

На правах рукописи



Сенюкова Ольга Викторовна

**Разработка алгоритмов семантической
сегментации и классификации биомедицинских
сигналов низкой размерности на основе
машинного обучения**

Специальность 05.13.11 – математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2012

Работа выполнена на кафедре автоматизации систем вычислительных комплексов факультета вычислительной математики и кибернетики Московского государственного университета имени М.В. Ломоносова.

Научный руководитель: *кандидат физико-математических наук,
доцент
Баяковский Юрий Матвеевич*

Официальные оппоненты: *доктор физико-математических наук,
профессор
Рычагов Михаил Николаевич*

*кандидат физико-математических наук,
доцент
Галатенко Владимир Владимирович*

Ведущая организация: *Институт прикладной математики
имени М.В. Келдыша Российской академии наук*

Защита состоится 19 октября 2012 г. в 11 часов на заседании диссертационного совета Д 501.001.44 при факультете вычислительной математики и кибернетики Московского государственного университета имени М.В. Ломоносова по адресу: 119991, ГСП-1, Москва, Ленинские горы, 2-й учебный корпус, факультет ВМК, аудитория 685. Желающие присутствовать на заседании диссертационного совета должны сообщить об этом за два дня по тел. (495)939-30-10 (для оформления заявки на пропуск).

С диссертацией можно ознакомиться в Фундаментальной библиотеке МГУ имени М.В. Ломоносова. С текстом автореферата можно ознакомиться на официальном сайте ВМК МГУ <http://cs.msu.ru> в разделе «Наука» – «Работа диссертационных советов» – «Д 501.001.44».

Автореферат разослан «___» сентября 2012 г.

Заместитель председателя
диссертационного совета,

профессор



В.М. Круглов

Общая характеристика работы

Объект исследования и актуальность работы

Алгоритмы автоматического анализа сигналов постепенно внедряются во всевозможные области человеческой деятельности, в частности, медицину и биологию. В настоящее время существуют алгоритмы, способные классифицировать данные и распознавать объекты на изображениях. В биомедицинских приложениях это находит применение для автоматизации процесса описания данных и предварительной диагностики, что особенно актуально при больших объемах данных и сложности визуального анализа самих данных. Рост вычислительных мощностей способствует появлению более точных алгоритмов.

В рамках диссертационной работы рассматриваются биомедицинские сигналы низкой размерности — двумерные изображения и одномерные сигналы.

Существующие биомедицинские технологии позволяют получать достаточно качественные изображения внутренних органов в виде трехмерных изображений или набора последовательных двумерных срезов. Одна из актуальных задач, связанных с обработкой полученных таким образом изображений, — семантическая сегментация, т.е. выделение на изображениях объектов, относящихся к заданному классу. Это могут быть, например, задачи сегментации изображений на анатомические структуры, выделения патологических образований.

В задаче сегментации изображения участка тела на анатомические структуры может использоваться атлас — набор изображений усредненной модели участка тела с анатомической разметкой. Но при этом необходимо учитывать индивидуальные особенности субъекта и трансформировать разметку из атласа в соответствии с этими особенностями.

Задача выделения патологических образований характерна тем, что для патологии сложнее построить модель. Все существующие атласы описывают здоровые органы и ткани. Патологические образования могут сильно варьироваться по форме и другим характеристикам, и зачастую нельзя сказать ничего конкретного об их местоположении и взаимном расположении.

Задачи, связанные с обработкой одномерных сигналов, возникают, в частности, в кардиологии и включают в себя классификацию данных на случай здорового пациента и случаи различных заболеваний, а также индикацию различных физиологических состояний.

Большинство существующих алгоритмов используют машинное обучение. Но для обучения классификаторов обычно необходимы большие обучающие базы, которые требуют тщательной разметки экспертом. Особенно трудоемким этот процесс оказывается при разметке изображений. Кроме того, в некоторых задачах получение большого объема самих данных для обучаю-

щей базы либо невозможно, либо сопряжено с существенными трудностями.

В диссертационной работе предложены новые алгоритмы семантической сегментации и классификации, основанные на машинном обучении, которые упрощают обучение классификатора либо за счет введения новой эффективной процедуры разметки обучающей базы, либо за счет возможности обучения по одному прецеденту. Разработанные алгоритмы были применены к конкретным биомедицинским задачам и продемонстрировали точность либо выше аналогов, либо сравнимую с аналогами. Эти алгоритмы также применимы и к другим задачам обработки изображений.

