

ЗМЕНШЕННЯ ПРОСТОРУ ТЕКСТУРНИХ ОЗНАК ГІСТОЛОГІЧНИХ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ МЕТОДУ ГОЛОВНИХ КОМПОНЕНТ

Abstract. Texture is an important feature for histological image classification. In this paper, we analyze four textural features that are used to classify objects in histological image. A reduction of feature space using the method of principal components is made.

Актуальність.

Для діагностування злякисних новоутворень використовуються цифрові гістологічні зображення (ГЗ). Однією із задач аналізу ГЗ є класифікація тканин різних типів. Класифікацію тканин можна представити у вигляді текстурної сегментації. Результатом такої сегментації є розбиття зображення на однорідні області, що відповідають різним класам об'єктів. В найпростішому випадку класами є міжклітинна рідина (фон), цитоплазма клітин, ядра клітин, міжклітинна з'єднувальна тканина, перерізи судин та ін. Об'єктом наших досліджень є гістологічні зображення.

Для класифікації об'єктів на гістологічних зображення застосовують різноманітні статистичні текстурні ознаки [1-4]. Програмні засоби аналізу ГЗ, наприклад, CellProfiler обчислюють великий набір ознак мікрооб'єктів, таких як розмір, форма і зафарбовування, інтенсивність пікселів, текстура та ін.

Текстурна сегментація зображень володіє відомою складністю [5]. Це пояснюється статистичною різноманітністю текстур, великою кількістю розроблених ознак та параметризацією алгоритмів. Ці труднощі долаються ручним відбором характерних ознак. Тому актуальною є задача витягнення з вхідних багатовимірних даних набору ознак, інформативних з точки зору подальшого вирішення задачі класифікації. Тому предметом даної роботи є методи зменшення розмірності простору текстурних ознак.

Розглядається задача зменшення простору текстурних ознак з метою отримання компактної множини інформативних ознак. Формування простору текстурних ознак має низку істотних недоліків. Велика кількість ознак призводить до високих часових витратах на обробку даних, великих обсягів пам'яті, необхідних для зберігання інформації, а також до необхідності збору великої кількості прецедентів для впевненого відновлення прихованих залежностей в багатовимірному просторі. Іншим серйозним недоліком отриманого простору ознак є той факт, що близькі в просторі ознак об'єкти не відповідають одним і тим же класам. Виконання гіпотези компактності є одним з основних вимог для більшості методів класифікації. Методи зменшення розмірності даних дозволяють отримувати представлення вибірок

в маломірних просторах, що володіють рядом хороших властивостей.

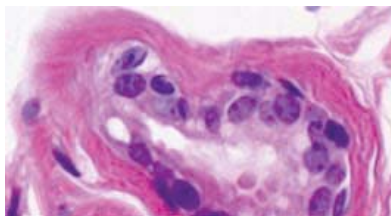
Постановка задачі

Задача полягає в зменшенні розмірності простору текстурних ознак гістологічних зображень з допомогою методу головних компонент.

2. Ознаки гістологічних і цитологічних зображень

Текстура в загальному випадку – це функція просторової зміни розподілу інтенсивностей пікселів. Зображення текстури [6] просторово однорідне, і зазвичай містить повторювані структури, часто з деякими випадковими змінами (наприклад, зміною позиції, орієнтації або кольору). При описі текстури виділяють такі її властивості: однорідність, щільність, грубість або зернистість, шершавість, регулярність, лінійність, напрямленість. Така різноманітність ознак зумовлює появу великої кількості моделей текстури.

Об'єкти на ГЗ розбиті на три класи: фон, ядро клітини, з'єднувальна тканина. На рис. 1,а наведено приклад ГЗ та відповідне йому класифіковане зображення на рис.1,б. Чорним кольором позначено фон, сірим позначено з'єднувальну тканину, білим – ядра клітин.



а) Вихідне зображення

б) Розбиття на класи

Рис. 1. Приклад класифікації гістологічного зображення

Для класифікації областей використовуємо текстурні ознаки на основі матриць розподілу рівнів сірого. Обчислимо такі ознаки: контраст, несхожість, однорідність, другий кутовий момент (ДКМ). В таблиці 1 представлено характеристики текстурних ознак для кожного класу об'єктів.

