

UDK 656.61.052:621.43

## ПІДХОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ФАКТОРУ СУДНОВОДІЯ В ЕРГАТИЧНИХ СИСТЕМАХ УПРАВЛІННЯ РУХОМ СУДНА

*Кочубей П.О., Носов П.С.*

*Херсонська державна морська академія  
(Україна)*

**Вступ.** Сучасний розвиток морського транспорту стикається з викликами та завданнями, серед яких особливе місце займає питання раціонального управління судном. Актуальність цього дослідження полягає у визначенні та аналізі ролі людського фактора у процесах прийняття рішень на судах, особливо під час виконання складних маневрів у непередбачуваних ситуаціях та під час проходження навчально-тренажерної підготовки [1].

Враховуючи глобальні виклики, такі як забезпечення екологічної безпеки та підвищення енергоефективності морського транспорту, потрібні дослідження що мають як теоретичне значення, так і практичну цінність для галузі морських перевезень задля покращення безпеки мореплавства, зниження ризиків аварій та сприянню ефективної експлуатації судових технічних систем [2].

**Основний матеріал дослідження.** Розвиток обчислювальних потужностей за останнє десятиліття став каталізатором значних досягнень у галузі комп'ютерного бачення, відкривши нові горизонти для інновацій та автоматизації. Цей прогрес забезпечує трансформацію багатьох сфер, зокрема автомобільної промисловості, медицини, а також низки інших галузей, де технології комп'ютерного зору сприяли підвищенню ефективності, точності та безпеки процесів. У контексті морського транспорту ці досягнення створюють нові можливості для інтеграції передових технологій, що підкреслює важливість подальших досліджень у цьому напрямку для забезпечення його сталого розвитку та підвищення безпеки судноплавства.

Інтеграція комп'ютерного бачення в судноводіння відкриває нові перспективи для більш глибокого аналізу взаємодії штурмана та технічного обладнання містку. Це дозволяє отримати нові дані, які раніше залишалися поза увагою або потребувало використання численних непростих у впровадженні сенсорів, таких як електрокардіографи (ECG) або окуляри для відстеження руху очей, що ускладнювало широкомасштабне застосування таких методів. Використання комп'ютерного бачення усуває ці обмеження, роблячи процес більш практичним і доступним.

Особлива увага приділяється вивченню фізіологічного та психологічного стану морських штурманів, що безпосередньо впливає на їх здатність приймати оптимальні рішення в реальному часі [3]. Перевантаження когнітивної або фізіологічної складової може призводити до помилок у судноводінні, що створює загрозу для безпеки судна, екіпажу та навколишнього середовища. Застосування RGB-камер для моніторингу та аналізу діяльності морських штурманів дозволить ідентифікувати ключові показники, пов'язані з когнітивним станом, поведінковими особливостями та загальною продуктивністю.

Ідентифікація індивідуальних особливостей штурманів, особливо під час навчально-практичної підготовки зокрема тренажерної, їх реакції на зовнішні подразники та впливу стресових факторів є складним і багатоаспектним завданням, яке потребує системного та структурованого підходу [4]. Для забезпечення ефективності дослідження головним завданням роботи є розробка інтелектуальної системи для визначення поточного рівня когнітивного навантаження штурмана на основі аналізу рисунку обличчя, включно з виявленням мікровиразів, що свідчать про емоційний стан, стрес чи втому.

Наразі експеримент проводиться в навчальних лабораторіях Херсонської державної морської академії з використанням навігаційного симулятора *Navi-Trainer Professional*

5000, розробленого компанією Wärtsilä. Цей симулятор повністю відповідає вимогам Міжнародної морської організації (ІМО) та Конвенції STCW (Міжнародні стандарти підготовки, дипломування та несення вахти для моряків). Симулятор оснащений трьома моніторами, які відображають елементи навігаційних систем, включаючи основні параметри управління судном, ситуаційну обізнаність та інформацію про судноплавну обстановку (Рис.1).



