



DOI 10.28925/2663-4023.2024.24.363375

УДК 004.4, 004.6, 004.9

Зубик Людмила Володимирівна

к.п.н., доцент, доцент кафедри програмних систем і технологій
Київський національний університет імені Тараса Шевченка, Київ, Україна
ORCID ID: 0000-0002-2087-5379
zubyk.liudmyla@knu.ua

Пужай-Черета Сергій Костянтинович

провідний науковий співробітник науково-дослідного відділу наукового центру
Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
ORCID ID: 0000-0001-6972-3301
serzaz95@gmail.com

Сапельников Олександр Олександрович

старший науковий співробітник науково-дослідного відділу наукового центру
Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
ORCID ID: 0009-0007-5667-1662
aleksandrsapela@gmail.com

Калугін Денис Сергійович

старший науковий співробітник науково-дослідного відділу наукового центру
Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
ORCID ID: 0000-0003-4821-3681
killer007ck@gmail.com

Котляр Максим Олександрович

молодший науковий співробітник науково-дослідного відділу наукового центру
Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
ORCID ID: 0009-0007-8970-3683
makskot4238@gmail.com

АНАЛІЗ МЕТОДІВ І АЛГОРИТМІВ РОЗПІЗНАВАННЯ ТА ІДЕНТИФІКАЦІЇ ЗОБРАЖЕНЬ ЗА ЇХ ОКРЕМИМИ ФРАГМЕНТАМИ

Анотація. Розглянуто теоретичні основи побудови систем прийняття рішень на базі результатів розпізнавання зображень, що супроводжуються текстами. Наведено орієнтовну структуру системи розпізнавання зображень. В основу систем розпізнавання зображень покладено виокремлення текстових надписів на наявних фотографіях, їх попередня обробка, виокремлення ізольованих ділянок на зображенні, виконання математичних операцій над окремими групами пікселів для приведення їх до відомих форм і порівняння з ними. Виконано опис різних методів попередньої обробки зображень. Проведено аналіз доцільності застосування таких методів бінаризації зображень як адаптивна бінаризація Бредлі-Рота, медіанна фільтрація, фільтрація Гауса, методи збалансованих гістограм і дисперсій класу, дискримінантний аналіз, логістична, пробіт регресія тощо. Розглянуто різні алгоритми розбиття зображення на окремі ділянки з метою їх подальшого розпізнавання, серед них алгоритм ковзаючого середнього, алгоритм оцінки ймовірності знаходження об'єкту у виокремленій ділянці, що базується на аналізі меж, Category-independent object proposals, Constrained Parametric Min-Cuts, Multiscale combinatorial grouping, Selective Search тощо. Виконано порівняння різних реалізацій алгоритмів обробки зображень для забезпечення ефективного розпізнавання, класифікації та ідентифікації зображень. Удосконалення окремих реалізацій алгоритмів обробки зображень дозволяє скоротити час їх обробки, що важливо для роботи з великими масивами даних. Основна увага в дослідженні зосереджена на виборі найбільш ефективних методів для розпізнавання надписів на зображеннях,



удосконаленні алгоритмів, що їх реалізують, з метою побудови систем розпізнавання, спрямованих на обробку великих масивів даних.

Ключові слова: розпізнавання зображень; методи обробки зображень; алгоритми обробки зображень; методи ідентифікації; програмне забезпечення.

ВСТУП

Завдання розпізнавання зображень спрямовані на класифікацію вихідних даних за допомогою виділення істотних ознак, що притаманні саме цим даним. Для вирішення завдань розпізнавання зображень можна використовувати різні методи, серед яких — перебирати вигляди об'єктів під різними кутами, масштабами, зсувами тощо; шукати контур об'єкта і досліджувати його властивості; використовувати штучні нейронні мережі різних конфігурацій. Останній з підходів вимагає великої кількості зразків для конкретної задачі розпізнавання з правильними відповідями, що не завжди є у наявності, або ж такої структури нейронної мережі, що потребує тривалої розробки і буде чітко налаштована під специфіку конкретної поставленої задачі. Часто в якості завдання ставиться розпізнавання не всього зображення, а лише його певної частини, наприклад, текстового надпису на ньому, що передбачає вибір загального підходу до створення програмного забезпечення та розробку і реалізацію відповідних окремих алгоритмів класифікації.

Постановка проблеми. Сучасний світ характеризується появою величезної кількості неструктурованої цифрової інформації, обробку якої доцільно доручати системам штучного інтелекту. Сюди можна віднести багато різноманітних завдань, які постають перед суспільством як за мирного часу, так і в умовах різноманітних військових конфліктів, а саме — розшифровка супутникових знімків, розпізнавання облич людей та їх пошук у соціальних мережах, ідентифікація й класифікація розміщених у мережі зображень тощо. Для розробки автоматизованих систем прийняття рішень, які будуть спроможні забезпечити швидко і якісну обробку значних обсягів цифрових даних, необхідно використовувати певні схеми їх побудови й алгоритми, що забезпечать очікувані результати.

