

Мультимодальная ансамблевая нейросетевая система обнаружения рака кожи на основе анализа гетерогенных дерматологических данных*

У. А. Ляхова, П. А. Ляхов

Северо-Кавказский федеральный университет,
Российская Федерация, 355017, Ставрополь, ул. Пушкина, 1

Для цитирования: Ляхова У. А., Ляхов П. А. Мультимодальная ансамблевая нейросетевая система обнаружения рака кожи на основе анализа гетерогенных дерматологических данных // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2024. Т. 20. Вып. 2. С. 231–243.

<https://doi.org/10.21638/spbu10.2024.208>

В настоящее время рак кожи служит одной из ведущих причин смертности в мире. Его диагностика на ранних этапах имеет решающее значение для увеличения потенциальной выживаемости. Поэтому важна разработка высокоточных интеллектуальных систем вспомогательной диагностики для выявления рака кожи на ранних стадиях. Ансамблевое обучение — один из актуальных и перспективных методов повышения точности систем интеллектуальной классификации за счет уменьшения дисперсии и вариативности прогнозов отдельных составляющих общей системы. Представлена ансамблевая интеллектуальная система анализа гетерогенных дерматологических данных на основе мультимодальных нейронных сетей. Точность разработанной ансамблевой системы составила 85.92 %, что на 1.85 процентных пункта выше по сравнению со средней точностью отдельных мультимодальных архитектур для классификации гетерогенных дерматологических данных. Предложенная система может использоваться как высокоточный вспомогательный диагностический инструмент, помогающий принять медицинское решение, что позволит повысить шанс раннего выявления пигментных онкопатологий.

Ключевые слова: мультимодальная нейронная сеть, ансамблевая нейронная сеть, машинное обучение, гетерогенные данные, дерматологические изображения, пигментные новообразования кожи, рак кожи, меланома.

1. Введение. Согласно данным Всемирной организации здравоохранения (ВОЗ), рак является ведущей причиной смертности в мире, а рак кожи занимает пятое место среди наиболее распространенных форм онкопатологий [1]. Ранняя диагностика рака кожи имеет решающее значение для увеличения потенциальной выживаемости, однако диагностика на ранних стадиях сложна и субъективна из-за схожих морфологических признаков злокачественных и доброкачественных пигментных новообразований. Средняя точность метода визуальной диагностики пигментных новообразований зависит от подготовки и квалификации врача-дерматолога и составляет 65–75 % [2]. Несмотря на то, что гистопатологический анализ считается «золотым стандартом» диагностики рака кожи, данный метод относится к инвазивным и может быть неубедительным в пограничных случаях, когда расхождения в диагнозе,

* Исследование в п. 2 выполнено в Северо-Кавказском центре математических исследований в рамках соглашения с Министерством науки и высшего образования Российской Федерации (соглашение № 075-02-2024-1451), в п. 3 — за счет гранта Российского научного фонда № 23-71-10013 (<https://rscf.ru/project/23-71-10013/>). Авторы благодарят Северо-Кавказский федеральный университет за помощь в рамках проекта поддержки малых научных групп и отдельных ученых.

© Санкт-Петербургский государственный университет, 2024

поставленном разными патологоанатомами, могут составлять до 25 % [3]. Поэтому в настоящее время актуальна разработка высокоточных интеллектуальных систем, которые могут использоваться в качестве вспомогательных диагностических инструментов для выявления рака кожи на ранних стадиях [4]. Ансамблевое обучение является одним из актуальных и перспективных методов повышения точности интеллектуальной классификации [5]. Он объединяет выходные сигналы нескольких классификаторов для получения общего решения путем голосования. Ансамблевые методы дают возможность уменьшать дисперсию и вариативность прогнозов, при этом значительно повышая точность нейросетевой классификации [6].

Цель исследования — разработка ансамблевой мультимодальной нейросетевой системы распознавания пигментных новообразований кожи за счет объединения мультимодальных архитектур анализа гетерогенных дерматологических данных в ансамблевую модель взвешенного голосования. Представленная система позволит уменьшить количество ошибочных классификаций и значительно повысить надежность, благодаря использованию как их разнообразных подходов к классификации данных, так и разнородных дерматологических данных и этапа предварительной очистки волосяных структур для подготовки визуальных данных к дальнейшему извлечению признаков. Новизна предлагаемой ансамблевой интеллектуальной системы заключается в одновременном применении метода предварительной очистки волосяных структур на дерматологических изображениях и ансамбля мультимодальных архитектур анализа гетерогенных данных.

2. Ансамблевая система распознавания пигментных новообразований кожи на основе мультимодальных нейронных сетей. Ансамблевое обучение искусственного интеллекта в области дерматологии представляет собой различные методы для создания надежной и точной комбинированной модели путем агрегирования базовых нейросетевых архитектур для классификации дерматологических данных [7]. Ансамблевые модели объединяют результаты отдельных систем искусственного интеллекта для повышения точности за счет использования их разнообразных подходов к классификации данных. Каждая отдельная интеллектуальная система имеет свои преимущества в методах извлечения признаков и при этом ограниченную способность фиксировать разброс выходных данных. Объединение нескольких интеллектуальных систем в ансамблевую модель может повысить надежность результатов классификации, а также уменьшить систематическую ошибку и дисперсию [8]. Наиболее распространенная стратегия в создании ансамблевых моделей — вычисление среднего значения прогнозов, сделанных каждой моделью на аналогичном наборе данных. Ансамбль усреднения в равной степени объединяет прогнозы нескольких обученных моделей и генерирует средние комбинированные прогнозы [9]. Однако для обеспечения более высокой производительности необходимо учитывать разнообразие применяемых в ансамбле нейросетевых моделей и точность каждого выбранного классификатора. К возможным методам реализации разнообразия интеллектуальных систем для ансамблевого обучения относится создание нейросетевых моделей с использованием различных наборов оптимизаторов [10].

