

ЗАДАЧА КЛАСИФІКАЦІЇ БІОЛОГІЧНИХ ОБ'ЄКТІВ

ã Грицик В., Пелих Н., 2009

Описуються деякі алгоритми і методи, які були застосовані авторами для аналізу зображення клітин при попередній обробці зображень, знаходження і виділення клітин, їх класифікації.

In the article are described algorithms and methods that were applied for analysis of images, pattern recognition, and classification.

1. Вступ

Переглянувши перелік пропозицій програмних продуктів з класифікації клітин, які сьогодні пропонують провідні світові фірми (Цейс, Відеотест, Олімпус...) [1] ми побачимо потребу суспільства в автоматизації систем класифікації. Однак, розроблення програм для проведення досліджень у цій галузі – одна з найскладніших задач. Це видно з тих самих пропозицій: провідні розробники пропонують класифікацію лише в “ідеальних” умовах – без завод і з однорідним полем уваги. Чому ж накладають такі умови?

Розглянемо проблему обробки зображень (розпізнавання та класифікації) на прикладі задачі класифікації ракових клітин.

2. Постановка задачі

Мета роботи –ослідити зображення клітин, отриманих за допомогою мікроскопа заданим методом і реалізувати адаптивну програму для їх дослідження [2, 5].

Алгоритм та відповідна програма мають виконувати такі операції:

1. Знаходити та виділяти на зображенні усі можливі клітини на полі уваги;
2. Здійснювати класифікаційний аналіз кожної клітини: розпізнати форму клітини (кругла чи подовгаста), знайти співвідношення кольорів, зробити висновок за отриманими результатами;
3. Розрахувати співвідношення круглих клітин до подовгастих.

3. Застосований підхід

При автоматичному аналізі візуальних об'єктів однією з найбільших і найважливіших проблем є підбір такого алгоритму обробки, який би підходив для різних класів зображень [3]. Труднощі полягають у тому, що методика отримання зображення передбачає обробку „умовно” однакових за густиною, масштабом, вертикальним розрізом (відповідно) образів [4]. „Умовно” означає, що для електронної обробки при прийнятті рішення щодо приналежності до об'єкта, фону чи завади важливий кожний піксел, а людина інтегрує миттєво. Тому зазвичай підбирають алгоритм для одного класу.

Розробляючи алгоритм розпізнавання об'єктів на полі уваги (рис. 1), розв'язували такі задачі:

- якість зображення (від якості залежатиме перепад між фоном та клітинами, а також наявність сторонніх перешкод);
- тип/клас зображення;
- за якими складовими кольору клітин відбуватиметься перетворення;
- визначення приналежності до того чи іншого кольору.
- можуть зустрічатися конгломерати (злиті клітини) і потрібно вирішити, чи можливо їх розділити;

- визначення форми клітини;
- сторонні перешкоди, що можуть бути частиною клітини.

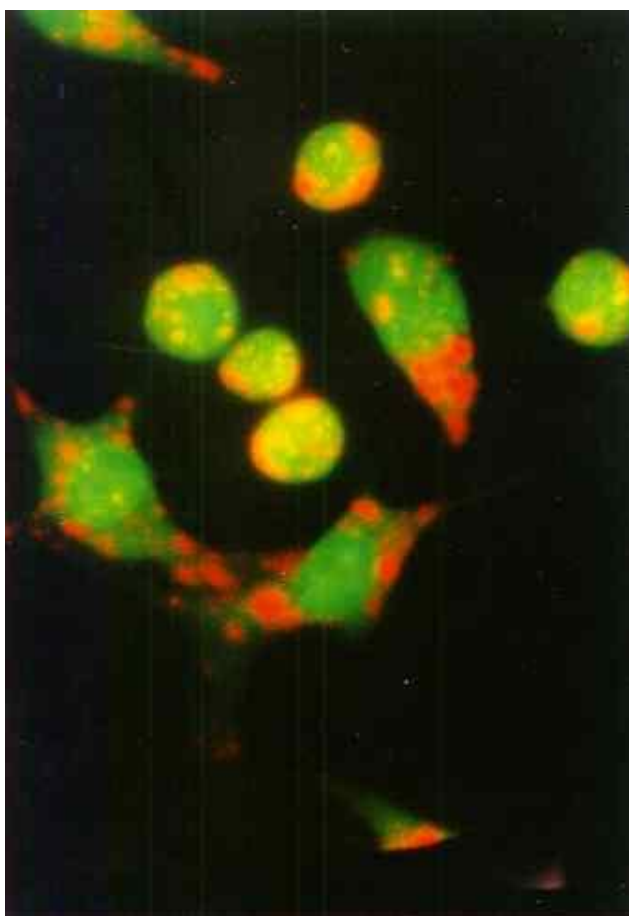


Рис. 1. Приклад поля уваги з різними класами об'єктів на ньому

У розглядуваній нами задачі важливими є дві ознаки зображень: форма клітини та її колір. омінуючі кольори – червоний і зелений.