Цель диссертационной работы

Целью диссертационной работы является разработка и программная реализация алгоритмов семантической сегментации и классификации биомедицинских сигналов низкой размерности.

Научная новизна

Разработанный в рамках диссертации алгоритм выделения на изображениях однородных объектов с нечеткими границами и похожими характеристиками, основанный на классификации изолиний функции интенсивности, является новым и превосходит аналоги по точности в рамках поставленной задачи.

Алгоритм разбиения изображений на заданный набор областей, неоднородных по интенсивности, является новым. В отличие от аналогов, он может работать при наличии единственного обучающего размеченного изображения за счет введения нового метода расчета априорной пространственной информации и демонстрирует точность, сравнимую с аналогами для похожих задач.

Алгоритм классификации с обучением по одному прецеденту основан на новом методе “декомпозиции ансамблей классификаторов”. В отличие от существующих алгоритмов классификации с обучением по одному прецеденту, алгоритм не требует наличия классификаторов для других классов, построенных по полноценным обучающим базам, а требует лишь наличие обобщенного двухклассового классификатора. В рамках поставленной медицинской задачи алгоритм превосходит аналоги по точности.

Практическая значимость

Предложенный алгоритм выделения на изображениях однородных объектов с нечеткими границами и похожими характеристиками позволяет выделять широкий класс патологических образований на медицинских изображениях. Процедура разметки обучающей базы является более быстрой и удоб-

ной, чем у аналогов: вместо попиксельной ручной разметки необходимо среди автоматически построенных изолиний функции интенсивности указать те из них, которые являются границами искомых объектов. Программная система, разработанная на основе построенного алгоритма, используется в НИИ Неотложной Детской Хирургии и Травматологии (НИИ НДХиТ).

Разработанный алгоритм разбиения изображений на заданный набор областей, неоднородных по интенсивности, может применяться для сегментации изображений гистологических срезов на анатомические структуры, так как он устойчив к перепадам интенсивности между соседними пикселями внутри областей. Алгоритм может работать при наличии единственного обучающего размеченного изображения. Это чрезвычайно важно для экспериментов, где процесс получения и обработки данных представляется особенно трудоемким, например, получение изображений гистологических срезов мозга мыши. Разработка алгоритма проводилась совместно с НИИ Нормальной Физиологии имени П.К. Анохина (НИИ НФ).

Разработанный алгоритм классификации с обучением по одному прецеденту при наличии обобщенного двухклассового классификатора может использоваться для классификации данных на подклассы этих двух классов. Например, если имеется классификатор “больной/здоровый”, построенный по кардиоинтервалограммам и являющийся ансамблем классификаторов, то на базе него можно построить классификатор для конкретного заболевания или физиологического состояния по одному обучающему примеру. Это особенно актуально для задач определения редких заболеваний и конкретных физиологических состояний, для которых трудно или невозможно подготовить большую обучающую базу. Разработанный алгоритм планируется для использования прежде всего в профессиональном спорте, в частности, для определения состояний перетренировки и оптимального функционирования спортсмена.

Апробация работы

Основные результаты диссертации докладывались и обсуждались на:

- 19-й международной конференции по компьютерной графике и зрению “GraphiCon’2009”, Россия, Москва, 2009;
- 12-й всероссийской научно-технической конференции “Нейроинформатика-2010”, Россия, Москва, 2010;
- 20-й международной конференции по компьютерной графике и зрению “GraphiCon’2010”, Россия, Санкт-Петербург, 2010;

- 7-й международной научно-практической конференции по психологии спорта и физической культуры “Рудиковские чтения - 2011”, Москва, Россия, 2011;
- 21-й международной конференции по компьютерной графике и зрению “GraphiCon’2011”, Россия, Москва, 2011;
- международном симпозиуме “Computational Models for Life Sciences” (CMLS-11), Япония, Тояма, 2011;
- 6-й международной конференции “Computational Intelligence and Bioinformatics” (CIB 2011), США, Питтсбург, 2011;
- международной конференции “International Conference on Computer and Computational Intelligence” (ICCCI 2011), Таиланд, Бангкок, 2011;
- семинаре лаборатории Компьютерной Графики и Мультимедиа факультета ВМиК МГУ под руководством доц. Ю.М. Баяковского;
- семинаре кафедры Автоматизации Систем Вычислительных Комплексов под руководством чл.-корр. РАН Л.Н. Королева;
- международном симпозиуме “Нейроимиджинг и магнитоэнцефалография: фундаментальные исследования и клиническая практика”, Россия, Москва, 2012;
- семинаре направления “Программирование” им. М.Р. Шура-Бура Института Прикладной Математики им. М.В. Келдыша.
- семинаре Исследовательского центра Samsung Electronics Co., Ltd.

