

ВИКОРИСТАННЯ АУГМЕНТАЦІЇ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ НАВЧАННЯ МОДЕЛІ ШТУЧНОГО ЗОРУ ДЛЯ ВИЗНАЧЕННЯ РАКУРСУ СУДНА

Пашенко О. Л., старший викладач кафедри електронних комплексів судноводіння
Національного університету «Одеська морська академія», м. Одеса, Україна,
e-mail: olena.p.3333@gmail.com, ORCID: 0000-0002-0462-0130.

Дослідження присвячене аналізу впливу методів аугментації даних на ефективність тренування моделей комп'ютерного зору для визначення ракурсу суден у морському середовищі. У роботі розглядаються особливості застосування геометричних, кольорних та просторово-локальних трансформацій, враховуючи специфіку морських зображень, таких як наявність горизонту, мінливі умови освітлення та оклюзії. Актуальність дослідження зумовлена обмеженою кількістю даних і дисбалансом класів у наборах морських зображень, що ускладнює досягнення високої точності класифікації. Для експериментів використано базу даних із 1742 зображень, включаючи приклади без об'єктів розпізнавання. Для проведення дослідження була створена модель на базі найлегшої YOLOv8n. Тренування моделі відбувалося з використанням різних типів аугментацій. Результати показують, що показник $mAP50 = 0,61229$ без використання аугментації може бути покращений завдяки використанню деяких кольорних аугментацій та геометричних трансформацій ($mAP50$ до 0,64649). Водночас такі трансформації можуть і погіршити результати, знижуючи якість через спотворення орієнтації суден. Отримані висновки мають практичне значення для систем автономного судноводіння, сприяючи підвищенню точності розпізнавання в реальних умовах.

Ключові слова: аугментація даних; комп'ютерний зір; ракурс суден; глибоке навчання; геометричні трансформації; кольорні аугментації; просторово-локальні трансформації; YOLOv8n; $mAP50$; $mAP50-95$; precision; recall.

DOI: 10.33815/2313-4763.2025.2.31.052-063

Вступ. Сучасний розвиток штучного інтелекту та комп'ютерного зору відкриває нові можливості для автоматизації процесів у сфері морської навігації. Одним із важливих напрямів є створення систем, здатних не лише виявляти судна на зображеннях, а й визначати їх ракурс, що дозволяє прогнозувати траєкторію руху та підвищувати рівень безпеки на морі. Визначення ракурсу судна є одним із важливих завдань, оскільки від нього залежить ефективність ухвалення навігаційних рішень у режимі реального часу, особливо в умовах інтенсивного морського трафіку.

Упродовж останніх років моделі штучного зору [1, 2], зокрема YOLOv8, зарекомендували себе як високопродуктивні рішення для задач виявлення й класифікації об'єктів. Вони поєднують швидкодію та високу точність, що робить їх придатними для інтеграції в бортові та дистанційні системи підтримки навігації. Проте ефективність таких моделей значною мірою залежить від різноманітності та якості навчальних даних [3–5]. Морське середовище характеризується зміною освітлення та погодних умов [6]. Це призводить до того, що модель, навчена на обмеженому наборі даних, може втрачати точність у реальних умовах експлуатації.

Одним із поширених способів подолання цієї проблеми є використання аугментацій – штучного збільшення різноманіття даних шляхом застосування кольорних, геометричних та інших трансформацій. Такі методи традиційно застосовуються для покращення узагальнюючої здатності моделей і зменшення перенавчання. Водночас у задачі визначення ракурсу судна аугментації мають особливе значення, оскільки неправильний підбір трансформацій може призвести не до підвищення, а навпаки – до зниження точності моделі. Це зумовлює необхідність цілеспрямованого дослідження впливу окремих типів аугментацій на результати роботи моделі в умовах морського середовища.

Постановка проблеми. Безвідмовне функціонування судноплавної інфраструктури безпосередньо залежить від стабільної роботи екіпажів суден. Точне визначення ракурсу суден у морському середовищі є однією з задач для забезпечення безпеки та ефективності

навігаційних систем. Моделі комп'ютерного зору, такі як YOLOv8, потребують різноманітних і збалансованих навчальних даних для досягнення високих показників ефективності. Однак набори даних для морських зображень часто обмежені за обсягом і характеризуються дисбалансом класів, де деякі ракурси представлені значно меншою кількістю прикладів. Крім того, морське середовище ускладнює завдання через мінливі погодні умови, зміну освітлення та часткову оклюзію об'єктів.

Аугментація даних може компенсувати ці недоліки шляхом трансформації зображень, але неправильний вибір трансформацій, таких як віддзеркалення чи обертання, може призвести до спотворення орієнтації судна, що критично для задачі класифікації ракурсу. Таким чином, виникає проблема: які методи аугментації є оптимальними для підвищення продуктивності моделі в умовах обмежених даних, зберігаючи при цьому фізичну достовірність морських сцен, і як кількісно оцінити їхній вплив на метрики якості моделі.

Аналіз останніх досліджень та публікацій. Дослідження в галузі комп'ютерного зору, особливо в задачах із обмеженою кількістю даних, активно використовують методи аугментації для підвищення продуктивності моделей глибокого навчання. Ці методи дозволяють штучно розширювати тренувальні набори, покращуючи узагальнюючу здатність моделей і зменшуючи ризик перенавчання.

