ОБЗОР ПОРОГОВОГО МЕТОДА СЕГМЕНТАЦИИ И ПОДСЧЕТА КЛЕТОК НА ПРЕПАРАТАХ КРОВИ ЧЕЛОВЕКА

OVERVIEW OF THE THRESHOLD METHOD FOR SEGMENTATION AND CELL COUNTING ON HUMAN BLOOD PREPARATIONS

УДК 004

Васильев И.Н., студент-магистрант, 2 курс Институт Информатики и Вычислительной техники, кафедра «Программное обеспечение»Ижевский государственный технический университет

Россия, г. Ижевск

Перевозчиков Д.Е., студент-магистрант, 2 курс Институт Информатики и Вычислительной техники, кафедра «Программное обеспечение» Ижевский государственный технический университет

Россия, г. Ижевск

Vasiliev I. N., it.ilya97@gmail.com

Perevozchikov D. E., it.ilya97@gmail.com

Аннотация

Рассмотрен простой в исполнении и главное быстрый метод сегментации, локализации и подсчета клеток крови на изображении препарата крови человека. В этом методе используется пороговое значение для отделения клеток от других частей изображения. Мы также используем преобразование Хафа для кругов, чтобы найти центр белых клеток(лейкоцитов). Определение местоположения и подсчет красных клеток (эритроцитов) осуществляется с помощью сопоставления шаблонов. Используется поиск локальных максимумов, маркировка и вычисление среднего значения, чтобы сжать области, полученные после применения преобразования Хафа или

сопоставления шаблонов до одного пикселя в качестве однозначного местоположения каждой области. Предлагаемый метод очень быстр и точно вычисляет количество и расположение лейкоцитов. Он также способен обнаруживать и подсчитывать красные клетки с небольшой ошибкой.

Abstraction

A simple in execution and, most importantly, fast method of segmentation, localization and counting of blood cells on the image of a human blood sample is considered. This method uses a threshold value to separate cells from other parts of the image. We also use the Hough transform for circles to find the center of the white cells (leukocytes). Locating and counting red cells (erythrocytes) is done using pattern matching. Local maximum search, labeling, and average computation are used to compress the regions obtained after applying the Hough transform or pattern matching to one pixel as the unambiguous location of each region. The proposed method is very fast and accurately calculates the number and location of leukocytes. It is also capable of detecting and counting red cells with little error.

Ключевые слова: Подсчет клеток крови, сегментация клеток крови, преобразование Хафа для кругов, сопоставление шаблонов, Пороговое значение.

KeyWords: Blood Cell Counting, Blood Cell Segmentation, Hough Transform for Circles, Pattern Matching, Threshold.

1. Введение

Визуальное исследование изображений крови является важным инструментом диагностики, профилактики и лечения пациентов. Ручные операции с этими изображениями очень медленные и неточные. Таким образом, возникла потребность в автоматизации таких операций, и многие автоматизированные операции вводятся для обнаружения, классификации и измерения объектов в гематологической цитологии. На самом деле анализ образцов крови предполагает окрашивание образца для исследования под микроскопом. Изображения, полученные с помощью микроскопа, могут быть

автоматически или полуавтоматически обработаны с помощью методов компьютерного зрения для классификации и подсчета клеток. До сих пор для этого было предложено несколько автоматических и полуавтоматических схем задача [1]~[6]. Например, в работе [1] был предложен быстрый и полностью автоматизированный метод сегментации и идентификации границ различных типов клеток. [2] предлагает метод сегментации ядра лейкоцитов, основанный на ортогонализации Грама-Шмидта. В [3] мы видим метод подсчета эритроцитов, основанный на импульсно-связанной нейронной сети и авто-волне. [4] предлагает метод дифференциального подсчета лейкоцитов. Он делает отделение ядра от других частей с помощью порогового значения из-за разницы интенсивности между частями белых клеток и фоном и разделение цитоплазмы с помощью серии морфологических эрозий и дилатаций.

В нашем методе разделения ядра мы используем пороговое значение, как в [4]. Мы также используем преобразование Хафа для кругов для определения местоположения белых клеток и сопоставления шаблонов для определения местоположения и подсчета красных клеток. Чтобы объяснить предложенный метод с полной детализации, рассмотрим пример изображения и опишем метод шаг за шагом. Как мы увидим, наш метод очень быстр, и единственной значительной трудоемкой операцией является вычисление взаимной корреляции между исходным изображением и шаблонами. Время обработки для нашего метода составляет примерно 6 секунд на изображение, что сопоставимо со временем обработки, заявленным в работе [1]. Метод, предложенный в [1], полностью автоматизирован, но наш метод является полуавтоматическим и требует некоторой ручной настройки параметров и оптимизации. Предлагаемый метод в основном пред ставляет собой комбинацию общих схем обработки изображений и компьютерного зрения, но порядок и тип конкретных схем, используемых на каждом этапе, делают его новым подходом.

