

ПРОЕКТИРОВАНИЕ WEB-ПРИЛОЖЕНИЯ ЭКСПРЕСС-ДИАГНОСТИКИ МЕЛАНОМЫ НА ОСНОВЕ АЛГОРИТМА ШТОЛЬЦА

Коробкова Тамара Александровна

Преподаватель, ФГБОУ ВО Самарский государственный
медицинский университет Минздрава России;
аспирант, федеральное государственное автономное
образовательное учреждение высшего образования
«Самарский национальный исследовательский
университет имени академика С.П. Королева»

DESIGNING A WEB APPLICATION FOR EXPRESS DIAGNOSTICS OF MELANOMA BASED ON THE STOLZ ALGORITHM

T. Korobkova

Summary. The article discusses the process of designing a web application for express diagnosis of melanoma using computer vision algorithms and a neural classifier. Based on the dermatoscopic method for determining V. Stolz's melanoma, an image with a skin neoplasm is pre-processed: a neoplasm is highlighted in the image; parameters for Stolz's formula and a total dermatoscopic score are calculated, according to which a preliminary diagnosis is calculated. Next, the optimization problem is solved to clarify the boundaries of the areas of differentiation of neoplasms and adjust the coefficients of the Stolz formula, a balanced set for training a neural network is generated. Based on the obtained parameter values, a trained neural network (multilayer perceptron) classifies the neoplasm as benign or melanoma. The approbation was carried out on two data sets, one of which is balanced. During testing, the average accuracy of the algorithm on two independent datasets is 86 %.

Keywords: melanoma, Stolz algorithm, computer vision in medicine, multilayer perceptron, diagnosis of skin tumors, ABCD rule, melanoma diagnostic app.

Аннотация. В статье рассматривается процесс проектирования web-приложения экспресс-диагностики меланомы с помощью алгоритмов компьютерного зрения и нейронного классификатора. На основе дерматоскопического метода определения меланомы В. Штольца изображение с кожным новообразованием предварительно обрабатывается: выделяется новообразование на изображении; рассчитываются параметры для формулы Штольца и общий дерматоскопический балл, по которому вычисляется предварительный диагноз. Далее решается задача оптимизации для уточнения границ областей дифференциации новообразований и корректировка коэффициентов формулы Штольца, генерируется сбалансированный набор для обучения нейронной сети. По полученным значениям параметров обученная нейронная сеть (многослойный перцептрон) классифицирует новообразование как доброкачественное или меланому. Апробация проводилась на двух наборах данных, один из которых является сбалансированным. При проведении испытаний средняя точность алгоритма на двух независимых наборах данных составляет 86 %.

Ключевые слова: меланомы, алгоритм Штольца, компьютерное зрение в медицине, многослойный перцептрон, диагностика кожных новообразований, правило ABCD, приложение для диагностики меланомы.

Введение

Российскими учеными и разработчиками программного обеспечения создаются различные web и мобильные приложения для диагностики кожных новообразований, предлагаются различные алгоритмы для прототипов систем диагностики, например, мобильное приложение «Про родинки» (Нижний Новгород) [10], в работах [5] и [6] предложены варианты экспериментальных образцов мобильных систем скрининга пигментных новообразований кожи при поддержке Российского фонда фундаментальных исследований (РФФИ).

Для проектирования своего web-приложения экспресс-диагностики «MelanomaDiagnosisRT» в основу был положен алгоритм ABCD для диагностики меланомы, предложенный Р.Фридманом в 1985 году. Этот алгоритм оценивает пигментные образования кожи по четырем параметрам: асимметрия (A), неровные границы (B), неравномерность окраски (C) и диаметр более 6 мм (D)

[15]. Сумма произведений факторов веса и относительных единиц каждого признака можно определяет общее количество баллов (TDS — total dermoscopy score): $TDS = (A * 1.3) + (B * 0.1) + (C * 0.5) + (D * 0.5)$. Интерпретация величины общего дерматоскопического балла приведена в таблице 1.

Таблица 1.

