

DOI: <https://doi.org/10.15276/ict.03.2026.13>

УДК 004.93

Детектування об'єктів без навчання на основі тензорно керованої дифузії з енергетичною регуляризацією

Мороз Дмитро Володимирович¹⁾

ORCID: <https://orcid.org/0009-0002-1082-5033>; mordmytro@stud.op.edu.ua

Любченко Віра Вікторівна¹⁾

ORCID: <https://orcid.org/0000-0002-4611-7832>; lvv@op.edu.ua. Scopus Author ID: 56667638800

¹⁾ Національний університет «Одеська політехніка», пр. Шевченка, 1. Одеса, 65044, Україна

АНОТАЦІЯ

У роботі розв'язано задачу підвищення точності детектування об'єктів без навчання (unsupervised detection) на цифрових зображеннях, отриманих в умовах неоднорідного освітлення та високого рівня шуму, коли застосування нейромережових детекторів є неможливим або недоцільним через відсутність розмічених даних, обмеження обчислювальних ресурсів або вимоги до математичної інтерпретованості результатів. Запропоновано двоетапний метод, який поєднує геометричну фільтрацію на основі тензорно-керованої анізотропної дифузії та подальшу локалізацію контурів із використанням класичних детекторів. Локальна геометрія зображення описується тензором структури, власні вектори якого спрямовують дифузію вздовж контурів об'єктів, що дозволяє «зшивати» розриви ліній без розмиття меж. Для стабілізації еволюційного процесу модель сформульована як задача мінімізації енергетичного функціонала з членом відповідності даним, у межах якої досліджено три альтернативні енергетичні регуляризатори: загальну варіацію, логарифмічний потенціал Перона-Маліка та нелінійний лапласіан. Проведено порівняльні експерименти на серії підводних зображень затонулих об'єктів із різкими перепадами локальної освітленості та адитивним гаусовим шумом. Якість оцінено згідно з динамікою індексу збереження меж. Встановлено, що регуляризація на основі загальної варіації забезпечує найшвидше придушення шуму на початкових ітераціях, однак спричиняє ефект «сходинок» і різку деградацію індексу збереження меж за тривалого еволюційного процесу; моделі Перона-Маліка та обмеження Ліпшиця виходять на стабільне плато. За критерієм точності позиціонування контурів (міра Пратта) модель обмеження Ліпшиця показала найкращий результат серед усіх розглянутих методів, включно з нелокальними середніми та білатеральною фільтрацією. Показано, що індекс збереження меж придатний як індикатор динаміки процесу, але не як метрика підсумкової якості детектування. Доведено, що застосування запропонованої фільтрації перед детектором Кенні дозволяє отримати замкнені зв'язні контури об'єктів без хибних меж у зонах плавного градієнта яскравості. Метод є повністю детермінованим, не потребує навчальних даних і придатний для вбудованих систем комп'ютерного зору.

Ключові слова: детектування об'єктів; анізотропна дифузія; тензор структури; енергетична регуляризація; загальна варіація; модель Перона-Маліка; нескінченний лапласіан; індекс збереження меж; обробка підводних зображень

Для цитування: Мороз Д. В., Любченко В. В. «Детектування об'єктів без навчання на основі тензорно керованої дифузії з енергетичною регуляризацією». *Інформатика. Культура. Техніка.* 2026; Том 3 №1 (3): 155–165. DOI: <https://doi.org/10.15276/ict.03.2026.13>

Актуальність. Детектування об'єктів на цифрових зображеннях є однією з фундаментальних задач комп'ютерного зору, яка знаходить широке застосування в робототехніці, системах автономного керування транспортом, медичній діагностиці, підводній археології та аналізі супутникових знімків. Сучасний етап розвитку цієї галузі характеризується домінуванням методів глибокого навчання: двоетапних та одноетапних згорткових детекторів [1], зокрема сімейства YOLO [2], та архітектур на основі трансформерів [3]. Попри високу точність локалізації об'єктів, такі підходи мають суттєві обмеження: вони критично залежать від наявності великих обсягів розмічених навчальних даних, потребують значних обчислювальних ресурсів (GPU/TPU) і часто працюють як «чорна скринька», що ускладнює інтерпретацію та математичну верифікацію отриманих результатів [4]. Навіть універсальні моделі з можливістю роботи «без прикладів» (zero-shot), такі як Segment Anything [5], потребують масштабного передтренування на мільярдах масок і залишаються ресурсомісткими на етапі виведення. Це зумовлює науковий інтерес до методів детектування без навчання (unsupervised/zero-shot methods), що базуються на апараті математичної фізики, диференціальних рівняннях у частинних похідних (PDE) та варіаційному численні.

Аналіз останніх досліджень і публікацій. Головною перешкодою для класичних детекторів без навчання – детектора Кенні [6], порогового аналізу Оцу [7] чи сегментації

активними контурами – є чутливість до високочастотного шуму, неоднорідності освітлення (градієнтів яскравості) та наявності локальних розривів у контурах об'єктів [8]. Традиційні лінійні методи фільтрації (такі як згладжування Гауса) усувають шум ціною втрати чіткості меж і геометричної деформації об'єктів. Для розв'язання цієї суперечності використовують нелінійну анізотропну дифузію [9] та варіаційні методи мінімізації загальної варіації [10].