Мультимодальное машинное обучение представляет собой интеллектуальный анализ информации из разных модальностей [11]. Оно позволяет извлекать признаки из гетерогенных данных для определения неочевидной взаимосвязи между обрабатываемыми данными и результатами классификации [12]. В области дерматологии наиболее распространенными типами данных являются изображения пигментных новообразований кожи и такие метаданные как пол, возраст, локализация пигмент-

ного поражения на теле пациента. В качестве основного метода диагностирования онкопатологий используется визуальное клиническое исследование кожи. При этом статистические метаданные пациентов также могут указывать на риск развития злокачественных форм пигментных новообразований кожи [13]. Всесторонний анализ разнородных данных позволяет повысить точность и надежность результатов первичной классификации за счет поиска дополнительных связей между визуальными объектами исследований и метаданными как у медиков и специалистов, так и у систем искусственного интеллекта. Наиболее оптимальной интеллектуальной архитектурой для анализа статистических метаданных пациентов является многослойный линейный перцептрон с прямой связью [14]. Для обработки дерматологических изображений наиболее подходящий тип нейронных сетей — сверточные нейронные сети (СНС) [15].

Предлагаемая мультимодальная интеллектуальная система (МИС) анализа гетерогенных дерматологических данных состоит из сверточных слоев предварительно обученной нейронной сети и линейного перцептрона, которые параллельно извлекают признаки из данных различных модальностей. В результате обработки визуальных многомерных данных при помощи СНС происходит создание карт извлеченных признаков следующим образом:

$$I_j^k = f \left(\sum_i^{m^{k-1}} I_i^{k-1} w_{ij}^k + p_j^i \right),$$

где m^{k-1} — количество карт объекта в слое $k-1$; I_i^{k-1} — выход слоя активации; w_{ij}^k — весовые коэффициенты фильтра маски; p_j^i — коэффициент смещения.

В линейном перцептроне нейроны выполняют суммирование полученного вектора входных метаданных \mathbf{m} и коэффициента смещения p , формируя выходной сигнал:

$$M^k = f \left(\sum_i^k m_i w_i + p \right),$$

здесь w_i — весовой коэффициент нейрона.

На слое конкатенации производится объединение построенной карты признаков I_j^k и выходного сигнала M^k :

$$C = \sum_i \sum_j \sum_n I_j^k w_{ijn}^{(1)} + \sum_{i=1}^n M^k w_{in}^{(2)},$$

где $w_{ijn}^{(1)}$ — весовой коэффициент для обработки карт признаков I_j^k дерматологических визуальных данных; $w_{in}^{(2)}$ — весовой коэффициент для обработки извлеченных признаков M^k входного вектора метаданных.

Активация последнего слоя мультимодальной нейросетевой системы производится при помощи функции *softmax*. Полученное выходное распределение вероятностей между диагностическими категориями сравнивается с истинным распределением при помощи функции кросс-энтропийных потерь, в результате которого происходят постепенное запоминание истинных векторов и минимизация потерь при обучении. Регулирование значений весовых коэффициентов линейного перцептрона и сверточных слоев осуществляется при помощи выбранного метода оптимизации для более точной сходимости функции потерь в область глобального минимума. Стохастический

градиентный спуск (SGD) и адаптивная оценка моментов (Adam) в настоящее время являются наиболее распространенными алгоритмами оптимизации. При этом Yogi и Diffgrad относятся к наиболее передовым методам оптимизации процесса обучения, которые позволяют достичь наибольшей точности классификации.

Все МИС, обученные с применением разных параметров, объединяются в ансамблевую модель с использованием выбранного правила комбинирования при помощи усреднения. Карта прогноза ансамблевой модели строится путем суммирования значений прогнозов отдельных моделей для каждого D случая в наборе данных. Прогноз категории N для каждого случая вычисляется по функции $\Delta(S_k^{(x)}, t^N)$, которая сопоставляет прогнозируемое значение каждой отдельной модели с соответствующей меткой категории следующим образом:

$$\Delta(S_k^{(x)}, t^N) = \begin{cases} 1, & \text{если } S_k^{(x)} \in t^1, \\ 2, & \text{если } S_k^{(x)} \in t^2, \\ 3, & \text{если } S_k^{(x)} \in t^3, \\ 4, & \text{если } S_k^{(x)} \in t^4, \\ \dots & \\ N, & \text{если } S_k^{(x)} \in t^N. \end{cases}$$

Если $S_k^{(x)}$ — прогноз МИС, обученной с помощью оптимизатора k , то случай $D^{(x)}$ принадлежит категории N . Общее количество прогнозов, полученных от отдельных моделей для категории N , можно представить так:

$$A_n^{(x)} = \sum_{k=1}^n \Delta(S_k^{(x)}, t^n), \quad 1 \leq n \leq N.$$

Прогноз ансамблевой модели рассчитывается на основе правила взвешенного голосования:

