



DOI: <https://doi.org/10.15688/mpcm.jvolsu.2019.3.5>

УДК 004.89  
ББК 55.6

Дата поступления статьи: 01.04.2019  
Дата принятия статьи: 24.06.2019

## УПРАВЛЯЕМЫЙ КЛАССИФИКАТОР В ДИАГНОСТИКЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ ПО ДАННЫМ МИКРОВОЛНОВОЙ РАДИОТЕРМОМЕТРИИ<sup>1</sup>

**Татьяна Владимировна Замечник**

Кандидат медицинских наук, доцент кафедры патофизиологии, клинической патофизиологии,  
Волгоградский государственный медицинский университет  
tvzamechnic.61@mail.ru  
пл. Павших Борцов, 1, 400131 г. Волгоград, Российская Федерация

**Александр Георгиевич Лосев**

Доктор физико-математических наук, профессор кафедры математического анализа и теории функций,  
Волгоградский государственный университет  
alexander.losev@volsu.ru, matf@volsu.ru  
<https://orcid.org/0000-0002-1072-8375>  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Александр Юрьевич Петренко**

Студент кафедры фундаментальной информатики и оптимального управления,  
Волгоградский государственный университет  
fiou@volsu.ru  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Аннотация.** Работа выполнена в рамках направления, занимающегося вопросами разработки консультативной экспертной системы диагностики рака молочной железы по данным микроволновой радиотермометрии. А именно, обсуждаются вопросы построения многомерных термометрических признаков диагностики и оценивается эффективность их применения. Также в статье

обсуждается проблема разработки управляемого классификатора на базе двумерных термометрических признаков. Доказана эффективность предлагаемого подхода при определении степени выраженности температурных аномалий, и, как следствие, эффективность метода в определении наличия рака молочных желез у обследуемых пациенток.

**Ключевые слова:** микроволновая радиотермометрия, интеллектуальный анализ данных, высокоинформативные признаки, рак молочной железы, консультативные интеллектуальные системы.

### Введение

В России, начиная с 1980 г., рак молочной железы (РМЖ) стал ведущей онкологической патологией у женщин. Показатели заболеваемости РМЖ за последние 20 лет возросли на 40 процентов и имеют дальнейшую тенденцию к росту. Современные данные литературы свидетельствуют об отсутствии единой точки зрения на причины возникновения доброкачественных и злокачественных заболеваний молочных желез. Основная роль отводится дисгормоналам. Имеет значение также и наследственность, и хронический стресс, и возраст старше 40 лет, и особенности питания, и даже принадлежность к женскому полу. Сложность оценки этих факторов диктуют необходимость строгих правил и оптимизации при проведении комплексного обследования. Необходимо учитывать, что именно нарушение нейрогуморальной регуляции ведет к активации пролиферативных процессов в тканях молочных желез. Поэтому использование методов функциональной визуализации необходимы в комплексном исследовании молочной железы.

Одним из наиболее перспективных для повышения эффективности ранней дифференциальной диагностики видится метод микроволновой радиотермометрии (РТМ). Он основан на измерении собственного электромагнитного излучения тканей человека в микроволновом диапазоне длин волн и инфракрасном диапазоне, что позволяет неинвазивно выявлять температурные аномалии на глубине нескольких сантиметров [3].

Метод микроволновой радиотермометрии признан эффективным и безопасным способом обследования молочных желез женщин любого возраста. Последнее имеет крайне важное значение, так как в последние годы прослеживается тенденция роста доброкачественных заболеваний молочных желез, которые диагностируются у каждой 4-й женщины в возрасте до 30 лет. У пациенток старше 40 лет различные патологические состояния молочных желез выявляются в 60 процентах случаях. Эти заболевания молочных желез объединяются общим термином мастопатия. Самой распространенной формой мастопатии является фиброзно-кистозная болезнь (ФКБ), которая также весьма неоднородна в рентгенологических и гистологических проявлениях. Ее формы значительно отличаются между собой по анатомическим признакам, клиническому проявлению и, главное, той опасности малигнизации, которая вынуждает рассматривать мастопатию как предраковое заболевание. Основным методом объективной оценки состояния молочных желез - рентгеномаммография, имеет высокие ложноотрицательные результаты (по данным разных авторов от 12 до 24 процентов) в выявлении опухолей при хорошо развитой, плотной железистой паренхиме и на фоне выраженных инволютивных изменений, когда на маммограмах хорошо распознаются образования размером более 1 см.

При комплексном подходе с использованием микроволновой радиотермометрии диагностика существенно улучшается. В течение последних двух десятилетий данный ме-

тод получил распространение в различных областях медицины (маммология, онкология, флебология, хирургия и др.). Одновременно возник и ряд проблем. Одна из них заключается в разработке методов, алгоритмов и программных средств обработки, качественного и количественного анализа термометрической и другой медицинской информации о пациенте.