Для подолання цих обмежень у сучасній обробці зображень застосовують тензор структури (structure tensor), який дозволяє оцінити локальну геометрію та напрямок граней і побудувати справді анізотропну (керовану напрямком) дифузію [11]. Проте відкритим залишається питання оптимального вибору енергетичного регуляризатора, який би стабілізував процес дифузії в умовах складного освітлення, не допускаючи деградації та згладжування ключових геометричних маркерів об'єктів.

Теоретичні основи нелінійної дифузійної фільтрації закладено в роботі П. Перона та Дж. Маліка [9], де скалярний коефіцієнт провідності залежить від модуля локального градієнта. Варіаційний напрям був сформований моделлю Рудіна-Ошера-Фатемі [10], а узагальнення скалярної провідності до матричної (тензорної) та когерентно-підсилювальна дифузія, керована власними векторами тензора структури, запропоновані Й. Вайкертом [11]. З огляду на появу глибокого навчання цей клас методів не втратив актуальності, а навпаки — переживає оновлення. Варіаційні регуляризатори, що ґрунтуються на принципах загальної варіації, розвинуто в стабільні гібридні схеми розв'язання обернених задач (total deep variation) [12]. У [13] показано глибокі структурні відповідності між чисельними схемами нелінійної дифузії та сучасними нейронними архітектурами (ResNet, U-Net). Це підтверджує роль PDE-методів як математичного каркаса для інтерпретаційного аналізу зображень. Сучасний стан методів виділення країв, включно з оцінкою їхньої чутливості до шуму та неоднорідного освітлення, систематизовано в огляді [8].

Окрему групу становлять методи, що ґрунтуються на варіаційному численні. Поняття абсолютно мінімізуювальних ліпшицевих продовжень і пов'язаний з ними ∞ -лапласіан уведено Г. Аронссоном [14]. В останні роки цей апарат переживає відродження в межах Ліпшицевого навчання (Lipschitz learning) на графах, для якого доведено рівномірну збіжність дискретних ∞ -гармонічних продовжень до неперервного розв'язку [15]. Водночас у сучасних оглядах фільтрації та видалення шуму [16] порівняння проводиться переважно між класичними локальними/нелокальними фільтрами та нейромережевими методами. В той же час, порівняльний аналіз поведінки різних енергетичних регуляризаторів саме в складі тензорно-керованої дифузії висвітлено недостатньо.

Це є особливо важливим для зображень із суттєво неоднорідним освітленням, характерним для підводного відео, де деградація контрасту, розсіювання та поглинання світла є основними факторами зниження якості виявлення об'єктів [17]. Дослідження стану розробки систем підводного детектування [18] та аналіз відповідних датасетів [19] підтверджують, що домінуючі нейромережеві архітектури є дуже чутливими до масштабу розмічених даних і якості зображень після видалення шуму. Тому, існує необхідність дослідити потенціал детермінованих методів без навчання, який досі не розкритий повною мірою.

Метою роботи є розроблення детермінованого методу підвищення точності виділення та локалізації контурів об'єктів на зображеннях із шумом і неоднорідним освітленням на основі тензорно-керованої дифузії з енергетичною регуляризацією та адаптивним критерієм завершення ітераційного процесу.

Для досягнення поставленої мети сформульовано наступні задачі.

1. Розробити математичну модель тензорно-керованої дифузії, інтегровану з механізмом локального аналізу за допомогою тензора структури, та реалізувати чисельні схеми для трьох типів енергетичних регуляризаторів.

2. Провести порівняльні експерименти на тестових зображеннях із різними характеристиками контрасту та освітленості, використовуючи кількісну метрику – індекс збереження меж (Edge Preservation Index, β).

3. Проаналізувати динаміку збіжності моделей та визначити їхню ефективність для подальшого детектування контурів, сформулювати критерій автоматичного зупинення ітераційного процесу фільтрації за динамікою індексу β .

Тензорно керована дифузія з енергетичною регуляризацією. Запропонований метод детектування об'єктів без навчання складається з двох послідовних етапів:

1. Геометрична фільтрація – локальне згладжування текстури й шуму з одночасним підсиленням граней за допомогою нелінійної тензорної дифузії.

2. Сегментація/локалізація контурів – виділення меж об'єктів на очищеному та структурованому зображенні за допомогою класичних порогових операторів.

Схема процесу детектування об'єктів без навчання показана на Рис. 1.



Рис. 1. Схема запропонованого процесу детектування об'єктів без навчання

Джерело: розроблено авторами

Для того, щоб процес дифузії адаптувався до форми об'єктів, необхідно описати локальну поведінку градієнта зображення. Звичайний вектор градієнта $\nabla I = (I_x, I_y)^T$ є нестабільним в умовах шуму. Для стійкого аналізу геометрії використовуємо концепцію тензора структури (Structure Tensor) J_ρ [11], який є симетричною напіввизначеною матрицею 2×2 :

$$J_\rho(x, y) = K_\rho * (\nabla I_\sigma \otimes \nabla I_\sigma) = K_\rho * \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}, \quad (1)$$

де: $I_\sigma = K_\sigma * I_0$ – зображення, згладжене ядром Гауса K_σ з масштабом диференціювання σ для придушення мікрошуму; $*$ – операція згортки; K_ρ – інтегруюче вікно Гауса з масштабом $\rho > \sigma$, яке усереднює напрямки градієнтів у сусідніх пікселях.