Интерпретация значения
общего дерматоскопического балла

TDS	Интерпретация
<4,75	Доброкачественные меланоцитарные новообразования;
4,75–5,45	Подозрительные новообразования
>5,45	Новообразования с высоким подозрением на меланому;

Цель и дизайн исследования

Целью исследования является разработка web-приложения и уникального алгоритма для эффективной классификации меланомы на изображениях.

Разработка web-приложения включает в себя 2 этапа: разработка алгоритма классификации по изображению и разработка web-приложения.

Разработка алгоритма классификации по изображению включает в себя следующие этапы: выделение новообразования на изображении; расчет параметров для формулы Штольца; расчет TDS (вычисление диагноза); решение задачи оптимизации для уточнения границ областей дифференциации новообразований и корректировка коэффициентов формулы Штольца; генерация сбалансированного набора для обучения нейронной сети; классификация нейронной сетью двух наборов изображений по параметрам Штольца; постановка диагноза.

Данные для исследования

Все операции по работе с изображениями реализованы с помощью алгоритмов компьютерного зрения и библиотеки с открытым исходным кодом OpenCV [12]. Набор данных для разработки и апробации предлагаемых методов состоит из 127 фотографий (62 изображения меланомы, 65 — доброкачественные новообразования), сделанных в ГБУЗ «Самарский областной клинический онкологический диспансер» (далее — СамараОнко) и набора The Human Against Machine with 10000 training images (HAM10000) включающим в себя 10015 изображений (меланома — 1113 изображений, другие заболевания — 8902 изображения).

Выделение новообразований на изображениях

К каждому изображению из наборов применялись поочередно 12 предложенных методов, с перебором порогового значения от 10 до 180 с шагом 10. Целью проведенной работы было максимально точное выделение объекта-новообразования для дальнейшего исследования.

1. Выделение контура новообразования из исходного изображения по пороговому значению.
2. Выделение контура новообразования из исходного изображения методом Canny (при пороговых значениях $minVal=80$, $maxVal=200$).
3. Удаление фона изображения с помощью морфологических операций дилатации и эрозии, выделение контура по пороговому значению.
4. Удаление фона изображения с помощью морфологических операций дилатации и эрозии, выделение контура методом Canny;
5. Цветовая сегментация с применением морфологических операций (операция закрытия), выделение контура по пороговому значению;
6. Цветовая сегментация с применением морфологических операций (операция закрытия), выделение контура методом Canny;

7. Цветовая сегментация с применением морфологических операций (на основе предположения диапазона оттенков объекта и фона MIN-MAX), выделение контура по пороговому значению;
8. Цветовая сегментация с применением морфологических операций (на основе предположения диапазона оттенков объекта и фона MIN-MAX), выделение контура методом Canny;
9. Применение билатерального фильтра с последующим выделением контура новообразования из исходного изображения по пороговому значению;
10. Применение билатерального фильтра с последующим выделением контура новообразования из исходного изображения методом Canny;
11. Повышение резкости изображения с последующим выделением контура новообразования по пороговому значению;
12. Повышение резкости изображения с последующим выделением контура новообразования методом Canny.

Результаты работы алгоритма приведены в таблице 2.

Таблица 2.

Результаты работы алгоритма выделения новообразований на изображениях

	НАМ 10000 (ед.)	СамОнко (ед.)
Общее количество изображений в наборе	10015	127
Ошибка выделения на всем наборе	291 (2,9 %)	3 (2,36 %)
Ошибка по нозологии меланома	31 (2,78 %)	2 (3,2 %)

В ходе апробации предложенной методики выделения объектов по фото было отбраковано менее 3 % от общего числа изображений по причине несоответствия установленным требованиям. А именно, новообразование на фото было слишком маленьким и занимало менее 5 % площади, либо слишком большим и соприкасалось с границами изображения, либо располагалось не по центру фото, либо небольшое контрастное значение между поверхностью кожи и новообразованием.

Вычисление признаков алгоритма Штольца и расчет TDS

После выделения новообразования на изображении выполнено вычисление признаков и расчет TDS.

