

УДК 004.932: 004.83: 004.8.032.26

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ КЛАССИФИКАЦИИ ФОРМЕННЫХ ЭЛЕМЕНТОВ КРОВИ

Томакова Р.А., Филлист С.А., Жилин В.В., Борисовский С.А.

ФГБОУ ВПО «Юго-Западный государственный университет», Курская государственная сельскохозяйственная академия, Курск, e-mail: Tomakova@rambler.ru

В работе рассмотрена структурно-функциональная организация интеллектуальной системы классификации форменных элементов крови на основе анализа изображений мазков периферической крови. В качестве математического аппарата для построения классифицирующих моделей использовались морфологический анализ, Фурье-анализ и нейросетевое моделирование. Сформирована база данных, которая содержит базу изображений, базу моделей нейронных сетей, предназначенных для сегментации и классификации изображений и базу обучающих выборок. Разработанное программное обеспечение для интеллектуальной системы классификации форменных элементов крови характеризуется гибким управлением процессом интерактивного взаимодействия программного обеспечения, с лицом, принимающим решение, осуществляется посредством интерфейсных окон и всплывающих меню. Реализация программных модулей анализа и классификации изображений осуществлена в среде MATLAB 7.10 в виде модуля *MedImProc*. Используются следующие пакеты этой программной среды: Image Processing Toolbox и Neural Network Toolbox.

Ключевые слова: сегментация изображений, RGB-коды пикселей, контур, дескрипторы Фурье, морфологические операторы, информативные признаки, нейронные сети

SOFTWARE INTELLECTUAL SYSTEM OF CLASSIFICATION OF FORMED ELEMENTS OF BLOOD

Tomakova R.A., Filist S.A., Zhilin V.V., Borisovskiy S.A.

South-West State University, Kursk State Agricultural Academy, Kursk, e-mail: Tomakova@rambler.ru

The paper considers the structural-functional organization of the intellectual system of classification of blood cells on the basis of the analysis of images of peripheral blood smears. As a mathematical apparatus for building of classification models were used morphological analysis, Fourier analysis and neural network modeling. A database that contains a database of images, base models of neural networks intended for segmentation and classification of images and a database of training samples. Developed software for intellectual system of classification of the formed elements of the blood characterized by flexibility in the management of process of interaction of the software with the person making the decision, is carried out through interface Windows and pop-up menus. Development of the software modules of the analysis and classification of images made in MATLAB 7.10 as a module *MedImProc*. Use the following packages of this software environment: Image Processing Toolbox and Neural Network Toolbox.

Keywords: image segmentation, RGB codes pixels, contour, Fourier descriptors, morphological operators, informative characteristics, neural networks

Проблемы построения интеллектуальных систем, особенности построения специального математического обеспечения, выбор модели и учет неопределенности находятся в центре внимания различных исследователей [3].

В настоящее время интенсивно развиваются интеллектуальные системы, предназначенные для автоматической обработки медицинских изображений. На практике при обнаружении форменных элементов крови сталкиваются с основным и типичным для медицинских приложений компьютерного зрения препятствием – большой вариабельностью изображений, с которыми приходится иметь дело [4].

Форменные элементы крови могут быть классифицированы по двум независимым группам признаков. К первой группе относятся цветовые показатели. Несмотря на подробное описание стандартной окраски мазка, сформулировать количественные критерии нелегко. Цветовые характеристики изменяются непрерывно, а восприятие цветов зависит от множества факторов – ос-

вещения, наследственности и возраста данного наблюдателя и т.д.

Ко второй группе относятся геометрические параметры элемента: размер и форма. Но эти параметры не определяют однозначно принадлежность форменного элемента к тому или другому классу.

Учитывая вышесказанное, программное обеспечение, реализующие классификацию форменных элементов крови, построено таким образом, что позволяет осуществлять независимый анализ форменных элементов крови по двум группам признаков с возможностью последующей агрегации получаемых решений.

Программное обеспечение для интеллектуальной системы классификации форменных элементов крови выполнено в среде Matlab 7.10 R2010a в виде модуля *MedImProc* и использует следующие пакеты этой программной среды: Image Processing Toolbox и Neural Network Toolbox [2].

Модульная структура разработанного программного обеспечения представлена на рис. 1.

