

А.В. Калининченко, Т.А. Муштина

КЛАССИФИКАЦИЯ СТАДИЙ ГЛАУКОМЫ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ

Научный руководитель: д-р. мед. наук, доц. Т.В. Качан

Кафедра глазных болезней

Белорусский государственный медицинский университет, г. Минск

**УЗ «3-я городская клиническая больница имени Е.В. Клумова», г. Минска*

A.V. Kalinichenko, T.A. Mushtina

CLASSIFICATION OF GLAUCOMA STAGES USING MACHINE LEARNING

Tutor: Doctor of Medical Sciences, associate professor T.V. Kachan

Department of Eye Diseases

Belarusian State Medical University, Minsk

**3rd City Clinical Hospital named after E.V. Klumov, Minsk*

Резюме. Для стадирования ПОУГ построены четыре модели машинного обучения. На тестовой выборке точность моделей составила 84,6–92,3%, лучший результат показал Random Forest. На основе топ-6 наиболее значимых признаков (MD, возраст, PSD, sup 90°, толщина RNFL в темпорально-нижнем секторе, толщина GCL в нижнем секторе) создана упрощённая модель, сохранившая точность 92,3% и может быть рекомендована для клинического использования.

Ключевые слова: первичная открытоугольная глаукома, машинное обучение, Random Forest, стадирование, ОКТ.

Resume. For POAG staging, four machine learning models were developed. Test accuracy ranged from 84.6% to 92.3%, with Random Forest achieving the best result. A simplified model based on the top-6 features (MD, age, PSD, sup 90°, inferotemporal RNFL thickness, inferior GCL thickness) preserved 92.3% accuracy and can be recommended for clinical use.

Keywords: primary open-angle glaucoma, machine learning, Random Forest, staging, OCT.

Актуальность. Точное стадирование критически важно для выбора лечения и прогноза. Традиционный подход требует комплексной оценки периметрии, офтальмоскопии и ОКТ. Но даже глаукомные эксперты часто расходятся в интерпретации: межэкспертное согласие при оценке диска зрительного нерва варьируется от умеренного до низкого – значение карра составляет 0,34–0,68 [1,2]. А при решении вопроса, нуждается ли пациент в направлении к специалисту, эксперты соглашались друг с другом лишь в 60% случаев [3]. Эта диагностическая неоднородность и создаёт предпосылки для разработки объективных автоматизированных инструментов стадирования.

В последние годы методы машинного обучения активно внедряются в офтальмологию для автоматизации диагностики и прогнозирования течения заболеваний [4]. Однако большинство существующих моделей либо используют большое количество входных параметров (до 52 точек полей зрения), либо требуют специализированного оборудования, что ограничивает их практическую применимость в рутинной клинической работе.

Цель: разработать и валидировать модель классификации стадий ПОУГ с использованием методов машинного обучения, а также определить минимальный

набор наиболее информативных признаков, достаточный для сохранения высокой точности прогноза.

Задачи:

1. Создать базу данных мультимодальных показателей пациентов с глаукомой различных стадий и здоровых лиц.
2. Разработать и сравнить модели машинного обучения (логистическая регрессия, SVM, XGBoost, Random Forest) для классификации стадий глаукомы.
3. Выявить ключевые признаки, определяющие стадию заболевания, и отобрать минимальный набор для сохранения высокой точности.
4. Оценить диагностическую значимость упрощённой модели (на 6 параметрах) для возможного внедрения в клиническую практику в качестве инструмента поддержки принятия решений.

Материалы и методы. Данное ретроспективное исследование было проведено на базе УЗ «3-я городская клиническая больница имени Е.В. Клумова». В анализ включены 63 глаза 46 пациентов с верифицированным диагнозом ПОУГ. Критериями включения являлись: наличие полного офтальмологического обследования (тонометрия, периметрия, ОКТ); отсутствие сопутствующей патологии, способной повлиять на параметры (диабетическая ретинопатия, возрастная макулярная дегенерация, миопия и гиперметропия высокой степени). Распределение по стадиям (классификация А.П. Нестерова): I стадия (начальная) – 33 глаза, II стадия (развитая) – 19 глаз, III стадия (далекозашедшая) – 11 глаз.

Исходная база данных содержала 49 переменных, включая демографические (возраст, пол), тонометрические (БТМ, Маклаков), периметрические (MD, PSD), ОКТ-параметры (толщина слоя нервных волокон – RNFL, ганглиозного слоя – GCL в шести секторах, параметры диска зрительного нерва). Качество полученных томограмм контролировалось по показателю силы сигнала, который составлял не менее 30 единиц. Процесс построения модели включал этап предобработки данных с заменой пропущенных значений, кодированием категориальных признаков и масштабированием численных переменных. Итоговый набор признаков составил 53 числовые переменные. Данные стратифицированно разделены на обучающую (80%, $n=50$) и тестовую (20%, $n=13$) выборки с сохранением пропорций стадий.