Огляд методів аугментації, представлений у [7], підкреслює їхню універсальність для обробки зображень та відео. Автори зазначають, що аугментація підвищує точність на тестових даних шляхом імітації реальних умов. Основні групи методів аугментації включають: геометричні, колірні, просторово-локальні, генеративні підходи для створення синтетичних зображень [8–10]. Генеративні методи, описані в [11], дозволяють створювати реалістичні синтетичні морські сцени за допомогою ігрових рушіїв (game engines) або GAN, що зменшує потребу в зборі реальних даних.

Дослідження [5] показують, що в морських умовах деякі геометричні трансформації, такі як віддзеркалення чи обертання, можуть знижувати якість класифікації, порушуючи фізичний сенс сцени, наприклад, орієнтацію горизонту чи напрям руху судна. Аналогічно, у [10] підкреслюється, що методи, такі як випадкове стирання, є ефективними для подолання оклюзій у морських зображеннях, тоді як віддзеркалення може вводити модель в оману.

Таким чином, сучасні дослідження підкреслюють важливість адаптації аугментацій до специфіки задачі, особливо в морському середовищі, де неправильні трансформації можуть погіршити результати. Перспективи подальших досліджень пов'язані з розробкою адаптивних стратегій аугментації та вдосконаленням генеративних методів.

Мета та задачі дослідження. У роботі [12] були опубліковані дані щодо роботи моделі на власному наборі даних. Через незбалансованість класів, відсутність достатньої кількості прикладів для деяких ракурсів, а також складності виявлення великих суден на передньому плані та інше, модель потребувала донавчання. Основна мета цього дослідження полягає у визначенні впливу різних типів аугментацій – колірних, геометричних та просторово-локальних – на ефективність доповнення даних під час навчання моделей комп'ютерного зору.

Для досягнення поставленої мети задачами дослідження є проведення аналізу різних видів аугментацій у комп'ютерному зорі з особливим акцентом на їх застосуванні в морських умовах, навчання моделі YOLOv8n із використанням різних типів аугментацій та без них для порівняння результатів, вплив кожного методу аугментації на ключові метрики моделі, а також визначення найбільш доцільних трансформацій, що підвищують стійкість до варіативних умов зйомки без спотворення критично важливих просторових ознак сцени.

Основна частина

1. Аугментація

1.1 Аугментація та її різноманітня

Аугментація даних відіграє ключову роль у підвищенні ефективності моделей глибокого навчання. Вона дозволяє розширити навчальний набір шляхом генерації нових варіацій на основі наявних зображень без необхідності збору додаткових даних. Це, у свою

чергу, зменшує ризик перенавчання, оскільки модель не фокусується виключно на унікальних особливостях вихідних даних. Крім того, аугментація сприяє покращенню здатності моделі до узагальнення, що дає змогу ефективніше розпізнавати об'єкти в умовах, що відрізняються від навчальних. У результаті підвищується загальна точність розпізнавання валідаційних і тестових даних, що є критично важливим для практичного застосування в реальних сценаріях.

Як вже зазначалося, аугментація – це багатофункціональний засіб для збільшення різноманіття навчальних зразків. Багато сфер знань не мають доступу до великого об'єму даних. Залежно від того, що саме повинна розпізнавати модель штучного зору, доцільно підбирати відповідні види аугментації для штучного урізноманітнення набору. Можна виділити наступні види аугментації:

- Геометричні трансформації: горизонтальне або вертикальне віддзеркалення, випадкове обертання, масштабування, вирізання, зсув. Ці операції призначені для зміни просторового положення об'єктів.

- Зміни кольору: регулювання яскравості, контрасту, насиченості, відтінку, додавання шуму, розмиття тощо. Вони імітують різні умови освітлення та шуму камери.

- Випадкове затемнення/вирізання, що покращує стійкість моделі до часткової неповноти об'єкта (оклюзії).

- Поєднання зображень дозволяє комбінувати зображення в одне ціле, шляхом усереднення або накладання фрагментів. Це створює нові «змішані» приклади та зменшує чутливість до конкретних пікселів.

- Генеративні підходи: створення синтетичних кадрів (відеозаписів чи зображень) за допомогою GAN, 3D-симуляторів або інших моделей генерації. Це перспективний напрямок, особливо коли базові методи мають свої обмеження [13].

Вибір відповідних методів аугментації має базуватися на специфіці конкретного застосування та характеристиках набору даних.

1.2 Методи аугментації зображень для задачі розпізнавання ракурсу суден

Як зазначалося раніше, застосування аугментації дозволяє створювати модифіковані копії наявних зображень (наприклад, шляхом обертання, відображення чи зміни кольору), що сприяє кращому узагальненню моделей на нових даних. Але при таких перетвореннях варто враховувати характер морського середовища: завжди присутній горизонт, море знаходиться унизу, а небо зверху. Ці особливості повинні залишатися в аугментованих зображеннях. Як зазначено в [10], при невеликому обсязі даних і специфічних умовах «фонова» експертиза є критичною при виборі аугментацій. Тобто саме знання особливостей зображення морського середовища (хвилі на воді, метеорологічні опади, перепади освітлення в денний та нічний час тощо) допоможе обирати ефективні методи трансформацій.

У цій роботі автором запропоновані три групи аугментацій, а саме: аугментації кольорового простору, геометричні трансформації та просторово-локальні перетворення. Інші стандартні аугментації знижують якість навчання, а саме обертання, віддзеркалення, композиційні трансформації було виключено, оскільки вони можуть змінювати просторову структуру або напрямок об'єкта, що є критичним у задачах визначення ракурсу. Наприклад, дзеркальне відображення об'єктів є недоцільним при навчанні моделі штучного зору для визначення ракурсу суден, оскільки повністю спотворює результати дослідження.

