

applications [Текст] / R. S. Das, Y. K. Agrawal // *Vibrational Spectroscopy*. – 2011. – Vol. 57. – P. 163–176.

2.Chen, S. Recovery of Raman spectra with low signal-to-noise ratio using Wiener estimation [Текст] / S. Chen, X. Lin, C. Yuen [и др.] // *Optics Express*. – 2014. – Vol. 22, № 10. – P. 12102–12114. — URL: <https://opg.optica.org/oe/fulltext.cfm?uri=oe-22-10-12102>

Мельникова Ангелина Андреевна, студ. гр. 6402-120304D, gelik.work@mail.ru.
Христофорова Юлия Александровна, к.физ.-мат.н., доцент каф. ЛБС khristoforova.yua@ssau.ru.

УДК 543.424.2; 004.93

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ СПЕКТРАЛЬНЫХ ПРИЗНАКОВ НОВООБРАЗОВАНИЙ КОЖИ

Е.О. Сальникова, Ю.А. Христофорова

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

Ключевые слова: метод ANOVA, линейный дискриминантный анализ (LDA), метод PLS-DA, кросс-валидация (K-fold), метод ALS, *ex vivo* спектры рамановского рассеяния кожи, рак кожи.

Рак кожи является одним из наиболее распространенных онкологических заболеваний в мире, представляя серьезную медико-социальную проблему. В структуре заболеваемости преобладают немеланомные формы – базальноклеточный и плоскоклеточный рак, однако наибольшую опасность представляет меланома, характеризующаяся высокой агрессивностью и метастатическим потенциалом. Традиционным «золотым стандартом» диагностики остается инвазивная биопсия с последующим гистологическим исследованием [1]. Несмотря на высокую точность, этот метод сопряжен с рисками для пациента, такими как рубцевание, повреждение нервов и потенциальное распространение опухолевых клеток. Кроме того, низкая специфичность отбора новообразований для биопсии приводит к значительному числу неоправданных хирургических вмешательств, поскольку большинство удаляемых образований оказываются доброкачественными [2]. В связи с этим существует необходимость в разработке методов предварительной диагностики рака кожи.

Целью настоящей работы является построение бинарных классификаторов для анализа рамановских спектров кожи с помощью алгоритмов машинного обучения для установления возможности предварительного определения рака кожи.

Для реализации поставленной задачи были использованы *ex vivo* спектры кожи пациентов с гистологически подтвержденными диагнозами злокачественных и доброкачественных кожных новообразований. Спектры были сняты с помощью портативного прибора с длиной волны

зондирующего излучения 785 нм как в месте новообразования, так и в местах непораженной кожи для каждого пациента. Все спектры в диапазоне от 265 до 2000 см^{-1} были подвергнуты предварительной обработке по следующему алгоритму: сглаживание с помощью фильтра Савицкого-Голея с шириной окна 15 и 2-ым порядком полинома, нормализация по среднеквадратичному отклонению. Для сравнения качества классификации в зависимости от предварительной обработки был добавлен этап удаления базовой линии методом чередующихся наименьших квадратов (ALS). На рисунке 1 представлены все спектры после предобработки в первом и втором случае соответственно.

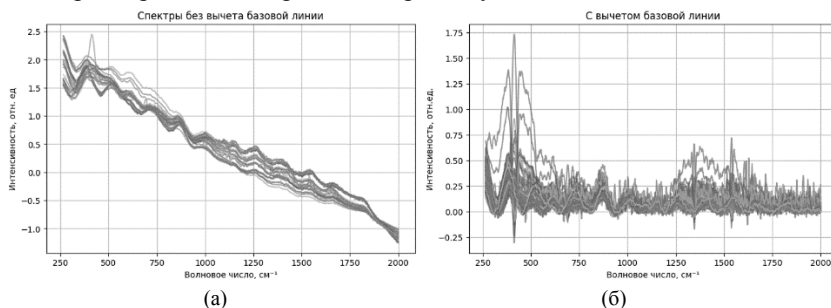


Рисунок 1 – Спектры рамановского рассеяния кожи после предварительной обработки без удаления базовой линии (а) и с удалением (б)

Далее для построения классификатора были построены две модели. Первая с использованием метода ANOVA с последующим линейным дискриминантным анализом (LDA), а вторая с использованием PLS-DA. Обе модели предполагали кросс-валидацию (K-fold) по 10 фолдам. Классификация была проведена для классов «норма» в количестве 183 спектров и «патология» в количестве 194 спектров. Результаты классификации для спектров с первым алгоритмом предобработки представлены на рисунке 2 для моделей на основе ANOVA с LDA и PLS-DA, а для спектров с удаленной базовой линией и аналогичными моделями на рисунке 3.

