

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-25>

УДК 004.932.2

Седінкін Олександр Анатолійович<sup>1</sup>, студент

<https://orcid.org/0009-0000-6500-4887>

Деркач Марина Володимирівна<sup>1,2</sup>, к.т.н., доцент

<https://orcid.org/0000-0001-8977-2776>

Скарга-Бандурова Інна Сергіївна<sup>2</sup>, д.т.н., професор

<https://orcid.org/0000-0003-3458-8730>

Матюк Данило Сергійович<sup>1</sup>, студент

<https://orcid.org/0000-0001-5851-8433>

<sup>1</sup>Східноукраїнський національний університет імені Володимира Даля, м. Київ, Україна

<sup>2</sup>Тернопільський національний технічний університет імені Івана Пулюя, м. Тернопіль, Україна

## СИСТЕМА ДЛЯ ВІДСТЕЖЕННЯ РУХУ ОЧЕЙ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Седінкін О.А., Деркач М.В., Скарга-Бандурова І.С., Матюк Д.С. Система для відстеження руху очей на основі машинного навчання. У статті розглянута актуальна технологія визначення точки фіксації погляду, що надає можливість інтерактивної взаємодії у різних сферах, зокрема у вдосконаленні комп'ютерних інтерфейсів. Розроблена система для відстеження руху очей ґрунтується на методи HOG та лінійному алгоритмі машинного навчання SVM й реалізована на мові програмування Python з підключенням бібліотек: OpenCV, Dlib, numpy. Представлена система виділяє зіниці на зображенні обличчя людини, визначає координати зіниць лівого і правого ока, визначає напрямок погляду людини.

**Ключові слова:** система, розпізнавання обличчя, машинне навчання, технологія, зіниці, око.

Sedinkin O., Derkach M., Skarga-Bandurova I., Matiuk D. Eye tracking system based on machine learning. The article discusses current technology for determining the gaze fixation point, which provides possibility of interactive interaction in various areas in computer interfaces. The developed eye tracking system is based on the HOG method and linear SVM and is implemented in Python with libraries: OpenCV, Dlib, numpy. The presented system selects the pupils in the image of a person's face, determines the coordinates of the pupils of the left and right eye, and determines the direction of the person's gaze.

**Keywords:** system, face recognition, machine learning, technology, pupils, eye.

**Постановка проблеми.** Вивчення траєкторії руху очного яблука стає ключем до зрозуміння процесів розпізнавання зорових образів та когнітивних механізмів, що охоплюють всю глибину внутрішнього світу людського розуму [1]. Крім того, відстеження руху очей відкриває нові перспективи для розробки новаторських методів взаємодії між людиною та технічними системами, надаючи можливість впливу на них через найпростіший інструмент — погляд. Саме тому, актуальність технологій визначення точки фіксації погляду стає основною науково-технічною дилемою, яка стимулює багаторічні дослідження та інноваційні розробки. Ця технологія надає можливість інтерактивної взаємодії у різних сферах, зокрема в ігровій індустрії, розробці інтерфейсів користувача [2], віртуальній та доповненій реальності, аналізі користувацьких пристроїв та додатків, а також в галузі кібербезпеки та ідентифікації [3]. Технологія відстеження руху очей на основі машинного навчання відіграє ключову роль у вдосконаленні інтерфейсів та взаємодії з комп'ютером.

**Аналіз останніх досліджень і публікацій.** Останнім часом застосування технології відстеження руху очей значно поширюється на різноманітні області та відкриває нові перспективи для розвитку цифрового світу. Відомі підходи для вимірювання та аналізу рухів очей включають електроокулографію (ЕОГ), електромагнітні методи, методи відеоокулографії, тощо.

