

А.А. Дмитриев, С.В. Кучерявский

Применение методов многомерного анализа данных для диагностики ранних стадий болезни Альцгеймера по томограммам мозга

A.A. Dmitriev, S.V. Kucheryavsky

Multivariate Data Analysis to Diagnose the Alzheimer's Disease using Computerized Tomographic Imaging

В настоящей работе предложен способ анализа изображений мозга пациентов с различной степенью болезни Альцгеймера, получаемых с помощью магнитно-резонансной томографии. Метод основан на проекционном подходе, который позволяет находить скрытые (латентные) переменные в исходном пространстве признаков, непосредственно коррелирующих с искомой величиной. Предварительное изучение изменений в структуре мозга, вызванных развитием заболевания, позволило выделить признаки, наиболее релевантные для диагностики болезни. Полученные признаки использовались для построения классификационных моделей, определяющих наличие или отсутствие болезни у пациента.

Ключевые слова: болезнь Альцгеймера, обработка и анализ изображений, формальное моделирование, анализ многомерных данных.

Введение. Методы обработки и анализа изображений широко применяются для решения различных задач в современной науке и технике. Они дают возможность не только количественно описать визуальные характеристики объектов исследования, но и связать их с другими свойствами (в том числе с физическими, химическими, функциональными характеристиками и т.п.), так или иначе зависящими от структурных особенностей этих объектов.

Основное достоинство использования анализа изображений заключается в неинвазивности и наглядности метода. Кроме того, в последнее время созданы аппаратные решения, позволяющие получать уникальные изображения на различных масштабных уровнях, в том числе и внутренней структуры объектов. Такие методики существенно расширяют круг задач, решаемых с помощью аппарата обработки и анализа изображений.

Один из наиболее ярких примеров развития и применения визуальных средств анализа на практике – медицинская диагностика. Использование электромагнитного излучения в различных диапазонах, а также явлений магнитного резонанса (МР), особенностей распространения звуковых волн в различных частях тела человека позволило создать широкий спектр устройств, отображающих внутреннюю структуру организма. Все это в свою оче-

The present work proposes a new method of analysis of MRI tomograms typical to patients with early stages of Alzheimer's disease. The method is based on soft modeling approach, which allows us to define latent variables in initial space of features correlated with unknown quantity.

The prior study of changes in a brain structure during disease development allowed us to determine features relevant to disease diagnostics. Obtained results are used for constructing classification models which can define whether the patient is sick or not.

Key words: Alzheimer's disease, image treatment and analysis, formal modeling, multivariate data analysis.

редь дало толчок к развитию математических средств обработки медицинских изображений, а также методов их анализа для более точной и быстрой диагностики различных болезней [1].

Задача ранней диагностики заболеваний – одна из главных проблем современной медицины. Своевременная постановка диагноза позволяет обеспечить быстрое лечение больного, принять меры для предупреждения болезни, а также замедлить ее развитие, что особенно важно в случае, если недуг имеет неизлечимую форму. К такому роду заболеваний относится болезнь Альцгеймера – одна из самых тяжелых форм деменции у человека, которая характеризуется потерей психической активности, памяти и функциональной деятельности. Анализ изображений мозга, полученных при помощи магнитно-резонансной томографии, позволяет с определенной точностью диагностировать болезнь на поздних стадиях развития, однако надежного способа обнаружения признаков начального этапа заболевания по МР-томограммам не существует. В связи с этим актуальной задачей является определение количественных признаков изображений головного мозга, которые бы коррелировали с наличием у пациентов ранней стадии болезни Альцгеймера и могли быть использованы для разработки методов автоматической диагностики заболевания.

Для установления зависимости между характеристиками изображения и стадией развития болезни в настоящей работе было решено использовать принципы формального моделирования, когда взаимосвязь между признаками и изучаемым явлением описывается математическим соотношением, но может не носить физического смысла [2]. Наиболее предпочтителен среди таких способов метод проекций на латентные структуры. Основной алгоритм его работы заключается в представлении группы образцов с различной степенью болезни на плоскости в пространстве скрытых переменных. Ориентация плоскости выбирается так, чтобы максимально отличить группы объектов с различной степенью заболевания. Для этого при построении модели учитывается взаимосвязь между количественными признаками образцов и характерной для них степенью болезни [3].