$$A_n^{(ENS)} = \max A_n^{(x)}.$$

Таким образом, рассматриваемому случаю присваивается категория, получившая максимальное количество голосов от каждой нейросетевой модели. На рис. 1 представлена предлагаемая ансамблевая мультимодальная нейросетевая система анализа гетерогенных дерматологических данных на основе сверточной архитектуры Densenet_161. В результате предварительной обработки статистических метаданных формируется вектор входных признаков, включающий в себя 14 переменных, указанных в блоке «Дерматологические данные». Изображения также проходят этап предварительной обработки, в результате которого визуальные данные подаются на вход выбранной СНС. После получения карт признаков при помощи сверточных слоев формируется выходной сигнал, состоящий из распределения вероятностей по 10 распознаваемым категориям. Аналогичным образом формируется выходной сигнал из линейной нейросетевой архитектуры, состоящий также из 10 значений, соответствующих количеству распознаваемых категорий. На слое конкатенации производится объединение разномодальных сигналов. В результате классификации входных данных из вероятностного распределения выбирается максимальное значение, характерное диагностической категории. Все прогнозируемые категории каждой МИС суммируются, и методом взвешенного голосования ансамблевой модели выбирается категория, которой было присвоено наибольшее количество голосов.

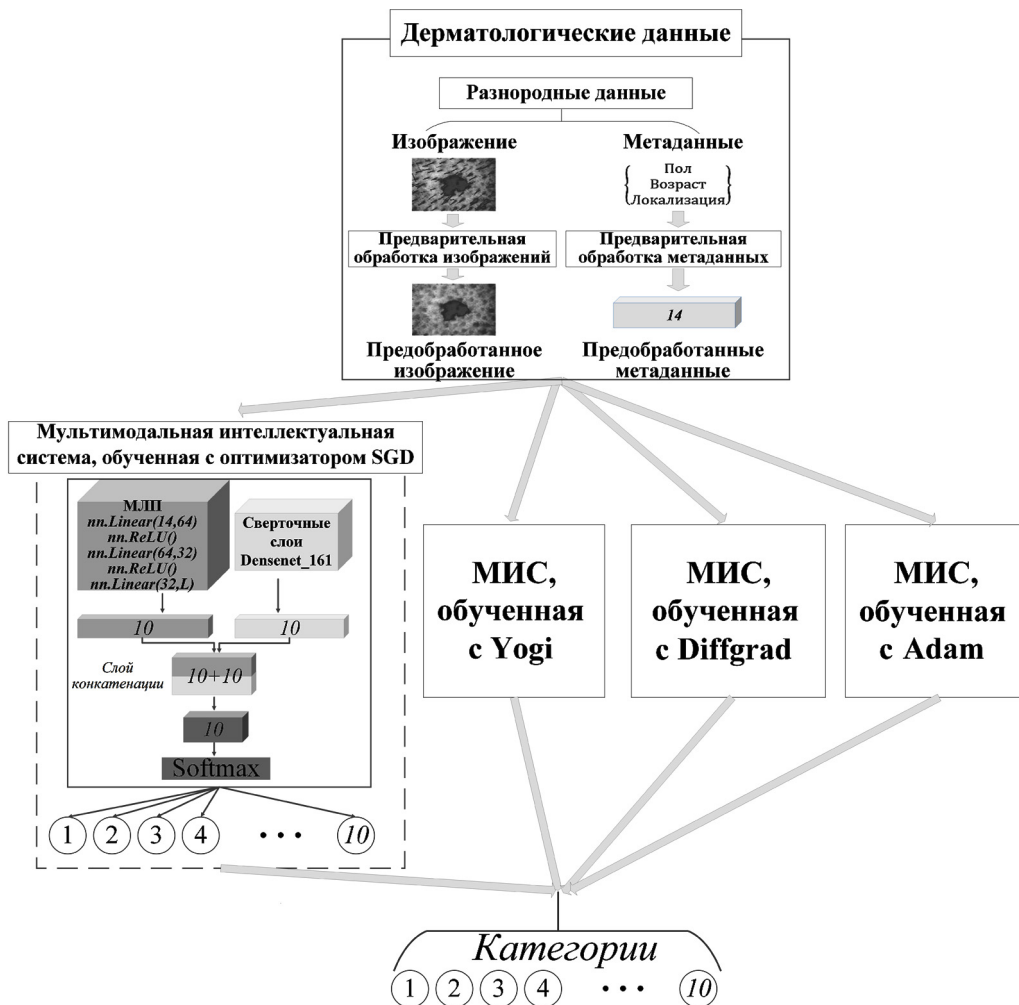


Рис. 1. Схема ансамблевой модели из мультимодальных интеллектуальных систем на основе архитектуры Densenet_161

3. Материалы и методы. Для практического моделирования использовались данные из открытого архива The International Skin Imaging Collaboration (ISIC). Отобранная база данных для обучения и тестирования состоит из 41 725 дерматологических изображений пигментных новообразований кожи, связанных с метаданными пациентов и гистопатологически подтвержденными диагнозами. Все данные разделены на 10 диагностически значимых категорий. Распределение отобранных дерматологических данных по диагностически значимым категориям представлено на рис. 2. Выбранные дерматологические категории подразделяются на группы «злокачественные» и «доброкачественные» и располагаются в порядке возрастания риска и тяжести течения заболевания. Поскольку актинический кератоз является внутриэпителиальной дисплазией кератиноцитов и характеризуется как преинвазивный плоскоклеточный рак (рак кожи стадии 0), то данная категория относится к группе «злокачественных» пигментных поражений кожи [16].