В ходе медицинских обследований и анализа данных микроволновой радиотермометрии специалистами был выявлен ряд качественных признаков рака молочной железы (см., например, [5; 8]). В том числе у пациенток с патологией отмечается наличие повышенной величины термоасимметрии между одноименными точками молочных желез; повышенный разброс температур между отдельными точками в пораженной молочной железе; повышенная разница температур сосков; повышенная температура соска в пораженной молочной железе по сравнению со средней температурой молочной железы с учетом возрастных изменений температуры; соотношение кожной и глубинной температур и некоторые другие. Появились и первые работы, посвященные анализу областей изменения функций, описывающих наличие аномалий поведения температурных полей внутри органов человека [5]. Также появились и первые исследования, посвященные применению методов интеллектуального анализа данных и разработки алгоритмов классификации в данной области [1].

Заметим, что в настоящее время при проведении медико-биологических исследований методы интеллектуального анализа данных чаще всего применяются для решения задач классификации и прогноза. Однако традиционные подходы в применении ряда весьма популярных методов машинного обучения в медицинских экспертных системах, например искусственных нейронных сетей, генетических алгоритмов, эволюционного программирования и некоторых других подходов, имеют ограничения в силу фундаментальных свойств и специфики предметной области. А именно, использование указанных выше способов вызывает крайнее затруднение при попытке объяснения и обоснования врачу-диагносту полученного результата. В настоящее время подавляющее большинство специалистов отмечает те общие свойства, которыми должна обладать современная экспертная диагностическая система. Во-первых, экспертные системы должны учитывать особенности лечебнодиагностических технологий и обрабатываемых медицинских данных. Во-вторых, должны иметь возможность динамического развития с учетом накопленного опыта, знаний и расширения экспертной базы. В-третьих, решение, выдаваемое системой, не должно быть окончательным, а должно носить рекомендательный характер и давать возможность специалисту критически оценивать решение системы и принимать свое окончательное решение.

Уже на первом этапе применения микроволновой радиотермометрии выяснилось, что используя только значения температур в различных точках молочной железы определить наличие заболевания практически невозможно. Значительно более информативным оказался анализ не значений температур, а некоторых их отношений. В частности, это показатели, характеризующие изменение температуры в некоторых основных направлениях (от поверхности молочной железы внутрь, перпендикулярно поверхности железы, особенно в районе соска, также от соска в радиальном направлении к границе молочной железы и тела). В их числе можно указать величину термоасимметрии между «зеркально» расположенными точками правой и левой молочных желез, соотношение кожной и внутренних температур («внутренний градиент»), скорость изменения внутреннего градиента в некоторых «характерных» направлениях и ряд других характеристик (см. [5]). Наборы функций, описывающие данные эффекты, с достаточно высокой эффективно-

стью определяют наличие аномалий поведения температурных полей внутри молочных желез. Одним из наиболее существенных моментов подобного подхода к построению признакового пространства является нахождение и описание характерных для некоторого класса пациенток подобластей множеств изменения функций, описывающих указанные выше эффекты. Чаще всего данные подобласти называют «информативными». Естественно, при таком подходе, подобные подобласти являются множествами на числовой прямой. Однако больший эффект дает применение многомерных термометрических признаков (см. [2]). Их построение основано на построении признакового пространства с помощью вектор-функций. Соответственно, «информативные» подобласти уже являются множествами евклидова пространства соответствующей размерности.

Данная статья посвящена развитию методов применения многомерных термометрических признаков в диагностике рака молочных желез.

### 1. Многомерные признаки

Применяемая в настоящее время методика диагностики рака молочной железы на основе данных микроволновой радиотермометрии основана на оценке функционального состояния тканей путем измерения внутренней температуры на глубине до 5 см. и температуры кожи. Обследование пациентки начинается с измерения температур в опорных точках, расположенных, первая — в центре грудной клетки сразу под и между молочными железами, вторая — непосредственно под мечевидным отростком. Далее измерения температур проводятся в 10 точках на каждой железе, и в аксиллярной области, после чего производится визуализация полученных данных, результаты которой выводятся на монитор (см. рис. 1).

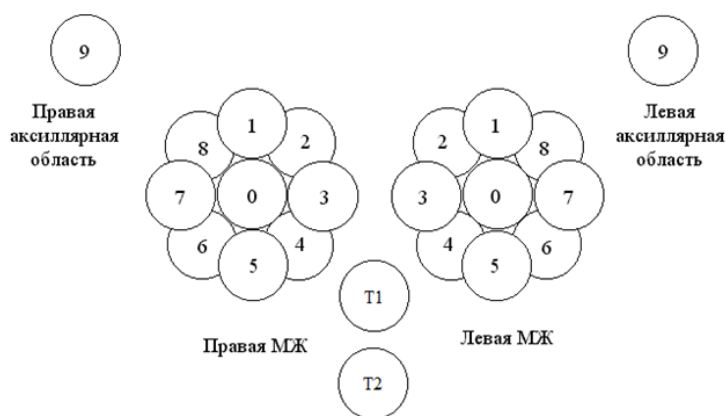


Рис. 1. Схема обследования молочной железы

Перейдем к формальному описанию задачи поиска информативных признаков (см. [2; 4; 6]). Введем следующие обозначения.

