

УДК 004.932.2

## ВЫЯВЛЕНИЕ ПРИЗНАКОВ БИОЛОГИЧЕСКИХ КЛЕТОК НА ИЗОБРАЖЕНИЯХ И ВИДЕО НА ОСНОВЕ АЛГОРИТМА ИЕРАРХИЧЕСКОЙ ВРЕМЕННОЙ ПАМЯТИ

Кондратенко В.А., Спицын В.Г.

ФГАОУ ВО «Национальный исследовательский Томский политехнический университет»,  
Томск, e-mail: vyacheslav\_kondratenko@mail.ru

В статье описывается использование формализованных отличительных признаков биологических клеток в фармакологии и смежных практических областях. Представлено описание алгоритма, способного выявить и формализовать отличительные признаки клеток. Проведена классификация возможных подпопуляций клеток. Описания возможных популяций могут быть получены только для фиксированной клеточной культуры и условий выращивания. Описание клеток может включать несколько параметров: форма, размер, текстура. Применение указанных параметров необходимо, чтобы на последующих этапах анализа можно было определить, является ли клетка типичной для данной культуры и условий или клетка является аномальной. Одним из алгоритмов, удовлетворяющих вышеописанным требованиям, является алгоритм иерархической временной памяти. Данный алгоритм может быть успешно применён для распознавания клеток в видеоряде, но его использование в распознавании клеток на изображениях затруднительно. Выявление описаний клеток вызывает большой интерес и имеет практическую ценность.

**Ключевые слова:** алгоритм, фармакология, распознавание изображений, биологические клетки, иерархическая временная память

## IDENTIFICATION OF BIOLOGICAL CELLS FEATURES ON THE IMAGES AND VIDEO USING HIEARARCHICAL TEMPORAL MEMORY ALGORITHM

Kondratenko V.A., Spitsyn V.G.

National Research Tomsk Polytechnic University, Tomsk, e-mail: vyacheslav\_kondratenko@mail.ru

This paper describes how pharmacology and related practical areas can use formal distinctive features of biological cells. Presents the description of algorithm, which can detect and formalize the distinctive features of biological cells. Described possible sub-populations of biological cells. Descriptions of possible sub-populations can be obtained only for fixed cell culture and growing conditions. Cell description can contains several parameters like size, texture, shape. These cell descriptions could be used for determine which cell is typical for current culture or cell is anomaly. Hierarchical temporal memory algorithm is one of algorithms, which meet necessary requirements. This algorithm can provide good results with identifying of biological cells on the video, but identifying of biological cells on the image is not so successfully. Identification of biological cells description is of practical value.

**Keywords:** algorithm, pharmacology, image recognition, biological cells, hierarchical temporal memory

В фармакологической отрасли часто возникает необходимость наблюдения за биологическими клетками. Наблюдение может производиться путём видеосъёмки. Встречаются ситуации, когда удобнее производить фотосъёмку материала, а потом анализировать фотографии. В обоих случаях анализ заключается в следующем:

- 1) выявление ядер клеток;
- 2) выявление цитоплазмы клеток;
- 3) подсчёт клеток;
- 4) отличительные особенности клеток;
- 5) активность клеток (реализуем только на видео ряде).

Проведение такого анализа человеком весьма затруднительно. Поэтому для анализа используется специализированное программное обеспечение.

### Описание проблемы

Изображения могут сильно отличаться друг от друга. Отличия возникают в силу

того, что биологический материал, клеточная культура, в присутствии вируса и лекарственного компонента часто подвержен биологическим процессам, связанным с ростом, размножением, заражением, выздоровлением и отмиранием клеток. Аппаратура и метод визуализации (флуоресценция белков) это ещё один источник разнообразия. Настройка аппаратуры и типа микроскопии сильно влияет на яркостной диапазон, контраст, фокус и уровень шума. Вариации условий эксперимента: время инкубации, поставщик лекарственных компонент, флуоресцентный белок – могут привести к совершенно разным результатам.

Артефакты аппаратуры, которые иногда встречаются на фотографиях, являются дополнительной проблемой во время распознавания клеток (рис. 1) [8].

