



Б. С. Лукашук

Національний лісотехнічний університет України, Львів, Україна

SCALESENSE: УДОСКОНАЛЕНИЙ МЕТОД АВТОМАТИЧНОГО ВИЗНАЧЕННЯ МАСШТАБУ ЛІНІЙКИ НА ЗОБРАЖЕННЯХ ДЛЯ ВИРІШЕННЯ ЗАВДАННЯ ОЦІНЮВАННЯ ГЕОМЕТРИЧНИХ РОЗМІРІВ ПЕВНИХ ОБ'ЄКТІВ

Виявлення відповідності пікселів реальним одиницям на зображенні є важливим кроком для подальшого аналізу геометричних характеристик об'єктів. Запропоновано удосконалений ітераційний метод для автоматичного визначення масштабу лінійки на зображенні, який базується на розпізнаванні цифр. Розроблено та протестовано отримані результати на зображеннях із біомедицинської сфери. Встановлено можливість застосування методу і до інших галузей, зокрема тих, які згадані у проаналізованих схожих дослідженнях – криміналістиці, ветеринарії, музейній справі тощо. Проаналізовано першу версію методу та встановлено деякі важливі недоліки в його алгоритмах, що зокрема дало змогу ефективніше опрацювати зображення із лінійками, що містять двоцифрові числа. Навчено та інтегровано у метод нову версію моделі для розпізнавання цифр – YOLOv7, що дало змогу виправити проблеми із оберненими зображеннями. Збільшено кількість опрацьованих методом зображень до 90 %. Досліджено альтернативні методи групування тестових сегментів, що є фінальним кроком методу. Опрацьовано експериментальні дані при застосуванні різних методів групування (кластеризація DBSCAN (англ. *Density-Based Spatial Clustering of Applications With Noise*), медіана, модифікований z-критерій та міжквартильний розмах) та похибки, порівняно із вручну вимірними значеннями. Отримано медіану похибки 4,2–4,4 %, у разі використання медіани та кластеризації методом DBSCAN, а при окремих конфігураціях методу DBSCAN – 3,1–3,7 %. Розроблено вебсторінку з демонстраційною версією методу для залучення більшої кількості даних та практичного використання методу у вирішенні реальних завдань. Надано можливість вибору одного з чотирьох методів групування та їх параметрів (для DBSCAN) та зображення для тестування методу. Реалізовано візуалізацію випадкового тестового сегменту на завантаженому зображенні для зручнішого суб'єктивного оцінювання отриманого результату користувачем. Запропоновано: провести додаткове навчання моделей виявлення цифр; реалізувати підтримку опрацювання чисел, з будь-якою кількістю цифрових символів; врахувати деякі з підходів, використаних в аналогічних методах; реалізувати REST API для виклику методу.

Ключові слова: цифрове оброблення зображень; вимірювання рани; опорний об'єкт; визначення об'єктів; визначення цифр; виявлення об'єктів у реальному часі; комп'ютерне моделювання; геометричний аналіз; біомедицинські зображення.

Вступ / Introduction

Завдання визначення геометричних розмірів певних об'єктів на зображеннях є досить поширеним і добре відомим у багатьох сферах. Різні методи дають змогу швидко і з достатньою точністю вирішити його. Однак їх застосування зазвичай вимагає знання відстані до об'єкта, а також параметрів об'єктива фотоапарата, принаймні фокусної відстані. Якщо ця інформація недоступна, таке завдання можна вирішити за допомогою еталонного об'єкта з відомими розмірами та відстанню до нього. Еталонний об'єкт потрібно розташовувати поруч із об'єктом, що цікавить, він повинен міститися в одній площині і на однаковій відстані від камери. У більшості практичних випадків використовуються наближені ручні методи, які характеризуються великою кількістю суб'єктивних похибок. Наприклад, у медицині лікарі ставлять вимірну лінійку біля рани па-

цієнта. Це дає змогу оцінити розмір досліджуваного об'єкта поглянувши на лінійку та оцінивши у такий спосіб приблизні геометричні характеристики. На сьогодні існують різні архіви і різні бази даних з величезною кількістю таких зображень не тільки з медичної галузі. Основна проблема полягає в тому, що оцінювання проводиться вручну і є дуже суб'єктивним та неточним. Це створює важливе завдання – точно й автоматично оцінити розмір певних об'єктів на зображеннях, якщо зображення містять еталонну лінійку. Очевидно, що таке завдання розбивається як мінімум на два підзавдання. Суть першого полягає в тому, що необхідно ідентифікувати числовий масштаб лінійки, а, водночас, визначити ціну всіх поділок, нанесених на цю лінійку. Друге завдання полягає у визначенні геометричних розмірів об'єкта дослідження порівнянням його з уже визначеною шкалою лінійки. Для вирішення першого завдання розроблено метод автоматичного визначення масштабу

Інформація про автора:

Лукашук Богдан Сергійович, аспірант, кафедра інформаційних технологій.

Email: bohdan.lukashchuk@gmail.com; <https://orcid.org/0009-0002-2365-4923>

Цитування за ДСТУ: Лукашук Б. С. ScaleSense: удосконалений метод автоматичного визначення масштабу лінійки на зображеннях для вирішення завдання оцінювання геометричних розмірів певних об'єктів. Науковий вісник НЛТУ України. 2023, т. 33, № 3. С. 90–96.

Citation APA: Lukashchuk, B. S. (2023). ScaleSense: an improved method of automatic identification of ruler scales on images to solve the problem of estimating the geometric dimensions of certain objects. *Scientific Bulletin of UNFU*, 33(3), 90–96.

