



Московский государственный университет имени М.В. Ломоносова

Факультет вычислительной математики и кибернетики

Кафедра математической физики

Тлеубаев Адиль

**Разработка и экспериментальное исследование
математических методов автоматизации анализа
микроскопических изображений нейронов**

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Научный руководитель:

к.ф-м.н., доцент

И.Б. Гуревич

Москва, 2018

Содержание

Введение	3
Постановка задачи	5
Метод идентификации нейронов на МИН	7
Этап 1. Предварительная обработка МИН	7
Этап 2. Анализ МИН	7
Этап 3. Построение описания объектов, представленных на анализируемых изображениях	8
Этап 4. Классификация изображений и объектов на них представленных	9
Этап 5. Распознавание	9
Описание признаков, основанных на моментах	11
Результаты	14
Результаты классификации	14
Результаты идентификации	15
Результаты идентификации по группам	15
Заключение	20
Литература	21

Введение

Работа посвящена:

- постановке и решению математических задач, связанных с разработкой метода, обеспечивающего идентификацию изображений нейронов на микроскопических изображениях срезов головного мозга экспериментальных животных с заданной точностью;
- программной реализации разработанного метода;
- исследованию разработанного метода и его программной реализации в вычислительных экспериментах.

Одной из важнейших задач нейронаук является разработка экспериментальных моделей нейродегенеративных заболеваний (создание моделей гибели нейронов в определенных частях мозга). Данные модели будут использоваться для разработки новых методов и технологий, для диагностики и лечения нейродегенеративных заболеваний.

Разработка моделей могла бы продвигаться гораздо быстрее и была бы экономически эффективнее при снижении временных и материальных затрат на морфологические исследования за счет автоматизации и оптимизации методов обработки и анализа экспериментального материала на основе применения современных алгебраических и математических теорий анализа изображений и распознавания образов. Показано, что автоматизация анализа микроскопических изображений нейронов (МИН) и их отростков позволяет снизить материальные затраты на порядок, а временные затраты на два порядка [10, 11].

Анализ МИН является достаточно сложной многофакторной задачей. Решение данной задачи приведет к устранению различного рода проблем, связанных с построением моделей НДЗ, автоматизацией морфологических исследований, анализом экспериментальных материалов и прочих задач.

Как правило, процесс обнаружения и выделения нейронов на гистологических микроскопических изображениях включает следующие основные этапы: предварительная обработка, сегментация и классификация. Необходимость в предварительной обработке МИН вызвана наличием шума, низким разрешением МИН и неоднородностью контраста, связанной с погрешностями при окрашивании срезов. Для улучшения качества МИН обычно используют стандартные операции обработки изображений – сглаживание, обратную свертку, морфологическую фильтрацию и некоторые другие [14, 16].

В опубликованных работах, посвященных автоматическому или полуавтоматическому выделению нейронов на МИН, для сегментации, в основном, используются алгоритмы, основанные на пороговой обработке (например, [2]), морфологических операциях (например, модели Поттса [15]), методе водоразделов [17] или активных контурах [12]. Последующая классификация проводится, например, с помощью байесовских процедур, анализа главных компонент или методов машинного обучения [1].

Важным этапом анализа МИН срезов мозга является морфологическая характеристика выделенных нейронов. В последние годы определен широкий спектр специфических морфологических параметров для эффективной математической характеристики морфологии нейронов, в том числе их ядер и отростков [14].

Авторами доклада был ранее разработан метод идентификации нейронов, представленный в статье [10]. Данный метод основан на использовании методов математической морфологии. К изображению применялись методы фильтрации и нормализации, далее проводился этап бинаризации изображения. Бинаризованное изображение сегментировалось, для каждого сегмента проводилось вычисление признаков гистограмм ориентированных градиентов (HOG-признаки) [5]. Классификация сегментов на нейрон/фон производилась при помощи метода опорных векторов (SVM) с линейным ядром [3].

Разработанный ранее метод показывал хорошие результаты для представленного на данный момент класса изображений нейронов. Подготовка тканей головного мозга экспериментальных животных является трудоемким процессом, и не всегда проходящим одинаково. В ходе экспериментов были выделены 5 групп исходных изображений, визуально отличающиеся друг от друга. Для определенных групп изображений реализованный ранее метод показывал неприемлемые результаты.

В данной статье представлен новый автоматический метод выделения нейронов на микроскопических изображениях дофаминергических (ДА-ергических) нейронов срезов головного мозга экспериментальных животных. Метод основывается на совместном использовании математической теории распознавания образов и анализа изображений, теории информации и методов математической статистики. В работе описаны результаты оценки эффективности работы метода и сравнение его работы с представленным ранее [10].

Работа состоит из Введения, 4 параграфов, Заключения и Списка литературы.

В разделе «Постановка задачи» представлена постановка задачи, для которой разрабатывается метод, производится анализ исследуемых изображений.

В разделе «Метод идентификации нейронов на МИН» рассматривается метод идентификации нейронов на МИН, дается его пошаговое описание.

В разделе «Результаты» проводится анализ и обсуждение результатов, полученных в ходе работы разработанного метода.

В Заключении приведены основные результаты работы и направления дальнейших исследований.

Постановка задачи

Основная задача работы заключается в выделении нейронов на цифровых микроскопических изображениях дофаминергических (ДА-ергических) нейронов срезов головного мозга экспериментальных животных.

Для данной работы использованы изображения компактной части черной субстанции (КЧС) срезов головного мозга мышей (визуально отличающиеся изображения, полученные в разных экспериментах показаны на Рис. 1, Рис. 2). Изображения были получены в Лаборатории гормональных регуляций ФГУП Института биологии развития РАН им. Н.И. Кольцова. МИН в зависимости от группы имеют размеры либо 4080 на 3072 пикселей, либо 2040 на 1536 пикселей (разрешение снимков одинаковое).