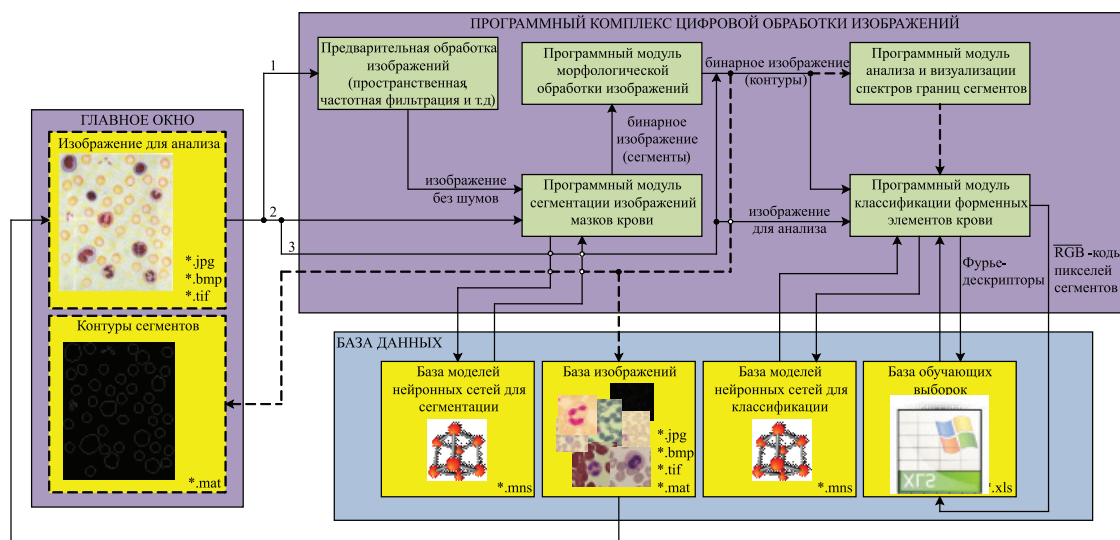


Рис. 1. Модульная структура программного обеспечения для гематологического анализа

В процессе выполнения программы автоматически формируется база данных, состоящая из четырех блоков: первый представляет собой базу изображений (в формате *.bmp, *.jpg, *.tif, *.mat); второй – базу моделей нейронных сетей (в формате *.mns), предназначенных для сегментации изображений; третий – базу моделей нейронных сетей (в формате *.mns), предназначенных для классификации сегментов изображения; четвертый – базу обучающих выборок (в формате *.xls).

База изображений содержит цветные микроскопические изображения мазков крови, которые необходимо либо анализировать (определять количество лейкоцитов и эритроцитов, присутствующих на них), либо формировать по ним обучающие выборки. База изображений также содержит бинарные изображения контуров выделенных сегментов (в формате *.mat). Изображение для анализа выбирается из базы данных изображений.

Прежде чем перейти к анализу межклеточных соотношений в мазке периферической крови, его изображение необходимо сегментировать, то есть обнаружить все форменные элементы крови на изображении и представить их пиксели «единицами», а пиксели сегментов изображения, не относящихся к форменным элементам крови, представить «нулями».

Эта процедура соответствует маршруту «2», показанному на рис. 1. Если требуется препарирование анализируемого изображения, то выбирается маршрут «1». После выполнения процедур препарирования обработка продолжается по маршруту «2».

Модуль программного обеспечения сегментации микроскопических изображений

мазков крови состоит из двух программных блоков:

1) блок нейросетевой классификации пикселей изображения;

2) блок морфологического анализа сегментов изображения.

Программное обеспечение блока нейросетевой классификации пикселей изображения в процессе обучения осуществляет формирование моделей нейронных сетей для сегментации фотографий изображений мазков периферической крови.

Основными технологическими операциями при этом являются: выбор размера маски (окна) для формирования обучающей выборки; выбор цветового канала, по которому осуществляется обучение; формирование обучающих выборок для класса «плазма» и классов «форменные элементы крови» путем перемещения окна и выбора соответствующей области на изображении; настройка параметров нейронной сети; обучение нейронной сети; сохранение модели нейронной сети в базу данных.

Модель нейронной сети для сегментации представляет собой файл с расширением *.mns, структура данных которого включает в себя как параметры нейронной сети, так и характеристики обучающей выборки, на которой были получены эти параметры нейронной сети.

