

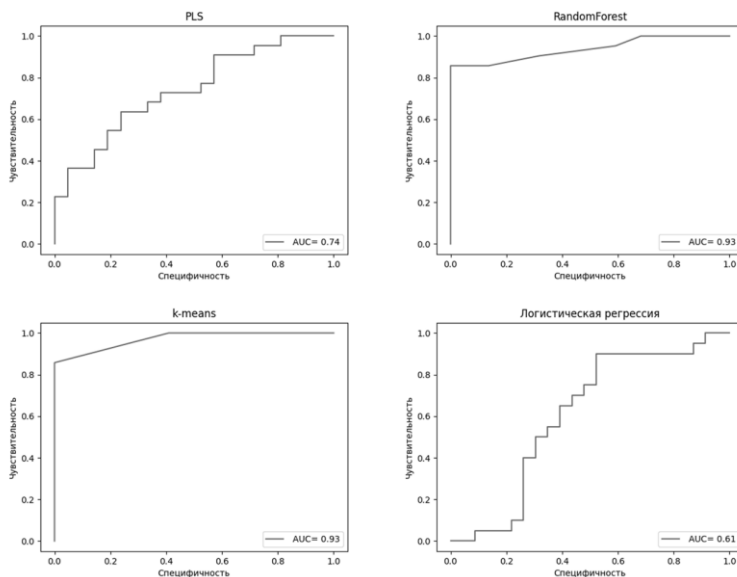


Рисунок 1 – ROC-кривые для методов, используемых в работе

Анализ рамановских спектров сыворотки крови с привлечением методов машинного обучения позволил выделить ключевые спектральные полосы, которые могут указывать на наличие определенных химических веществ в крови человека, ассоциированных с миеломной болезнью. Результаты могут быть использованы в медицине при диагностике миеломы для анализа больших объемов данных за короткое время, что позволит предсказывать вероятность наличия миеломы.

#### Список использованных источников

1. Горская, Н.И. Оптические исследования крови в видимом спектральном диапазоне / Н. И. Горская // Успехи современного естествознания. – 2007. – № 12–1. – С. 48.
2. Silver nanoparticles-based substrate for blood serum analysis under 785 nm laser excitation / S. Z. Al-Sammaraie, L. A. Bratchenko, E. N. Tytikova [et al.] // Journal of Biomedical Photonics & Engineering. – 2022. – Vol. 8, № 1. – P. 010301.

Горбачева Полина Владимировна, студ. гр. 6231-120404D, polinag63@mail.ru.  
Матвеева Ирина Александровна, к.т.н., доцент каф. ЛБС, matveeva.ia@ssau.ru.

УДК 543.424.2:616-71

## АУГМЕНТАЦИЯ РАМАНОВСКИХ СПЕКТРОВ ДЛЯ ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ

К.М. Гуськова

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

**Ключевые слова:** рамановские спектры, новообразования кожи, аугментация, классификация, машинное обучение.

В задачах классификации зачастую количества исходных данных бывает недостаточно для эффективной многоклассовой классификации. Однако можно применить методы аугментации для искусственного увеличения выборки. В работе проведена аугментация рамановских спектров новообразований кожи с помощью различных методов [1].

В исходной выборке присутствуют следующие виды новообразований: дерматофиброма – DF, гемангиома – H, себорейный кератоз – SK, базальноклеточная карцинома – BCC, плоскоклеточная карцинома – SCC, злокачественная меланома – MM, пигментный невус – PN, папиллома – HPV.

Аугментация данных – это процесс искусственного генерирования новых данных на основе существующих, который используется преимущественно для обучения новых моделей машинного обучения. Использование данного метода помогает моделям лучше обобщать данные, снизить переобучение и повысить устойчивость, так как модель учится распознавать объекты в разных условиях, и при этом не требуется сбор новых реальных данных [2].

В ходе выполнения работы рассматривались два метода аугментации данных:

- 1) с помощью искажений спектра (аугментация с шумом, с масштабированием, со сдвигом, со сглаживанием);
- 2) с помощью метода разрешения многомерных кривых (MCR).

На рисунке 1 представлен пример получения аугментированных рамановских спектров с помощью искажений спектра (с шумом, с масштабированием, со сдвигом, со сглаживанием).

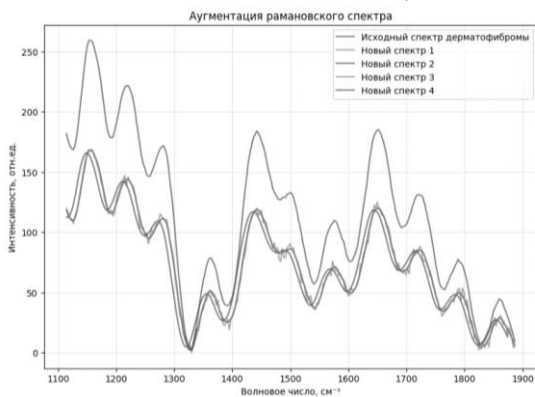


Рисунок 1 – Аугментированные рамановские спектры дерматофибromы

На рисунке 2 представлен пример аугментации рамановского спектра

с помощью метода MCR. Смысл этого метода заключается в том, что исходный спектр раскладывается на компоненты и их концентрации (вклад). Далее для каждого класса строятся диаграммы размаха и генерируются новые случайные концентрации по равномерному закону. Далее происходит восстановление по сгенерированным концентрациям новых спектров.



Рисунок 2 – Аугментированный рамановский спектр дерматофибромы

Искажения спектра изменяют существующие спектры через прямые трансформации: добавление шума, масштабирование интенсивности, сдвиги по волновому числу и сглаживание. Эти операции имитируют экспериментальные вариации (шум детектора, калибровка). MCR-аугментация использует математическое разложение исходных спектров на чистые компоненты и их концентрации. Новые спектры генерируются восстановлением с случайными концентрациями по равномерному закону, моделируя биохимические вариации тканей.

Применение методов аугментации позволит увеличивать количество спектров в выборке, которая будет применяться в дальнейших исследованиях для обучения алгоритма многоклассовой классификации новообразований кожи.

#### Список использованных источников

- 1.State-of-the-Art in Skin Fluorescent Photography for Cosmetic and Skincare Research: From Molecular Spectra to AI Image Analysis / K. Chekanov, D. Danko, T. Tlyachev [et al.] // Life. – 2024. – Vol. 14, № 10. – P. 1271.
- 2.Deep learning data augmentation for Raman spectroscopy cancer tissue classification / M. Wu, S. Wang, S. Pan [et al.] // Scientific reports. – 2021. – Vol. 11, № 1. – P. 23842.

Гуськова Карина Михайловна, студ. гр. 6231-120404D, karinaguskova2002@mail.ru.