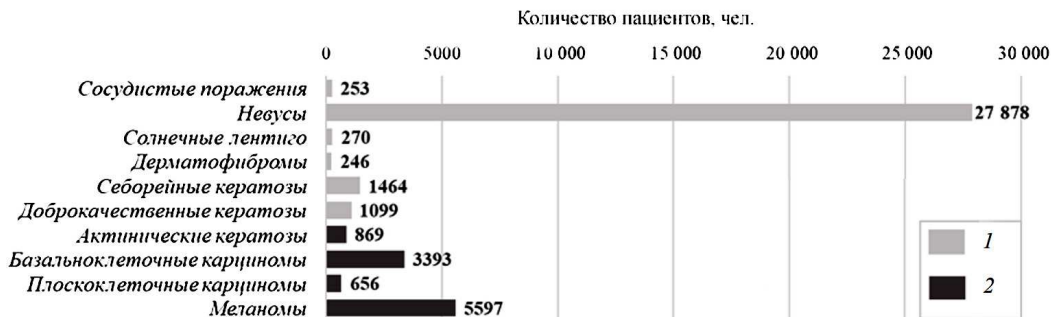


Рис. 2. График распределения отобранных дерматологических данных по диагностическим категориям: доброкачественным (1) и злокачественным (2)

Метаданные для каждого изображения представляют собой информацию о возрасте пациента, половой принадлежности и данные о локализации пигментного новообразования на теле. На этапе предварительной обработки метаданных фактор «Возраст» был разделен на четыре группы в соответствии с возрастной классификацией, принятой ВОЗ. Таким образом, вариабельность параметра «Возраст» была снижена с 18 до 4 возможных значений. Распределение дерматологических данных по различным статистическим факторам представлено на рис. 3, а–в.

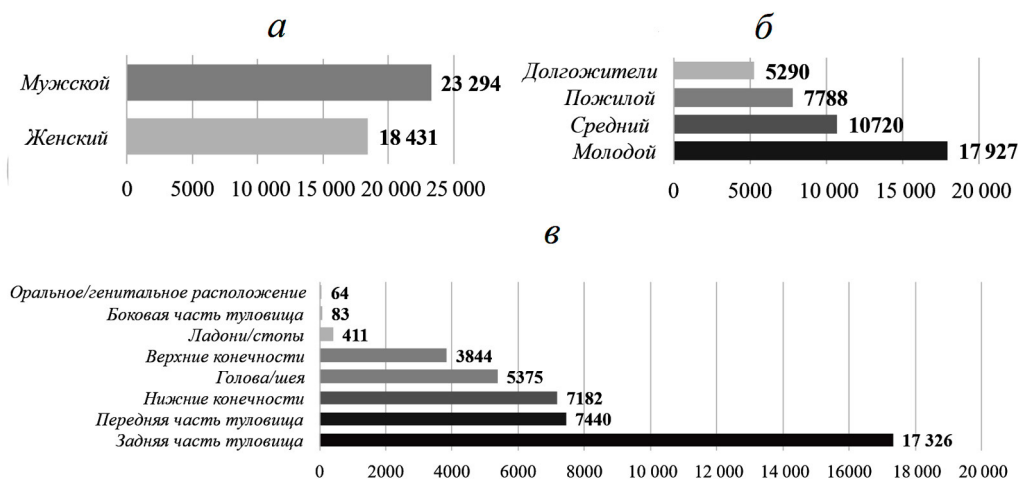


Рис. 3. Распределение отобранных дерматологических данных по статистическим факторам пациентов: по полу (а), по возрасту (б), по локализации пигментного образования на теле пациента (в)

Практическое моделирование проводилось на компьютере Intel(R) Core(TM) i5-8500 с частотой 3,00 ГГц с 16 Гб оперативной памяти и 64-битной операционной системой Windows 10 с использованием языка программирования высокого уровня Python 3.11.0. Для обучения мультимодальных нейросетевых систем и моделирования ансамблевой системы применялся фреймворк машинного обучения Pytorch совместно с графическим процессором (GPU) на базе видеочипсета NVIDIA GeForce GTX 1050TI, для обработки статистических данных — библиотеки NumPy, Pandas

и ScikitLearn, для визуализации данных — библиотека Matplotlib. Для моделирования всех мультимодальных интеллектуальных систем анализа дерматологических данных в качестве архитектуры для извлечения признаков из визуальных данных использовалась СНС Densenet_161. Выбранная сверточная архитектура была предварительно обучена на наборе естественных изображений ImageNet. В качестве методов для оптимизации процесса обучения были выбраны SGD, Adam, Yogi и Diffgrad.

Предварительная обработка визуальных данных состояла из преобразования дерматологических изображений до требуемого размера 224×224 пикселей и применении метода очистки от волосных структур, представленного в работе [17]. Предварительная обработка метаданных заключалась в преобразовании всех статистических факторов в вектор входных признаков при помощи метода one hot encoding. В соответствии с золотым стандартом машинного обучения для моделирования дерматологическая база данных была разделена на обучающие, валидационные и тестовые данные в процентном соотношении 60:20:20 (25 035 обучающих, 8345 валидационных и 8345 тестовых случаев). В табл. 1 представлено распределение дерматологических изображений по категориям в обучающей, валидационной и тестовой выборках.

Таблица 1. Распределение дерматологических изображений по категориям в обучающей, валидационной и тестовой выборках

Категория	Выборка		
	обучающая	валидационная	тестовая
Сосудистые поражения	151	51	51
Невусы	16 730	5574	5574
Солнечные лентиго	162	54	54
Дерматофибромы	146	50	50
Себорейные кератозы	878	293	293
Доброкачественные кератозы	659	220	220
Актинические кератозы	521	174	174
Базальноклеточные карциномы	2035	679	679
Плоскоклеточные карциномы	394	131	131
Меланомы	3359	1119	1119

Набор данных был разделен на обучающую, валидационную и тестовую выборки таким образом, чтобы соотношение визуальных данных в каждой категории было пропорциональным как для обучающего, так и для валидационного и тестового наборов. К обучающей выборке визуальных данных применялись аффинные преобразования, такие как отражение, вращение, перемещение, масштабирование и др., для минимизации риска возникновения переобучения моделей. Размер пакета входных данных составлял 8 примеров. Для всех оптимизаторов была выбрана одинаковая стандартная скорость обучения 0.001. На рис. 4 представлены графики точности при обучении и валидационной проверки МИС на основе СНС Densenet_161.

