmytro@stud.op.edu.ua
ORCID ID: 0009-0002-1082-5033

Liubchenko Vira, Doctor of Science, Professor
National University "Odesa Polytechnic"
Shevtchenko Ave., 1, Odesa, Ukraine, 65044
Tel.: +380-50-392-21-50, email: lvv@op.edu.ua
ORCID ID: 0000-0002-4611-7832

УДК 004.93

Інтеграція дифузійних моделей та глибокого навчання для виявлення і відстеження об'єктів в підводних зображеннях // Мороз Д.В., Любченко В.В.
// Вісник НТУ "ХПІ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПІ". – 2025. – № 2 (14). – С. 62 – 75.

Складні оптичні властивості води призводять до спотворення кольору, розмиття та низького контрасту, що ускладнює виявлення та відстеження об'єктів. У поєднанні з дефіцитом розмічених даних підводних зображень це серйозно обмежує продуктивність систем виявлення та відстеження об'єктів на основі глибокого навчання. Досліджується застосування дифузійних моделей для попередньої обробки підводних зображень виявлення та відстеження об'єктів за допомогою глибокого навчання. Табл.: 1. Бібліогр.: 22 назв.

Ключові слова: підводні зображення, відновлення зображень, моделі дифузії, глибоке навчання.

UDK 004.93

Integrating diffusion models and deep learning for object detection and tracking in underwater images // Moroz D.V., Lyubchenko V.V. // Herald of the National Technical University "KhPI". Series of "Informatics and Modeling". – Kharkov: NTU "KhPI". – 2025. – № 2 (14). – P. 62 – 75.

The complex optical properties of water lead to color distortion, blurring, and low contrast, which makes object detection and tracking difficult. Combined with the scarcity of labeled underwater image data, this severely limits the performance of deep learning-based object detection and tracking systems. The application of diffusion models for preprocessing underwater images for object detection and tracking using deep learning is investigated. Tabl.: 1. Refs.: 22 item.

Keywords: underwater images, image restoration, diffusion models, deep learning.