Публикации

Автором опубликовано 18 научных работ, в том числе 12 по теме диссертации: 2 статьи в рецензируемых журналах, рекомендованных ВАК [1–3], 7 статей в сборниках трудов международных конференций [4–10], 1 статья в сборнике трудов всероссийской конференции [11], а также 2 тезисов доклада [12, 13].

Структура и объем диссертации

Диссертация состоит из введения, 4 глав, заключения и библиографии. Общий объем диссертации составляет 133 страницы, включая 34 рисунка. Библиография включает 125 наименований.

Содержание работы

Во введении обоснована актуальность диссертационной работы, сформулированы цели и задачи, аргументирована научная новизна исследований, показана практическая значимость полученных результатов. Описана структура диссертации.

В первой главе приводятся примеры биомедицинских сигналов низкой размерности и особенности их обработки. Дается обзор существующих алгоритмов сегментации биомедицинских изображений и классификации одномерных медицинских сигналов.

Биомедицинские изображения – это обычно изображения внутренних органов, полученные с помощью специальной аппаратуры, например, изображения магнитно-резонансной томографии, компьютерной томографии, гистологических срезов. Одномерные медицинские сигналы в основном используются в кардиологии. Множество работ посвящено автоматическому анализу кардиоинтервалограмм – временных рядов, представляющих собой график изменения длины интервалов между последующими сокращениями сердца с течением времени.

В настоящее время для решения задач семантической сегментации и классификации биомедицинских сигналов активно применяется машинное обучение. Общая постановка задачи машинного обучения заключается в следующем. Имеется множество объектов, $x_i, i = 1, \dots, m$, каждый из которых описан с помощью набора признаков $f_j(x), j = 1, \dots, n$:

$$\mathbf{x} = \{f_1(x), \dots, f_n(x)\}. \quad (1)$$

Вектор \mathbf{x} называется *вектор-признаком* объекта x .

Каждому объекту x , заданному своим вектор-признаком \mathbf{x} , поставлена в соответствие метка класса $y_k, k = 1, \dots, l$:

$$X^m = \{(\mathbf{x}_1, y_{i_1}), \dots, (\mathbf{x}_m, y_{i_m})\}. \quad (2)$$

Множество таких пар называется *обучающей выборкой*.

Требуется по обучающей выборке построить классификатор – алгоритм, осуществляющий отображение $\mathbf{X} \rightarrow \mathbf{Y}$, который может поставить в соответствие произвольному объекту x , заданному вектор-признаком $\mathbf{x} \in \mathbf{X}$, метку класса $y \in \mathbf{Y}$.

В задаче семантической сегментации изображений происходит классификация пикселей. В вектор-признак пикселя могут входить его интенсивности на различных модальностях магнитно-резонансной томографии, если данные мультимодальные¹, а также расстояния до различных анатомических струк-

¹ De Boer R., Vrooman H. A., Van der Lijn F. et al. White matter lesion extension to automatic brain tissue segmentation on MRI // Neuroimage, 2009. Vol. 45, pp. 1151-1161

тур, информация об окрестности пикселя, информация о симметрии и другие признаки ^{2 3}.

В задаче анализа кардиоинтервалограмм классифицируется весь исследуемый временной ряд. Алгоритмы, основанные на машинном обучении, в качестве признаков используют выходы различных линейных и нелинейных индикаторов, рассчитанных по временному ряду ^{4 5}.

Во второй главе описываются разработанные алгоритмы семантической сегментации и классификации.

В первом разделе описывается предложенный новый алгоритм выделения на изображениях однородных объектов с нечеткими границами и похожими характеристиками.

