

УДК 004.93.12

DOI:10.25729/ESI.2024.35.3.004

Анализ снимков кожных новообразований с применением комбинированной архитектуры сверточных нейронных сетей

Милантьев Сергей Андреевич^{1,2}, Святкина Виталия Игоревна², Бессмертный Игорь Александрович¹, Зайченко Кирилл Вадимович²

¹Университет ИТМО, Россия, Санкт-Петербург, nevdev@niuitmo.ru

²ИАП РАН, Россия, Санкт-Петербург

Аннотация. В данном исследовании рассматривается возможность применения комбинированных архитектур сверточных нейронных сетей для анализа кожных новообразований. Разработаны архитектуры моделей для извлечения дополнительных признаков, связанных с характером формы кожных новообразований. Проведена оптимизация моделей, а также архитектуры, с целью минимизации ошибок первого и второго рода для редких кожных новообразований. В процессе обучения использовались наборы данных ISIC2017-2020, MED-NODE, SD-198, 7-Point Criteria Database, Light Field Image Dataset of Skin Lesions, PH², датасет ИАП РАН. Для обучения классификационных моделей использовались оптимизатор AdamW, функции потерь FocalLoss и scheduler CosineAnnealingWarmRestarts. Для обучения сегментационных моделей применялась функция потерь BCEDice. Оценка моделей проводилась с использованием взвешенных классификационных метрик, таких, как Recall, Precision и F1-score. В качестве рассматриваемой концепции моделей учитывалась их устойчивость на этапе валидации. Модели, использующие дополнительные сверточные нейронные сети для извлечения признаков формы новообразования, демонстрируют более высокие показатели метрик, а также имеют меньшую сумму ошибок первого и второго рода для редких заболеваний по сравнению с обычными классификационными моделями. Полученные результаты могут найти применение в анализе медицинских задач с дисбалансом данных в обучающем наборе данных.

Ключевые слова: кожные новообразования, сверточные нейронные сети, анализ кожных поражений, дисбаланс классов, многоспектральная обработка изображений

Цитирование: Милантьев С.А. Анализ снимков кожных новообразований с применением комбинированной архитектуры сверточных нейронных сетей / С.А. Милантьев, В.И. Святкина, И.А. Бессмертный, К.В. Зайченко // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 44-56. – DOI: 10.25729/ESI.2024.35.3.004.

Введение. Ранняя диагностика злокачественных новообразований кожи является актуальным исследовательским направлением в области дерматологии. Особое внимание уделяется классификации редких кожных поражений в задаче многоклассовой классификации. Анализ редких кожных новообразований представляет собой важное направление исследований, поскольку эти патологии являются малоизученными и вызывают особый интерес у исследователей. В настоящее время предпринимаются значительные усилия для развития новых методов и технологий с целью повышения точности распознавания и классификации редких кожных новообразований. В частности, применение методов искусственного интеллекта становится все более значимым при разработке инновационных подходов к диагностированию, что позволяет улучшить качество и надежность получаемой информации, а также обеспечить высокую точность в классификации редких кожных новообразований.

Редкие заболевания с точки зрения применения методов искусственного интеллекта представляют собой малочисленные классы в наборе данных, что приводит к значительному дисбалансу в распределении классов. Дисбаланс данных является естественной особенностью многих наборов данных и приводит к низким значениям метрик качества модели для редких классов. Тем не менее, существует ряд подходов для снижения ошибок при дисбалансе классов, хотя даже их применение не обеспечивает идеальную модель для классификации разнообразных изображений, таких, как медицинские снимки кожных новообразований. В связи с

этим, основной целью данного исследования является применение сверточных нейронных сетей для улучшения показателей метрик классификации редких кожных новообразований.

1. Методы и подходы. С каждым годом методы машинного и глубокого обучения играют все более значимую роль в классификации кожных поражений, обеспечивая точность и скорость автоматизированной диагностики. На сегодняшний день для классификации новообразований кожи известны следующие подходы:

1. Классические методы извлечения и классификации признаков. Среди ранних подходов выделяются методы получения гистограмм ориентированных градиентов (HOG), выявления локальных бинарных паттернов (LBP) и анализа цветовых гистограмм новообразований. Эти признаки классифицировались с помощью алгоритмов, таких, как метод опорных векторов (SVM), алгоритм ближайших соседей (KNN) и многослойные перцептроны (MLP). Несмотря на устаревание данных подходов и их ограниченную точность по сравнению с современными методами, они все еще встречаются в ряде современных исследований для решения специфических задач.
2. Применение сверточных нейронных сетей и трансферного обучения. С 2017 года, после появления крупных публичных соревнований, таких, как ISIC на платформе Kaggle, сверточные сети (CNN) с трансферным обучением стали доминировать в области анализа кожных новообразований. Эти сети, предобученные на больших наборах данных (например, ImageNet или других медицинских датасетах), показали высокие показатели метрик при решении медицинских задач с дообучением на специализированных наборах данных изображений кожных поражений.
3. Мультимодальные методы. Мультимодальные модели, которые объединяют различные типы данных, такие, как изображения, метаданные о пациенте и текстовые описания диагнозов, обеспечивают значительное улучшение качества классификации. Разработка мультимодальных подходов продолжается, и они часто демонстрируют лучшие показатели метрик в сравнении с классическими сверточными нейронными сетями. В мультимодальный подход входят и модели из данного исследования.
4. Методы на основе Vision Transformers (ViT). Введение Vision Transformers внесло значительный вклад в анализ изображений, так как они способны выявлять глобальные закономерности, анализируя зависимость между всеми частями изображения одновременно, а не только в пределах рецептивного поля, как это делают CNN. Благодаря этому, ViT более эффективно выявляют сложные паттерны, что особенно важно для диагностики кожных заболеваний.
5. Использование генеративных моделей. Генеративные модели GAN (Generative Adversarial Networks) и диффузионные модели находят применение для создания синтетических данных, особенно при ограниченности доступных изображений редких классов кожных новообразований. Генерация качественных синтетических изображений позволяет улучшить качество обучения моделей за счет увеличения объема данных и уменьшения дисбаланса классов.

Начальной точкой данного исследования является анализ данных, полученных с использованием метода мультиспектральной обработки изображений. Метод мультиспектральной обработки изображений предусматривает использование последовательностей изображений, где каждый элемент освещается группами диодов разных длин волн. Для получения таких последовательностей используется специальное устройство, разработанное в ИАП РАН под руководством Зайченко Кирилла Вадимовича.