Таблиця 1

Вихідні текстурні ознаки

	Фон		Тканина		Ядро	
	Середнє	СКВ	Середнє	СКВ	Середнє	СКВ
Контраст	12,32	61,04	10,32	47,94	19,999	81,35
Несхожість	0,7593	2,88	0,82196	2,5363	1,147	3,58
Однорідність	0,0246	0,093	0,0315	0,0841	0,0175	0,0515
ДКМ	0,004518	0,0253	0,00329	0,009556	0,00201	0,00575

1. Зменшення простору ознак

Метод головних компонент (principal component analysis, PCA) є найпростішим методом зменшення розмірності в даних. Ідея методу полягає в пошуку в вихідному просторі гіперплощини заданої розмірності з подальшим проектуванням вибірки на дану гіперплощину. При цьому вибирається та гіперплощина, помилка проектування даних на яку є мінімальною в відносно суми квадратів відхилень (СКВ).

Нехай є деяка вибірка об'єктів $X = \{\mathbf{x}_n\}_{n=1}^N$, $\mathbf{x}_n \in R^D$. Задача зменшення розмірності полягає в отриманні представлення цієї вибірки в просторі меншої розмірності $T = \{\mathbf{t}_n\}_{n=1}^N$, $\mathbf{t}_n \in R^d$. Розмірність $d < D$, але в окремих випадках d може і збігатися з D . Зменшення розмірності в описі даних забезпечує:

- зменшення обчислювальних затрат при обробці даних;
- зменшення часу навчання, оскільки чим менше кількість ознак, тим менше потрібно об'єктів для певненого відновлення прихованих залежностей в даних і тим вище якість відновлення подібних залежностей;
- стиснення даних для більш ефективного зберігання інформації;
- візуалізація вибірки багатомірних даних у дво-, тривимірному просторі;
- витягнення нових ознак отримані в результаті перетворення, може надавати значний внесок при подальшому вирішенні задач розпізнавання.

3. Експериментальні дослідження

Для експериментальних досліджень використано мову програмування Python, бібліотеки алгоритмів машинного навчання sklearn, бібліотеку роботи з матрицями numpy. Узагальнений алгоритм зменшення простору ознак такий:

- нормування даних, оскільки текстурні ознаки мають різні діапазони значень;
- отримання власних векторів і власних значень з коваріаційної матриці або кореляційної матриці;
- сортування власних значень в порядку зменшення і вибірка d власних векторів, які відповідають d найбільшим власним значенням;
- побудувати проекційну матрицю W з обраних d власних векторів;
- перетворити вихідний масив даних X з допомогою W , щоб отримати d -мірний підпростір ознак T .

Експериментальні дані містять вимірювання текстурних ознак для кожної точки тестового гістологічного зображення. Дані розбиті на три класи: фон, ядро клітини, з'єднувальна тканина. Вихідний набір даних представлено у вигляді таблиці 3.1. Кожен експеримент проводиться для ковзаючого вікна розміру 9 пікселів та вектору зсувів 4 пікселі.

Таблиця 2

Фрагмент тестових даних

№	Контраст	Несхожість	Однорідність	ДКМ	Клас
1	10,12	0,533	0,022	0,0035	ядро

В таблиці 3 предствлено ознаки класів у двовимірному просторі ознак. На рис. 2 наведено результуючий двовимірний простір ознак точок зображення.

Таблиця 3

Двовимірний простір текстурних ознак

	Фон		Тканина		Ядро	
	Середнє	СКВ	Середнє	СКВ	Середнє	СКВ
Ознака 1	0.0307044	1,434	0,03738	0,91787	-0,204712	0.8214
Ознака 2	0.0162723	1,071	-0,02395	0,87924	0,053651	1.3928

