

УДК 616-092

## Классификация гистологических препаратов хряща коленного сустава по степени тяжести гонартроза с помощью глубокого обучения на основе YOLOv8

Бромберг А.А.

Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)»  
141701, Московская область, Долгопрудный, Институтский пер., д. 9

**Целью исследования** было разработать модель машинного обучения классификации гонартроза по степени тяжести на основании изображений костно-хрящевых образцов, окрашенных Сафранином O+FastGreen (SafO) и гематоксилин-эозином (HE).

**Материалы и методы.** Использована предобученная нейросеть YOLOv8m-cls пакета Ultralytics. Работа выполнена на языке Python 3.9. Машинное обучение проводилось на наборе гистологических данных Исследования хряща коленного сустава.

**Результаты.** Разработаны модели машинного обучения для классификации костно-хрящевых образцов по степени тяжести гонартроза. Достигнуты целевые метрики *accuracy\_top1* соответственно 95,7% для модели, обученной на HE-окрашенных изображениях и 94,3% – на SafO-окрашенных изображениях.

**Заключение.** YOLOv8 обладает достаточной предиктивной способностью в решении задач патоморфологии. Разработанные модели для классификации гонартроза могут быть применены в клинической практике для автоматизации и ускорения работы патологоанатомов и гистологов. Необходимо дальнейшее повышение качества моделей.

**Ключевые слова:** гонартроз; YOLO; глубокое обучение; патоморфология.

**Для цитирования:** Бромберг А.А. Классификация гистологических препаратов хряща коленного сустава по степени тяжести гонартроза с помощью глубокого обучения на основе YOLOv8. *Патогенез.* 2024; 22(2): 35-38.

**DOI:** 10.25557/2310-0435.2024.02.35-38

**Для корреспонденции:** Бромберг Анастасия Алексеевна, e-mail: a.bromberg@bestdoctor.ru

**Финансирование:** Исследование не имеет спонсорской поддержки

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Поступила:** 30.05.2024.

## Deep learning-based prediction of knee osteoarthritis severity score using histopathological images of bone-cartilage specimens and YOLOv8

Bromberg A.A.

The Moscow Institute of Physics and Technology  
Institutskiy Pereulok 9, Dolgoprudny of Moscow Region 141701, Russian Federation

This study **aims** to classify the histopathological HE- and SafO-stained images of bone-cartilage specimens automatically, using YOLOv8 by severity.

**Materials and methods.** The pre-trained YOLOv8m-cls neural network of the Ultralytics package was used. Python 3.9. was used. Models were trained on a Human Knee Cartilage Histopathology Assessment Dataset.

**Results.** Machine learning models were built to score the severity levels of knee osteoarthritis using the histopathological images of the bone-cartilage specimens. The *accuracy\_top1* performance was 95.7% for the model trained on HE-images and 94.3% - on SafO images.

**Conclusion.** YOLOv8 can be used in histopathology with high performance. Models can be used by researchers to automate the histopathology severity scoring of knee osteoarthritis. The dataset upscaling and the model modifying could help to improve the quality of predictions.

**Key words:** osteoarthritis; histopathological; YOLO; Deep learning.

**For citation:** Bromberg A.A. [Deep learning-based prediction of knee osteoarthritis severity score using histopathological images of bone-cartilage specimens and YOLOv8]. *Patogenez [Pathogenesis].* 2024; 22(2): 35-38. (in Russian)

**DOI:** 10.25557/2310-0435.2024.02.35-38

**For correspondence:** Bromberg Anastasia Alekseevna, e-mail: a.bromberg@bestdoctor.ru

**Funding.** The study had no sponsorship.

**Conflict of interest.** The authors declare no conflict of interest.

**Received:** 30.05.2024.

## Введение

Существует две наиболее распространённые шкалы для комплексной патоморфологической оценки изменений хряща при остеоартрите: HHGS (Histologic/histochemical grading system) [1] и OARSI (Osteoarthritis Research Society International) [2]. Результаты по шкале OARSI имеют очень высокую корреляцию с оценками по HHGS. У каждой шкалы есть ограничения. Оценка по HHGS не имеет значений для полной потери хрящевой ткани с обнажением субхондральной кости и может занимать больше времени, в основном из-за необходимости исследования образцов на большом увеличении. Оценка по шкале OARSI затруднена в переходных зонах между областями с разной степенью тяжести [3, 4].