З метою підвищення стійкості моделі до незначних коливань у кольоровому представленні зображень, спричинених варіативністю умов зйомки, та розширення бази даних, у процесі навчання було застосовано кольорові аугментації у вигляді випадкового регулювання відтінку (Hue Shift), насиченості (Saturation Shift) та яскравості (Brightness Shift). Регулювання відтінку особливо ефективне для задач комп'ютерного зору в умовах змінного природного освітлення, коли вигляд об'єктів може суттєво змінюватися. Техніка регулювання насиченості дозволяє моделі краще адаптуватися до різноманітних умов зйомки, зокрема до змін погоди чи налаштувань камери. При навчанні моделей, які повинні працювати за реальних умов освітлення закладається зміна яскравості. На рис. 1

представлені випадкові значення параметрів зміни відтінку, насиченості та яскравості прийнятні для розпізнавання об'єктів у морському середовищі.

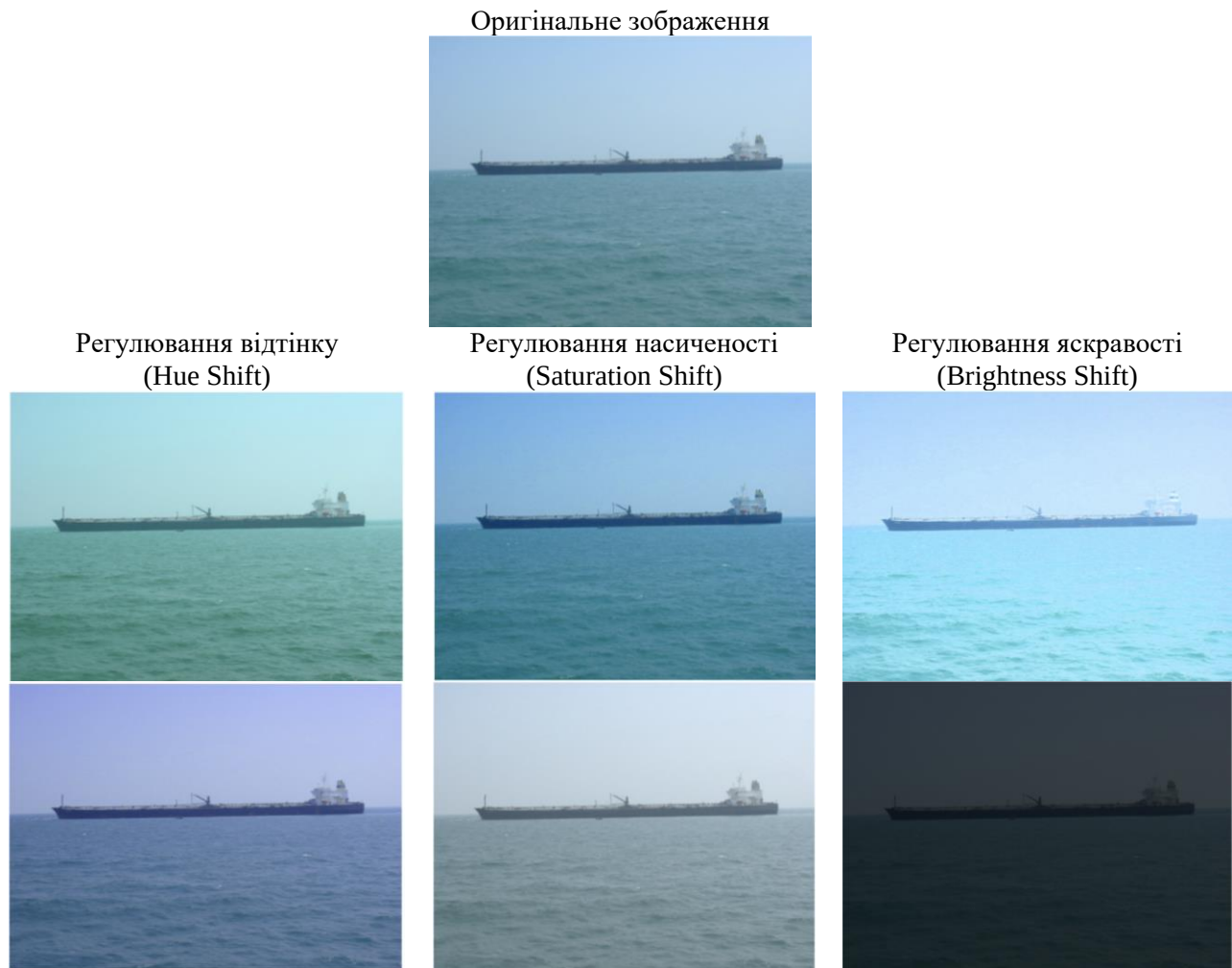


Рисунок 1 – Приклади аугментацій кольорового простору

Геометричні трансформації є ключовим інструментом аугментації даних при навчанні моделей комп'ютерного зору. Серед основних методів – зсув (translate), масштабування (scaling), зсув по осі (shear), перспективне викривлення (perspective) (рис. 2). Для цілей розпізнавання суден, а саме їх ракурсу, зсув (translate) є одним із базових типів геометричних аугментацій, який полягає у випадковому зміщенні зображення вздовж горизонтальної та вертикальної осей. У задачах розпізнавання суден на морській поверхні цей метод дозволяє моделі ефективно адаптуватися до зміни положення об'єкта в кадрі, що є типовим при зйомці з борту судна. Це імітує зміщення судна в полі зору камери внаслідок хитамиці, маневрування або змін кута спостереження, забезпечуючи більш стійке розпізнавання незалежно від положення.