A — асимметрия. Максимальная цена критерия = 2, предполагается отсутствие симметрии по обеим осям. Экспериментальным путем установлен порог допустимой погрешности 5 %:

Симметрия относительно оси x:

$$\frac{|AB - DC|}{AB + DC} * 100\% < 5\%; \tag{1}$$

Симметрия относительно оси y:

$$\frac{|AD - BC|}{AD + BC} * 100\% < 5\%; \tag{2}$$

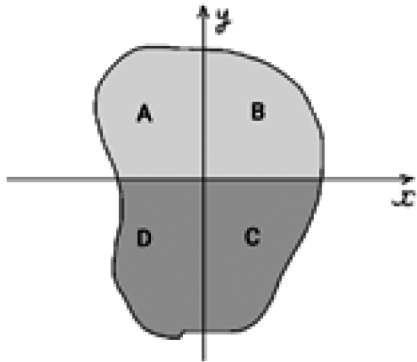


Рис. 1. Расчет симметричности новообразования

B — границы. Максимальная цена критерия = 8. Через точки максимума и минимума ($x_{max}, x_{min}, y_{max}, y_{min}$) проведем 4 прямые, параллельные осям координат. Точки пересечения этих прямых с осями будут вершинами вписанного в построенный прямоугольник эллипса.

I_d — разница длин контуров рассчитывается как отношение разности длины контура эллипса (I_{el}) и контура выделенного объекта (I_{con}) к длине контура эллипса:

$$I_d = \frac{|I_{el} - I_{con}|}{I_{el}}. \tag{3}$$

Экспериментальным путем (для наборов HAM1000 и СамараОнко) были получены максимальная (I_{dmax}) и минимальная (I_{dmin}) разница длин контуров. Полученный диапазон разделим на 8 равных участков, значение I_d будет попадать в один из таких участков, порядковый

номер которого будет соответствовать значению критерия *B*.

C — распределение цвета. Максимальная цена критерия = 6. Значения оттенка в OpenCV варьируются от 0 до 179. Условно палитра разделяется на отдельные группы оттенков, по 10 единиц на каждую группу. Если при подсчете цветов их оказалось более 6 или 6, то критерий *C* имеет максимальное значение — 6 баллов. Если цветов менее 6, то за наличие в наборе цвета добавляется по одному баллу к критерию *C*. Мультиколерность меланомы часто является одним из признаков, на которые врачи обращают внимание при дерматологическом осмотре или оценке образований на коже.

D — дерматоскопические структуры. Максимальная цена критерия = 5. Для анализа и оценки паттернов выделенного объекта необходимо найти цвета, соответствующие дерматоскопическим структурам злокачественных новообразований. С помощью масок определяются следующие цвета: черный (черные точки по периферии); коричневый, от светло-коричневого до темно-коричневого; голубой (бело-голубая пелена); серый (множественные голубые и серые точки); белый.

Таблица 3.

Результаты классификации по TDS наборов HAM10000 и СамараОнко

	HAM10000 (ед.)	HAM10000 (%)	Самара Онко (ед.)	Самара Онко (%)
Всего случаев	9724	100	124	100
Правильно рассчитанные случаи	5236	53,84	73	58,87
Подозрительные случаи	3150	32,39	29	23,39
Неверно спрогнозированные случаи	1338	13,76	16	12,9

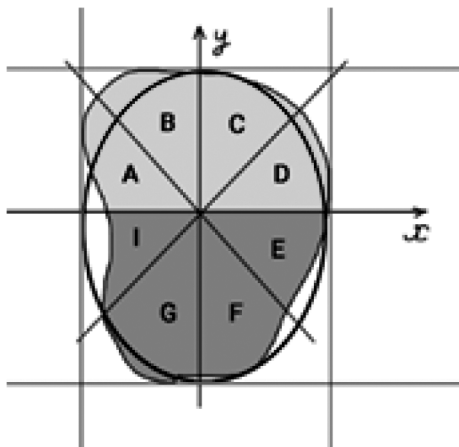


Рис. 2. Расчет количества сегментов новообразования с четким контуром

Наличие более 200 точек цвета добавляет к критерию 1 балл. Значение 200 было получено в ходе проведенных экспериментов и является минимально достаточным для утверждения наличия признака.

При подсчете метрик значения, попавшие в подозрительный диапазон, следует относить к правильным, так как они являются гипердиагностикой. Результаты классификации представлены в таблице 3.