Метод мультиспектральной обработки позволяет извлекать пространственную информацию о характере новообразования. Для этого осуществляется получение последовательности

монохромных изображений, на каждом из которых представлено распределение интенсивности светового излучения на определенной длине волны.

Таким образом, применение нескольких групп освещения с различными длинами волн позволяет получить дополнительную спектральную информацию, которая дополняет уже имеющуюся пространственную информацию.

Предполагается, что форма поражения является важным фактором, определяющим различия при классификации кожных новообразований. В данном исследовании не использовались данные, полученные с использованием метода мультиспектральной обработки изображений, поэтому было решено провести ряд экспериментов с целью улучшения показателей метрик в задаче многоклассовой классификации кожных новообразований с помощью сверточных нейронных сетей. Особое внимание уделяется улучшению показателей метрик классификации для редких кожных поражений, таких, как актинический кератоз (АК), базалиома (BKL), дерматофиброма (DF), сосудистые поражения (VASC) и сквамозно-клеточная карцинома (SCC).

1.1. Данные. В данном исследовании была поставлена задача многоклассовой классификации для дальнейшего анализа и обработки данных. Для проведения поставленных экспериментов были использованы наборы данных, представленные в таблице 1.

Таблица 1. Используемые наборы данных

Название датасета	Количество изображений	Задача	Приватный набор	Метаданные
ISIC-2017 [1], ISIC-2018 [2-3]	15414	Сегментация	-	-
ISIC-2019 [1-2, 4]	33579	Классификация	-	+
ISIC-2020 [5]	42373	Классификация	-	+
MED-NODE [6]	170	Классификация	-	-
SD-198 [7]	6582	Классификация	+	-
7-Point Criteria Database [8]	2019	Классификация	+	+
Light Field Image Dataset of Skin Lesions [9]	114	Классификация	+	-
PH ² [10-11]	200	Классификация	-	+
RAS Skin Dataset	43	Классификация	+	+

Кроме того, набор данных ISIC состоит из других датасетов, таких, как MSK [1], содержащий изображения различных размеров, HAM10000 [2] с изображениями размером 600x450 пикселей и BCN20000 [4] с изображениями размером 1024x1024 пикселей.

Для обучения state-of-the-art сегментационных сверточных нейронных сетей в данном исследовании использовались два датасета: ISIC-2017 [1] и ISIC-2018 [2-3]. Другие датасеты были использованы для обучения классификаторов. В процессе предобработки некоторых изображений выполнялось удаление черных областей, маркировок и других медицинских атрибутов. Кроме того, для предотвращения повторений при обучении использовалось перцептивное хеширование для удаления дубликатов изображений.

Полученный набор данных содержит метаданные пациентов и был сформирован по 9 классам, основываясь на классификации, представленной в датасете ISIC-2019 [1-2, 4]. Любые другие классы кожных поражений, лежащих вне набора классов из ISIC-2019, распределены по иерархической структуре классов заболевания.

Для приватного тестирования моделей была создана отдельная подвыборка из 5066 изображений. Подмножества выборки для обучения, валидации и тестирования были строго определены по индексам в генераторе данных с целью строгого контроля состава каждого подмножества.

Несмотря на относительно большой размер и разнообразие выборки, в ней проявляется значительный дисбаланс классов. Распределение классов в выборке представлено на диаграмме, показанной на рисунке 1.

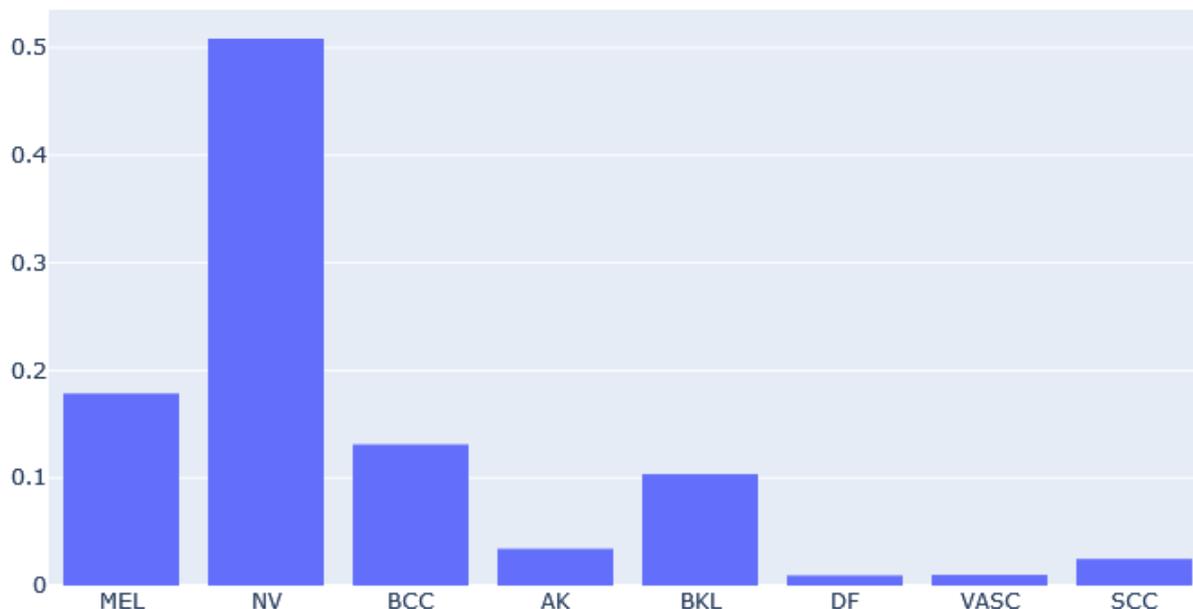


Рис. 1. Распределение классов кожных новообразований

На диаграмме представлены сокращенные названия классов новообразований. Ниже представлен список соответствия классов и их значимые черты:

- меланома (MEL) - участки темного цвета с частичной кровоточивостью;
- меланоцитарный невус (NV) - небольшой выпуклый участок коричневого цвета;
- базальноклеточная карцинома (BCC) - черные базальные клетки;
- актинический кератоз (AK) - участок с выраженным пигментом, покрытый чешуйками;
- кератоз, базалиома (BKL) - темное пигментированное пятно с выраженными краями;
- дерматофиброма (DF) - пятно различного спектра окраски, но преимущественно коричневого цвета;
- сосудистое поражение (VASC) - разрастание кровеносных сосудов;
- сквамозно-клеточная карцинома, спиналиома (SCC) - представляет собой язву, покрытую чешуйчатым налетом.