Этот метод состоит из трех частей. В первой части выполняется некоторая предварительная обработка изображения для того, чтобы

подготовить его к применению следующих операторов, и в следующих двух этапах будет использован ряд методов обработки изображений для сегментации и подсчета белых и красных клеток.

2. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИЗОБРАЖЕНИЯ

Как видно из рис. 1, образец исходного изображения имеет низкий контраст, и в нем есть некоторые нежелательные периодические линии, которые особенно видны в правой нижней части (ортант IV) изображения. Это обычные артефакты в этих типах изображений. В частности, артефакты останутся и после процесса обнаружения контура клеток, который мы собираемся сделать позже, и могут сделать изображение края очень грязным. Чтобы устранить линии, сначала мы должны понять, что это за шум. Они кажутся эффектами синусоидального шума на гистограмме. Если мы посмотрим на абсолютное значение БПФ-преобразования изображения на рис. 2, мы видим в нем две яркие точки.

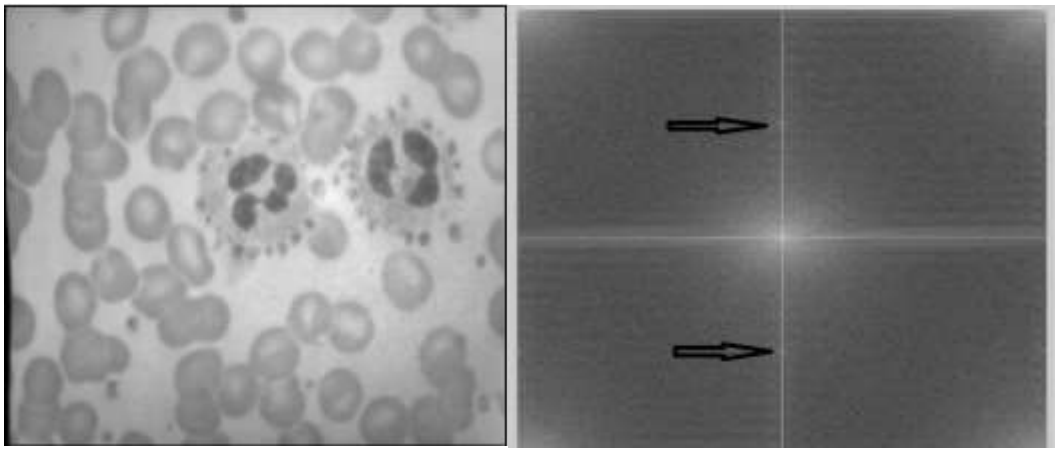

Рис.1. Исходное изображение клеток Рис.2. Быстрое преобразование фурье

Эти точки соответствуют частоте этого одночастотного шума. Чтобы удалить этот шум, мы фильтруем изображение с помощью низкочастотного фильтра Баттерворта порядка 9 и частоты среза ¼. Чтобы улучшить контраст, далее делаем стандартную гистограмму выравнивания по изображению. Полученное изображение можно увидеть на рис. 3.

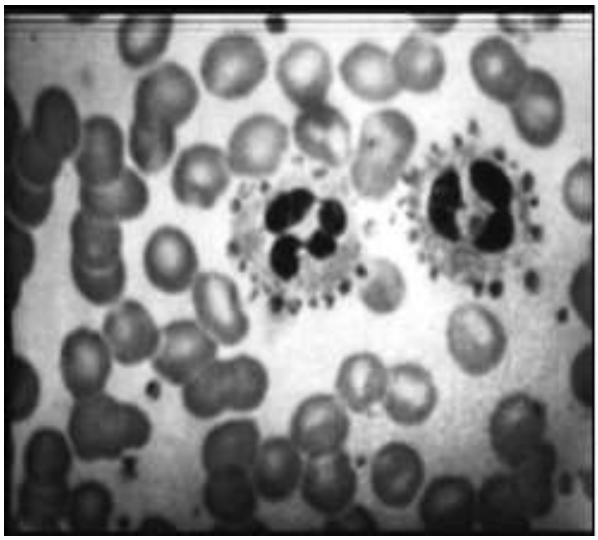

Рис.3. Изображение после обработки

Как видно из рис. 3, большая часть нежелательных линий устранена, а оставшиеся находятся в верхней и нижней частях изображения и не вызовут серьезных проблем в процессе сегментации и подсчета клеток. Это изображение готово к применению основного метода, и мы объясним предложенный метод в следующих подразделах, выполнив ряд процессов на этом изображении.

