

Распознавание патологий по критерию сопряженности с подпространствами эталонных изображений

В.И. Коршиков

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
vlad73453@gmail.com

В.А. Фурсов

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
fursov@ssau.ru

Аннотация—Решается задача распознавания патологий на медицинских изображениях головного мозга, полученных с помощью магнитно-резонансной томографии (МРТ). Предлагается технология распознавания с разбиением векторных пространств классов на подпространства. Получены зависимости вероятности распознавания от числа подпространств. Приведены результаты экспериментов.

Ключевые слова—МРТ-изображения, распознавание патологий, критерий сопряженности

I. ВВЕДЕНИЕ

Одним из важнейших факторов эффективного лечения раковых опухолей головного мозга является своевременное и корректное их выявление. В настоящее время магнитно-резонансная томография (МРТ) нашла широкое применение при диагностировании опухолей головного мозга, так как она позволяет радиологам получить большое количество информации об опухоли [1]. Однако, с распространением данного инструмента, особую актуальность приобретает вопрос интерпретации данных для определения конкретного типа опухоли. Наиболее распространенными типами опухолей являются: глиальные опухоли (глиомы), менингиомы и опухоли гипофиза [2].

Проблема интерпретации данных связана, с одной стороны, с большой вариативностью данных, с другой стороны, с малыми различиями между изображениями различных патологий. Попытки формирования системы признаков из наборов геометрических фигур, к сожалению, слишком трудоемки и часто не приводят к хорошему результату. Мы предлагаем технологию распознавания с разбиением векторных пространств классов на подпространства, основанную на определении близости к этим подпространствам по критерию сопряженности.

II. ПОСТАНОВКА ЗАДАЧИ

Предполагается, что осуществлена предварительная обработка изображений, в частности, выполнены центрирование и масштабирование (приведение к одинаковым размерам). Каждому изображению (путем развертки по строкам) поставлен в соответствие вектор размерности $N \times 1$, а из множества эталонных векторов исключены почти линейно-зависимые векторы.

Множество N -векторов, сформированных из эталонных изображений каждого класса (распознаваемой патологии), образует векторное

подпространство, заданное в виде $N \times m$ -матрицы $X = [x_1, x_2, \dots, x_m]$ (для определенности полагаем, что число эталонных векторов в каждом классе равно m).

Показатель сопряженности $R(x_k)$ $N \times 1$ -вектора x_k с векторным пространством, заданным в виде матрицы X определяется как

$$R(x_k) = \frac{x_k^T X [X^T X]^{-1} X^T x_k}{x_k^T x_k}. \quad (1)$$

Критерий сопряженности (1) является показателем близости вектора x_k , представляющего распознаваемое изображение, к пространству эталонных векторов класса. Удобство состоит в том, что в матрицу могут быть включены все эталонные векторы класса, при этом показатели близости $R(x_k)$ являются скалярными характеристиками. В работе [3] рассматривался метод повышения точности распознавания за счет итерационного разбиения классов на два подкласса для достижения искомого количества четных подклассов.

Мы ставим задачу разработки основанного на применении показателей сопряженности метода разбиения пространства, образованного векторами каждого класса, на любое (в т. ч. нечетное) число подпространств.

III. ОПИСАНИЕ МЕТОДА КЛАСТЕРИЗАЦИИ

В настоящей технологии вначале пошагово определяются центры подклассов по критерию минимальной сопряженности (максимальной удаленности) с пространством, образованным векторами-центрами, выбранными не предшествующих этапах. Затем к найденным центрам подклассов по критерию максимальной сопряженности (близости) поэтапно присоединяются остальные эталонные векторы.

На первом шаге определения центров подклассов осуществляется поиск в обучающем наборе каждого класса пары векторов с наименьшим показателем критерия

$$r(x_i, x_j) = (x_i^T x_j) / (\|x_i\| \times \|x_j\|). \quad (2)$$

Эти два вектора признаются центрами первых двух подклассов. Следующий центр (третьего) подкласса определяется как вектор, имеющий минимальную сопряженность, (в смысле формулы (1)) с подпространством, образованным двумя найденными векторами. При определении третьего центра подкласса матрица X имеет размерность $N \times 2$, для

определения четвертого центра – $N \times 3$ и т.д. до достижения требуемого количества центральных векторов подклассов.

Теперь к ним последовательно присоединяются оставшиеся свободными в этом классе эталонные векторы. Для присоединения к каждому центру по одному (первому) вектору используется максимальное значение критерия (2). В результате на основе каждого центра образуется подпространство двух векторов. На следующем шаге к этим подпространствам, образованным первыми двумя векторами, по критерию максимальной сопряженности присоединяется по одному (третьему) вектору из свободных векторов класса. При этом матрицы X подпространств будут иметь размерность $N \times 2$. На третьем шаге присоединения четвертого вектора – $N \times 3$ и т.д. до завершения распределения всех свободных векторов. При этом возможна ситуация, когда некоторое число эталонных векторов останется нераспределенным.