Цель настоящей работы – определение релевантных признаков на изображениях головного мозга, пригодных для разработки методов автоматической диагностики болезни Альцгеймера. Процесс решения задачи можно представить в виде нескольких последовательных этапов. На первом из них происходило улучшение качества томограмм мозга. Из преобразованных изображений выделялись релевантные признаки болезни, которые использовались в дальнейшем для классификации и предсказания образцов. Для проведения экспериментов использовались две группы по 25 образцов. Первая состояла из пациентов с мягкой формой деменции, во вторую были включены здоровые объекты. Для каждого образца было получено 175 томограмм головного мозга.

Применение метода Фурье-дескрипторов при описании контура на изображении. В современной медицинской практике для диагностики заболевания на ранней стадии медицинский персонал наиболее часто использует оценку изменения размера желудочковой системы на МРТ-изображении мозга. Обычно данный морфологический объект на томограмме представляет собой область черного цвета, которая может быть выделена из изображения методом сегментации по порогу яркости. Однако использование напрямую предложенного преобразования не позволяет качественно описать объект, так как неизбежно вовлечет в рассмотрение другие сегменты изображения, которые в дальнейшем могут служить причиной ошибок при классификации образцов с различной степенью заболевания [4]. Для устранения данного недостатка было предложено выделить из изображения только контур желудочковой системы и описать его методом Фурье-дескрипторов. Данная методика заключается в представлении точек границы объекта на томограмме, в виде ряда комплексных чисел, вещественная и мнимая часть которых определяется положением пикселей в двумерной системе координат, наложенной на изображение [5]. Полученный ряд подвергается дискретному преобразованию Фурье, которое для ко-

нечной последовательности длиной K примет следующий вид.

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-\frac{i2\pi uk}{K}}$$

Комплексные коэффициенты $a(u)$ называются Фурье-дескрипторами границы, основным их свойством является то, что они трансформируются при видоизменении контура, вызванного развитием заболевания.

Производя операцию по расчету коэффициентов для всей серии изображений, формировалась матрица X , строками которой являлись образцы с различной степенью заболевания. Каждому образцу соответствовал его вектор морфологических признаков. Полученному массиву сопоставлялся вектор - столбец степеней болезни y . В структуре данных у пациенты с мягкой степенью болезни отмечались единицей, тогда как здоровые образцы были маркированы -1 . Оба набора данных участвовали в классификации методом проекций на латентные структуры. Результат моделирования можно представить в виде двух графиков. График счетов показывает расположение образцов на плоскости, образованной в новом пространстве двумя скрытыми переменными, характеризующими степень заболевания. Для полученной модели на рисунке 1 образцы, принадлежащие разным степеням болезни, поделились четко на две непересекающиеся группы, что говорит о хорошей классифицирующей способности.

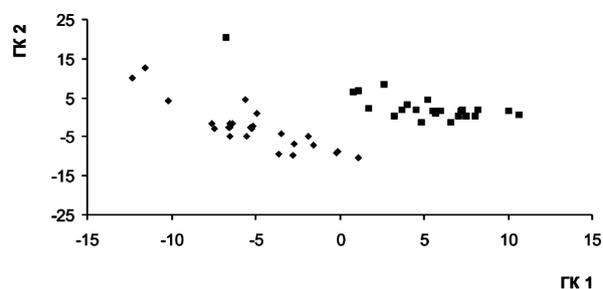


Рис. 1. Разделение на две группы образцов на графике счетов

График предсказанных против измеренных значений позволяет судить о предсказывающей способности полученной модели для тестового набора. В идеальном случае тестовые и калибровочные образцы с обнаруженным заболеванием должны располагаться вдоль оси ординат около 1, а здоровые пациенты – в области -1 .

Для построенной модели на рисунке 2 показано пересечение двух групп образцов, что позволяет судить об ошибках в предсказании тестового набора. Таким образом, описание контура методом Фурье-дескрипторов позволяет достаточно хорошо классифицировать пациентов по степени болезни Альцгеймера, но характерным недостатком построенной модели является низкая предсказательная способность.