Ключевой идеей алгоритма является построение изолиний функции интенсивности для некоторого диапазона интенсивностей, среди которых есть линии, являющиеся контурами искомых объектов. Изолиния уровня h функции $f(x, y)$ есть геометрическое место точек (x, y) , таких что $f(x, y) = h$.

Контурами искомых объектов могут быть только замкнутые изолинии. Для идентификации этих изолиний среди всех построенных замкнутых изолиний используется двухклассовая классификация методом опорных векторов с нелинейным ядром, предложенным Б. Босером, И. Гийоном и В. Вапником в 1992 году ⁶. Метка класса может принимать одно из двух значений: $y_i \in \{+1, -1\}$. Метод основан на построении гиперплоскости, максимизирующей отступ между элементами обоих классов из обучающей выборки (2):

$$f(\mathbf{x}) = \sum_{i=1}^m \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b. \quad (3)$$

Веса α_i не равны 0 только для *опорных* векторов – элементов \mathbf{x}_i , принадлежащих выпуклой оболочке одного из классов.

В качестве ядровой функции K используется Гауссова радиальная базисная функция (RBF):

$$K(\mathbf{x}, \mathbf{y}) = \exp\left\{-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}\right\}. \quad (4)$$

² Geremia E., Menze B. H., Clatz O. et al. Spatial decision forests for MS lesion segmentation in multi-channel MR images // Proceedings of MICCAI, 2010. Vol. 13, pp. 111-118

³ Scully M., Anderson B., Lane T. et al. An Automated Method for Segmenting White Matter Lesions through Multi-level Morphometric Feature Classification with Application to Lupus // Frontiers in Human Neuroscience, 2010. Vol. 4, no. 27

⁴ Kampouraki A., Manis G., Nikou C. Heartbeat Time Series Classification With Support Vector Machines // IEEE Trans. on Information Technology in Biomedicine, 2009. Vol. 13, no. 4, pp. 512-518

⁵ Jovic A., Bogunovic N. Classification of biological signals based on nonlinear features // Proceedings of MELECON, 2010. Pp. 1340-1345

⁶ Boser B. E., Guyon I. M., Vapnik V. N. A training algorithm for optimal margin classifiers // Proc. of the 5th Conference on Computational Learning Theory. 1992. Pp. 144-152

Каждая замкнутая изолиния задается вектор-признаком \mathbf{x} из 6 признаков: морфологические признаки фигуры, ограниченной изолинией (площадь, компактность, периметр, эксцентриситет) и признаки области изображения, в которой находится изолиния, основанные на интенсивности (разность интенсивностей вдоль внешней и внутренней границ замкнутой изолинии, отношение средней интенсивности внутри замкнутой линии к средней интенсивности исследуемого региона).

При классификации методом опорных векторов используется линейная комбинация ядер с отдельным ядром по каждому признаку ⁷.

С целью повышения устойчивости алгоритма пользователем указываются прямоугольные регионы на изображении, которые содержат объекты интереса.

Во втором разделе приводится описание разработанного автоматического алгоритма разбиения изображений на заданный набор областей, неоднородных по интенсивности, при наличии единственного обучающего размеченного изображения.

Изображение, которое необходимо разметить, будем называть *тестовым изображением*. Размеченное изображение, по которому алгоритм будет обучаться, назовем *обучающим изображением*.

По обучающему изображению строится многоклассовый классификатор рандомизированный решающий лес ⁸, который может присваивать пикселям метки областей. Для того чтобы алгоритм успешно справлялся с неоднородностями областей и извлекал из них полезную информацию, обучающее и тестовое изображения разбиваются на суперпиксели (мелкие сегменты) с помощью алгоритма сдвига среднего (mean shift). Для классификации пикселей используется два типа признаков: признаки суперпикселя, которому принадлежит пиксель ⁹, и априорная пространственная информация, для расчета которой используется новый метод.