Предварительно обработанные дерматологические изображения из обучающей выборки подавались на вход архитектуры СНС Densenet_161, закодированный вектор признаков метаданных — на вход многослойной линейной нейросетевой архитектуры. На слое конкатенации производилось объединение разномодальных данных, полученных на выходных слоях Densenet_161, и линейной нейросетевой архитектуры. Объединенный сигнал подавался на вход слоя *softmax* для определения вероятностного соотношения прогнозируемых меток и выбора максимального значения. В процессе обучения использовался алгоритм принудительной остановки, при котором прекращалось обучение МИС, если в течение нескольких эпох отсутствовало увеличение валидационной точности. Таким образом, каждая МИС обучалась в те-

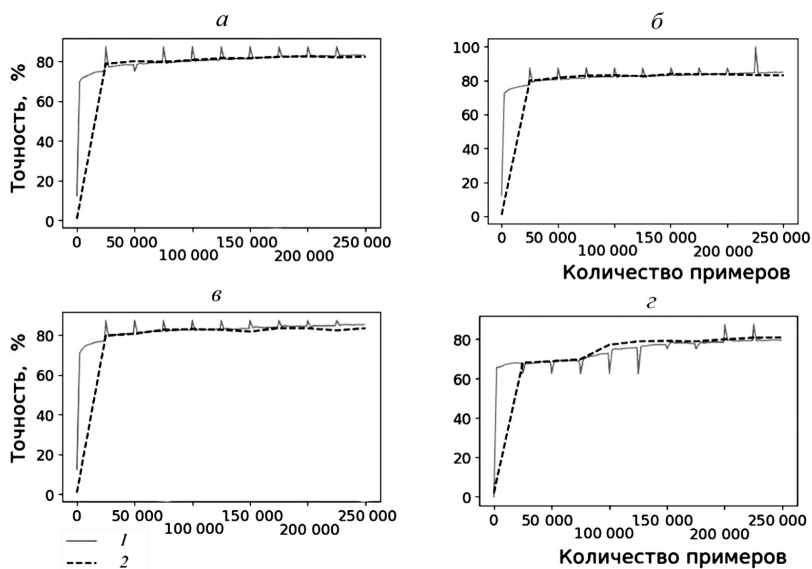


Рис. 4. Графики тренировочной (1) и валидационной (2) точности при проверке МИС на основе оптимизаторов: Adam (а), Diffgrad (б), Yogi (в), SGD (г)

чение 10 эпох. При большом количестве эпох наблюдалось выраженное переобучение каждой из предложенных мультимодальных систем.

4. Результаты моделирования. Для получения результатов тестирования ансамблевой модели все прогнозируемые категории каждой МИС суммировались, и методом взвешенного голосования для каждого случая выбиралась категория, которой было присвоено наибольшее количество голосов. На рис. 5 показаны примеры карт значимости для предлагаемых МИС.

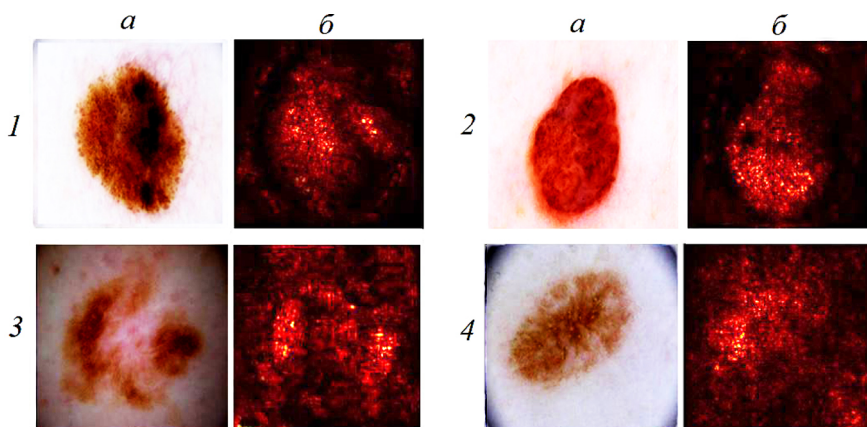


Рис. 5. Тестовые дерматологические изображения (а) и примеры карт значимости, полученные после тестирования МИС на основе архитектуры Densenet_161 в результате интеллектуального анализа (б) МИС, обученные при помощи оптимизаторов: 1 – Adam, 2 – Diffgrad, 3 – Yogi, 4 – SGD.

В табл. 2 представлены результаты оценки при тестировании всех обученных МИС на основе архитектуры Densenet_161 для классификации пигментных новообразований кожи. В качестве количественных методов тестовой статистической оценки моделей были выбраны специфичность (Sp), коэффициент корреляции Мэттьюса (MCC — Matthews correlation coefficient [18]), частота ложноотрицательных результатов (FNR — false negative rate), частота ложноположительных результатов (FPR — false positive rate).