Аналіз останніх досліджень і публікацій. Завдання побудови автоматизованих систем розпізнавання в реальному часі, які надаватимуть користувачам інформацію, зібрану в соціальних мережах або пристроями IoT чи певними іншими засобами і опрацьовану програмно, приведену до очікуваного вигляду, завжди були актуальними, а в умовах воєнного стану тим більше [1]. З кожним новим днем з'являється певна кількість нових методів і алгоритмів, які можна успішно використовувати в архітектурах таких систем. У дослідженнях різних авторів, присвячених методам обробки оцифрованих даних було розглянуто такі методи обробки зображень як адаптивна бінаризація Бредлі-Рота, медіанна фільтрація, фільтрація Гауса, методи збалансованих гістограм і дисперсій класу, дискримінантний аналіз, логістична і пробіт регресія тощо [2]. Авторами розглянуто і проаналізовано наявні методи й алгоритми опрацювання зображень, виконано їх порівняння й підбір з метою подальшого використання для розпізнавання та класифікації зображень на основі супроводжуваних їх текстових надписів (рис. 1).



Рис. 1. Зразки зображень з надписами для розпізнавання

Метою статті є висвітлення підходів до побудови систем автоматизованої класифікації оцифрованих зображень для забезпечення прийняття обґрунтованих рішень.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

1. Структура систем розпізнавання зображень

Процеси розпізнавання зображень полягають у виявленні закономірностей у даних за допомогою певних алгоритмів і їх використанні для подальших дій, таких як класифікація тощо. Системи розпізнавання зображень зазвичай будують з таких блоків [4]:

- **Вхідне зображення** — подається у вигляді масиву чисел, які характеризують колір та/або яскравість у кожній точці зображення.
- **Примітивний опис** — виокремлення низькорівневих ознак, що не пов'язані з контекстом. Ці ознаки описують такі локальні атрибути зображення, як яскравість, текстура, колір, лінії, кути, межі тощо. Блок спрямований на приведення зображення до вигляду, зручного для наступної обробки. Примітивний опис вважається прийнятним, якщо він перетворює ознаки в систему координат, де числова відстань між ознаками відповідає різниці людського сприйняття цих двох ознак.
- **Символьний опис** — виокремлення більш складних ознак на основі примітивного опису, таких як форми об'єктів та самі об'єкти. На цьому етапі відбувається сегментація зображення, в основу якої закладається прийняття рішення про поділ сцени однорідні області.
- **Рівень семантичної інтерпретації** — відповідає за надання сенсу об'єктам, виокремленим на попередніх етапах. Система формує гіпотезу про зображення, що розпізнається, і очікує її підтвердження з нижчих рівнів.
- **Результуючий опис** — система генерує опис результатів обробки зображення на основі результатів попередніх рівнів.

Підходи до розпізнавання можуть суттєво відрізнятися від наведених і залежать від поставленого завдання.



2. Методи обробки зображень

2.1. Попередня підготовка зображень до обробки

Попередня обробка зображень призначена для підвищення точності його розпізнавання. Вона може включати підвищення різкості чи контрастності, видалення шумів та інші методи опрацювання. Один з базових методів, який використовують для обробки вхідного зображення, якщо це не суперечить предметній галузі, є вилучення зайвої інформації. Зменшення складності зображення можна отримати за рахунок здійснення переходу від кольорового зображення до відтінків сірого. Зображення у відтінках сірого є чорно-білим, і має лише один канал кольору замість трьох. Значення відтінку сірого також має назву «яскравість пікселя», і розраховують за стандартною формулою:

$$B = 0,3 * R + 0,59 * G + 0,11 * B \quad (1)$$

де B — значення яскравості, а R , G , B — значення кожного з трьох каналів кольору пікселя вхідного зображення.

Після приведення зображення до одноканального вигляду, його обробка займатиме менше часу і алгоритмічно буде простішою для реалізації.

2.2. Вилучення шумів на цифрових зображеннях

Цифровий шум — це випадкові та випадково розташовані зміни яскравості або кольору пікселів, що відсутні на реальному об'єкті. Цифровий шум помітний на зображенні у вигляді накладеної маски з пікселів. Шум завжди є небажаним побічним продуктом, який лише додає зайву інформацію. Нині існує багато способів зменшення шуму. При цифровій обробці зображень використовується **просторове шумозниження**. Виділяють наступні методи:

- адаптивна фільтрація;
- медіанна фільтрація;
- математична морфологія;
- Гаусове згладжування;
- методи на основі дискретного вейвлет-перетворення;
- метод головних компонент;
- анізотропна дифузія;
- фільтри Вінера тощо.

Адаптивна бінаризація Бредлі-Рота. Метод простий у реалізації і достатньо швидкий, проте підходить не для будь-яких зображень. Частіше за все це фото з рівномірними гистограмами, які насправді мають нерівномірне освітлення. Гистограми з маркуванням зазвичай мають дві чітко виражені групи пікселів за яскравістю. Алгоритм працює таким чином: для кожного пікселя зображення розраховується середня яскравість сусідніх йому пікселів в межах вікна певного розміру, що може бути встановлене користувачем як $1/8$ ширини зображення або інакше. Поріг — це добуток розрахованої середньої яскравості і константи k , що також встановлюється користувачем частіше за все як 15% від середнього значення яскравості пікселів у ділянці (зазвичай знаходиться в межах від 0,15 до 1). Якщо піксель перевищує поріг, то йому присвоюється значення 1, інакше — 0.