Рис.1. Компактная часть черной субстанции (КЧС) среза головного мозга мыши

Обнаружение и выделение нейронов на двухмерных микроскопических изображениях срезов мозга усложняется следующими факторами [9]:

- помимо нейронов на МИН присутствуют объекты, которые нейронами не являются (частички грязи, погрешности окрашивания, складки ткани, кровеносные сосуды, синапсы, аксоны и другие части соединений нейронов);
- связи между нейронами (тонкие нити в большом количестве) образуют сложную текстуру, на фоне которой выделение нейронов затрудняется;
- нейроны на МИН могут очень сильно отличаться друг от друга как по размеру, так и по форме;
- нейроны могут быть повреждены в процессе подготовки препаратов, что влияет на их форму, и, как следствие, приводит к большому числу разнотипных клеток на срезе;

- нейроны могут «слипаться» или налагаться друг на друга;
- МИН содержат участки мозга с различным количественным распределением нейронов на единицу площади;
- микроскопические изображения различных срезов мозга сильно различаются по контрасту и яркости.



Рис.2. Компактная часть черной субстанции (КЧС) среза головного мозга мыши

Разработанный метод предназначен для выделения множества мелких информативных протяженных объектов с хорошо различимым по яркости включением овальной формы на микроскопических изображениях фронтальных срезов КЧС. Результатом работы метода являются выделенные области изображения, содержащие нейроны.

Новый математический метод идентификации нейронов должен одинаково эффективно работать на различных группах изображений.

Метод идентификации нейронов на МИН

Теоретическим фундаментом для разработки методов анализа МИН являлся Дескриптивный подход к анализу и пониманию изображений И.Б. Гуревича [10]. В соответствии с его концепцией математические методы представляются в виде специализированных алгоритмических схем. Такие схемы включают основные этапы извлечения информации из изображений:

- 1) предварительную обработку (улучшение качества изображений, удаление несущественных деталей и фрагментов, статистическую и логическую фильтрацию);
- 2) анализ изображений (обнаружение объектов и выделение их контуров, сегментацию, выбор и выделение признаков, описывающих состав и содержание изображения, поиск, выделение и построение отдельных объектов, отображающих специфику анализируемых изображений и др.);
- 3) построение описания объектов, представленных на анализируемых изображениях;
- 4) классификацию изображений и объектов на них представленных;
- 5) распознавание.

Алгоритмическая схема, реализующая разработанный метод выделения и идентификации объектов, включает следующие этапы.

Этап 1. Предварительная обработка МИН

При разработке нового метода идентификации нейронов проводился анализ работы построенного ранее метода [10]. Основным недостатком метода был этап бинаризации: часть нейронов идентифицировались как фон; многие участки фона идентифицировались как потенциальные нейроны; неправильно определялась граница нейрона - в его компоненту связности добавлялись участки фона. Сегментация, основанная на выделении связанных областей бинарного изображения, обеспечивала выделение около 70% нейронов для проведения следующего этапа классификации. В результате чего далее на этапе распознавания допускалось большое количество ошибок первого и второго рода.

В новом методе в общем виде предварительная обработка изображений не проводилась. В частном случае описания объектов, представленных на анализируемых изображениях, с помощью НОГ дескрипторов проводилось сравнение работы метода на исходных изображениях и на МИН с нормализацией.

Этап 2. Анализ МИН

На этапе анализа МИН применялся метод скользящего окна для выделения областей, содержащих объекты. Исходное изображение разбивалось скользящим окном на квадратные области разных масштабов: 24, 36, 48 и 60 пикселя. Шаг сдвига окна при этом: 6, 9, 12 и 15 пикселей. Пример выборки для классификатора нейронов методом скользящего окна приведен на Рис. 3, Рис. 4.

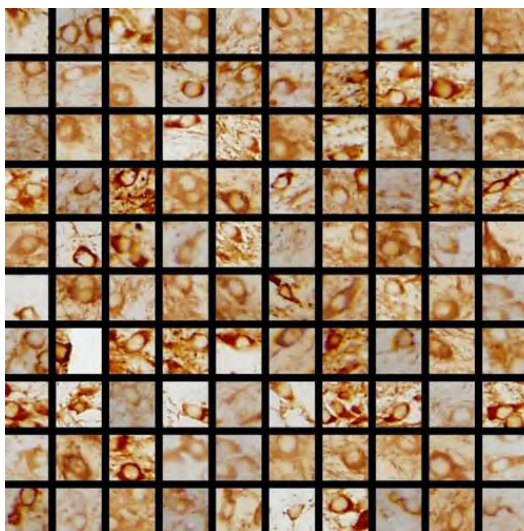


Рис. 3. Примеры объектов типа «нейрон»

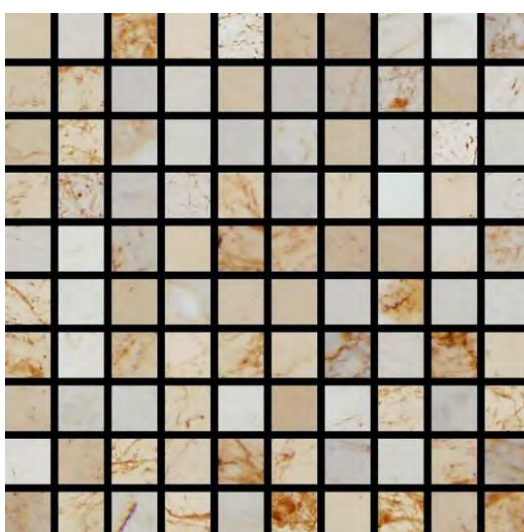


Рис. 4. Примеры объектов типа «фон»

Этап 3. Построение описания объектов, представленных на анализируемых изображениях

Для описания объектов, представленных на анализируемых изображениях, были выбраны следующие широко применяемые в области распознавания и анализа изображений и специфичные признаки:

- интенсивности цветовых каналов изображения;
- дескриптор HOG (histogram of oriented gradients) – гистограмма ориентированных градиентов. Алгоритм получения данных признаков, основанный на разбиении изображения на определенное число зон, с последующим вычислением в каждой зоне преобладающего направления градиента [5];
- гистограмма локальных бинарных шаблонов [6];
- признаки, полученные с помощью метода главных компонент из яркостей пикселей [13];
- признаки, основанные на моментах (специфичные признаки, разработанные для решения данной задачи).

