

Захожай О.І., Крохмаль А.В.

## МЕТОД І ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СИМВОЛЬНОГО ПЕРЕТВОРЕННЯ РАСТРОВИХ ЗОБРАЖЕНЬ

*В статті розглянуто новий метод і програмне забезпечення для перетворення растрових зображень в символний вигляд для систем обробки і аналізу візуальної інформації. Відомо, що обробка растрових зображень, особливо високої роздільної здатності, вимагає значних об'ємів пам'яті для зберігання, а також обчислювальної потужності комп'ютерних систем. Представлений підхід дозволяє візуально представити інформативні об'єкти на растровому зображенні символами, які за виглядом максимально співпадають з відповідними фрагментами початкового зображення. З урахуванням малого обсягу для збереження символних даних, таке технічне рішення дозволяє забезпечити компактне зберігання інформативних ознак растрових зображень, а також спростити процедуру їхнього співставлення під час розпізнавання. Компактне представлення алфавіту класів також дозволяє значно знизити часову складність інформаційного процесу класифікації об'єктів на зображенні. Представлене програмне забезпечення, в якому реалізовано новий метод перетворення растрових зображень, може бути легко інтегроване до будь яких прикладних систем розпізнавання візуальних образів, а також системи підтримки прийняття рішень на основі візуальної інформації.*

**Ключові слова:** перетворення растрових зображень в символний вигляд, розпізнавання растрових зображень, програмне забезпечення обробки та аналізу зображень, методи і алгоритми попередньої обробки зображень, інформаційні системи та технології, моделі представлення візуальних даних.

**Вступ.** Методи і засоби обробки візуальної інформації знаходять широке використання в сучасних технічних і технологічних системах, системах контролю та управління різноманітними об'єктами [1-4]. Це пов'язано, насамперед, з тим, що людина в оточуючому світі сприймає інформацію переважно у візуальній формі (порядку 80% інформації через органи візуального сприйняття). Це все стимулює стрімкий розвиток технічних засобів реєстрації візуальної інформації, які останнього часу значно удосконалилися і можуть використовуватися в різноманітних умовах спостереження та з різним рівнем деталізації [2].

Однак, одночасно з вдосконаленням технічних засобів реєстрації відеоінформації, необхідне вдосконалення методологічної та алгоритмічної бази процесів обробки візуальних даних для вирішення широкої номенклатури прикладних задач [1-5]. Особливої уваги заслуговує прикладне використання систем обробки візуальної інформації для розпізнавання образів, вирішення класифікаційних задач. Відповідно, напрям досліджень, спрямований на удосконалення методів та програмних засобів обробки растрових зображень, який розглядається в роботі є, безсумнівно, актуальним.

**Аналіз питання та постановка задачі.** Складність реалізації прикладних систем обробки візуальної інформації пов'язана, насамперед, з тим, що існує необхідність у використанні великих обсягів пам'яті для зберігання відеоданих, а також порівняно високої обчислювальної потужності для обробки цих масивів даних [6]. Таким чином, ефективність реалізації засобів обробки даних напряму залежить від розміру даних, що надаються для обробки, а також часової складності самих алгоритмів обробки.

Таким чином, важливим аспектом покращення методів і засобів аналізу візуальної інформації є забезпечення прийняття рішення на основі обмеженої кількості але найбільш інформативних даних. В цьому випадку може бути досягнуто достатній рівень достовірності при мінімальних витратах часу та обсягів пам'яті.

Існують різні підходи для виділення найбільш інформативних ознак з растрового зображення. Найбільш розповсюджені з них:

- перетворення кольорового растрового зображення у відтінки сірого кольору;
- порогова чи адаптивна бінаризація;
- виділення контурів на зображенні;
- кольорова фільтрація;
- морфологічні перетворення;
- скелетонізація;
- сегментація.

Усі зазначені підходи дозволяють певним чином зменшити розмірність інформаційного поля аналізу, однак, найбільший ефект забезпечує підхід, запропонований в [6] та оснований на символному перетворенні растрових зображень. Згідно цього підходу, растрове зображення замінюється на послідовність текстових символів, які за виглядом максимально подібні для відповідного фрагмента зображення. Співставлення кожного символу з числовим кодом дозволяє компактно записати таке представлення зображення з використанням мінімальних обсягів пам'яті. Подальше співставлення символів під час порівняння з еталонами також значно спрощує отримання класифікаційного рішення.