Зміна розміру об'єктів (scaling) на зображенні дозволяє моделі краще розпізнавати судна, незалежно від їхньої відстані до камери. Така трансформація сприяє адаптації до різних масштабів, без втрати ключової інформації. Цей параметр відповідає за діапазон масштабування, регулює ступінь збільшення або зменшення розміру, забезпечуючи різноманітність навчального набору. Це особливо актуально для систем автономного керування, де судна можуть знаходитись на різній відстані від камери, що впливає на їхній візуальний розмір у кадрі.

Зсув по осі (shear) і перспективне викривлення (perspective) імітують реальні деформації форми суден, спричинені кутом огляду або рухом камери. Вони дозволяють навчити модель коректно розпізнавати об'єкти навіть при геометричних викривленнях, що

особливо актуально для динамічних сцен. Цей тип аугментації особливо важливий при знаходженні камери під нахилом, де такі об'єкти, як будівлі чи судна, можуть сприйматися стиснутими або витягнутими. Завдяки перспективній трансформації модель отримує здатність адаптуватися до змін оглядової точки та зберігати точність у реальних умовах [14–15].

Оригінальне зображення



Зсув (Translate)



Масштабування (Scaling)



Зсув по осі (Shear)



Перспективне викривлення (Perspective)



Випадкове стирання (Erasing)



Рисунок 2 – Приклади геометричної та просторово-локальної аугментацій

Такі класичні види аугментації як горизонтальне або вертикальне віддзеркалення (flip) та повороти (rotation) не можна використовувати в моделях розпізнавання ракурсу суден. Попри те, що ці трансформації допомагають моделям краще узагальнювати випадки, будь-які оберти чи віддзеркалення змінюють ніс на корму і навпаки, що плутає в результаті модель.

Одним із поширених методів просторово-локальної аугментації є випадкове стирання (erasing) (рис. 2). Суть методу полягає у випадковому маскуванні прямокутної області зображення, що імітує часткову втрату візуальної інформації внаслідок оклюзії, шуму, бліків чи фонових перешкод, притаманних реальним морським умовам. Застосування даної техніки дозволяє зменшити перенавчання моделі на локальних або найбільш конкретних ознаках (наприклад, форма надбудови чи характер носової частини), змушуючи її враховувати ширший контекст зображення, включно з навколишньою водною поверхнею та формою корпусу.

2. Експериментальна частина

2.1 Оновлена база даних для тренування моделі штучного зору

Для проведення навчання моделі штучного зору, у початковій конфігурації, було сформовано власний набір даних, що включав 925 зображень. Кожен об'єкт на зображеннях класифікувався за одним з восьми ракурсів: forward, port bow, port side, starboard bow, starboard quarter, starboard side, stern, port quarter. Дані були розподілені пропорційно на три групи: тренування (70%), валідація (20%) та тестування (10%). Такі пропорції є стандартними в машинному навчанні для задач комп'ютерного зору, оскільки вони дозволяють виділити переважну частину даних на тренування моделі. Розподіл є приблизним, через те, що навчальні дані зображають реальні сцени на воді і могли містити декілька суден на одному зображенні і розділити їх пропорційно неможливо. Фактичні пропорції в таблиці можуть незначно відрізнятися через округлення або необхідність зберегти цілісність зображень.

Як зазначалося в роботі [12] вищезазначений набір даних потребував доопрацювання. Через велику кількість хибнопозитивних спрацювань, набір даних було доповнено negative examples, які зображають різноманітня навколишнього середовища та портової інфраструктури без суден. Наявність хибненагивних спрацювань сигналізувала про значний дисбаланс класів, що призводило до ігнорування тих класів, які представлені меншою кількістю об'єктів (stern, forward, port quarter).

Оновлена база даних складалася з 1742 зображень (3442 об'єкти), 115 з яких є negative examples, тобто зображеннями без анотацій, які зображають навколишнє середовище. Таблиця 1 відображає зміни між початковим та оновленим набором даних. Варто зазначити, що в таблиці наводиться не кількість зображень, а кількість об'єктів для розпізнавання, тобто зображення може містити декілька суден з різного ракурсу.

Таблиця 1 – Зміна кількості об'єктів розпізнавання для початкового та оновленого набору даних

№ класу	Назва класу	Тренування	Валідація	Тестування	Тренування	Валідація	Тестування	Загальна кількість об'єктів
0	forward	105	26	17	266	70	29	365
1	port bow	164	42	19	291	77	64	432
2	port side	372	112	52	494	166	63	723
3	starboard bow	153	29	17	285	66	37	388
4	starboard quarter	109	33	13	223	65	33	321
5	starboard side	273	94	41	443	84	50	577
6	stern	72	17	6	150	49	22	221
7	port quarter	110	21	16	213	55	32	300
	negative examples				68	33	14	115
		1358	374	181	2433	665	344	3442

Варто відзначити, що попри доповнення бази даних, все одно спостерігався суттєвий дисбаланс кількості прикладів серед окремих класів. Зокрема, найчисельніший клас port side (494) у тренувальній вибірці мав у понад три рази більше прикладів, ніж найменш представлений stern (150). Такий дисбаланс серед класів навчальної вибірки є типовою проблемою для практичних задач комп'ютерного зору. Зазначений дисбаланс спричиняє схильність моделі до перенавчання на більш представлених класах, ігноруючи менш чисельні. Це, у свою чергу, призводить до низької точності (precision) і повноти (recall) на менш чисельних класах, зростання помилок класифікації (confusion), а також погіршення якості загальної поведінки моделі.

