

К.ф-м.н. доц. Третиник В.В., магістрант Ковріжкін О.С.

**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**

КОНЦЕПТУАЛЬНА МОДЕЛЬ ГІБРИДНОЇ СИСТЕМИ ДЛЯ ДЕТЕКЦІЇ ТА КЛАСИФІКАЦІЇ КОНТУРІВ НА ОСНОВІ ТРАДИЦІЙНИХ МЕТОДІВ ТА ГЛИБОКОГО НАВЧАННЯ

Abstract

Violeta Tretynyk, Associate Professor; Oleksandr Kovrizhkin, student
A hybrid system for contour detection and classification based on traditional methods and deep learning

This paper addresses the problem of balancing computational performance and accuracy in computer vision systems for geometric shape detection. Traditional methods (e.g., Canny edge detection, Hough Transform) are fast but highly sensitive to image noise. Conversely, deep learning models (e.g., U-Net) are robust to noise but are computationally expensive. This work proposes the conceptual and structural model of an adaptive hybrid system. The system features an "adaptive router" that analyzes image complexity and dynamically selects one of two processing pipelines: a "fast path" (Canny-Hough-SVM) for simple images or a "robust path" (U-Net) for complex, noisy data. This approach aims to optimize the trade-off between speed and accuracy for heterogeneous data streams.

Вступ

Виявлення та класифікація геометричних контурів є фундаментальною задачею комп'ютерного зору, що має критичне значення для таких галузей, як промислова автоматизація (контроль якості), робототехніка (навігація) та медична візуалізація. Існуючі підходи до вирішення цієї задачі можна умовно поділити на дві категорії [1].

З одного боку, це традиційні (класичні) методи, такі як детектор країв Canny [2] у поєднанні з перетворенням Хафа (Hough Transform). Їхньою беззаперечною перевагою є висока швидкість виконання та низькі вимоги до обчислювальних ресурсів (CPU). Однак, як показує практика, ці методи є "крихкими" — їхня точність катастрофічно падає за наявності на зображенні шуму, поганого освітлення або часткового перекриття об'єктів.

З іншого боку, сучасні методи на основі глибокого навчання (Deep Learning), зокрема архітектури семантичної сегментації, як-от U-Net [3], демонструють надзвичайну робастність. Вони здатні виділяти коректні контури навіть на зображеннях з високим рівнем шуму, ігноруючи

перешкоди. Платою за таку точність є високі обчислювальні витрати: навчання таких моделей вимагає великих анотованих наборів даних, а їхня робота (inference) часто потребує потужних графічних процесорів (GPU), що робить їх надлишковими для обробки простих, незашумлених зображень.

Таким чином, виникає актуальна науково-прикладна проблема: існує розрив між швидкими, але ненадійними традиційними методами, та надійними, але обчислювально "важкими" методами глибокого навчання. Жоден з "чистих" підходів не пропонує оптимального рішення для систем, що мають працювати з гетерогенними даними (сумішшю простих і складних зображень) в умовах, близьких до реального часу.

Постановка задачі

Метою даної роботи є підвищення загальної ефективності процесу детекції та класифікації контурів геометричних фігур шляхом розробки концептуальної та структурної моделі гібридної системи з адаптивною логікою.

Для досягнення поставленої мети необхідно вирішити наступні задачі:

1. Провести аналіз переваг та недоліків традиційних (Canny, Hough, SVM) та глибоких (U-Net) методів у контексті задачі.
2. Розробити концептуальну модель гібридної системи, що включає два окремі конвеєри обробки ("швидкий" та "робастний") та модуль прийняття рішень ("адаптивний маршрутизатор").
3. Сформулювати математичне та методичне забезпечення для ключових компонентів системи.
4. Обґрунтувати вибір програмних засобів для реалізації прототипу системи.

Концептуальна модель гібридної системи

Для вирішення вищезазначеної проблеми пропонується розробка гібридної системи, архітектура якої базується на принципі адаптивної маршрутизації (adaptive routing). Замість того, щоб обробляти всі зображення єдиним "важким" конвеєром U-Net, система спочатку проводить швидкий аналіз складності вхідних даних. На основі цього аналізу приймається рішення про вибір одного з двох шляхів обробки.

1. "Швидкий шлях" (Fast Path). Призначений для обробки простих, висококонтрастних та незашумлених зображень. Він використовує обчислювально легкий традиційний конвеєр: а. Застосування детектора Canny для отримання бінарної карти країв. б. Використання перетворення Хафа для виявлення примітивів (ліній). с. Формування вектору ознак фіксованої довжини (напр., на основі моментів H_u , кількості ліній,

співвідношення сторін). d. Класифікація вектору ознак за допомогою "легкої" моделі машинного навчання (наприклад, SVM або Random Forest).

2. "Робастний шлях" (Robust Path). Призначений для складних випадків (зашумлені, низькоконтрастні зображення, накладання фігур). a. Зображення подається на вхід натренованої моделі семантичної сегментації (U-Net). b. Модель генерує чисту попиксельну маску контурів, ігноруючи шум. c. Отримана маска обробляється (напр., апроксимується полігоном) для визначення типу фігури.