Пусть  $t_i^j, i = 0, \dots, 9$  — температуры  $i$ -ой точки правой молочной железы  $j$ -ой пациентки в диапазоне РТМ;

$t_i^j, i = 10, \dots, 19$  — температура  $[i - 10]$  точки правой молочной железы  $j$ -ой пациентки в диапазоне ИК;

$t_i^j, i = 20, \dots, 29$  — температура  $[i - 20]$  точки левой молочной железы  $j$ -ой пациентки в диапазоне РТМ;

$t_i^j, i = 30, \dots, 39$  — температура  $[i - 30]$  точки левой молочной железы  $j$ -ой пациентки в диапазоне ИК;

$t_{40}^j = T1, t_{41}^j = T2$  — глубинные температуры опорных точек  $j$ -ой пациентки;

$t_{42}^j = T1, t_{43}^j = T2$  — кожные температуры опорных точек  $j$ -ой пациентки.

Тогда обучающую выборку можно представить в виде матрицы

$$S = \begin{pmatrix} t_0^1 & \dots & t_{43}^1 \\ \vdots & \ddots & \vdots \\ t_0^k & \dots & t_{43}^k \\ t_0^{k+1} & \dots & t_{43}^{k+1} \\ \vdots & \ddots & \vdots \\ t_0^n & \dots & t_{43}^n \end{pmatrix}$$

где  $t_i^j$  при  $j = 1, \dots, k$  — температуры молочных желез пациенток класса «Не рак»,  $t_i^j$  при  $j = k + 1, \dots, n$  — температуры молочных желез пациенток класса «Рак».

Как указывалось выше, в ходе медицинских обследований и анализа данных микроволновой радиотермометрии был выявлен ряд качественных признаков рака молочной железы. В том числе у пациенток с патологией отмечается наличие повышенной величины термоасимметрии между одноименными точками молочных желез; повышенный разброс температур между отдельными точками в пораженной молочной железе; повышенная разница температур сосков; повышенная температура соска в пораженной молочной железе по сравнению со средней температурой молочной железы с учетом возрастных изменений температуры; соотношение кожной и глубинной температур и некоторые другие. Процессе исследований был сформирован ряд гипотез о поведении температурных полей внутри молочных желез. Во-первых, это гипотеза «зеркальной» симметрии температурных полей правой и левой молочных желез здоровых пациенток. Она исходит из симметричного внутреннего строения парных органов человека. Для анализа данных на предмет выполнения данной гипотезы возможно использование функционалов самого разного вида. В данной работе применялись функционалы следующих видов:

$$Pr6 = \max\{|t_0 - t_{20}|, |t_1 - t_{21}|, \dots, |t_8 - t_{21}|\}$$

и

$$Pr5 = \max\{|t_{10} - t_{30}|, |t_{11} - t_{31}|, \dots, |t_{18} - t_{31}|\}.$$

Во-вторых, предполагается использование гипотезы о малом разбросе температур в здоровой молочной железе. Для анализа данных на предмет выполнения данной гипотезы также возможно использование функционалов самого разного вида. В данной работе применялись функционалы следующих видов:

$$Pr1 = \max\{|t_{10} - t_{11}|, |t_{10} - t_{12}|, \dots, |t_{10} - t_{18}|, |t_{30} - t_{31}|, |t_{30} - t_{32}|, \dots, |t_{30} - t_{38}|\}$$

и

$$Pr2 = \max\{|t_0 - t_1|, |t_0 - t_2|, \dots, |t_0 - t_8|, |t_{20} - t_{21}|, |t_{20} - t_{22}|, \dots, |t_{20} - t_{28}|\}.$$

Отдельный интерес вызывает исследование, так называемого, «внутреннего градиента». Для исследования последнего можно использовать функционал следующего вида:

$$Pr3 = \max\{|t_0 - t_{10}|, |t_1 - t_{11}|, \dots, |t_8 - t_{18}|, |t_{20} - t_{30}|, |t_{21} - t_{31}|, \dots, |t_{28} - t_{38}|\}.$$

Высокую эффективность показал признак, описывающий изменение внутреннего градиента в радиальном направлении ([5]). Он может быть описан функционалом вида

$$Pr4 = \max\{ |(t_0 - t_{10}) - \frac{1}{3} \sum_{i=1}^3 t_i - t_{i+10}|, |(t_0 - t_{10}) - \frac{1}{3} \sum_{i=2}^4 t_i - t_{i+10}|, \dots, \\ |(t_0 - t_{10}) - \frac{(t_8 - t_{18}) + (t_1 - t_{11}) + (t_2 - t_{12})}{3}|, |(t_{20} - t_{30}) - \frac{1}{3} \sum_{i=21}^{23} t_i - t_{i+10}|, \\ |(t_{20} - t_{30}) - \frac{1}{3} \sum_{i=22}^{24} t_i - t_{i+10}|, \dots, |(t_{20} - t_{30}) - \frac{(t_{28} - t_{38}) + (t_{21} - t_{31}) + (t_{22} - t_{32})}{3}| \}.$$

В дальнейшем предлагается рассматривать данные функции в комплексе. Пусть  $f^q$  —  $q$ -я исследуемая вектор-функция, рассматриваемая на множестве векторов  $\Omega^q = \{ (t_0^j, t_1^j, \dots, t_{43}^j) \}$ . Здесь  $j = 1, 2, \dots, n$ . Обозначим  $f_j^q$  значение  $q$ -й моделирующей вектор-функции на температурных данных  $j$ -го пациента.

Назовем высокоинформативным признаком тройку  $(f^q, V, X)$ , где  $X$  — «информативная» область множества значений вектор-функции  $f^q$ ,  $V = I(f^q, X)$  — информативность признака. Опишем подробно данные понятия.