Признаки, основанные на моментах, построены в соответствии с концепцией, что прямоугольные участки изображения рассматриваются в качестве элементов пространств, заданных на соответствующих прямоугольниках функциями. В качестве признаков используются коэффициенты разложения указанных функций по базисам соответствующих пространств.

Говоря более точно, участок с пикселем $(a + 1, b + 1)$ в левом верхнем углу и пикселем $(a + n, b + n)$ в правом нижнем углу задавался набором моментов

$$M_k = \int_a^{a+n} \int_b^{b+n} \rho_k\left(\frac{x-a}{n}, \frac{y-b}{n}\right) f(x, y) dx dy, \quad k \in \overline{1, 14} \quad (1)$$

где f – функция, значения которой на всевозможных квадратах видов $(a - 1, a) \times (b - 1, b)$ совпадают с яркостями пикселей (a, b) исходного изображения. Подробнее данные признаки описаны в следующем разделе.

Этап 4. Классификация изображений и объектов на них представленных.

На этапе классификации использовались следующие алгоритмы:

- Метод k ближайших соседей (англ. kNN, k-nearest neighbors algorithm) - метрический алгоритм для автоматической классификации объектов. Основным принципом метода ближайших соседей является то, что объект присваивается тому классу, который является наиболее распространённым среди соседей данного элемента [18];
- Случайный лес (англ. Random forest) – алгоритм машинного обучения, заключающийся в использовании комитета (ансамбля) решающих деревьев. Решающее дерево – классификатор, представленный в виде дерева. На ребрах дерева решения записаны признаки, от которых зависит целевая функция, в листьях записаны значения целевой функции, а в остальных узлах - атрибуты, по которым различаются случаи [4];
- XGBoost (Extreme Gradient Boosting – Экстремальный Градиентный Бустинг) [7];
- Логистическая регрессия - метод линейной классификации, позволяющий оценить апостериорные вероятности принадлежности объектов классам; при этом апостериорное распределение вероятностей основано на логистической (сигмоидной) функции [8].

На данном этапе все окна, получаемые с одного изображения, проходят этап классификации, в результате чего производится отбор окон, содержащих потенциальные нейроны.

Этап 5. Распознавание

На этапе распознавания проводилась обработка результатов, полученных на этапе классификации. Так как один объект попадает в несколько окон различного размера, то на этапе классификации окна, содержащих объект или его части, классифицировались как потенциальные нейроны. Поэтому проводился отбор окон, для предотвращения дублирования выделения одного объекта несколькими окнами. Для этого использовались значения вероятности принадлежности окна к определенному классу. Для окон, классифицированных как потенциальный нейрон проводился отбор, при котором в

результатирующее множества попадали окна имеющие высокую вероятность отношения к классу нейронов и имеющих пересечение по площади не более 30% с окнами уже находящимися в результирующем множестве.

Проведенная классификация позволила выбрать наиболее удачную пару: признаки+классификатор, которая применялась к исходным изображениям для идентификации нейронов.

Описание признаков, основанных на моментах

С математической точки зрения монохромное прямоугольное изображение размера $n \times m$ допускает рассмотрение в качестве заданной на множестве $[0, n] \times [0, m]$ конечной борелевской меры, для которой величина меры прямоугольников вида $[a - 1, a] \times [b - 1, b]$ даёт яркости соответствующих пикселей. С учётом обычных для встречающихся в приложениях ограничений на допустимую яркость, класс математических объектов, моделирующих изображения, может быть дополнительно сужен с пространства мер до пространства неотрицательных квадратично интегрируемых функций. В частности, предполагая распределение яркости равномерным внутри каждого пикселя, заданному матрицей $\{F_{ab}\}_{1 \leq a \leq n; 1 \leq b \leq m}$ монохромному изображению мы сопоставляем кусочно-постоянную функцию класса $L_2([0, n] \times [0, m])$ вида $f(x, y) = F_{[x]+1, [y]+1}$, где символом $[x]$ обозначена целая часть вещественного числа x .

Перейдём к рассмотрению вопроса о получении гипотетически пригодного для целей распознавания набора характеристик изображения. Мы ограничимся при этом случаем единичного квадрата $[0, 1] \times [0, 1]$, имея в виду, что любое квадратное изображение может быть сведено к изображению на единичном квадрате при помощи преобразования подобия.

Хорошо известно, что система многочленов $\{x^k y^l\}_{k \in \mathbb{N}, l \in \mathbb{N}}$ является фундаментальной в пространстве $L_2([0, 1] \times [0, 1])$. Соответственно, упорядочив указанную систему согласно лексикографическому порядку пар вида $\{k + l, l\}$ и проведя стандартную процедуру ортогонализации, мы получим ортонормированный базис $\{\rho_k\}_{k \in \mathbb{N}}$ пространства $L_2([0, 1] \times [0, 1])$. Векторы этого базиса имеют следующий вид:

$$\begin{aligned}
 \rho_0(x, y) &= 1, \\
 \rho_1(x, y) &= \sqrt{12} * \left(x - \frac{1}{2}\right), \\
 \rho_2(x, y) &= \sqrt{12} * \left(y - \frac{1}{2}\right), \\
 \rho_3(x, y) &= \sqrt{180} * \left(x^2 - x - \frac{11}{6}\right), \\
 \rho_4(x, y) &= 12 * \left(xy - \frac{x}{2} - \frac{y}{2} + \frac{1}{4}\right), \\
 \rho_5(x, y) &= \sqrt{180} * \left(y^2 - y - \frac{11}{6}\right), \\
 \rho_6(x, y) &= \sqrt{2800} * \left(x^3 - \frac{3x^2}{2} - \frac{27x}{5} + \frac{59}{20}\right), \\
 \rho_7(x, y) &= \sqrt{2160} * \left(x^2 y - \frac{x^2}{2} - xy - \frac{x}{2} - \frac{11y}{6} + \frac{11}{12}\right),
 \end{aligned} \tag{2}$$