В списке выше приведено соответствие классов и их значимые особенности, представляющие классификацию из датасета ISIC. Все классы, не вошедшие в этот список и взятые из других датасетов, были распределены в соответствии с иерархической структурой кожных новообразований.

В процессе обучения сверточных нейронных сетей для обеспечения дополнительной регуляризации использовалась аугментация данных. Для тренировочной и валидационной выборок применялись следующие преобразования с определенными вероятностями:

- $p=1$: выполнялись изменение размера 224×224 и нормализация;
- $p=0.5$: применялась операция cutout (1-5) размером не более 16×16 пикселей;
- $p=0.3$: применялось одно из следующих преобразований: ElasticTransform, GridDistortion и OpticalDistortion;
- $p=0.25$: выполнялись повороты, сдвиги, вращения и масштабирования (на 90° , по горизонтали, по вертикали);
- $p=0.15$: применялось одно из следующих преобразований: motion blur и gaussian blur;
- $p=0.05$: выполнялось изменение яркости и контрастности.

Такие преобразования позволяют варьировать и дополнять исходные данные, а также улучшать обобщающую способность модели и значительно снижать вероятность переобучения, обеспечивая регуляризацию модели. Вероятности преобразований были выбраны экспериментально.

1.2. Предлагаемый подход. В данной работе был проведен анализ кожных новообразований с использованием комбинированной архитектуры, состоящей из сегментационных и классификационных моделей сверточных нейронных сетей. Целью комбинированной модели является извлечение дополнительной информации о форме новообразований для решения проблемы классификации редких новообразований. Для классификации кожных новообразований были предложены несколько вариантов комбинированных моделей:

1. Первая модель включает в себя параллельное объединение сегментационной сверточной нейронной сети, классификационной сверточной нейронной сети и полносвязной нейронной сети для учета метаданных пациента. Результаты сегментационной модели подаются через серию $MBConv$ блоков перед объединением, а полученные векторы признаков подаются в несколько линейных слоев обработки.
2. Вторая модель основана на совместном использовании сегментационной и классификационной сверточных нейронных сетей. Результаты сегментационной модели передаются как вход и выход классификационной модели. Аналогично первой модели, результирующая матрица сегментационной модели проходит через серию $MBConv$ блоков перед объединением с результатами классификационной модели, и полученные векторы признаков подаются в несколько линейных слоев обработки.
3. Третья модель реализует комбинированный подход, включающий некоторые свойства предыдущих моделей. Кроме того, в данном варианте добавляется большее количество линейных слоев для каждого выхода модели.

Такой вектор исследований связан с анализом данных, полученных с использованием метода мультиспектральной обработки изображений, при котором данные представляют собой последовательности изображений. Каждый элемент в этих последовательностях формируется путем освещения группами диодов с различными длинами волн. Такие последовательности были получены с использованием специальной камеры, разработанной под руководством Зайченко К. В. ИАП РАН. Благодаря данному методу можно рассматривать кожные новообразования в многомерном пространстве.

В представляемом исследовании данные, полученные с использованием метода мультиспектральной обработки, пока не были использованы, но основной упор сделан на подготовку базы для будущего включения таких данных. Форму новообразования предполагается рассматривать в качестве ключевого фактора.

Первая версия, используемая в исследовании, является базовой моделью, предназначенной для сравнения с последующими улучшениями. В последующих версиях модели была добавлена сеть для извлечения признаков при сегментации.

Во второй версии модели результат работы сверточной нейронной сети для сегментации, обработанный с использованием порога, добавлялся в качестве дополнительного 4-го канала к исходному изображению и подавался на вход классификатора.

Третья версия модели основана на параллельной концепции, в которой исходное изображение обрабатывается параллельно с использованием отдельных сегментационной и классификационной сверточных нейронных сетей.

Четвертая версия модели частично объединяет идеи предыдущих двух версий, используя комбинированный подход для обработки исходного изображения.

2. Обучение и оптимизация моделей. В ходе экспериментов и обучения моделей использовались дополнительные техники для предотвращения переобучения. В первую очередь

были применены L2-регуляризация для обеспечения устойчивости моделей и взвешивание классов с целью учета дисбаланса данных. Оптимизация параметров моделей проводилась путем оценки на кросс-валидации и на валидационной выборке для получения надежных результатов.

Собранный набор данных заранее разделен на тренировочную, валидационную и тестовую выборки. Каждому изображению был присвоен индекс, указывающий его принадлежность к определенной выборке. Размер приватной тестовой выборки составил 5066 образцов. Для каждого образца выполнялись операции нормализации и изменения размера до 224x224.

Для обучения всех моделей в данном исследовании применялись оптимизатор AdamW (с регуляризационным коэффициентом $\text{weight_decay} = 1e-4$) и функция потерь FocalLoss (с $\gamma=2$ и коэффициентом α получена на основе распределения классов). В дополнение к этому проводились эксперименты с использованием оптимизаторов SGD Нестерова с моментумом, равным 0.9 и Adam, в том числе с оптимизацией по кросс-энтропии и кастомными функциями потерь, основанных на кросс-энтропии.

В процессе исследования использовался сервис wandb для сохранения показателей метрик, производительности системы и артефактов модели. Модели оценивались и сохранялись каждые 5 эпох в локальном хранилище, а также один раз в процессе всего обучения в облачном хранилище. Обучение моделей осуществлялось с использованием следующих графических ускорителей: NVIDIA GeForce RTX 4090 (1 шт.), NVIDIA GeForce RTX 3090 Ti (2 шт.), NVIDIA GeForce RTX 3060 (1 шт.) и NVIDIA GeForce RTX 2080 (1 шт.).

2.1. Архитектура SkinNet-v1. Как было отмечено ранее, в качестве базовой модели была использована модель, состоящая из сверточной нейронной сети и полносвязной сети для анализа кожных поражений с учетом метаданных. Последняя представлена на рисунке 2.

