

ИСПОЛЬЗОВАНИЕ ТЕКСТУРНЫХ ХАРАКТЕРИСТИК ДЛЯ ДЕТЕКТИРОВАНИЯ ПАТОЛОГИЙ ДИССЕМНИРОВАННОГО ТУБЕРКУЛЁЗА ЛЁГКИХ

Костин К.А.¹, Лассе Бах²

Аксёнов С.В.¹

¹Национальный исследовательский Томский политехнический университет

²Орхусский университет, Дания

kak@tpu.ru

По данным Всемирной организации здравоохранения туберкулёз, наряду с другими заболеваниями дыхательных путей, входит в рейтинг десяти самых распространённых причин смерти по всему миру по состоянию на 2015 год [1]. Наиболее высокая заболеваемость данным заболеванием наблюдается в развивающихся странах [2], в России туберкулёз является лидирующим по заболеваемости среди инфекционных заболеваний: в среднем 94,6 человек на 10 000 населения [3].

Одной из основных проблем борьбы с туберкулёзом является сложность его ранней диагностики – современные методы позволяют обнаружить туберкулёз и определить его клиническую форму зачастую только на стадиях, когда болезнь уже значительно прогрессирует. На сегодняшний день, одними из наиболее точных и информативных методов диагностики являются лучевые методы, в частности, компьютерная томография (КТ) [4]. Однако эффективность их применения в значительной степени зависит от квалификации и опыта медицинского специалиста [5]. В этой связи, большую актуальность имеет задача разработки компьютерной системы медицинской диагностики для автоматического детектирования и классификации различных форм туберкулёза лёгких.

В настоящее время уже существует множество исследований, посвящённых выявлению основных характеристик различных форм туберкулёза лёгких по данным КТ [6; 7], а также автоматическому обнаружению признаков туберкулёза по текстурным характеристикам [8] и морфологическим характеристикам патологий лёгких [9]. Однако задача классификации формы туберкулёза по выделяемым характеристикам остаётся нерешённой.

Целью данного исследования, как этапа разработки системы компьютерной диагностики, является оценка эффективности применения методов классификации, основанных на текстурных признаках изображения КТ, для детектирования патологий диссеминированной формы туберкулёза лёгких.

Для проведения исследования были использованы персонифицированные данные КТ пациентов, с установленной формой диссеминированного туберкулёза. Примеры данных показаны на рис. 1.

В работе [10], описывающей, разработанный авторами предварительный этап обработки данных КТ, приводится описание алгоритма сегментации, позволяющего выделить участки изображения, содержащие исключительно ткани лёгких.

Полученные результаты в данной работе позволяют выделить образцы текстуры патологий ткани лёгких, свойственных диссеминированному туберкулёзу, а также образцы здоровой ткани для проведения эксперимента по классификации. Примеры образцов приведены на рис. 1.

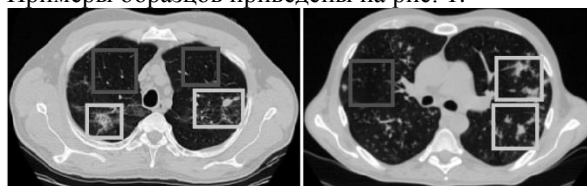


Рис. 1. Примеры данных КТ. Рамками показаны примеры текстурных образцов, выделенных для эксперимента: светлые – патологии диссеминированного туберкулёза, тёмные – здоровые ткани.

В качестве метода классификации в данной работе был применён классификатор на основе деревьев решений, обученный на текстурных признаках, полученных с помощью алгоритма *Gray Level Co-Occurrence Matrix (GLCM)* [11]. Алгоритм *GLCM* основывается на анализе фрагмента изображения (текстуры) с помощью окна определённого размера. Интенсивности цвета пикселей, попавших в окно, преобразуются в целочисленные значения яркости в соответствии с выбранным числом уровней квантования. По полученным значениям строятся матрицы смежности уровней яркости пикселей, используя которые рассчитываются такие статистические характеристики текстуры как контраст, энергия, гомогенность и т.д. Определение основных параметров алгоритма для анализа снимков КТ является одной из задач проводимого исследования.

Используя полученные характеристики был разработан классификатор на основе деревьев решений. Классификатор был протестирован на данных компьютерной томографии 200 пациентов, среди которых присутствовали снимки, содержащие патологии диссеминированного туберкулёза и здоровые ткани лёгких. При этом тестирование было произведено с использованием характеристик *GLCM*, полученных для различных значений параметров размеров окна анализа и числа уровней квантования яркости. Результаты тестирования, демонстрирующие точность классификации, отображены на трёхмерном графике – рис. 2.

