

МЕТОДЫ И АЛГОРИТМЫ ПОСТРОЕНИЯ КОМПЬЮТЕРНОЙ СИСТЕМЫ ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ ЧЕЛОВЕКА НА ОСНОВЕ ИММУНОГИСТОХИМИЧЕСКОГО АНАЛИЗА

Батлук Н.С.

Донецкий национальный технический университет

Наиболее актуальным вопросом медицины сегодня является как можно более ранняя диагностика патологий с целью придания врачам времени для проведения исследований и правильного лечения. А наиболее ранняя диагностика – диагностика на низшем уровне организации, а именно на клеточном. Но, как известно, самые важные задачи сложно выполнимы. Каждый день в мире появляются новые методики по обработке изображений гистологических срезов и изучению морфологических признаков разнообразных клеток, совершенствуются микроскопы и программное обеспечение для работы с изображениями на персональных компьютерах. Работа ученых биохимиков также не заставляет себя долго ждать, и непрерывно продолжается процесс усовершенствования иммуногистохимического анализа - метода выявления и точной локализации того или иного клеточного или тканевого компонента (антигена) *in situ* с помощью иммунологических и гистохимических реакций[1]. Основной проблемой при создании автоматизированной системы диагностики заболеваний человека на основе иммуногистохимического анализа является точность выявления клетки и определения ее параметров. Поэтому разработка такой КС, которая сможет помогать врачам в постановке максимально точного диагноза, сокращая при этом затраты времени, является важной и актуальной задачей.

Целью данной работы является разработка автоматизированной подсистемы диагностирования различных заболеваний человека по изображениям гистологических срезов, полученных с помощью цифрового микроскопа и проведения иммуногистохимического окрашивания. Для достижения поставленной цели необходимо решить следующие задачи:

- поиск и оценка точности существующих методов обработки изображений гистологических срезов;
- разработка алгоритма определения патологических признаков клеток на изображениях гистологических срезов;
- выбор наиболее подходящего метода для реализации поставленной задачи;
- реализация выбранных методов на ЭВМ.

Основным нововведением работы является использование различных алгоритмов и методов обработки изображений гистологических срезов для наиболее точного анализа, используются методы зарубежных ученых, ранее не применяемых в нашей стране. Сравнение полученных результатов позволит сделать вывод о состоянии исследуемого участка ткани и оценить точность и скорость каждого из подходов. Разработанное ПО предполагается использовать в клинической лаборатории «Новая диагностика» при ИНВХ им.Гусака для диагностирования заболеваний человека, в основном новообразований различных органов.

Среди основных методов, которые были использованы в ходе разработки программного обеспечения для обработки изображений гистологических срезов, метод отсечения по цвету, метод Прюита, пороговая сегментация, метод Лапласа, морфологические операции и другие.

1. Адаптивный метод установления порогов и сегментация

Сегментированное изображение вычисляется через порог на градиентном изображении, который вычисляется из первоначального изображения. Градиент изображения вычисляется с использованием метода Прюита. Первоначальное изображение фильтруется с помощью двух масок, по одной на каждое направление, где каждая маска усиливает различия в соответствующем направлении.

Фильтр Прюита задается матрицей

$$M_1^{prewitt} = \frac{1}{3} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}. \quad (1)$$

Фильтры, приближающие оператор производной по y -координате $\frac{\partial}{\partial y}$, получаются путем транспонирования матриц. Это приводит к двум новым изображениям, I_x и I_y , каждое из которых является границей пиксельных вариаций в X и Y направлениях, соответственно. После этого, мы положим, что

$$I_{mag} = (I_x^2 + I_y^2)^{\frac{1}{2}} \quad (2)$$

для расположения каждого пикселя. Пиксели с градиентом значения выше определенного порога β теперь будут помечены, как области переднего плана. Решающая роль в шаговой сегментации заключается в том, чтобы автоматически найти порог β . Градиентные значения устанавливаются Релеевой функцией распределения,

$$f_R(x) = \frac{x}{\sigma^2} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}, \quad (3)$$

где σ является параметром, который варьирует, пока наиболее подходящий не будет найден, а X - значение пикселя. Смысл этой процедуры заключается в том, что градиентные значения пикселей изображения будут распределены по закону Релея в соответствии с предположениями о том, что такие значения I_x и I_y являются идентичными, независимыми, и, как правило нормально (по закону Гаусса) распределены[3].

2. Медианный фильтр. Медианный фильтр реализует нелинейную процедуру подавления аддитивного и импульсного шумов на изображении. Медианный фильтр представляет собой скользящее по полю изображения окно W , охватывающее нечетное число отсчетов. Центральный отсчет заменяется медианой всех элементов изображения, попавших в окно. Медианой дискретной последовательности x_1, x_2, \dots, x_L для нечетного L называют такой ее элемент, для которого существуют $(L - 1)/2$ элементов, меньших или равных ему по величине, и $(L - 1)/2$ элементов, больших или равных ему по величине. Другими словами, медианой является средний по порядку член ряда, получающегося при упорядочении исходной последовательности. Например, $\text{med}(20, 10, 3, 7, 7) = 7$. Двумерный медианный фильтр с окном W определим следующим образом:

$$\hat{X}(n_1, n_2) = \underset{W}{\text{med}} y(n_1, n_2) = \underset{W}{\text{med}} [y(n_1 + k_1, n_2 + k_2) : (k_1, k_2) \text{ из области } W] \quad (4)$$

Среди медианных фильтров размерностью 3×3 наиболее распространены следующие:

$$W_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad W_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad W_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad W_4 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (5)$$

3. Оператор Лапласа

Для повышения резкости изображений используется такой пространственный фильтр как оператор Лапласа.