По полученным моделям нейронной сети программный модуль реализует бинарную классификацию пикселей на классы «форменный элемент» – «плазма».

В процессе сегментации изображения программный модуль осуществляет выборку из базы моделей нейронных сетей необходимой модели, наиболее подходящей по цветовому окрасу к исследуемому изо-

бражению. Это осуществляется в интерактивном режиме. Такая процедура не увеличивает существенно время анализа, так как исследуется серия фотоснимков с одинаковым цветовым окрасом. На вход нейронной сети подается не RGB-код пикселя, а среднее значение кодов, попавших в окно. Поэтому модель нейронной сети подбирается не только по цветовому окрасу, но и по размеру окна, в котором определяются цветовые характеристики пикселя (средние, моды и т.п. RGB-кодов пикселей, попавших в окно).

В результате работы блока сегментации изображений мазков крови получают изображение, пример которого представлен на рис. 2, а. Такие изображения содержат объекты (ложные сегменты), образованные неправильно классифицированными пикселями. Причем характерной особенностью этих сегментов является то, что их размеры значительно меньше форменных элементов крови, поэтому для получения качественного бинарного изображения, содержащего контуры только форменных элементов крови, используется блок морфологической обработки изображения. Поэтому после сегментации выбранного изображения осуществляется дополнительная обработ-

ка этого изображения посредством модуля морфологической обработки изображения.

На изображении, представленном на рис. 2, видно, что ядра эритроцитов при сегментации были отнесены к классу «плазма», поэтому для их заполнения применим морфологический оператор «imfill», посредством которого заполняют «отверстия» единицами на бинарном изображении. Применение этого морфологического оператора приведено на рис. 2, б. Использование морфологического оператора размывания «imopen» приводит к исчезновению мелких объектов (размер которых значительно меньше форменных элементов крови) и зубчатых элементов границ на изображении. При этом можно использовать структурообразующий элемент в форме плоского шестиугольника «octagon» с $R = 3$ (R – расстояние от центра элемента до стороны шестиугольника, измеренное вдоль горизонтальной или вертикальной оси). Результат применения данного морфологического оператора представлен на рис. 2, в. Для выделения границ бинарных объектов на изображении используем морфологический оператор «bwperim». Результат применения данного морфологического оператора представлен на рис. 2, г.

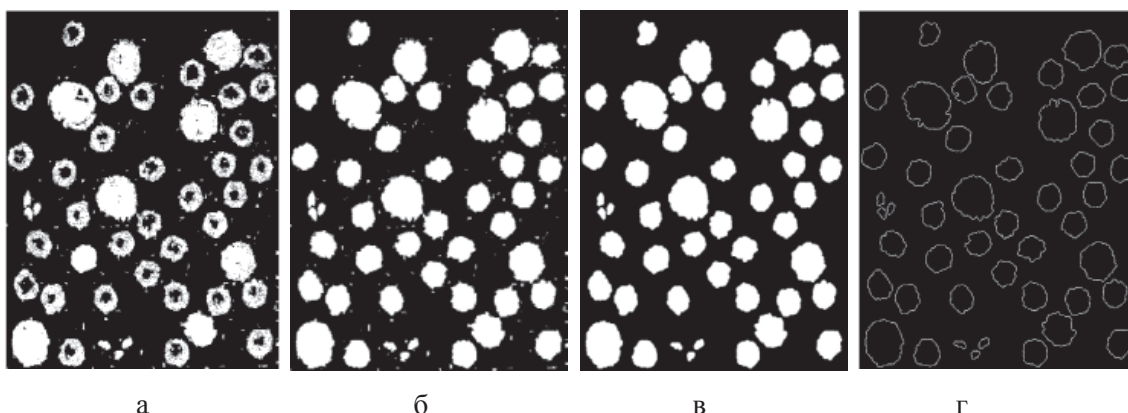


Рис. 2. Результаты работы морфологических операторов «imfill», «imopen» и «bwperim»

В результате выполнения такой последовательности морфологических операторов получается бинарное изображение, представленное на рис. 2, г, которое используется для работы программного модуля классификации форменных элементов крови. На выходе этого модуля морфологической обработки получаем координаты границ выделенных сегментов, которые поступают на вход модуля анализа и визуализации спектров границ сегментов. Данный модуль позволяет определить и визуализировать действительную, мнимую и амплитудную составляющую спектра контура, анализи-

ровать зависимость изменение формы сегментов от числа используемых дескрипторов Фурье [5].