Рисунок 1 — Навігаційний тренажер як експериментальна база

Для збору даних використовується вебкамера Logitech HD Pro Webcam C920, яка розташована на рівні очей штурмана, як показано на зображенні. Камера закріплена так, щоб забезпечити оптимальний кут огляду, дозволяючи фіксувати рухи голови, обличчя та очей під час виконання завдань. Ця позиція камери гарантує, що всі ключові мікровирази та фізіологічні показники будуть захоплені з максимальною точністю.

Для обробки зібраних відеоданих використовується мова програмування Python, яка є стандартом у наукових дослідженнях завдяки своїй універсальності та широкій системі бібліотек. Однією з ключових бібліотек, що застосовується в цьому проєкті, є OpenCV (Open Source Computer Vision Library), яка надає потужні інструменти для роботи із зображеннями та відео. OpenCV використовується для попередньої обробки відео, такої як нормалізація кольорів, видалення шумів та підготовка кадрів до подальшого аналізу. Завдяки інтеграції Python та OpenCV, обробка даних відбувається ефективно та в реальному часі, що забезпечує високу точність аналізу.

**Опис експерименту.** Експеримент спрямовано на аналіз поведінки штурмана в умовах різного рівня когнітивного навантаження шляхом виконання двох типів завдань: легкого та складного. Легке завдання імітує рутинну роботу в спокійних умовах, коли погода без хвиль і вітру, навігаційна обстановка характеризується мінімальним рухом, а курсанти підтримують заданий курс і швидкість без необхідності виконання складних маневрів. Цей сценарій створює умови з мінімальним когнітивним навантаженням, що дозволяє оцінити базовий стан штурмана. Складні завдання, навпаки, представляють ситуації, що вимагають більшої уваги та швидкого прийняття рішень. Курсанти стикаються з несприятливими погодними умовами, такими як поривчастий вітер та помірні хвилі, а також інтенсивним рухом суден у вузькій протоці, на прикладі Босфор. Вони повинні реагувати на аварійні сигнали, виконувати маневри для уникнення зіткнень та вирішувати проблеми з навігаційним обладнанням. Такі сценарії моделюють високий рівень когнітивного навантаження, характерного для реальних умов судноводіння.

**Аналіз та обробка відеоряду.** Аналіз відеоданих починається з обробки кожного

кадру для визначення зон інтересу, таких зон обличчя як очі, рот і положення голови. Для цього використовується MobileNetV2 [5], згортна нейронна мережа імплементована в бібліотеці MediaPipe [6] яка забезпечує швидке і точне виявлення ділянок обличчя у кадрах відео.

На основі ключових точок обличчя, визначених за допомогою MediaPipe, розраховуються три основні метрики: EAR (Eye Aspect Ratio), LAR (Lip Aspect Ratio) і Head Yaw [7–9]. EAR використовується для моніторингу частоти моргань, що є важливим показником втоми, а LAR аналізує зміни положення губ, які сигналізують про стрес або емоційний стан. Метрика Head Yaw оцінює поворот голови штурмана вліво чи вправо, що дозволяє зрозуміти, на який монітор він спрямовує свою увагу під час виконання завдань. Ці метрики слугують додатковими входними даними для згорткової нейронної мережі (CNN), яка також фіксує просторові ознаки з кожного кадру відеоряду.

Було використано комбінований підхід: застосування CNN та мануально пороховані метрики, виходячи із гіпотези, що такий підхід дасть об'єктивні результати, бо саме ці метрики є критичними у визначенні когнітивного стану судноводія у режимі реального часу. Схему обробки даних представлено на Рисунку 2.



Рисунок 2 — Схема аналізу ідентифікаторів по даним відеоряду

### Мануальні метрики

EAR (Eye Aspect Ratio) використовується для аналізу частоти моргань, яка є індикатором втоми. EAR розраховується за формулою (1):

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \cdot \|p_1 - p_4\|}, \quad (1)$$

де  $p_1$ - $p_6$  — ключові точки на вколо очей (Рис. 3).