Оператор Лапласа (лапласиан) - дифференциальный оператор, который действует в линейном пространстве гладких функций и обозначается символом Δ . Функции F он ставит в соответствие функцию

$$\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \dots + \frac{\partial^2}{\partial x_n^2} \right) F \quad (6)$$

Оператор Лапласа часто обозначается таким способом ∇^2 , т.е. в виде скалярного произведения оператора Набла на себя. Оператор Лапласа эквивалентен также последовательному взятию операций градиента и дивергенции:

$$\Delta = \nabla \cdot \nabla = \text{div grad} \quad (7)$$

Таким образом, значения оператора Лапласа в точке могут пониматься как плотность источников (стоков) потенциального векторного поля в этой точке.

4. Морфологические операции

Для достижения эффекта сглаживания бинарных растровых изображений и удаления отдельных дефектов используется комбинация базовых операторов математической морфологии - эрозии и расширения. Данный метод относится к группе нелинейных методов цифровой обработки изображений.

Пусть подмножество S - совокупность точек изображения, которые принадлежат интересующему объекту B - структурируют элемент. Операция эрозии определяется в математической морфологии следующим образом:

$$\mathcal{E}_B(S) = S - B = \{B_h \subseteq S \mid h \in E\} \quad (8)$$

где $\mathcal{E}_B(S)$ означает "эрозию подмножества S структурируют элементом B ";

E - дискретная плоскость;

h - вектор, определяющий каждую позицию B на E ;

B_h - перенесение B вдоль вектора h ;

$$B_h = \{b + h \mid b \in B\} \quad (9)$$

Во время процедуры эрозии элементы не удаляются с изображения, а лишь обозначаются для удаления после окончания всей процедуры.

Расширение определяется как:

$$\mathcal{D}_B(S) = S \oplus B = \{s + b \mid s \in S; b \in B\} \quad (10)$$

где $\mathcal{D}_B(S)$ означает "расширения подмножества S структурируют элементом B ".

Последовательное использование эрозии и расширения ведет к сглаживанию бинарных изображений[2].

Результат обработки существенно зависит от двух факторов: конфигурации структурированного элемента и его размера. Как известно, выбор конфигурации является скорее эмпирическим и интуитивным процессом. Протестировав разные комбинации элементов, мы пришли к выводу, что в наибольшей степени поставленной задаче отвечает следующая: структурированный элемент для эрозии является совокупностью далеких соседей опорного пикселя, а для расширения - полное безразличие соседей опорного пикселя.

Что касается размера структурированного элемента, то очевидно, что операции эрозии и расширения должны использовать один и тот же размер, иначе толщина линий объекта изменится после окончания процесса обработки.

С помощью описанных выше методов было разработано два алгоритма для реализации поставленной задачи. Результатами работы программы является количественный подсчет положительно заряженных лейкоцитов на изображении гистологического среза (клетки, мембраны которых окрашены в коричневый цвет), а также дополнительные параметры, как площадь клеток, относительная площадь, периметр. Программный продукт выполнен на языке Visual Basic 6.0. Интерфейс программы представлен на рис. 1.

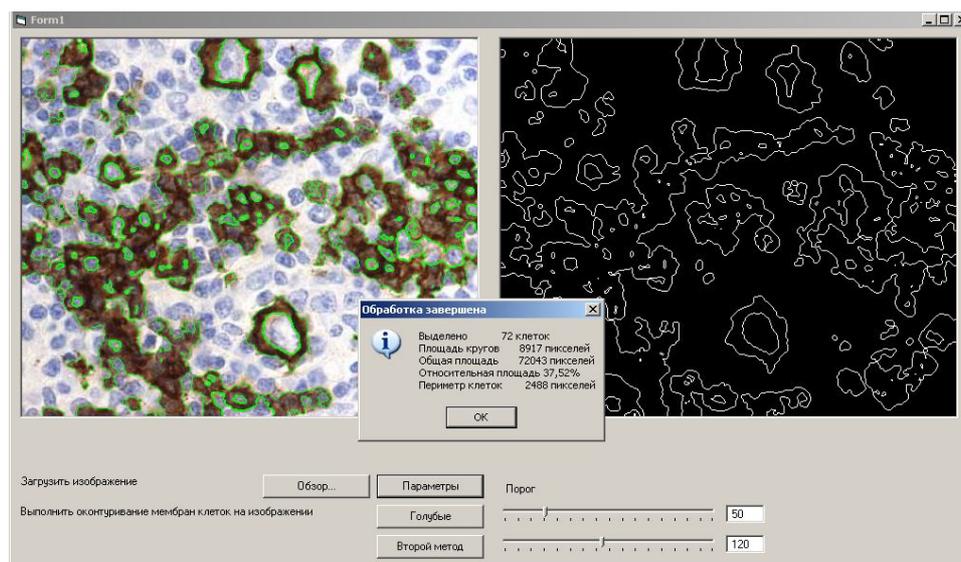


Рисунок 1 - Интерфейс ПО по обработке изображений

Для установления точности работы программного продукта проведем следующий эксперимент. Выберем 5 независимых экспертов, которые проведут ручной подсчет количества клеток на десяти различных изображениях гистологических срезов. Результаты приведены в табл. 1.

