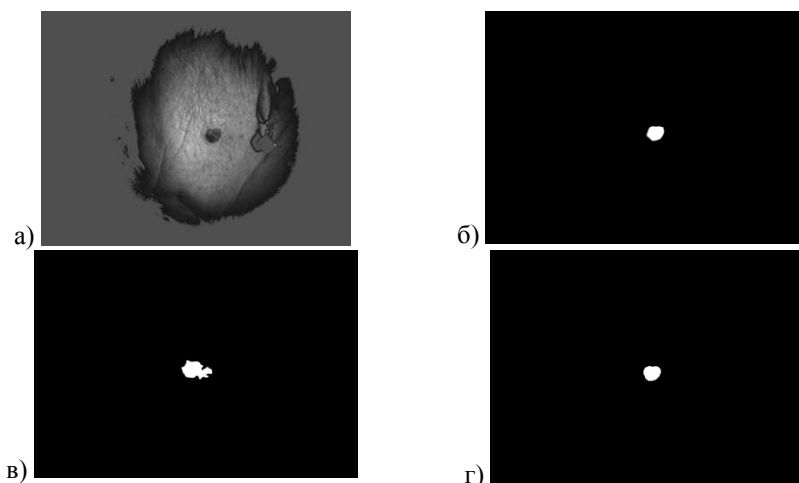




а) оригинал изображения, прошедший предобработку; б) результат Оцу;
в) результат U-Net; г) эталон

Рисунок 1 – Результаты сегментации тестового изображения

Таблица 1 – Сравнение эффективности методов сегментации на тестовой выборке

Метод сегментации	Средний Dice Score	Особенности
Метод Оцу	0,56	Требует высокой контрастности
Нейронная сеть U-Net	0,42	Склонна к переобучению на малых выборках
Гибридный подход	0,53	Пересечение масок Оцу и U-Net

Список использованных источников

1.Huang, H.-Y. A Review of recent advances in Computer-Aided Detection Methods Using Hyperspectral Imaging Engineering to detect skin cancer / H.-Y. Huang, Y.-P. Hsiao, R. Karmakar [et al.] // Cancers. – 2023. – Vol. 15, № 23. – P. 5634.

2.Лукашик, Д. В. Анализ современных методов сегментации изображений / Д. В. Лукашик // Экономика и качество систем связи. – 2022. – № 2. – С. 57-65.

Неповиннов Артём Павлович, студ. гр. 6304-010302D, nepovinnov.ap@ssau.ru.
Матвеева Ирина Александровна, к.т.н., доцент каф. ЛБС, matveeva.ia@ssau.ru.

УДК 004.891.3:543.424.2:616-71

СИНТЕЗ РАМАНОВСКИХ СПЕКТРОВ С ПОМОЩЬЮ МЕТОДА РАЗРЕШЕНИЯ МНОГОМЕРНЫХ КРИВЫХ

М.С. Павлова, И.А. Матвеева

«Самарский национальный исследовательский университет имени
академика С.П. Королёва», г. Самара

Ключевые слова: рамановская спектроскопия, многомерное разрешение кривых, синтез спектров, Gaussian Mixture Model, машинное обучение, бинарная классификация, новообразования кожи.

Рамановская спектроскопия является перспективным неинвазивным методом анализа биологических тканей, позволяющим получать информацию о молекулярном составе исследуемых объектов. Изменения в рамановских спектрах могут отражать патологические процессы и использоваться для диагностики новообразований кожи [1]. Однако применение методов машинного обучения часто ограничено небольшим объемом экспериментальных данных, что делает актуальной разработку методов синтеза рамановских спектров для расширения обучающих выборок.

В работе использовались рамановские спектры кожи, полученные *in vivo* при исследовании новообразований кожи, подробно описанных в [2]. Для синтеза применяется разложение рамановских спектров методом разрешения многомерных кривых (MCR). Число компонент определялось на основе анализа сингулярных значений. Рассматривались разложения на 8 компонент для интерпретации спектров и на 30 компонент для более точного описания данных и синтеза спектров. Для каждой группы спектров (здоровая кожа и новообразования кожи) строилась статистическая модель распределения концентраций компонент с использованием Gaussian Mixture Model. На ее основе генерировались новые значения концентраций, после чего синтетические спектры восстанавливались как линейная комбинация спектральных профилей компонент. Для повышения реалистичности синтетических данных дополнительно использовался шум, полученный из остатков MCR-разложения. Примеры синтетических спектров представлены на рисунке 1.

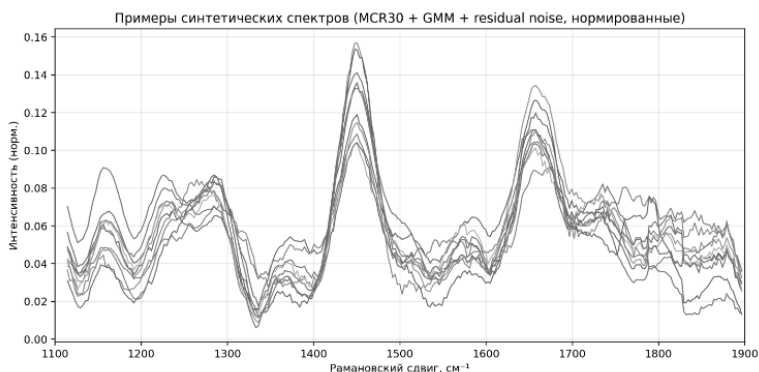
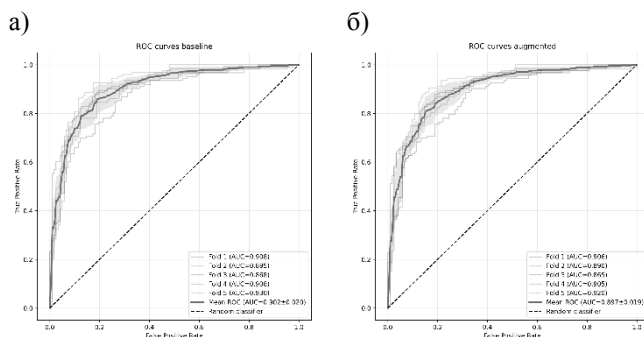