Тому, задачею представленого дослідження стала розробка нового методу символічного перетворення растрового зображення та його програмна реалізація для подальшого використання в інформаційних системах розпізнавання візуальної інформації.

До нового методу апріорно були зазначені наступні вимоги:

- вхідними даними є довільне растрове фотографічне зображення;
- вихідними даними є текстовий документ із отриманим зображенням у символічному представленні, що є зручним для візуального сприйняття людиною;
- метод повинен дозволити виділяти на зображенні інформативний об'єкт на фоні;
- неінформативними даними є фон на якому відображається об'єкт;
- розмірність інформаційного поля даних повинна бути мінімальною завдяки перетворенню зображення у символічне представлення за принципом співставлення структури;
- найбільш інформативними ознаками об'єкту на зображенні є його контури;
- найменш інформативними ознаками об'єкту на зображенні є колір та текстура;
- якість виділення контурів має мінімально залежати від освітленості та контрастності зображення;
- час на обробку зображення повинен бути мінімізований шляхом обрізки зображення перед його подальшою обробкою;
- метод повинен бути зручним для ручного налаштування параметрів обробки, таких як константи фільтрації, порогове значення бінаризації чи контури обрізки.

**Вирішення задачі.** У розробленому методі обробка зображення відбувається в 6 етапів. Структурна схема перетворень та зміни зображення на кожному етапі наведено на рисунку 1.



Рисунок 1 – Перетворення зображення на кожному етапі реалізації метода

Надалі наводиться опис кожного етапу в запропонованому методі на прикладі перетворення фотографічного зображення людини.

**Обрізка.** На даному етапі вхідне зображення обрізається до розмірів людини чи її обличчя. Це дозволяє видалити із зображення непотрібні об'єкти та обмежити розміри зображення для подальшої обробки з метою зменшення обчислювального навантаження. Для можливості гнучкого налаштування, дана стадія може бути пропущена, виконана автоматично чи виконана вручну. Перший варіант може бути застосований користувачем в ситуації, коли вхідне зображення вже обрізане до розмірів досліджуваного об'єкта. Ручна обрізка повинна здійснюватися через інтерактивний графічний інтерфейс програмного забезпечення. Сутність автоматичної обрізки полягає у розпізнаванні досліджуваного об'єкта на зображенні та обрізці останнього по координатах області, в якій цей об'єкт знаходиться. Тобто, в цьому випадку, використовується додаткова система розпізнавання образів, яка повинна мати достатньо високу швидкодію для того, щоб її використання було доцільним. У запропонованому методі було використано класифікатор на основі каскадів Хаара [1] для визначення людини чи людського обличчя. Однак даний метод класифікації відзначається невисокою точністю ідентифікації людини. Для запобігання цьому, може бути застосований класифікатор на основі гістограм орієнтованих градієнтів, який показує більшу точність в контексті представлені прикладної задачі. Вхідне зображення та результати автоматичної обрізки обличчя та людини наведено на рисунку 2.

**Видалення фону.** Даний етап може бути пропущеним за умови, що фон на зображенні вже видалено чи він є однорідним. Також для цього етапу доцільно передбачити можливість його виконання вручну або автоматично. Ручне видалення фону відбувається за допомогою інтерактивного інтерфейсу, який пропонує користувачу відмітити ділянки зображення для залишення і видалення. Також доцільно передбачити можливість автоматичного видалення фону без участі людини. Це реалізовується на основі нейронної мережі для семантичної сегментації зображень [2]. Таким чином, з фотографії вирізаються лише сегменти, що відповідають зображенням людей на ній. Попередній етап обрізки дозволяє максимально зменшити обчислювальне навантаження на етапі сегментації. Альтернативою семантичної сегментації є використання нейронних мереж

для пошуку об'єктів на передньому плані. Відмінністю у методах є те, що останній дозволяє вирізати будь-який об'єкт переднього плану незалежно від його класу приналежності. Коли фон видалено, зображення додатково обрізається до розмірів об'єкта для видалення зайвих областей, що відповідають фону. Після цих двох етапів отримується зображення з максимально зменшеною розмірністю, до якого будуть застосовані перетворення наступних етапів реалізації метода.