3. "Адаптивний маршрутизатор" (Adaptive Router). Це ключовий керуючий компонент, що активується першим. Він виконує швидку евристичну оцінку якості вхідного зображення (наприклад, аналіз дисперсії Лапласіана або аналіз кількості дрібних, фрагментованих контурів, отриманих від Canny). Якщо зображення ідентифіковано як "просте", воно відправляється на "швидкий шлях". Якщо воно ідентифіковано як "складне", система активує "робастний шлях".

Математичне та програмне забезпечення

Для реалізації описаної моделі було обрано відповідне математичне та програмне забезпечення. У "швидкому шляху", для реалізації детектора Canny та перетворення Хафа, а також для виділення ознак (моменти H_u , площа, периметр) було обрано бібліотеку OpenCV. Для модуля класифікації було обрано два методи для порівняльного аналізу: метод опорних векторів (SVM) та ансамблевий метод випадковий ліс (Random Forest).

Метод опорних векторів (SVM) [4] вирішує задачу оптимізації, знаходячи оптимальну гіперплощину (w, b) , що максимізує зазор між класами:

$$\min_{w,b} \frac{1}{2} \|\omega\|^2 \quad \text{за умови} \quad y_i(w \cdot x_i + b) \geq 1, \quad \forall i = 1, \dots, N$$

де x_i – вектор ознак, y_i – мітка класу. Для нелінійних даних використовується "ядерний трюк" (kernel trick), зокрема, радіальна базисна функція (RBF):

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2)$$

Випадковий ліс (RF) є ансамблем з K дерев рішень, що навчаються на бутстреп-вибірках. Фінальне рішення приймається шляхом "голосування" (majority vote):

$$\hat{y} = \operatorname{argmax}_{c \in C} \left(\sum_{k=1}^K I(y_k = c) \right)$$

де $\hat{y}_k$ – прогноз k -го дерева, c – клас, I – індикаторна функція. Обидва методи реалізуються за допомогою бібліотеки Scikit-learn, оскільки вони демонструють високу точність на структурованих векторах ознак при мінімальних обчислювальних витратах на етапі прогнозування [4].

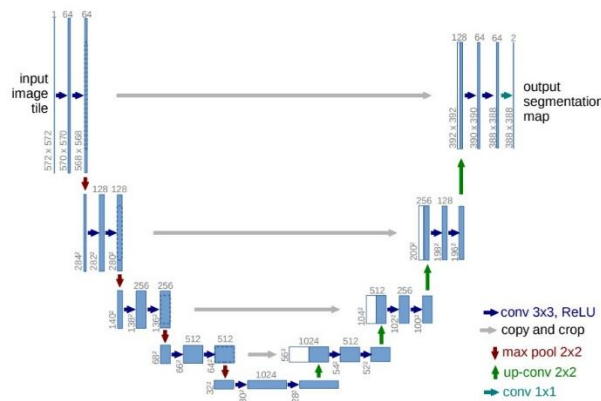


Рис. 1. Архітектура U-net (приклад для зображення 32×32 пікселі) [3].

Для "робастного шляху" було обрано архітектуру U-Net (Рис. 1). Для її реалізації використовується фреймворк TensorFlow з Keras API. Для ефективного навчання моделі на задачі виявлення тонких контурів (що є типовою проблемою класового дисбалансу, де пікселі "фону" значно переважають пікселі "контуру"), стандартна функція втрат (бінарна перехресна ентропія) доповнюється коефіцієнтом Дайса (Dice Loss). Такий підхід дозволяє моделі краще фокусуватися на коректній сегментації цільових пікселів контуру.

Для об'єднання всіх компонентів в єдину систему використовується мова програмування Python, яка виступає як інтеграційна платформа для бібліотек OpenCV, Scikit-learn та TensorFlow.

Висновки

У ході даного дослідження було вирішено проблему вибору між швидкістю та робастністю при детекції геометричних контурів. Проведено аналіз, який показав фундаментальні обмеження "чистих" традиційних підходів (чутливість до шуму) та "чистих" глибоких підходів (обчислювальна надлишковість). Запропоновано концептуальну модель гібридної системи з адаптивною маршрутизацією, що поєднує "швидкий шлях" (Canny-Hough-SVM) та "робастний шлях" (U-Net). Обґрунтовано вибір математичних методів (Canny, SVM, U-Net) та програмних засобів (Python, OpenCV, Scikit-learn, TensorFlow), що є методичною основою для реалізації прототипу. Запропонована архітектура дозволяє створити систему, що в середньому працює значно швидше, ніж "чисті" рішення на базі U-Net, але при цьому зберігає їхню високу точність та робастність при обробці складних зашумлених зображень.

Перспективи подальших досліджень включають удосконалення модуля "адаптивного маршрутизатора" шляхом заміни евристичних методів

на легку згорткову нейронну мережу, а також оптимізацію архітектури U-Net для "робастного шляху" з метою її прискорення.

Література

1. Huang, Q., & Huang, J. (2023). Comprehensive review of edge and contour detection: from traditional methods to recent advances. *Displays*, 79, 102462.
2. Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6), 679-698.
3. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234-241). Springer.
4. Gonzalez, R. C., & Woods, R. E. (2009). *Digital Image Processing* (3rd ed.). Pearson.
5. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2961-2969).