Аналіз власних значень ($\mu_1 \geq \mu_2 \geq 0$) та відповідних власних векторів v_1, v_2 тензора J_ρ дозволяє класифікувати локальну структуру зображення.

Замість скалярного коефіцієнта дифузії ми вводимо матрицю дифузії (тензор) $D(J_\rho)$, яка конструюється на основі власних векторів v_1, v_2 тензора структури:

$$D = \lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T. \quad (2)$$

Коефіцієнти дифузії λ_1 та λ_2 визначають швидкість згладжування в напрямках v_1 та v_2 відповідно. Для збереження меж встановимо $\lambda_2 = 1.0$ – дифузія вздовж контуру завжди

залишається максимальною, що дозволяє «зшивати» розриви ліній, $\lambda_1 = g(\mu_1)$ – дифузія впоперек контуру блокується в точках із високим градієнтом.

Функція фільтрації $g(\mu_1)$ задається схемою Вайкерта:

$$g(\mu_1) = \begin{cases} 1, & \mu_1 = 0 \\ 1 - \exp\left(-\frac{C}{(\mu_1 / \alpha)^4}\right), & \mu_1 > 0 \end{cases}, \quad (3)$$

де α – поріг контрастності, що автоматично розраховується як 80%-й квантиль розподілу локальних градієнтів; $C > 0$ – параметр жорсткості фільтрації; μ_1, μ_2 – власні значення тензора.

Для стабілізації процесу фільтрації та щоб запобігти втраті контрасту при великій кількості ітерацій, математична модель формулюється як задача мінімізації енергетичного функціонала:

$$E(I) = \int_{\Omega} \left(\frac{\gamma}{2} (I - I_0)^2 + \Phi(|\nabla I|) \right) dx dy, \quad (4)$$

де перший член утримує розв'язок I поблизу вхідного зображення I_0 , а функція штрафу $\Phi(|\nabla I|)$ задає характер регуляризації. Рівняння Ейлера-Лагранжа для цього функціонала визначає еволюційне рівняння дифузії:

$$\frac{\partial I}{\partial t} = \operatorname{div}(D \cdot \nabla I) - \gamma(I - I_0) + R(I). \quad (5)$$

Будемо досліджувати три класичні сучасні форми члена регуляризації $R(I) = \operatorname{div}\left(\Phi'(|\nabla I|) \frac{\nabla I}{|\nabla I|}\right)$.

Модель повної варіації TV (Total Variation або L_1 -регуляризація) штрафує сумарну довжину ізоліній, дозволяючи різкі перепади яскравості:

$$\Phi(s) = s \Rightarrow R_{TV}(I) = \alpha \cdot \operatorname{div}\left(\frac{\nabla I}{\sqrt{|\nabla I|^2 + \delta^2}}\right), \quad (6)$$

де $\delta = 10^{-5}$ – параметр регуляризації для уникнення сингулярності у точках з нульовим градієнтом.

Модель Перона–Маліка (PM або логарифмічний штраф) вводить критерій, який послаблює штраф для аномально великих градієнтів:

$$\Phi(s) = \frac{K^2}{2} \ln\left(1 + \frac{s^2}{K^2}\right) \Rightarrow R_{PM}(I) = \alpha \cdot \operatorname{div}\left(\frac{1}{1 + \frac{|\nabla I|^2}{K^2}} \nabla I\right). \quad (7)$$

Модель обмеження Ліпшиця (нелінійний ∞ -лапласіан) [14], [15] виконує регуляризацію на основі мінімізації L_{∞} -норми градієнта. Нескінченний лапласіан виникає як граничний оператор для p -лапласіана при $p \rightarrow \infty$. На відміну від моделей із фіксованим p , цей підхід пов'язаний з мінімізацією максимальної величини градієнта та абсолютно мінімізувальними ліпшицевими продовженнями.

Нелінійний ∞ -лапласіан інтерполує проміжки між контурами з мінімально можливою варіацією швидкості зміни яскравості:

$$R_{inf}(I) = \alpha \cdot \frac{(\nabla I)^T \cdot H(I) \cdot \nabla I}{|\nabla I|^2 + \delta^2} = \alpha \cdot \frac{I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}}{I_x^2 + I_y^2 + \delta^2}, \quad (8)$$

де $H(I)$ – матриця Гессе зображення I .

Рівняння (5) розв'язувалося явною скінченно-різницевою схемою з апроксимацією дивергентного члена за стандартним 9-точковим шаблоном. Замість складного неперервного диференціального рівняння простір і час розбиваються на сітку (точки з певним кроком).

Крок $dt = 0.1$ обрано з умови стійкості явної схеми

$$\Delta t \leq \frac{h^2}{C \cdot D_{\max}},$$

де C – константа, яка для 9-точкового шаблону залежить від вагових коефіцієнтів. Обчислювальна складність одного повного проходу становить $O(N \cdot P)$, де N – кількість ітерацій, P – кількість пікселів, що робить метод придатним для реалізації на вбудованих платформах без графічних прискорювачів.

Експериментальні дослідження і аналіз результатів. Для оцінки ефективності запропонованих моделей регуляризації проведено серію чисельних експериментів на сформованій тестовій вибірці з 16 підводних зображень, переважно сцен затонулих суден. Вибірка охоплює фрагменти корабельних конструкцій, вантаж у трюмах, внутрішні приміщення суден, а також сцени з дайверами та морською фауною, що забезпечує різноманіття геометричних структур і умов освітлення. Роздільна здатність зображень становить 600×400 – 600×899 пікселів; сцени характеризуються значними перепадами локальної освітленості. До кожного кадру додавався випадковий адитивний гаусів шум із нульовим середнім та середньоквадратичним відхиленням $\sigma_{\text{noise}} = 0.05$.