Для классификации использовались четыре алгоритма: логистическая регрессия, SVM с радиальным ядром, XGBoost и Random Forest. Все модели обучались на стандартизированных синтезированных данных. Для Random Forest с целью предотвращения переобучения был проведён подбор гиперпараметров методом GridSearchCV (5-кратная кросс-валидация). Экспериментально отобраны следующие значения: $n_estimators = 100$, $max_depth = 10$, $min_samples_split = 10$, $min_samples_leaf = 1$.

Нормальность распределения количественных признаков оценивалась с помощью теста Шапиро-Уилка. Для сравнения групп по стадиям использован непараметрический критерий Краскела-Уоллиса. Корреляционный анализ выполнен с коэффициентом Пирсона. Все вычисления проведены в среде Python (библиотеки pandas, scikit-learn, xgboost, matplotlib, seaborn).

Результаты и их обсуждение. Клиническая характеристика выборки: средний возраст – 65 ± 11 лет, женщины составили 71%. Ввиду ненормального распределения большинства параметров (критерий Шапиро–Уилка, $p < 0,05$) для межгруппового сравнения стадий ПОУГ применяли непараметрический критерий Краскела–Уоллиса. По всем анализируемым признакам (MD, PSD, толщина RNFL, соотношение экскавации к диску) выявлены статистически значимые различия между стадиями ($p < 0,001$). Данный факт обосновывает корректность классификационной задачи.

Корреляционный анализ (рисунок 1) выявил наибольшую связь со стадией заболевания для среднего отклонения (MD, $r = 0,82$), секторальных периметрических порогов ($r = 0,65–0,78$), а также параметров оптической когерентной томографии (ОКТ) в нижних и темпоральных секторах ($r = 0,63–0,70$).

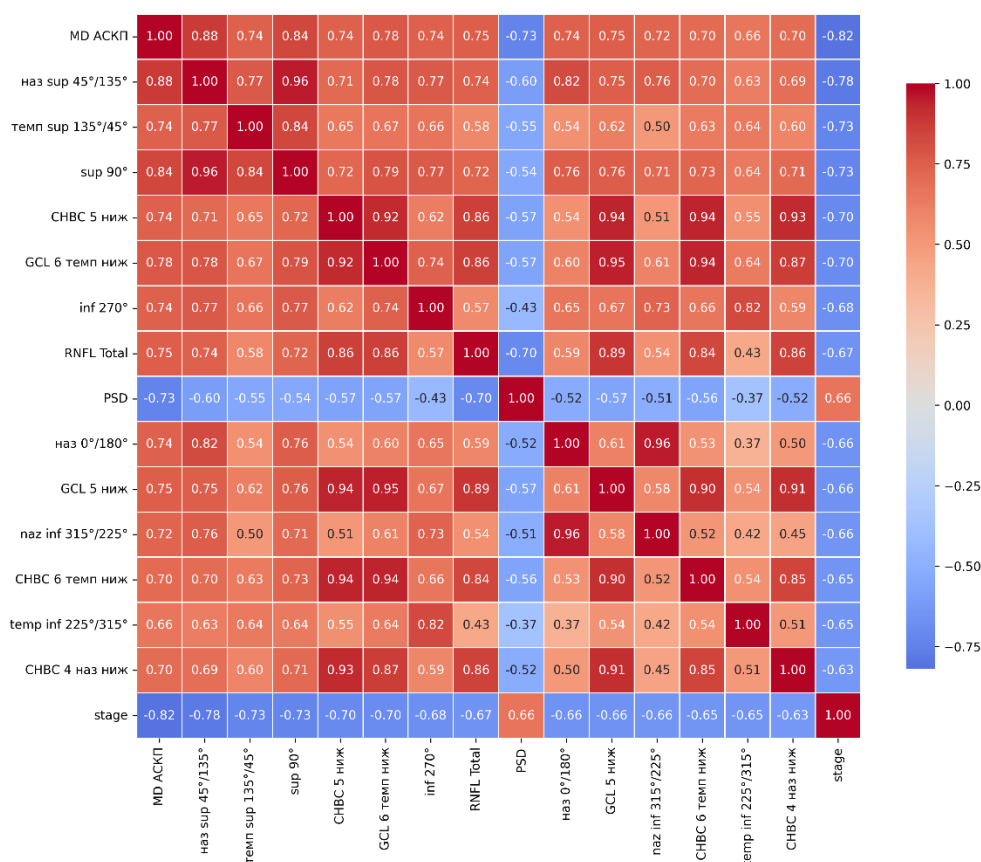


Рис. 1 – Корреляционная матрица топ-15 признаков

Для классификации стадий ПОУГ были обучены четыре модели: логистическая регрессия, SVM с радиальным ядром, XGBoost и Random Forest. Точность всех моделей на тестовой выборке варьировала от 84,6% до 92,3%. Наилучший результат показал Random Forest с оптимизированными гиперпараметрами, его кросс-валидационная точность (5-кратная) составила $82,0\% \pm 16,7\%$, что свидетельствует об устойчивости модели. Анализ матрицы ошибок для Random Forest продемонстрировал высокую специфичность: все 7 глаз с I стадией классифицированы верно, 3 из 4 глаз со II стадией – верно (один ошибочно отнесён к I), оба глаза с III стадией распознаны безошибочно.