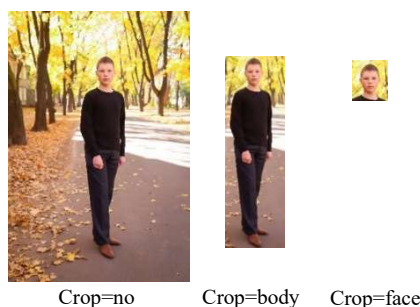


Рисунок 2 – Ілюстрація результатів застосування різних режимів обрізки інформативного об'єкта на зображенні

*Фільтрація за допомогою двостороннього фільтра.* На цьому етапі зображення піддається фільтрації для видалення дрібних об'єктів, текстур, цифрового шуму, артефактів стиснення з метою запобігання надлишковості даних у вихідному зображенні. Двосторонній фільтр був обраний через його властивість залишати контури предметів на зображенні чіткими і згладжувати лише дрібні деталі. Чіткість контурів є необхідною умовою для коректної роботи наступного етапу перетворень – гаусівської бінаризації.

*Бінаризація з адаптивним пороговим значенням за Гаусом.* На цьому етапі необхідно зменшити розрядність кольору зображення до 1 біту на піксель, зважаючи на необхідність збереження контурів зображення та видалення інформації про кольори частин об'єкта. Отже, для досягнення цієї мети було застосовано метод адаптивної бінаризації за Гаусом. Даний метод дозволяє бінаризувати зображення зі збереженням контурів та видаленням кольору в їх середині. Також метод забезпечує релевантний результат при різних рівнях освітленості та контрастності об'єктів. Також можливе тонке налаштування бінаризації шляхом підбору відповідних констант.

*Стоншення темних ділянок до товщини в 1 піксель.* Цей етап необхідний для коректного виконання подальшого перетворення інформативного об'єкта в символічне представлення. Результатом гаусівської бінаризації є контури неоднакової товщини, яку необхідно вирівняти для спрощення процесу перетворення у символічне представлення. Для цього було використано морфологічну операцію стоншення, так як на відміну від скелетонізації, вона дає лінії, найбільш схожі на контури вхідного зображення. Традиційно, в морфологічних операціях, зображенням вважаються білі ділянки, а фоном – чорні. Тому, зображення, фактично, інвертується двічі: перед та після перетворення. В результаті отримується зображення чорними лініями товщиною в 1 піксель на білому фоні. Цей етап є останнім для попередньої обробки растрового зображення перед його трансформацією в символічне представлення.

*Перетворення у символічне представлення.* Цей етап є ключовим в даному методі. На вході мається бінарне зображення з товщиною ліній в 1 піксель, а на виході отримується масив строкових даних, що є символічним представленням вхідного зображення. Перетворення відбувається за структурним принципом. Сутність метода полягає у проході по зображенню вікном 64x64 пікселі та співставленні структури підматриці пікселів центральної області вікна одному з 411 символів кодової таблиці Shift JIS. Для здійснення цього співставлення використано згорткову нейронну мережу (CNN) [3], архітектура якої представлена на рисунку 3.

Нейронну мережу було натреновано на наборі даних, що складається з символів текстових документів, які містять приклади символічних малюнків ASCII Art у відповідності до зображень, отриманих рендерингом цих документів. Усі символічні малюнки створені шрифтом 12-pt MS PGothic з 2-піксельними пробілами. Тренування здійснювалося на 90% всього набору даних, 10% використовувалось для валідації. Для аугментації даних використано горизонтальний і вертикальний зсуви та гаусівський шум розмірністю в 1 піксель.

В результаті отримується текстовий файл із зображенням в символічному представленні, який при перегляді у текстовому редакторі, зі встановленим вищезгаданим шрифтом, буде візуально сприйматися людиною як вхідне зображення. Результат перетворення представлений на рисунку 4.