Такая рутинная работа патологоанатомов и гистологов может быть ускорена и упрощена при использовании технологий искусственного интеллекта (ИИ), в частности, модели обнаружения объектов и сегментации изображений в реальном времени YOLOv8. YOLOv8 – новейшая версия YOLO, вышедшая в 2023 году. Это глубокая нейронная сеть, которая использует инновационный одноэтапный подход к обработке входных изображений. Согласно базам рецензируемых изданий PubMed и Scopus, YOLOv8-cls не применялась на медицинских данных. Для анализа микроскопических изображений с высокой эффективностью использовались предыдущие версии YOLO: например, для обнаружения и подсчёта клеток колоректального рака и лейкоцитов [5, 6].

**Цель:** разработать модель классификации гонартроза по степени тяжести на основании изображений костно-хрящевых образцов, окрашенных Сафранином O+FastGreen и гематоксилин-эозином.

## Материалы и методы исследования

Версия YOLOv8 ultralytics для классификации YOLOv8-cls предобучена на наборе данных ImageNet [7]. Набор данных Исследования хряща коленного сустава состоит из гистологических изображений интраоперационного материала от 90 пациентов. Всем пациентам выполнено первичное тотальное эндопротезирование коленного сустава в связи с первичным гонартрозом. Получено 180 костно-хрящевых образцов.

Сразу после получения образцы фиксировали в течение 48 часов при 4°C, готовили парафиновые срезы толщиной 5 мкм и окрашивали свежеприготовленным красителем гематоксилин-эозин (HE) или Сафранин O+FastGreen (SafO).

Два смежных среза изучили в 5-кратном увеличении. Проводилась оценка по шкалам HHGS и OARSI тремя исследователями по 3 раза с перерывом минимум в 3 месяца. Все изображения оцифрованы и представлены в формате tiff. Метаданные в формате XML к каждому пациенту включают возраст, пол, сторону хирургического вмешательства, среднюю оценку по шкалам HHGS, OARSI [8].

Для разделения наблюдений на классы по степени тяжести использованы оценки по OARSI. Общепринятая шкала соответствия оценки по OARSI и степени тяжести гонартроза отсутствует, поэтому для сопоставления оценок по OARSI и степени тяжести использована регрессионная модель соответствия оценок OARSI и HHGS, разработанная V.P. Mantripragada с соавт. [4] (**табл. 1**):

В наборе данных преобладают пациенты с изменениями, соответствующими лёгкой и умеренной степени тяжести гонартроза (**рис. 1**).

С помощью модуля train\_test\_split библиотеки Scikit-learn сформированы тренировочная, валидационная и тестовая выборки с применением стратификации в соответствии распределению по классам. Распределение данных составило 70%, 20% и 10% соответственно (**табл. 2**). Наборы данных преобразованы к формату YAML.

Основные метрики производительности YOLOv8 – это accuracy\_top1 и accuracy\_top5. Первая метрика подразумевает абсолютную точность предсказания, top5\_accuracy указывает на наличие правильного ответа в 5 предложенных алгоритмом ответах, отсортированных по убыванию вероятности. В качестве основной метрики оценки качества выбрана accuracy\_top1. Accuracy\_top5 не имеет смысла для выбранной задачи, так как представлено 4 класса. Для повышения качества обучения применялись аугментации (**рис. 2**).

## Результаты исследования

Достигнута accuracy\_top1: 95,7% для HE- и 94,3% для SafO-окрашенных изображений. Модели обладают меньшей точностью для раннего гонартроза. Точ-

Таблица 1

Соответствие оценок по HHGS, OARSI и степени тяжести гонартроза (по [4])

Степень тяжести гонартроза	HHGS	OARSI
Ранняя	< 2	< 3,4
Лёгкая	2-5	3,4-8,6
Умеренная	6-9	8,6-15,4
Тяжёлая	10-14	15,4-24,0

Таблица 2

Количественное распределение изображений в разных выборках

	Набор данных, окрашенных HE	Набор данных, окрашенных SafO
Тренировочная выборка	245	245
Валидационная выборка	70	70
Тестовая выборка	36	35

ность модели на тестовом наборе данных сопоставима с валидацией: 94,4% для HE- (100% – ранний, 95,7% – лёгкий, 83,3% – умеренный, 100% – тяжёлый), 91,4% для SafO-окрашенных изображений (66,6% – ранний, 95,7% – лёгкий, 100% – умеренный).

### Обсуждение

Разведочный анализ набора данных Исследования хряща коленного сустава выявил несбалансированность по оценкам OARSI, и недостаточный объём данных с тяжёлой степенью гонартроза.