Априорная пространственная информация – вектор априорных вероятностей принадлежности пикселя $X = (x, y)$ каждой из областей $s_i, i = 1, \dots, n$, исходя только из его местоположения:

$$L = [P(a(X) = s_1), P(a(X) = s_2), \dots, P(a(X) = s_n)], \quad (5)$$

где $a(X)$ – это функция отображения пространства пикселей изображения в пространство меток областей. Предполагая, что обучающее и тестовое изображения имеют похожую конфигурацию – одинаковый набор областей, расположение границ которых различается в определенных пределах, вероятности

⁷ Lampert C. H. Kernel Methods in Computer Vision // Foundations and Trends in Computer Graphics and Vision, 2009. Vol. 4, no. 3. Pp. 193-285

⁸ Breiman L. Random Forests // Machine Learning. 2004. Vol. 45, no. 1. Pp. 5–32

⁹ Sudakov S., Barinova O., Velizhev A., Konushin A. Semantic segmentation of road images based on cascade classifiers // Proceedings of ISPRS. 2008. Pp. 601–604.

можно рассчитать следующим образом:

$$P(a(X) = s_i) = \begin{cases} 1, & a(Y) = s_i, \forall Y \in U(X) \\ 0, & \nexists Y \in U(X) : a(Y) = s_i \\ \frac{1}{\sum s_j}, & \exists Y_1 : a(Y_1) = s_i \wedge \dots \wedge \exists Y_m : a(Y_m) = s_k, \\ & Y_1, \dots, Y_m \in U(X) \end{cases} \quad (6)$$

$U(X) = \{Y : |Y - X| \leq R\}$, т.е. $U(x)$ есть круглая окрестность радиуса R с центром в X .

Из (6) следует, что если в окрестность пикселя входят несколько областей, то априорная вероятность принадлежности пикселя каждой из них распределяется поровну между этими областями. Благодаря этому результат сегментации становится ближе к требуемому результату и дальше от разметки обучающего изображения, чем в случае распределения априорных вероятностей между областями в зависимости от доли пикселей этих областей, попавших в окрестность.

Перед применением классификатора к тестовому изображению, оно подвергается процедуре предварительной обработки.

Для уточнения результата классификации с учетом взаимного расположения пикселей вероятностный выход классификатора встраивается в Марковское случайное поле. Задача нахождения наиболее вероятного распределения меток областей на изображении сводится к задаче минимизации функционала энергии:

$$E(T|X) = \sum_{(i,j) \in N} E_{ij}(t_i, t_j) + \sum_i E_i(x_i, t_i) \rightarrow \min_T, \quad (7)$$

где x_j это наблюдаемые переменные — пиксели, а t_i — скрытые переменные, их метки областей.

Унарный член в (7) отвечает за меру соответствия пикселя каждой из областей:

$$E(x, t) = -\log_{10} C, \quad (8)$$

где C — вероятностный выход рандомизированного решающего леса.

Бинарный член в (7) отражает сходство меток классов соседних пикселей. В настоящей работе используется модель Поттса, где несовпадение меток классов соседних пикселей штрафует константным значением 1:

$$E_{ij}(t_i, t_j) = 1 - \delta(t_i, t_j) \quad (9)$$

В третьем разделе описан разработанный алгоритм классификации, который может работать при наличии единственного обучающего примера.

Алгоритм основан на новом методе “декомпозиции ансамблей классификаторов”. Ансамбль классификаторов представляет собой набор некоторых базовых классификаторов $h_t(\mathbf{x})$ с соответствующими весами α_t . В задаче двухклассовой классификации для отнесения прецедента \mathbf{x} к одному из классов $+1$ или -1 , как правило, используется знак суммарного выхода ансамбля классификаторов $H(\mathbf{x})$:

$$H(\mathbf{x}) = \sum_{t=1}^T \alpha_t h_t(\mathbf{x}) / \sum_{t=1}^T \alpha_t. \quad (10)$$

Тем не менее, полезная информация о выходах каждой из составляющих $h_t(\mathbf{x})$ остается неиспользованной. Например, в случае алгоритма усиления слабых классификаторов AdaBoost каждый из классификаторов $h_t(\mathbf{x})$ является экспертом в определенной области пространства признаков. Предложенный метод “декомпозиции ансамблей классификаторов” позволяет учесть эту информацию.

После применения ансамбля классификаторов (обобщенного двухклассового классификатора) к обучающему примеру некоторого класса можно построить вектор “декомпозиции ансамбля”, который можно рассматривать в качестве описания обучающего примера:

$$D(\mathbf{x}) = [\alpha_1 h_1(\mathbf{x}), \alpha_2 h_2(\mathbf{x}), \dots, \alpha_T h_T(\mathbf{x})]. \quad (11)$$

Для того чтобы классифицировать новый объект \mathbf{x}_1 , достаточно сравнить вектор декомпозиции ансамбля для этого объекта $D_1 = D(\mathbf{x}_1)$ с вектором декомпозиции ансамбля для обучающего примера $D_0 = D(\mathbf{x}_0)$. Если $\|D_1 - D_0\|_2 < \delta$, где δ – пороговое значение, рассчитанное в ходе экспериментов, то объекты можно считать принадлежащими одному и тому же классу.

