

ПОСТРОЕНИЕ И АНАЛИЗ СИГНАТУР ОБЪЕКТОВ БИОМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Ф. А. Сиверский, Л. В. Калацкая

Белорусский государственный университет, Минск, Беларусь

Искусственные нейронные сети позволяют создавать автоматизированные системы для решения задач управления, распознавания образов, диагностики заболеваний, автоматического анализа документов и многих других приложений [1].

Главным их отличием от других систем, например таких, как экспертные системы, является то, что нейросети в принципе не нуждаются в заранее известной модели, а строят ее сами только на основе предъявляемой информации [2].

Обработка непосредственно изображений или контуров объектов нейронными сетями часто связана с большой размерностью входного вектора нейросети. Для уменьшения размерности входного вектора и выделения ключевых признаков объектов применяются сигнатуры [3], представляющие собой одномерную функцию, например, зависимость расстояния от центра масс контура до его границы от угла.

В работе предлагается алгоритм выделения контура и построения сигнатуры изображений, а также проводится сравнение нейросетевой классификации с использованием сигнатур и контурных представлений объектов биомедицинских изображений.

В исследовании использована перевиваемая культура клеток крови K562, опухолевых клеток из клеток предшественников эритроцитов и гранулоцитов. Клетки облучались полторы минуты ультрафиолетовым светом с длиной волны 365 нм для того, чтобы вызвать разрывы ДНК. После облучения живые клетки окрашивали флюорохромом акридиновый оранжевый, который встраивается в двойную спираль ДНК.

Особенностью данных изображений является наличие относительно яркого ядра и тусклого тела клетки. Поэтому создание контура клетки состоит из двух частей. Для построения внешнего контура клетки применяется пороговая обработка. При построении контура ядра клетки используется гамма-коррекция для увеличения контрастности ядра [4]. Затем два контура совмещаются. Для построения сигнатуры используется только внешний контур. Таким образом, алгоритм предварительной обработки изображений включает:

- линейное контрастирование;

- пороговую обработку;
- создание внешнего контура;
- гамма-коррекцию;
- создание контура ядра клетки;
- создание сигнатуры объекта.

На рис. 1 представлены результаты предварительной обработки. При этом изображения контуров инвертированы, так как при подаче на вход нейронной сети пиксели, принадлежащие линии контура, имеют значения 1, фона - 0.