ЕОГ є біомедичними сигналами, зазвичай які отримують від ЕЕГ-електродів і використовують для виявлення руху очей і моргання [4]. Так, наприклад, дослідження [5] зосереджено на зменшенні кількості електродів за допомогою методів вибору ознак без будь-яких наслідків для статистичних показників ефективності ЕЕГ-пристроїв для прогнозування стану очей, досягнута точність склала 96,3% за допомогою класифікатора KStar (K\*). А у дослідженні [6] було навчено дев'ять алгоритмів машинного навчання, таких як штучні нейронні мережі, лінійний дискримінантний аналіз, дерево рішень, метод k-найближчих сусідів, наївний баєсів класифікатор і метод опорних векторів, тощо, на наборі даних, створеному шляхом виділення ознак ЕЕГ-сигналів.

В той час, як відеоокулографія заснована на записі положення очей за допомогою відеокамер і може виконуватися кількома техніками: відеоокулографія межі лімба/райдужки і склери, зінична відеоокулографія й відеоокулографія відображень зіниці та рогівки. Для прикладу у статті [7]

автори досліджують як можна передбачити стан очей (відкриті чи закриті) шляхом вимірювання мозкових хвиль за допомогою камери під час вимірювання ЕЕГ, протестовано 42 різні алгоритми машинного навчання. А у дослідженні [8] запропонували нову бінокулярну систему виявлення зіниці та погляду, яка використовує дистанційну камеру з роздільною здатністю Full HD і використовує алгоритми обробки зображень LabVIEW.

**Метою статті** є розроблення системи відстеження руху очей на основі машинного навчання для визначення положення зіниць та напрямку погляду людини на зображенні.

**Вирішення проблеми.** У центрі сприйняття світу людською свідомістю знаходиться заплутана мережа зорової системи, що складається з інтегрованого комплексу ока, нервових шляхів та зорового аналізатора головного мозку. Існує 52 поля, що визначив німецький невропатолог Корбініан Бродман завдяки поділу кори великих півкуль на цитоархітектонічні зони, згруповані в 11 гістологічних регіонів, так як ці ділянки мають різну структуру та виконують різні функції. Результат зорової активності проходить через бічні колінчасті ядра таламусу, а потім досягає зорової кори, тобто частини кори головного мозку, що відіграє важливу роль в обробці візуальної інформації. Обидві півкулі мозку містять зорову кору. Зорова кора лівої півкулі отримує сигнали від правого зорового поля, а зорова кора в правій півкулі отримує сигнал від лівого поля зору. Зорова кора розташована в потиличній частці, в задній частині черепа. Частина зорової кори, яка отримує сенсорні входи від таламуса, називається первинною зоровою корою, або першою візуальною зоною (відповідає полю Бродмана 17). Частина екстрастріарної кори, яка складається з зорових областей другої, третьої, четвертої і п'ятої, відповідає полям Бродмана 18 і 19. Стійка активність у зоровій корі передається з потилично-скроневої ділянки задньої частини мозку назад у префронтальну кору. І навіть після зникнення візуальної активності, мозкова активність значно знижується, але зберігається. Так працює мозок людини, що дає можливість миттєво розпізнати будь-який об'єкт.

Що ж стосується технічного зору, то комп'ютери на сьогодні не здатні до узагальнень високого рівня і для задач розпізнавання [9], в тому числі задачі відстеження руху очей, потрібно побудувати систему, що складається з кількох етапів розпізнавання окремих об'єктів і послідовно поєднує декілька алгоритмів машинного навчання.

Звісно існують сервіси розпізнавання й порівняння обличчя на зображенні/відео, такі як, наприклад, Microsoft Azure Face API, Google Cloud Vision API та Amazon Rekognition Image, але це комерційні проєкти.

Одним з головних критеріїв роботи подібних сервісів, є час, за який система розпізнає обличчя на зображенні/відео, особливо для інтерфейсів, де в режимі реального часу будується людино-машинна взаємодія. Результат порівняння за часом опрацювання фото згаданих вище сервісів і розробленої системи показано в таблиці 1.