В диссертационной работе в качестве ансамбля классификаторов используется алгоритм усиления слабых классификаторов AdaBoost, где в роли слабых классификаторов выступают индикаторы из нелинейной динамики – детрендовый анализ флуктуаций и мультимасштабная энтропия.

В третьей главе дается экспериментальная оценка разработанных алгоритмов.

В первом разделе описывается тестирование предложенного алгоритма выделения однородных объектов, который был применен к задаче выделения очагов поражения головного мозга на изображениях магнитно-резонансной томографии, предоставленных НИИ НДХиТ. На тех же данных для тех же регионов интереса предложенный алгоритм сравнивался с алгоритмом на основе метода кластеризации ISODATA¹⁰. Размер всех изображений – 512×512 пикселей.

¹⁰ Hillary F. G., Biswal B. B. Automated Detection and Quantification of Brain Lesions in Acute Traumatic Brain Injury Using MRI // Brain Imaging and Behavior, 2009. Vol. 3, pp. 111-122

Результаты тестирования и сравнения приведены в таблице 1 и на рисунке 1 а). Все показатели рассчитывались для изолиний/кластеров.

Таблица 1. Сравнение алгоритмов: ISODATA и ISOLINES (предложенный алгоритм).

Алгоритм	Чувствительность	Избирательность	Мера сходства Дайса
ISODATA	0,64	0,68	0,60
ISOLINES	0,90	0,88	0,89

Как видно из рисунка 1 б), использовать для классификации изолинии лучше, чем края или регионы.

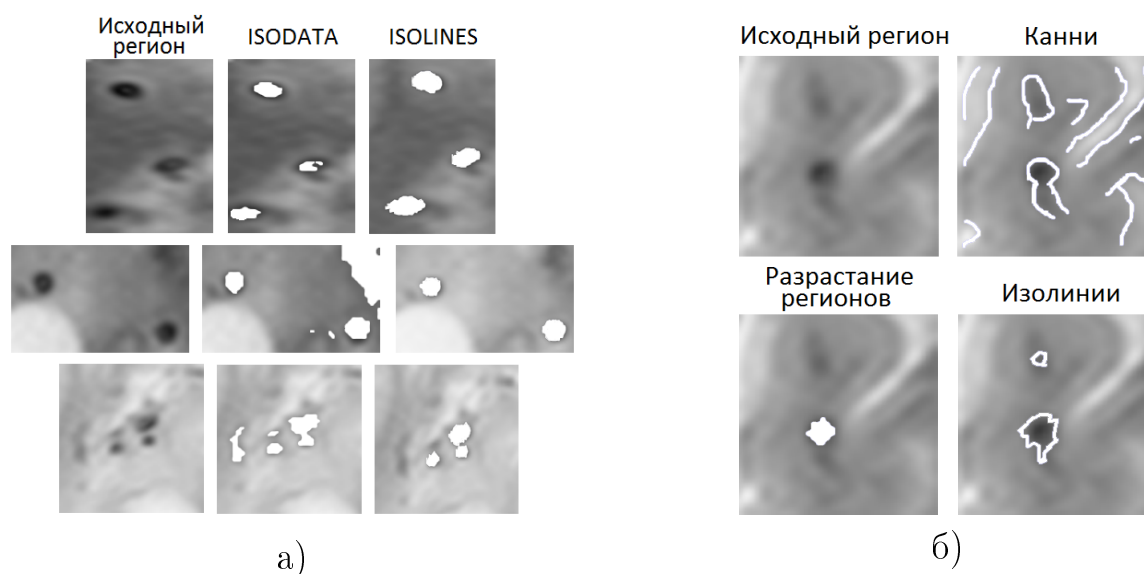


Рис. 1. Сравнение предложенного алгоритма (ISOLINES) с другими алгоритмами.