Усі основні ознаки для конкретного типу клітин наведено у таблиці.

Тип клітини	Форма клітини	Місткість кольору, %	
		Червоний	Зелений
Жива клітина	Подовгаста	-	-
Вмирає	Кругла	75 – 100	0 – 25
Розмножується		50 – 75	25 – 50
Жива клітина		Усі інші	

Для знаходження клітин на зображенні потрібно це зображення привести до стандартизованого вигляду. Приймемо за стандартні ті зображення, які мають однаковий вигляд. Наприклад, чорні об'єкти на білому фоні. Саме до такого вигляду зображень приводитиме нижче описаний алгоритм.

У цьому алгоритмі використовується два поняття – густина і перепад.

Густина – це кількість пікселів певної яскравості.

Перепадом вважатимемо значну різницю між густинами яскравостей.

Для прикладу взято вхідне зображення, представлене на рис. 2, а. Як бачимо, фон займає більшу частину зображення і є однорідним. Клітини яскравіші і значно відрізняються від фону.

Отже, задача полягає в знаходженні верхньої межі фону за яскравістю, тобто вище цієї межі можуть знаходитись корисні об'єкти.

Візьмемо масив із 255 елементів. Порядковий номер кожного із елементів становить одиницю яскравості за цим кольором. Значення елементів масиву і визначатиме густину одиниці яскравості. Оскільки фон є однорідним і темним, то насиченість пікселів буде найбільшою у першій чверті масиву. Після одержання такого масиву для зручності і простоти розіб'ємо його на 25 частин. Отримаємо масив із 25 елементів, значення яких визначається за формулою:

$$E_i^* = \sum_{j=i*10}^{(i+1)*10} E_j \quad i = \overline{0,24},$$

де E_i^* – елементи нового масиву, E_j – елементи вхідного масиву.

Для знаходження перепаду потрібно знайти максимальне значення густини. Задача зводиться до пошуку максимального елемента у новоутвореному масиві. Цей елемент є початком просування масивом для знаходження перепаду. В цьому прикладі за значний перепад було взято різницю між сусідніми густинами в 3 рази. Фіксуємо елемент масиву, що в 3 рази менший від максимального. Якщо наступний елемент також значно відрізняється від фіксованого, то фіксуємо його. В протилежному випадку можемо вважати, що верхню межу фону знайдено.

У результаті однозначно знаходимо фон і об'єкти, а саме: всі пікселі, яскравіші за знайдену межу – це вірогідні клітини, темніші – фон. Результатом буде матриця значення елементів якої 1 або 0, відповідно точка клітини – 1 чи фону – 0.

Графік перепаду густини за яскравістю представлено на рис. 3. Оскільки важлива роль у цьому зображенні належить червоному і зеленому кольорам (на рис.3 відповідно червона і зелена криві), то перетворення велося за цими складовими. Верхню межу фону за складовою R взято 50 одиниць по яскравості, за складовою G – 40 одиниць.

Верхню межу фону у вищеописаному алгоритмі знаходимо в адаптивному режимі.