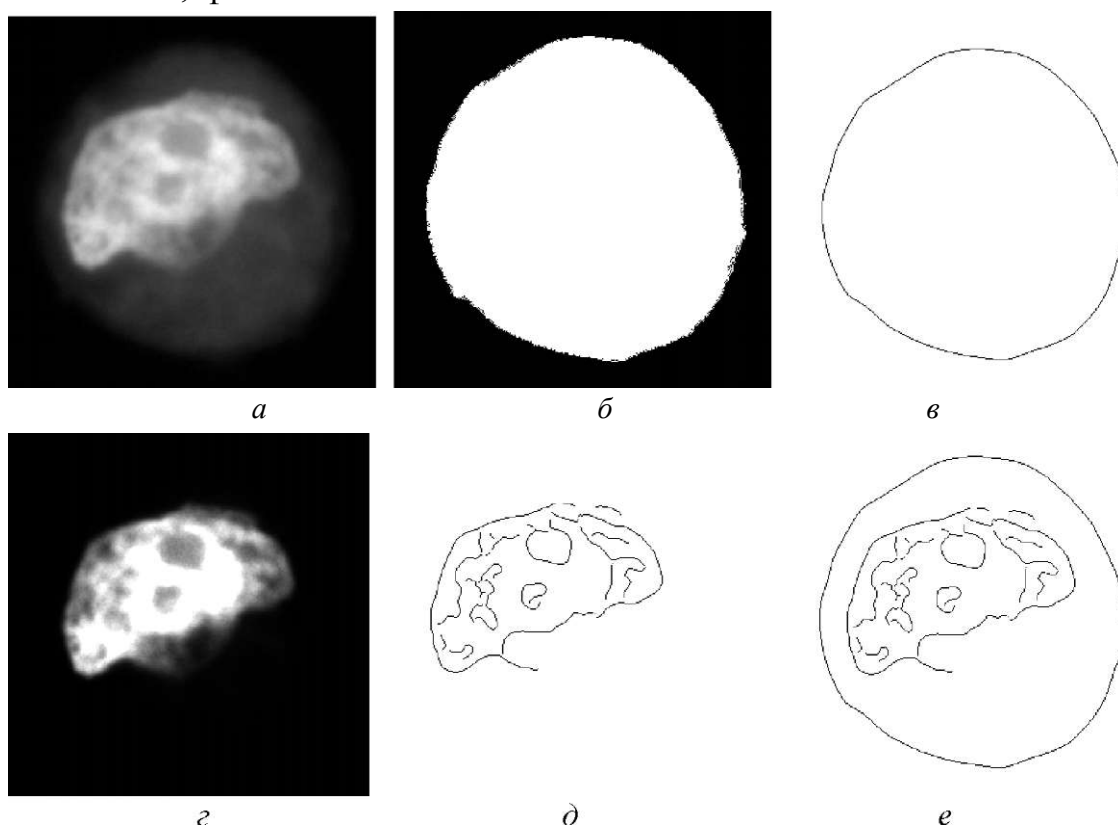


Рис. 1. Схема алгоритма подчеркивания контура:

а - исходное изображение, *б* - пороговая обработка, *в* - внешний контур клетки, *г* - гамма-коррекция, *д* - контур ядра клетки, *е* - полный контур клетки

Из-за особенностей обрабатываемых биомедицинских изображений сигнатура должна обладать рядом свойств, таких как независимость от расположения контура, независимость от линейных размеров, независимость от поворота объекта.

Выбранная для обработки изображений сигнатура представляет собой одномерную функцию - зависимость расстояния от центра массы до границы контура и угла. Чтобы избежать зависимости сигнатуры от линейного смещения, начальной точкой выбирается центр массы, так как

при смещении объекта он так же будет смещаться. Для обеспечения независимости сигнатуры от масштаба производится нормировка на длину периметра контура, так как периметр линейно зависит от размеров исходного изображения. Инвариантность к повороту объекта обеспечивается путем циклического смещения значений сигнатуры так, чтобы они начинались с минимального или максимального значения. Для обеспечения устойчивости алгоритма к разрывам контура и множественным границам проводится аппроксимация тех значений сигнатуры, которые попадают на разрыв контура и не были определены, а также выбираются средние значения множественных краев.

Таким образом, алгоритм построения сигнатуры включает в себя:

- определение центра масс контура;
- построение прямых из центра до границы объекта с фиксированным шагом по углу;
- аппроксимацию неопределенных значений;
- нормировку полученных значений делением их на длину периметра;
- циклический сдвиг значений так, чтобы они начинались с минимального (максимального).

К недостаткам сигнатуры такого типа можно отнести то, что для получения сигнатур клеток используется только внешний контур, при этом теряется информация о внутренней структуре клетки.

Для построения сигнатур, учитывающих особенности анализируемых объектов изображений, в системе MATLAB [5] реализована процедура, вычисляющая сигнатуру с задаваемым шагом по углу. Для исследуемых клеток был выбран шаг в 2 градуса, следовательно, полученная сигнатура содержит 180 элементов. Примеры контуров различных классов клеток представлены на рис. 2.

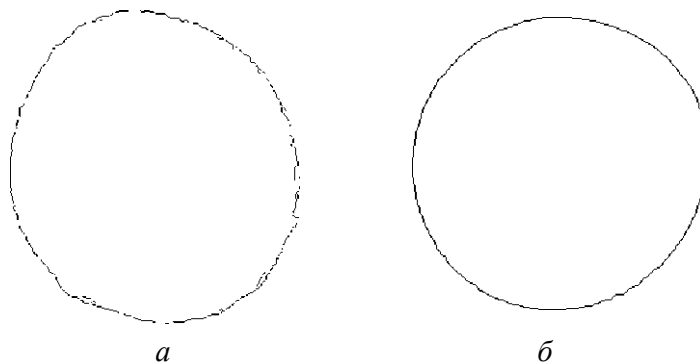


Рис. 2. Примеры контуров различных классов клеток:
а - облученная клетка, б - необлученная клетка

Сигнатуры контуров (рис. 2) приведены на рис. 3. Для облученной клетки сигнатура представлена под номером 1, для необлученной - 2.