3. ПРЕДЛАГАЕМЫЙ СПОСОБ ПОЛУЧЕНИЯ ЛЕЙКОЦИТОВ

Как показано на рис. 3, ядро каждой белой клетки имеет значительно меньшую интенсивность, чем другие части изображения. Таким образом, как указано в [4], используя пороговое значение, мы можем отделить эти части от других частей. После порогового значения и преобразования изображения серого уровня в двоичное с помощью метода Otsu [7] для нахождения оптимального уровня порогового значения мы получаем изображение, подобное тому, что мы видим на рис. 4.



Рис.4. Изображение в бинарном виде после применения порогового значения.

Чтобы подсчитать количество ядер лейкоцитов (которое равно количеству лейкоцитов) и локализовать их, сначала мы помечаем каждую белую область на основе критерия связности 8. Мы исключаем некоторые ненужные части на полях изображения при маркировке. Для каждого региона мы выбираем среднюю точку. Номер строки средней точки определяется как среднее значение максимального и минимального числа строк пикселей в области. Аналогичным образом определяется и номер столбца. Поскольку каждое ядро может иметь разные части и все они должны учитываться как одно ядро, мы рассматриваем средние точки, которые составляют не более 60 пикселей (приблизительная длина радиуса белых клеток на этом изображении для его конкретного масштаба) далеко, как одно ядро и присваиваем им одну и ту же метку в процессе переклейки. После повторной маркировки количество областей равно количеству лейкоцитов.

Для определения местоположения белых клеток мы находим приблизительный центральный пиксель для каждого из них. Чтобы найти такой центральный пиксель, мы ищем центр окружностей для окружностей радиуса 60 пикселей в краевом изображении с помощью преобразования Хафа. Чтобы выполнить такой поиск, нам нужна область поиска для каждой белой клетки. Чтобы найти эту область, мы используем средние точки различных областей в белых

клетках. Фактически, мы находим максимальное и минимальное число строк и устанавливаем их как нижние и верхние строки прямоугольной области поиска.

Мы следуем тому же подходу для нахождения максимального и минимального номера столбца. После нахождения областей поиска мы извлекаем изображение контура с помощью метода «Детектор границ Канни». Мы можем видеть изображение контура клеток на рис. 5.

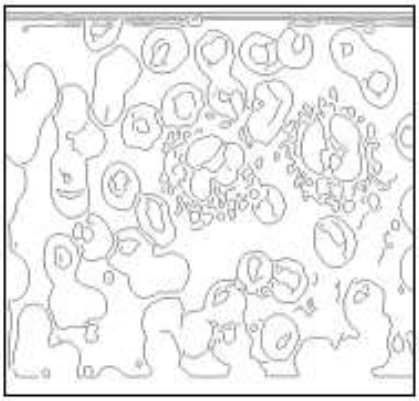

Рис.5. Полученные контуры клеток.

Как мы уже говорили, мы пытаемся найти центр ядер, применяя Преобразование Хафа для кругов на последнем изображении. Мы должны заметить, что могут быть некоторые поддельные области, которые выжили после порогового значения, которые не принадлежат ни одному ядру. Эти области могут быть исключены, когда мы ищем центры белых клеток, потому что ни один пиксель в них не получит вес как представитель центрального пикселя. Поэтому мы не вводим их в виде ядра и не вычитаем их количество из общего числа белых клеток.

Чтобы найти части, принадлежащие цитоплазме, мы можем присвоить новый цвет частям расположенных кругов, которые не были признаны ядром. Мы видим результаты на рис. 6. Число белых клеток, вычисленное как 2, является правильным.

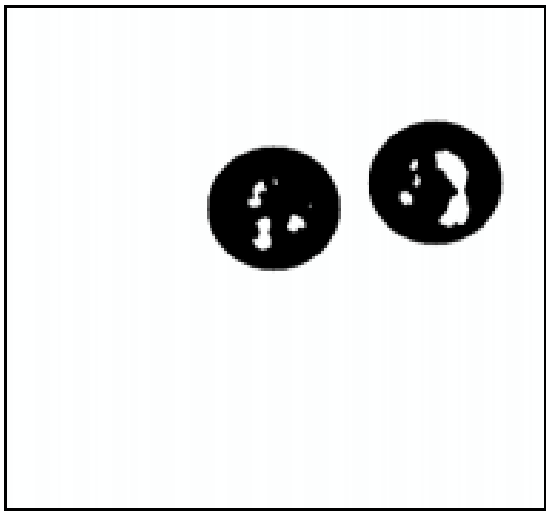

Рис.6. Определенные преобразованием Хафа области лейкоцитов.