Медіанна фільтрація і математична морфологія — це одні з найбільш поширених методів, які використовуються для обробки зображень з метою зменшення рівня шуму. Медіанна фільтрація зменшує рівень зашумленості зображення у відтінках сірого, особливо позбувається шуму виду «сіль-перець», тоді як математична морфологія



покликана покращити загальний вигляд об'єктів на бінаризованих зображеннях, шляхом видалення занадто малих об'єктів або заповнення пропусків у інших об'єктах.

Медіанна фільтрація є досить поширеним методом, що використовується для очищення зображень від шуму та імпульсних перешкод. Медіанний фільтр замінює кожен елемент зображення на середнє значення окремих елементів зображення, що містяться в межах квадратного вікна. Перевагами застосування цього фільтру для вирішення задач обробки зображень у порівнянні з іншими фільтрами є його простота, ефективність, швидкість обробки даних і збереження чіткості контурів на фото. Розмір вікна медіанного фільтру є вхідним параметром, і впливає на силу згладжування зображення. Найбільш широко використовувані розміри вікон — це квадрати зі стороною 3, 5, 7 або 9 пікселів залежно від кількості шуму на фото.

Математична морфологія. Морфологічні методи спрямовані на опис форми об'єктів. Враховуючи це, для обробки окремих ділянок зображення, які підлягають розпізнаванню й спрямовані на ідентифікацію об'єктів, доцільно використовуватися саме математичну морфологію. Вхідними даними для бінарної морфології є вхідне бінарне зображення і деякий структурний елемент — тобто, інше і зазвичай менше бінарне зображення. Структурний елемент застосовується до всіх пікселів вхідного зображення.

Основними операціями бінарної морфології є нарощування, ерозія, замикання і розмикання.

Операція нарощування спрямована на збільшення об'єкту на зображенні. Кожного разу, коли початок координат структурного елемента співпадає з пікселем вхідного зображення, до всього структурного елемента застосовується перенос і операція диз'юнкції з відповідними пікселями вхідного зображення. Результати операції записуються у результуюче бінарне зображення, яке спочатку заповнюється нулями.

При виконанні операції ерозії, якщо одиничний піксель структурного елемента співпадає з одиничним пікселем вхідного зображення, то виконується операція кон'юнкції центрального пікселя структурного елемента з відповідним пікселем вихідного зображення. Операцію ерозії застосовують для вилучення малих об'єктів і шумів. Але вона, звісно, має недолік — зменшує всі об'єкти на зображенні. Його можна компенсувати, якщо після операції ерозії застосувати операцію нарощування з таким же структурним елементом.

В цьому полягає суть операції розмикання. Вона відсіює всі елементи зображення, менші за структурний елемент, але при цьому не дозволяє занадто зменшити великі об'єкти. Операція замикання схожа до розмикання, але спочатку вона застосовує операцію нарощування до зображення, а вже потім ерозії. Це дозволяє позбутися невеликих отворів у об'єктах.

Гаусове згладжування. Альтернативним варіантом для медіанної фільтрації є фільтр Гауса [2]. На відміну від неї, цей метод передбачає використання вікна певного розміру із заздалегідь розрахованими значеннями комірок, які є сталими для всього зображення. Вхідними даними є розмір ковзаючого вікна, в межах якого будуть розглядатися пікселі, та значення стандартного відхилення σ .

Значення комірок матриці розраховуються, за функцією Гауса:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

де x та y — координати комірки вікна, а σ — стандартне відхилення.



2.3. Методи бінаризації

При виконанні операції бінаризації зображення частина пікселів обирається в якості пікселів переднього плану, який представляє об'єкти, що розпізнаються, а решта — в якості фону. Значення порогу між фоном та переднім планом часто обирається вручну. Але, враховуючи специфіку зображень, які у більшості випадків отримуються з камер спостереження, тобто з відеопотоку, крім того, фіксація часто здійснюється при недостатньому освітленні, тому поріг краще обирати динамічно. Існує кілька методів бінаризації ділянок зображення, кожен із них на вхід отримує попередньо оброблене зображення у відтінках сірого, а на виході продукує бінаризоване зображення. Розглянемо деякі з найбільш поширених.

Метод збалансованої гистограми. Іншим методом бінаризації зображення є метод збалансованої гистограми [3]. Він не має жодних вхідних параметрів, по суті щоразу ітеративно «зважує» гистограму, розраховуючи яка з її частин є «важчою», і переміщуючи вагу з важчої частини, до легшої. Таким чином, відбувається поступове вирівнювання гистограми, як на вагах. Алгоритм гарно працює саме у тих випадках, коли на гистограмі чітко виражені два кластери, тому його можна розглядати як основний для досягнення поставленої мети.