Под информативностью понимается числовая функция, которая показывает, насколько хорошо указанный признак отделяет класс «больных» от класса «здоровых». При построении признаков использовались статистическая, эвристическая, энтропийная информативность, их различные комбинации, а также некоторые другие функции (см. [5]). Приведем некоторые примеры.

Статистическая информативность, вычисляемая по формуле

$$ST(f^q, x) = -\ln\left(\frac{C_k^h C_{n-k}^s}{C_{k+(n-k)}^{h+s}}\right)$$

где  $n$  — количество векторов с температурными данными пациенток в обучающей выборке,  $k$  — количество векторов с температурными данными пациенток класса «не рак» в обучающей выборке,  $h$  — количество молочных желез класса «не рак», для которых  $f^q \in X$ , а  $s$  — количество молочных желез класса «рак», для которых  $f^q \in X$ .

Эвристическая информативность, вычисляемая по формуле

$$HR(f^q, x) = \frac{\max(\frac{k}{h}, \frac{n-k}{s})}{\min(\frac{k}{h}, \frac{n-k}{s})}$$

Энтропийная информативность, вычисляемая по формуле

$$EN(f^q, x) = H\left(\frac{k}{k+(n-k)}, \frac{n-k}{k+(n-k)}\right) - \\ - \frac{k+(n-k)-h-s}{k+(n-k)} \cdot H\left(\frac{k-h}{k+(n-k)-h-s}, \frac{(n-k)-s}{k+(n-k)-h-s}\right)$$

где  $H(q_0, q_1) = -q_0 \log_2(q_0) - q_1 \log_2(q_1)$  — математическое ожидание количества информации.

Далее, пусть

$$I_s = \sup_{X_\alpha} I(f^q, X),$$

где  $X_\alpha$  — все возможные подобласти соответствующего многомерного пространства, на которых указанная информативность имеет смысл. «Информативной» областью множества значений вектор-функции  $f^q$  будем называть подмножество многомерного евклидова пространства, обеспечивающее достижение (или приближение с заданной точностью) величины  $I_s$ . Для  $j$ -ых молочных желез признак  $(f_j^q, V, X)$  считается выполненным, если  $f_j^q \in X$ .

Оценка эффективности алгоритмов классификации, в частности и диагностических методов в медицине, традиционно выполняется при помощи следующих показателей.

Чувствительность (sensitivity), также известная, как полнота (recall), вычисляется по формуле:

$$Sens = \frac{TP}{TP + FN},$$

где  $TP$  — количество молочных желез класса «рак», которые были классифицированы как «рак»,  $FN$  — количество молочных желез класса «рак», которые не были классифицированы как «рак».

Специфичность (specificity), вычисляется по формуле:

$$Spec = \frac{TN}{TN + FP},$$

где  $TN$  — количество молочных желез класса «не рак», которые не были классифицированы как «рак»,  $FP$  — количество молочных желез класса «не рак», которые были классифицированы как «рак».

Точность (precision), вычисляется по формуле:

$$Prec = \frac{TP}{TP + FP}.$$

Достоверность (accuracy), вычисляется по формуле:

$$Acc = \frac{TP + TN}{TP + FP + FN + TN}.$$

В качестве метрики эффективности классификации обычно используются различные варианты  $F$ -мер, в частности — среднее гармоническое между точностью ( $Prec$ ) и чувствительностью ( $Sens$ ):

$$F_1 = 2 \frac{Prec \cdot Sens}{Prec + Sens}.$$

Учитывая, что в медицинской диагностике наиболее значимыми являются параметры  $Sens$  и  $Spec$ , было решено в качестве метрики эффективности классификации использовать среднее геометрическое между  $Sens$  и  $Spec$ .

В одномерном случае, когда  $X$  является подмножеством  $R$ , а моделирующая функция  $f^q$  является скалярной функцией, технология получения и применения соответствующих признаков детально разработана (см. [5]), реализован соответствующий модуль, который встроен в консультативную информационную систему. С геометрической точки зрения задача заключается в том, чтобы найти отрезок действительной оси, на котором



расположены значения функции  $f^q$  подавляющего большинства «здоровых» пациентов из обучающей выборки и небольшого количества «больных». Для размерности 2 и выше задача становится нетривиальной. Некоторые из возможных вариантов нахождения «информативных» подобластей описаны в [2] и [4].

К настоящему времени моделирующие функции одномерных признаков достаточно хорошо описаны и исследованы (см., например, [5]). При формировании признакового пространства моделирующие функции разбиваются на несколько классов, описывающих их отклик на наличие тех или иных аномалий температурных полей. В качестве моделирующих функций для построения двумерных признаков целесообразно взять небольшое количество функций, каждая из которых адекватно описывает тот или иной класс, моделирующий соответствующий медицинский признак. В данной статье будут применяться функции  $Pr1, Pr2, \dots, Pr6$ , описанные выше.

Выбирая пары из полученных шести функций, получаем 15 двумерных признаков. Обучающая выборка состояла из 73 «здоровых» и 56 «больных» пациентов. Эффективность проверялась на тестовой выборке из 36 «здоровых» и 27 «больных» пациентов. Для каждого двумерного признака вычислялась чувствительность  $Sens$  (доля верно диагностированных больных) и специфичность  $Spec$  (доля верно диагностированных здоровых).