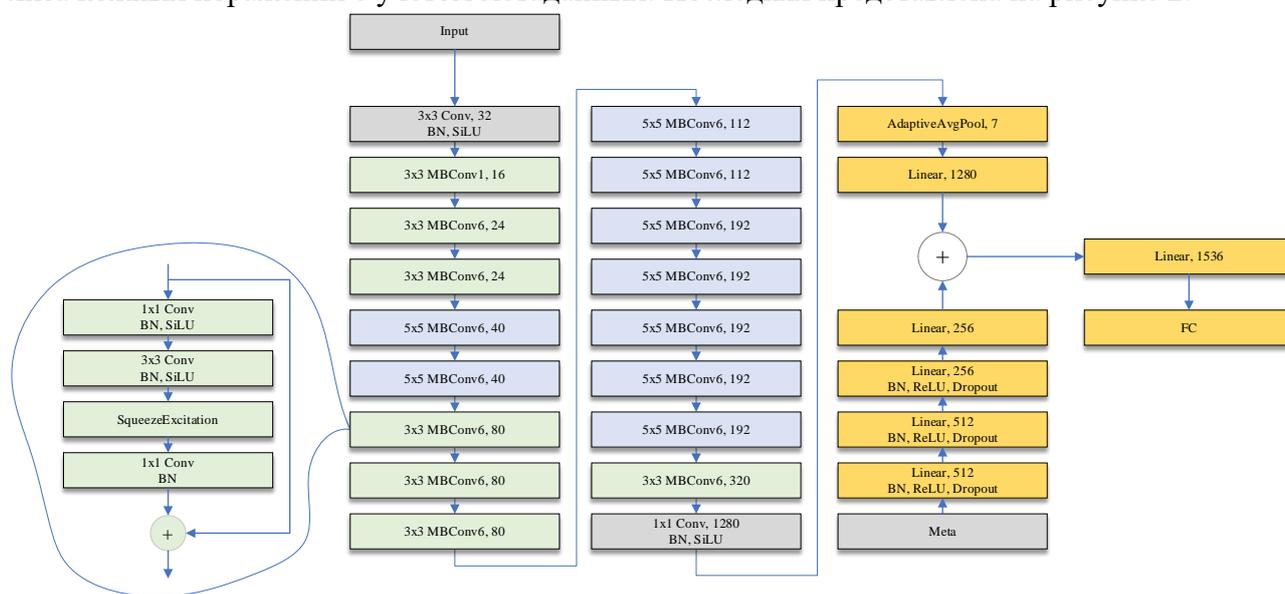


Рис. 2. Архитектура модели SkinNet-v1

В качестве архитектуры сверточной нейронной сети была выбрана модифицированная EfficientNet-B0, продемонстрировавшая высокую эффективность в предыдущих экспериментах и исследованиях, выполненных на аналогичных наборах данных [12]. Модель характеризуется устойчивыми и высокими результатами по ключевым метрикам на этапах кросс-валидации и тестовом наборе данных, а также низкими вычислительными затратами, что стало важным критерием при выборе, учитывая ограниченные вычислительные ресурсы. Помимо EfficientNet-B0, были также оценены такие архитектуры, как ResNet, ResNeXt-101, Inception-ResNet-v2, Inception-ResNet-v3, FixResNext-101, NoisyStudent, MobileNet, SENet154, ViT и ряд других.

После ее обучения веса сверточной нейронной сети были заморожены, затем производилось обучение полносвязной нейронной сети для получения признаков из метаданных. Каждый слой этой сети включает пакетную нормализацию, функцию активации ReLU, в первых трех слоях также применяется Dropout с вероятностью 0.2. Вектор признаков, полученный из нейронной сети по метаданным, объединялся с признаками, полученными из сверточной нейронной сети, и подавался в однослойную сеть с 1536 нейронами, результат которой используется для классификации.

2.2. Архитектура SkinNet-v2. Для решения задачи сегментации был исследован ряд архитектур. На основе метрик, полученных на тестовом наборе данных, и вычислительной сложности была выбрана архитектура Split-Attention U-Net, построенной по архитектурному принципу encoder-decoder. Данная сеть обладает высокими показателями метрик и отличается относительно низкой вычислительной сложностью за счет использования блока Split-Attention в первом слое. В блоке Split-Attention применяются операции свертки Depth-wise Convolution, Point-wise Convolution.

В дополнение к данной архитектуре были проведены эксперименты с несколькими моделями, включая U-Net, R2U-Net, Attention U-Net, Attention R2U-Net, TransUNet, TransCep-tion, TransFuse, TransDeepLab, MALUNet, DCSAU-NET и MISSFormer. На тестовом наборе данных наилучшие результаты по метрикам продемонстрировали MALUNet (mIoU = 0.79759) и MISSFormer (mIoU = 0.78715). Модель Split-Attention U-Net (mIoU = 0.77408) продемонстрировала сопоставимые показатели метрик, однако значительно превосходит другие архитектуры по скорости обучения благодаря последовательному применению свёрточных блоков Depth-wise и Point-wise. Операция Depth-wise свертки выполняется независимо для каждого входного канала, используя ядра размером 5x5, что уменьшает вычислительную сложность, так как применяет свертку только по глубине входного тензора, отдельно для каждого канала. Операция Point-wise свертки с использованием ядра размером 1x1 выполняется по всем пространственным размерам входного тензора после Depth-wise свертки для агрегации признаков с разных каналов входного тензора.

Каждый последующий encoder-блок завершается операцией max pooling размером 2x2, в то время как каждый decoder-блок начинается с операции upsampling, которая последовательно восстанавливает оригинальный размер входного фрейма. Skip connections применяются для объединения признаков (feature maps) из блока декодера с соответствующими признаками из блока энкодера, что способствует интеграции низкоуровневой и высокоуровневой информации. Завершающий слой модели представляет собой свертку 1x1, что позволяет формировать бинарные или многоклассовые маски для сегментации кожных новообразований.

На рисунке 3 сегментационная сеть обозначена красной областью. Входное изображение подается на вход сегментационной сети. Затем результат конкатенируется с входным изображением в виде четвертого канала и подается на вход классификатора SkinNet-v1. При объединении к результату сегментационной сверточной нейронной сети применяется пороговая функция, которая формирует бинарную маску для кожного поражения.

По показателям метрик данная модель работает хуже, однако количество неправильно предсказанных редких поражений уменьшилось на 4%.