Таблица 2. Результаты статистической оценки при тестировании каждой МИС на основе архитектуры Densenet_161 и предложенной ансамблевой модели

Модель	Точность, %	Sp	MCC	FNR	FPR
МИС_Adam	83.70	0.9819	0.6828	0.1630	0.0181
МИС_Diffgrad	83.94	0.9822	0.6904	0.1606	0.0178
МИС_Yogi	84.10	0.9823	0.6917	0.1590	0.0177
МИС_SGD	84.52	0.9828	0.7032	0.1548	0.0172
Ансамбль МИС	85.92	0.9844	0.7268	0.1408	0.0156

В результате тестирования наилучший показатель точности распознавания пигментных новообразований кожи составил 85.92 % и был получен при помощи разработанной ансамблевой модели взвешенных голосов на основе обученных мультимодальных систем. Предложенная ансамблевая модель на 1.40 процентных пункта точнее по сравнению с наилучшей по точности МИС, которая была обучена с использованием оптимизатора SGD, и на 1.85 процентных пункта точнее по сравнению со средней точностью всех обученных МИС (84.06 %). Наилучшие показатели Sp, MCC были получены при тестировании разработанной ансамблевой модели на основе различных МИС и составили 0.9844 и 0.7268 соответственно. Для таких мер статистической оценки разработанных интеллектуальных систем, как FNR и FPR, наилучшие показатели составили 0.1408 и 0.0156 и также были получены при тестировании разработанной ансамблевой модели взвешенных голосов на основе обученных мультимодальных систем. Приведенные результаты показывают, что объединение МИС в ансамблевую модель позволяет значительно повысить точность и надежность классификации пигментных новообразований кожи по сравнению с эффективностью каждой отдельной интеллектуальной системы.

5. Заключение. Представлена ансамблевая интеллектуальная система анализа гетерогенных дерматологических данных. Ее точность на основе мультимодальных архитектур со сверточными слоями Densenet_161 составила 85.92 %. Система состоит из нескольких мультимодальных архитектур, которые анализируют разнообразные дерматологические данные, представленные изображениями пигментных новообразований кожи и статистическими метаданными пациентов. Объединение МИС в ансамблевую модель позволило минимизировать возможные ошибки и дисперсию при классификации у каждой отдельной нейросетевой архитектуры, а также значительно повысить надежность за счет использования их разнообразных подходов к классификации данных. Разработанная ансамблевая модель взвешенных голосов на основе мультимодальных архитектур позволила добиться более высокой точности по сравнению с результатами тестирования каждой отдельной мультимодальной системы, а также по сравнению с результатами точности аналогичных ансамблевых систем распознавания злокачественных новообразований кожи [19–21]. В работе [19] представлена ансамблевая система на основе СНС AlexNet, VGGNet и GoogLeNet, тестовая точность которой составила 83.85 % для четырех диагностических катего-

рий, что на 2.07 процентного пункта ниже по сравнению с точностью предложенной ансамблевой интеллектуальной системы анализа гетерогенных дерматологических данных на основе МИС. Авторы работы [20] предложили ансамблевую модель на основе сверточных архитектур Resnet-50, VGG-16, Densenet, Mobilenet, Inceptionv3 и Xception, тестовая точность которой равна 78.00 %, что на 7.92 процентного пункта ниже по сравнению с тестовой точностью предложенной взвешенной ансамблевой модели на основе мультимодальных архитектур. Авторы [21] предложили различные варианты ансамблевых моделей, состоящих из двух архитектур, обученных на одном оптимизаторе. Для ансамбля из СНС DenseNet-201 и MobileNet, обученных с оптимизатором Adam, точность составила 84.43 %, что на 1.49 процентного пункта ниже, чем у предложенной ансамблевой системы на основе МИС. Для архитектур DenseNet-201 и ResNet-50, обученных с использованием оптимизатора Adam, тестовая точность также составила 84.43 % для 25 эпох, что на 1.49 процентного пункта ниже по сравнению с разработанной взвешенной ансамблевой моделью анализа гетерогенных дерматологических данных.

Основным ограничением представленной ансамблевой нейросетевой системы на основе мультимодальных архитектур для анализа гетерогенных дерматологических данных является то, что медики и специалисты могут применять систему только как вспомогательный диагностический инструмент. Такая система не может быть использована в качестве самостоятельного метода диагностики из-за возможного риска возникновения ложноотрицательных ошибок. Однако ее можно считать высокоточным вспомогательным инструментом, помогающим принять медицинское решение, что позволит сократить потребление финансовых и трудовых ресурсов, задействованных в медицинской отрасли, а также повысит шанс раннего выявления пигментных онкопатологий.