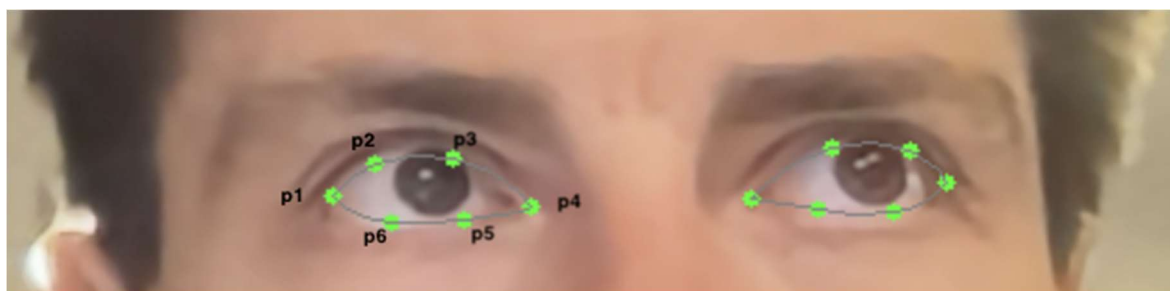


Рисунок 3 — Автоматизована ідентифікація контуру очей судноводія

LAR (Lip Aspect Ratio) дозволяє оцінювати емоційний стан через зміну положення губ (Рис. 4) (2):

$$LAR = \frac{\|q_2 - q_8\|}{q_1 - q_4}, \quad (2)$$

$q_1, q_2, q_4, q_8$  — ключові точки навколо губ.

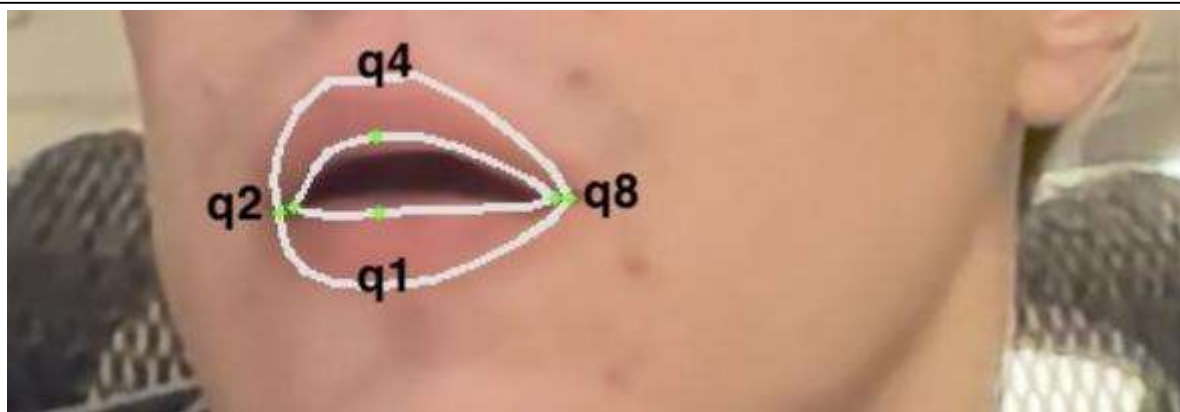


Рисунок 4 — Автоматизована ідентифікація контуру рота

Head Yaw (Кут повороту голови)  $\theta$  визначається за допомогою ключових точок обличчя (очі, кінчик носа) (Рис. 5), як (3):

$$\theta = \arctan 2(\text{nose\_vector}_x, \text{horizontal\_vector}_x) \cdot \frac{180}{\pi}. \quad (3)$$

Горизонтальний вектор (4):

$$\text{horizontal\_vector} = \begin{bmatrix} \text{righth\_eye}_x - \text{lefteye}_x \\ \text{righth\_eye}_y - \text{lefteye}_y \end{bmatrix}. \quad (4)$$

Вектор носу (5):

$$\text{nose\_vector} = \begin{bmatrix} \text{nose\_tip}_x - \text{eye\_midpoint}_x \\ \text{nose\_tip}_y - \text{eye\_midpoint}_y \end{bmatrix}. \quad (5)$$

Середня точка між очима (6):

$$\text{eye\_midpoint} = \frac{\text{left\_eye} + \text{right\_eye}}{2}. \quad (6)$$

