

Воронцов С.А.

*Научный руководитель: к. т. н., доц. каф. ИС 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*

Критерии распознавания доброкачественных заболеваний на маммограммах

Заболевания молочных желез представляют особую опасность для женщины [1]. Маммографический скрининг – это единственный способ выявить заболевание молочной железы на ранних стадиях, когда заболевание поддается успешному лечению [2].

Маммограмма [3] – это полутоновое изображение, на котором в краниокаудальной проекции отображаются фон, область грудной мышцы, область молочной железы. Благодаря предварительной подготовке малоконтрастных маммограмм и их дальнейшей сегментации, на снимке можно выделить подозрительные области. Задача классификации найденной патологии является одной из основных задач области компьютерного зрения [4]. Решением этой задачи является подтверждение гипотезы о принадлежности выделенной области к опухоли, а затем определение класса патологии (доброкачественная и злокачественная опухоль). Поскольку маммограмма является текстурным изображением, то любой маммографический снимок может быть охарактеризован вектором текстурных признаков.

В работе были исследованы следующие классы признаков: признаки первого порядка, признаки матрицы смежности уровня серого или Gray-Level Co-Occurrence Matrix (GLCM), признаки матрицы зон размерности серого, признаки матрицы распространения серого, признаки матрицы разности соседних пикселей, признаки матрицы зависимостей серого.

Матрица смежности уровня серого или Gray-Level Co-Occurrence Matrix (GLCM) [5] – описывает характеристики изображения второго порядка, путем подсчета встречающихся комбинаций двух пикселей под определенным углом.

В основе матрицы зон размерности уровня серого или Gray Level Size Zone Matrix (GLSZ) [5] лежит подсчет количества пикселей с одинаковой яркостью в любом направлении. На основе этих данных строится матрица, где ось Y – интенсивность яркости, а ось X – максимальный размер зоны, на которой лежат пиксели с одинаковой яркостью.

Матрица распространения уровней серого или Gray Level Run Length Matrix (GLRLM) [5] является разновидностью матрицы GLSZ, только лишь с тем условием, что подсчет количества одинаковых яркостей изображений идет по одному из направлений – 0°, 45°, 90° или 135°.

Матрица разности соседних пикселей серого или Neighboring Gray Tone Difference Matrix (NGTDM) [5] высчитывается путем разности яркости текущего пикселя и средней яркости пикселей внутри окна некоторого радиуса. Такая матрица в результате будет иметь следующие столбцы: i – градация яркости; n_i – количество яркости i на текстуре; p_i – вероятность появления пикселя i на текстуре; s_i – сумма разности по модулю яркости i и суммы его соседей в некотором промежутке n .

Матрица зависимости уровней серого или Gray Level Dependence Matrix (GLDM) [5] высчитывается путем подсчета одинаковых пикселей в определенном радиусе. Результирующая матрица будет иметь ось X – количество соседей, и ось Y – уровни яркости. Элемент матрицы – количество одинаковых соседей X с центром Y в исходном изображении.

Исследования изображений обучающей выборки, сформированной по маммограммам из базы MIAS[3] показали, что в вектор признаков, характеризующих опухоль должны входить перечисленные ниже признаки:

- по матрице GLCM выбираются: контраст, корреляция, разность среднего, разность энтропии, разность дисперсии, однородность, инвертированная разница момента, нормализованная разница момента, информационная мера корреляции, инвертированная дисперсия, энтропия, максимальный корреляционный коэффициент, сумма энтропии;

- по матрице GLDM - неравномерность зависимостей, нормализованная неравномерность зависимостей, дисперсия зависимостей, выраженность высоких зависимостей;

- по матрице GLRM - выраженность малых отрезков со слабой яркостью, выраженность малых отрезков с высокой яркостью, выраженность малых отрезков, дисперсия отрезков, процент занимаемого отрезком места на изображении, энтропия отрезка, выраженность больших отрезков со слабой яркостью, нормализованная неравномерность отрезков;

- по матрице NGTDM – сложность.

Признаки матрицы GLSZ были полностью отброшены как нерелевантные.

Литература

1. Подгорнова Ю.А., Садыков С.С. Исследование алгоритмов сегментации для выделения микрокальцинов на маммографических снимках // Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений. Распознавание - 2019. сборник материалов XV Международной научно-технической конференции. 2019. С. 140-142.

2. Подгорнова Ю.А., Жизняков А.Л., Садыков С.С. Повышение контраста маммограмм, содержащих области рака молочной железы, на фоне жировой инволюции // Прикаспийский журнал: управление и высокие технологии. 2020. № 2 (50). С. 136-147.

3. Podgornova Y.A., Sadykov S.S. Detection of malignant breast tumors on the background of fibrocystic breast disease // CEUR Workshop Proceedings. IPERS-ITNT 2018 - Proceedings of the International Conference on Information Technology and Nanotechnology - Session: Image Processing and Earth Remote Sensing. 2018. С. 177-183.

4. Тымчук А.И. Классификация текстур на основе матриц смежности и нейронной сети // Перспективы науки. 2020. № 7 (130). С. 54-58.

5. Sakai A., Onishi Yu., et al. A method for the automated classification of benign and malignant masses on digital breast tomosynthesis images using machine learning and radiomic features // Radiological Physics and Technology. 2020. no. 13, pp. 27–36. DOI: <https://doi.org/10.1007/s12194-019-00543-5>