Литература

1. *Apalla Z., Lallas A., Sotiriou E., Lazaridou E., Ioannides D.* Epidemiological trends in skin cancer // *Dermatol. Pract. Concept.* 1885. Vol. 7. Iss. 2. P. 1.
2. *Sinz C., Tschandl P., Rosendahl C., Akay B. N., Argenziano G., Blum A., Kittler H.* Accuracy of dermatoscopy for the diagnosis of nonpigmented cancers of the skin // *Journal of Acad. Dermatol.* 2017. Vol. 77. Iss. 6. P. 1100–1109.
3. *Lodha S., Saggarr S., Celebi J. T., Silvers D. N.* Discordance in the histopathologic diagnosis of difficult melanocytic neoplasms in the clinical setting // *Journal of Cutan Pathol.* 2008. Vol. 35. Iss. 4. P. 349–352.
4. *Haggemuller S., Maron R. C., Hekler A., Utikal J. S., Barata C., Barnhill R. L., Brinker T. J.* Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts // *Eur. Journal of Cancer.* 2021. Vol. 156. P. 202–216.
5. *Polikar R., Zhang C., Ma Y.* Ensemble Learning // *Ensemble Machine Learning.* 2012. P. 1–34.
6. *Qureshi A. S., Roos T.* Transfer learning with ensembles of deep neural networks for skin cancer detection in imbalanced data sets // *Neural Process Lett.* 2023. Vol. 55. Iss. 4. P. 4461–4479.
7. *Raza R., Zulfiqar F., Tariq S., Anwar G. B., Sargano A. B., Habib Z.* Melanoma classification from dermoscopy images using ensemble of convolutional neural networks // *Mathematics.* 2021. Vol. 10. Iss. 1. P. 26–43.
8. *Kausar N., Hameed A., Sattar M., Ashraf R., Imran A. S., Abidin M. Z. U., Ali A.* Multiclass skin cancer classification using ensemble of fine-tuned deep learning models // *Applied Sciences.* 2021. Vol. 11. Iss. 22. P. 10593–10608.
9. *Lu Y., Zhang L., Wang B., Yang J.* Feature ensemble learning based on sparse autoencoders for image classification // *Proceedings of the International Joint Conference on Neural Networks.* 2014. P. 1739–1745.
10. *Tang E. K., Suganthan P. N., Yao X.* An analysis of diversity measures // *Machine Learning.* 2006. Vol. 65. Iss. 1. P. 247–271.

11. Baltrušaitis T., Ahuja C., Morency L. P. Multimodal machine learning // IEEE Trans. Pattern Anal. Mach. Intell. 2019. Vol. 41. Iss. 2. P. 423–443.
12. Liu K., Li Y., Xu N., Natarajan P. Learn to combine modalities in multimodal deep learning // arXiv preprint. arXiv:1805.11730. 2018.
13. Kurtansky N. R., Dusza S. W., Halpern A. C., Hartman R. I., Geller A. C., Marghoob A. A., Marchetti M. A. An epidemiologic analysis of melanoma overdiagnosis in the United States, 1975–2017 // Journal of Investigative Dermatology. 2022. Vol. 142. Iss. 7. P. 1804–1811.
14. Hohn J., Hekler A., Kriehoff-Henning E., Kather J. N., Utikal J. S., Meier F., Brinker T. J. Integrating patient data into skin cancer classification using convolutional neural networks: systematic review // Journal of Medical Internet Research. 2021. Vol. 23. Iss. 7. P. 20708–20723.
15. Sriwong K., Bunrit S., Kerdprasop K., Kerdprasop N. Dermatological classification using deep learning of skin image and patient background knowledge // International Journal of Machine Learning and Computing. 2019. Vol. 9. Iss. 6. P. 862–867.
16. Siegel J. A., Korgavkar K., Weinstock M. A. Current perspective on actinic keratosis: a review // British Journal of Dermatology. 2017. Vol. 177. Iss. 2. P. 350–358.
17. Lyakhov P. A., Lyakhova U. A., Nagornov N. N. System for the recognizing of pigmented skin lesions with fusion and analysis of heterogeneous data based on a multimodal neural network // Cancers. 2022. Vol. 14. P. 1819–1836.
18. Chicco D., Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation // BMC Genomics. 2020. Vol. 21. Iss. 1. P. 1–13.
19. Harangi B., Baran A., Hajdu A. Classification of skin lesions using an ensemble of deep neural networks // Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2018. P. 2575–2578.
20. Akter M. S., Shahriar H., Sneha S., Cuzzocrea A. Multi-class skin cancer classification architecture based on deep convolutional neural network // 2022 IEEE International Conference on Big Data. Proceedings. 2022. P. 5404–5413.
21. Keerthana D., Venugopal V., Nath M. K., Mishra M. Hybrid convolutional neural networks with SVM classifier for classification of skin cancer // Biomedical Engineering Advances. 2023. Vol. 5. P. 100069–100103.

Статья поступила в редакцию 2 октября 2023 г.

Статья принята к печати 12 марта 2024 г.

Контактная информация:

Ляхова Ульяна Алексеевна — аспирант, мл. науч. сотр.; uljahovs@mail.ru

Ляхов Павел Алексеевич — канд. физ.-мат. наук, доц.; ljahov@mail.ru

Multimodal ensemble neural network system for skin cancer detection on heterogeneous dermatological data*

U. A. Lyakhova, P. A. Lyakhov

North-Caucasus Federal University, 1, Pushkina ul., Stavropol, 355017, Russian Federation

For citation: Lyakhova U. A., Lyakhov P. A. Multimodal ensemble neural network system for skin cancer detection on heterogeneous dermatological data. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2024, vol. 20, iss. 2, pp. 231–243. <https://doi.org/10.21638/spbu10.2024.208> (In Russian)

* The research in section 2 was supported in the North-Caucasus Center for Mathematical Research under an agreement with the Ministry of Science and Higher Education of the Russian Federation (agreement N 075-02-2024-1451), the research in section 3 was supported by the Russian Science Foundation, project N 23-71-10013 (<https://rscf.ru/project/23-71-10013/>). The authors thank the North-Caucasus Federal University for assistance within the framework of the project to support small scientific groups and individual scientists.

Today, skin cancer is one of the leading causes of death in the world. Diagnosing skin cancer early is critical to increasing potential survival. Therefore, it is relevant to develop high-precision intelligent auxiliary diagnostic systems for detecting skin cancer in the early stages. Ensemble learning is one of the current and promising methods for increasing the accuracy of intelligent classification systems by reducing the dispersion and variability of predictions of individual components of the overall system. The work proposes an ensemble intelligent system for analyzing heterogeneous dermatological data based on multimodal neural networks. The accuracy of the developed ensemble system was 85.92 %, which is 1.85 percentage points higher than the average accuracy of individual multimodal architectures for classifying heterogeneous dermatological data. The developed system can be used as a high-precision auxiliary diagnostic tool to help make a medical decision, which will increase the chance of early detection of pigmented oncological pathologies.