Для дальнейшего исследования наиболее интересным способом построения «информативных» областей, нам показался алгоритм «сужающихся эллипсов» (см. [2]). Он состоит из следующих шагов:

- *Шаг 1:* Строится эллипс минимальной площади, внутри которого лежат точки всех, соответствующие «здоровым» пациенткам из обучающей выборки, с помощью известного алгоритма Хачияна [7].
- *Шаг 2:* Для каждой точки считается сумма расстояний от нее до всех остальных.
- *Шаг 3:* Исключается точка с наибольшей суммой.
- *Шаг 4:* Перестраивается эллипс по оставшимся точкам.
- *Шаг 5:* Сужается построенный эллипс. Оставляя неизменным центр эллипса и направление его главных осей, варьируются его полуоси. В результате получается двупараметрическое семейство эллипсов, среди которых выбирается эллипс с наибольшей эффективностью. Его сохраняют.
- *Шаг 6:* Если не достигнуто минимальное число точек, по которому можно построить эллипс, из оставшихся точек исключается точка с наибольшей суммой расстояний и переходят к Шагу 4.

Из сохраненных на шаге 5 эллипсов выбирается эллипс с наибольшей эффективностью.

В [2] приведены примеры работы алгоритма. Лучшие признаки дают эффективность порядка 80–85 процентов.

## 2. Медицинское обоснование двумерных термометрических признаков

При проведении исследований было обследовано 192 пациента. Все пациенты женского пола в возрасте 24–95 лет, имели предварительное клиническое, рентгенологическое, ультразвуковое, некоторые цитологическое исследование и определенный диагноз. Всех пациентов разделили на обучающую и тестовую группы в соотношении 2 и 1.



Каждая группа включала пациентов с диагнозами «Без патологии молочных желез», «ФКБ», «Подозрение на рак на фоне ФКБ», «РМЖ». В каждой группе было проведено РТМ обследование «слепым методом» (без знания предварительного диагноза).

В результате применения двумерных признаков были получены следующие результаты. В зону эллипса попали преимущественно точки, соответствующие здоровым пациенткам. В зону прямоугольника, соответствующего последовательному применению одномерных признаков, попали точки, соответствующие и больным и здоровым пациенткам. При этом внутри прямоугольника, но вне эллипса, попали в основном точки, соответствующие больным пациенткам. В двух выборках (тестовой и обучающей) хотя бы по одному признаку из 15-ти 82 больных пациента оказались в зоне здоровых (внутри эллипса), 78 больных вне эллипса, но внутри прямоугольника, 65 здоровых вне эллипса, но внутри прямоугольника и 55 здоровых в зоне больных.

Далее мы выделили для исследования следующие группы пациенток:

1. Больные пациенты, попавшие в группу здоровых (соответствующие точки которых находятся внутри эллипса);
2. Больные пациентки, точки которых находятся вне эллипса, но внутри прямоугольника;
3. Здоровые пациентки, попавшие в группу больных (соответствующие точки вне эллипса и квадрата);
4. Здоровые пациенты, соответствующие точки которых не принадлежат эллипсу, но принадлежат прямоугольнику;
5. «Истинно больные» (больные, соответствующие точки которых лежат вне прямоугольника);
6. «Истинно здоровые» (здоровые, соответствующие точки которых находятся внутри эллипса).

В каждой из этих групп выделили пациентов, которые попали в указанные области более чем по 5-ти признакам.

1. Больные пациенты, попавшие в группу здоровых (соответствующие точки которых находятся внутри эллипса) – 29 человек;
2. Больные пациентки, точки которых находятся вне эллипса, но внутри прямоугольника – 17 человек;
3. Здоровые пациентки, попавшие в группу больных (соответствующие точки вне эллипса и квадрата) – 14 человек;
4. Здоровые пациенты, соответствующие точки которых не принадлежат эллипсу, но принадлежат прямоугольнику – 4 человека;
5. «Истинно больные» (больные, соответствующие точки которых лежат вне прямоугольника) – 51 человек;
6. «Истинно здоровые» (здоровые, соответствующие точки которых находятся внутри эллипса) – 37 человек.

Далее мы рассмотрели частоту признаков в каждой группе. В группе «истинно больных» чаще других подтверждались характерные для больных признаки, которые характеризуют изменения глубокой температуры и внутреннего градиента температур.

В то же время больные, точки которых находятся в пограничной области (вне эллипса, но внутри прямоугольника) имели низкую частоту указанных выше признаков, но высокую частоту признаков, базирующихся на изменении ИК температур. Пациенты этой группы — женщины в возрасте от 50 до 91 года, находятся в менопаузе, средний диаметр молочной железы 25,3 см. Можно предположить, что инволюция собственно железы и увеличение доли жировой ткани в области молочной железы привели к тому, что характерные для больных признаки, связанные с внутренним градиентом отошли на второй план, что и снизило качество диагностики. Больные, точки которых находятся внутри эллипса при высокой частоте характерных для здоровых признаков, также имели низкую частоту признаков, связанных с внутренними температурами. У «истинно здоровых» подтверждались все признаки, характерные для здоровых с высокой частотой. Тогда как у здоровых, точки которых находятся в пограничной области (вне эллипса, но в прямоугольнике) с высокой частотой проявлялись признаки, характеризующие изменение ИК температуры. Возраст пациентов данной группы от 23 до 50 лет, средний размер железы 17,5 см, с низким и нормальным индексом массы. Повышение ИК температуры и изменения признаков с ней связанных в данной группе пациентов, объясняется активным функционированием молочной железы и малым ее размером, а также отсутствием значительного слоя подкожной жировой ткани в области железы. В данном случае ИК температура увеличивается не только из-за влияния температуры активно функционирующей железы, но и грудных мышц.