Распознав белые клетки и определив их расположение, мы можем вычесть эти части из предварительно обработанного изображения (рис. 3) и получить новое изображение, содержащее только красные клетки. Такое изображение мы видим на рис. 7.

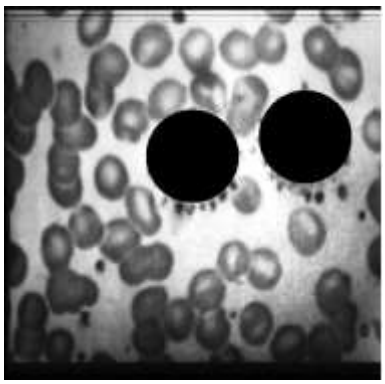

Рис.7. Изображение, содержащее только эритроциты.

4. ПРЕДЛАГАЕМЫЙ СПОСОБ ПОЛУЧЕНИЯ ЭРИТРОЦИТОВ

Для того чтобы найти и подсчитать количество красных клеток в обработав изображение, мы выбираем пять красных клеток в качестве шаблонов (пронумерованные клетки на рис. 8) и вычислить взаимную корреляцию между обработанным изображением и ними. Затем мы линейно объединяем полученные изображения процесса корреляционных вычислений, чтобы получить изображение, в котором для каждой красной клетки есть пиковое значение. Для этого надо оптимизировать коэффициенты. Кажется, что

красные клетки, которые значительно не пересекаются с другими частями, могут быть лучшими шаблонами, такими как шаблоны № 36 и 46. Таким образом, мы увеличиваем их влияние на результирующее изображение, выбирая для них большие коэффициенты, как мы определяем как $Res_0=Res_1+Res_{28}+1.2\times Res_{36}+Res_{45}+1.2\times Res_{46}$

Здесь Res_i - результирующее изображение после вычисления корреляции с номером шаблона “i”, а Res_0 – окончательное изображение после линейного комбинирования. Мы можем видеть изображение Res_0 на рис. 9.

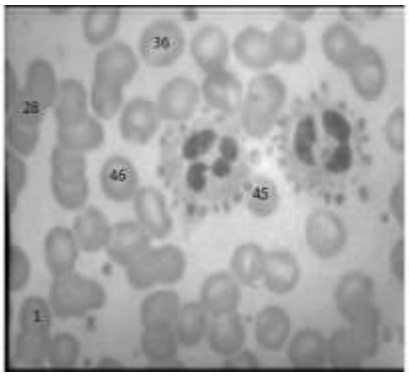

Рис.8. Эритроциты, помеченные в качестве шаблона.

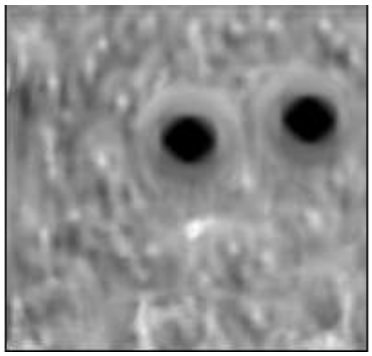

Рис.9. Результирующее избражение сумм корреляций.

На этом изображении в месте расположения красных клеток мы видим пиковые области, а не пиковые пиксели. Чтобы подсчитать количество этих областей и найти их центральный пиксель, мы сначала используем пороговое значение для удаления некоторых не максимальных пикселей. После этого мы находим локальный максимум в каждой области, чтобы уменьшить площадь пика все больше и больше. Затем мы помечаем сморщенные участки и находим

их количество, это и является количеством красных клеток. Для определения местоположения мы вводим местоположение средней точки каждой области как среднее местоположение между всеми пикселями в этой области.

5. РЕЗУЛЬТАТЫ

Если мы присвоим красный цвет красным клеткам, желтый-цитоплазме, голубой - ядру и синий - фону, используя предложенный метод, мы будем распознавать и разделять различные части изображения, как показано на рис. 10. Красные клетки, которые использовались в качестве шаблонов, нумеруются на изображении. Количество распознанных эритроцитов составляет 49, что близко к фактическому числу (46). Мы также применили наш метод к изображению, изображенному на рис. 11, и полученное изображение изображено на рис.12. Как мы видим на рис. 12, есть некоторые пиксели, распознанные как ядро, но не получившие голосов после применения преобразования Хафа. Таким образом, полученное число белых клеток является истинным числом, равным 2. В этом случае распознанное число эритроцитов равно 52, что совпадает с фактическим числом.