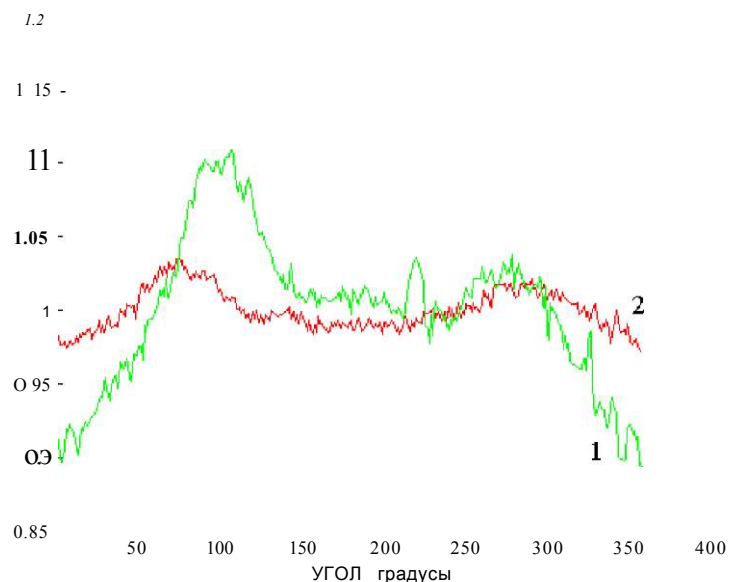


Рис. 3. Примеры сигнатур различных классов клеток:
1 - для облученной клетки, 2 - для необлученной

Полученные значения элементов сигнатур находятся вблизи единицы, поэтому перед их подачей на вход нейросети они предварительно приводятся к диапазону $[-1; 1]$, что является оптимальным для персептрона с биполярной логистической функцией активации.

Моделирование нейросети производится в системе MATLAB с использованием приложения Neural Network Toolbox [2]. Для классификации с использованием сигнатур клеток нейронная сеть представляет собой многослойный персептрон с одним скрытым слоем и биполярной логистической функцией активации [1]. Число входных нейронов сети равно 180, скрытых - 32, выходных - 2. Первый выходной нейрон соответствует контрольной клетке, а второй облученной, за результат классификации принимается большее из значений выходов нейронов. При использовании самого контура клеток как входного вектора, нейронная сеть содержит 14400 входных нейронов и 60 нейронов скрытого слоя. Для обучения и тестирования работы нейросети использовалось 76 клеток (38 контрольных и 38 облученных).

Результаты сравнения работы двух нейросетей с различным представлением входных данных приведены в таблице.

При использовании сигнатуры контура все неправильно распознанные клетки относятся ко второму классу, то есть, это клетки после облучения. Данный результат является следствием того, что при создании сигнатуры используется только внешний контур клетки, и теряется информация об ее внутренней структуре.

Сравнительная характеристика нейросетей

Параметр	Использование контура	Использование сигнатуры
Размер входного вектора	14400	180
Количество эпох обучения	65	2300
Время обучения, с	13,6	7,1
Точность распознавания, %	87,5	85,7

Количество эпох обучения при использовании сигнатур значительно выше, так как при обучении сети методом обратного распространения ошибки постоянно попадают локальные минимумы, выйти из которых удается только с помощью импульса. В целом же, время обучения второй нейросети значительно меньше.

Ключевым критерием эффективности использования сигнатур является точность правильно классифицированных объектов. Средние точности распознавания приблизительно равны для обеих нейросетей, то есть, использование только внешнего контура для построения сигнатур не привело к уменьшению точности распознавания. При этом использование всего контура в качестве входного вектора приводит к неправильной классификации клеток в обоих классах, а при использовании сигнатур - только в одном классе.

Таким образом, в работе представлен разработанный и реализованный способ построения сигнатур биомедицинских изображений, инвариантных к повороту и масштабу, и позволяющих значительно уменьшить размерность входного вектора нейронных сетей, используемых для распознавания объектов изображений.

Применение подобных сигнатур позволяет сократить время обучения нейросети. Также преимуществом данной сигнатуры является возможность изменения шага дискретизации угла, что позволяет увеличить точность представления сигнатур изображений.

Литература

1. Осовский, С. Нейронные сети для обработки информации / Пер с пол. И. Д. Рудинского. М. : Финансы и статистика, 2002.
2. Организация и обучение искусственных нейронных сетей : учеб. пособие / Авт.-сост. Л. В. Калацкая, В. А. Новиков, В. С. Садов. - Мн. : БГУ, 2003.
3. Гонсалес, Р. Цифровая обработка изображений в среде MATLAB. М.: Техносфера, 2006.
4. Компьютерный анализ и синтез изображений: Курс лекций. / Авт.-сост. Л. В. Калацкая. Мн. : БГУ, 2006.
5. Интерней-адрес: <http://matlab.exponenta.ru>.