Для компенсації виявленого дисбалансу було прийнято рішення про використання ваг класів (class weights) відносно найбільш представленого класу – тобто класу 2, який має найбільше прикладів (494). Зокрема, для кожного класу було обчислено ваговий коефіцієнт, обернено пропорційний до кількості прикладів у відповідному класі, відповідно до формули:

$$w_i = \frac{N_{\max}}{N_i},$$

де w_i – ваги для класу i , $N_{\max}$ – максимальна кількість об'єктів серед усіх класів, N_i – кількість об'єктів для класу i .

У результаті розрахунку були отримані наступні вагові коефіцієнти для класів від 0 до 7 класу, відповідно (табл. 2). Це дозволило дати більшу вагу мало представленим класам, для того, щоб модель не ігнорувала їх у процесі навчання.

Таблиця 2 – Розрахунок вагових коефіцієнтів для кожного класу об'єктів

Клас	Кількість об'єктів для тренування	Розрахунок	Ваги
0	266	494 / 266 = 1.857	1.86
1	291	494 / 291 = 1.697	1.70
2	494	494 / 494 = 1.000	1.00
3	285	494 / 285 = 1.733	1.73
4	223	494 / 223 = 2.215	2.22
5	443	494 / 443 = 1.115	1.11
6	150	494 / 150 = 3.293	3.29
7	213	494 / 213 = 2.319	2.32

2.2 Тренування моделі штучного зору з використанням різних типів аугментації

За основу, для проведення тренування, була обрана модель, отримана в результаті раніше опублікованого дослідження [12]. Отримані результати свідчать про недонавчання моделі, загальна кількість епох тренування – 165 епох, що є недостатнім для ефективної роботи моделі.

Для виявлення найбільш продуктивної стратегії навчання в цій статті були порівняні різні типи аугментації. Навчання моделі при використанні одного типу аугментації відбувалося протягом 60 епох. Фіксовані параметри для всіх експериментів з тренування моделі YOLOv8n наведені в таблиці 3. Ці параметри забезпечують уніфікованість тренувального процесу: $epochs = 60$ визначає кількість ітерацій; $batch = 4$ – розмір пакета даних для оптимізації пам'яті; $imgsz = 702$ – роздільну здатність зображень, адаптовану до морських сцен; $optimizer = Adam$ з $lr0 = 1e-05$, $lrf = 0.2$ та $momentum = 0.937$ – для ефективної мінімізації втрат; $weight_decay = 0.0005$ та $dropout = 0.0$ – для регуляризації без перенавчання; $box = 0.03$, $cls = 0.7$ та $dfl = 1.2$ – ваги функцій втрат для балансу між локалізацією, класифікацією та розподілом; $warmup_epochs = 3.0$ – для поступового збільшення швидкості навчання. Аугментаційні параметри (degrees, flipud, fliplr, mosaic,

mixup, cutmix, copy_paste, auto_augment) встановлено на 0, щоб ізолювати вплив на тестову аугментацію. Слід зазначити, що тип аугментації підбирався відповідно до задачі, а саме навчання моделі штучного зору, яка визначає ракурс судна в морському середовищі.

Таблиця 3 – Спільні вихідні параметри для всіх проведених експериментальних тренувань

<i>Параметр</i>	<i>Значення</i>	<i>Параметр</i>	<i>Значення</i>
epoch	60	df1	1.2
batch	4	pose	12.0
imgsz	702	dropout	0.0
optimizer	Adam	degrees	0.0
seed	42	flipud	0.0
lr0	1.0e-05	fliplr	0.0
lrf	0.2	mosaic	0.0
momentum	0.937	mixup	0.0
weight_decay	0.0005	cutmix	0.0
warmup_epochs	3.0	copy_paste	0.0
box	0.03	auto_augment	null
cls	0.7		

У таблиці 4 наведені параметри аугментації, що змінювалися: для кожного тренування один параметр набував вказаного значення, а інші дорівнювали нулю. Це дозволяло оцінити ізолюваний вплив кольорних (hsv_h, hsv_s, hsv_v), геометричних (translate, scale, shear, perspective) та просторово-локальних (erasing) аугментацій. Ці значення підібрано для симуляції варіативності морського середовища (освітлення, позиції суден, оклюзії) без спотворення ключових ознак, сприяючи підвищенню стійкості моделі до реальних умов зйомки.