Во втором разделе приведено описание тестирования алгоритма разбиения изображений на заданный набор областей, который был применен к задаче сегментации гистологических срезов мозга мыши на анатомические структуры. Рассматривается 17 анатомических структур. Алгоритм был протестирован на изображениях из Алленовского атласа мозга мыши <http://mouse.brain-map.org/atlas/index.html> на 10 парах соседних срезов. Размер всех изображений – 270 × 204 пикселей.

Средняя попиксельная точность по всем анатомическим структурам превышает 86%. Полученные результаты сопоставимы с результатами работы существующих алгоритмов сегментации мозга мыши. В частности, алгоритм Bae, 2009 ¹¹ демонстрирует точность порядка 80%.

Алгоритм также был протестирован на изображениях, предоставленных НИИ НФ. Результаты приведены на рисунке 2.

¹¹ Bae M., Pan R., Wu T., Badea A. Automated Segmentation of Mouse Brain Images Using Extended MRF // Neuroimage. 2009. Vol. 46, no. 3. Pp. 717–725

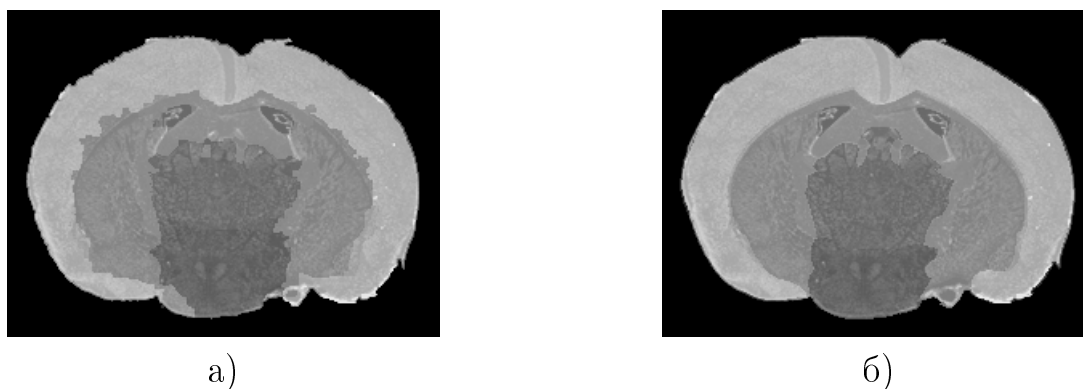


Рис. 2. Сравнение результатов работы предложенного алгоритма с ручной разметкой: а) автоматическая разметка; б) ручная разметка.

В третьем разделе приведено описание тестирования разработанного алгоритма классификации с обучением по одному прецеденту, который был применен к задаче классификации коротких временных промежутков кардиоинтервалограмм (около 15 минут). Для тестирования была использована база данных <http://www.physionet.org/physiobank>. В качестве обучающего примера использовались кардиоинтервалограммы пациентов с различными видами аритмии. Обобщенный классификатор “больной/здоровый” был обучен на данных от здоровых пациентов, пациентов с аритмией и застойной сердечной недостаточностью.

Как видно из рисунка 3, предложенный подход демонстрирует существенно лучшие результаты, чем индикаторы из нелинейной динамики в чистом виде.

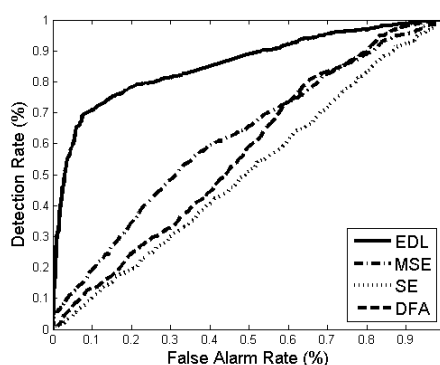


Рис. 3. ROC-кривые для предложенного метода (EDL) и индикаторов из нелинейной динамики (SE – энтропия выборки, MSE – мультимасштабная энтропия, DFA – детрендовый анализ флуктуаций).

В пятой главе описывается программная реализация разработанных алгоритмов.

Разработанные программы могут использоваться на персональных компьютерах. Алгоритм выделения однородных объектов с нечеткими граница-

ми и похожими характеристиками реализован на языке C#. На базе предложенного алгоритма разработано интерактивное приложение. Алгоритм разбиения изображений на заданный набор областей, неоднородных по интенсивности, при наличии единственного обучающего размеченного изображения реализован на языке C++. Алгоритм классификации данных с обучением по одному прецеденту реализован в среде MATLAB.