В работе получено, что для спектров без вычета базовой линии классификатор ANOVA с LDA демонстрирует ROC AUC, равный $0,536 \pm 0,073$, а PLS-DA — $0,698 \pm 0,114$. Для спектров с удалением базовой линии методом ALS ANOVA с LDA показал ROC AUC $0,563 \pm 0,055$, а PLS-DA — $0,672 \pm 0,096$.

Классификатор на основе PLS-DA показал статистически значимый результат (на основе парного t-теста при $p = 0,05$) как для спектров с вычетом базовой линии, так и без, поэтому можно сказать, что он лучше справляется с классификацией этих спектров. В случае же сравнения одинаковых моделей для спектров с разными предобработками нельзя сказать, какой алгоритм предобработки является лучшим, так как различия

между моделями статистически не значимо.

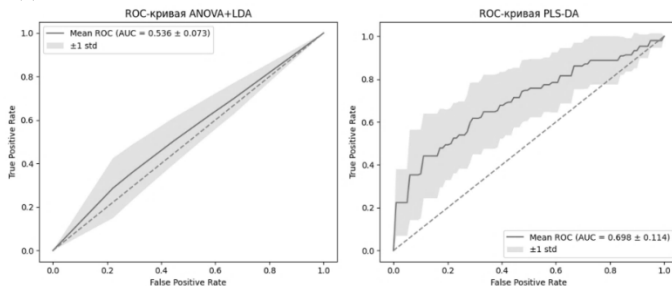


Рисунок 2 – ROC-кривые моделей на основе ANOVA с LDA и PLS-DA для спектров с без вычета базовой линии

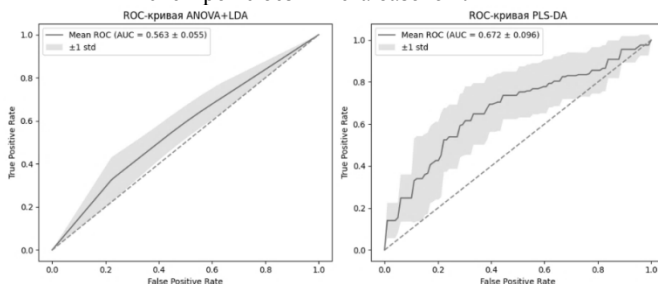


Рисунок 3 – ROC-кривые моделей на основе ANOVA с LDA и PLS-DA для спектров с вычетом базовой линии

Список использованных источников

1. Чупров И.Н., Сыдилов А.А., Заславский Д.В., Насыров Р.А. Дерматогнокопатология / Иллюстрированное руководство для врачей, 2021.
2. English D.R. Factors influencing the number needed to excise: excision rates of pigmented lesions by general practitioners, 2004.

Сальникова Екатерина Олеговна, студ. гр. 6402-120304D, fume890@outlook.com.
Христофорова Юлия Александровна, к.физ.-мат.н., доцент каф. ЛБС khristoforova.yua@ssau.ru.

УДК 543.424.2; 004.93

АНАЛИЗ ВЛИЯНИЯ АУГМЕНТАЦИИ ДАННЫХ НА СТРУКТУРУ РАМАНОВСКИХ СПЕКТРОВ СЫВОРОТКИ КРОВИ В ПРОСТРАНСТВЕ ГЛАВНЫХ КОМПОНЕНТ

А.И. Дьяченко, Ю.А. Христофорова
«Самарский национальный исследовательский университет имени академика С.П. Королёва», г. Самара

Ключевые слова: PCA, аугментация, спектры рамановского рассеяния, дисперсия.