Для забезпечення відтворюваності використано тестову вибірку, розміщену у відкритому репозиторії [21].

Якість фільтрації контурів оцінювалася на кожній ітерації за допомогою індексу збереження меж EPI (Edge Preservation Index) β [20], який визначається як нормований коефіцієнт кореляції високочастотних (лапласіанових) складових еталонного та відфільтрованого зображень:

$$\beta = \frac{\sum_{\Omega} (\Delta I_0 - \overline{\Delta I_0})(\Delta I - \overline{\Delta I})}{\sqrt{\sum_{\Omega} (\Delta I_0 - \overline{\Delta I_0})^2 \cdot \sum_{\Omega} (\Delta I - \overline{\Delta I})^2}}, \quad (9)$$

де Δ – оператор Лапласа, а риска зверху позначає середнє значення в області Ω .

На початковому етапі дифузії модель TV демонструє найкращі показники збереження меж (червона крива на Рис. 2).

Це пояснюється властивістю L_1 -мінімізації миттєво згладжувати дрібномасштабний високочастотний шум, зберігаючи при цьому загальну контрастність великих перепадів яскравості. Але після 15-ї ітерації спостерігається різкий спад індексу β . Фізичною причиною цього є ефект «сходинок». За умов перепадів освітленості, де градієнт яскравості змінюється плавно, TV-регуляризатор починає квантувати півтони, створюючи штучні плоскі зони з неіснуючими різкими межами.

Нелінійні методи демонструють якісно іншу динаміку – повільне, але стабільне зростання індексу β з виходом на стійке плато після 30-ї ітерації. Така поведінка свідчить про досягнення стаціонарного стану, коли процеси згладжування фону та утримання меж врівноважують один одного. Модель Перона–Маліка ефективно фіксує виражені контури

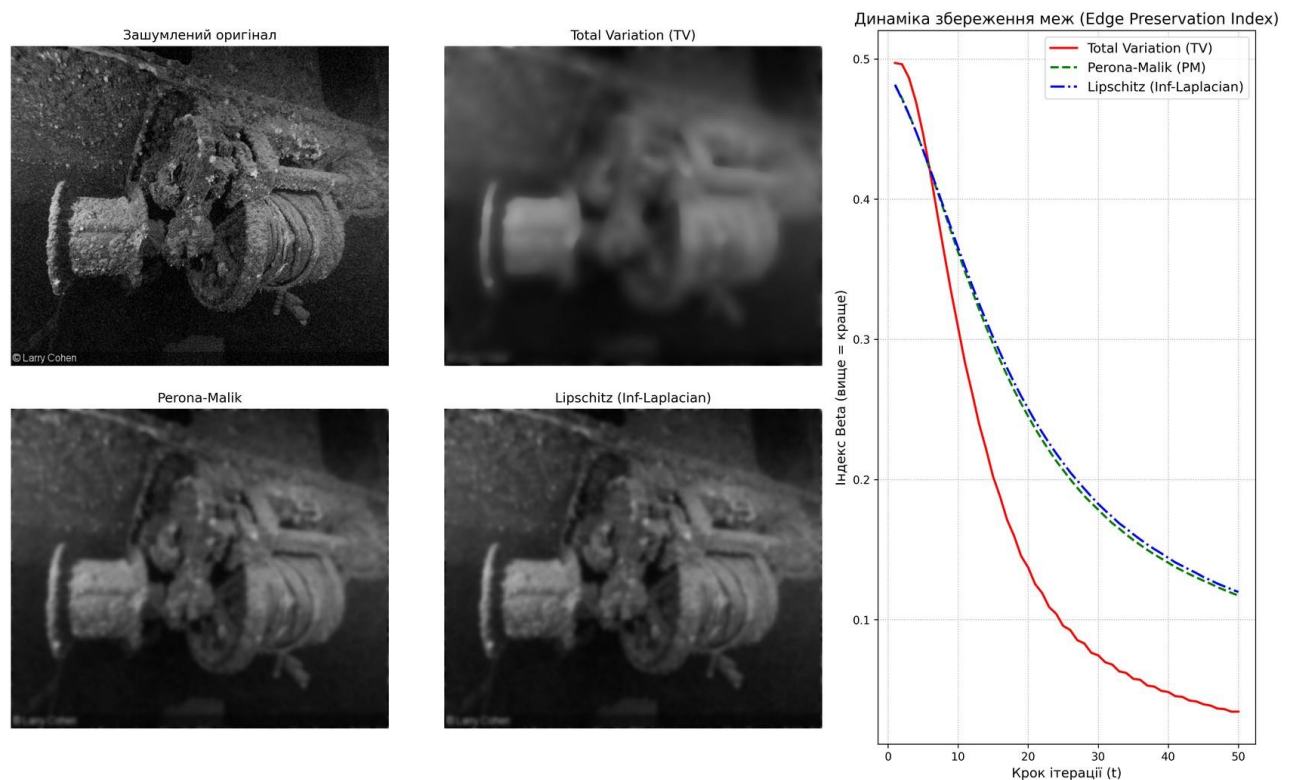


Рис. 2. Результати фільтрації зображення «лебідка» (високий локальний контраст) та динаміка індексу β

Джерело: отримано авторами на основі відкритого набору даних [21]

об'єкта завдяки гасінню дифузії поперек меж. Модель обмеження Ліпшиця показала найкращий збалансований результат на плато. Завдяки мінімізації максимального градієнта, цей метод не створює сходинок на плавних перепадах яскравості, а натомість делікатно "натягує" мінімально гладку поверхню між ключовими контурами [14], [15]. Власні вектори тензора структури спрямовують ∞ -Лапласіан вздовж ліній, що дозволяє буквально "зшивати" розриви контурів навіть у затінених і слабо освітлених ділянках зображення.