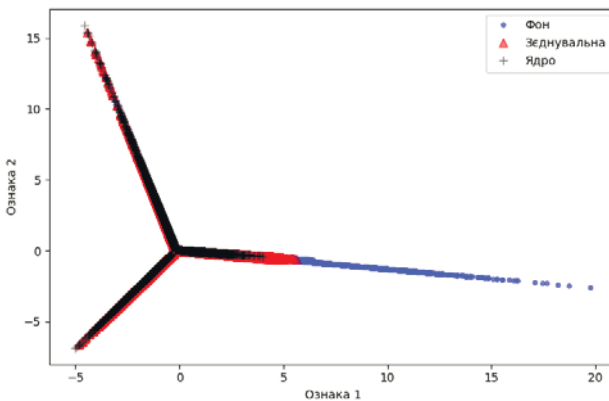


Рис. 2 Двовимірний простір текстурних ознак

Як видно з графіку частина пікселів всіх трьох класів накладається. Це зумовлено принципом розрахунку текстурних ознак в ковзаючому вікні, що приводить до ефекту розмиття границь окремих об'єктів на зображенні. Проте в області обмеженій [0; 20] по першій ознаці і [0;-3] по другій ознаці вдалося досягти розбиття простору на три класи.

Висновки

Досліджено 10 зображень та 4 текстурних ознаки на основі матриць розподілу рівнів сірого. Здійснено зменшення простору текстурних ознак до двовимірного з допомогою методу головних компонент.

1. *Rajkovic N.* Comparison of Monofractal, Multifractal and gray level Co-occurrence matrix algorithms in analysis of Breast tumor microscopic images for prognosis of distant metastasis risk / N. Rajković, D. Kolarević, K. Kanjer, N. T. Milosević, D. Nikolić-

- Vukosavljević, M. Radulovic // Biomedical Microdevices - Springer, 2016. - Vol. 5, Num. 18. - P. 83
2. Saito A. A novel method for morphological pleomorphism and heterogeneity quantitative measurement: Named cell feature level co-occurrence matrix / A. Saito et al. // Journal of Pathology Informatics – 2016 - Vol. 7, Num. 36
3. Mohanty A. K. Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram / A. K. Mohanty, S. Beberta, S. K. Lenka // International Journal of Engineering Research and Applications - Vol. 1, Issue 3, pp.687-693
4. Bieri M. "Quantitative analysis of Alzheimer plaques in mice using virtual microscopy" / M. Bieri, A. Wethmar, N. Wey. - First European Workshop on Tissue Imaging and Analysis, - pp 31-38.
5. Singh M. SAR Image Classification Using PCA and Texture Analysis / Mandeep Singh, Gunjit Kaur // Information Technology and Mobile Communication: International Conference, AIM 2011, Nagpur, Maharashtra, India, April 21-22, 2011. Proceedings - Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. – P. 435-439.
6. Мельник Г. М. Інформаційна технологія аналізу структурних текстур для опрацювання зображень ауто- та ксеногенних тканин / Г. М. Мельник // Вісник Хмельницького національного університету - 2014. - № 6 (217). - С. 132-141
7. Berezsky O.M. "Vision-based medical expert system" / O.M. Berezsky, K.M. Berezska, Yu.M. Batko, G.M. Melnyk. - 6th International Scientific and Technical Conference "Computer Sciences and Information Technologies"(CSIT'2011, Lviv, Ukraine, November 16-19), - pp 49-50
8. Berezsky O.M. Segmentation of Cytological and Histological Images of Breast Cancer Cells / O.M. Berezsky, Yu.M. Batko, G.M. Melnyk, S.O. Verbovyu, L. Haida. - Proceedings of the 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2015), 24-26 September 2015, Warsaw, Poland. – Warsaw, 2015. – V.1. – P. 287-292

Поступила 26.09.2016р.

УДК 004.2

І.Г. Цмоць, м.Львів

І.В. Ігнатєв, А.Р. Боднар, м.Тернопіль

ПАРАЛЕЛЬНІ АЛГОРИТМИ ТА НВІС-СТРУКТУРИ ДЛЯ МЕДІАННОЇ ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ В РЕАЛЬНОМУ ЧАСІ

Abstract. Especially effective is the median filter in image processing video stream to identify objects. This median filtering must be performed quickly, which is very advisable to carry on GPU. Existing hardware solutions in the public domain not found. It is therefore necessary to develop a median filter and its graph algorithm which would be implemented on the GPU that have made it possible to speed up the filtration results in real time.