Рисунок 1 – Примеры синтетических рамановских спектров, восстановленных на основе MCR-компонент ($k = 30$)

Для оценки диагностической информативности спектров была решена задача бинарной классификации рамановских спектров кожи на два класса:

здоровая кожа и новообразования кожи. В качестве классификатора использовалась логистическая регрессия. Качество классификации оценивалось методом стратифицированной пятикратной кросс-валидации по метрикам ROC AUC, ассигасу, F1-score, чувствительности и специфичности. Среднее значение ROC AUC составило около 0,90, а значения ассигасу и F1-score — около 0,83. Влияние синтетических спектров на качество классификации исследовалось путем добавления синтетических данных только в обучающую выборку каждого фолда кросс-валидации. ROC-кривые для классификации без использования синтетических данных и при их добавлении приведены на рисунке 2.

Сравнение количественных метрик классификации приведено в таблице 1. Полученные результаты показывают, что использование синтетических спектров не приводит к заметному улучшению или ухудшению качества классификации по сравнению с обучением только на экспериментальных данных, что подтверждает безопасность синтеза.



а) без использования синтетических данных; б) с использованием синтетических спектров
Рисунок 2 – ROC-кривые модели бинарной классификации рамановских спектров кожи

Таблица 1 – Сравнение метрик бинарной классификации рамановских спектров

Режим обучения	ROC AUC	Accuracy	F1-score	Sensitivity	Specificity
Без синтетических данных	0,90 ± 0,02	0,83 ± 0,03	0,83 ± 0,03	0,83 ± 0,03	0,84 ± 0,04
С использованием синтетических спектров	0,90 ± 0,02	0,83 ± 0,03	0,83 ± 0,03	0,83 ± 0,02	0,83 ± 0,05

В работе реализован метод синтеза рамановских спектров кожи на основе разложения MCR и статистического моделирования распределения концентраций компонент. Показано, что полученные синтетические спектры воспроизводят основные характеристики экспериментальных

данных и могут использоваться для расширения обучающих выборок при анализе рамановских спектров кожи методами машинного обучения.

Список использованных источников

1.Khristoforova, Y. Raman-based techniques in medical applications for diagnostic tasks: a review / Y. Khristoforova, L. Bratchenko, I. Bratchenko //International Journal of Molecular Sciences. – 2023. – Vol. 24, № 21. – P. 15605.

2.Multidimensional Analysis of Dermoscopic Images and Spectral Information for the Diagnosis of Skin Tumors / I. A. Matveeva, A. I. Komlev, O. I. Kaganov [et al.] // Journal of Biomedical Photonics and Engineering 2024. – Vol. 10, № 1. – P. 010307.

Павлова Мария Сергеевна, студ. гр. 6403-010302D, mariaparamonova@mail.ru.

Матвеева Ирина Александровна, к.т.н., доцент каф. ЛБС, matveeva.ia@ssau.ru.

УДК 004.891.3:543.424.2:616-71

ФОРМИРОВАНИЕ ПРИЗНАКОВ ДЛЯ ЗАДАЧИ ИДЕНТИФИКАЦИИ ХРОНИЧЕСКОЙ СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТИ

В.В. Новикова

«Самарский национальный исследовательский университет имени академика С.П. Королёва», г. Самара

Ключевые слова: рамановская спектроскопия, хроническая сердечная недостаточность, машинное обучение, многомерное разрешение кривых, классификация, формирование признаков

Диагностика хронической сердечной недостаточности (ХСН) на ранних стадиях осложнена отсутствием уникальных симптомов, а анализ на стандартные маркеры (NT-proBNP) дорогостоящий и требует большого объема крови. Перспективной альтернативой является метод рамановской спектроскопии, позволяющий анализировать состав крови даже по одной капле [1]. Однако из-за сложности и визуальной схожести спектров возникает необходимость в разработке алгоритмов автоматизированного извлечения признаков и классификации [2].

В работе использованы рамановские спектры сыворотки и плазмы крови 236 пациентов (199 с ХСН и 37 человек группы контроля). Измерения выполнялись на спектрометрическом комплексе EnSpectr R785, интегрированным с микроскопом ADF U300. Для удаления флуоресцентного фона и фильтрации шумов реализован алгоритм, включающий коррекцию базовой линии (алгоритм arPLS), сглаживание фильтром Савицкого-Голея и векторную нормализацию. На рисунке 1 представлен результат работы алгоритма.

Основным методом формирования признаков выбрано разложение спектральной матрицы на компоненты методом разрешения многомерных