Метод Оцу. Одним з найбільш ефективних порогових методів вважається метод, описаний Нобуюкі Оцу [6]. На відміну від попереднього, він замість поняття ваги класів оперує поняттям дисперсії класів. Метод Оцу [6] в деяких випадках є швидшим, тому що достатньо лише одного проходження по гистограмі. Алгоритм ітеративно обходить кожне значення яскравості $t = [0; 255]$ на гистограмі p , умовно розділяючи гистограму на два класи на кожній ітерації.

У багатьох випадках обидва методи (збалансованої гистограми і Оцу) працюють добре. Проте метод Оцу краще використовувати для бінаризації у випадках, коли метод збалансованої гистограми не працює, тобто гистограма для методу Оцу має бути бімодальною.

Для пришвидшення роботи методу, можна розбити гистограму на рівні за розмірами інтервали по осі значень яскравості OX на гистограмі, й ітеративно обходити ці інтервали. Перший клас позначається індексом b (*background*), оскільки до нього відносяться темніші пікселі й він є фоном зображення на даній ітерації. Відповідно, другий клас позначають індексом f (*foreground*). Для кожного з двох класів на даній ітерації розраховуються значення ваги класу w , середньозважене середнє класу m , та дисперсії класу σ^2 за формулами:

$$w_b = \frac{\sum_{i=0}^{t-1} p(i)}{N}, \quad w_f = \frac{\sum_{i=t}^{255} p(i)}{N} = 1 - w_b, \quad (3)$$

де $p(i)$ — це кількість пікселів з яскравістю i , N — загальна кількість пікселів на зображенні, t — поточне значення яскравості.

$$m_b = \frac{\sum_{i=0}^{t-1} i * p(i)}{\sum_{i=0}^{t-1} p(i)}, \quad m_f = \frac{\sum_{i=t}^{255} i * p(i)}{\sum_{i=t}^{255} p(i)} \quad (4)$$

$$\sigma_b^2 = \frac{\sum_{i=0}^{t-1} p_i * (i - m_b)^2}{\sum_{i=0}^{t-1} p(i)}, \quad \sigma_f^2 = \frac{\sum_{i=0}^{255} p_i * (i - m_f)^2}{\sum_{i=t}^{255} p(i)} \quad (5)$$

Далі обчислюють значення дисперсії всередині класу (англ. *within class variance*), який охоплює все зображення:

$$\sigma^2 = w_b * \sigma_b^2 + w_f * \sigma_f^2, \quad (6)$$

Дискримінантний аналіз — різновид багатовимірною аналізу, призначеного для рішення задач розпізнавання образів. Використовується для прийняття рішення про те, які змінні розділюють, тобто «дискримінують» певні набори даних. Дискримінантний аналіз є близьким до дисперсійного і регресійного аналізів, які також намагаються виразити одну із залежних змінних у вигляді лінійної комбінації інших показників. Однак, у двох інших методів залежна змінна є числовою величиною, в той час як у дискримінантному аналізі це категорійна змінна. Більш подібними до дискримінантного аналізу є **логістична** і **пробіт регресія**. Ці та інші методи використовуються переважно у тих випадках, коли не припускається нормальний розподіл незалежних змінних, що є основним припущенням методу дискримінантного аналізу.

Метод класифікації природних розривів (або Дженкса) використовує алгоритм для групування значень у класах, розділених різними точками розриву. Цей метод краще використовувати з нерівномірно розподіленими даними, але вони не повинні бути перекошеними до будь-якого кінця розподілу. На рис. 2 показана класифікація природних розривів для даних щільності населення округу США 1997 р.