Таблиця 1. Порівняння часу, витраченого для пошуку обличчя

Назва сервісу	Face ID	Microsoft Azure Face API	Google Cloud Vision API	Amazon Recognition Image	Розроблена система
Час опрацювання, с	0,77	3,08	6	1,77	1,67

Отримавши прийнятний результат розпізнавання обличчя на першому етапі, для наступного етапу реалізації розробки побудовано систему для визначення положення зіниць та напрямку погляду людини.

Тепер безпосередньо представимо розроблену систему для відстеження руху очей, що реалізовано на мові програмування Python з підключенням бібліотек: OpenCV, Dlib, numpy, й ґрунтується на методі HOG (Histogram of Oriented Gradients, гістограма спрямованих градієнтів) та лінійному алгоритмі машинного навчання SVM (Support Vector Machines, метод опорних векторів):

1. Основна ідея HOG полягає в наступних кроках:

Крок 1. Поділити зображення на маленькі з'єднані клітинки.

Крок 2. Обчислити гістограму для кожної клітинки.

Крок 3. Об'єднати всі гістограми разом, щоб сформувати одну унікальну для кожного обличчя гістограму.

2. SVM є алгоритмом класифікації, що визначає гіперплощину, яка робить поділ між класами. Функцію прийняття рішень SVM можна виразити як:

$$f(x) = \text{sign} \left( \sum_{i=1}^N a_i y_i K(x_i, x) + b \right),$$

де  $N$  — розмір навчальних даних,  $K$  — функція ядра, яка вимірює подібність між  $x_i$  (опорний вектор) і  $x$  (значення ознак),  $a_i$  — множник Лагранжа,  $y_i$  представляє клас належності кожного даного ( $\pm 1$ ), а  $b$  — числова константа.

Лінійний SVM: функція ядра виражається як внутрішній добуток опорного вектора та значень ознак:

$$K(x_i, x) = x_i^T \cdot x.$$

Фактично, HoG застосовується для обчислення дескрипторів, що використовуються для навчання лінійної SVM, яка виявляє обличчя.

Алгоритм функціонування системи зображено на рис. 1. Спочатку створюється відеопотік, у циклі зчитуються кадри відеопотоку один за одним, після чого кожен кадр перетворюється в чорно-біле зображення. Потім відбувається пошук та виявлення обличчя на кадрі завдяки методу HOG та лінійному SVM.

Наступний крок - аналіз кожного кадру для розпізнавання обличчя та ініціалізація очей на зображенні. Спочатку знаходимо обличчя, після чого визначаємо ключові точки (орієнтири) обличчя за допомогою навченої моделі. Ці орієнтири включають точки навколо очей, що використовується для визначення положень очей та зіниць, й містять порогове значення для бінаризації зображення очей для визначення райдужної оболонки ока. Під час знаходження орієнтирів також калібруємо алгоритм, для покращення результатів. Далі обробляємо основний кадр з виділеними зіницями і визначаємо напрям погляду користувача.

Для перевірки положення зіниць використовується метод, завдяки якому можна отримати цілі значення координат  $x$  та  $y$ , що вказують на координати центра для зіниць лівого і правого ока:

Лістинг 1. Метод визначення положення зіниць

```
def pupils_located(self):  
    try:  
        int(self.eye_left.pupil.x)  
        int(self.eye_left.pupil.y)  
        int(self.eye_right.pupil.x)  
        int(self.eye_right.pupil.y)  
        return True  
    except Exception:  
        return False
```

Потім візуалізуємо кожен зіницю на кадрі, щоб перевірити, що зіниці були знайдені та правильно проаналізовані. Система призначена для розпізнавання руху очей однієї людини в кадрі.

**Результати досліджень.** Розроблена система протестована за допомогою камери мобільного телефону Redmi 9 Pro, що розташована перед людиною у фіксованому положенні навпроти її голови.

Характеристики камери: матриця 64 мегапікселі; роздільна здатність: 9248 x 6920; сенсор: Samsung Bright S5KGW1; діафрагма:  $f/ 1.89$ ; розмір пікселю: 0.80  $\mu\text{m}$ ; розмір датчику: 1/1.72"; зчитування кадру: до 60 кадрів при записуванні відео 1080 роздільності.