Данные о повышении глубокой температуры в области сосков в сочетании с термоасимметрией глубоких и поверхностных температур обеих МЖ используемые при диагностике в тестовой выборке дали результаты с высокой чувствительностью. Термоасимметрия и повышение глубокой температуры в области соска выделялись нами и ранее как характеристические признаки. Однако эффективность диагностики при парном использовании этих признаков значительно возросла. Ранее мы выделяли значимость для диагностики и такого признака как изменение внутреннего градиента в области соска МЖ. При расширении функции признака до изменения внутреннего градиента температур во всех точках обеих молочных железах и второй производной этих изменений мы получили новые значимые признаки, которые при использовании в паре с другими вышеназванными признаками увеличивает чувствительность диагностики до 92,5 процентов. Таким образом, диагностика при использовании комбинации функций, отражающих аномальные ассиметрии температур в различных направлениях в обеих МЖ может зафиксировать множество незначительных изменений, которые в своей совокупной оценке укажут на развитие РМЖ с высокой степенью достоверности.

### 3. Управляемый классификатор

Предлагаемый выше подход к построению признакового пространства позволяет решить проблему качества алгоритмов классификации. Как указывалось выше, чаще всего в машинном обучении в качестве показателей информативности используют чувствительность (sensitivity), также известную как полноту (recall), специфичность (specificity), точность (precision), достоверность (accuracy). При этом в медицине наиболее значимыми являются первые две характеристики. Для оценки качества моделей и сравнения различных алгоритмов классификации практически во всех случаях используются различные метрики. При этом выбор метрики оказывает достаточно существенное влияние на соотношение чувствительности и специфичности. Повышая одну величину

ну, мы автоматически уменьшаем другую. Чаще всего в качестве метрик используются различные варианты  $F$ -мер (в общем случае  $F_\beta$ -мер), в частности — среднее гармоническое между точностью и чувствительностью, коэффициент корреляции Мэтьюса и т.д. Более того, за счет выбора параметра можно управлять соотношением чувствительности и специфичности, что крайне важно при проведении диагностических мероприятий. Выбору метрик качества классификации посвящено весьма большое количество исследований. Практически во всех исследованиях управление качеством классификации производится за счет выбора соответствующих параметров метрики. Это является вполне логичным и обоснованным, если в процессе эксплуатации не будет регулярно возникать необходимость в изменении соотношения чувствительности и специфичности. Однако, в медицине это не так. При диагностических мероприятиях, проводимых в различных группах пациентов с использованием экспертной системы, вполне может возникнуть задача изменения соотношения чувствительности и специфичности, причем желательно, чтобы это был управляемый врачом параметр. Конечно, можно было бы с помощью изменения параметра настроить несколько классификаторов, но при таком подходе в случае изменения предполагаемого диагностического решения, совершенно непонятно, из-за чего это произошло.

Основная идея данного исследования заключается в управлении качеством алгоритмов классификации в диагностике рака молочной железы за счет деформации (в первом приближении, за счет сжатия или расширения) «характерных» областей многомерных диагностических признаков. Отметим, что предлагаемый подход имеет ряд преимуществ. Во-первых, блок обоснования диагностического решения не потребует изменений. Во-вторых, если при разных соотношениях чувствительности и специфичности произошло изменение предлагаемого диагностического решения, легко выявляется, какой именно диагностический признак повлиял на это изменение и насколько значимыми являются выявленные температурные аномалии. Последнее представляет для врача-диагноста особый интерес, и, что самое главное, само по себе управление качеством алгоритмов классификации явно увеличит эффективность ранней диагностики рака молочной железы.

Учитывая то, что при диагностике проводится анализ измеренных температур, в стандартной методике диагностики по данным микроволновой радиотермометрии предлагается строить заключение по степени выраженности температурных аномалий с описанием проблемных областей. Чем выше уровень тепловых аномалий, тем выше вероятность выявления рака молочной железы, методами, определяющими структурные изменения в ткани молочной железы. Все РТМ-заключения можно разделить на шесть групп по уровню выраженности тепловых процессов.

Th0 (УВТП-0) — Тепловые изменения не выявлены.

Th1(УВТП-1) — Характерных признаков рака молочной железы не выявлено. При этом на фоне умеренно выраженных тепловых изменений, наблюдается участок понижения температуры (имеется термоасимметрия между одноименными точками правой и левой молочных желез) который не объясняется расположением молочной железы на грудной клетке. Соотношение уровней внутренней и кожной температур носит постоянный характер.