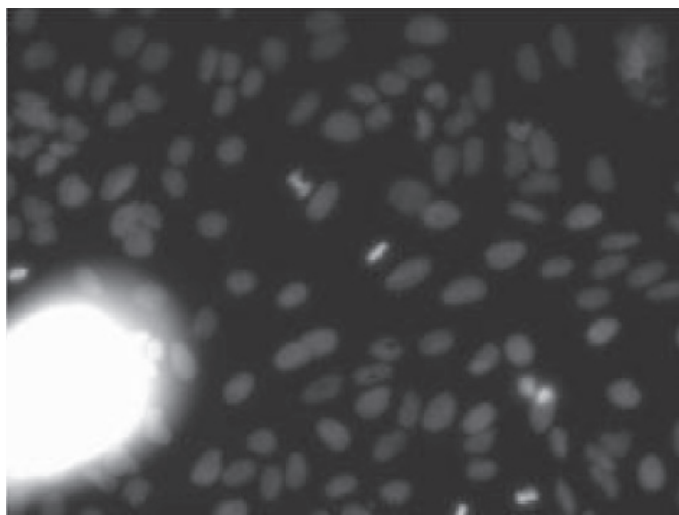


Рис. 1. Пример фотографии с артефактом аппаратуры

Одним из результатов анализа является выделение отличительных особенностей клеток. То есть описание возможных подпопуляций клеток. Такие описания могут быть получены только для фиксированной клеточной культуры и условий выращивания. Описание клеток может включать несколько параметров: форма, размер, текстура. Это описание необходимо, чтобы на последующих этапах анализа можно было определить, является ли клетка типичной для данной культуры и условий или клетка является аномальной [6]. Таким образом, выявление описаний клеток вызывает большой интерес и имеет практическую ценность. Визуализация получаемых описаний также имеет практическую ценность [1].

#### Существующие решения

На сегодняшний момент в области анализа клеточного материала имеется несколько известных программных продуктов: Fiji, CellProfiler, Icy. Это весьма мощные инструменты, имеющие возможность обнаружения ядра и цитоплазмы клетки. Некоторые инструменты позволяют строить 3D модель. К преимуществам этих систем также можно отнести внушительную базу расширений, которые могут упростить работу с системой или добавить новый функционал. Но тем не менее данные системы имеют сложный интерфейс, а также нуждаются в тонкой настройке под каждый отдельный эксперимент. Можно сделать вывод, что вышеописанные продукты используют алгоритмы, которые слишком быстро реагируют на изменения в биологическом материале. Таким обра-

зом, необходим алгоритм, удовлетворяющий следующим требованиям:

- 1) устойчивость к большому разнообразию биологического материала;
- 2) дообучаемость;
- 3) простота настройки;
- 4) возможность работы с видеорядом;
- 5) возможность выделения описания подпопуляций клеток.

#### Алгоритм НТМ

Одним из алгоритмов, удовлетворяющих вышеописанным требованиям, является алгоритм «НТМ» (Hierarchical Temporal Memory – Иерархическая временная память). НТМ можно использовать для извлечения описания подпопуляций клеток и в дальнейшем использовать эти описания в других алгоритмах, например в нейронных сетях.

Сеть НТМ представляет собой иерархию, состоящую из регионов. Регион НТМ – функциональная единица памяти, основной строительный блок. По мере продвижения вверх по этой иерархии всегда присутствует конвергенция данных, когда многие элементы дочернего (нижнего) региона соединяются на одном элементе родительского (верхнего) региона. Наличие обратных связей позволяет информации разделяться обратно при движении вниз по уровням иерархии. На рис. 2 [7] представлена упрощенная диаграмма четырех регионов НТМ, организованных в четырехуровневую иерархию, а также показаны потоки информации внутри уровня, между уровнями и извне/наружу для всей иерархии [2].

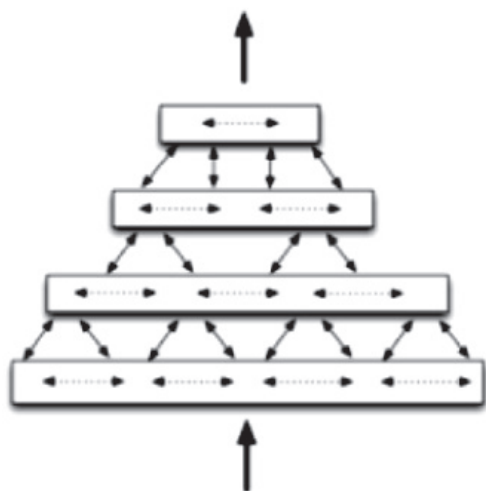


Рис. 2. Упрощенная диаграмма четырех регионов НТМ

Высокая эффективность – одно из преимуществ иерархической организации. Иерархическая организация существенно сокращает время обучения и необходимые объемы памяти, так как на более высоких уровнях шаблоны, выученные на каждом уровне иерархии, используются многократно в комбинациях [4]. Регионы НТМ представляют собой слои сильно взаимосвязанных клеток, организованных в колонки.

Формирование в регионе НТМ пространственного распределенного представления входных данных является первым этапом в работе алгоритма.