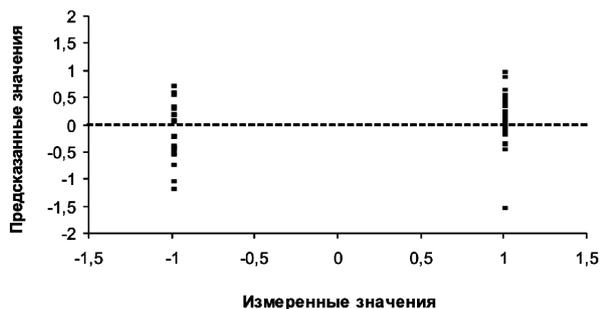


Рис. 2. График предсказанных против измеренных значений

Анализ морфологии головного мозга методами фрактальной геометрии. Ошибки в предсказании тестового набора для образцов с признаками Фурье-дескрипторов способствовали исследованию второго морфологического признака болезни, а именно изменения количества вещества головного мозга. Для количественного описания данной характеристики томограмму предполагалось разделить на цветовые составляющие, характерные для каждого из видов мозгового вещества, и воспользоваться аппаратом фрактальной геометрии. В данном способе для любого изображения были рассчитаны коэффициенты фрактальной размерности, которые позволяли судить о степени изрезанности вещества на плоском изображении. Для вычисления коэффициентов использовался клеточный метод расчета [6]. Алгоритм его работы заключается в последовательном наложении на изображение вещества сетки с ячейками определенной площади и подсчете той их части, в которую попали сегменты изображения мозга. При этом, изменяя масштаб сетки в сторону увеличения размера ячейки, можно получить линейную зависимость логарифма числа сегментов изображения от логарифма размера ячейки, тангенс угла наклона которой и укажет на коэффициент фрактальной размерности [7].

Вектор признаков для классификации формировался из найденных коэффициентов размерности для всей серии срезов, на которых представлено изображение мозга. График счетов для полученной ПЛС модели, изображенный на рисунке 3, показал образование двух линейно разделимых групп образцов, соответствующих различным степеням болезни.

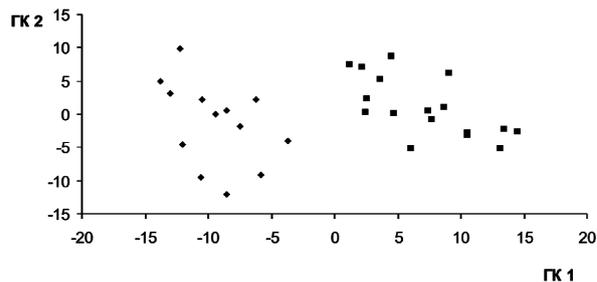


Рис. 3. График счетов с признаками фрактальной размерности

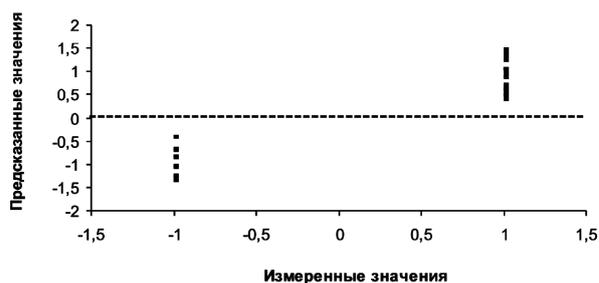


Рис. 4. Зависимость предсказанных от измеренных значений для полученной модели

График предсказанных от измеренных значений, представленный на рисунке 4, также показал правильное распределение образцов по группам. В отличие от предыдущего эксперимента облака точек вдоль оси Y не перекрывались, а группировались около соответствующего им класса заболевания. Полученный результат позволил сделать вывод о хорошей предсказательной способности построенной модели.

Заключение. По результатам работы выделены морфологические признаки головного мозга у пациентов с ранней стадией болезни Альцгеймера, пригодные для разработки методов автоматической диагностики этого заболевания. Представлены методы количественного описания предложенных характеристик. На примере анализа расширения желудочковой системы мозга и изменения количества мозгового вещества проведена классификация образцов с различной степенью деменции методом проекций на латентные структуры. Произведен анализ способности полученных моделей предсказывать наличие болезни у тестовых образцов.

Библиографический список

1. Хорнак Д. П. Основы МРТ. – М., 2007.
2. Эсбенсен К. Анализ многомерных данных. – Черноголовка, 2005.
3. Martens H., Naes T. Multivariate Calibration. John Wiley & Sons, 2001.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М., 2006.
5. Прэтт У. Цифровая обработка изображений. – М., 1982.
6. Fisher Y. Fractal Image Compression. – California, 1994.
7. Федер Е. Фракталы. – М., 1991.