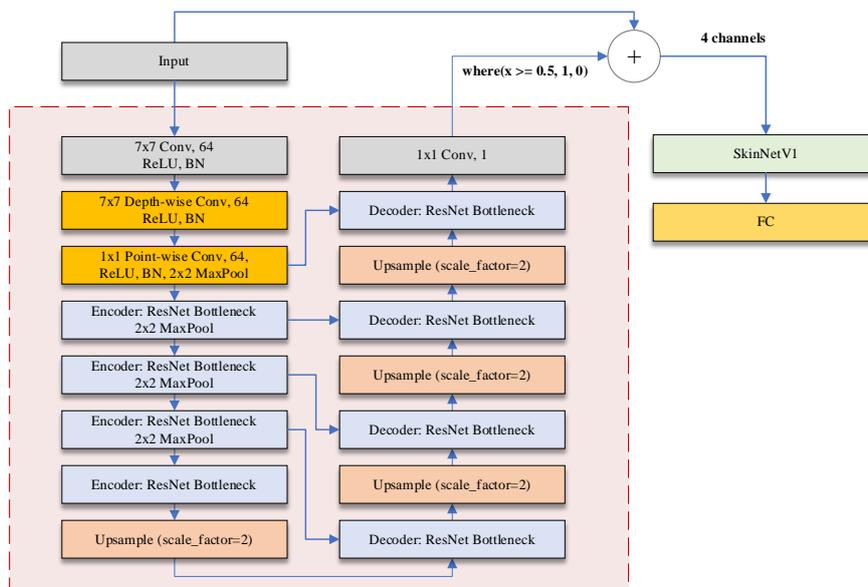


Рис. 3. Архитектура модели SkinNet-v2

2.3. Архитектура SkinNet-v3. На рисунке 4 представлена третья версия модели, отличительной особенностью которой является внедрение небольшой сверточной сети для обработки вывода сегментационной сети, обозначенной фиолетовой областью.

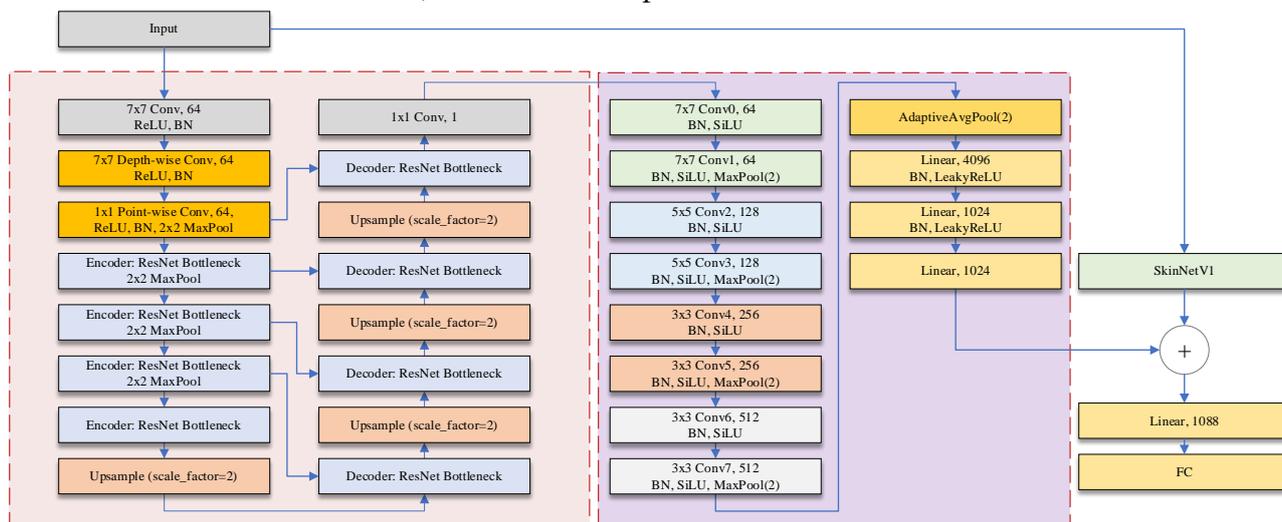


Рис. 4. Архитектура модели SkinNet-v3

Исходное входное изображение подается на вход классификационной сети, а результат конкатенируется с 64 признаками из блока сегментационной сети. Метрики данной модели незначительно, но продемонстрировали улучшение по сравнению с предыдущими моделями. Результат сегментационной сети, обозначенной светло-красным областью, в данном случае представляет из себя изображение поражения, которое было выделено с использованием бинарной маски, аналогично семантической сегментации.

Классификационная сверточная сеть, предназначенная для обработки вывода сегментационной сети, состоит из двух сверток 7x7, двух сверток 5x5, четырех сверток 3x3 и нескольких линейных слоев. После каждого сверточного слоя применяется операция MaxPool 2x2. Результат сети, обозначенной фиолетовой областью, является вектором признаков размерности 1024. Исходное входное изображение подается на вход классификационной сети, а результат конкатенируется с 64 признаками из блока сегментационной сети. Показатели метрик данной модели незначительно, но продемонстрировали улучшение по сравнению с предыдущими моделями.

2.4. Архитектура SkinNet-v4. Данная модель является наиболее сложной в плане вычислительной мощности в серии проведенных экспериментов. С точки зрения концепции, она является комбинацией второй и третьей версий моделей (рис.5).

Однако вместо дублирования модели первой версии для обработки результатов сегментации была добавлена более модифицированная сверточная сеть на основе EfficientNet-B3. Было замечено, что метрики улучшились в среднем на 1% по сравнению с базовой версией.