Критерій автоматичного зупинення ітераційного процесу. Виявлена закономірність еволюції індексу β , як короточасного піку з подальшим спадом для TV-моделі та монотонний вихід на плато для моделей Перона–Маліка й обмеження Ліпшиця, дозволяє сформулювати єдине правило зупинення, придатне для всіх трьох регуляризаторів. Позначимо $\beta_{\max}(t) = \max_{s \leq t} \beta(s)$ як найкраще значення індексу, досягнуте до кроку t .

Тоді номер ітерації зупинення визначається як

$$t^* = \min \left\{ t \geq t_{\min} : \beta_{\max}(t) - \beta(t) > \delta \cdot \beta_{\max}(t) \vee 0 \leq |\beta(t) - \beta(t - \tau)| < \varepsilon_{\beta} \tau \right\}, \quad (10)$$

де перша умова фіксує відносне падіння індексу від досягнутого піку на частку δ , а друга – вихід на плато. Друга умова є односторонньою: вимагається, щоб приріст β за останні τ ітерацій був малим за модулем і невід'ємним за знаком. Ця асиметрія принципова: симетрична умова $|\beta(t) - \beta(t - \tau)| < \varepsilon_{\beta} \tau$ спрацьовує також на висхідній гілці кривої, де приріст малий, і призводить до передчасного зупинення процесу до досягнення оптимуму. За результатами експериментів рекомендовано $\delta = 0.05$; $\varepsilon_{\beta} = 2 \cdot 10^{-3}$ (як частка поточного значення β); $\tau = 5$; $t_{\min} = 5$. Значення β обчислюється на кожній ітерації (складність $O(P)$), а перевірка умов (10) зводиться до порівняння скалярів. Суттєвий розкид t^* між кадрами (від 5 до 24 для моделі Перона–Маліка) відображає відмінності в рівні контрасту та освітленості тестових сцен.

Для оцінки практичної придатності результатів до відфільтрованих зображень у точці зупинення t^* застосовано класичний детектор контурів Кенні [6] з фіксованими порогами (40/120).

Результати експериментів, наведені в Таблиці 1, показують, що:

1) після TV-регуляризації контурна карта є найбільшньою ($FOM = 0.311$) – разом із шумом усунуто значну частину справжніх меж, а в зонах плавного градієнта освітленості ефект «сходинок» породжує хибні прямолінійні контури;

2) після моделі Перона–Маліка межі контрастних об'єктів виділені чітко ($FOM = 0.684$), проте контури дрібних елементів та об'єктів у тіні місцями розірвані на окремі сегменти;

3) після моделі обмеження Ліпшиця досягнуто найвищої точності позиціонування контурів ($FOM = 0.758$), зокрема в зонах низького контрасту, де інші моделі втрачають межі. Кількість зв'язних компонентів при цьому залишається значною (1038), що вказує на збереження дрібних деталей контурної карти разом із залишковою фрагментацією; повне усунення фрагментації потребує додаткових морфологічних операцій.

Таблиця 1. Зведені результати експериментів по тестових зображеннях (середнє $\pm$ СКВ)

Метод	PSNR, дБ	SSIM	FOM	Компонент	β	t^*	Час, мс
Без фільтрації	26.51 $\pm$ 0.38	0.547 $\pm$ 0.117	0.527 $\pm$ 0.220	4382 $\pm$ 1698	0.462 $\pm$ 0.118	–	0
Гауса ($\sigma=1.5$)	28.13 $\pm$ 2.51	0.785 $\pm$ 0.085	0.310 $\pm$ 0.100	340 $\pm$ 298	0.326 $\pm$ 0.034	–	9 $\pm$ 4
Білатеральний	29.25 $\pm$ 3.43	0.762 $\pm$ 0.099	0.247 $\pm$ 0.096	318 $\pm$ 279	0.698 $\pm$ 0.106	–	29 $\pm$ 13
Non-local means (NLM)	32.02 $\pm$ 2.73	0.853 $\pm$ 0.039	0.613 $\pm$ 0.125	1144 $\pm$ 1072	0.734 $\pm$ 0.075	–	521 $\pm$ 229
TV (на t^*)	28.23 $\pm$ 2.78	0.641 $\pm$ 0.072	0.311 $\pm$ 0.104	335 $\pm$ 277	0.419 $\pm$ 0.072	5 $\pm$ 0	7507 $\pm$ 2524
Перона–Маліка (на t^*)	30.27 $\pm$ 1.95	0.773 $\pm$ 0.041	0.684 $\pm$ 0.153	941 $\pm$ 822	0.509 $\pm$ 0.067	14 $\pm$ 7	7506 $\pm$ 2758
Ліпшиця (на t^*)	30.02 $\pm$ 1.75	0.751 $\pm$ 0.033	0.758 $\pm$ 0.138	1038 $\pm$ 825	0.478 $\pm$ 0.078	10 $\pm$ 4	7585 $\pm$ 2772

Джерело: розроблено авторами

Показники β , PSNR і SSIM розраховуються відносно чистого еталона, FOM (Pratt's Figure of Merit) і кількість зв'язаних компонент – по карті Кенні з фіксованими порогами (40/120).