**Апробація результатів досліджень.** Запропонований метод був реалізований в консольному додатку, так як такий варіант, в подальшому, можна максимально просто застосувати в складі комплексів програмного забезпечення для певних прикладних застосувань щодо обробки та розпізнавання візуальної інформації. Вхідним файлом для додатку є зображення JPG, PNG або BMP. Результат роботи зберігається у вихідну папку: текстовий документ TXT та зображення PNG, утворене рендерингом отриманого тексту. За допомогою додаткових параметрів можна обрати режим автоматичної обрізки (обрізка зображення людини або лица чи відсутність обрізки), застосувати автоматичне видалення фону (в поточній реалізації налаштоване на ідентифікацію обличчя людей), змінити розмір зображення, обрати більш швидку модель нейронної мережі. Також була створена графічна оболонка для консольної програми. Вона являє собою програму з графічним інтерфейсом, що використовує графічні елементи керування для полегшення процесу перетворення. Консольна програма і

графічна оболонка були зібрані в єдиний інсталяційний пакет, що встановлює консольну програму за замовчуванням та надає можливість користувачу обрати установку графічної оболонки в якості додаткового компонента. Реалізація програмного забезпечення з різними типами інтерфейсу зображено на рисунку 5.

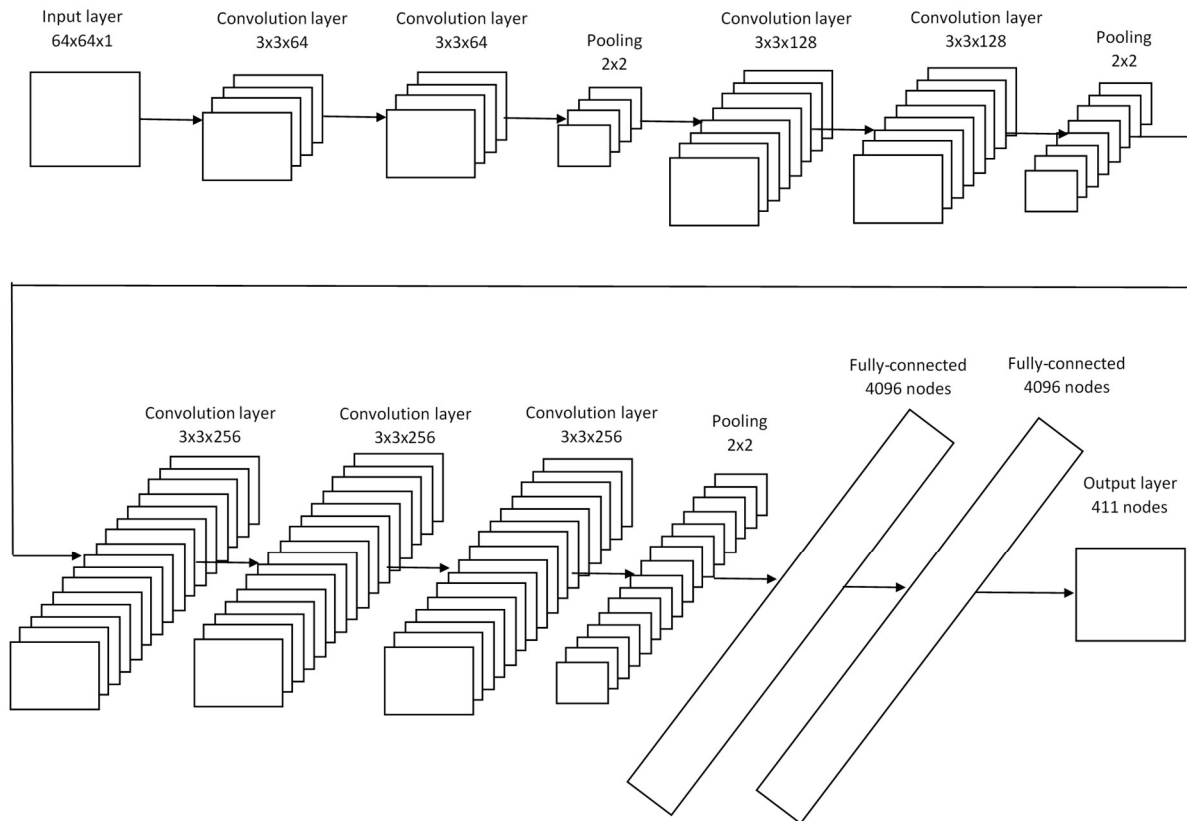


Рисунок 3 – Архітектура згорткової нейронної мережі для співставлення матриці пікселів текстовим символам