Th2 (УВТП-2) — Характерных признаков рака молочной железы не выявлено. При этом на фоне умеренно выраженных тепловых изменений, наблюдается участок повышения температуры (имеется термоасимметрия между одноименными точками правой и левой молочных желез), который не объясняется расположением молочной железы на

грудной клетке. Соотношение уровней внутренней и кожной температур носит постоянный характер.

Th3 (УВТП3) — Группа риска. Высокий уровень и внутренней и кожной температур, который не объясняется структурными особенностями молочных желез. Соотношение уровней внутренней и кожной температур носит постоянный характер.

Th4 (УВТП4) — Подозрение на рак. Возможен любой из ниже перечисленных признаков.

- Наблюдается термоасимметрия, заметно (в 1,5 и более раз) отличающаяся от остальных ТА. Соотношение уровней внутренней и кожной температур носит постоянный характер.
- На высоком уровне температур наблюдается системный наклон, то есть температуры в одной из желез выше, чем в другой. Соотношение уровней внутренней и кожной температур может носить постоянный характер.
- На совмещенной термограмме наблюдается либо сходимость, либо локальный участок, где кожная температура выше внутренней.

Th5 (УВТП5) — Термограмма характерная для заболевания (острое воспаление, рак) Возможен любой из ниже перечисленных признаков и их совокупность.

- Наблюдается термоасимметрия, заметно (в 1,5 и более раз) отличающаяся от остальных ТА. На совмещенной термограмме наблюдается либо сходимость, либо локальный участок, где кожная температура выше внутренней.
- На температурном поле внутренних температур наблюдается красное пятно в виде «колокола». На совмещенной термограмме наблюдается либо сходимость, либо локальный участок, где кожная температура выше внутренней.
- На высоком уровне температур наблюдается системный наклон, температуры в одной из желез выше, чем в другой. На совмещенной термограмме наблюдается либо сходимость, либо локальный участок, где кожная температура выше внутренней.

Далее применялся следующий алгоритм действий. Равномерно сужая первоначальные эллипсы, то есть понижая чувствительность и повышая специфичность, мы следили за изменением предполагаемого диагностического решения. У 70 процентов пациенток с Th0, которым первоначально был поставлен ошибочный диагноз «рак», он не подтвердился. При этом, естественно, новых ошибочных диагнозов не появилось. У 70 процентов пациенток с Th3 и Th4 первоначально поставленный диагноз «рак» подтверждался при повышении специфичности до 95 процентов. У пациенток, больных РМЖ, с Th1 и Th2, первоначальный диагноз подтвердился примерно в 50 процентах случаев. Таким образом можно констатировать, что применение управляемого классификатора при проверке первоначального диагноза вполне имеет право на существование.

### **ПРИМЕЧАНИЕ**

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФИ и Администрации Волгоградской области, проект № 18-47-340004 р\_а «Управление качеством классификации в диагностике рака молочной железы по многомерным данным микроволновой радиотермометрии».

## СПИСОК ЛИТЕРАТУРЫ

1. Алгоритмы принятия решений в консультативной интеллектуальной системе диагностики молочных желез / А. В. Зенович, В. А. Глазунов, А. С. Опарин, Ф. Г. Примаченко // Вестник Волгоградского государственного университета. Серия 1, Математика. Физика. — 2016. — № 6 (37). — С. 141–154. — DOI: 10.15688/jvolsu1.2016.6.13.
2. Алгоритмы формирования двумерных признаков диагностики заболеваний молочных желез методами комбинированной радиотермометрии / А. В. Зенович, Н. А. Батулин, Д. А. Медведев, А. Ю. Петренко // Математическая физика и компьютерное моделирование. — 2018. — Т. 21, № 4. — С. 44–56. — DOI: 10.15688/mpcm.jvolsu.2018.4.4.
3. Веснин, С. Г. Современная микроволновая радиотермометрия молочных желез / С. Г. Веснин, М. А. Каплан, Р. С. Авакян // Опухоли женской репродуктивной системы. — 2008. — № 3. — С. 28–33.
4. Интеллектуальный анализ многомерных термометрических данных в медицинской диагностике / А. Г. Лосев, А. В. Зенович, О. А. Бочкарев, В. В. Левшинский // Вестник Волгоградского государственного университета. Серия 1, Математика. Физика. — 2016. — № 5 (36). — С. 150–161. — DOI: 10.15688/jvolsu1.2016.5.13.
5. Лосев, А. Г. Интеллектуальный анализ данных микроволновой радиотермометрии в диагностике рака молочной железы / А. Г. Лосев, В. В. Левшинский // Математическая физика и компьютерное моделирование. — 2017. — Т. 20, № 5. — С. 49–62. — DOI: 10.15688/mpcm.jvolsu.2017.5.6.
6. Проблемы измерения и моделирования тепловых и радиационных полей в биотканях: анализ данных микроволновой термометрии / А. Г. Лосев, А. В. Хоперсков, А. С. Астахов, Х. М. Сулейманова // Вестник Волгоградского государственного университета. Серия 1, Математика. Физика. — 2015. — № 6 (31). — С. 31–71. — DOI: 10.15688/jvolsu1.2015.6.3.
7. Khachiyan, L. G. Rounding of Polytopes in the Real Number Model of Computation / L. G. Khachiyan // Mathematics of Operations Research. — 1996. — Vol. 21, № 2. — P. 307–320.
8. Modern Microwave Thermometry for Breast Cancer / S. Vesnin, A. K. Turnbull, J. M. Dixon, I. Goryanin // Journal of Molecular Imaging and Dynamics. — 2017. — Vol. 7, № 2 — DOI: 10.4172/2155-9937.1000136.