Результати кількісного порівняння дозволяють розмежувати три різні аспекти якості, які в попередніх дослідженнях часто ототожнюють.

Точність позиціонування контурів. За критерієм FOM модель обмеження Ліпшиця показала найкращий результат серед усіх розглянутих методів – 0.758 проти 0.684 у моделі Перона–Маліка, 0.613 у нелокальних середніх та 0.311 у TV-моделі. Перевага є статистично значущою за парним критерієм Вілкоксона (Таблиця 2): над моделлю Перона–Маліка – на 14 кадрах із 16 ($p=0.004$), над рештою методів – на 15-16 кадрах із 16 ($p<0.001$).

Фрагментованість контурних карт. Кількість зв'язних компонентів у моделі Ліпшиця перевищує показники TV та білатерального фільтра. Проте це порівняння некоректне без урахування FOM: TV-модель і білатеральний фільтр втрачають справжні межі об'єктів разом із шумом. На рівні FOM модель Ліпшиця має кількість компонент, порівнянну з нелокальними середніми (1038 проти 1144), забезпечуючи при цьому суттєво вищу точність позиціонування. Отже, показник фрагментованості слід інтерпретувати лише в парі з FOM.

Індекс β як індикатор динаміки. Прямі значення β у табл. 2 демонструють обмеження цієї метрики. Причина полягає в самій побудові індексу (9): він є кореляцією лапласіанових

Таблиця 2. Парне порівняння моделей обмеження Ліпшиця з альтернативними методами за критерієм FOM ($n = 16$, парний критерій Вілкоксона)

Порівняння методів	Середня різниця FOM	Кадрів на користь Ліпшиця	p
Ліпшиць vs PM	+0.074	14 з 16	0.004
Ліпшиць vs TV	+0.447	16 з 16	< 0.001
Ліпшиць vs NLM	+0.145	15 з 16	< 0.001
Ліпшиць vs Білатеральний	+0.511	16 з 16	< 0.001
Ліпшиць vs Гаус	+0.448	16 з 16	< 0.001

Джерело: розроблено авторами

складових і тому штрафувє будь-яке придушення високочастотного вмісту незалежно від того, чи покращується при цьому локалізація меж. Тому в даній роботі β використовується не як показник якості детектування, а виключно як *індикатор динаміки* еволюційного процесу, а підсумкова якість оцінюється за FOM.

Обчислювальна вартість. Дифузійні моделі потребують 7.5 ± 2.7 с на кадр, що на порядок перевищує витрати нелокальних середніх (0.52 с) і на два порядки перевищує витрати білатерального фільтра (0.03 с). Це є платою за детермінованість та ітераційний характер методу. Тому метод орієнтований на застосування, де інтерпретованість і точність меж важливіші за швидкодію, а також на платформи без графічних прискорювачів.

Оцінка впливу власного шуму вихідних зображень. Вхідні кадри, отримані камерами в умовах слабкого освітлення, містять залишковий власний шум. Його рівень оцінено за методом Іммеркаера, що ґрунтується на згортці зображення з лапласіановою маскою другого порядку: середньоквадратичне відхилення власного шуму становить $\sigma_{own} = 0.002 \dots 0.021$ при медіані 0.008, тобто в середньому у 6.4 раза менше за доданий синтетичний шум $\sigma_{noise} = 0.05$. Отже, у сумарному шумовому навантаженні домінує контрольована синтетична компонента, а еталонне зображення (кадр до додавання шуму) є достатньо чистим для коректного порівняння методів. Водночас слід враховувати, що абсолютні значення PSNR і SSIM дещо занижені для всіх методів однаково, оскільки еталон не є ідеально безшумним.

Висновки. У роботі розв'язано задачу підвищення точності детектування об'єктів без навчання в умовах складного й неоднорідного освітлення та високого рівня завад. На основі проведених теоретичних та експериментальних досліджень сформульовано такі висновки.

1. Запропоновано інтегровану математичну модель нелінійної фільтрації, яка поєднує аналітичну потужність тензора структури та анізотропну дифузію. Проведено порівняльний аналіз чисельних схем для трьох енергетичних регуляризаторів: варіаційного, логарифмічного та геометричного.

2. Доведено перевагу регуляризації на основі обмеження Ліпшиця для задач виділення меж без навчання за критерієм точності позиціонування контурів. Встановлено, що індекс збереження меж не є придатною метрикою підсумкової якості детектування, оскільки штрафувє будь-яке придушення високочастотного вмісту незалежно від коректності локалізації меж. Запропоновано критерій автоматичного зупинення ітераційного процесу.

3. Застосування запропонованої тензорно-керованої дифузії з регуляризацією Ліпшиця перед класичними детекторами дозволяє суттєво підвищити точність локалізації меж у зонах плавного перепаду яскравості без використання нейронних мереж та етапу навчання.