Рис. 1. Алгоритм системи для відстеження руху очей на основі машинного навчання  
Результати тестування для визначення координат центрів зіниць та напрямку погляду  
вліво/вправо представлені на рисунку 2.

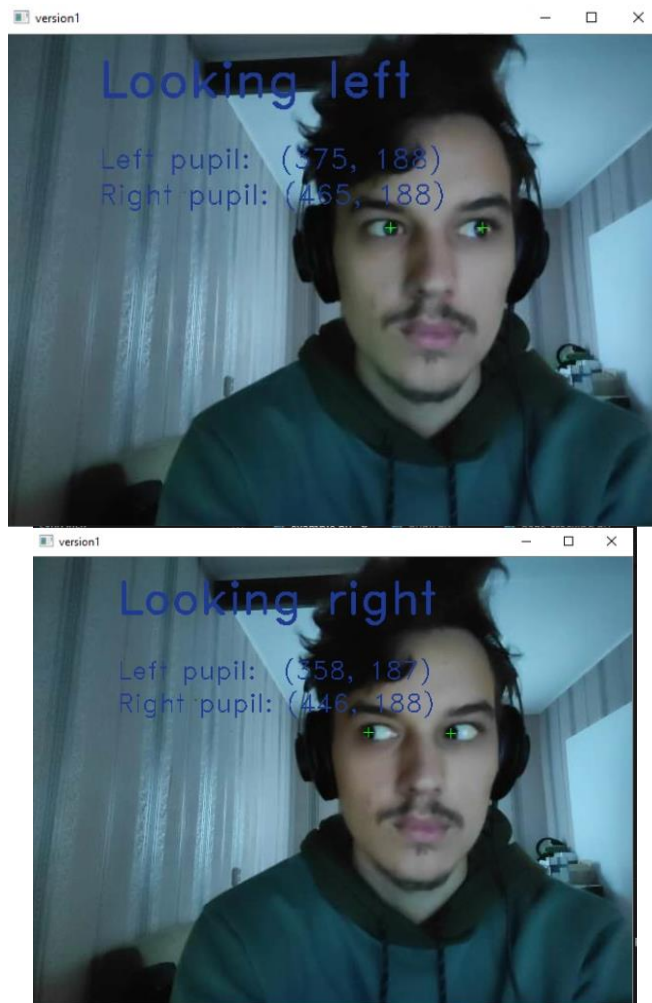


Рис. 2 – Визначення напрямку погляду

Проведено тестування системи в стані заплющених очей та в умовах слабкої освітленості приміщення (рис. 3).

Результати тестування підтверджують, що система для відстеження руху очей:

- виділяє зіниці на зображенні обличчя людини;
- визначає координати зіниць лівого і правого ока;
- визначає напрямок погляду людини.