Входная информация для региона НТМ представлена в виде большого количества бит. В первую очередь регион НТМ конвертирует свой вход в пространственное разреженное представление. К примеру, при 60% активных (значение входа равно 1) битах на входе, во внутреннем представлении может быть порядка 2% активных бит. Каждая колонка в регионе подключена к своей уникальной области входных битов (области никогда полностью не совпадают для двух разных колонок) [3]. В результате различные входные данные региона приводят к различным уровням активации колонок. В то же время колонки с более высоким уровнем активации подавляют (не дают возможности стать активными) близлежащие колонки с меньшим уровнем активации. Данное подавление происходит в некоторой окрестности. Варьируя радиус этой окрестности, мы можем контролировать количество одновременно активных колонок. Распределенное представление текущих входных сигналов кодируется тем, какие колонки являются активными, а какие нет, после подавления. Таким образом, при несущественном изменении входа некото-

рые из колонок будут получать себе на вход немного больше, или немного меньше активных битов, но итоговое множество активных битов не должно сильно измениться. Таким образом, похожие входные данные (которые имеют существенное множество общих активных битов) будут отображаться в относительно стабильное множество активных колонок. Такое поведение алгоритма обуславливает его устойчивость к небольшим изменениям во входных данных [5].

Вторым этапом является формирование представления входных данных в контексте предыдущих входов.

Данный этап осуществляет конвертирование вышеописанного колончатого представления текущего входа региона в новое представление, учитывающее предыдущее состояние региона, которое называется путём активации не целой колонки а только некоторой части клеток колонки. Клетки колонки могут быть пассивными и активными. Таким образом, активируя различные клетки колонки, можно представить один и тот же вход региона в различных контекстах. Для формирования такого представления входного сигнала используется следующее правило. Если входной сигнал является неожиданным для колонки, то есть не предсказанным заранее, то активируются все клетки колонки. Но если в активируемой колонке находятся клетки в состоянии предсказания, то только эти клетки будут активированы в колонке.

Третий этап заключается в формировании предсказания на основе входа и контекста предыдущих входов.

При изменении входного сигнала с течением времени изменяется и набор активных клеток в регионе. При активации любой клетки формируются соединения с клетками, которые были активны в предыдущий момент времени и находятся в пределах некоторого расстояния от активной клетки. Скорость формирования этих связей зависит от установленной скорости обучения. Суть формирования предсказания заключается в отслеживании клеткой активации этих связей. То есть когда эти соединения становятся активными, можно предположить, что клетка тоже вскоре активируется. Такая клетка переходит в состояние предсказания. Таким образом, прямой входной сигнал в регион активирует некоторое подмножество клеток, что в свою очередь приводит другое подмножество клеток в состояние предсказания. На рис. 3 [7] представлен регион с активными клетками (серые) и с клетками в состоянии предсказания (чёрные), белые клетки являются не активированными и не находящимися в состоянии возбуждения, т.е. пассивными.

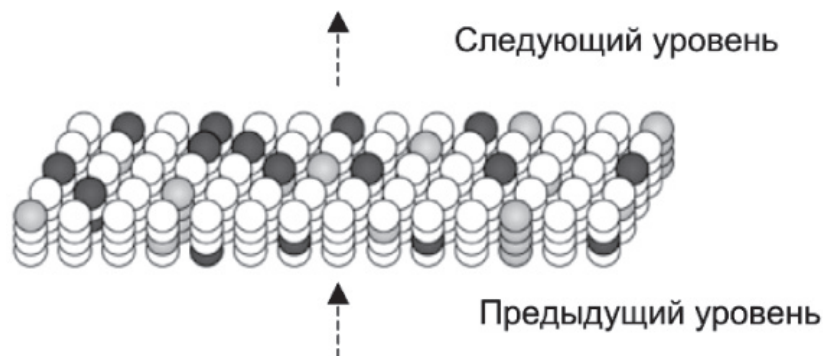


Рис. 3. Регион с активными клетками (серые) и клетками в состоянии предсказания (тёмно-серые)

Формирование таких предсказаний не ограничивается одним шагом. Такие предсказания могут формироваться на несколько шагов вперёд. Выходом региона является вектор всех активных (и находящихся в состоянии предсказания) клеток. Таким образом, при увеличении количества шагов, на которые будет формироваться предсказание, увеличится стабильность выхода региона.

### Заключение

Данная статья описывает потенциальную пользу от использования формализованных отличительных признаков биологических клеток в фармакологии и смежных практических областях. Представлено описание алгоритма, способного выявить и формализовать отличительные признаки клеток. Такие признаки могут быть получены только для фиксированной клеточной культуры и условий выращивания. Описание клеток может включать несколько параметров: форма, размер, текстура. Это описание необходимо, чтобы на последующих этапах анализа можно было определить, является ли клетка типичной для данной культуры и условий или клетка является аномальной. Одним из алгоритмов, удовлетворяющих вышеописанным требованиям, является алгоритм иерархической временной памяти. Данный алгоритм может быть успешно применён для распознавания клеток в видеоряде, но его использование в распознавании клеток на изображениях требует выполнения ряда условий. Это связано с тем, что алгоритм основывается на выявлении шаблонов во входном потоке данных. Качество распоз-

навания на изображениях можно значительно повысить, уменьшив временной интервал между получением изображений или получить ряд изображений, изменяя исходное изображение, к примеру используя вращение. Таким образом, можно сделать вывод, что предложенный в работе способ описания клеток и алгоритм их распознавания целесообразно применять при поиске аномальных клеток в биологических структурах.