Оптимизация границ по алгоритму Штольца и генерация набора для обучения нейронного классификатора

В целях оптимизации алгоритма было выполнено уточнение коэффициентов и границ по набору СамараОнко. Исходя из того, что сдвиг границ к центру, т.е. сокращение диапазона подозрительных новообразований приведет к увеличению числа ошибок, а расширение диапазона хотя и даст улучшение метрик, приведет к увеличению изображений, классифицируемых как подозрительные, было установлено ограничение на количество определения подозрительных новообразований (Δ_{unk}) — не более 25 %, шаг смещения границы 1 % от фактора веса, количество смещений (n_{step}) = 6.

Таким образом, при максимальном значении точности представим полученные по набору СамараОнко данные как функцию (F_1) нескольких переменных:

$$F_1 = \max_{pre} (F(q_1, q_2, k_a, k_b, k_c, k_d)) \quad (4)$$

q_1 — верхняя граница доброкачественного новообразования;

q_2 — нижняя граница злокачественного новообразования;

k_a, k_b, k_c, k_d — коэффициенты (факторы веса) критериев по Штольцу.

В результате из 531441 комбинации был выбран вариант с наибольшим значением точности и высокой чувствительностью, где $q_1 = q_2 = 4,785$; $k_a = 1,339$; $k_b = 0,101$; $k_c = 0,52$; $k_d = 0,52$.

На основе полученных границ и коэффициентов сгенерирован сбалансированный набор данных, равное количество доброкачественных новообразований и меланом, представляющий собой числовой набор параметров ABCD и бинарную отметку принадлежности к классу.

Классификация заболевания нейронной сетью по критериям Штольца

Обученной нейронной сетью были классифицированы рассчитанные параметры новообразований из на-

боров HAM10000 и СамараОнко. В качестве классификатора был взят многослойный перцептрон. Результаты приведены в таблице 4.

Таблица 4.

Результаты диагностирования нейронной сетью по критериям Штольца наборов HAM10000 и СамараОнко

	HAM 10000 (ед.)	HAM 10000 (%)	Самара-Онко (ед.)	Самара-Онко (%)
Всего случаев	9724	100	124	100
Правильно спрогнозированные случаи	8475	87,16	106	85,48
Меланома (верно/всего)	958/1082	88,54	57/60	95

Оценка качества модели

Для оценки используемого нейронного классификатора использованы наиболее популярные характеристики, такие как матрица ошибок, точность модели, полнота, чувствительность и специфичность, рассчитаны доля правильных ответов и взвешенное гармоническое среднее полноты и точности. Результаты приведены в таблице 5.

Таблица 5.

Характеристики оценки модели нейронного классификатора

	HAM10000	СамараОнко
Истинно положительные результаты (TP)	958	57
Ложноотрицательные результаты (FN)	124	3
Истинно отрицательные результаты (TN)	7517	49
Ложноположительные результаты (FP)	1125	15
Точность (Pre) (%)	45,99	79,16
Полнота (Recall) (%)	88,54	95
Чувствительность (Se) (%)	88,54	95
Специфичность (Sp) (%)	86,98	76,52
Доля правильных ответов алгоритма (accuracy)	87,16	85,48
F-мера	75,95	86,35

Разработка web-приложения

Web-приложение построено с использованием серверов и состоит из клиентской части и серверной частей.

Серверная часть, включает в себя обработку изображений, модель нейронной сети, реализована на языке Java, с использованием библиотеки для работы с компьютерным зрением OpenCV и платформы Neugorph, библиотека с открытым исходным кодом под лицензией Apache 2.0.