<https://doi.org/10.36930/40330313>

лінійки на зображеннях для вирішення завдання оцінювання геометричних розмірів певних об'єктів [9] (далі згадуваний, як ScaleSense). Подано розвиток цього методу з урахуванням результатів першого дослідження та виправлення помилок в оригінальному алгоритмі.

Об'єкт дослідження – удосконалення методу автоматичного визначення масштабу лінійки на зображеннях для вирішення завдання оцінювання геометричних розмірів певних об'єктів.

Предмет дослідження – підходи до вдосконалення методу автоматичного визначення масштабу лінійки на зображеннях для вирішення завдання оцінювання геометричних розмірів певних об'єктів.

Мета роботи – удосконалити метод ScaleSelect задля збільшення ймовірності виявлення шкали лінійки на зображенні.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

1. Проаналізувати аналогічні дослідження для порівняння підходів та результатів;
2. Описати першу версію методу, встановити її недоліки, реалізувати їх виправлення у другій версії та оцінити кількість опрацьованих зображень;
3. Визначити та застосувати альтернативні методи групування тестових сегментів, оцінити та порівняти точність ScaleSelect із впровадженням цих методів;
4. Розробити демоверсію методу для кращого наочного оцінювання та практичного застосування.

Варто зауважити, що у цій роботі основну увагу приділено зображенням із медичної сфери через важливість завдання вимірювань ран, які використовуються для перевірки ефективності лікування та під час виконання хірургічних втручань. Проте цей метод також можна застосовувати в інших галузях, зокрема археології, музейній справі, криміналістиці тощо.

Аналіз останніх досліджень та публікацій. Як відомо [?], визначення масштабу лінійки є важливим прикладним завданням у багатьох сферах, усі супутні роботи намагаються вирішити цю проблему для застосування у певній практичній області. Більшість авторів прагнуть досягти високої точності шляхом виявлення лінійки та її кута з подальшим виявленням позначок лінійки та роблячи припущення на підставі цієї інформації.

Абгір Бгалерао та інші [3] вирішують проблему визначення шкали лінійки в області криміналістичних зображень. Як і в моїй роботі, вони роблять це, щоб зменшити обсяг рутинної ручної роботи. Автори визначають розташування лінійки на зображенні, а потім використовують спектральний аналіз, щоб визначити напрямки поділок на лінійці. Вони використовують двоступінний підхід, де спочатку визначається орієнтація лінійки за допомогою двовимірного дискретного перетворення Фур'є (ДПФ) на підставі точки-кандидата, а після цього створюється модель зображення для оцінювання градування лінійки. Модель лінійки базується на синусоїді, яка створюється за допомогою певної орієнтації, довжини хвилі та фази.

Дмитро Коновалов та інші [8] повідомляють про метод, який вирішує завдання визначення масштабу лінійки для автоматичного масштабування зображень риб, який може бути застосований для племінного виробництва. Автори досліджують проблему виявлення лінійки або її частини, але не визначають кількості пікселів на градацію лінійки. Метод складається з декількох етапів: за допомогою набору кроків розділити зоб-

раження на блоки однакового розміру (застосовуючи сітку з заданим кроком); для кожного блоку застосувати ДПФ для виявлення блоків, які містять частини лінійки, де багато паралельних ліній визначають лінійку; об'єднати ці блоки-кандидати, щоб отримати остаточний результат.

Казуакі Уеда та інші [13] також використовують двовимірні ДПФ для виявлення позначок шкали. Аналіз Фур'є також використовує Майкл Герман та співавтори [5] для виявлення реальних розмірів старовинних монет для музейних каталогів. На відміну від попередніх, Джун Сун [7] досліджує різні оператори для сегментації шляхом виявлення країв, і використовує їх для виявлення центрів позначки шкали лінійки. Автор обмежується тільки зображеннями лінійки, яка займає практично всю площу зображення та не аналізує більш складні приклади.

Майкл Теллахун та інші [11] представляють підхід, де спершу проводиться сегментація зображення лінійки із зображення за допомогою попередньо навченої згорткової нейронної мережі, після чого адаптивне порогове значення та накладання шаблонів використовуються для визначення просторових частот, які відповідають градуванням лінійки. Після цього виконується калібрування, де градуванням лінійки ставиться у відповідність очікувана кількість пікселів.

На відміну від більшості пов'язаних робіт, ScaleSense не потребує розрахунків частотних характеристик зображення для виявлення градування лінійки, натомість використовує цифрові позначки на поверхні лінійки, які зазвичай розміщують біля великих поділок і вказують на відстань від першої лінії градування до спостережуваної.

За декілька місяців після публікації першої версії методу ScaleSense Чжіон Сю та інші [15] опублікували метод визначення шкали лінійки для використання в галузі стоматологічних зображень. Автори пропонують комплексний підхід, який покриває різні типи лінійок і містить такі кроки:

1. Сегментація зображення лінійки із загального зображення.
2. Визначення цифр на лінійці та розрахунок відстані між їх центрами для визначення шкали лінійки.
3. У разі невдачі або відсутності цифр на зображенні – виділення лінійки прямокутником, використовуючи метод [1].
4. Знаходження градувань лінійки через частотний аналіз у відтинках сірого.
5. Знаходження шкали лінійки, задаючи сталу відстань між двома найменшими градуваннями лінійки.