Таблиця 4 – Перелік параметрів які змінювалися для кожного тренування

<i>Параметр</i>	<i>Англійський відповідник</i>	<i>Значення</i>
Регулювання відтінку	hsv_h	0.005
Регулювання насиченості	hsv_s	0.5
Регулювання яскравості	hsv_v	0.7
Регулювання зсуву	translate	0.3
Регулювання масштабування	scale	0.5
Зсув по осі	shear	10.0
Перспективне викривлення	perspective	0.05
Випадкове стирання	erasing	0.4

Основні результати та їх обговорення. У дослідженні для реалізації процесу аугментації використано бібліотеку Albumentations, яка є високопродуктивним інструментом для обробки зображень у Python та забезпечує широкий спектр різних трансформацій. Такі аугментації, як розмиття (Blur, MedianBlur), перетворення до відтінків сірого (ToGray), а також локальне вирівнювання гістограми (CLAHE), були застосовані автоматично на завершальному етапі обробки зображень. Ймовірність застосування кожного з цих перетворень становила 1%. Застосування цих трансформацій не було явно задане у конфігураційних параметрах експерименту, оскільки вони активуються автоматично, як частина внутрішньої логіки системи Ultralytics YOLO. За замовчуванням під час тренування активується певний перелік слабких аугментацій з метою підвищення варіативності навчального набору. У цьому дослідженні основний акцент був зроблений на контролі параметрів геометричної та кольорної аугментації, які задавалися вручну, тоді як автоматично

додані перетворення не розглядалися як ключові чинники впливу. Результати дослідження наведені в таблиці 5:

Таблиця 5 – Порівняння різних технік аугментації

Параметр	Початкові параметри моделі	Тренування моделі без додаткової аугментації	Регулювання відтінку	Регулювання насиченості	Регулювання яскравості	Регулювання зсуву	Регулювання масштабування	Регулювання зсуву по осі	Регулювання перспективного викривлення	Регулювання випадкового стирання
epoch	25	19	31	36	12	44	57	52	47	19
train/box_loss	0.0045	0.00192	0.00157	0.00147	0.00241	0.0022	0.00249	0.00309	0.00875	0.00192
train/cls_loss	0.69631	0.53541	0.45063	0.42623	0.77714	0.65095	0.71796	0.68118	5.74881	0.53541
train/dfl_loss	0.93709	0.67876	0.65252	0.64642	0.72514	0.7162	0.73957	0.8066	1.96891	0.67876
precision	0.3964	0.66866	0.65382	0.6307	0.64021	0.70257	0.6617	0.66745	0.28919	0.66866
recall	0.41739	0.52828	0.5677	0.56586	0.57797	0.5388	0.60364	0.54362	0.12445	0.52828
mAP50	0.37135	0.61229	0.61404	0.60985	0.61082	0.63799	0.64649	0.60522	0.1013	0.61229
mAP50-95	0.30399	0.4918	0.49507	0.48686	0.49069	0.51033	0.52609	0.40495	0.03576	0.4918
val/box_loss	0.00585	0.00335	0.00339	0.00344	0.00337	0.00326	0.0031	0.00521	0.00874	0.00335
val/cls_loss	2.1155	1.87315	1.91049	1.90416	1.82186	2.028194	1.65527	2.06874	4.46031	1.87315
val/dfl_loss	1.12038	0.83368	0.87164	0.88611	0.83792	0.83175	0.80121	1.022	1.77471	0.83368

Для оцінки якості моделі були використані метрики mean Average Precision (mAP) при різних порогах перекриття (IoU – Intersection over Union), а саме mAP50 та mAP50-95, що є загальноприйнятими показниками у задачах розпізнавання об'єктів. Порівняння отриманих результатів показало, що базове тренування без додаткових аугментацій забезпечило mAP50 = 0.61229 та mAP50-95 = 0.4918, що стало орієнтиром для оцінки впливу інших методів. Регулювання відтінку, насиченості та яскравості дало близькі значення, де mAP50 перебував у межах 0.60985–0.61404, а mAP50-95 – 0.48686–0.49507, що свідчить про помірний позитивний ефект кольорових трансформацій. Найкращі результати отримано під час застосування геометричних аугментацій: зсув забезпечив mAP50 = 0.63799, тоді як масштабування дало максимальне значення mAP50 = 0.64649 та mAP50-95 = 0.52609. Навпаки, перспективне викривлення призвело до різкого погіршення показників – mAP50 знизився до 0.1013, а precision впав до 0.28919, що свідчить про несумісність або занадто велике значення цього параметру для задач визначення ракурсу судна. Випадкове стирання показало mAP50 на рівні 0.61229, що дозволяє використання даного виду аугментації для розглянутих задач. Таким чином, найбільш продуктивними виявилися трансформації, які імітують природні зміни положення судна в кадрі та кольорові аугментації, тоді як методи, що порушують геометрію сцени, не можуть бути рекомендовані для застосування. Варто зазначити, що при тренуванні моделей штучного зору рекомендовано використовувати комбінації різних видів аугментації, тому найбільшу ефективність варто очікувати шляхом комбінування прийнятних технік аугментації.

Висновки. Аугментація даних є потужним інструментом для покращення навчання моделей комп'ютерного зору та підвищення їх стійкості до варіативності даних. Однак її застосування слід ретельно адаптувати до конкретного завдання. Аналіз різних видів аугментацій, з акцентом на морські умови, показав, що колірні (регулювання відтінку, насиченості, яскравості), геометричні (зсув, масштабування, зсув по осі, перспективне викривлення) та просторово-локальні (випадкове стирання) трансформації впливають на