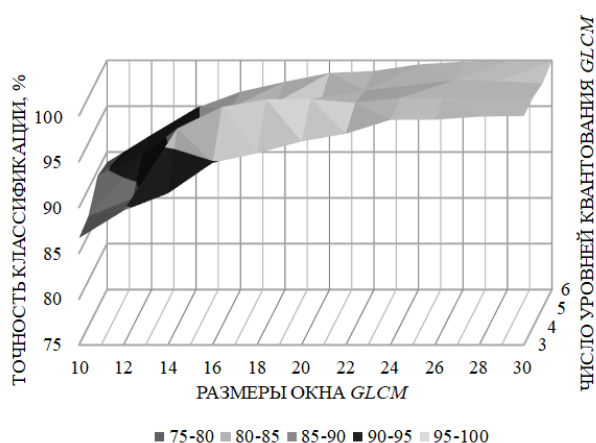


Рис. 2. График функции точности классификации от размеров окна анализа и числа уровней квантования яркости алгоритма *GLCM*

На рисунке видно, что при увеличении размеров окна анализа алгоритма, увеличивается точность классификации. Точность растёт, пока значение размера окна не достигает 20-22 пикселей. После этого порога, её значение изменяется незначительно. Это говорит о том, что именно при таких установленных значениях данного параметра, классификатору удастся с высокой степенью вероятности различить патологии диссеминированной формы туберкулёза лёгких. Также, на представленном графике можно заметить то, что изменение числа уровней квантования алгоритма оказывает незначительное влияние на точность при значениях размера окна, больших 20 пикселей. Т.к. высокие значения данного параметра влекут за собой извлечение большего количества текстурных признаков и усложнения алгоритма классификации, то это даёт основания выбрать значения данного параметра, равным не более четырёх.

Эти результаты показывают состоятельность выбранной модели классификации диссеминированной формы туберкулёза лёгких по текстурным характеристикам, полученным на данных КТ. При этом тестирование позволяет оценить параметры выбранного алгоритма извлечения характеристик, которые позволяют получить наиболее точные результаты классификации.

По результатам проделанной работы можно заключить, что статистические текстурные характеристики, полученные с помощью метода *GLCM* являются достаточно репрезентативными для достижения точности бинарной классификации диссеминированной формы туберкулёза с вероятностью 95 %. В дальнейшем планируется исследовать возможности разработанного классификатора для различения диссеминированной формы туберкулёза и других заболеваний лёгких, семиотика которых на данных КТ является значительно схожей: это такие заболевания как инфильтративный туберкулёз, саркоидоз и пневмония.

Данные для исследования предоставлены Национальной академией наук Белоруссии. Исследование выполнено при поддержке гранта РФФИ №16-47-700289.

Список использованных источников

- 10 ведущих причин смерти в мире. Центр СМИ // Всемирная организация здравоохранения. 2017. URL: http://www.who.int/mediacentre/factsheets/fs_310/ru/ (дата обращения: 28.09.2017).
- Доклад о глобальной борьбе с туберкулезом. Туберкулез // Всемирная организация здравоохранения. 2016. URL: http://www.who.int/tb/publications/global_report/ru/ (дата обращения: 28.09.2017).
- Tuberculosis surveillance and monitoring in Europe 2016. The Russian Federation // World Health Organization. 2016. URL: http://www.euro.who.int/_data/assets/pdf_file/0018/310806/TB-surveillance-report-2016-The-Russian-Federation.pdf (дата обращения: 28.09.2017).
- Отраслевые и экономические показатели противотуберкулёзной работы в 2012 – 2013 г.г. Аналитический обзор основных показателей и статистические материалы / Т.Ч. Касаева [и др.]: под ред. С.А. Стерликова – М.: РИО ЦНИИОИЗ, 2014. – 72 с.
- Berlin L. Accuracy of Diagnostic Procedures: Has It Improved Over the Past Five Decades? / L. Berlin // *AJR*. – 2007. – Vol. 188. – P. 1173–1178.
- Cardoso C.A. Disseminated tuberculosis: Challenges in diagnosis / C.A. Cardoso, T.F. Garcia, P.R. Cachado // *Int. J. Case Rep. Imag.* – 2016 – Vol. 7, Issue 4 – Pp. 228-231.
- Feng F. Radiological characterization of disseminated tuberculosis in patients with AIDS / F. Feng, G. Xia, Y. Shi, Z. Zhang // *Radiology of Infectious Diseases* – 2016 – Vol. 3, Issue 1 – Pp. 1-8.
- Devan L. Automated Texture-Based Characterization of Fibrosis and Carcinoma Using Low-Dose Lung CT Images / L. Devan, R. Santosham, R. Hariharan // *Int. J. of Imag. Sys. and Tech.* – 2014 – 10.1002/ima.22077.
- Ramya R. Automatic tuberculosis screening using canny Edge detection method / R. Ramya, P.S. Babu // *ICECS* – 2015 – Pp. 282-285.
- Костин К.А., Аксёнов С.В. Детектирование патологий туберкулёза лёгких по данным компьютерной томографии // Научная сессия ТУСУР – 2017: Материалы международной научно-технической конференции студентов, аспирантов и молодых ученых: в восьми частях, Томск, 10-12 Мая, 2017. - Томск: ТУСУР, 2017 - часть 4, С. 212-215.
- Haralick R.M. Texture features for image classification / R.M. Haralick, R. Bosley // *Proceedings "3th Earth Res. Tech. Satel. Symp."* - 1973 - Vol. SP-351 – Pp. 1929-1969.