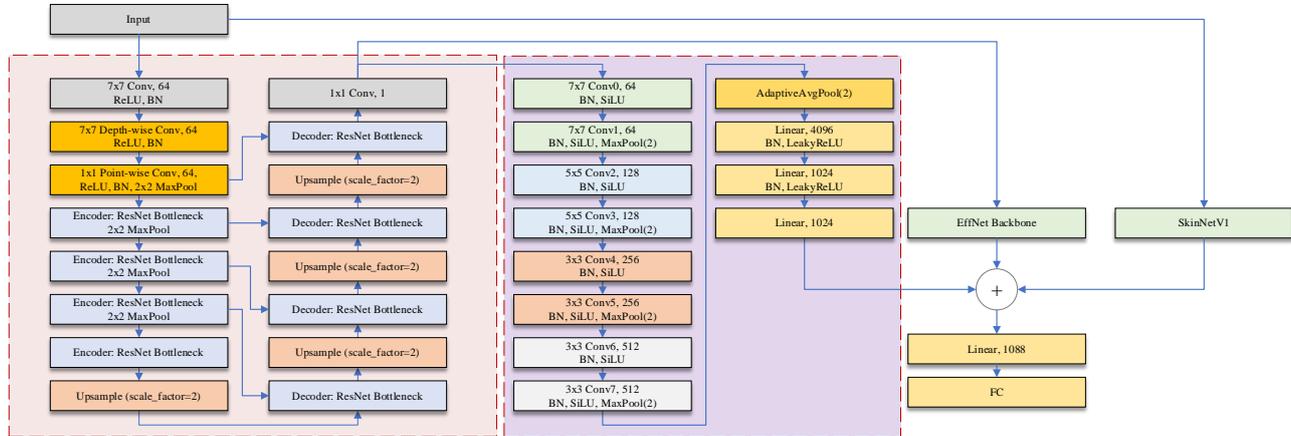


Рис. 5. Архитектура модели SkinNet-v4

3. Оценка моделей. Стоит отметить, что в данном исследовании, помимо тренировочной и валидационной выборок, была сформирована тестовая выборка, состоящая из 5066 изображений, которая моделирует оценку процесса эксплуатации моделей на новых данных. Тестовая выборка аналогичным образом строго предопределялась по индексам. В задаче классификации оценка моделей выполнялась с использованием матрицы ошибок и взвешенных метрик Recall, Precision и F1-score, а также обычных метрик Specificity, Recall, Precision и F1-score. Для лучшего понимания внесенных изменений в архитектуру моделей в результатах приведены расширенные расчеты метрик для каждого из классов. Для оценки сегментационных моделей использовались метрики IoU и Dice, результаты которых лежат вне рамок данного исследования.

Модели, в том числе и ансамбли, были оценены на тестовом наборе данных, а также с использованием стратифицированной кросс-валидации. Кроме того, в данном исследовании особое внимание уделялось оценке суммарных ошибок первого и второго рода на основе матрицы ошибок классификации как для всех типов новообразований, так и для редких кожных поражений (AK, BKL, DF, VASC, SCC).

Результаты модели SkinNet-v1 представлены в таблице 2. Суммарное значение FP+FN для всех новообразований составляет 655, в то время как для редких поражений это значение равно 199. Эти значения будут использоваться в качестве отправной точки для дальнейшего анализа. Результаты модели SkinNet-v2 представлены в таблице 3. Суммарное значение FP+FN для всех типов новообразований составляет 676, что на 3% меньше по сравнению с моделью SkinNet-v1. Для редких поражений значение FP+FN составляет 191. Несмотря на то, что устойчивая к дисбалансу классов метрика Recall улучшилась, взвешенные метрики показывают обратную тенденцию.

Таблица 2. Показатели метрик модели SkinNet-v1

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9586	0.8374	0.8067	0.8217	0.1433	0.1381	0.1406
NV	0.9359	0.9328	0.9410	0.9368	0.4876	0.4918	0.4897
BCC	0.9889	0.8294	0.9263	0.8752	0.1190	0.1329	0.1256
AK	0.9851	0.5789	0.5756	0.5773	0.0195	0.0194	0.0195

BKL	0.9673	0.8004	0.7028	0.7484	0.0705	0.0619	0.0659
DF	0.9982	0.7826	0.7826	0.7826	0.0071	0.0071	0.0071
VASC	0.9988	0.8491	0.8824	0.8654	0.0089	0.0092	0.0091
SCC	0.9923	0.6944	0.6637	0.6787	0.0148	0.0141	0.0145
Total	0.9781	0.7881	0.7851	0.7858	0.8707	0.8746	0.8719

Таблица 3. Показатели метрик модели SkinNet-v2

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9554	0.8186	0.7922	0.8052	0.1407	0.1362	0.1384
NV	0.9293	0.9347	0.9354	0.9350	0.4886	0.4889	0.4887
BCC	0.9871	0.8248	0.9137	0.8670	0.1171	0.1297	0.1230
AK	0.9849	0.6045	0.5912	0.5978	0.0211	0.0207	0.0209
BKL	0.9710	0.7652	0.7167	0.7402	0.0669	0.0627	0.0647
DF	0.9976	0.8261	0.7451	0.7835	0.0075	0.0068	0.0071
VASC	0.9978	0.8679	0.8070	0.8364	0.0091	0.0084	0.0087
SCC	0.9939	0.7248	0.7248	0.7248	0.0156	0.0156	0.0156
Total	0.9771	0.7958	0.7783	0.7862	0.8666	0.8689	0.8673

Результаты модели SkinNet-v3 представлены в таблице 4. Суммарное значение FP+FN для всех типов новообразований составляет 625 (улучшение на 7,5% по сравнению с предыдущей моделью), а для редких поражений – 164.

Таблица 4. Показатели метрик модели SkinNet-v3

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9562	0.8053	0.7916	0.7984	0.1380	0.1356	0.1368
NV	0.9321	0.9490	0.9388	0.9439	0.4962	0.4909	0.4935
BCC	0.9894	0.8456	0.9297	0.8857	0.1200	0.1319	0.1257
AK	0.9877	0.5954	0.6319	0.6131	0.0203	0.0216	0.0209
BKL	0.9762	0.8000	0.7660	0.7826	0.0711	0.0680	0.0695
DF	0.9968	0.8043	0.6981	0.7475	0.0073	0.0063	0.0068
VASC	0.9980	0.7547	0.8000	0.7767	0.0079	0.0084	0.0081
SCC	0.9929	0.7407	0.6957	0.7175	0.0158	0.0148	0.0153
Total	0.9787	0.7869	0.7815	0.7832	0.8766	0.8776	0.8767

При сравнении первых трех моделей по метрике Recall можно предположить, что вторая модель является наиболее точной, в то время как третья модель имеет наименьшее значение данной метрики. Однако при расчете взвешенных метрик становится ясно, что лучшей моделью является SkinNet-v3. Если сравнивать модели по сумме ошибок первого и второго рода, можно сделать вывод, что третья модель имеет наименьшее количество ошибок, как для всех типов новообразований, так и для редких поражений.

Результаты модели SkinNet-v4 представлены в таблице 5. Суммарное значение FP+FN для всех типов новообразований составляет 561, что является улучшением на 10% по сравнению с предыдущей моделью, в то время как для редких поражений значение FP+FN составляет 130.

