**Применение методов искусственного интеллекта к анализу инфракрасных спектров мочи в урологическом скрининге**

***Павлеев Иван Александрович1,2***  
*Аспирант*  
*1. Московский государственный университет имени М. В. Ломоносова,  
Физический факультет, Москва, Россия  
E–mail: pavleev.ia16@physics.msu.ru*

*2. МИРЭА – Российский технологический университет, Москва, Россия*

Одна из актуальных проблем в медицинской физике – отсутствие простого и быстрого способа ежедневного получения информации о здоровье человека. Современные технологии неинвазивной диагностики активно развиваются за счёт интеграции спектроскопии с методами машинного обучения [1]. Моча, содержащая информацию о состоянии внутренних органов и метаболизме и отличающаяся простотой получения, – становится идеальным биоматериалом для такого анализа [2]. В настоящем исследовании предлагается методика урологического анализа на основе инфракрасной спектроскопии и хемометрики. Основная идея работы заключается в получении оптических спектров мочи в диапазоне 410–940 нм с использованием 18-канального спектрометра и последующей обработки данных алгоритмами снижения размерности и регрессии. Проблема высокой размерности спектральных данных, обусловленная корреляцией между соседними длинами волн, решается с применением методов снижения размерности, таких как метод главных компонент (PCA) и метод частичных наименьших квадратов (PLS) [3, 4].

Актуальность применения инфракрасной спектроскопии подтверждается исследованием [5] показавшим, что в диапазонах 1350–1800 и 2050–2375 нм спектральные сигнатуры таких компонентов, как мочевина, креатинин, глюкоза, белок и кетоновые тела, достаточно уникальны и могут использоваться для количественного анализа с точностью, сопоставимой с традиционными лабораторными методами. Данное исследование подчёркивает преимущество спектрального анализа мочи за счёт отсутствия необходимости в реагентах и простоты подготовки проб. Диагностический потенциал и ценность инфракрасной спектроскопии в урологическом скрининге демонстрирует исследование [6], в ходе которого было показана возможность применения инфракрасной спектроскопии образцов мочи для диагностики хронической болезни почек.

В рамках настоящего исследования были измерены оптические спектры пропускания мочи 500 пациентов с использованием портативного 18-канального спектрометра, работающего в диапазоне 410–940 нм. Для каждого образца производилось последовательное измерение спектров при освещении двумя диодами - белым и инфракрасным. Полученные спектры сопоставлялись с результатами общего клинического анализа мочи, где для каждого параметра фиксировались отклонения от нормальных значений. Для установления корреляций между спектральными данными и клиническими параметрами применялся метод PLS, позволяющий построить модель регрессии. На обучающей выборке модель определяла оптимальное линейное преобразование, максимизирующее ковариацию между матрицами спектральных данных и отклонениями клинических показателей. Применение ступенчатой функции к результатам регрессионной модели позволяло интерпретировать их в виде бинарного распределения: отклонение/норма. Результаты работы модели оценивались по метрикам чувствительности и специфичности. В ходе обучения модели проводилась максимизация гармонического среднего данных параметров. Эффективность такого подхода демонстрируется следующими результатами [7].

В результате исследования были получены показатели чувствительности и специфичности определения отклонения от нормы для 19 параметров клинического анализа мочи. Так, для определения аномальных концентраций белка, глюкозы, нитритов и уробилиногена была получена чувствительность свыше 80%, что свидетельствует о высокой диагностической ценности метода для выявления патологических концентраций данных молекул. Также модель продемонстрировала высокую точность определения таких параметров, как клетки крови: эритроциты и тромбоциты, эпителиальные клетки, цилиндры и бактерии, где значения гармонического среднего метрик чувствительности и специфичности превышали 74%. В то же время для таких параметров, как кетоновые тела и дрожжевые клетки были получены сравнительно низкие значения чувствительности – 66% и 50% соответственно. Также не получилось обучить модель для определения аномального содержания билирубина и сперматозоидов в моче – из-за недостаточной статистики по данным параметрам. Эти факторы указывают на необходимость проведения дополнительных исследований для оптимизации модели.

Наибольший интерес представляют результаты по определению нитритов: чувствительность 99% и специфичность 92%. Точное определение малых концентраций нитритов имеет высокую диагностическую ценность, так как их присутствие в моче свидетельствует об инфекции мочевыводящих путей. Предложенный метод не требует расходных компонентов и специальных лабораторных условий для регистрации спектров мочи, что делает его удобным для массового скрининга. Однако для его валидации необходимы дополнительные исследования, включающие расширенную выборку пациентов с инфекциями мочевыводящих путей и дополнительный контроль диеты добровольцев. Таким образом, полученные результаты подтверждают перспективность использования инфракрасной спектроскопии в сочетании с методами искусственного интеллекта в урологическом скрининге.

**Литература**

[1] Goecks J. et al. How machine learning will transform biomedicine //Cell. – 2020. – Т. 181. – №. 1. – С. 92-101.

[2] Simerville J. A., Maxted W. C., Pahira J. J. Urinalysis: a comprehensive review //American family physician. – 2005. – Т. 71. – №. 6. – С. 1153-1162.

[3] Pearson K. Principal components analysis //The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. – 1901. – Т. 6. – №. 2. – С. 559.

[4] Martens H., Naes T. Multivariate calibration. – John Wiley & Sons, 1992.

[5] Pezzaniti J. L. et al. Preliminary investigation of near-infrared spectroscopic measurements of urea, creatinine, glucose, protein, and ketone in urine //Clinical biochemistry. – 2001. – Т. 34. – №. 3. – С. 239-246.

[6] Oliver K. V. Infrared spectroscopy as a clinical diagnostic method for detection of disease states: developments and applications in kidney diseases and cancer diagnoses : дис. – UCL (University College London), 2015.

[7] Balashov, I. S., Poimanov, Y. M., Egorenkov, M. V., Pavleev, I. A., Nesmiyanov, P. P., Samokhodskaya, L. M., Grunin, A. A., Fedyanin, A. A. Medical urine analysis method based on Vis-NIR optical spectroscopy using machine learning algorithms //Emerging Topics in Artificial Intelligence (ETAI) 2021. – SPIE, 2021. – Т. 11804. – С. 118041W.