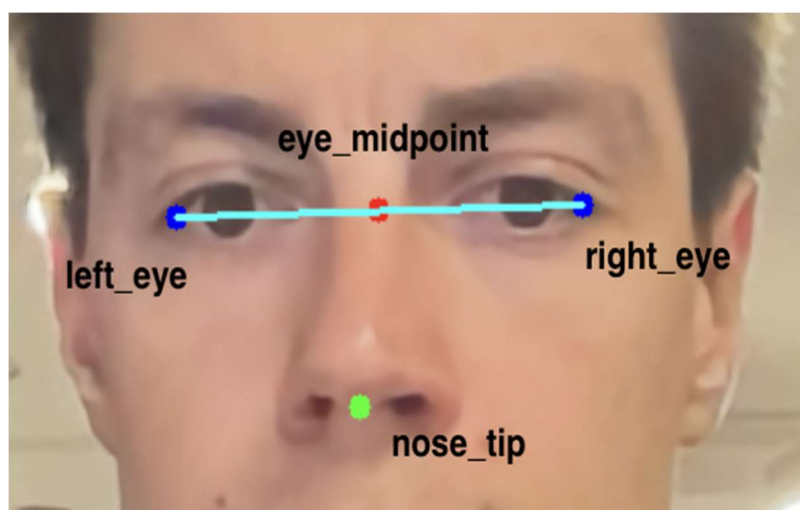


Рисунок 5 — Процес CNN-фіксування просторових ознак

У рамках дослідження, CNN використовуються для визначення просторових ознак із відеокадрів. Основною операцією в CNN є згортка, яка дозволяє виділяти ключові аспекти зображення, такі як межі, текстури та форми [10]. Схематична робота згортної мережі позначена на Рисунку 6.

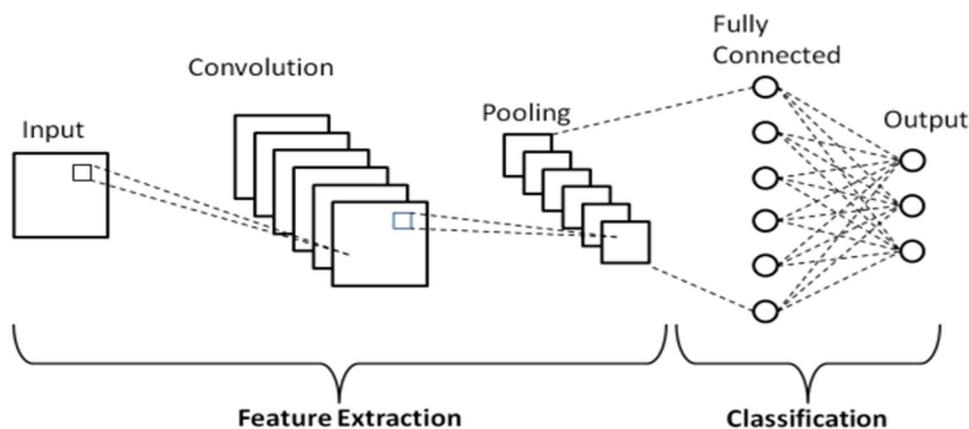


Рисунок 6 — Схема роботи CNN

Згортка (Convolution) обчислює вихідний тензор ознак  $F$  за допомогою ядра згортки (Kernel) [11] та вхідного зображення  $I$  (Input) (Рис. 7) (7):

$$F(i, j) = \sum_{m=1}^M \sum_{n=1}^N K(m, n) \cdot I(i + m, j + n), \quad (7)$$

де  $I(i + m, j + n)$  - пікселі сусідньої області вхідного зображення,  
 $K(m, n)$  — ядро згортки,  
 $F(i, j)$  — значення виходу (Output) для точки  $(i, j)$ .