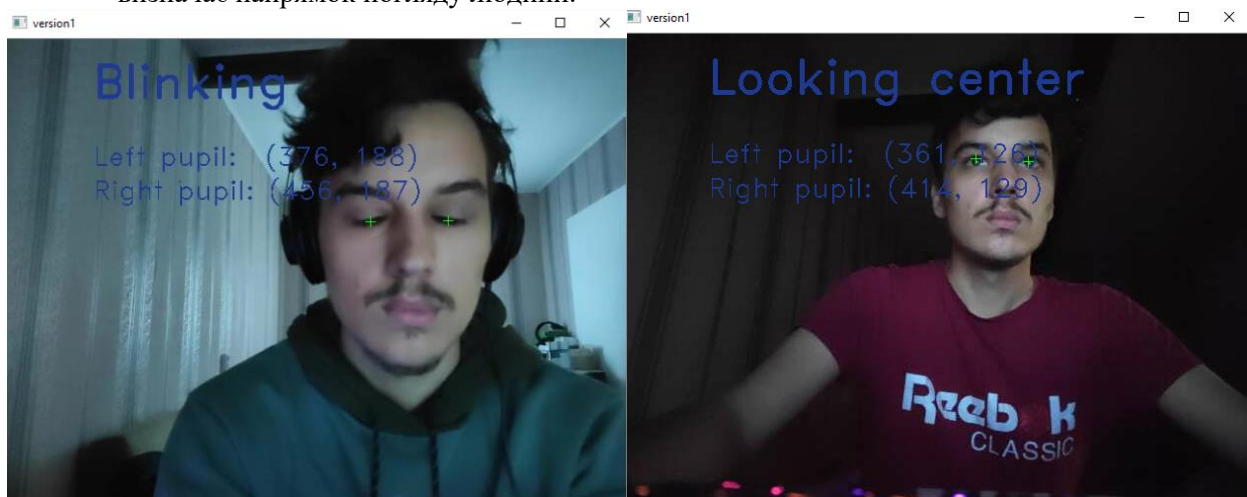


Рис. 3 – Результати тестування системи

Крім того, визначено, що система розпізнає стан заплющених очей, та однаково добре працює при помірній та слабкій освітленості приміщення. При цьому виявлено недоліки системи: при надмірній освітленості приміщення система може не виявити очі, а також при швидкому русі голови система не встигає обробляти результати і перестає визначати напрям погляду та координати зіниць. В подальшому планується вирішення цих проблем та покращення точності розпізнавання шляхом збільшення частоти кадрів камери та збільшення набору даних для навчання моделі.

**Висновки.** Розроблено систему для відстеження руху очей на основі машинного навчання. Система здатна виділяти зіниці на зображенні обличчя людини, визначати координати зіниць лівого і правого ока, визначати напрямок погляду людини. Для виявлених недоліків запропоновано методи їх вирішення. З перспективних напрямів розвитку, можна очікувати застосування систем відстеження руху очей для особистісної ідентифікації, оскільки характер рухів очей має індивідуальну природу, схожу на почерк, але неможливий для точної імітації. Ще один значний напрям застосування подібних систем полягає у комп'ютерних інтерфейсах, наприклад, можливість керувати курсором при роботі з ПК без допомоги рук, що допоможе людям з обмеженими можливостями після травм чи захворювань продовжувати жити повноцінним життям.

#### Список бібліографічного опису

1. Mai, T.D.T., Phung, TN. (2023). Evaluating the Performance of Some Deep Learning Model for the Problem of Emotion Recognition Based on EEG Signal. In: Nghia, P.T., Thai, V.D., Thuy, N.T., Son, L.H., Huynh, VN. (eds) *Advances in Information and Communication Technology. ICTA 2023. Lecture Notes in Networks and Systems*, vol 847. Springer, Cham. [https://doi.org/10.1007/978-3-031-49529-8\\_19](https://doi.org/10.1007/978-3-031-49529-8_19).
2. М. В. Деркач, Д. С. Матюк. Дослідження технології ЕЕГ на прикладі реалізованого нейрокомп'ютерного інтерфейсу. Сучасні інформаційні технології та системи в управлінні [Електронний ресурс]: Зб. матеріалів IV Міжнар. наук.-практ. конф. молодих вчених, аспірантів і студентів; 20–21 квітня 2023 р. Київ: КНЕУ, 2023. С. 81 – 83.
3. Jayarathne I, Cohen M, Amarakeerthi S. (2020) Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio. *PLoS ONE* 15(9): e0238872. <https://doi.org/10.1371/journal.pone.0238872>.
4. Д. С. Матюк, М. В. Деркач. Оцінка спектральної щільності потужності ЕЕГ сигналу. Актуальні задачі сучасних технологій : зб. тез доповідей XII міжнар. наук.-практ. конф. Молодих учених та студентів, (Тернопіль, 6-7 грудня 2023). М-во освіти і науки України, Терн. націон. техн. ун-т ім. І. Пулюя [та ін.]. Тернопіль: ФОП Паляниця В. А., 2023. С. 404-405.
5. Uwisengeyimana, J. de D., Khalid Al-Salihi, N., & Ibriki, T. (2016). Statistical Performance Effect of Feature Selection Techniques on Eye State Prediction Using EEG. *International Journal of Statistics in Medical Research*, 5(3), 224–230. <https://doi.org/10.6000/1929-6029.2016.05.03.9>.
6. Ramirez-Arias, F.J.; Garcia-Guerrero, E.E.; Tlelo-Cuautle, E.; Colores-Vargas, J.M.; Garcia-Canseco, E.; López-Bonilla, O.R.; Galindo-Aldana, G.M.; Inzunza-González, E. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies* 2022, 10, 79. <https://doi.org/10.3390/technologies10040079>.
7. Rösler, Oliver and David Suendermann. "A First Step towards Eye State Prediction Using EEG." (2013).
8. Durna Y, Ari F. Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images. *Applied Sciences*. 2017; 7(5):498. <https://doi.org/10.3390/app7050498>.
9. Y. Boltov, I. Skarga-Bandurova and M. Derkach, "A Comparative Analysis of Deep Learning-Based Object Detectors for Embedded Systems," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1156-1160, doi: 10.1109/IDAACS58523.2023.10348642.

