DOI 10.36074/grail-of-science.20.02.2026.121

МУЛЬТИМОДАЛЬНІ МЕТОДИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ У ПІДВОДНОМУ СЕРЕДОВИЩІ

Шумейко Костянтин Павлович
магістр, аспірант
Національний університет «Одеська політехніка», Україна

Постановка проблеми. Розпізнавання об'єктів у підводному середовищі супроводжується низкою фундаментальних проблем, зумовлених фізичними властивостями води та специфікою підводної зйомки. Ці чинники обмежують застосовність стандартних методів комп'ютерного зору без їх цільової адаптації. З точки зору машинного навчання додатковим обмеженням є висока варіативність підводних умов і обмежена доступність розмічених даних. Підводні зображення істотно відрізняються залежно від глибини, прозорості води, типу освітлення та географічного регіону. Сукупність фізичних, геометричних і обчислювальних чинників робить задачу розпізнавання об'єктів під водою суттєво складнішою і зумовлює необхідність розроблення спеціалізованих методів обробки, що враховують специфіку підводного середовища.

Аналіз досліджень та публікацій. У дослідженнях з підводного розпізнавання об'єктів значна увага приділяється оптичним методам, заснованим на адаптації сучасних детекторів загального призначення (Faster R-CNN, SSD, YOLO та їх модифікацій) до специфіки підводного середовища. Основні проблеми пов'язані з деградацією зображень унаслідок поглинання та розсіювання світла, зниженням контрастності та вираженими колірними спотвореннями. У низці робіт підкреслюється, що наявні набори даних часто мають обмежене різноманіття сцен і умов, що призводить до слабкої узагальнювальної здатності навчених моделей [1]. Систематичний аналіз цих обмежень і можливих шляхів удосконалення всього процесу обробки даних підводного розпізнавання — від формування даних до архітектур нейронних мереж — представлено в [1]. Для уніфікації порівняння алгоритмів було запропоновано зведений тест DUO, який об'єднав і повторно анотовано наявні набори даних, що дало змогу коректно зіставляти різні методи в межах єдиного експериментального протоколу [2]. Окремий напрям досліджень пов'язаний із розпізнаванням об'єктів за даними гідроакустичних сенсорів, таких як forward-looking sonar та синтетична апертурна сонарна зйомка (SAS). Акустичні зображення характеризуються високою стійкістю до оптичних завад і забезпечують більшу дальність виявлення, однак мають низьку просторову роздільну здатність і високий рівень шуму. Низка робіт показала, що використання згорткових нейронних мереж і методів перенесення навчання

 Авторські права захищені | Creative Commons Attribution-ShareAlike 4.0 International License

дозволяє суттєво підвищити якість детекції об'єктів на сонарних зображеннях порівняно з класичними ознаковими підходами [3]. Додаткові дослідження, виконані європейськими науковими групами, демонструють ефективність спеціалізованих архітектур для SAS-даних і підкреслюють важливість урахування геометричних і статистичних особливостей акустичних зображень під час навчання моделей [4]. Активно розвивається напрям, що передбачає спільне використання оптичних і акустичних даних як взаємодоповнювальних джерел інформації. У таких підходах оптичні сенсори забезпечують високу деталізацію форми та текстури об'єктів на малих дистанціях, тоді як акустичні системи підвищують надійність виявлення в умовах поганої видимості. Сучасні роботи пропонують мультимодальні архітектури з об'єднанням даних на рівні ознак або рішень, приділяючи особливу увагу задачам синхронізації, просторово-часового узгодження та калібрування сенсорів [5]. Показано, що акустико-оптичне злиття даних дозволяє підвищити стійкість і точність розпізнавання підводних об'єктів у складних і змінних умовах експлуатації, що підтверджується дослідженнями, опублікованими в європейських журналах з робототехніки та сенсорних систем[5, 6].

Постановка задачі. Розроблення та дослідження методів розпізнавання об'єктів у підводному середовищі на основі мультимодальних даних, що передбачають спільне використання інформації від оптичних і акустичних сенсорів у поєднанні з алгоритмами попередньої обробки та інтелектуального аналізу даних. Основна увага приділяється підвищенню точності, стійкості та надійності розпізнавання об'єктів в умовах обмеженої видимості, шумів і змінних характеристик водного середовища за рахунок ефективного об'єднання різнорідних даних та адаптації алгоритмів обробки до специфіки підводних умов.