Кроме этих двух маршрутов обработки данных, предусмотрен маршрут «3», которым можно воспользоваться в случае, если обрабатываемое исходное изображение уже было сегментировано и в базе данных хранится файл с сегментированным изображением, который имеет то же имя, что и анализируемое изображение, но с расширением *.mat.

Все три маршрута обработки данных заканчиваются модулем классификации

форменных элементов крови. В нем формируются информативные признаки «по геометрии» и «по цвету», и с выхода которого считываются межклеточные соотношения в мазке.

При обучении для модуля сегментации изображений выходными данными являются модели нейронной сети, которые поступают с данного программного модуля в базу данных моделей нейронных сетей для сегментации. Для модуля классификации форменных элементов крови выходными данными являются модели нейронной сети и обучающие выборки, которые поступают с данного программного модуля в базу данных моделей нейронных сетей для классификации и в базу данных обучающих выборок соответственно [1].

Интерфейсное окно разработанного специального программного обеспечения для

получения этих моделей включает четыре пункта горизонтального меню: «Файл», «Изображение», «Методы обработки изображений» и «Помощь».

Работа программы начинается с меню «Файл», которое позволяет вызвать из базы данных необходимое изображение. Анализируемое изображение располагается в основном поле экрана. Меню «Изображение» служит для вспомогательных операций, а основные программные модули специального программного изображения собраны в список группы «Методы обработки изображений».

На рис. 3 показана диаграмма, иллюстрирующая сравнительные характеристики показателей качества классификации лейкоцитов и эритроцитов (средние значения) для ПК «ДиаМорф», ПО «Цито» и разработанного программного обеспечения.

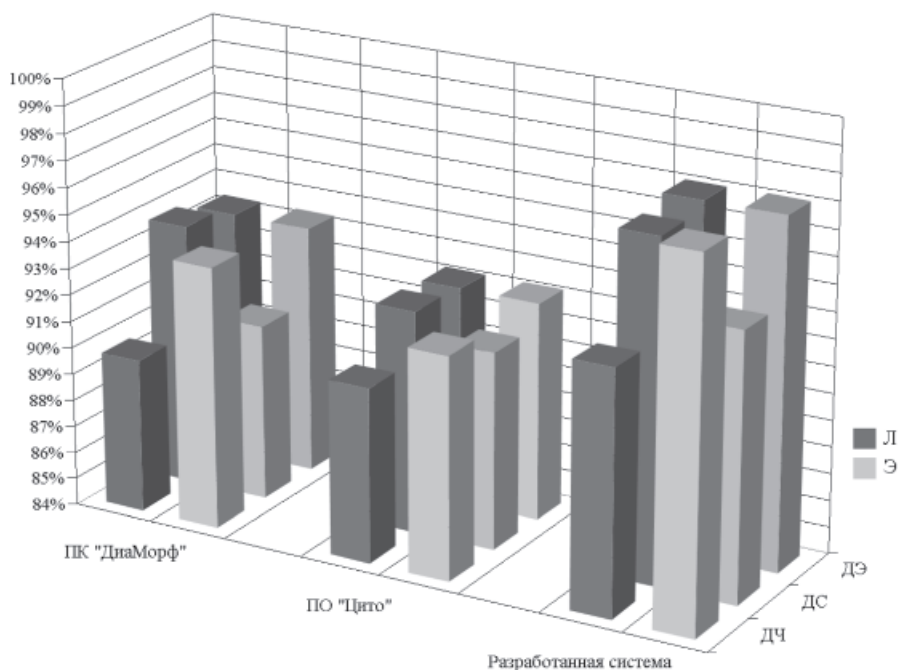


Рис. 3. Сравнительные характеристики показателей качества классификации лейкоцитов и эритроцитов

Экспериментальные исследования показывают, что во всех рассмотренных системах показатель диагностической чувствительности (ДЧ) при классификации эритроцитов выше, чем показатель ДЧ при классификации лейкоцитов, это связано с пропусками лейкоцитов, принятых за скопления эритроцитов.

Также следует отметить, что представленное специальное программное обеспечение характеризуется меньшим числом ложных тревог (разрушенные клетки, скопления тромбоцитов, грязь, пятна краски,

ошибочно принятые за лейкоциты) и превосходит известные по показателям качества классификации на 4...6%. Кроме того, представленное программное обеспечение позволяет полностью исключить человека из процедуры классификации форменных элементов крови.