IV. ОПИСАНИЕ И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Целью эксперимента является определение типа опухоли на МРТ-изображениях, полученных в аксиальной проекции. В работе использовались изображения из набора данных Brian Tumor Classification (MRI), содержащего снимки глиальных опухолей, менингиомы и опухоли гипофиза (рис. 1).

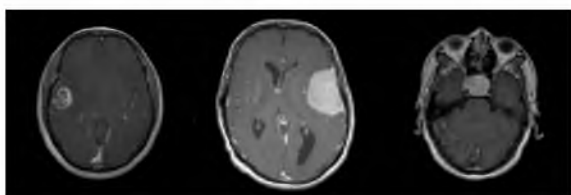


Рис. 1. Примеры изображений глиомы, менингиомы и опухоли гипофиза из набора Brian Tumor Classification (MRI).

Для подготовки экспериментальных данных к тестированию, были сформированы наборы, содержащие 300, 375 и 450 изображений. Каждый набор разбит на выборки: 80% обучающая выборка и 20% – тестовая. Поскольку размер изображений в наборе различается, все изображения приведены к размеру 256x256 пикселей, проведено выравнивание рабочей области полученных изображения по центру, затем они были построчно развернуты в векторы.

В ходе эксперимента было проведено сравнение точности классификации тестовой выборки на наборах данных разного размера. Для каждого набора выявлено количество подпространств, при котором достигнута наибольшая точность измерений (рис. 2).

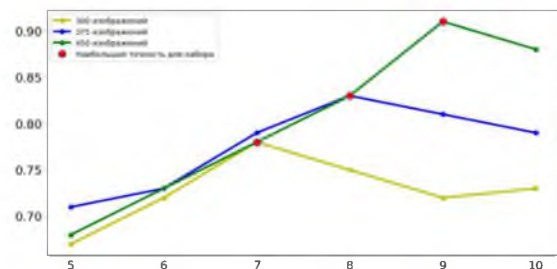


Рис. 2. График зависимости точности классификации от числа подклассов для наборов разного размера.

Как видно из представленного графика, с увеличением числа изображений в наборе, количество подпространств, при котором достигается наиболее высокая точность, также увеличивается. Как видно из результатов, представленных в таблице I, при увеличении числа объектов оценки, вероятность их корректного распознавания увеличивается.

Таблица I. СРАВНЕНИЕ ЛУЧШЕЙ ТОЧНОСТИ КЛАССИФИКАЦИИ ДЛЯ КАЖДОГО НАБОРА

Объем	Вероятность распознавания			
	глиома	менингиома	гипофиз	средняя
300	0,8	0,75	0,8	0,78
375	0,72	0,84	0,92	0,83
450	0,87	0,9	0,97	0,91

В случае использования в качестве меры близости показателя сопряженности, разделимость обеспечивается за счет включения в подклассы векторов признаков, имеющих малые значения показателя сопряженности со своим классом [4].

Следует заметить, что значительное увеличение числа подклассов может снижать точность классификации, так как наиболее удаленные центры подклассов могут присоединять к себе эталоны из смежных классов. Для обеспечения высокой точности необходима настройка, обеспечивающая компромисс между количеством используемых эталонных изображений и числом подклассов в каждом классе.

V. ЗАКЛЮЧЕНИЕ

В настоящей статье впервые применен критерий сопряженности в задаче распознавания медицинских изображений. Результаты экспериментов подтвердили, что, при достаточном количестве эталонных изображений метод, основанный на вычислении показателей сопряженности при разбиении классов на подклассы способен обеспечить хорошую точность классификации на относительно небольшом наборе обучающих изображений. К преимуществам данного метода следует отнести устойчивость к изменениям ракурса и простота настройки технологии. При появлении новых эталонных изображений они просто включаются в матрицы соответствующих классов.

ЛИТЕРАТУРА

- [1] Huang, M. Brain tumor segmentation based on local independent projection-based classification / M. Huang, W. Yang, Y. Wu [et al.] // IEEE transactions on biomedical engineering. – 2014. – Vol. 10(61) – P. 2633-2645.
- [2] Парфенов, В.Е. Сборник лекций по актуальным вопросам нейрохирургии / В.Е. Парфенов, Д.В. Свистов, Б.В. Мартынов – СПб.: Элби-СПб, 2008. – 456 с.
- [3] Жердев, Д.А. Распознавание объектов на радиолокационных изображениях с использованием показателей сопряженности и опорных подпространств / Д.А. Жердев, Н.И. Казанский, В.А. Фурсов // Компьютерная оптика. – 2015. – Т. 39, №. 2. – С. 255-264.
- [4] Фурсов, В.А. Тематическая классификация гиперспектральных изображений по показателю сопряженности / В.А. Фурсов, С.А. Бибилов, О.А. Байда // Компьютерная оптика. – 2014. – Т. 38, №. 1. – С. 154-158.