Keywords: multimodal neural network, ensemble neural network, machine learning, heterogeneous data, dermatological images, pigmented skin lesions, skin cancer, melanoma.

References

1. Apalla Z., Lallas A., Sotiriou E., Lazaridou E., Ioannides D. Epidemiological trends in skin cancer. *Dermatol. Pract. Concept.*, 1885, vol. 7, iss. 2, pp. 1.
2. Sinz C., Tschandl P., Rosendahl C., Akay B. N., Argenziano G., Blum A., Kittler H. Accuracy of dermatoscopy for the diagnosis of nonpigmented cancers of the skin. *Journal of Acad. Dermatol.*, 2017, vol. 77, iss. 6, pp. 1100–1109.
3. Lodha S., Saggarr S., Celebi J. T., Silvers D. N. Discordance in the histopathologic diagnosis of difficult melanocytic neoplasms in the clinical setting. *Journal of Cutan Pathol.*, 2008, vol. 35, iss. 4, pp. 349–352.
4. Haggemuller S., Maron R. C., Hekler A., Utikal J. S., Barata C., Barnhill R. L., Brinker T. J. Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts. *Eur. Journal of Cancer.*, 2021, vol. 156, pp. 202–216.
5. Polikar R., Zhang C., Ma Y. Ensemble Learning. *Ensemble Machine Learning*, 2012, pp. 1–34.
6. Qureshi A. S., Roos T. Transfer learning with ensembles of deep neural networks for skin cancer detection in imbalanced data sets. *Neural Process Lett.*, 2023, vol. 55, iss. 4, pp. 4461–4479.
7. Raza R., Zulfiqar F., Tariq S., Anwar G. B., Sargano A. B., Habib Z. Melanoma classification from dermoscopy images using ensemble of convolutional neural networks. *Mathematics*, 2021, vol. 10, iss. 1, pp. 26–43.
8. Kausar N., Hameed A., Sattar M., Ashraf R., Imran A. S., Abidin M. Z. U., Ali A. Multiclass skin cancer classification using ensemble of fine-tuned deep learning models. *Applied Sciences*, 2021, vol. 11, iss. 22, pp. 10593–10608.
9. Lu Y., Zhang L., Wang B., Yang J. Feature ensemble learning based on sparse autoencoders for image classification. *Proceedings of the International Joint Conference on Neural Networks*, 2014, pp. 1739–1745.
10. Tang E. K., Suganthan P. N., Yao X. An analysis of diversity measures. *Machine Learning*, 2006, vol. 65, iss. 1, pp. 247–271.
11. Baltrušaitis T., Ahuja C., Morency L. P. Multimodal machine learning. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019, vol. 41, iss. 2, pp. 423–443.
12. Liu K., Li Y., Xu N., Natarajan P. Learn to combine modalities in multimodal deep learning. *arXiv preprint. arXiv:1805.11730*, 2018.
13. Kurtansky N. R., Dusza S. W., Halpern A. C., Hartman R. I., Geller A. C., Marghoob A. A., Marchetti M. A. An epidemiologic analysis of melanoma overdiagnosis in the United States, 1975–2017. *Journal of Investigative Dermatology*, 2022, vol. 142, iss. 7, pp. 1804–1811.
14. Hohn J., Hekler A., Kriehoff-Henning E., Kather J. N., Utikal J. S., Meier F., Brinker T. J. Integrating patient data into skin cancer classification using convolutional neural networks: systematic review. *Journal of Medical Internet Research*, 2021, vol. 23, iss. 7, pp. 20708–20723.
15. Sriwong K., Bunrit S., Kerdprasop K., Kerdprasop N. Dermatological classification using deep learning of skin image and patient background knowledge. *International Journal of Machine Learning and Computing*, 2019, vol. 9, iss. 6, pp. 862–867.
16. Siegel J. A., Korgavkar K., Weinstock M. A. Current perspective on actinic keratosis: a review. *British Journal of Dermatology*, 2017, vol. 177, iss. 2, pp. 350–358.

17. Lyakhov P. A., Lyakhova U. A., Nagornov N. N. System for the recognizing of pigmented skin lesions with fusion and analysis of heterogeneous data based on a multimodal neural network. *Cancers*, 2022, vol. 14, pp. 1819–1836.
18. Chicco D., Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 2020, vol. 21, iss. 1, pp. 1–13.
19. Harangi B., Baran A., Hajdu A. Classification of skin lesions using an ensemble of deep neural networks. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2018, pp. 2575–2578.
20. Akter M. S., Shahriar H., Sneha S., Cuzzocrea A. Multi-class skin cancer classification architecture based on deep convolutional neural network. *2022 IEEE International Conference on Big Data. Proceedings*, 2022, pp. 5404–5413.
21. Keerthana D., Venugopal V., Nath M. K., Mishra M. Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. *Biomedical Engineering Advances*, 2023, vol. 5, pp. 100069–100103.

Received: October 2, 2023.

Accepted: March 12, 2024.

Authors' information:

Ulyana A. Lyakhova — Postgraduate Student, Junior Researcher; uljahovs@mail.ru

Pavel A. Lyakhov — PhD in Physics and Mathematics, Associate Professor; ljahov@mail.ru