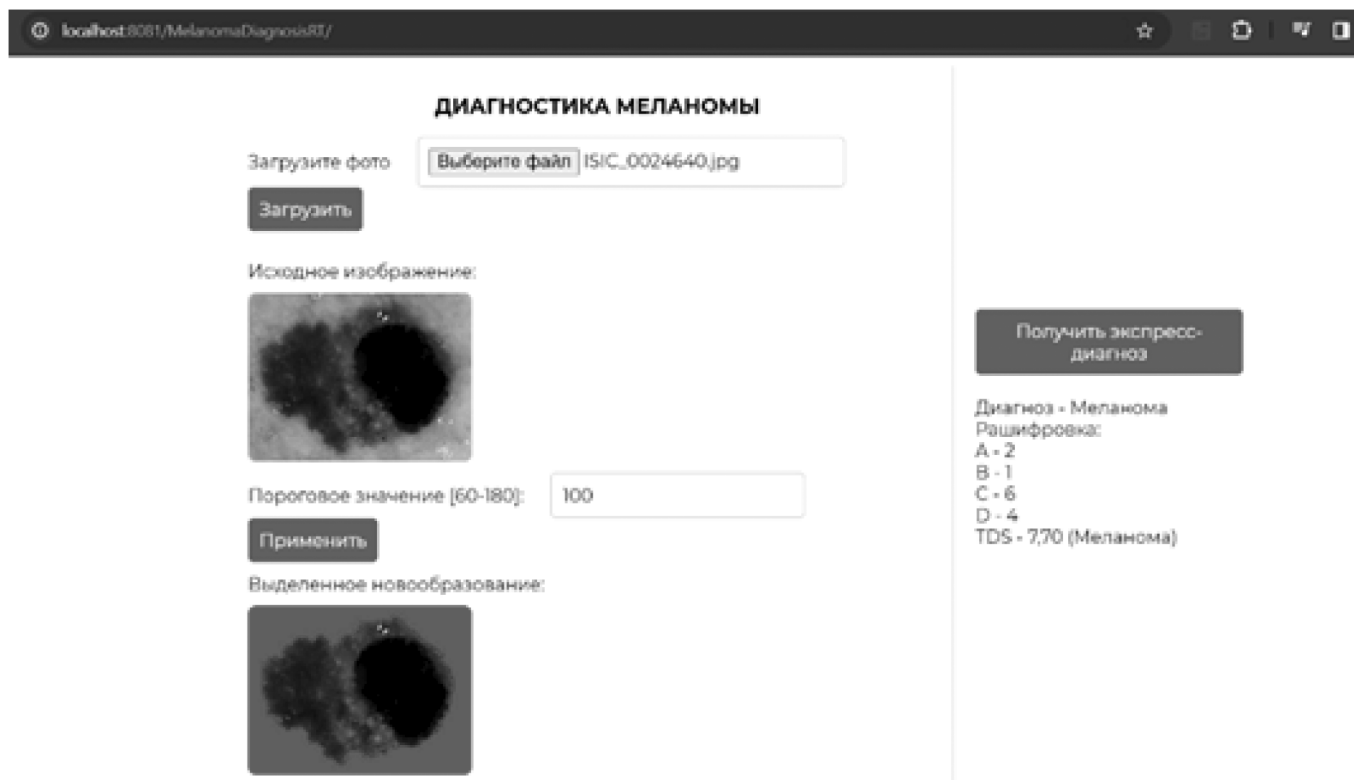


Рис. 3. Интерфейс приложения

Интерфейс приложения представлен на рисунке 3, содержит область загрузки изображения, область выбора порогового значения и поле просмотра выделяемого изображения, область с диагнозом.

Заключение

Разработанное web-приложение прежде всего основано на медицинском диагностическом алгоритме и берет за основу «оцифровку» визуального осмотра больного, при этом анализ изображения составляет пару секунд.

Предлагаемый алгоритм автоматизированной диагностики меланомы на основе дерматоскопического метода Штольца показал хорошие результаты на реальных наборах изображений. Пониженная точность модели на наборе NAM10000 обусловлена несбалансированностью классов (соотношение классов 1 к 9).

Высокое значение F-меры указывает на хороший баланс между чувствительностью и специфичностью модели, модель склонна лучше обнаруживать меланомы — истинно положительные случаи (TP). Высокая чувствительность является желательной в задачах, где снижение ложноотрицательных результатов (пропущенных положительных случаев) критически важно, даже за счет увеличения числа ошибок первого рода.

Полнота модели составила 88,54 % и 95 % и говорит о том, что в модели минимизированы случаи пропусков заболевания.

Web-приложение экспресс-диагностики меланомы «MelanomaDiagnosisRT», описанное в статье, было зарегистрировано в Реестре программ для ЭВМ 15 марта 2024 г. № 2024616064 и внедрено в учебный процесс в ФГБОУ ВО СамГМУ Минздрава России на кафедре медицинской физики, математики и информатики.