#### References

1. Mai, T.D.T., Phung, TN. (2023). Evaluating the Performance of Some Deep Learning Model for the Problem of Emotion Recognition Based on EEG Signal. In: Nghia, P.T., Thai, V.D., Thuy, N.T., Son, L.H., Huynh, VN. (eds) *Advances in Information and Communication Technology. ICTA 2023. Lecture Notes in Networks and Systems*, vol 847. Springer, Cham. [https://doi.org/10.1007/978-3-031-49529-8\\_19](https://doi.org/10.1007/978-3-031-49529-8_19).
2. M. V. Derkach, D. S. Matiuk. Research of EEG technology on the example of a implemented neurocomputer interface. *Modern information technologies and management systems [Electronic resource]: Collection. materials IV International Scientific and Practical Conference of Young Scientists, Postgraduate Students and Students; April 20–21, 2023.* Kyiv: KNEU, 2023. P. 81–83.
3. Jayarathne I, Cohen M, Amarakeerthi S (2020) Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio. *PLoS ONE* 15(9): e0238872. <https://doi.org/10.1371/journal.pone.0238872>.
4. D. S. Matiuk, M. V. Derkach. Estimation power spectral density of EEG signal. *Actual problems of modern technologies: book of abstracts of the XII International scientific and practical conference of young researchers and*

- students, (Ternopil, December, 6th-7th, 2023). Ministry of Education and Science of Ukraine, Ternopil Ivan Puluj National Technical University [and other.]. Ternopil: PE Palianytsia V.A., 2023. P. 404-405.
5. Uwisengeyimana, J. de D., Khalid Al\_Salihy, N., & Ibrikci, T. (2016). Statistical Performance Effect of Feature Selection Techniques on Eye State Prediction Using EEG. *International Journal of Statistics in Medical Research*, 5(3), 224–230. <https://doi.org/10.6000/1929-6029.2016.05.03.9>.
  6. Ramírez-Arias, F.J.; García-Guerrero, E.E.; Tlelo-Cuautle, E.; Colores-Vargas, J.M.; García-Canseco, E.; López-Bonilla, O.R.; Galindo-Aldana, G.M.; Inzunza-González, E. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies* 2022, 10, 79. <https://doi.org/10.3390/technologies10040079>.
  7. Rösler, Oliver and David Suendermann. "A First Step towards Eye State Prediction Using EEG." (2013).
  8. Durna Y, Ari F. Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images. *Applied Sciences*. 2017; 7(5):498. <https://doi.org/10.3390/app7050498>.
  9. Y. Boltov, I. Skarga-Bandurova and M. Derkach, "A Comparative Analysis of Deep Learning-Based Object Detectors for Embedded Systems," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1156-1160, doi: 10.1109/IDAACS58523.2023.10348642.