

Подгорнова Ю.А.

*Муромский институт (филиал) федерального государственного образовательного учреждения высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»
602264, г. Муром, Владимирская обл., ул. Орловская, 23*

Анализ текстурных признаков при разработке алгоритма распознавания заболеваний молочной железы по маммограммам

Маммография является одним из самых распространенных способов диагностики заболеваний молочной железы[1]. Однако не всегда на маммографических снимках можно точно определить диагноз и назначить правильное лечение пациенту.

Для повышения точности постановки корректного диагноза используются компьютерные системы[2] для распознавания заболеваний молочной железы на маммографических снимках.

Для того, чтобы проанализировать, какие признаки наиболее эффективны для распознавания, необходимо протестировать их на большом количестве маммографических снимков.

Целью данной работы является анализ признаков, используемых при распознавании кисты молочной железы на маммограммах.

Рассмотрим признаки, которые могут быть использованы для распознавания новообразований на изображении:

- 1) Гистограммные признаки первого порядка[3];
- 2) Статистические признаки Харалика второго порядка[4].

Методы отбора признаков обычно делят на 4 категории: фильтры, обёртки, встроенные и гибридные.

1) При таком подходе оценивается важность признаков только на основе свойственных им характеристик, без привлечения алгоритмов обучения. Эти методы работают быстрее и требуют меньше вычислительных ресурсов по сравнению с методами «обертками». Если для моделирования статистической корреляции между признаками не хватает объема данных, тогда фильтры могут давать результаты хуже, чем обёртки. В отличие от обёрток, такие методы менее склонны к переобучению. Они широко используются для работы с данными высокой размерности, когда методы обертки требуют слишком больших вычислительных мощностей.

2) При таком подходе оценивается эффективность подмножества признаков, учитывая финальный результат примененного алгоритма обучения (например, каков прирост точности при решении задачи классификации). В этой комбинации поисковой стратегии и моделирования может использоваться любой алгоритм обучения.

3) К группе встроенных методов относятся алгоритмы, которые одновременно обучают модель и отбирают признаки. Обычно это реализуют с помощью l_1 -регуляризатора (sparsity regularizer) или условия, которое ограничивает некоторые признаки.

- 4) Гибридные методы объединяют в себе симбиоз первых трёх категорий.

Было принято решение отбирать признаки средствами корреляционного анализа. Главными задачами корреляционного анализа являются выявление зависимости или сходства между какими-либо сигналами. Найденные в ходе анализа признаки, которые наиболее не коррелируют между собой, будут использованы для классификации изображения.

На основе полученных значений для классификации изображений были отобраны такие признаки как: среднее значение, эксцесс для классификации маммограмм с доброкачественными новообразованиями и сумма энтропий, разница дисперсий для классификации маммограмм со злокачественными новообразованиями.

Литература

- 1 Садыков С.С., Буланова Ю.А., Захарова Е.А. Алгоритм выделения области кисты на малоконтрастных маммограммах // Информационные технологии. 2013. № 8. С. 53-57.
- 2 Буланова Ю.А. Экспертно-аналитическая система обработки и анализа маммограмм // Прикаспийский журнал: управление и высокие технологии. 2014. № 1 (25). С. 092-102.

3 Чабан Л.Н. Методы и алгоритмы распознавания образов в автоматизированном дешифрировании данных дистанционного зондирования: учебное пособие. – М.: миигаик, 2016, – 94 с.

4 Haralick R. M. Textural Features for Image Classification / R. M. Haralick, K. Shanmugan, I. Dinstein. //IEEE Transactions on systems, man and cybernetics, volume SMC-3. – IEEE, 1979 – №6. – P. 610-621. – ISSN 0018-9472