$$\rho_8(x, y) = \sqrt{2160} * \left(xy^2 - \frac{y^2}{2} - xy - \frac{y}{2} - \frac{11x}{6} + \frac{11}{12} \right),$$

$$\rho_9(x, y) = \sqrt{2800} * \left(y^3 - \frac{3y^2}{2} - \frac{27y}{5} + \frac{59}{20} \right),$$

$$\rho_{10}(x, y) = 210 * \left(x^4 - 2x^3 - \frac{75x^2}{7} + \frac{82x}{7} - \frac{179}{70} \right),$$

$$\rho_{11}(x, y) = \sqrt{33600} * \left(x^3y - \frac{x^3}{2} - \frac{3x^2y}{2} + \frac{3x^2}{4} - \frac{27xy}{5} + \frac{27x}{10} + \frac{59y}{20} - \frac{59}{40} \right),$$

$$\rho_{12}(x, y) = 180 * \left(x^2y^2 - x^2y - \frac{11x^2}{6} - xy^2 + xy + \frac{11x}{6} - \frac{11y^2}{6} + \frac{11y}{6} - \frac{23}{36} \right),$$

$$\rho_{13}(x, y) = \sqrt{33600} * \left(xy^3 - \frac{y^3}{2} - \frac{3xy^2}{2} + \frac{3y^2}{4} - \frac{27xy}{5} + \frac{27y}{10} + \frac{59x}{20} - \frac{59}{40} \right),$$

$$\rho_{14}(x, y) = 210 * \left(y^4 - 2y^3 - \frac{75y^2}{7} + \frac{82y}{7} - \frac{179}{70} \right),$$

Функциональный ряд $\sum_{k=0}^{\infty} c_k \rho_k$, где символом c_k обозначены координаты

$$c_k \rightleftharpoons \int_0^1 \int_0^1 f(x, y) \rho_k(x, y) dx dy, \quad (3)$$

сходится в пространстве $L_2([0,1] \times [0,1])$ к исходной функции f , и потому последовательность $\{c_k\}_{k=0}^{\infty}$ даёт в совокупности полное описание исходного изображения. Для целей распознавания эта последовательность как целое, однако, заведомо является избыточной. Характер предметной области (задачи, связанные с обнаружением структур, визуально близких к кольцам) позволяет сделать априорное предположение о возможности ограничиться рассмотрением набора координат, отвечающих базисным многочленам порядка ≤ 4 , то есть конечного списка $\{c_k\}_{k=0}^{14}$. Действительно, распределение яркости с минимальным значением на эллипсе

$$(X - X_0)^T A (X - X_0) = 1, \quad X \rightleftharpoons \begin{pmatrix} x \\ y \end{pmatrix}, \quad X_0 \rightleftharpoons \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

где A - произвольным образом фиксированная положительная 2×2 -матрица, задаётся многочленом четвёртого порядка вида

$$f(x, y) = [(X - X_0)^T A (X - X_0)]^2.$$

Именно такой набор параметров и был использован в описываемом методе. Говоря более точно, за основу был положен несколько видоизменённый набор величин

$$\zeta_k \rightleftharpoons \frac{c_k}{\sqrt{0,01 * c_0^2 + \sum_{l=1}^{14} c_l^2}}, \quad k = \overline{1, 14} \quad (4)$$

при построении, которого произведено нивелирование величины средней яркости и общей контрастности исходного изображения.

Искомые координаты c_k для отнесённого к модельному квадрату $[0,1] \times [0,1]$ выделенного внутри произвольного изображения $f \in L_2([0,n] \times [0,m])$ участка со стороной N и координатами левого верхнего угла (a, b) очевидным образом удовлетворяют равенствам

$$c_k = \frac{1}{N^2} \int_a^{a+N} \int_b^{b+N} \rho_k \left(\frac{x-a}{N}, \frac{y-b}{N} \right) f(x, y) dx dy \quad (5)$$

Соответственно, они допускают вычисление как линейные комбинации моментов вида

$$M_{kl}(a, b, N) \rightleftharpoons \int_a^{a+N} \int_b^{b+N} x^k y^l f(x, y) dx dy \quad (6)$$

с коэффициентами, представляющими собой некоторые многочлены от параметров a, b и N^{-1} . Поэтому вопрос об эффективном вычислении координат c_k сводится к вопросу эффективного вычисления моментов $M_{kl}(a, b, N)$. Естественное решение последнего вопроса состоит в использовании модифицированного интегрального представления изображения, при котором каждому пикселу (a, b) ставится в соответствие набор величин

$$M_{kl}(a, b) \rightleftharpoons \int_0^a \int_0^b x^k y^l f(x, y) dx dy, \quad k + l \leq 4. \quad (7)$$

При этом моменты (6) допускают быстрое вычисление по формулам

$$M_{kl}(a, b, N) = M_{kl}(a + N, b + N) + M_{kl}(a, b) - M_{kl}(a + N, b) - M_{kl}(a, b + N).$$

Следует заметить, что осуществлённый нами выбор в качестве базисных функций именно многочленов не в последнюю очередь обусловлен возможностью определения соответствующих характеристик произвольных квадратных участков на основе быстрых операций над некоторым единым интегральным представлением исходного изображения.

Результаты

Прежде чем приступить к применению метода идентификации нейронов на МИН, было проведено тестирование различных признаков и классификаторов на специально подготовленных данных. Из снимков, относящихся к различным группам было выделено около 4 тысяч изображений нейронов и около 11 тысяч изображений фона различного размера. На этих данных проводилось тестирование метода.

Результаты классификации

В таблице 1 приведены результаты сравнения работы метода при различных признаках и классификаторах. В столбцах таблицы указаны используемые признаки, в строках парами указаны классификатор и результаты использования данного классификатора с признаками. В ячейках указаны следующие значения: AUC-ROC, ошибка II рода, ошибка I рода.