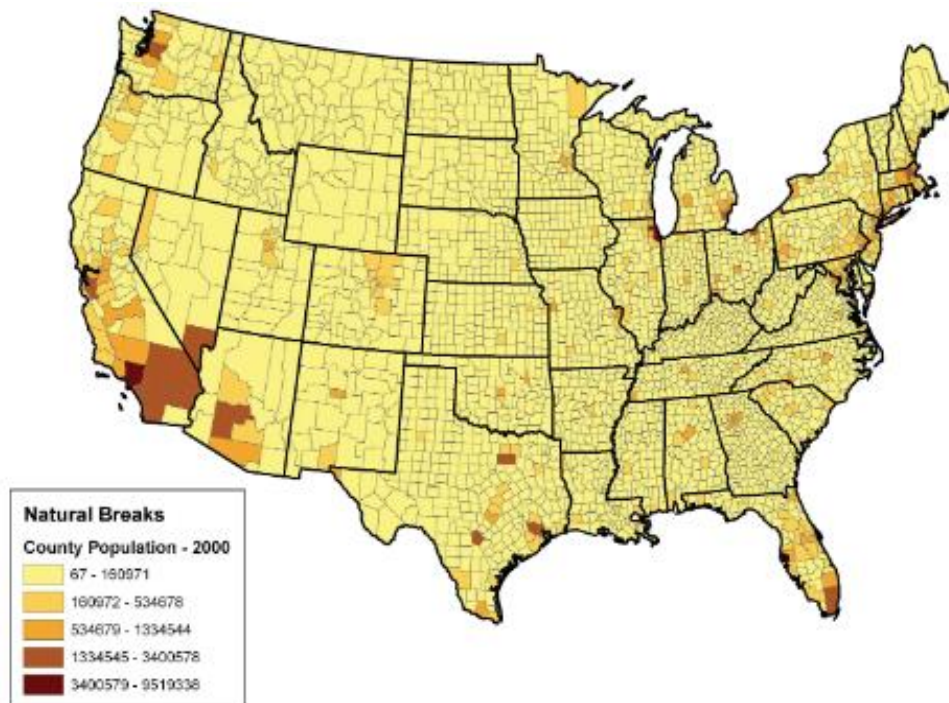


Рис. 2. Класифікація природних розривів для даних щільності населення округу США 1997 р.

Одним з потенційних недоліків є те, що цей метод може створювати класи, які містять широко різні діапазони чисел. Відповідно, клас 1 характеризується діапазоном трохи більше 150 000, тоді як клас 5 характеризується діапазоном понад 6 000 000. У таких випадках часто корисно або «налаштувати» класи після класифікації, або змінити мітки на деяку порядкову шкалу, наприклад «малий, середній або великий». Останній приклад, зокрема, може привести до більш зрозумілої глядачеві карти. Другим недоліком є той факт, що може бути важко порівняти дві або більше карт, створених методом класифікації природних розривів, оскільки діапазони класів настільки специфічні для кожного набору даних. У цих випадках набори даних, які можуть бути не надто розрізненими, можуть відображатися так у вихідній графіці.

3. Алгоритми виокремлення ділянок

Одним з найпростіших алгоритмів виокремлення ділянок вважається **алгоритм ковзаючого вікна** (англ. *exhaustive search* або *sliding window approach*). Він використовує «ковзаюче» вікно певного розміру для виокремлення ділянок на зображенні та подальшого їх аналізу з метою локалізації об'єкту. Щоб уникнути випадків, коли об'єкт, що розпізнається, потрапив до різних ділянок, і в результаті ігнорується, вікна зазвичай «накладаються» одне на одне. Такий підхід є дуже складним з обчислювальної точки зору, оскільки вимагає пошуку об'єкта в тисячах ділянок невеликого розміру.

Іншим методом для виокремлення ділянок є **алгоритм вимірювання «об'єктності» вікон зображення**. Він використовує алгоритм ковзаючого вікна, але вже кількісно визначає ймовірність знаходження об'єкту у виокремленій ділянці. Для цього він проводить межу між об'єктами у просторі й елементами фону. Проте для успішного виконання операції об'єкт повинен бути добре видимим у просторі (контрастним), або мати суттєві відмінності від фону, наприклад, білий літак на синьому небі (рис. 3).



Рис. 3. Приклад достатньо контрастного об'єкту

Існує ще декілька відомих алгоритмів для виокремлення ділянок, серед них Category-independent object proposals [7], Constrained Parametric Min-Cuts [8], Multiscale combinatorial grouping [10], Selective Search [24] тощо.

Алгоритм вибіркового пошуку є поєднанням двох інших алгоритмів — алгоритму ковзаючого вікна та сегментації. Як і алгоритм сегментації, він використовує структуру зображення для управління процесом пошуку. Як і алгоритм вичерпного пошуку, він прагне охопити всі можливі місця розташування об'єктів.



До переваг цього методу належать:

- швидкість роботи;
- висока точність;
- стійкість до шумів.

Метод не прив'язується до фіксованого співвідношення сторін, що надає йому ще більшої гнучкості. Вибірковий пошук працює наступним чином. Спочатку відбувається сегментація вхідного зображення. В результаті генерується набір сегментів, кожен з яких містить лише один об'єкт. Після цього схожі за певними ознаками сегменти ітеративно об'єднуються у більші ділянки, формуючи все більші фрагменти на зображенні. Операція об'єднання продовжується, доки все зображення не складе одну цілу ділянку. Усі ділянки, які були утворені шляхом злиття інших ділянок, запам'ятовуються, і є ділянками, які алгоритм повертає як результат своєї роботи.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Забезпечення осіб, які приймають критично важливі рішення, достовірною інформацією та виконання цього у обмежені часові інтервали завжди було надзвичайно актуальним. Процес розпізнавання комбінованих зображень на основі їх окремих фрагментів з їх подальшим ідентифікацією та інтерпретацією результатів є складним і багатоетапним. В умовах сьогодення він потребує автоматизації у зв'язку з постійною появою значних обсягів неструктурованих даних та необхідністю їх швидкого опрацювання.

Методи обробки даних постійно вдосконалюються, існуючі алгоритми доповнюються новими блоками, а також постійно з'являються нові, побудовані з урахуванням інших підходів, алгоритмів обробки даних, технологій тощо.