Выводы

1. Предложена процедура измерения межклеточных соотношений в периферической крови, включающая двухступенчатую обработку изображения. На первой ступени

обрабатывают изображение, полученное в результате сегментации исходного черно-белого изображения с помощью морфологического оператора «эрозия». На второй ступени – посредством использования морфологического оператора «дилатация». На каждой ступени решение принимается на основе агрегации двух оценок, полученных в результате анализа цветного и черно-белого изображения. Окончательное решение принимается на основе сопоставления решений, принятых на первой и на второй ступенях обработки, что позволяет использовать гибридные технологии для анализа изображений.

2. Разработано специализированное программное обеспечение, предназначенное для автоматизированного анализа микроскопических фотографий сложноструктурируемых изображений, которое реализовано в среде Matlab 7.10. Предложенное программное обеспечение характеризуется гибким управлением процессом интерактивного взаимодействия программного обеспечения, с лицом, принимающим решение, осуществляется посредством интерфейсных окон и всплывающих меню. Разработанное программное обеспечение позволяет осуществлять процессы сегментации и классификации сложноструктурируемых изображений на основе полного технологического цикла синтеза многоуровневых моделей нейронных сетей, включающего процессы формирования обучающих выборок, вычисления параметров нейронных сетей и определения диагностической эффективности, полученных решающих правил.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации в рамках государственного контракта № 14 514 11 4089.

Список литературы

1. Борисовский С.А. Нейросетевые модели с иерархическим пространством информативных признаков для сег-

ментации плохоструктурированных изображений / С.А. Борисовский, А.Н. Брежнева, Р.А. Томакова // Биомедицинская радиоэлектроника. – 2010. – № 2. – С. 49–53.

2. Гонсалес Р. Цифровая обработка изображений в среде MATLAB / Р. Гонсалес, Р. Вудс, С. Эддинс. – М.: Техносфера, 2006. – 616 с.

3. Подвальный С.Л. Интеллектуальные системы моделирования: принципы разработки / С.Л. Подвальный, Т.М. Леденева // Системы управления и информационные технологии. – 2013. – Т.51. – № 1. – С. 4–10.

4. Томакова Р.А. Метод обработки и анализа сложноструктурируемых изображений на основе встроенных функций среды MATLAB / Р.А. Томакова, С.А. Филлист // Вестник Читинского государственного университета. – 2012. – № 1 (80). – С. 3–9.

5. Анализ гистологических изображений посредством морфологических операторов, синтезированных на основе преобразования Фурье и нейросетевого моделирования / С.А. Филлист, Томакова Р.А., Горбатенко С.А. [и др.] // Биотехносфера. – 2010. – № 3(9). – С. 54–60.

References

1. Borisovsky, S. A. Neuronet Models with Hierarchical Space of Informative Signs for Segmentation of Badly Structured Images / S. Borisovsky, A.N. Bregneva, R.A. Tomakova // Biomedical radio electronics. 2010. no. 2. pp. 49–53.

2. Gonzalez, R. Digital image processing using MATLAB/R. Gonzalez, R. Woods, S. Eddins. M: Technosphere. 2006. 616 p.

3. Podvalny, S.L. Intellectual systems modeling: principles of development / С.Л. Basement, Т.М. Ledeneva // control Systems and information technologies. 2013. T.51, no. 1. pp. 4–10.

4. Tomakova, R.A. Method and Analysis of Complex-Structured Images Processing on the Basis of Built-in Functions of Environment MATLAB / R.A. Томакова, S.A. Filist//Bulletin of Chita state University. 2012. no. 1 (80). pp. 3–9.

5. Filist, S.A. Analysis of Histological Images Using Morphological Operators, sSynthesized on the Basis of Fourier and Neural Network Modeling / S.A. Filist, Tomakova R.A., Gorbatenko S.A. [and other] // Biotechnosphere. 2010. no. 3(9). pp. 54–60.

Рецензенты:

Мишустин В.Н., д.м.н., профессор, Курский государственный медицинский университет, г. Курск;

Бурмака А.А., д.т.н., профессор, главный научный сотрудник НИЦ, ФГУП 18 ЦНИИ, г. Курск.

Работа поступила в редакцию 15.08.2013.