В заключении представлены основные результаты работы. Основные результаты работы состоят в следующем:

1. Разработан алгоритм выделения на изображениях однородных объектов с нечеткими границами и похожими характеристиками. Алгоритм значительно упрощает процедуру разметки обучающей базы для классификатора. Показано превосходство над аналогами по точности в рамках задачи выделения очагов поражения на изображениях МРТ мозга.
2. Разработан автоматический алгоритм разбиения изображений на заданный набор областей, неоднородных по интенсивности. В отличие от аналогов, алгоритм работает при наличии единственного обучающего размеченного изображения. Алгоритм применим к задачам сегментации изображений гистологических срезов.
3. Разработан алгоритм классификации с обучением по одному прецеденту, который не требует наличия классификаторов для других классов. Показано превосходство над аналогами по точности в рамках задачи классификации кардиоинтервалограмм.
4. Разработаны программы, реализующие предложенные алгоритмы.

Список публикаций

1. Senyukova O., Lukin A., Vetrov D. Automated Atlas-Based Segmentation of NISSL-Stained Mouse Brain Sections Using Supervised Learning // Programming and Computer Software. 2011. Vol. 37, no. 5. Pp. 245–251.
2. Сенюкова О., Лукин А., Ветров Д. Автоматическая сегментация срезов мозга мыши, окрашенных по NISLL, основанная на обучении с учителем по разметке из атласа (перевод с английского языка статьи из Programming and Computer Software) // Программирование. 2011. Vol. 5. Pp. 39–48.
3. Сенюкова О., Галанин В. Выделение областей интереса на основе классификации изолиний // Программные продукты и системы. 2012. Vol. 1. Pp. 54–58.

4. Сеньюкова О. Визуализация данных об экспрессии генов на поверхности коры и гиппокампа мозга мыши // Труды конференции ГрафиКон. 2009. Pp. 372–375.
5. Senyukova O., Lukin A., Vetrov D. Automated Atlas-Based Segmentation of Nissl-Stained Mouse Brain Slices // Proceedings of GraphiCon. 2010. Pp. 92–96.
6. Senyukova O., Galanine V., Krylov A. et al. Diffuse Axonal Injury Lesion Segmentation Using Contouring Algorithm // Proceedings of GraphiCon. 2011. Pp. 84–87.
7. Gavrishchaka V., Senyukova O. Robust algorithmic detection of the developed cardiac pathologies and emerging or transient abnormalities from short periods of RR data // Proceedings of CMLS. Vol. 1371. 2011. Pp. 215–224.
8. Senyukova O., Gavrishchaka V. Ensemble decomposition learning for optimal utilization of implicitly encoded knowledge in biomedical applications // Proceedings of CIB. 2011. Pp. 69–73.
9. Gavrishchaka V., Senyukova O., Koepe M., Kryuchkova A. Multi-objective physiological indicators based on complementary complexity measures: application to early diagnostics and prediction of acute events // Proceedings of ICCCI. 2011. Pp. 95–106.
10. Senyukova O., Gavrishchaka V. Diagnostics of complex and rare abnormalities using ensemble decomposition learning // Proceedings of ICCCI. 2011. Pp. 19–26.
11. Сеньюкова О., Кошушин А., Ветров Д., Анохин К. Поверхностная визуализация данных об экспрессии генов в коре головного мозга и гиппокампе мозга мыши // Труды конференции Нейроинформатика. Vol. 1. 2010. Pp. 249–257.
12. Gavrishchaka V., Senyukova O., Ulanova O., Monin A. Physiological meta-indicators for professional sports applications: express diagnostics, overtraining detection, and quantification of individual zones of optimal functioning // Proceedings of VII Int. scientific and practical conference for memory of P.Roudik. 2011. Pp. 5–7.
13. Петряйкин А., Ахадов Т., Сеньюкова О., Крылов А. Алгоритм идентификации очагов диффузно-аксонального повреждения у больных с черепно-мозговой травмой // Труды международного симпозиума “Нейроимиджинг и магнитоэнцефалография: фундаментальные исследования и клиническая практика”. 2012.