Виклад основного матеріалу. Попередня обробка підводних зображень спрямована на компенсацію деградацій, спричинених поглинанням і розсіюванням світла, колірними зсувами та нерівномірним освітленням. На практиці широко застосовуються методи корекції колірного балансу, нормалізації інтенсивності та підсилення локального контрасту, зокрема баланс білого і CLAHE, завдяки їхній обчислювальній ефективності. Складніші підходи базуються на фізичних моделях формування підводного зображення, таких як моделі Кошмідера та Джаффе–МакГламарі, які дозволяють враховувати внесок розсіяного світла та частково відновлювати контраст сцени.

Попередня обробка акустичних даних спрямована на зменшення впливу шумів і спотворень, характерних для сонарних зображень. Типові методи включають фільтрацію адитивного та спекл-шуму, нормалізацію амплітудних характеристик і придушення фонових відбиттів. Для виділення інформативних ознак застосовуються як класичні методи обробки сигналів, так і підходи на основі глибокого навчання. Водночас акустичні дані характеризуються низькою просторовою роздільною здатністю, високим рівнем шуму та обмеженою здатністю передавати форму й тип об'єктів, що ускладнює подальше розпізнавання при використанні лише цього типу сенсорної інформації.

Для спільної обробки оптичних і акустичних даних можна застосовуватися нейронні мережі з мультимодальною архітектурою, у яких кожен тип даних

All rights reserved | Creative Commons Attribution-ShareAlike 4.0 International License

обробляється окремою підмережею з подальшим об'єднанням ознак. Злиття інформації може виконуватися на рівні ознак або рішень із використанням механізмів уваги чи спільних латентних представлень, що дозволяє нейронній мережі враховувати взаємодоповнюваність модальностей і підвищувати точність та стійкість розпізнавання

Висновки та пропозиції. Використання одиничних сенсорних даних не забезпечує необхідного рівня надійності розпізнавання об'єктів у складних підводних умовах, тоді як мультимодальний підхід дозволяє суттєво підвищити стійкість і інформативність системи за рахунок взаємодоповнюваності оптичних і акустичних даних. Такий підхід дає змогу компенсувати недоліки окремих сенсорів, а також передбачає розроблення спеціалізованих алгоритмів обробки, адаптованих до змінних умов водного середовища та обмежень.

Список використаних джерел:

[1] Li, J., Liu, Z., Chen, Y., & Wang, Y. (2022). Rethinking general underwater object detection: Datasets, challenges and solutions. *Neurocomputing, 491*, 144–160. https://doi.org/10.1016/j.neucom.2022.03.067
[2] Islam, M. J., Xia, Y., & Sattar, J. (2021). Detecting underwater objects: A benchmark and dataset (DUO). *arXiv preprint*, arXiv:2106.05681. https://arxiv.org/abs/2106.05681
[3] Hurtós, N., Cufí, X., Petillot, Y., & Salvi, J. (2017). Deep learning for forward-looking sonar image recognition. *arXiv preprint*, arXiv:1709.02601. https://arxiv.org/abs/1709.02601
[4] Reed, S., Hunter, A., & Petillot, Y. (2018). Automatic target recognition on synthetic aperture sonar imagery using deep learning. *IET Radar, Sonar & Navigation, 12*(6), 1281–1289. https://doi.org/10.1049/iet-rsn.2017.0451
[5] Li, Y., Wang, S., Zhao, J., & Xu, H. (2023). Underwater acoustic–optical fusion detection using deep learning. *Journal of Field Robotics, 40*(6), 1285–1302. https://doi.org/10.1002/rob.22432
[6] Negahdaripour, S., Firoozfam, P., & Sabzmeydani, P. (2020). Multimodal optical–acoustic sensor fusion for underwater object detection. *Robotics and Autonomous Systems, 124*, 103392. https://doi.org/10.1016/j.robot.2019.103392

MULTIMODAL METHODS FOR OBJECT RECOGNITION IN THE UNDERWATER ENVIRONMENT

Kostiantyn Shumeiko
Master's degree holder, PhD student
Odesa Polytechnic National University, Ukraine

 Авторські права захищені | Creative Commons Attribution-ShareAlike 4.0 International License