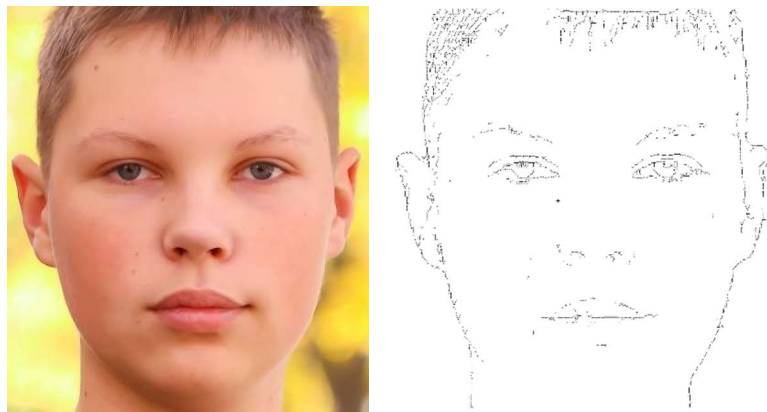


Рисунок 4 – Результат перетворення растрового зображення в символічне представлення за запропонованим методом

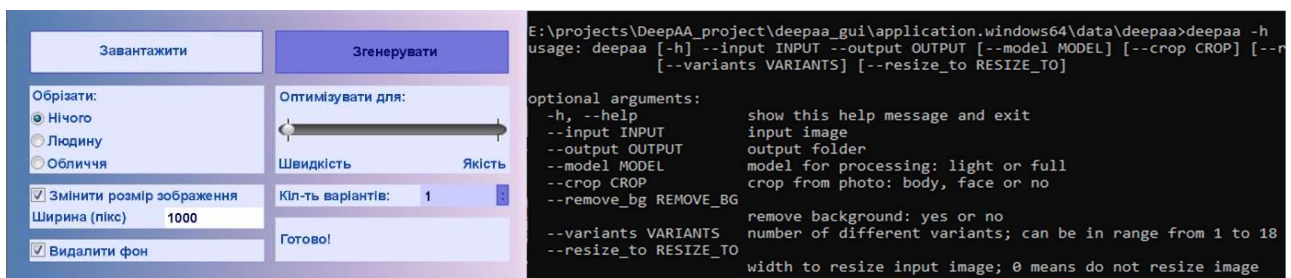


Рисунок 5 – Реалізація програмного забезпечення символічного перетворення растрових зображень з різними типами інтерфейсу

Ще одним варіантом реалізації програмного забезпечення для перетворення зображень на символічні малюнки став чат-бот Telegram. Він має командний інтерфейс керування, можливі команди відображаються при першому запуску бота. Алгоритми, використані для створення чат-бота, оптимізовані для паралельної обробки кількох зображень одночасно в системах з обмеженими ресурсами пам'яті.

Програма використовує фреймворк CUDA для прискорення обчислень у нейронних мережах, якщо в системі інстальовано графічний процесор, що підтримує цю технологію. Інакше, використовуються бібліотеки для обчислень на основі OpenBLAS, оптимізовані для використання центрального процесора [4].

Програма створена мовою програмування Python в середовищі MS Visual Studio Code. При розробці були використані фреймворки та бібліотеки у вільному доступі:

- Keras with Tensorflow backend;
- OpenCV-4;
- Pillow.

Використані апаратні засоби:

- процесор з підтримкою інструкцій AVX/AVX2;
- графічний процесор з підтримкою технології CUDA;

**Висновки.** Розроблено ефективний та гнучкий метод та програмне забезпечення перетворення растрового зображення у символічне представлення. Серед переваг цього метода можна виділити наступне.

1) Висока ефективність при обрізці зображення до розмірів об'єкта на ньому. При видаленні фону зображення максимально обрізається таким чином, щоб повністю видалити «рамку» з фонових пікселів навколо об'єкта. Крім того етап обрізки дозволяє мінімізувати час на видалення фону та знизити обчислювальне навантаження на комп'ютерну систему.