ЛИТЕРАТУРА

1. Гаранина О.Е., Самойленко И.В., Шлишко И.Л., Клеменова И.А., Незнахина М.С., Демидов Л.В. Неинвазивные методы диагностики опухолей кожи и их потенциал применения для скрининга меланомы кожи: систематический обзор литературы. Медицинский совет. 2020;(9):102–120. doi: 10.21518/2079-701X-2020-9-102-120.
2. Гаранина О.Е., Шлишко И.Л., Клеменова И.А., Ускова К.А., Миронычева А.М., Дардык В.И., Ласьков В.Н. Искусственный интеллект: как работает и критерии оценки // Consilium Medicum. 2021. №8. URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-kak-rabotaet-i-kriterii-otsenki> (дата обращения: 07.12.2023).

3. Гаврилов Д.А., Закиров Э.И., Гамеева Е.В., Семенов В.Ю., Александрова О.Ю. Автоматизированная диагностика меланомы кожи на основе математической модели искусственной сверточной нейронной сети. *Research'nPracticalMedicine Journal*. 2018;5(3):110–116. <https://doi.org/10.17709/2409-2231-2018-5-3-11>
4. Козачок А.В., Спирин А.А., Козачок Е.С. Обзор методов раннего обнаружения меланомы. *Труды ИСП РАН*, том 34, вып. 4, 2022 г., стр. 241–250. 10.15514/ISPRAS-2022-34(4)-17
5. Кудрин К.Г., Римская Е.Н., Аполлонова И.А., Николаев А.П., Черномырдин Н.В., Святославов Д.С., Давыдов Д.В., Решетов И.В. Ранняя диагностика меланомы кожи с применением нескольких изображающих систем // *Оптика и спектроскопия*. — 2020. — том 128. — вып. 6 DOI:10.21883/OS.2020.06.49416.53–20
6. Ляхов П.А. Система нейросетевой классификации пигментных новообразований кожи с предварительным удалением волос на фотографиях / П.А. Ляхов, У.А. Ляхова // *Компьютерная оптика*. — 2021. — Т. 45, № 5. — С. 728–735. — DOI: 10.18287/2412-6179-CO-863.
7. Минкина О.В., Дворников А.С., Скрипкина П.А., Оганесян Л.В., Палагина В.С., Иванова К.С. Новообразования кожи: современные представления о неинвазивных возможностях и перспективах диагностики. *Профилактическая медицина*. — 2020; — 23(6): Вып. 1:120–128. <https://doi.org/10.17116/profmed202023061120>
8. Мошкин А.В. Чувствительность и специфичность как клинические индикаторы качества лабораторных исследований. *Лабораторная служба*. 2020;9(4):5–6.
9. Неретин Е.Ю., Куколкина В.В. Возможности диагностики меланомы кожи с применением экспертной системы // *Поволжский онкологический вестник*. 2014. №3. URL: <https://cyberleninka.ru/article/n/vozmozhnosti-dianostiki-melanomy-kozhi-s-primeneniem-ekspertnoy-sistemy> (дата обращения: 06.12.2023).
10. ПроРодинки [Электронный ресурс]. — URL: <https://prorodinki.ru/> (Дата обращения 08.12.2023)
11. Прохоренок Н.В., ОренCV и Java. Обработка изображений и компьютерное зрение. — СПб.: БХВ-Петербург, 2018. — 320 с.
12. Распознавание злокачественных новообразований на самых ранних стадиях. диагностика на fotofinder [Электронный ресурс]. — URL: <https://spbkvd1.ru/services/raspoznvanie-zlokachestvennyh-novoobrazovanii-na-samyh-rannih-stadiyah>
13. Томакова Р.А., Дзюбин И.А., Брежнев А.В. Метод и алгоритм обучения сверточной нейронной сети, предназначенной для интеллектуальной системы распознавания меланомы // *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение*. 2022; 12(1): 65–83
14. ABCD Rule [Stolz et al. *Eur J Dermatol* 1994] [Electronic resource]. URL: <https://dermoscopy.org/consensus/2b.asp> (Дата обращения к ресурсу 22.11.2023)
15. Argenziano G. 3-point checklist of dermoscopy. Available from: <https://dermoscopy-ids.org/wp-content/uploads/2015/05/3pt.tu-torial.pdf> (Дата обращения 06.12.2023).