### Список литературы

1. Аврамчук В.С., Гергет О.М., Лунева Е.Е. Разработка компонента визуализации биомедицинских данных на основе технологии OPENGL // Известия Волгоградского государственного технического университета. – 2013. – Т. 17. – № 14 (117). – С. 28–31.
2. Блейкли С. Об интеллекте / С. Блейкли, Д. Хоккинс // Издательский дом «Вильямс». – Москва – Санкт-Петербург – Киев, 2007. – 128 с.
3. Болотова Ю.А., Кермани А.К., Спицын В.Г. Распознавание символов на цветном фоне на основе иерархической временной модели с предобработкой фильтрами Габора // Электромагнитные волны и электронные системы. – 2012. – Т. 17. – № 1. – С. 14–18.
4. Болотова Ю.А., Спицын В.Г., Фомин А.Э. Применение модели иерархической временной памяти в распознавании изображений // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 60–63.
5. Bolotova Y.A., Spitsyn V.G. Analysis of Hierarchically-Temporal Dependencies for Handwritten Symbols and Gestures Recognition // 7th International Forum on Strategic Technology (IFOST – 2012): Proceedings: in 2 vol., Tomsk, September 18–21, 2012. – Tomsk: TPU Press, 2012 – Vol. 1 – P. 596–601.
6. Gerget O.M., Devjatyh D.V. Information System for Health State Diagnostics // CSIT'2013. – Уфа, 2013. – С. 126–131.
7. Hierarchical Temporal Memory. [электронный ресурс]. – Режим доступа: <http://numenta.com/assets/pdf/whitepapers/hierarchical-temporal-memory-cortical-learning-algorithm-0.2.1-ru.pdf> (дата обращения: 25.04.2015).
8. Workflow and metrics for image quality control in large-scale high-content screens. [электронный ресурс] сайт. – URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3593271/figure/F3/> (дата обращения: 25.04.2015).

**References**

1. Avramchuk V.S., Gerget O.M., Luneva E.E. Razrabotka komponenta vizualizacii biomedicinskih dannyh na osnove tehnologii OPENGL // *Izvestija Volgogradskogo gosudarstvennogo tehniceskogo universiteta*. 2013. T. 17. no. 14 (117). pp. 28–31.
2. Blejksli S. Ob intellekte / S. Blejksli, D. Hokkins // *Izdatelskij dom «Viljams»*. Moskva-Sankt-Peterburg-Kiev 2007. 128 p.
3. Bolotova Ju.A., Kermani A.K., Spicyn V.G. Raspoznavanie simvolov na cvetnom fone na osnove ierarhicheskoj vremennoj modeli s predobrabotkoj filtrami Gabora // *Jeletromagnitnye volny i jelektronnye sistemy*. 2012. T. 17. no. 1. pp. 14–18.
4. Bolotova Ju.A., Spicyn V.G., Fomin A.Je. Primenenie modeli ierarhicheskoj vremennoj pamjati v raspoznavanii izobrazhenij // *Izvestija Tomskogo politehnicheskogo universiteta*. 2011. T. 318. no. 5. pp. 60–63.
5. Bolotova Y.A., Spitsyn V.G. Analysis of Hierarchically-Temporal Dependencies for Handwritten Symbols and Gestures Recognition // *7th International Forum on Strategic Technology (IFOST – 2012): Proceedings: in 2 vol., Tomsk, September 18–21, 2012. Tomsk: TPU Press, 2012 Vol. 1. pp. 596–601.*
6. Gerget O.M., Devjatyh D.V. Information System for Health State Diagnostics // *CSIT2013 Ufa*, 2013. pp. 126–131.
7. Hierarchical Temporal Memory, Available at: <http://numenta.com/assets/pdf/whitepapers/hierarchical-temporal-memory-cortical-learning-algorithm-0.2.1-ru.pdf> (accessed 25 April 2015).
8. Workflow and metrics for image quality control in large-scale high-content screens, Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3593271/figure/F3/> (accessed 25 April 2015).

**Рецензенты:**

Тузовский А.Ф., д.т.н., профессор, ФГАОУ ВО «Национальный исследовательский Томский политехнический университет» Министерства образования и науки РФ, г. Томск;

Кориков А.М., д.т.н., профессор, заведующий кафедрой АСУ факультета систем управления, ФГБОУ ВПО «Томский государственный университет систем управления и радиоэлектроники», г. Томск.