Таблица 1

Количество клеток в расчетах 5-ти экспертов

№ изобр.	1 эксп.	2 эксп.	3 эксп.	4 эксп.	5 эксп.	M	D	СКО
1	88	79	82	84	85	83,6	11,3	3,4
2	27	25	29	33	32	29,2	11,2	3,3
3	9	12	11	10	12	10,8	1,7	1,3
4	100	102	110	105	110	105,4	20,8	4,6
5	50	48	51	51	50	50	1,5	1,2
6	20	18	22	22	21	20,6	2,8	1,7
7	120	127	125	128	129	125,8	12,7	3,6
8	22	21	19	22	20	20,8	1,7	1,3
9	99	95	98	100	95	97,4	5,3	2,3
10	350	360	358	355	359	356,4	16,3	4,0

Предположим, что усредненное значение подсчетов пяти экспертов (математическое ожидание M) является точным количеством клеток на изображении.

Теперь представим данные расчетов с использованием двух разработанных алгоритмов подсчета клеток и сравним их с эталонным.

Таблица 2

Определение количества клеток алгоритмами программы

№ избр.	1	2	3	4	5	6	7	8	9	10
1 алг.	65	17	6	80	32	14	70	15	58	289
2 алг.	72	26	10	107	43	19	118	20	83	349

Проиллюстрируем графически приближенность данных методов расчета к эталонному значению.

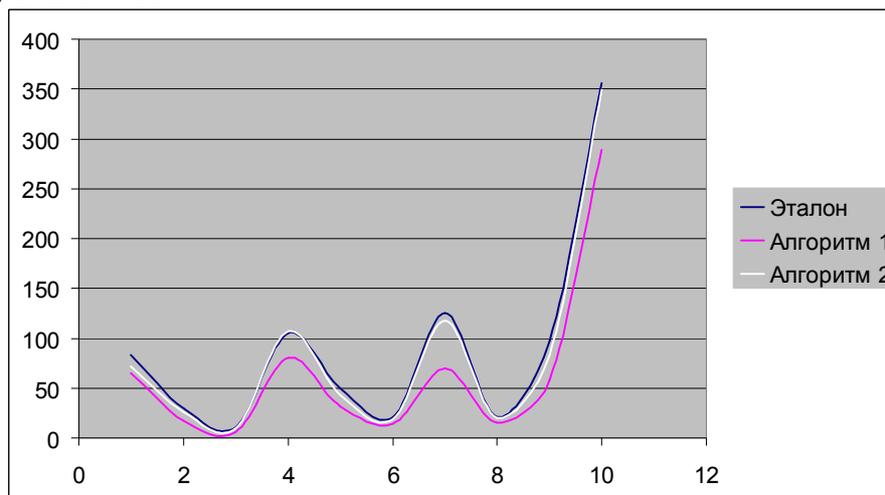


Рисунок 2 – Количество клеток, установленное различными методами

Как видно на рис.2, алгоритм №2 с использованием пороговой сегментации и метода Лапласа оказался наиболее эффективен в проблеме оконтуривания и установления количества клеток на изображении, чем алгоритм №1, использующий отсечение по цвету и метод Прюита. В заключении хотелось бы отметить, что не смотря на довольно многочисленные существующие ПО по обработке изображений, рассматриваемая проблема остается актуальной и по сей день, поскольку невозможно избрать один алгоритм для решения различных проблем, которых особенно много в области медицины. Надеюсь, что данная программа будет полезна для работы сотрудников лаборатории и ускорит процесс диагностики заболеваний.

Использованная литература:

- [1] В.К. Гринь, А.А. Селезнев «Основы иммуногистохимии. Часть 1. Антитела»
- [2] [http://www.newdiagnostics.dn.ua/index.php?option=com_content&task=view&id=18&Itemid=17&&lang=ru]: Научная статья.
- [3] О.Л. Коневский «Адаптивная морфологическая обработка бинарных контуров» // Электронный журнал "Исследовано в России", 4, 1722-1731, 2001.
- [4] Mats Kvarnström, Katarina Logg, Alfredo Diez, Kristofer Bodvard, Mikael Käll «Image analysis algorithms for cell contour recognition in budding yeast» [http://www.opticsinfobase.org/DirectPDFAccess/351802CD-BDB9-137E-C9C3F4794EF34E6D_170494.pdf?da=1&id=170494&seq=0&CFID=37376861&CFTOKEN=71773669]: Scientific article.