Таблица 5. Показатели метрик модели SkinNet-v4

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9612	0.8088	0.8116	0.8102	0.1386	0.1391	0.1388
NV	0.9368	0.9539	0.9429	0.9484	0.4986	0.4928	0.4957
BCC	0.9894	0.8695	0.9323	0.8998	0.1250	0.1340	0.1293
AK	0.9886	0.6491	0.6647	0.6568	0.0219	0.0224	0.0222
BKL	0.9766	0.8184	0.7717	0.7943	0.0720	0.0679	0.0699

DF	0.9988	0.8696	0.8696	0.8696	0.0079	0.0079	0.0079
VASC	0.9992	0.8491	0.9000	0.8738	0.0089	0.0094	0.0091
SCC	0.9952	0.7870	0.7798	0.7834	0.0168	0.0166	0.0167
Total	0.9807	0.8257	0.8341	0.8295	0.8897	0.8902	0.8897

В ходе экспериментов подобных результатов удалось достичь путем параллельного использования модели EffcientNet-V3 в качестве второго классификатора и с модификацией некоторых сверточных слоев с их последующим дообучением.

4. Результаты. В таблице 6 представлены показатели метрик, полученные разработанными моделями на тестовом наборе данных. Напомним, что основной целью нашего исследования является улучшение показателей метрик при классификации редких новообразований.

Таблица 6. Общие показатели метрик

Model	All: FP+FN	Rares: FP+FN	Specificity	WRe-call	WPrecision	WF1-score
SkinNet-v1	655	199	0.9781	0.8707	0.8746	0.8719
SkinNet-v2	676	191	0.9771	0.8666	0.8689	0.8673
SkinNet-v3	625	164	0.9787	0.8766	0.8776	0.8767
SkinNet-v4	561	130	0.9807	0.8897	0.8902	0.8897

Согласно результатам, наиболее эффективной для достижения этой цели оказалась четвертая версия модели. Данная модель не только снизила количество неверных классификаций редких новообразований, но также привела к значительному снижению числа ошибок для всех классов и, как следствие, к улучшению всех представленных метрик.

Из таблицы видно, что каждая последующая версия модели демонстрировала лучший потенциал в анализе редких новообразований. Например, SkinNet-v4 показывает лучшие показатели метрик с точки зрения Specificity и взвешенных Recall, Precision и F1-score. Кроме того, данная модель характеризуется наименьшим количеством ложноположительных и ложноотрицательных результатов как для всех классов, так и для редких образцов.

Заключение. В данном исследовании был представлен метод использования комбинированной модели с расширенным извлечением дополнительной информации о форме кожных новообразований с целью решения проблемы анализа редких новообразований. Тем самым, это позволяет более точно классифицировать эти редкие кожные поражения.

Был проведен ряд экспериментов, направленных на минимизацию ошибок первого и второго рода для редких кожных заболеваний (AK, BKL, DF, VASC, SCC).

На основе этих экспериментов были сформулированы четыре версии модели, приведенные в данной статье.

В дальнейших исследованиях планируется:

- проверить гипотезу исследования в других медицинских областях;
- оценить влияние генерации изображений редких классов кожных поражений с помощью модели на основе генеративно-состязательных сетей и операции MixUp на показатели метрик классификации базовой модели;
- проверить гипотезу исследования на лабораторных крысах с искусственно выведенными кожными новообразованиями.

Все это позволит улучшить показатели метрик моделей для классификации медицинских заболеваний, обученных на данных с сильным дисбалансом данных.

Благодарности. Работа была поддержана Минобрнауки РФ, госзадание № 075-01157-23-00, тема № FZZM-2022-0011.

Список источников

1. Codella N. et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI). IEEE 15th International symposium on biomedical imaging (ISBI 2018), 2018, pp. 168-172, DOI: 10.1109/ISBI.2018.8363547
2. Tschandl P., Rosendahl C., Kittler H. The HAM10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161, 2018, DOI: 10.1038/sdata.2018.161.
3. Codella N., Rotemberg V., Tschandl P. et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC), 2018, arXiv: 1902.03368 [cs.CV]
4. Combalia M., Codella N., Rotemberg V., et al. BCN20000: Dermoscopic lesions in the wild, arXiv:1908.02288, 2019, DOI: 10.48550/arXiv.1908.02288
5. Rotemberg V., Kurtansky N., Betz-Stablein B., et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34. 2021, DOI: 10.1038/s41597-021-00815-z
6. Giotis I., Molders N., Land S., Biehl M., Jonkman M.F., Petkov N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 2015, vol. 42, pp. 6578-6585.
7. Sun X., Yang J., Sun M., Wang K. A benchmark for automatic visual classification of clinical skin disease images. ECCV: Springer, 2016, vol. 9910, pp. 206–222.
8. Kawahara J., Daneshvar S., Argenziano G., Hamarneh G. Seven-point checklist and skin lesion classification using multi-task multi-modal neural nets. IEEE Journal of biomedical and health informatics (IEEE JBHI) special issue on Skin lesion image analysis for Melanoma Detectio, 2019, vol. 23, no. 2, pp. 538-546, DOI: 10.1109/JBHI.2018.2824327.
9. Faria S.M.M., et al. Light field image dataset of skin lesions. IEEE 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2019, pp. 3905–3908, DOI:10.1109/EMBC.2019.8856578.
10. Barata C., Ruela M., et al. Two systems for the detection of melanomas in Dermoscopy images using texture and color features. IEEE systems journal, 2013, 8(3), pp. 965–979.
11. Mendonça T., Ferreira P. M., Marques J., Marcal A., Rozeira J. PH² - A dermoscopic image database for research and benchmarking. IEEE 35th international conference of the IEEE engineering in medicine and biology society, Osaka, Japan, 2013, DOI:10.1109/EMBC.2013.6610779.
12. Milantev S., Bessmertnyi I., et al. Skin lesion analysis using ensemble of CNN with dermoscopic images and metadata. CEUR workshop proceedings, 2021, vol. 2893.