Таблица 1. Сравнение классификаторов (В ячейке таблицы три числа: точность, ошибка II рода, ошибка I рода сверху вниз соответственно. Столбцы «И» и «Н» означают «Исходные» и «Нормализованные» соответственно.)

Интенсивность и цветовых каналов	HOG-дескриптор		LBP- гистограм ма	Данные сокращённой размерности	Моменты
	И	Н			
Метод k ближайших соседей					
0.94945	0.92540	0.89688	0.82754	0.95046	0.969
0.05087	0.02585	0.01167	0.26939	0.05087	0.038
0.18407	0.28606	0.38723	0.21475	0.17827	0.147
Случайный лес					
0.97749	0.95263	0.94688	0.88284	0.96561	0.979
0.12010	0.06422	0.06588	0.22518	0.16930	0.023
0.04892	0.15837	0.18325	0.18656	0.04063	0.169
XGBoost					
0.97559	0.95581	0.94951	0.89322	0.97293	0.975
0.08256	0.07673	0.07673	0.21100	0.08340	0.041
0.07794	0.13681	0.15174	0.18988	0.08126	0.153
Логистическая регрессия					
0.94078	0.96292	0.94795	0.90225	0.94226	0.895
0.11009	0.06922	0.07506	0.18348	0.10341	0.075
0.14013	0.12106	0.15091	0.18407	0.14013	0.326

Лучшие результаты по точности показала пара признаки, основанные на моментах, и случайный лес, данная пара показала наилучшее значение AUC-ROC, низкое значение ошибки 1 рода и приемлемые значения ошибки 2 рода.

Результаты идентификации

Для тестирования метода идентификации нейронов было выбрано 60 изображений, с размеченными на них нейронами, относящиеся к 5 группам снимков, имеющих различие между собой.

Среди тестовых изображений была выбрана контрольная выборка, на которой было проведено тестирование всех признаков с классификатором случайный лес. Результаты идентификации нейронов на контрольной выборке приведены в таблице 2. В столбцах таблицы указаны используемые признаки, в строках указаны два значения: отношение числа найденных нейронов к общему числу найденных объектов; число найденных нейронов к числу размеченных нейронов на изображении.

Таблица 2. Результаты идентификации нейронов на контрольной выборке изображений

	Интенсивности цветовых каналов	НОГ- дескриптор	LBP- гистограмма	Данные сокращённой размерности	Моменты
Нейронов из найденных объектов	0.6262	0.5608	0.4270	0.5157	0.6824
Найдено нейронов	0.9328	0.8543	0.6065	0.9356	0.7939

Пара - признаки, основанные на моментах, и случайный лес показавшая лучший результат при классификации, при идентификации показала средние результаты. Но показало самый высокий результат отношения найденных нейронов к общему числу найденных объектов, это говорит о том, что метод совершает наименьшую ошибку II рода по сравнению с другими признаками.

Результаты идентификации по группам

Для тестирования метода идентификации нейронов было выбрано 60 изображений, с размеченными на них нейронами, относящиеся к 5 группам снимков, имеющих различие между собой.

Для каждой группы изображений был проведен анализ результатов, ошибок, совершаемых методом и влияние группы изображения на результат.

Характеристики группы изображений:

- Группа 1 (Рис. 5) – изображения характеризуются выраженным переходом от тканей к границам нейронов, с просматриваемым ядром нейрона. В среднем на изображениях данной группы содержится около 80 нейронов.
- Группа 2 (Рис. 6) – изображения характеризуются плохо различимым переходом от тканей к границам нейронов, ядрами нейронов, сливающимися с окружающей его тканями. Изображения данной группы содержат скудное число нейронов, не превышающих 50 нейронов на изображение.
- Группа 3 (Рис. 7) – изображения характеризуются выраженным переходом от тканей к границам нейронов, ядрами нейронов, сливающимися с границами нейронов. Также на данных изображениях наблюдается высокая плотность нейронов, что приводит к специальному случаю - «слипшиеся нейроны». В среднем на изображениях содержится более 120 нейронов.
- Группа 4 (Рис. 8) – изображения характеризуются плохо различимым переходом от тканей к границам нейронов, частичным отсутствием границ нейронов. В среднем на изображениях содержится более 80 нейронов.
- Группа 5 (Рис. 9) – изображения характеризуются плохо различимым переходом от тканей к границам нейронов, но с хорошо просматриваемым ядром нейрона. В среднем на изображениях содержится более 80 нейронов.