На цю роботу звернено особливу увагу, оскільки вона теж стосується зображень з медичної сфери, була опублікована незадовго після першої публікації методу ScaleSense, окрім цього пропонує використовувати цифрові анотації градувань лінійки, проте має декілька переваг та недоліків у порівнянні зі ScaleSense.

Результати дослідження та їх обговорення / Research results and their discussion

Метод ScaleSense 2 можна розділити на декілька окремих кроків, де результат попереднього кроку продукує вхідні дані для наступного:

- 1) розпізнавання цифр на зображенні;
- 2) розпізнавання двоцифрових чисел;
- 3) сортування чисел за зростанням;

- 4) виявлення чисел, що містяться на одній прямій;
- 5) створення та групування тестових сегментів;
- 6) виявлення масштабу лінійки.

Така послідовність кроків розроблена після детального аналізу першої версії методу [9]. Зокрема, у ньому крок 2 – Розпізнавання двоцифрових чисел та крок 4 – Виявлення чисел, що знаходяться на одній прямій, відбувалися одночасно для зменшення кількості ітерацій. А крок 3 – Сортування чисел за зростанням слідував кроку 1 – Розпізнавання цифр на зображенні. Така послідовність спричиняла помилку, оскільки, якщо припустити такий результат кроку 1: [8, 9, 1, 0, 1, 1], можна побачити, що останні 4 цифри формують двоцифрові числа – [6, 10] відповідно, якщо відсортувати цифри до моменту виділення чисел отримаємо [0, 1, 1, 1, 8, 9], що значно зменшить шанс визначення шкали лінійки та точність. Також винесено крок 2 із кроку 4, що дало змогу зробити код більш зрозумілим та структурованим. На кроці 4 спрощено умову зупинки пошуку, про що описано у розділах нижче.

Розпізнавання цифр на зображенні. Загальна ідея цього кроку полягає у виконанні завдання виявлення об'єкта, щоб розпізнати та ідентифікувати цифри, що є на зображенні. Гіпотеза полягає в тому, що принаймні деякі з виявлених цифр розташовані над або під шкалою на лінійці та можуть бути використані як точки відліку для визначення шкали.

Основним результатом цього кроку є набір обмежувальних рамок. Кожна обмежувальна рамка повинна бути представлена в певному форматі з властивостями ' x_{min} ', ' y_{min} ', ' x_{max} ', ' y_{max} ' і ' $name$ ' (рис. 1), де перші чотири маркери містять відповідні мінімальні та максимальні координати обмежувальної рамки, а властивість ' $name$ ' містить виявлену мітку цифри.

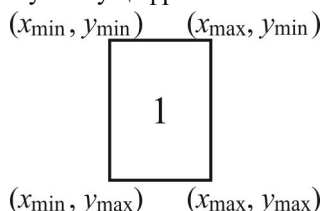


Рис. 1. Формат обмежувальної рамки виявленої цифри / Detected digit boundig box format

Суворих вимог до методу визначення цифр немає. Є тільки вимога до формату вихідних даних, яку описано вище. Отже, при подальшому розвитку методу ScaleSense можна допустити використання різних способів виявлення цифр та можливості їх конфігурації. У першій версії [9], де і описано мотивацію використати саме цю модель, було натреновано на власних даних та використано модель YOLOv5 [6], яка є однією з удосконалених версій оригінального алгоритму YOLO (англ. *You Only Look Once: Unified, Real-Time Object Detection*) [10]. Оскільки вже при роботі над першою публікацією передбачалася можливість застосування різних детекторів для першого кроку, для розроблення другої версії було використано Tesseract [12] – метод для оптичного розпізнавання символів, який часто використовується в оптичному скануванні сфотографованих документів, тексту на фотографіях тощо. Проте цей метод не здатен ефективно виявляти цифри на лінійках біомедичних зображень, отже, було прийнято рішення використовувати YOLOv5 та YOLOv7 [14] та порівняти результати ScaleSense з кожним з них.

У функції навчальних даних використано 184 зображення ран різними лінійками, зібраними з роботи [11], розмічені авторами вручну. Вся конфіденційна інформація на фото розмита. Розмір зображень 3140×1746 та роздільна здатність 72 точок на дюйм.

Оригінально поділ навчальних даних відбувався так:

- 159 зображень – тренувальна вибірка;
- 25 зображень – валідаційна вибірка.

Проте у висновках до першої версії вказано про недостатнє виявлення цифр на обернених зображеннях. Тому для навчання YOLOv7 моделі вибірка збільшено на 20 % способом додавання випадково обраних зображень із оригінальних і їх обороту на кут до 180 градусів. Отже, навчальна вибірка збільшилася до 196 зображень, а валідаційна – до 27.

Валідаційний та навчальний набори містять зображення з різними лінійками та різними ранами. Більшість зображень містили лінійки з масштабом до 9 см, але також присутні лінійки, які містять двозначні числа, проблему з якими вирішено у цій роботі, як описано на початку розділу. Оскільки лінійки на зображеннях різні, відповідно, і кількість представників різних цифр є різною, що ускладнює ефективне тренування.

Розпізнавання двоцифрових чисел. Оскільки модель навчена розпізнавати цифри, двоцифрові числа на лінійці можуть бути хибно сприйняті як дві цифри, які згодом формуватимуть один дуже короткий сегмент, що спричинить збільшення загальної похибки у разі його врахування. Для цього, на відміну від авторів [15], у методі ScaleSense реалізований алгоритм розпізнавання двоцифрових числа, який візуалізований на рис. 2.