У статті розглянуто й проаналізовано класичні методи і алгоритми, що використовуються для побудови сучасних систем розпізнавання зображень на основі аналізу їх фрагментів. Подальші дослідження будуть зорієнтовані на пошук ефективних алгоритмів розпізнавання зображень та їх ідентифікацію з акцентами на ідентифікацію фонових фрагментів, отриманих з наявних баз даних та соціальних мереж.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ivanytska, A., Zubyk, L., Gololobov, D., Isaienkov, Y., Grynkevych, G., & Bychkov, O. (2023). The system for recognizing useful information of the client's ID-card based on machine learning technologies. *CEUR Workshop Proceedings, Vol. 3687*, 115–120.
2. Mafi, M., Martin, H., Andrian, J., Barreto, A., Cabrerizo, M., & Adjouadi, M. (2019). A Comprehensive Survey on Impulse and Gaussian Denoising Filters for Digital Images. *Signal Processing, 157*, 236–260.
3. Mousavirad, S. J., & Ebrahimpour-Komleh, H. (2017). Multilevel image thresholding using entropy of histogram and recently developed population-based metaheuristic algorithms. *Evolutionary Intelligence, 10(1–2)*, 45–75. <https://doi.org/10.1007/s12065-017-0152-y>
4. Ma, G., & Yue, X. (2022). An improved whale optimization algorithm based on multilevel threshold image segmentation using the Otsu method. *Eng. Appl. Artif. Intell. 113*. <https://doi.org/10.1016/j.engappai.2022.104960>
5. Abd Elaziz, M., & Lu, S. (2019). Many-objectives multilevel thresholding image segmentation using knee evolutionary algorithm. *Expert Syst. Appl. 125*, 305–316.
6. Yadav, R., & Pandey, M. (2022). Image segmentation techniques: A survey. *Proceedings of data analytics and management: ICDAM 2021, 1*, 231–239.
7. Endres, I., & Hoiem, D. (2014). Category-Independent Object Proposals with Diverse Ranking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(2)*, 222–234. <https://doi.org/10.1109/TPAMI.2013.122>



8. Beideman, C., Chandrasekaran, K., & Chao, Xu. (2020). Licensed under Creative Commons License CC-BY Approximation, Randomization, and Combinatorial Optimization. *Algorithms and Techniques (APPROX/RANDOM 2020)*, 17, 1–21.
9. Aranguren, I., Valdivia, A., Pérez-Cisneros, M., Oliva, D., & Osuna-Enciso, V. (2022). Digital image thresholding by using a lateral inhibition 2d histogram and a mutated electromagnetic field optimization. *Multimed Tools Appl.* 81(7), 10023–10049.
10. Maire, M. & Yu, S. X. (2013). Progressive multigrid eigensolvers for multiscale spectral segmentation. *ICCV*, 2184–2191.
11. Nadimi-Shahraki, M. H., Taghian, S., Mirjalili, S., & Faris, H. (2020). Mtd: An effective multi-trial vector-based differential evolution algorithm and its applications for engineering design problems. *Appl. Soft. Comput.* 97. <https://doi.org/10.1016/j.asoc.2020.106761>
12. Houssein, E. H., Saad, M. R., Hashim, F. A., Shaban, H., Hassaballah, M. (2020). Lévy flight distribution: A new metaheuristic algorithm for solving engineering optimization problems. *Eng. Appl. Artif. Intell.* 94. <https://doi.org/10.1016/j.engappai.2020.103731>
13. Ahmadianfar, I., Bozorg-Haddad, O., & Chu, X. (2020). Gradient-based optimizer: A new metaheuristic optimization algorithm. *Inf. Sci.* 540, 131–159. <https://doi.org/10.1016/j.ins.2020.06.037>
14. Abdel-Basset, M., Mohamed, R., & Abouhawwash, M. (2022). A new fusion of whale optimizer algorithm with kapur's entropy for multi-threshold image segmentation: Analysis and validations. *Artif. Intell. Rev.* 55(8), 6389–6459.
15. Salehnia, T., MiarNaeimi, F., Izadi, S., Ahmadi, M., Montazerolghaem, A., Mirjalili, S., & Abualigah, L. (2024). A mtis method using a combined of whale and moth-flame optimization algorithms, 625–651.
16. Sharma, A., Chaturvedi, R., & Bhargava, A. (2022). A novel opposition based improved firefly algorithm for multilevel image segmentation. *Multimed Tools Appl.* 81(11), 15521–15544.
17. Chauhan, D., & Yadav, A. (2023). A crossover-based optimization algorithm for multilevel image segmentation. *Soft. Comput.* 1–33.
18. Thapliyal, S., & Kumar, N. (2024). Ascaeo: accelerated sine cosine algorithm hybridized with equilibrium optimizer with application in image segmentation using multilevel thresholding. *Evolving Syst.* 1–62
19. Liu, Q., Li, N., Jia, H., Qi, Q., & Abualigah, L. (2023). A chimp-inspired remora optimization algorithm for multilevel thresholding image segmentation using cross entropy. *Artif. Intell. Rev.* 56(1), 159–216.
20. Chen, Y., Wang, M., Heidari, A., Shi, B., Hu, Z., Zhang, Q., Chen, H., Mafarja, M., & Turabieh, H. (2022). Multi-threshold image segmentation using a multi-strategy shuffled frog leaping algorithm. *Expert Syst. Appl.* 194. <https://doi.org/10.1016/j.eswa.2022.116511>
21. Houssein, E. H., Abdalkarim, N., Hussain, K., & Mohamed, E. (2024). Accurate multilevel thresholding image segmentation via oppositional snake optimization algorithm: Real cases with liver disease. *Comput. Biol. Med.* 169.
22. Kumar, B. V., Oliva, D., & Suganthan, P. (2022). *Differential Evolution: From Theory to Practice*. Springer Singapore. <https://doi.org/10.1007/978-981-16-8082-3>
23. Reisenhofer, R., Bosse, S., Kutyniok, G., & Wiegand, T. (2018). A haar wavelet-based perceptual similarity index for image quality assessment. *Signal Process: Image Commun.* 61, 33–43.
24. Uijlings, J. R. R., van de Sande, K. E. A., Gevers, T. & Smeulders, A. W. M. (2013). Selective Search for Object Recognition. *Int. J. Comput. Vis.* 104, 154–171. <https://doi.org/10.1007/s11263-013-0620-5>