16. Babar, K.R. Dermatoscopy for melanoma and pigmented lesions / K.R. Babar, S.A. Christine // *Dermatol. Clin.* — 2012. — Vol. 30, № 3. P:413–434.
17. Budhiman A., Suyanto S., Arifianto A., «Melanoma Cancer Classification Using ResNet with Data Augmentation» 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2019, pp. 17–20, doi: 10.1109/ISRITI48646.2019.9034624.
18. Calzavara-Pinton, P., Longo, C., Venturini, M., Sala, R. and Pellacani, G. (2008), Reflectance Confocal Microscopy for In Vivo Skin Imaging. *Photochemistry and Photobiology*, 84: 1421–1430. <https://doi.org/10.1111/j.1751-1097.2008.00443.x>
19. Drexel Laboratory of Automated Systems [Electronic resource]. 2002. URL: http://www.pages.drexel.edu/~weg22/can_tut.html (Дата обращения 28.10.2023)
20. Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria / N.R. Abbasi., H.M. Shaw, D. S. Rigel [et al.] // *JAMA*. — 2004. — Vol. 292, № 22 P:2771–2776.
21. Geethu H., Shamna S., Kizhakkethottam J. Weighted Guided Image Filtering and Haze Removal in Single Image // *Procedia Technology*. — 2016. — 24. — p.1475–1482.
22. Haenssle H.A., Fink C. et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of oncology*, vol. 29, issue 8, 2018, pp. 1836–1842.
23. Hoffmann, K. et al. (1998). Optical Coherence Tomography (OCT) in der Dermatologie. In: Garbe, C., Rassner, G. (eds) *Dermatologie*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-71974-5_1
24. Kittler H. Evolution of the Clinical, Dermoscopic and Pathologic Diagnosis of Melanoma. *Dermatology Practical & Conceptual*, vol. 11, no. S1, 2021, pp. 1–10.
25. Nilsson T, Soderberg M, Lundqvist G, et al. A comparison of spiral computed tomography and latex agglutination D-dimer assay in acute pulmonary embolism using pulmonary arteriography as gold standard. *Scand Cardiovasc J* 2002;36:373–7.
26. Ngoo A, Finnane A, McMeniman E, et al. Fighting melanoma with smartphones: a snapshot of where we are a decade after app stores opened their doors. *Int J Med Inform*. 2018;118:99–112
27. Ohsaki M. et al. Confusion-matrix-based kernel logistic regression for imbalanced data classification // *IEEE Transactions on Knowledge and Data Engineering*. — 2017. — Т. 29. — No. 9. — С. 1806–1819.
28. Pakdemirli E. Artificial intelligence in radiology: friend or foe? Where are we now and where are we heading? *Acta Radiol Open*. 2019;8(2):2058460119830222.
29. Saghir U., Devendran V. A Brief Review of Feature Extraction Methods for Melanoma Detection. In *Proc. of the 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, 2021, pp. 1304–1307.
30. Thanh D.N., Prasath V.B. et al. Melanoma skin cancer detection method based on adaptive principal curvatures, colour normalisation and feature extraction with the ABCD rule. *Journal of Digital Imaging*, vol.33, issue 3, 2020, pp. 574–585
31. Thanh D.N.H., Hien N.N. et al. Adaptive Thresholding Skin Lesion Segmentation with Gabor Filters and Principal Component Analysis. *Advances in Intelligent Systems and Computing*, vol. 1125, 2019, pp. 811–820.
32. Thanh D.N.H., Hien N.N. et al. Automatic Initial Boundary Generation Methods Based on Edge Detectors for the Level Set Function of the Chan-Vese Segmentation Model and Applications in Biomedical Image Processing. *Advances in Intelligent Systems and Computing*, vol. 1125, 2019, 2018, pp. 171–181.
33. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Sci Data*. 2018;5:180161–180161. [PMC free article] [PubMed] [Google Scholar]