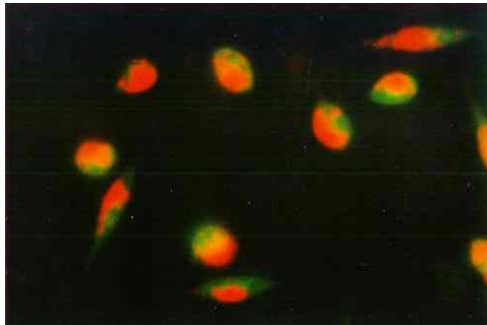


Рис. 2, а. Вхідне зображення

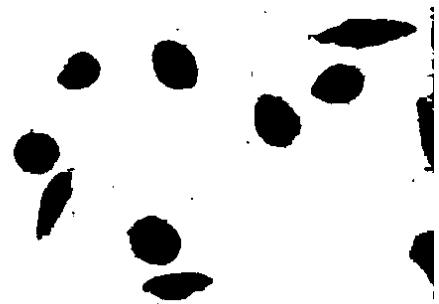


Рис. 2, б. Результат обробки

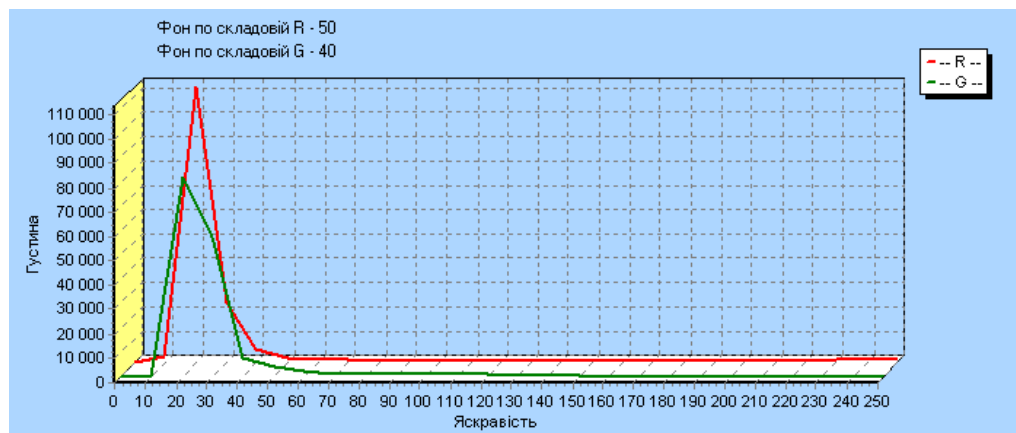


Рис. 3. Графік перепаду густини за яскравістю

Результат такого перетворення представлено на рис. 2, б де можемо спостерігати чітке виділення клітин на фоні. Також існують сторонні перешкоди, які будуть усунені при наступних кроках аналізу.

Як вже було вище сказано, дуже важливе значення для класифікації цього типу клітин має колір. Потрібно уважно визначити співвідношення кольорів, від якого залежатиме кінцевий результат.

Добре відомо, що при змішуванні червоного і зеленого кольорів утворюються різні відтінки жовтого. При класифікації задіяні лише червоний і зелений. Тому потрібно встановити чітку межу.

Межу знаходимо так: визначаємо яскравості червоного і зеленого кольорів і дивимось – якщо червоний переважає зелений, то відповідно вибираємо червоний, у протилежному випадку – зелений.

Результат представлено на рис. 4 (вхідне зображення взято з рис. 2, а).

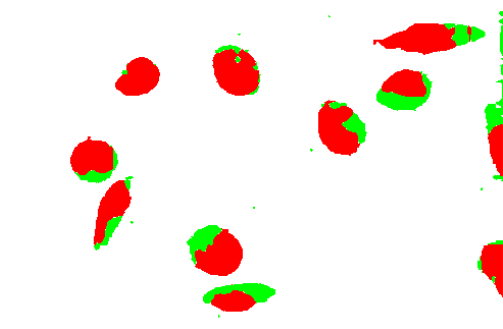


Рис. 4. Визначення яскравості кольорів

Добре виражені ці дві складові кольору визначаються однозначно. Коли проявляється жовтий колір, існує ймовірність неправильного віднесення до того чи іншого кольору, але це трапляється рідко. Крім того, вираження жовтого кольору також трапляється рідко, тому на результат це майже не впливає.

4. Висновок

У роботі здійснено постановку задачі та запропоновано алгоритм для розпізнавання та класифікації візуальних об'єктів – клітин, на полі уваги; наведено приклади реалізації обробки та розпізнавання і класифікації складних об'єктів. Проведено відповідний аналіз.

Важливою умовою, яка виконана, є відсутність попередньої обробки зображень, що пов'язано з недопустимим для цього класу задач послабленням інформаційних параметрів при застосуванні методів візуального покращання.

1. Грицик В.В., Влах М.А. Технічні та програмні засоби розпізнавання та аналізу зображень складних біологічних об'єктів. – Львів: ІТІС. 2005. Т.8. – №1. – С.17–28. 2. Фотографії та методика аналізу для побудови алгоритму та системи автоматизації експерименту, надані членом-кореспондентом НАН України, д.б.н. Р.С. Стойкою та працівником інституту біології клітини (м.Львів) О.Ю. Ключівською. 3. Грицик В.В. Оцінка якості передавання і комп'ютерна обробка даних образів // Доповіді Національної Академії наук України. – 2008. – №9. – «Інформатика та кібернетика». – С.43–48. 4. Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. Цифровая обработка изображений в информационных системах. – Новосибирск. – 2000. – 168 с. 5. Форсайт Д., Понс Ж. Компьютерное зрение; Современный подход. – М. –СПб. –К.: Изд. «Вильямс», 2004. – 926 с.