## REFERENCES

1. Zenovich A.V., Glazunov V.A., Oparin A.S., Primachenko F.G. Algorithms of Decision-Making in Intelligent Advisory System for Diagnostics of the Mammary Glands. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 1, Matematika. Fizika* [Science Journal of Volgograd State University. Mathematics. Physics], 2016, no. 6 (37), pp. 141-154. DOI: 10.15688/jvolsu1.2016.6.13.
2. Zenovich A.V., Baturin N.A., Medvedev D.A., Petrenko A.Yu. Algorithms for the Formation of Two-Dimensional Characteristic and Informative Signs of Diagnosis of Diseases of the Mammary Glands by the Methods of Combined Radio Thermometry. *Matematicheskaya fizika i kompyuternoe modelirovanie* [Mathematical Physics and Computer Simulation], 2018, vol. 21, no. 4, pp. 44-56. DOI: 10.15688/mpcm.jvolsu.2018.4.4.
3. Vesnin S.G., Kaplan M.A., Avakyan R.S. Sovremennaya mikrovolnovaya radiotermometriya molochnykh zhelez [Modern Microwave Radiometry of the Mammary Glands]. *Opukholy zhenskoy reproduktivnoy sistemy*, 2008, no. 3, pp. 28-33.
4. Losev A.G., Zenovich A.V., Bochkarev O.A., Levshinskiy V.V. Intellektualnyy analiz mnogomernykh termometricheskikh dannykh v meditsinskoj diagnostike [Multidimensional Thermometric Data Mining in Medical Diagnostics]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 1, Matematika. Fizika* [Science Journal of



Volgograd State University. Mathematics. Physics], 2016, no. 5 (36), pp. 150-161. DOI: 10.15688/jvolsu1.2016.5.13.

5. Losev A.G., Levshinskiy V.V. Intellektualnyy analiz dannykh mikrovolnovoy radiotermometrii v diagnostike raka molochnoy zhelezy [Data Mining of Microwave Radiometry Data in the Diagnosis of Breast Cancer]. *Matematicheskaya fizika i kompyuternoe modelirovanie* [Mathematical Physics and Computer Simulation], 2017, vol. 20, no. 5, pp. 49-62. DOI: 10.15688/mpcm.jvolsu.2017.5.6.

6. Losev A.G., Khoperskov A.V., Astakhov A.S., Suleymanova Kh.M. Problemy izmereniya i modelirovaniya teplovykh i radiatsionnykh poley v biotkanyakh: analiz dannykh mikrovolnovoy termometrii [Problems of Measurement and Modeling of Thermal and Radiation Fields in Biological Tissues: Analysis of Microwave Thermometry Data]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 1, Matematika. Fizika* [Science Journal of Volgograd State University. Mathematics. Physics], 2015, no. 6 (31), pp. 31-71. DOI: 10.15688/jvolsu1.2015.6.3.

7. Khachiyan L.G. Rounding of Polytopes in the Real Number Model of Computation. *Mathematics of Operations Research*, 1996, vol. 21, no. 2, pp. 307-320.

8. Vesnin S., Turnbull A.K., Dixon J.M., Goryanin I. Modern Microwave Thermometry for Breast Cancer. *Journal of Molecular Imaging and Dynamics*, 2017, vol. 7, no. 2 DOI: 10.4172/2155-9937.1000136.

## **GUIDED CLASSIFIER IN THE DIAGNOSIS OF BREAST CANCER ACCORDING TO MICROWAVE RADIOTHERMOMETRY**

**Tatyana Vladimirovna Zamechnik**

Candidate of Medical Sciences, Assistant Professor, Department of Pathophysiology,  
Clinical Pathophysiology,  
Volgograd State Medical University  
tvzamechnik.61@mail.ru  
Pl. Pavshih Borcov, 1, 400131 Volgograd, Russian Federation

**Alexander Georgievich Losev**

Doctor of Physical and Mathematical Sciences, Professor, Department of Mathematical  
Analysis and Theory of Functions,  
Volgograd State University  
alexander.losev@volsu.ru, matf@volsu.ru  
<https://orcid.org/0000-0002-1072-8375>  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Alexander Yuryevich Petrenko**

Student, Department of Fundamental Computer Science  
and the Optimal Control,  
Volgograd State University  
fiou@volsu.ru  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Abstract.** Current paper was performed in the framework of the direction dealing with the development of a consultative expert system for the diagnosis of breast cancer according to microwave radiometry data. Namely, the issues of



constructing multidimensional thermometric signs of diagnosis are discussed and the effectiveness of their use is assessed. The article also discusses the problem of developing a managed classifier based on two-dimensional thermometric features. Proved the effectiveness of the proposed approach in determining the severity of temperature anomalies, and, as a consequence, the effectiveness of the method in determining the presence of breast cancer in the examined patients.

**Key words:** microwave radiothermometry, data mining, highly informative signs, breast cancer, intelligent advisory systems.