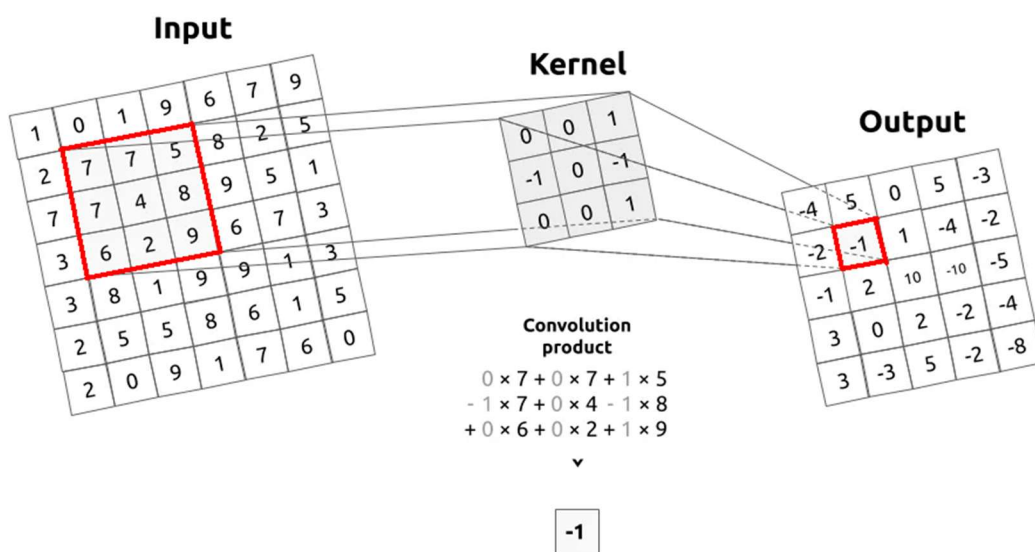


Рисунок 7 — Принцип роботи згортної мережі

Після згортки значення пропускаються через функцію активації ReLU, яка робить модель нелінійною. Це необхідно, щоб мережа могла моделювати складні залежності. В нашому випадку це складні вирази обличчя які можуть свідчити про втому або напругу: підняті щоки, опущені кінчики губ, тощо (8).

$$\text{ReLU}(x) = \max(0, x). \quad (8)$$

Наступний крок це Пулінг (Pooling Layer), що використовується для зменшення розмірності ознак тензора, зберігаючи найбільш релевантну інформацію. Найпоширеніший варіант — max-pooling:

Max-pooling вибирає найбільше значення в кожній області, зменшуючи розмір даних, підвищуючи стійкість до зсувів і скорочуючи обчислювальні витрати (9).

$$P(i, j) = \max_{m, n \in k} F(i + m, j + n). \quad (9)$$

Останнім кроком є Шари Об'єднання (Fully Connected Layers). Після кількох згорткових і пулінгових шарів результат (тензор ознак) перетворюється у вектор фіксованої довжини. Цей вектор передається через Fully Connected Layers, які виконують фінальну класифікацію (10):

$$y = W \cdot x + b, \quad (10)$$

де  $x$  — вектор ознак,  
 $W$  — матриця ваг,  
 $b$  — вектор зміщення,  
 $y$  — вихідний вектор.

Навчання відбувається шляхом оптимізації параметрів (ваг і зсувів) у всіх шарах за допомогою алгоритмів, таких як градієнтний спуск, а функція витрат визначає наскільки точним є передбачення моделі (11):

$$L = - \sum_{i=1}^N y_i \log(\hat{y}_i), \quad (11)$$

де  $N$  — кількість класів,  
 $y_i$  — справжня мітка,  
 $\hat{y}_i$  — ймовірність, передбачена моделлю для класу  $i$ .

Поєднання CNN для аналізу відеокadrів із вручну розрахованими метриками дозволяє створити багатовимірний підхід до аналізу когнітивного стану штурмана. Модель класифікує завдання, яке виконує судноводій, на легке або складне, спираючись на комбінацію просторових ознак із відео та ключових метрик, таких як EAR, LAR і Head Yaw.

Такий метод забезпечує точну оцінку стану штурмана в реальному часі, дозволяючи ідентифікувати рівень складності виконуваних завдань. Це створює потенціал для адаптивного управління робочим навантаженням штурмана, сприяючи підвищенню безпеки та ефективності судноводіння. А також створення алгоритмів прогнозування рівня втоми та готовності до виконання професійних обов'язків, що базуються на аналізі фізіологічних і поведінкових показників, отриманих у реальному часі.