Khader с соавт. применяли предварительно обученные нейросети DarkNet-19, MobileNet, ResNet101, NASNet для извлечения признаков и обучения клас-

сификатора с использованием метода опорных векторов на том же наборе данных. Результаты сопоставимы с разработанными моделями: ассигасу достигла 98,04% и 97,03% для HE и SafO соответственно [9]. Для повышения достоверности предсказаний необходимо увеличение объема данных и улучшение качества модели с учётом отличий в используемом оборудовании, методах подготовки и окраски препаратов в лабораториях. Доработанные модели глубокого обучения могут быть использованы в клинической практике для автоматизации рутинной морфометрической работы и снижения нагрузки на гистологов, патологоанатомов. Алгоритмы могут быть интегрированы в существующее программное обеспечение для обработки гистологических изображений.

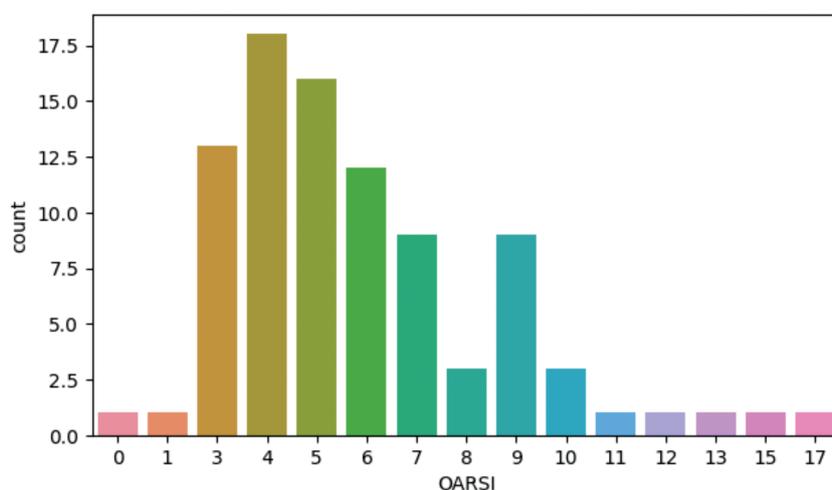


Рис.1. Распределение наблюдений в наборе данных по оценкам по OARSI

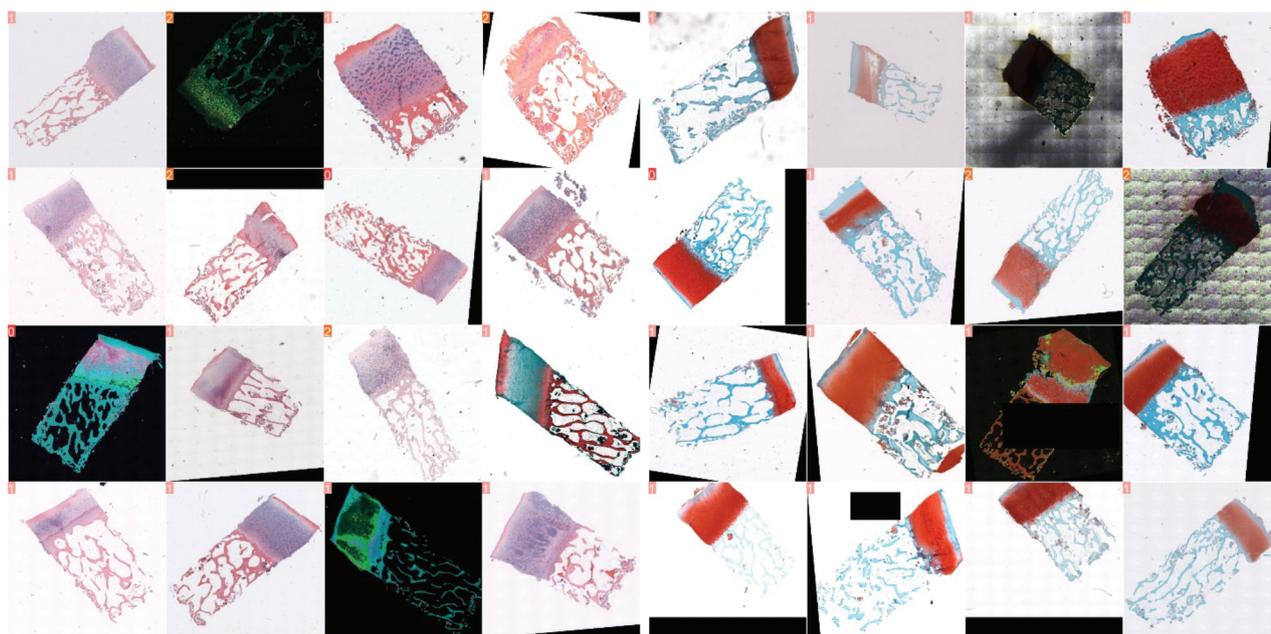


Рис. 2. Примеры изображений из тренировочных наборов данных с применением аугментаций.