Милантьев Сергей Андреевич. Аспирант факультета Программной инженерии и компьютерной техники университета ИТМО, младший научный сотрудник ИАП РАН. AuthorID: 1176348, SPIN: 5509-8366, ORCID: 0000-0002-1970-5217, nevdev@niuitmo.ru, Россия, Санкт-Петербург, Кронверкский проспект, 49

Святкина Виталия Игоревна. Младший научный сотрудник ИАП РАН, ORCID: 0000-0002-6456-8776, svyatkina.vi@edu.spbstu.ru, Россия, Санкт-Петербург, Ивана Черных, 31-33

Бессмертный Игорь Александрович. Профессор, доктор технических наук, университет ИТМО. AuthorID: 571032, SPIN: 4329-6492, ORCID: 0000-0001-6711-6399, bessmertny@itmo.ru, Россия, Санкт-Петербург, Кронверкский проспект, 49

Зайченко Кирилл Вадимович. Профессор, доктор технических наук, зав. лаб., главный научный сотрудник ИАП РАН. AuthorID: 107276, SPIN: 7951-3906, ORCID: 0000-0002-2881-4386, Россия, Санкт-Петербург, Ивана Черных, 31-33

UDC: 004.93.12

DOI:10.25729/ESI.2024.35.3.004

Image analysis of skin lesion using a combined convolutional neural network architecture

Sergey A. Milantev^{1,2}, Vitaliia I. Sviatkina², Igor A. Bessmertny¹, Kirill V. Zaichenko²

¹ITMO University, Russia, Saint Petersburg, nevdev@niuitmo.ru

²IAI RAS, Russia, Saint Petersburg

Abstract. This research explores the possibility of applying combined convolutional neural network architectures to analyze skin lesions. Model architectures have been developed to extract additional features related to the shape pattern of skin lesions. Optimization of the models, including the architecture, was performed in order to minimize

I and II types of errors for rare skin lesions. ISIC2017-2020, MED-NODE, SD-198, 7-Point Criteria Database, Light Field Image Dataset of Skin Lesions, PH2, IAP RAS were used in the training process. AdamW optimizer, FocalLoss functions and CosineAnnealingWarmRestarts scheduler were used to train classification models. The BCEDice loss function was used to train the segmentation models. The models were evaluated using weighted classification metrics such as Specificity, Recall, Precision and F1-score. The robustness of model architecture was considered during the validation phase. Models which are using additional convolutional neural networks for the skin lesion extraction shape features showed better metrics performance and also had lower sum of I and II type errors for rare lesions compared to conventional classification models. The results of this research can be used in analyzing medical problems with data imbalance in the training dataset.

Keywords: skin lesions, convolutional neural networks, skin lesion analysis, class imbalance, multispectral image processing

Acknowledgements: This work was supported by the Ministry of Science and Higher Education of the Russian Federation, state task no. 075-01157-23-00, topic no. FZZM-2022-0011.

References

1. Codella N. et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI). IEEE 15th International symposium on biomedical imaging (ISBI 2018), 2018, pp. 168-172, DOI: 10.1109/ISBI.2018.8363547
2. Tschandl P., Rosendahl C., Kittler H. The HAM10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161, 2018, DOI: 10.1038/sdata.2018.161.
3. Codella N., Rotemberg V., Tschandl P. et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC), 2018, arXiv: 1902.03368 [cs.CV]
4. Combalia M., Codella N., Rotemberg V., et al. BCN20000: Dermoscopic lesions in the wild, arXiv:1908.02288, 2019, DOI: 10.48550/arXiv.1908.02288
5. Rotemberg V., Kurtansky N., Betz-Stablein B., et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34. 2021, DOI: 10.1038/s41597-021-00815-z
6. Giotis I., Molders N., Land S., Biehl M., Jonkman M.F., Petkov N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 2015, vol. 42, pp. 6578-6585.
7. Sun X., Yang J., Sun M., Wang K. A benchmark for automatic visual classification of clinical skin disease images. ECCV: Springer, 2016, vol. 9910, pp. 206–222.
8. Kawahara J., Daneshvar S., Argenziano G., Hamarneh G. Seven-point checklist and skin lesion classification using multi-task multi-modal neural nets. IEEE Journal of biomedical and health informatics (IEEE JBHI) special issue on Skin lesion image analysis for Melanoma Detectio, 2019, vol. 23, no. 2, pp. 538-546, DOI: 10.1109/JBHI.2018.2824327.
9. Faria S.M.M., et al. Light field image dataset of skin lesions. IEEE 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2019, pp. 3905–3908, DOI:10.1109/EMBC.2019.8856578.
10. Barata C., Ruela M., et al. Two systems for the detection of melanomas in Dermoscopy images using texture and color features. IEEE systems journal, 2013, 8(3), pp. 965–979.
11. Mendonça T., Ferreira P. M., Marques J., Marcal A., Rozeira J. PH² - A dermoscopic image database for research and benchmarking. IEEE 35th international conference of the IEEE engineering in medicine and biology society, Osaka, Japan, 2013, DOI:10.1109/EMBC.2013.6610779.
12. Milantev S., Bessmertnyi I., et al. Skin lesion analysis using ensemble of CNN with dermoscopic images and metadata. CEUR workshop proceedings, 2021, vol. 2893.

Milantev Sergey Andreevich. PhD student of the faculty of software engineering and computer systems, ITMO university, junior researcher at IAI RAS, AuthorID: 1176348, SPIN: 5509-8366, ORCID: 0000-0002-1970-5217, nevdev@niuitmo.ru, Russia, St. Petersburg, Kronverksky Prospect, 49

Sviatkina Vitaliia Igorevna. Junior researcher at IAI RAS, ORCID: 0000-0002-6456-8776, sviatkina.vi@edu.spbstu.ru, Russia, St. Petersburg, Ivan Chernykh, 31-33

Bessmertny Igor Alexandrovich. Professor, Doctor of technical science, AuthorID: 571032, SPIN: 4329-6492, ORCID: 0000-0001-6711-6399, bessmertny@itmo.ru, Russia, St. Petersburg, Kronverksky Prospect, 49

Zaichenko Kirill Vadimovich. Professor, doctor of technical sciences, head of the laboratory at IAI RAS, AuthorID: 107276, SPIN: 7951-3906, ORCID: 0000-0002-2881-4386, Russia, St. Petersburg, Ivan Chernykh, 31-33

Статья поступила в редакцию 18.10.2023; одобрена после рецензирования 11.10.2024; принята к публикации 11.10.2024.

The article was submitted 10/18/2023; approved after reviewing 10/11/2024; accepted for publication 10/11/2024.