2) Можливість як автоматичного так і перетворення з тонким налаштуванням користувачем. Користувач може пропускати деякі кроки перетворення чи виконувати їх вручну. Також метод дає можливість користувачеві підлаштовувати окремі параметри з попереднім переглядом зображення на кожному кроці.

3) Якісне виділення інформативних контурів. Представлений метод дозволяє максимально ефективно виділити інформативні контури з мінімальною залежністю від яскравості та контрастності зображення, а також без надлишковості даних.

4) Істотне зменшення обсягів даних перетвореного зображення. При використанні символів 16x16 пікселів коефіцієнт стиснення у порівнянні з бінарним зображенням  $K \approx 3\%$ . Наприклад, бінарне зображення 160x160 пікселів займає 25600 бітів або 3200 байтів, тоді як у символічному представленні воно займатиме усього 100 байтів.

5) Результат перетворення візуально сприймається людиною при перегляді у будь-якому текстовому редакторі при встановленні необхідного шрифту.

6) Можливість крос-платформної реалізації програмного забезпечення як для Windows так і Linux платформ. Окрім того, є можливість використання графічних процесорів з CUDA-ядрами для прискорення обчислень (за його наявності у комп'ютерній системі).

7) Легкість інтеграції програмного забезпечення в різноманітні прикладні системи розпізнавання образів за рахунок використання консольної архітектури.

Серед недоліків запропонованого метода слід зазначити наступні.

1) Низька швидкість порівняно з методами без використання нейронної мережі.

2) В процесі перетворення зображення з символічного представлення у растрове виникають спотворення інформації через неповну відповідність вхідних пікселів символам.

3) Складності обробки зображень у символічному представленні. Для виконання деяких дій, таких як поворот зображення, необхідно виконати перетворення символічного зображення у растрове.

Серед перспективних напрямів розвитку та продовження представленого дослідження можна зазначити наступне:

- 1) Подальша оптимізація алгоритмів у напрямку зниження їхньої часової складності
- 2) Запровадження адаптивного методу виділення інформативних об'єктів на вхідному зображенні.
- 3) Створення кастомного символічного алфавіту з метою забезпечення максимальної візуальної подібності початкового та перетвореного зображення.

### Література

1. Wright J., Ma Y., Mairal J., Sapiro G., Huang T.S., Yan S. Sparse Representation for Computer Vision and Pattern Recognition // – NY: Proc. IEEE. – 2010. – Vol. 98. – Iss. 6. DOI: 10.1109/JPROC.2010.2044470
2. Zhang B., Gao Y., Zhao S., Liu J. Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High-Order Local Pattern Descriptor // – NY: Proc. IEEE. – 2010. – Vol. 19. – Iss. 2. DOI: 10.1109/TIP.2009.2035882
3. Захожай О. І. Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації й управління складними системами. Журнал «Проблеми інформаційних технологій» // – Х.: ХНТУ. – 2013. – № 01 (013). – с. 61-68.
4. Berclaz, J., Fleuret F., Turetken E., Fua P. Multiple object tracking using K-shortest paths optimization. IEEE Trans. Pattern Anal. Mach. Intell. // – NY: IEEE. – 2011. – vol. 33. – pp. 1806-1819.
5. Parker J. R., Terzidis K. Algorithms for image processing and computer vision. – Indianapolis: Wiley Pub. – 2011. – 506p.

6. Захожай О.І., Крохмаль А.В. Екстенціональний підхід до розпізнавання растрових зображень на основі їх символічних перетворень. Наукові вісті Давіського університету //– Київ.: СНУ. – 2023. – № 25 (04). DOI: 10.33216/2222-3428-2023-25-2
7. Face Recognition using Haar Cascade Classifier // International Journal for Modern Trends in Science and Technology 7(01) – January 2021 – pp. 85-87 -- DOI:10.46501/IJMTST070119.
8. Chen L-C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. NY: IEEE – 2017 – arXiv: 1606.00915v2.
9. Akiyama O. ASCII Art Synthesis with Convolutional Networks. NIPS 2017 //- Long Beach, California, USA. – 2017. – December 4-9. – Corpus ID: 19957044.
10. Wu Y., Liu L., Pu C., Cao W., Sahin S., Wei W., Zhang Q. A Comparative Measurement Study of Deep Learning as a Service Framework. IEEE Transactions on Services Computing // NY: IEEE. – 2019. – July 18. – pp. 1-15. DOI: 10.1109/TSC.2019.2928551