Точки P_1 і P_2 відповідають центру обмежувальної рамки виявленої цифри 1 і обмежувальної рамки цифри 0 відповідно, A_1B_1 – ширині цифри 1, а A_2B_2 – ширині цифри 0. Перевірка висоти також виконується так само, як і ширина.

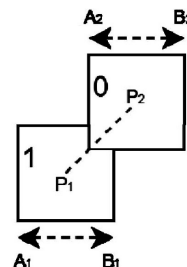


Рис. 2. Розпізнавання двозначних чисел / Two-digit number detection

Якщо $|A_1B_1| + |A_2B_2| > |P_1P_2|$ – ці дві цифри вважаються одним двозначним числом. Перевіряється як сума ширини цифр, так і сума висоти цифр, і якщо хоча б один показник довший, ніж відстань між центрами цих цифр – приймається рішення, що це двозначне число. Після цього дві цифри формують одне число і передаються на наступний крок сортування.

Виявлення чисел, що знаходяться на одній прямій. Суть методу полягає в наступному. Взявши відсортовані цифри та числа з попереднього кроку як вхідні дані та використовуючи прості рівняння лінійної алгебри та кілька ітерацій, метод формує точки-кандидати як вихідні дані для наступного кроку.

Алгоритм складається з кількох кроків:

- вибрати два опорні числа (D_1 і D_2) із знайдених;
- використовуючи (1), отримуємо рівняння прямої L_1 , яка проходить через центри (P_1 , P_2) чисел D_1 і D_2 , відповідно

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}; \quad (1)$$

- з решти цифр, використовуючи модифікацію алгоритму кидання променів, знаходимо ті, обмежувальні рамки яких перетинає L_1 (рис. 3),

де: D_1 і D_2 – опорні цифри (з мітками 1 і 4), P_1 і P_2 – їх центри, L_1 лінія, створена за допомогою (1), яка проходить через точки P_1 і P_2 ; x_1, x_2, y_1, y_2 – відповідні координати центрів D_1 і D_2 .

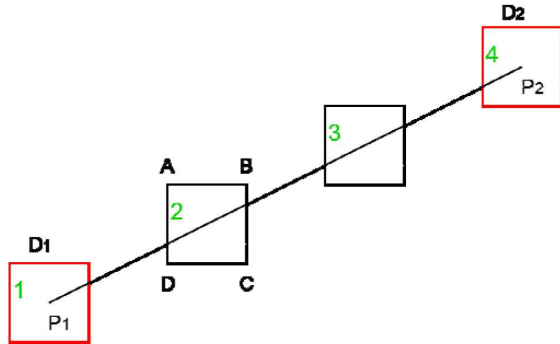


Рис. 3. Ілюстрація алгоритму кидання променя для виявлення чисел, що знаходяться на одній прямій між двома опорними точками / Illustration of the ray casting algorithm that detect numbers lying on the same line between two reference points

Наступним кроком виконується перевірка, чи L_1 перетинається з іншими цифрами, 2 і 3, відповідно до прикладу на рис. 3. До прикладу, цифра 2 представлена чотирма точками A, B, C, D , з яких за допомогою (1) утворюються ребра AB, BC, CD і DA . Розв'язок системи двох лінійних рівнянь вказує на точку перетину. Додатково проводиться перевірка, чи знайдена точка знаходиться в діапазоні точок A і D (належить відрізку AD). Те саме стосується решти цифр, які перевіряються. Тестування продовжується з різними опорними цифрами (або числами) D_1 і D_2 , поки не будуть виявлені ті, які знаходяться на одній прямій, що дає змогу відхилити помилково виявлені цифри (числа). Отримані цифри вважаються кандидатами на позначки шкали лінійки (рис. 4).

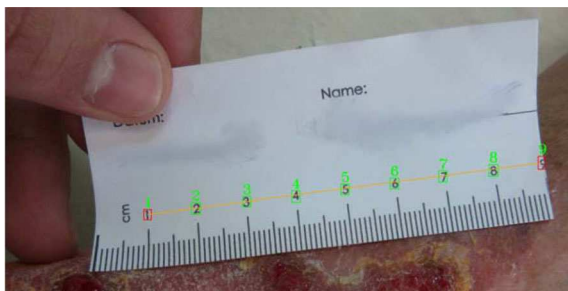


Рис. 4. Ілюстрація виявлення цифр, що знаходяться на одній прямій між двома опорними точками / Illustration of digits detection that lie on the same line between two reference points

Умова зупинки пошуку. Для того щоб визначити, що знайдено достатньо чисел, що знаходяться на одній прямій, у першій версії методу задається певна кількість ітерацій (5 ітерацій у версії алгоритму, з якої отримувалися експериментальні дані). Упродовж п'яти ітерацій відбувається пошук такої кількості цифр на одній прямій, щоб утворити хоча б 2 сегменти. На кожній з наступних ітерацій збільшується вимога до кількості знайдених чисел на одній прямій. Спершу половина з усіх визначених, потім 60, 70 і так до 100 %. Недолік полягає в тому, що 3 числа можуть сформувати тільки

2 сегменти, з яких неможливо створити групи сегментів чисельністю більше 1 у разі розбіжностей у значеннях відстаней, що є додатковим джерелом помилок. У ScaleSenseV2 підхід спрощений – пошук триває доти, доки або не знайдено хоча б 3 сегменти, тобто 4 числа, що знаходяться на одній прямій, або не перевірені всі комбінації чисел. Як і в попередньому випадку, переважно при першій ітерації знаходиться більше ніж 4 числа. Проте на зображеннях з дуже багатьма числами метод працюватиме довше, оскільки ітерації обходитимуть всі з них.