**Liudmyla Zubyk**

PhD, Associate Professor of Software Systems and Technologies Department

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

ORCID ID: 0000-0002-2087-5379

zubyk.liudmyla@knu.ua

Sergii Puzhai-Chereda

Leading Researcher of the Research Department

Ivan Kozhedub Kharkiv National University of the Air Force, Kharkiv, Ukraine

ORCID ID: 0000-0001-6972-3301

serzaz95@gmail.com

Oleksandr Sapelnykov

Senior Researcher of the Research Department

Ivan Kozhedub Kharkiv National University of the Air Force, Kharkiv, Ukraine

ORCID ID: 0009-0007-5667-1662

aleksandrsapela@gmail.com

Maksym Kotliar

Junior Researcher of the Research Department

Ivan Kozhedub Kharkiv National University of the Air Force, Kharkiv, Ukraine

ORCID ID: 0009-0007-8970-3683

maskkot4238@gmail.com

Denis Kalugin

Senior Researcher of the Research Department

Ivan Kozhedub Kharkiv National University of the Air Force, Kharkiv, Ukraine

ORCID ID: 0000-0003-4821-3681

killer007ck@gmail.com

ANALYSIS OF METHODS AND ALGORITHMS FOR RECOGNITION AND IDENTIFICATION OF IMAGES BY THEIR SEPARATE FRAGMENTS

Abstract. The theoretical foundations of building decision-making systems based on the results of image recognition accompanied by texts are considered. The approximate structure of the image recognition system is given. The basis of image recognition systems is the selection of text inscriptions on existing photos, their pre-processing, selection of isolated areas on the image, performance of mathematical operations on individual groups of pixels to bring them to known forms and comparison with them. The description of various methods of image preprocessing is performed. An analysis of the feasibility of using such methods of image binarization as adaptive Bradley-Roth binarization, median filtering, Gaussian filtering, methods of balanced histograms and class variances, discriminant analysis, logistic, probit regression, etc. was carried out. Different algorithms for dividing the image into separate areas for the purpose of their further recognition are considered. among them the moving average algorithm, the algorithm for estimating the probability of finding an object in a selected area based on boundary analysis, Category-independent object proposals, Constrained Parametric Min-Cuts, Multiscale combinatorial grouping, Selective Search, etc. A comparison of different implementations of image processing algorithms to ensure effective recognition, classification and identification of images is performed. Improvement of individual implementations of image processing algorithms allows to reduce their processing time, which is important for working with large data sets. The main focus of the research is on choosing the most effective methods for recognizing inscriptions on images, improving the algorithms that implement them, with the aim of building recognition systems aimed at processing large data sets.

Keywords: image recognition; image processing methods; image processing algorithms; methods of identification; Software.



REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Ivanytska, A., Zubyk, L., Gololobov, D., Isaienkov, Y., Grynkevych, G., & Bychkov, O. (2023). The system for recognizing useful information of the client's ID-card based on machine learning technologies. *CEUR Workshop Proceedings, Vol. 3687*, 115–120.
2. Mafi, M., Martin, H., Andrian, J., Barreto, A., Cabrerizo, M., & Adjouadi, M. (2019). A Comprehensive Survey on Impulse and Gaussian Denoising Filters for Digital Images. *Signal Processing, 157*, 236–260.
3. Mousavirad, S. J., & Ebrahimpour-Komleh, H. (2017). Multilevel image thresholding using entropy of histogram and recently developed population-based metaheuristic algorithms. *Evolutionary Intelligence, 10(1–2)*, 45–75. <https://doi.org/10.1007/s12065-017-0152-y>
4. Ma, G., & Yue, X. (2022). An improved whale optimization algorithm based on multilevel threshold image segmentation using the Otsu method. *Eng. Appl. Artif. Intell. 113*. <https://doi.org/10.1016/j.engappai.2022.104960>
5. Abd Elaziz, M., & Lu, S. (2019). Many-objectives multilevel thresholding image segmentation using knee evolutionary algorithm. *Expert Syst. Appl. 125*, 305–316.
6. Yadav, R., & Pandey, M. (2022). Image segmentation techniques: A survey. *Proceedings of data analytics and management: ICDAM 2021, 1*, 231–239.
7. Endres, I., & Hoiem, D. (2014). Category-Independent Object Proposals with Diverse Ranking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(2)*, 222–234. <https://doi.org/10.1109/TPAMI.2013.122>
8. Beideman, C., Chandrasekaran, K., & Chao, Xu. (2020). Licensed under Creative Commons License CC-BY Approximation, Randomization, and Combinatorial Optimization. *Algorithms and Techniques (APPROX/RANDOM 2020), 17*, 1–21.
9. Aranguren, I., Valdivia, A., Pérez-Cisneros, M., Oliva, D., & Osuna-Enciso, V. (2022). Digital image thresholding by using a lateral inhibition 2d histogram and a mutated electromagnetic field optimization. *Multimed Tools Appl. 81(7)*, 10023–10049.
10. Maire, M. & Yu, S. X. (2013). Progressive multigrid eigensolvers for multiscale spectral segmentation. *ICCV*, 2184–2191.
11. Nadimi-Shahraki, M. H., Taghian, S., Mirjalili, S., & Faris, H. (2020). Mtd: An effective multi-trial vector-based differential evolution algorithm and its applications for engineering design problems. *Appl. Soft. Comput. 97*. <https://doi.org/10.1016/j.asoc.2020.106761>
12. Houssein, E. H., Saad, M. R., Hashim, F. A., Shaban, H., Hassaballah, M. (2020). Lévy flight distribution: A new metaheuristic algorithm for solving engineering optimization problems. *Eng. Appl. Artif. Intell. 94*. <https://doi.org/10.1016/j.engappai.2020.103731>
13. Ahmadianfar, I., Bozorg-Haddad, O., & Chu, X. (2020). Gradient-based optimizer: A new metaheuristic optimization algorithm. *Inf. Sci. 540*, 131–159. <https://doi.org/10.1016/j.ins.2020.06.037>
14. Abdel-Basset, M., Mohamed, R., & Abouhawwash, M. (2022). A new fusion of whale optimizer algorithm with kapur's entropy for multi-threshold image segmentation: Analysis and validations. *Artif. Intell. Rev. 55(8)*, 6389–6459.
15. Salehnia, T., MiarNaeimi, F., Izadi, S., Ahmadi, M., Montazerolghaem, A., Mirjalili, S., & Abualigah, L. (2024). A mtis method using a combined of whale and moth-flame optimization algorithms, 625–651.
16. Sharma, A., Chaturvedi, R., & Bhargava, A. (2022). A novel opposition based improved firefly algorithm for multilevel image segmentation. *Multimed Tools Appl. 81(11)*, 15521–15544.
17. Chauhan, D., & Yadav, A. (2023). A crossover-based optimization algorithm for multilevel image segmentation. *Soft. Comput. 1–33*.
18. Thapliyal, S., & Kumar, N. (2024). Ascaeo: accelerated sine cosine algorithm hybridized with equilibrium optimizer with application in image segmentation using multilevel thresholding. *Evolving Syst. 1–62*
19. Liu, Q., Li, N., Jia, H., Qi, Q., & Abualigah, L. (2023). A chimp-inspired remora optimization algorithm for multilevel thresholding image segmentation using cross entropy. *Artif. Intell. Rev. 56(1)*, 159–216.
20. Chen, Y., Wang, M., Heidari, A., Shi, B., Hu, Z., Zhang, Q., Chen, H., Mafarja, M., & Turabieh, H. (2022). Multi-threshold image segmentation using a multi-strategy shuffled frog leaping algorithm. *Expert Syst. Appl. 194*. <https://doi.org/10.1016/j.eswa.2022.116511>
21. Houssein, E. H., Abdalkarim, N., Hussain, K., & Mohamed, E. (2024). Accurate multilevel thresholding image segmentation via oppositional snake optimization algorithm: Real cases with liver disease. *Comput. Biol. Med. 169*.
22. Kumar, B. V., Oliva, D., & Suganthan, P. (2022). *Differential Evolution: From Theory to Practice*. Springer Singapore. <https://doi.org/10.1007/978-981-16-8082-3>



23. Reisenhofer, R., Bosse, S., Kutyniok, G., & Wiegand, T. (2018). A haar wavelet-based perceptual similarity index for image quality assessment. *Signal Process: Image Commun.* 61, 33–43.
24. Uijlings, J. R. R., van de Sande, K. E. A., Gevers, T. & Smeulders, A. W. M. (2013). Selective Search for Object Recognition. *Int. J. Comput. Vis.* 104, 154–171. <https://doi.org/10.1007/s11263-013-0620-5>



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.