### References

1. Wright J., Ma Y., Mairal J., Sapiro G., Huang T.S., Yan S. Sparse Representation for Computer Vision and Pattern Recognition // – NY: Proc. IEEE. – 2010. – Vol. 98. – Iss. 6. DOI: 10.1109/JPROC.2010.2044470
2. Zhang B., Gao Y., Zhao S., Liu J. Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High-Order Local Pattern Descriptor // – NY: Proc. IEEE. – 2010. – Vol. 19. – Iss. 2. DOI: 10.1109/TIP.2009.2035882
3. Zakhzhay O. I. Information technology's patterns recognition in tasks of automation information's treatment and controlling of complex systems. Problems of information technologies //– Kherson.: KhNTU. – 2013. – No. 01 (013). – pp. 61-68.
4. Berclaz, J., Fleuret F., Turetken E., Fua P. Multiple object tracking using K-shortest paths optimization. IEEE Trans. Pattern Anal. Mach. Intell.//– NY:IEEE. – 2011. – vol. 33. – pp. 1806-1819.
5. Parker J. R., Terzidis K. Algorithms for image processing and computer vision. – Indianapolis: Wiley Pub. – 2011. – 506p.
6. Zakhzhay O.I., Krokhmal A.V. Extensional approach to the recognition of bitmap images based on their symbolic transformations. Naukovi Visti Dalivskogo Universytetu //– Kyiv.: VDEUNU. – 2023. – No. 25 (04). DOI: 10.33216/2222-3428-2023-25-2
7. Face Recognition using Haar Cascade Classifier // International Journal for Modern Trends in Science and Technology 7(01) – January 2021 – pp. 85-87 -- DOI:10.46501/IJMTST070119.
8. Chen L-C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. NY: IEEE – 2017 – arXiv: 1606.00915v2.
9. Akiyama O. ASCII Art Synthesis with Convolutional Networks. NIPS 2017 //- Long Beach, California, USA. – 2017. – December 4-9. – Corpus ID: 19957044.
10. Wu Y., Liu L., Pu C., Cao W., Sahin S., Wei W., Zhang Q. A Comparative Measurement Study of Deep Learning as a Service Framework. IEEE Transactions on Services Computing // NY: IEEE. – 2019. – July 18. – pp. 1-15. DOI: 10.1109/TSC.2019.2928551

*The article is devoted to solving the urgent issue of improving the methods and information technologies of raster image recognition for various applied applications. Today, there is a significant variety of methods for processing and recognizing raster images. The analysis of these methods and approaches presented in the article showed the feasibility of their use for certain applied applications. At the same time, there are no universal solutions. In addition, the processing of bitmap images is associated with the additional problem of using large amounts of memory for etalons storing, as well as excessive machine time consumption for processing large data sets. The article presents a new approach to increasing the raster images recognition speed and reducing the size of the data array that have been compared. This approach is based on the transformation of raster images into a symbolic form, followed by storage and comparison during recognition. According to the method, the pixel array of the image is divided into rectangular segments, after which each of them is replaced by an ASCII character that have a similar view. Further comparison is carried out not by individual pixels, but by symbol encodings of the image fragment. It is a further development of the extensional analysis approach. The new method makes it possible to ensure the compact storage of reference images for the classes alphabet, reduce memory usage for their storage, and also speed up classification.*

*Keywords: transformation of raster images into symbolic form, raster image recognition, image processing and analysis software, image preprocessing methods and algorithms, information systems and technologies, visual data representation models, informative signs selection methods and algorithms.*

**Захожай О.І.** - д.т.н., доцент, завідувач кафедри інформаційних технологій та програмування, Східноукраїнський національний університет імені Володимира Даля

**Крохмаль А.В.** - здобувач вищої освіти, Східноукраїнський національний університет імені Володимира Даля.