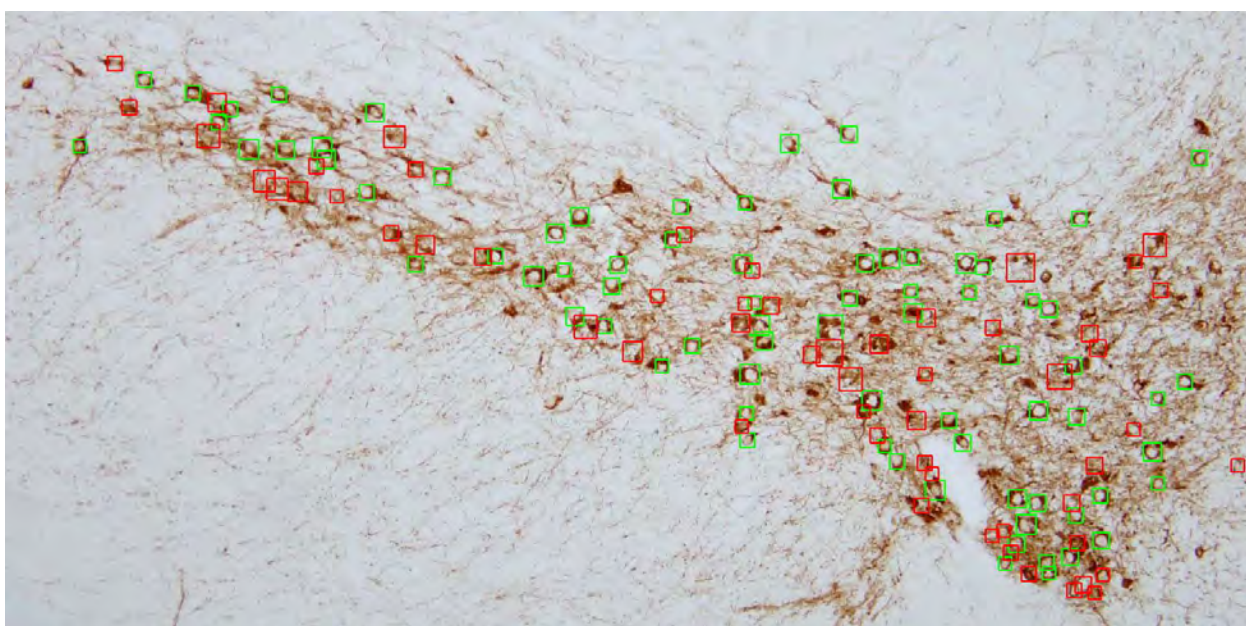


Рис. 5. Изображение Группы 1

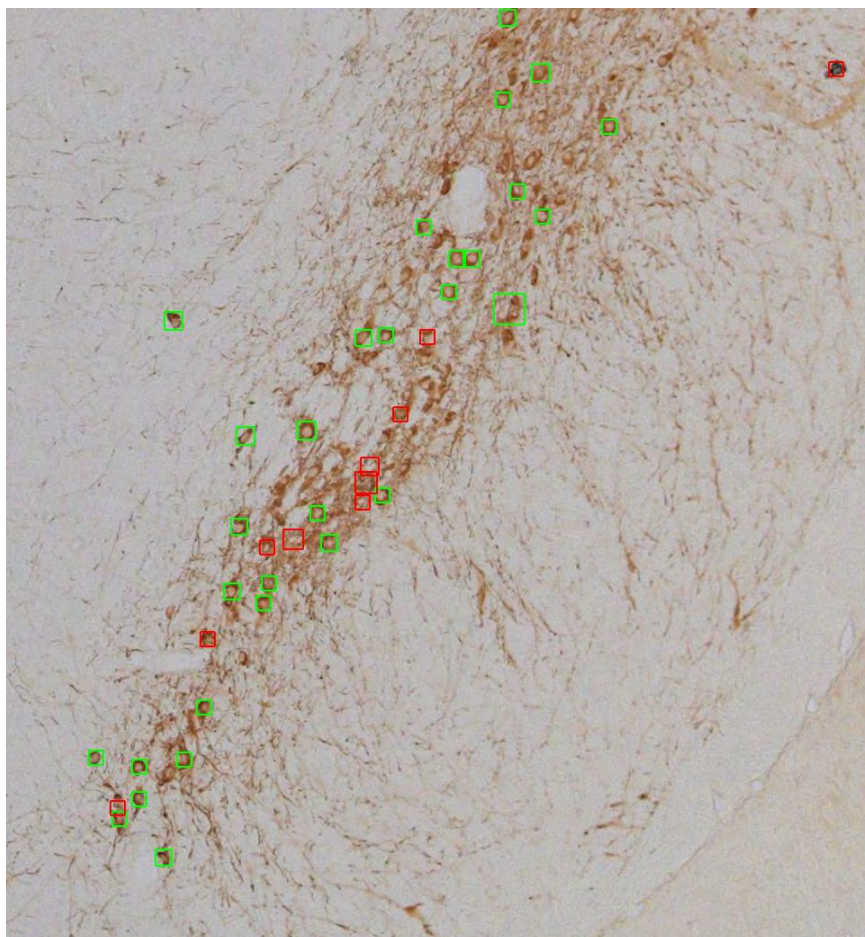


Рис. 6. Изображение Группы 2

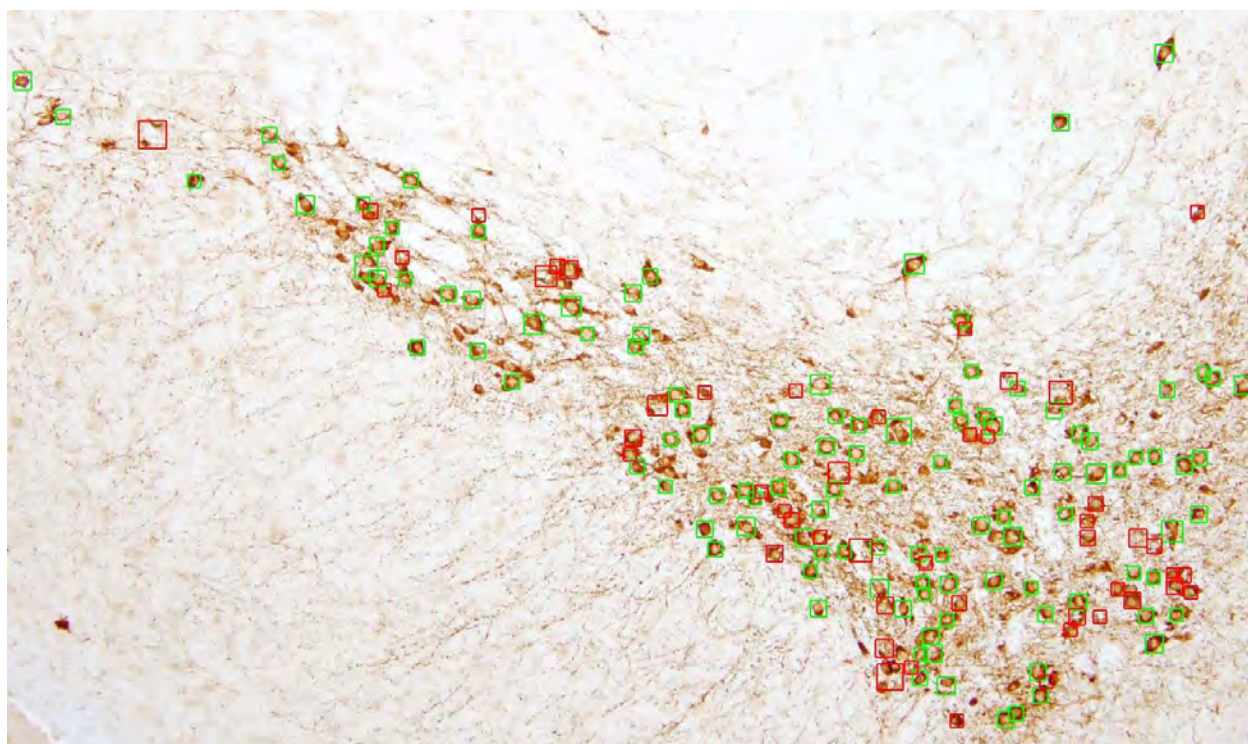


Рис. 7. Изображение Группы 3

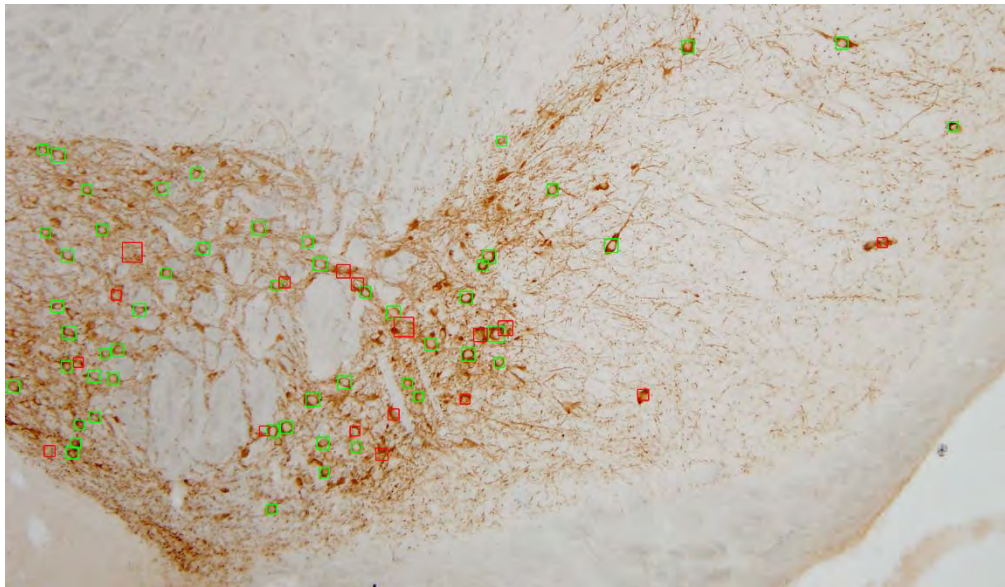


Рис. 8. Изображение Группы 4

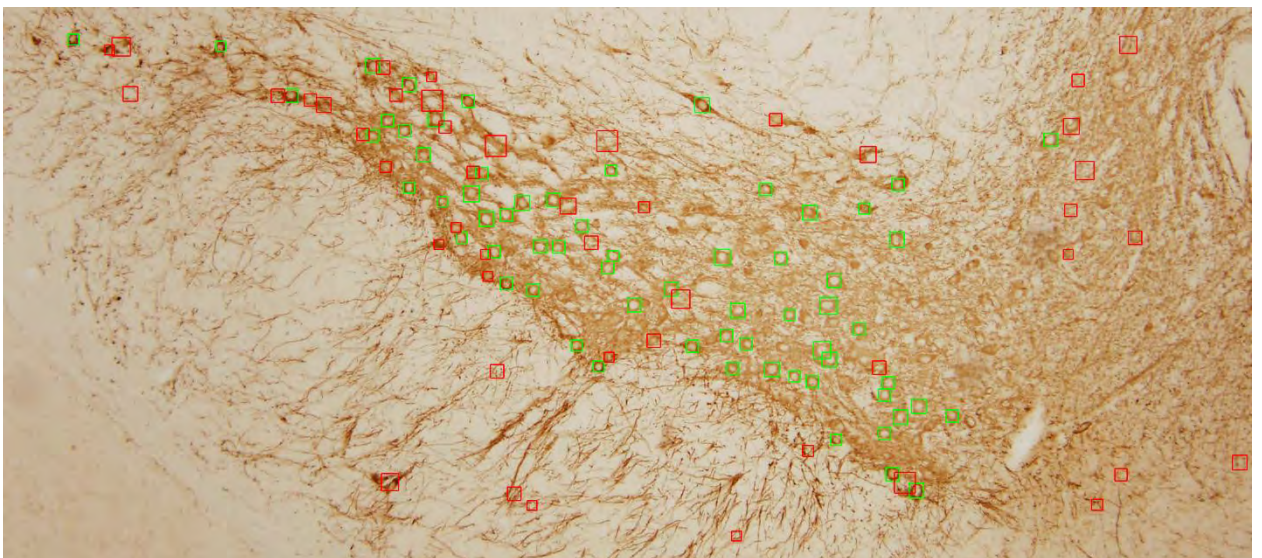


Рис. 9. Изображение Группы 5

На Рис. 5-9 изображены результаты применения метода идентификации нейронов на МИН на изображениях, относящихся к различным группам. На изображениях зелеными квадратами выделены найденные объекты, которые являются нейронами, красными квадратами выделены найденные объекты, которые не являются нейронами (участки фона, сгустки ткани и т.д.)

В силу плохой отличимости нейронов от тканевой составляющей и плохой видимости ядра на изображениях, относящихся к группам 2 и 4 метод идентификации нейронов допускает высокую ошибку 1 рода, в результате чего не на изображения данных групп находится в среднем не более 70% всех нейронов.

На изображениях, относящихся к группе 3, также возникают проблемы в результате особенностей изображения. В результате того что на данных изображениях наблюдается высокая плотность нейронов, приводящая к специальному случаю «слипшиеся нейроны», на данных изображениях часто наблюдается случаи, когда только один или несколько нейронов из группы слипшихся нейронов детектируются методом.