Створення та групування тестових сегментів і виявлення масштабу лінійки. У методі ScaleSense лінійка характеризується послідовністю цифр та чисел. На поточному кроці алгоритму на вхід надходить наростаючий список цифр і чисел, з яких утворюються тестові сегменти. Під тестовим сегментом розуміємо відрізок між центрами (P_1 та P_2) двох виявлених цифр, але тільки за умови послідовності цих цифр чи чисел. Проаналізувавши результат з рис. 4, видно, що тестові сегменти утворюються з пар цифр: (1, 2), (2, 3), (3, 4), (4, 5), (5, 6), (6, 7), (7, 8), (8, 9). Задавши послідовність виявлених цифр такою: (1, 3, 4, 5), отримуємо, що пара (1, 3) не утворить тестовий сегмент, оскільки 3 не є першим правим сусідом 1.

Основне джерело похибок цього методу – похибки визначення обмежувальних рамок цифри, тобто похибка алгоритму виявлення об'єктів. Також варто згадати про похибки друкування цифр на фізичній лінійці, проте ними можна впевнено знехтувати, враховуючи рівень розвитку сучасних верстатів. Для вирішення проблеми похибок визначення обмежувальних рамок цифри чи похибки, яка виникає при об'єднанні двох цифр в одне число, пропонується статистичний підхід. Замість того, щоб вибрати один випадковий сегмент та взяти його довжину в пікселях як відповідність 1 сантиметру при виявленні масштабу лінійки, проводиться пошук якомога більшої кількості сегментів та береться середнє арифметичне їх довжин як значення кількості пікселів на 1 сантиметр. На цьому етапі виникає проблема сегментів-викидів, тобто хибно-виявлених сегментів, довжина яких може бути значно меншою або значно більшою за всі інші, і відповідно значно впливати на кінцеве значення. У першій версії методу ScaleSense використано кластеризацію довжин сегментів методом DBSCAN (англ. *Density-Based Spatial Clustering of Applications With Noise*) [4], який враховує густину розподілу даних та підходить для таких задач. Із утворених кластерів обирався найбільший та обраховувалося його середнє арифметичне. Проте такий підхід можна вважати невиправдано складним, тому в цій роботі проаналізовано та порівняно додатково з ним три інших підходи для видалення сегментів-викидів:

- розрахунок медіани замість середнього арифметичного;
- модифікований Z-критерій;
- міжквартильний розмах (IQR).

Як і у першій публікації [9], як еталонні значення пікселів на сантиметр використано вручну виміряні співвідношення пікселів до сантиметрів з доступних та нами розмічених 184-х зображень. Для максимальної точності відстані вимірювалися між цифровими мітками, що знаходяться максимально близько до центру зображення, відсутні згини та лінійка розташована близько до ранової поверхні. Усі зображення мають ве-

лику роздільну здатність 3104×1746 пікселів, похибка ручного вимірювання варіюється в діапазоні [0-20] пікселів. Ці значення похибки ручного вимірювання є недостатньо точними, проте важливими для інтерпретації результатів роботи алгоритму. Для їх визначення проведено ручне вимірювання на кожній фотографії декілька разів та пораховано середнє відхилення від середнього значення.

Результати. На рис. 5 наведено приклад визначення шкали лінійки на зображенні алгоритмом ScaleSense.

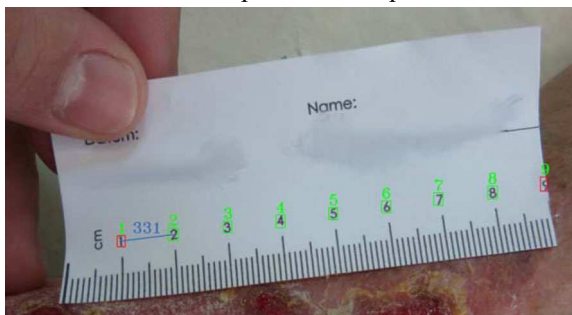


Рис. 5. Визначення кількості пікселів на 1 сантиметр методом ScaleSense / Pixel per 1 centimeter detection with the ScaleSense method

Для коректного оцінювання впливу змін, внесених до логіки роботи алгоритму, проведено серію тестів. У них порівнюємо не тільки першу версію методу з другою, а і чотири способи групування тестових сегментів та визначаємо їхні переваги та недоліки. Так, на рис. 6 наведено порівняльну діаграму кількості зображень, на яких метод зумів визначити шкалу лінійки.

За загальної кількості у 184 зображення, перша версія методу визначила шкалу лінійки тільки на 128 з них. Друга, використовуючи YOLOv5-154, а інтегрувавши YOLOv7, яка також була додатково натренована на зображеннях з випадковими оборотами – 167 зображень, що є більше, ніж 90 %. Зростання похибки між першою та другою версією методу сягає (якщо порівнювати за однакової моделі виявлення цифр) більше 14 %, а за різних – 21,2 %.