ефективність моделі по-різному. Результати тренування моделі YOLOv8n без аугментацій (mAP50 = 0.61229, mAP50-95 = 0.4918, precision = 0.66866, recall = 0.52828) порівняно з тренуваннями отриманими шляхом використання аугментації продемонстрували, що колірні трансформації дають помірне покращення: регулювання відтінку підвищує mAP50 до 0.61404 та mAP50-95 до 0.49507; регулювання насиченості – mAP50 до 0.60985 та mAP50-95 до 0.48686; регулювання яскравості – mAP50 до 0.61082 та mAP50-95 до 0.49069. Найкращі результати забезпечують геометричні трансформації, такі як зсув (mAP50 = 0.63799, mAP50-95 = 0.51033, precision = 0.70257, recall = 0.5388) та масштабування (mAP50 = 0.64649, mAP50-95 = 0.52609, precision = 0.6617, recall = 0.60364), що підвищують стійкість моделі до варіативних умов зйомки без спотворення просторових ознак. Навпаки, перспективне викривлення погіршує показники якості (mAP50 = 0.1013, mAP50-95 = 0.03576, precision = 0.28919, recall = 0.12445), а випадкове стирання та зсув по осі дають нейтральний або помірний ефект (mAP50 = 0.61229 та 0.60522 відповідно). Таким чином, найбільш доцільними трансформаціями є колірні аугментації та геометричні (зсув, масштабування), які зберігають фізичний сенс сцени та уникають нереалістичних спотворень.

Перспективи подальших досліджень. Подальші дослідження в цій галузі можуть бути спрямовані на розробку та впровадження комплексних систем автономного судноводіння, де інтеграція моделі комп'ютерного зору для визначення ракурсу суден стане одним з компонентів. Зокрема, планується технічна реалізація системи, що передбачає не лише теоретичне моделювання, але й підбір відповідних, актуальних станом на зараз, складових частин. Крім того, отримані в ході дослідження дані можуть бути активно використані при навчанні моделей штучного зору, призначених для роботи в морському середовищі, що має свої особливості.

У довгостроковій перспективі, такі розробки матимуть потенціал при впровадженні систем штучного зору в морській галузі, сприяючи підвищенню безпеки навігації та ефективності автономних суден.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Jiao L., Zhang F., Liu F., Yang S., Li L. & Feng Z. A survey of deep learning-based object detection. *IEEE Access*, 7, 128837–128868. 2019. <https://doi.org/10.1109/ACCESS.2019.2939201>.
2. Yang D., Solihin M. I., Ardiyanto I., Zhao Y., Li W., Cai B. & Chen C. A streamlined approach for intelligent ship object detection using EL-YOLO algorithm. *Scientific Reports*, 14, 15254. 2024. <https://doi.org/10.1038/s41598-024-64225-y>.
3. Langenkämper D., van Kevelaer R. & Nattkemper T. W. Strategies for Tackling the Class Imbalance Problem in Marine Image Classification. In Z. Zhang, D. Suter, Y. Tian, A. Branzan Albu, N. Sidère, & H. J. Escalante (Eds.), *Pattern Recognition and Information Forensics: ICPR International Workshops and Challenges* (pp. 26–36). 2019. Springer. https://doi.org/10.1007/978-3-030-05792-3_3.
4. McGuinness K. & O'Gara S. Comparing data augmentation strategies for deep image classification. In J. Courtney, C. Deegan, & P. Leamy (Eds.), *IMVIP 2019 Irish Machine Vision & Image Processing Conference proceedings*. Irish Pattern Recognition & Classification Society. <https://doras.dcu.ie/23666/>.
5. Wang H., Deng Z., Jiang C., Ren J. & Zhang S. A review of deep learning techniques for maritime object detection from optical remote sensing images. *Sensors*, 22(14), 5383. 2022. <https://doi.org/10.3390/s22145383>.
6. Li Y., Zhang X., Zhang Y., Wang R. & Li J. YOLO-SEA: An enhanced detection framework for multi-scale maritime targets in complex sea states and adverse weather. *Sensors*, 24(1), 243. 2024. <https://doi.org/10.3390/e27070667>.
7. Wang Z., Wang P., Liu K., Wang P., Fu Y., Lu C. T., Aggarwal C. C., Pei J. & Zhou Y. A comprehensive survey on data augmentation. 2024. arXiv preprint arXiv:2405.09591.

8. Shorten C. & Khoshgoftaar T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60. 2021. <https://doi.org/10.1186/s40537-019-0197-0>.
9. Xu M., Yoon S., Fuentes A. & Park D. S. A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognition*, 137, 109347. 2023. <https://doi.org/10.1016/j.patcog.2023.109347>.
10. Tan M., Langenkämper D. & Nattkemper T. W. The impact of data augmentations on deep learning-based marine object classification in benthic image transects. *Sensors*, 22(14), 5383. 2022. <https://doi.org/10.3390/s22145383>.
11. Cauli N. & Reforgiato Recupero D. Survey on videos data augmentation for deep learning models. *Future Internet*, 14(3), 93. 2022. <https://doi.org/10.3390/fi14030093>.
12. Пашенко О. Л., Піпченко О. Д. Розробка моделі штучного зору на базі алгоритму YOLO для визначення ракурсу судна. *Судноводіння (Shipping & Navigation)*. 2025. Вип. 38.
13. Cauli N. & Reforgiato Recupero D. Survey on Videos Data Augmentation for Deep Learning Models. *Future Internet*, 14(3), 93. 2022. <https://doi.org/10.3390/fi14030093>.
14. Mumuni A. & Mumuni F. Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. 2022. <https://doi.org/10.1016/j.array.2022.100258>.
15. Kumar T., Brennan R., Mileo A. & Bendeache M. Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions. *IEEE Access*, 12, 187536-187571. 2024. <https://doi.org/10.48550/arXiv.2301.02830>.