Перспективи подальших досліджень у цьому напрямку полягають у розробці локально-адаптивного кроку за часом $dt(x, y)$ для прискорення обчислень у фонових областях та інтеграції розробленого методу в мобільні системи комп'ютерного зору з низькою потужністю.

Конфлікт інтересів: Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого, авторського чи будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті

Фінансування: Дослідження проводилося без фінансової підтримки

Доступність даних: Усі дані доступні в цифровій або графічній формі в основному тексті рукопису

Використання засобів штучного інтелекту: Під час підготовки рукопису використовувався інструмент Claude Opus 4.8 штучного інтелекту (ШІ) для граматичної та стилістичної перевірки тексту. Усі фрагменти, опрацьовані за його допомогою, були перевірені й відредаговані автором. ШІ не використовувався для генерації тексту, формул і таблиць, формування наукових положень, отримання результатів або формулювання висновків. Наукові положення, результати та висновки є власним інтелектуальним внеском авторів

СПИСОК ЛІТЕРАТУРИ

1. Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M. & Lee, B. “A survey of modern deep learning based object detection models”. *Digital Signal Processing*. 2022; 126: 103514. DOI: <https://doi.org/10.1016/j.dsp.2022.103514>.
2. Terven, J., Córdova-Esparza, D.-M. & Romero-González, J.-A. “A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS”. *Machine Learning and Knowledge Extraction*. 2023; 5 (4): 1680–1716. DOI: <https://doi.org/10.3390/make5040083>.
3. Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S. & Shah, M. “Transformers in vision: A survey”. *ACM Computing Surveys*. 2022; 54 (10s): 200:1–200:41. DOI: <https://doi.org/10.1145/3505244>.
4. Zou, Z., Chen, K., Shi, Z., Guo, Y. & Ye, J. “Object detection in 20 years: A survey”. *Proceedings of the IEEE*. 2023; 111 (3): 257–276, <https://scopus.com/pages/publications/2-s2.0-85149303350>. DOI: <https://doi.org/10.1109/JPROC.2023.3238524>.
5. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P. & Girshick, R. “Segment Anything”. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris: IEEE. 2023. p. 4015–4026. DOI: <https://doi.org/10.1109/ICCV51070.2023.00371>.
6. Canny, J. “A computational approach to edge detection”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1986; PAMI-8 (6): 679–698. DOI: <https://doi.org/10.1109/TPAMI.1986.4767851>.
7. Otsu, N. “A threshold selection method from gray-level histograms”. *IEEE Transactions on Systems, Man, and Cybernetics*. 1979; 9 (1): 62–66. DOI: <https://doi.org/10.1109/TSMC.1979.4310076>.
8. Jing, J., Liu, S., Wang, G., Zhang, W. & Sun, C. “Recent advances on image edge detection: A comprehensive review”. *Neurocomputing*. 2022; 503: 259–271, <https://scopus.com/pages/publications/2-s2.0-85133611311>. DOI: <https://doi.org/10.1016/j.neucom.2022.06.083>.
9. Perona, P. & Malik, J. “Scale-space and edge detection using anisotropic diffusion”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990; 12 (7): 629–639. DOI: <https://doi.org/10.1109/34.56205>.
10. Rudin, L. I., Osher, S. & Fatemi, E. “Nonlinear total variation based noise removal algorithms”. *Physica D: Nonlinear Phenomena*. 1992; 60 (1–4): 259–268. DOI: [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
11. Weickert, J. “Coherence-enhancing diffusion filtering”. *International Journal of Computer Vision*. 1999; 31 (2–3): 111–127. DOI: <https://doi.org/10.1023/A:1008009714131>.
12. Kobler, E., Effland, A., Kunisch, K. & Pock, T. “Total deep variation: A stable regularization method for inverse problems”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2022; 44 (12): 9163–9180, <https://scopus.com/pages/publications/2-s2.0-85142104595>. DOI: <https://doi.org/10.1109/TPAMI.2021.3124086>.

13. Alt, T., Schrader, K., Augustin, M., Peter, P. & Weickert, J. “Connections between numerical algorithms for PDEs and neural networks”. *Journal of Mathematical Imaging and Vision*. 2023; 65 (1): 185–208. DOI: <https://doi.org/10.1007/s10851-022-01106-x>.
14. Aronsson, G. “Extension of functions satisfying Lipschitz conditions”. *Arkiv för Matematik*. 1967; 6 (6): 551–561. DOI: <https://doi.org/10.1007/BF02591928>.
15. Bungert, L., Calder, J. & Roith, T. “Uniform convergence rates for Lipschitz learning on graphs”. *IMA Journal of Numerical Analysis*. 2023; 43 (4): 2445–2495. DOI: <https://doi.org/10.1093/imanum/drac048>.
16. Elad, M., Kawar, B. & Vaksman, G. “Image denoising: The deep learning revolution and beyond – a survey paper”. *SIAM Journal on Imaging Sciences*. 2023; 16 (3): 1594–1654. DOI: <https://doi.org/10.1137/23M1545859>.
17. Raveendran, S., Patil, M. D. & Birajdar, G. K. “Underwater image enhancement: a comprehensive review, recent trends, challenges and applications”. *Artificial Intelligence Review*. 2021; 54 (7): 5413–5467, <https://scopus.com/pages/publications/2-s2.0-85107324533>. DOI: <https://doi.org/10.1007/s10462-021-10025-z>.
18. Xu, S., Zhang, M., Song, W., Mei, H., He, Q. & Liotta, A. “A systematic review and analysis of deep learning-based underwater object detection”. *Neurocomputing*. 2023; 527: 204–232, <https://scopus.com/pages/publications/2-s2.0-85149303350>. DOI: <https://doi.org/10.1016/j.neucom.2023.01.056>.
19. Fu, C., Liu, R., Fan, X., Chen, P., Fu, H., Yuan, W., Zhu, M. & Luo, Z. “Rethinking general underwater object detection: Datasets, challenges, and solutions”. *Neurocomputing*. 2023; 517: 243–256. DOI: <https://doi.org/10.1016/j.neucom.2022.10.039>.
20. Sattar, F., Floreby, L., Salomonsson, G. & Löfström, B. “Image enhancement based on a nonlinear multiscale method”. *IEEE Transactions on Image Processing*. 1997; 6 (6): 888–895. DOI: <https://doi.org/10.1109/83.585239>.
21. Moroz, D. V. & Liubchenko, V. V. “Experimental subset of 16 underwater images used in this study”. *Google Drive*. 2026. – Available at: <https://drive.google.com/drive/folders/1GFsQBfaSvmm0dNkTo7NIdeqffa5GPiYZ>. – [Accessed: 25.04.2026].