Основные результаты применения метода идентификации нейронов:

- наиболее часто метод идентифицирует нейроны среднего размера, на нейронах малого и большого размера допускается основная часть ошибок;
- в ряде случаев метод идентифицирует в качестве нейронов объекты изображения очень сильно похожие на нейроны, но которые не были размечены;
- при возникновении специального случая «слипшиеся нейроны», метод идентифицирует не все нейроны, входящие в группу слипшихся нейронов.

Для изображений нейронов и фона выделенных скользящим окном на МИН была проведена кластеризация. Оказалось, что некоторые размеченные нейроны очень схожи с объектами фона, содержащими сгустки ткани. Многочисленные встречи с экспертами также показали, что врач-эксперт порой не знает с 100% точностью относится тот или иной объект к нейронам.

Заключение

Использование информационных технологий и поддерживающих их АПК позволит увеличить производительность труда врачей-клиницистов и снизить затраты на подготовку экспериментального материала. Кроме того, наличие информационных технологий открывает возможность ставить и решать такие задачи научных исследований, решение которых теряет смысл при отсутствии возможности автоматизировать обработку и анализ получаемых результатов в силу их высокой и специфической информационной насыщенности, сложности и колоссальных объемов экспериментальных данных. В целом создание информационных технологий анализа медицинских и биологических изображений имеет определяющее значение для совершенствования медицинской диагностики в клинических учреждениях и существенному повышению качества медицинской помощи.

В работе был описан новый метод идентификации нейронов на микроскопических изображениях. Проведенные эксперименты показали, что метод показывает приемлемые результаты. Позволяющие за счет выделения на обрабатываемых изображениях нейронов и объектов окружения похожих на нейроны значительно снизить общий объем информации, необходимый проанализировать врачу-эксперту для принятия определенных решений и выводов.

Основные результаты работы:

1. Разработка метода идентификации нейронов на микроскопических изображениях, основанного на новой системе признаков.
2. Программная реализация разработанного метода.
3. Анализ признаков, используемых для идентификации нейронов, на микроскопических изображениях.
4. Формирование набора признаков, основанного на методе моментов.
5. Экспериментальные исследования и анализ результатов разработанного метода.

В дальнейшем планируется улучшить результаты за счет увеличения обучающей выборки при поступлении новых изображений. Улучшить результаты планируется также путем увеличения диапазона размеров окон для скользящего окна, что позволит лучше идентифицировать нейроны малого и большого размеров. Провести эксперименты по решению проблемы специального случая «слипшиеся нейроны». Предполагается также использовать результаты кластерного анализа для разработки мультиалгоритмических классификаторов, специфичных для разных классов нейронов.

Литература

- [1] Alavi A., Cavanagh B., Tuxworth G., Meedeniya A., Mackay-sim V. Automated classification of dopaminergic neurons in the rodent brain // Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, June 14-19, 2009. P. 81-88.
- [2] Benali A., Leefken I., Eysel U.T., Weiler E. A computerized image analysis system for quantitative analysis of cells in histological brain sections // Journal of Neuroscience Methods. 2003. Vol.125, № 1-2. P. 33-43.
- [3]
- [4] Bernd J. Digital Image Processing Hardcover. Springer; 7th ed. Edition, 2012
- [5] Breiman, Leo. «Random Forests». Machine Learning 45 (1), 2001. pp. 5–32.
- [6] Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection. 2005.
- [7] DC. He, L. Wang. "Texture Unit, Texture Spectrum, And Texture Analysis", Geoscience and Remote Sensing, IEEE Transactions on, vol. 28, 1990. pp. 509 - 512.
- [8] Friedman, J. H. "Greedy Function Approximation: A Gradient Boosting Machine." 1999.
- [9] Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning, 2nd edition. - Springer, 2009. - 533 pp.
- [10] I.B. Gurevich, Yu. I. Zhuravlev, A. A. Myagkov, Yu. O. Trusova, and V. V. Yashina. On Basic Problems of Image Recognition in Neurosciences and Heuristic Methods for Their Solution // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications. Pleiades Publishing, Ltd., 2015. - Vol. 25, No. 1. – P.132–160.
- [11] I.B. Gurevich, Yu.I. Zhuravlev, A. A. Myagkov, A.A. Kolacheva, Yu.O. Trusova, V.V. Yashina, T.S. Pronina, M.V. Ugryumov. Automation of analysis and evaluation of images in brain research: statement of problems, mathematical foundations, information technologies / Neurodegenerative diseases: from the genome to the whole organism: in 2 volumes. - Moscow: The scientific world, 2014.
- [12] I.B. Gurevich, Yu.I. Zhuravlev. Mathematical and informational aspects of brain research. Herald of the Russian Academy of Sciences, 2010. – Vol.80, № 5–6. - pp. 503–513.
- [13] Inglis A., Cruz L., Roe D.L., Stanley H.E., Rosene D.L., Urbanc B. Automated identification of neurons and their locations // Journal of Microscopy. 2008. Vol. 230. № 3 P. 339-52.
- [14] Jolliffe, I. T. Principal Component Analysis. Springer-Verlag, 1986. p. 487.
- [15] Meijering E. Neuron Tracing in Perspective // Cytometry. Part A. 2010. Vol. 77. №7. P. 693-704.
- [16] Peng S., Urbanc B., Cruz L., Hyman B.T., Stanley H.E. Neuron Recognition by Parallel Potts Segmentation // Proceedings of the National Academy of Sciences of the United States of America. 2003. Vol. 100. № 7. P. 3847-3852.
- [17] Sciarabba M., Serrao G., Bauer D., Arnaboldi F., Borghese N.A. Automatic Detection of Neurons in Large Cortical Slices // Journal of Neuroscience Methods. 2009. Vol. 182. P. 123-140.
- [18] Wang Y.-Y., Sun Y.-N., Chou-Ching K., Ju M.-S. Nerve Cell Segmentation via Multi-Scale Gradient Watershed Hierarchies // Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2006. P. 6698-6701.
- [19] Yu.I. Zhuravlev, V. V. Ryazanov, O. V. Sen'ko. RECOGNITION: Mathematical methods, software system, practical solutions. M.: Phasis, 2006.