REFERENCES

1. Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., & Feng, Z. (2019). A survey of deep learning-based object detection. *IEEE Access*, 7, 128837–128868. <https://doi.org/10.1109/ACCESS.2019.2939201>.
2. Yang, D., Solihin, M. I., Ardiyanto, I., Zhao, Y., Li, W., Cai, B., & Chen, C. (2024). A streamlined approach for intelligent ship object detection using EL-YOLO algorithm. *Scientific Reports*, 14, 15254. <https://doi.org/10.1038/s41598-024-64225-y>.
3. Langenkämper, D., van Kevelaer, R., & Nattkemper, T. W. (2019). Strategies for Tackling the Class Imbalance Problem in Marine Image Classification. In Z. Zhang, D. Suter, Y. Tian, A. Branzan Albu, N. Sidère, & H. J. Escalante (Eds.), *Pattern Recognition and Information Forensics: ICPR International Workshops and Challenges* (pp. 26–36). Springer. https://doi.org/10.1007/978-3-030-05792-3_3.
4. McGuinness, K., & O'Gara, S. (2019). Comparing data augmentation strategies for deep image classification. In J. Courtney, C. Deegan, & P. Leamy (Eds.), *IMVIP 2019 Irish Machine Vision & Image Processing Conference proceedings*. Irish Pattern Recognition & Classification Society. <https://doras.dcu.ie/23666/>.
5. Wang, H., Deng, Z., Jiang, C., Ren, J., & Zhang, S. (2022). A review of deep learning techniques for maritime object detection from optical remote sensing images. *Sensors*, 22(14), 5383. <https://doi.org/10.3390/s22145383>.
6. Li, Y., Zhang, X., Zhang, Y., Wang, R., & Li, J. (2024). YOLO-SEA: An enhanced detection framework for multi-scale maritime targets in complex sea states and adverse weather. *Sensors*, 24(1), 243. <https://doi.org/10.3390/s24010243>.
7. Wang, Z., Wang, P., Liu, K., Wang, P., Fu, Y., Lu, C. T., Aggarwal, C. C., Pei, J., & Zhou, Y. (2024). A comprehensive survey on data augmentation. *arXiv preprint arXiv:2405.09591*.
8. Shorten, C., & Khoshgoftaar, T. M. (2021). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>.
9. Xu, M., Yoon, S., Fuentes, A., & Park, D. S. (2023). A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognition*, 137, 109347. <https://doi.org/10.1016/j.patcog.2023.109347>.

10. Tan, M., Langenkämper, D., & Nattkemper, T. W. (2022). The impact of data augmentations on deep learning-based marine object classification in benthic image transects. *Sensors*, 22(14), 5383. <https://doi.org/10.3390/s22145383>.
11. Cauli, N., & Reforgiato Recupero, D. (2022). Survey on videos data augmentation for deep learning models. *Future Internet*, 14(3), 93. <https://doi.org/10.3390/fi14030093>.
12. Pashenko O. L., Pipchenko O. D. (2025). *Rozrobka modeli shtuchnoho zoru na bazi alhorytmu YOLO dlia vyznachennia rakursu sudna*. Sudnovodinnia (Shipping & Navigation). 38.
13. Cauli, N., & Reforgiato Recupero, D. (2022). *Survey on Videos Data Augmentation for Deep Learning Models*. *Future Internet*, 14(3), 93. <https://doi.org/10.3390/fi14030093>.
14. Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. <https://doi.org/10.1016/j.array.2022.100258>.
15. Kumar, T., Brennan, R., Mileo, A., & Bendeache, M. (2024). Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions. *IEEE Access*, 12, 187536-187571. <https://doi.org/10.48550/arXiv.2301.02830>.

Pashenko O. L. IMPACT OF DATA AUGMENTATION ON TRAINING COMPUTER VISION MODEL FOR SHIPS' ASPECT ANGLE DETECTION

Abstract. The study investigates the impact of targeted data augmentation strategies on the performance of a deep learning-based computer vision model designed to determine a ships' aspect angle in maritime scenes. The work addresses the challenge posed by limited and highly imbalanced datasets that are typical for maritime imaging, where variations in illumination, weather, and occlusion significantly affect recognition accuracy. The aim of the research is to evaluate which augmentation techniques contribute to improved robustness of YOLOv8n-based orientation classification without distorting the physical characteristics of marine scenes. The methodology combines controlled experiments with color, geometric and spatially local augmentation, applied to an extended dataset containing both annotated examples and negative samples. The scientific novelty lies in identifying the augmentation techniques that deliver the highest computer vision model performance according to the key metrics mAP50, mAP50-95, precision, and recall, while ensuring that these techniques enhance the performance of a model specifically designed for operation in maritime environments. The results demonstrate that color and geometry-preserving augmentations, such as translation and scaling, yield measurable improvements, while perspective distortions severely degrade performance due to the violation of spatial realism. The obtained results have practical significance for autonomous navigation systems, contributing to improved recognition accuracy under real operational conditions. The conclusions highlight that augmentation strategies must be carefully selected for tasks involving orientation-sensitive objects.

Key words: data augmentation; computer vision; ship aspect angle; YOLOv8n; geometric augmentation; color augmentation; spatial-local augmentation; deep learning; object detection.

© Пашенко О. Л.

Статтю прийнято до редакції 04.12.2025