Отримано 21.07.2026

Отримано після перегляду 26.08.2026

Прийнято 03.09.2026

DOI: <https://doi.org/10.15276/ict.03.2026.13>

УДК 004.93

Untrained object detection based on tensor-guided diffusion with energy regularization

Dmytro V. Moroz¹⁾

ORCID: <https://orcid.org/0009-0002-1082-5033>; mordmytro@stud.op.edu.ua

Vira V. Liubchenko¹⁾

ORCID: <https://orcid.org/0000-0002-4611-7832>; lvv@op.edu.ua. Scopus Author ID: 56667638800

¹⁾ Odesa Polytechnic National University, 1, Shevchenko Av. Odesa, 65044, Ukraine

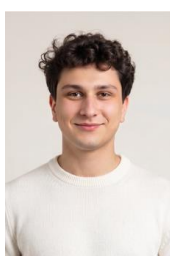
ABSTRACT

This paper addresses the problem of improving the accuracy of unsupervised object detection (unsupervised detection) in digital images acquired under conditions of uneven lighting and high noise levels, where the use of neural network detectors is impossible or impractical due to the lack of labeled data, computational resource constraints, or requirements for the mathematical interpretability of results. A two-stage method is proposed that combines geometric filtering based on tensor-controlled anisotropic diffusion with subsequent contour localization using classical detectors. The local geometry of the image is described by a structure tensor, whose eigenvectors direct diffusion along the contours of objects, allowing line breaks to be “stitched” together without

blurring the boundaries. To stabilize the evolutionary process, the model is formulated as a problem of minimizing an energy functional with a data-conformity term, within which three alternative energy regularizers are investigated: total variation, the Perona–Malik logarithmic potential, and the nonlinear Laplacian. Comparative experiments were conducted on a series of underwater images of sunken objects with sharp variations in local illumination and additive Gaussian noise. Quality was assessed based on the dynamics of the boundary preservation index. It was found that regularization based on total variation provides the fastest noise suppression during the initial iterations but causes a “stair-step” effect and a sharp degradation of the edge preservation index during prolonged evolution; the Perona–Malik and Lipschitz constraint models reach a stable plateau. Based on the criterion of contour positioning accuracy (Pratt’s measure), the Lipschitz constraint model showed the best result among all the methods considered, including nonlocal averages and bilateral filtering. It is shown that the boundary preservation index is suitable as an indicator of the process dynamics, but not as a metric of the final detection quality. It has been proven that applying the proposed filtering before the Kenny detector allows for obtaining closed, connected object contours without false boundaries in areas of smooth brightness gradients. The method is fully deterministic, requires no training data, and is suitable for embedded computer vision systems.

Keywords: object detection; anisotropic diffusion; structure tensor; energy regularization; total variation; Perona–Malik model; infinity Laplacian; edge preservation index; underwater image processing

ІНФОРМАЦІЯ ПРО АВТОРІВ



Мороз Дмитро Володимирович - аспірант каф. Інженерії програмного забезпечення. Національний університет «Одеська політехніка», пр. Шевченка, 1. Одеса, 65044, Україна

Галузь досліджень: Розпізнавання образів, інтелектуальний аналіз складних даних

Dmytro V. Moroz - Postgraduate Student, Department of Software Engineering. Odesa Polytechnic National University, 1, Shevchenko Av. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0009-0002-1082-5033>; mordmytro@stud.op.edu.ua

Research field: Pattern recognition, complex data analysis



Любченко Віра Вікторівна - доктор техніч. наук, професор каф. Інженерії програмного забезпечення, Національний університет «Одеська політехніка», пр. Шевченка, 1. Одеса, 65044, Україна

Галузь досліджень: Наука про дані, програмна інженерія, управління проєктами

Vira V. Liubchenko - Doctor of Engineering Sciences, Professor, Department of Software Engineering Odesa Polytechnic National University, 1, Shevchenko Av. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0000-0002-4611-7832>; lvv@op.edu.ua. Scopus Author ID: 56667638800

Research field: Data Science, Software Engineering, Project Management