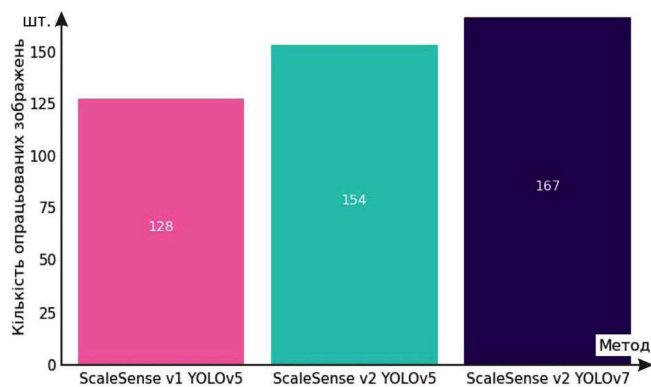


Рис. 6. Кількість зображень, опрацьованих методом / Number of images processed by the method

Збільшення кількості опрацьованих зображень спричинило зростання похибки. У табл. 1 наведено значення середньоарифметичної та медіанної похибки застосування відповідного методу за використання кожного з 4-х способів групування тестових сегментів, або ж відсутності групування, за використання медіани замість середньоарифметичного. Для першої версії методу є тільки дані щодо DBSCAN групування. У табл. 2 вказано середні значення похибок у відсотках від значень вручну виміряної кількості пікселів.

Табличні дані демонструють зростання середньоарифметичної абсолютної похибки разом зі збільшенням кількості опрацьованих зображень. При цьому значення медіани похибки є значно меншим, отже, здебільшого абсолютна похибка вимірювань не відрізняється від середньої похибки ручних вимірювань, яка коливається від 0 до 20 px, за нашими спостереженнями. З цього можна зробити висновок, що використання процедури групування не завжди вберігає від хибно позитивних тестових сегментів. На підставі отриманих результатів рекомендовано використовувати групування методом DBSCAN, яке є найбільш обчислювально складним, або ж медіану, яка є найпростішим способом з погляду витрат обчислювальних ресурсів, враховуючи результати тестування.

Табл. 1. Порівняння абсолютних значень помилок різних конфігурацій методу ScaleSense / Comparison of absolute error values of different configurations of the ScaleSense method

Груп. Метод	DBSCAN		Модифікований Z-критерій		IQR		Медіана	
	Медіана похибки, px	Середнє арифметичне похибки, px	Медіана похибки, px	Середнє арифметичне похибки, px	Медіана похибки, px	Середнє арифметичне похибки, px	Медіана похибки, px	Середнє арифметичне похибки, px
ScaleSense v1 YOLOv5	4	6,6	-	-	-	-	-	-
ScaleSense v2 YOLOv5	8	31,5	7	29,4	10	44,8	6	27,2
ScaleSense v2 YOLOv7	8	31,5	8	35,8	12	54,5	8	32

Табл. 2. Порівняння відносних значень помилок різних конфігурацій методу ScaleSense / Comparison of percentage error values of different configurations of the ScaleSense method

Груп. Метод	DBSCAN		Модифікований Z-критерій		IQR		Медіана	
	Медіана похибки, %	Середнє арифметичне похибки, %	Медіана похибки, %	Середнє арифметичне похибки, %	Медіана похибки, %	Середнє арифметичне похибки, %	Медіана похибки, %	Середнє арифметичне похибки, %
ScaleSense v1 YOLOv5	2,1	3,5	-	-	-	-	-	-
ScaleSense v2 YOLOv5	3,4	6,8	4,3	8,4	5,2	12,2	3,4	8,1
ScaleSense v2 YOLOv7	4,2	12,6	4,4	11,3	6,6	14,5	4,4	10,7

Також проведено декілька додаткових експериментів з різними конфігураціями методу DBSCAN. При задаванні параметра ϵ (найменша відстань для того, щоб вважати точки потенційними сусідами), що дорівнює 30, а мінімальної кількості точок у кластері 2, отримано відчутно менші значення абсолютної похибки, але, водночас, і меншу кількість опрацьованих зображень (140 та 145 у YOLOv5 та YOLOv7 версіях) (табл. 3).

Табл. 3. Значення похибок при конфігурації DBSCAN: $\epsilon = 30$; $\min_sample = 2$ / Error values for DBSCAN configuration: $\epsilon = 30$; $\min_sample = 2$

Метод \ Похибка	Медіана похибки, px	Середнє арифметичне похибки, px	Медіана похибки, %	Середнє арифметичне похибки, %
ScaleSense v2 YOLOv5	6	11,9	3,1	5,3
ScaleSense v2 YOLOv7	7	10	3,7	6,5

Можна зробити висновок, що хоч YOLOv7 дає змогу визначити більшу кількість цифр, ніж YOLOv5, проте створює менш точні обмежувальні рамки для них. У середньому для різних методів групування у 20 % розрахунків шкали лінійки методом ScaleSenseV2 похибка не переважає 2.5 px за роздільної здатності в 72 точки на дюйм (рис. 7).

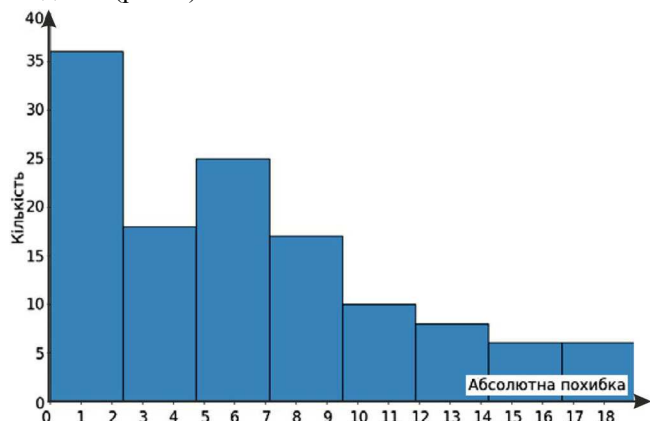


Рис. 7. Гістограма розподілу помилки частини даних для ScaleSenseV2 з використанням DBSCAN / Histogram of the error distribution of a portion of the data for ScaleSenseV2 using DBSCAN

Обговорення результатів дослідження. Автори [2] вирішують проблеми знаходження реальних розмірів об'єкта на зображенні. У вступі цієї роботи згадано, що за відомої фокусної відстані та відстані до об'єкта інтересу, задача визначення відповідності пікселів до реальних розмірів є вирішуваною, проте вказують на недоліки такого підходу та використовують метод з використанням еталонного об'єкта, де розміри потрібно вказувати користувачеві, що підкреслює практичну користь методу ScaleSelect та використання лінійки у ролі еталонного об'єкта.