**Висновок.** Отже, в рамках дослідження було запропоновано систему моніторингу виконання навігаційних завдань у реальному часі шляхом аналізу положення тіла оператора-судноводія, жестів та рухів у просторі для оцінки рівня залучення та його когнітивного навантаження.

У подальшій роботі планується розробка мультиагентної ергатичної системи, здатної одночасно ідентифікувати просторові фізіологічні прояви кількох штурманів одночасно, оцінювати їх когнітивний стан та передбачати рівень втоми, що є особливо важливим у контексті командної роботи навігаційної вахти.

Ці етапи можуть стати першою ланкою до розробки універсальної моделі ефективних дій оператора-судноводія, яка сприятиме ідентифікації оптимальних рішень і запобіганню негативним проявам фактору людини.

Запропоновані підходи можуть покращити процес тренажерної підготовки курсантів і делегатів судноплавних компаній.

## REFERENCES

1. Nosov P. S., Palamarchuk I.V., Safonov M.S., Novikov V.I. Modeling the manifestations of the human factor of the Maritime crew // Dnipropetrovsk National University of Railway Transport named after Academician V. Lazaryan (Dnipro) № 5 (77). — 2018. Pages 82-92. <https://doi.org/10.15802/stp2018/147937>.
2. Zinchenko, S., Kobets, V., Tovstokoryi, O., Nosov, P., & Popovych, I. (2023). Intelligent System Control of the Vessel Executive Devices Redundant Structure. In CEUR Workshop

Proceedings (Vol. 3403, Paper 44, pp. 582-594). CEUR-WS.org.

3. Zinchenko S.M., Nosov P.S., Popovych I.S. Control redundancy as a quantitative measure of maneuverability // Науковий вісник Херсонської державної морської академії: науковий журнал. – Херсон: Херсонська державна морська академія, 2021. № 3(21). С. 23-35.

4. Носов П.С., Тонконогий В.М., Яковенко О.Є. Застосування адаптивних функцій для впливу на модель знань студента // Тр. Одес. политехн. ун-та. Одеса: ОНПУ. Вып.1(25). 2006.— С. 118–122.

5. Sandler, Mark, et al. «MobilenetV2: Inverted residuals and linear bottlenecks». Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. <https://arxiv.org/abs/1801.04381>.

6. Lugaresi, Camillo, et al. «Mediapipe: A framework for building perception pipelines» preprint arXiv:1906.08172 (2019). <https://arxiv.org/abs/1906.08172>.

7. Saeed, Anwar & Al-Hamadi, Ayoub & Ghoneim, Ahmed. (2015) Head Pose Estimation on Top of Haar-Like Face Detection: A Study Using the Kinect Sensor. Sensors (Basel, Switzerland). 15. 20945-20966. DOI: 10.3390/s150920945 <https://www.researchgate.net/publication/281587953>

8. Mehta, Sukrit, et al. (2019) «Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio». Proceedings of international conference on sustainable computing in science, technology and management (SUSCOM), Amity University Rajasthan, Jaipur-India.

9. Sri Mounika, T. V. N. S. R., et al. (2022) «Driver drowsiness detection using eye aspect ratio (EAR), mouth aspect ratio (MAR), and driver distraction using head pose estimation». ICT Systems and Sustainability: Proceedings of ICT4SD 2021, Volume 1. Springer Singapore.

10. Pollák, Martin & Gabštur, Peter & Kočíško, Marek. (2024). Prediction of Errors in the Field of Additive Manufacturing Technology. TEM Journal. 849-855. DOI:10.18421/TEM132-01.

11. Canny Edge Detection from Scratch with Pytorch in Python URL: [https://www.codespeedy.com/canny-edge-detection-from-scratch-with-pytorch-in-python/#google\\_vignette](https://www.codespeedy.com/canny-edge-detection-from-scratch-with-pytorch-in-python/#google_vignette)