Оценка вклада признаков (важности) показала, что наибольшее влияние на решение модели оказывают MD, возраст, PSD, толщина RNFL в темпорально-нижнем секторе, толщина GCL в нижнем секторе, а также секторальные периметрические пороги (90° , 225° , 315°). Выявленные закономерности соответствуют известным патогенетическим механизмам глаукомы – первичному поражению аксонов в нижнем и темпоральном секторах с последующими характерными периметрическими изменениями.

На основе шести наиболее значимых признаков (MD, возраст, PSD, $\text{sup } 90^\circ$, толщина RNFL в темпорально-нижнем секторе, толщина GCL в нижнем секторе) была обучена упрощённая модель Random Forest. Выбор обусловлен тем, что данные параметры вносят основной вклад в снижение неопределённости при классификации, тогда как остальные 47 признаков дублируют эту информацию. Точность упрощённой модели составила 92,3%, что не уступает полной модели, что делает её пригодной для рутинного клинического применения.

Разработанная шестипризнаковая модель является ключевым практическим результатом исследования (рисунок 2). Она позволяет врачу получить точную вероятностную оценку стадии ПОУГ (92,3%) при вводе всего шести параметров в калькулятор, что экономит время, снижает риск ошибок и делает инструмент пригодным для рутинной клинической работы.

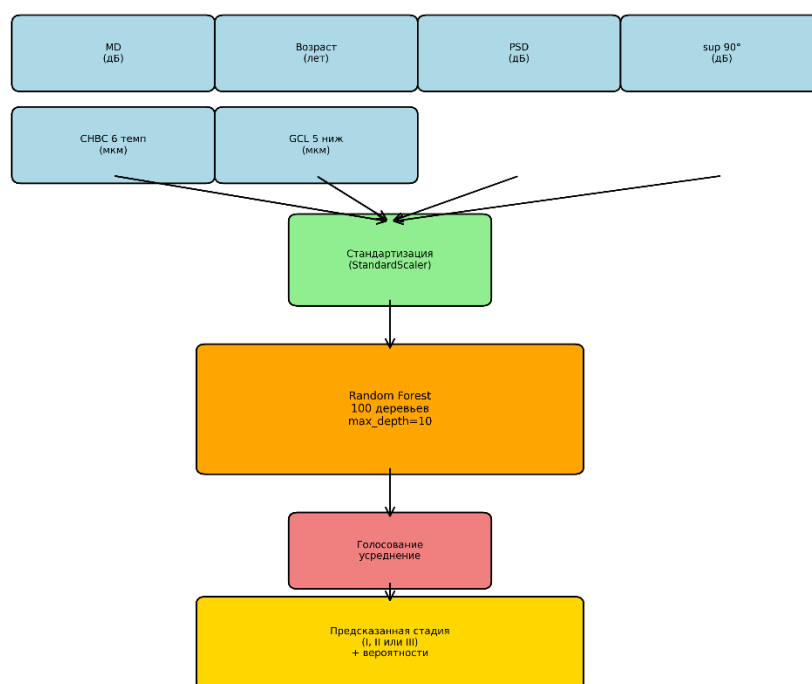


Рис. 2 – Схема работы модели

Выводы:

1. Разработана и валидирована модель классификации стадий глаукомы на основе ансамблевого метода Random Forest, достигшая точности 92,3% на тестовой выборке.
2. Оптимизация гиперпараметров ($n_estimators=100$, $max_depth=10$,

min_samples_split=10, min_samples_leaf=1) позволила улучшить обобщающую способность модели (кросс-валидационная точность 82,0%).

3. Наиболее значимыми предикторами явились MD, PSD, возраст, толщина RNFL в темпорально-нижнем секторе и толщина GCL в нижнем секторе.

4. Упрощённая модель, использующая всего шесть параметров, сохранила точность 92,3% и может быть рекомендована для создания клинического калькулятора стадии глаукомы.

5. Требуются дальнейшие исследования для внешней валидации и оценки влияния модели на клинические решения.

Литература

1. Clinical agreement among glaucoma experts in the detection of glaucomatous changes of the optic disk using simultaneous stereoscopic photographs / Azuara-Blanco A., Katz LJ, Spaeth GL [et al.] // Am J Ophthalmol. – 2003. – Vol. 136, No 5. – P. 949-950.

2. Correlation between Optic Nerve Parameters Obtained Using 3D Nonmydriatic Retinal Camera and Optical Coherence Tomography: Interobserver Agreement on the Disc Damage Likelihood Scale / Han JW, Cho SY, Kang KD [et al.] // J Ophthalmol. – 2014. – Vol. 2014. – P. 931738.

3. Optical Coherence Tomography Reading Centers Determine Glaucoma Referability / Stroka M. // Optometry Advisor. – 2022. – Vol. 2022. – P. 15.

4. Artificial intelligence in the prediction of glaucoma development and progression: A systematic review / Yang WYL, Wong HJ, Fu CE, Rojas-Carabali W, Agrawal R, Ang BCH [et al.] // Surv Ophthalmol. – 2026. – Vol. 71, No 1. – P. 235-247.