Автори робіт [11, 15] не наводять даних про порівняння їхнього методу визначення шкали лінійки з певними еталонними даними. Проте у роботі [15] кроку визначення цифр на зображенні передують кроки сегментації лінійки із зображення, після чого пошук цифр відбувається тільки на виділеному сегменті. Можна висунути гіпотезу, що інтеграція цього кроку у метод ScaleSense дасть змогу уникнути помилок та пришвидшить роботу методу на зображеннях, що містять багато цифр,

які не належать градуванням лінійки. Для перевірки цієї гіпотези необхідно зібрати хоча б декілька десятків відповідних зображень. Також потрібно передбачити достатню різноманітність зображень лінійки, при навчанні сегментаційної нейронної мережі, оскільки недостатня навченість штучної нейронної мережі тільки збільшить кількість помилок на зображеннях, які містять лінійки, нетипові для зображень, на яких мережа була навчена.

У роботі [7] Джун Сун вимірює повну довжину лінійки, а не конкретно відповідності одного сантиметра чи іншої одиниці, а у роботі [8] автори знаходять лінійку або її частину на зображенні, проте не визначають шкалу. Абгір Бгалерао та інші [3] подають найбільш точні дані про результати свого дослідження та зазначають, що найбільша похибка їхнього методу сягає 1,26 % на зображенні з роздільною здатністю 173 точки на дюйм та 0,1 % за роздільної здатності зображень > 300 точок на дюйм, а в середньому похибка не перевищує 2,5 пікселя. При цьому автори використовують не фотографії, а якісні скановані зображення лінійки без додаткових елементів фону та не вказують кількість зображень, які використовують для обрахунку похибки свого методу.

Майкл Герман та інші [5], використовуючи ДПФ для знаходження градувань лінійки та відстані між ними у пікселях, знаходять відповідність пікселів до реальних розмірів об'єкта. Перевіривши свій метод на вимірюванні розмірів 30-ти зображень із давніми монетами вони отримали середню відносну похибку 1 % (0,34 px) та максимальну похибку – 6,67 % (1,4 px), що переважає результати методу ScaleSelect.

Аналіз досліджень дає змогу сформулювати такі гіпотези для подальших удосконалень методу ScaleSelect:

- попередня сегментація лінійки використовуючи штучні нейронні мережі дасть змогу зменшити кількість цифр, що не належать градуванням лінійки;
- виявлення градувань лінійки і відстані між ними з використанням ДПФ дає змогу отримати результати дещо кращі ніж у методі ScaleSelect, тому варто розглянути поєднання підходів;
- роздільна здатність фотографій відіграє надзвичайно важливу роль, необхідно оцінити метод з використанням зображень кращої якості.

За результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – удосконалено метод ScaleSelect, що дало змогу збільшити ймовірність визначення шкали лінійки на зображенні на 21,2 %.

Практична значущість результатів дослідження – удосконалений метод дає змогу збільшити ймовірність автоматичного розрахунку реальної площі рани на зображенні шкіри людини за використання цього методу для визначення відповідності пікселів на зображенні до сантиметрів.

Висновки / Conclusion

За результатами виконаної роботи удосконалено першу версію методу, що дало змогу збільшити ймовірність визначення шкали лінійки на 21,2 % (від 128 до 167 опрацьованих зображень із 184 загалом). Зокрема, дещо детальніше:

- проаналізовано альтернативні підходи;

- виправлено нестійкість методу для виявлення цифр на обернутих зображеннях;
- з'ясовано причини помилок у роботі зі зображеннями, що містять двоцифрові числа на лінійках, та виправлено їх;
- доповнено метод ScaleSense новими методами групування та проведено порівняння результатів;
- спрощено загальну структуру методу;
- розроблено вебсайт з демонстрацією методу: <https://scale-sense.streamlit.app/>

Отримано медіану похибки 4,2 % порівняно з ручним вимірюванням зі 167 зображень, а абсолютне значення 8 пікселів, що знаходиться у межах можливих похибок ручного вимірювання – [0-20] пікселів.

Визначено, що серед досліджених методів групування DBSCAN та використання медіани дають найкращі результати, при цьому DBSCAN дає змогу конфігурувати параметри, водночас, краще налаштувати метод під конкретні дані, а метод медіани є найменш ресурсозатратним та найшвидшим.

Сформульовано гіпотези щодо подальших удосконалень відповідних методів, які сприятимуть збільшенню точності вимірювання.