Стоит учитывать, что шкалы HHGS и OARSI не учитывают ряд изменений при гонартрозе.

Mantripragada с соавт. провели исследование признаков гонартроза, не учитываемых HHGS и OARSI. Выявлены патологические изменения в глубоких зонах при сохранении хорошей структуры поверхности, что несовместимо с рассчитанными низкими степенями тяжести [4]. Такой полиморфизм может быть как проблемой для обучения алгоритмов компьютерного зрения, так и возможностью для развития диагностической точности. Машинное обучение может предложить новые закономерности и подходы к классификации: например, перспективно применение ансамблей или каскадов моделей для выделения участков с разными признаками на одном изображении.

### Заключение

YOLOv8-cls впервые применена для классификации гонартроза по степени тяжести на основании костно-хрящевых образцов, окрашенных HE и Safo. Достигнута высокая точность модели на наборе данных Исследования хряща коленного сустава. YOLOv8 обладает достаточной предиктивной способностью в решении задач патоморфологии. Для достоверности предсказаний необходим больший объем данных.

### Список литературы

1. Mankin H.J., Dorfman H., Lippiello L., Zarins A. Biochemical and metabolic abnormalities in articular cartilage from osteo-

- arthritic human hips. II. Correlation of morphology with biochemical and metabolic data. *J. Bone Joint Surg. Am.* 1971; 53(3): 523–537.
2. Pritzker K.P., Gay S., Jimenez S.A., Ostergaard K., Pelletier J.P., Revell P.A., Salter D., van den Berg W.B. Osteoarthritis cartilage histopathology: grading and staging. *Osteoarthritis Cartilage.* 2006; 14(1): 13–29. DOI: 10.1016/j.joca.2005.07.014
3. Pearson R.G., Kurien T., Shu K.S., Scammell B.E. Histopathology grading systems for characterisation of human knee osteoarthritis-reproducibility, variability, reliability, correlation, and validity. *Osteoarthritis Cartilage.* 2011; 19(3): 324–331. DOI: 10.1016/j.joca.2010.12.005
4. Mantripragada V.P., Piuizzi N.S., Zachos T., Obuchowski N.A., Muschler G.F., Midura R.J. High occurrence of osteoarthritic histopathological features unaccounted for by traditional scoring systems in lateral femoral condyles from total knee arthroplasty patients with varus alignment. *Acta Orthop.* 2018; 89(2): 197–203. DOI: 10.1080/17453674.2017.1398559
5. Haq I., Mazhar T., Asif R.N., Ghadi Y.Y., Ullah N., Khan M.A., Al-Rasheed A. YOLO and residual network for colorectal cancer cell detection and counting. *Heliyon.* 2024; 10(2): e24403. DOI: 10.1016/j.heliyon.2024.e24403
6. Chen Y., Zhang C., Chen B., Huang Y., Sun Y., Wang C., Fu X., Dai Y., Qin F., Peng Y., Gao Y. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases. *Comput. Biol. Med.* 2024; 170: 107917. DOI: 10.1016/j.combiomed.2024.107917
7. YOLOv8 – Ultralytics YOLO Docs. Available at: <https://docs.ultralytics.com/models/yolov8/> Retrieved: 30.05.2024
8. Mantripragada V.P., Piuizzi N.S., Muschler G.F., Erdemir A., Midura R.J. A comprehensive dataset of histopathology images, grades and patient demographics for human Osteoarthritis Cartilage. *Data Brief.* 2021; 37: 107129. DOI: 10.1016/j.dib.2021.107129
9. Khader A., Alquran H. Automated Prediction of Osteoarthritis Level in Human Osteochondral Tissue Using Histopathological Images. *Bioengineering (Basel).* 2023; 10(7): 764. DOI: 10.3390/bioengineering10070764

### Сведения об авторе:

Бромберг Анастасия Алексеевна — студентка 2 курса магистратуры Федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)»