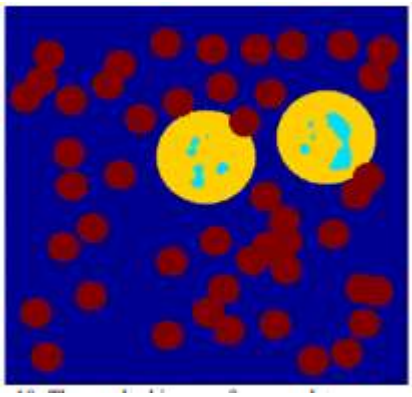

Рис.10.Результирующее изображении после сегментации.

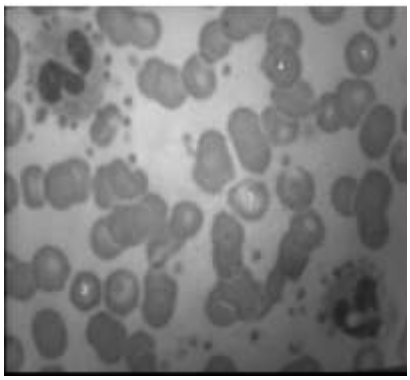

Рис.11. Тестовое изображение.

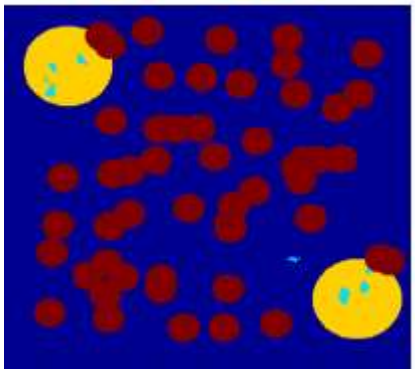

Рис.12. Результат сегментации на тестовом изображении

6. ЗАКЛЮЧЕНИЕ

Мы предложили быстрый метод сегментации и подсчета количества клеток крови на изображении. Как мы видели, этот метод показал почти идеальные результаты в подсчете и локализации лейкоцитов. Он также был способен распознавать искусственные области низкой интенсивности, вызванные плохим освещением, и удалять их из набора белых клеток. Кроме того, он мог обнаружить и подсчитать количество красных клеток с небольшой ошибкой. В дальнейшем его планируется применять для задач обучения нейронной сети.

Использованные источники:

1. Атлас по медицинской микробиологии, вирусологии и иммунологии: учеб. пособие для студентов мед. вузов / под ред. А.С. Быкова, А.А. Воробьева, В.В. Зверева. 2-е изд., доп. И перераб. М.: Мед. информ. агентство, 2008. 272 с.
2. Система автоматического распознавания лейкоцитов в мазке периферической крови / И.А. Пастушок [и др.] // Вестник медицинского института «РЕАВИЗ», № 1, 2018 г.
3. Angulo J., Flandrin G. Automated Detection of Working Area of Peripheral Blood Smears Using Mathematical Morphology // Analytical Cellular Pathology. 2003. Volume 25, №1, P. 37-49.
4. Дифференциальный подсчет лейкоцитов крови с использованием автоматической микроскопии и системы поддержки принятия решений на

основе искусственных нейронных сетей - оценка DiffMasterTM Octavia / Сволин Б.Г. [и др.] // Clinical and laboratory hematology. 2003. Volume 25, №3, P. 139–147.

Sources used:

1. Atlas of Medical Microbiology, Virology and Immunology: textbook. manual for students of medical universities /
1. ed. by A. S. Bykov, A. A. Vorobyov, V. V. Zverev. 2nd ed., add. And pererab. M.: Med. inform. agency, 2008. 272 p.
2. The system of automatic recognition of leukocytes in a smear of peripheral blood / I. A. Pastushok [et al.] / / Bulletin of the medical Institute "REAVIZ", No. 1, 2018.
3. Angulo J., Flandrin G. Automated determination of the working area of peripheral blood smears using mathematical morphology // Analytical cell pathology. 2003. Volume 25, No. 1, pp. 37-49.
4. Differential counting of blood leukocytes using automatic microscopy and decision support systems based on artificial neural networks-evaluation of DiffMasterTM Octavia / Svolin B. G. [et al.] / / Clinical and laboratory hematology. 2003. Volume 25, No. 3, pp. 139-147.