References

1. *A suite of minimal bounding objects*. (n.d.). Ww.mathworks.com. URL: <https://www.mathworks.com/matlabcentral/fileexchange/34767-a-suite-of-minimal-bounding-objects>.
2. Anil Singh Parihar, Gupta, M., Sikka, V., & Kaur, G. (2017). *Dimensional analysis of objects in a 2D image*. <https://doi.org/10.1109/icccnt.2017.8203937>
3. Bhalerao, A., & Reynolds, G. (2014). Ruler detection for autoscaling forensic images. *International Journal of Digital Crime and Forensics*, 6(1), 9–27. <https://wrap.warwick.ac.uk/63292/>
4. Deng, D. (2020). *DBSCAN Clustering Algorithm Based on Density*. 2020 7th International Forum on Electrical Engineering and Automation. URL: <https://www.semanticscholar.org/paper/DBSCAN-Clustering-Algorithm-Based-on-Density>.
5. Herrmann, M., Zambanini, S., & Kampel, M. (2009). Image Based Measurement of Ancient Coins. In *Computer Applications to Archaeology*. URL: <https://cvl.tuwien.ac.at/wp-content/uploads/2014/12/caa09.pdf>
6. Jocher, G., Stoken, A., Borovec, J., NanoCode012, ChristopherSTAN, Changyu, L., Laughing, tkianai, Hogan, A., lorenzo-mammanna, yxNONG, AlexWang1900, Diaconu, L., Marc, wanghaoyang0106, ml5ah, Doug, Ingham, F., Frederik, & Guilhen. (2020). *ultralytics/yolov5: v3.1 – Bug Fixes and Performance Improvements*. Zenodo. <https://zenodo.org/record/4154370#.ZHdLXqVBy3A>.
7. Jun, S. (2015). *Research on Image Measuring method based on monocular camera and Ruler*. Ww.atlantis-Press.com; Atlantis Press. <https://doi.org/10.2991/iemb-15.2015.30>
8. Konovalov, D., Domingos, J., Bajema, C., White, R. D., & Jerry, D. (2017). *Ruler Detection for Automatic Scaling of Fish Images*. Proceedings of the International Conference on Advances in Image Processing. URL: <https://www.semanticscholar.org/paper/Ruler-Detection-for-Automatic-Scaling-of-Fish-Konovalov-Domingos/ab9aa906580a18e2c4751d6d7d4b8bc5da5bfa77>.
9. Lukashchuk, B. (2022). Method of automatic identification of ruler scales on images to solve the problem of estimating the geometric dimensions of certain objects. <https://doi.org/10.1109/KhPIWeek57572.2022.9916468>
10. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You Only Look Once: Unified, Real-Time Object Detection*. ArXiv.org. <https://arxiv.org/abs/1506.02640>
11. Telahun, M., Sierra-Sossa, D., & Elmaghraby, A. S. (2020). Heuristic Analysis for In-Plane Non-Contact Calibration of Rulers Using Mask R-CNN. *Information*, 11(5), 259. <https://doi.org/10.3390/info11050259>
12. Tesseract: an Open-Source Optical Character Recognition Engine | Linux Journal. (n.d.). Ww.linuxjournal.com. <https://www.linuxjournal.com/article/9676>
13. Ueda, K., Baba, T., Nakagawa, Y., & Amano, K. (2005). *Detection of Scale Intervals in Digital Images*. <https://doi.org/10.1109/ICDE.2005.211>
14. Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *ArXiv:2207.02696 [Cs]*. URL: <https://arxiv.org/abs/2207.02696>.
15. Xue, Z., Yu, K. J., Pearlman, P. C., Chen, T.-C., Chun Hung Hua, Kang, C. G., Chien, C.-Y., Ming Hsui Tsai, Wang, C.-P., Chaturvedi, A. K., & Sameer Antani. (2022). *Extraction of Ruler Markings For Estimating Physical Size of Oral Lesions*. <https://doi.org/10.1109/icpr56361.2022.9956251>

B. S. Lukashchuk

Ukrainian National Forestry University, Lviv, Ukraine

SCALESENSE: AN IMPROVED METHOD OF AUTOMATIC IDENTIFICATION OF RULER SCALES ON IMAGES TO SOLVE THE PROBLEM OF ESTIMATING THE GEOMETRIC DIMENSIONS OF CERTAIN OBJECTS

The identification of pixel correspondence to real-world units in an image has been identified as a crucial step for further analysis of the geometric characteristics of objects in the image. An enhanced iterative method for the automatic determination of ruler scale on the image based on digit recognition is proposed. The developed results have been tested on images from the biomedical field. Applicability of this method has been established in other fields, including those mentioned in the analyzed related research such as forensics, veterinary medicine, museum studies, etc. The first version of the method has been analyzed, and significant algorithm flaws have been identified and corrected, particularly enabling more efficient processing of images with rulers containing two-digit numbers. A new digit recognition model – YOLOv7 – has been trained and integrated into the method, effectively correcting issues with inverted images. The number of images processed by the method has been increased to 90 %. Alternative methods of grouping test segments, which represent the method final step, have been investigated. Experimental data has been processed using various grouping methods, in particular, DBSCAN clustering, median, modified z-score, and interquartile range, and errors compared to manually measured values. In the case of using median and DBSCAN clustering methods, a median error of 4.2-4.4 % was achieved, while in specific configurations of the DBSCAN method, the error was 3.1-3.7 %. A web page with a demonstration version of the method has been developed to attract more data and practical applications in solving real problems. An option to choose one of four grouping methods, their parameters (for DBSCAN), and images for method testing has been provided. The visualization of a random test segment on the uploaded image has been implemented for a more straightforward subjective evaluation of the obtained result by the user. Future proposals include as follows: conduct additional training of digit detection models; implement support for processing numbers with any number of digital symbols; consider some of the approaches used in similar methods; implement REST API for invoking the method.

Keywords: digital image processing; wound measurement; reference object; object detection; digit recognition; YOLO; computer simulation; geometric analysis; biomedical imaging.