



УДК 57.087.3+581.331.2

## Изготовление стандарта для метода абсолютных концентраций в спорово-пыльцевом анализе с использованием искусственного интеллекта

А. Т. Галимов

Институт экологии растений и животных УрО РАН, ул. 8 марта, д. 202, г. Екатеринбург, 620144, Россия.

E-mail: galimov\_at@ipae.uran.ru; ORCID iD: <https://orcid.org/0000-0003-1766-7461>

**Ключевые слова:** абсолютные концентрации, аналитический стандарт, нейросеть, палеореконструкция, палинология.

**Аннотация.** Для количественного определения абсолютной концентрации пыльцы в геологических отложениях обычно используется метод маркерных частиц. Чаще всего в качестве маркеров применяют таблетки со спорами *Lycopodium* и алюмосиликатные сферы. Однако производство и приобретение сфер «Palynospheres» и таблеток *Lycopodium* сопряжено с рядом трудностей. В данной работе предложено решение, основанное на спорах *Lycopodium* sp., которые окрашиваются в яркие цвета, что позволяет эксплуатировать их при работе с теми пыльцевыми спектрами, в естественном составе которых присутствуют споры рода *Lycopodium*. Для определения концентрации спор в каждом аналитическом стандарте был обучен ИИ. Получены аналитические стандарты, пригодные для установления абсолютной концентрации пыльцы в спорово-пыльцевом анализе, территориально доступные для палинологов из РФ.

## Preparation of an analytical standard for the absolute concentration method in pollen analysis using artificial intelligence

A. T. Galimov

Institute of Plant and Animal Ecology, UB RAS, 8 Marta St., 202, Ekaterinburg, 620144, Russian Federation

**Keywords:** absolute concentrations, analytical standard, neural network, paleoreconstruction, palynology.

**Summary.** To quantify the absolute concentration of pollen in geological sediments, the marker grain method is commonly used. *Lycopodium* spore tablets and aluminosilicate spheres are the two most commonly used markers. However, the production and acquisition of “Palynospheres” and *Lycopodium* tablets are associated with a number of difficulties. This work proposes a solution based on spores of *Lycopodium* sp., which are painted in bright colors, which allows them to be used even when working with those pollen spectra in which spores of the genus *Lycopodium* are naturally present. It is proposed to use AI to determine the concentration of spores in each analytical standard. Analytical standards suitable for use in establishing the absolute concentration of pollen in spore-pollen analysis have been obtained and are geographically accessible to palynologists from the Russian Federation.

### Введение

Абсолютное количество микрофоссилий – показатель, с помощью которого возможно оценить пыльцевую продуктивность растений, на-

личие или отсутствие сортировки пыльцевых зерен при захоронении (Devis, 1966; Bonny, 1972; Faegri, Iversen, 1975). Поскольку абсолютная концентрация пыльцы не зависит от частоты

встречаемости пыльцевых зерен других таксонов, она часто отражает изменения окружающей среды более непосредственно, чем процентные значения. Добавление известного количества маркерных частиц, таких как споры *Lycopodium* и «Palynospheres» делает возможным использование метода абсолютных концентраций.

Применение экзотической пыльцы или спор (например, *Lycopodium* sp.) для абсолютного подсчета пыльцы берет свое начало в палинологии XX в. Впервые метод был описан в 1962 г. (Benninghoff, 1962); новые разработки появляются и сегодня (Kitaba, Nakagawa, 2017). На данный момент мировым стандартом являются таблетки, произведенные в университете Лунда по методу, описанному в 1971 г. (Stockmarr, 1971). В оригинальных методиках использование спор *Lycopodium* не рекомендовалось для образцов, в которых *Lycopodium* присутствует в естественном спектре, но при современном изготовлении аналитических стандартов споры предварительно проходят процедуру ацетолиза, что решает эту проблему. Несмотря на это, в некоторых случаях ископаемые споры *Lycopodium* близки по цвету к спорам из аналитического стандарта, и мы можем потерять некоторую палеоэкологическую информацию из-за желания иметь абсолютные концентрации пыльцы. Эту проблему можно решить, используя алюмосиликатные сферы (Kitaba, Nakagawa, 2017). Но возникает другая проблема – логистические трудности и цена при их приобретении.

Производство собственного аналитического стандарта сопряжено с двумя основными трудностями. Первая – это подбор частиц, сопоставимых по размерности и плотности с пыльцевыми зёрнами. При этом важным является создание частиц, контрастирующих с пыльцой и спорами составляющих естественный пыльцевой спектр. Наиболее удобным представляется окрашивание в цвета, не встречающиеся в природной окраске пыльцы и спор. Вторая проблема – это максимально точное определение числа частиц в каждом аналитическом стандарте. Используемый метод, основанный на счетчике Коултера, предполагает наличие дорогостоящего оборудования (Stockmarr, 1973), тогда как искусственный интеллект позволяет решить аналогичную задачу с помощью микроскопа и цифровой камеры.

Цель работы: предложить альтернативу имеющимся маркерам используемых в анализе микрофосилий, которая будет доступна для исследователей на территории РФ.

## Материалы и методы

В работе использованы споры *Lycopodium clavatum* L. Окрашивание спор проводилось по методике, предложенной Gardner, Gray (1976). Было взято 3 навески по 1 г спор и помещено в пластиковые конические центрифужные пробирки на 15 мл. В каждую пробирку было добавлено 2 мл дистиллированной воды и 2 капли моющего средства в качестве ПАВ для повышения смачиваемости гидрофобных спор. Пробирки были помещены на водяную баню и прогревались в течение 15 мин. После чего в каждую пробирку были добавлены красители: сафранин-Т, малахитовый зеленый, фиолетовый кристаллический, приготовленные из расчета 0,009 г красителя на 0,1 г этилового спирта 96 %. Окрашивание проводилось в три цвета, т. к. это позволяет применять аналитический стандарт, например, синего цвета, в случаях, когда при пробоподготовке пыльца окрашивается в красный цвет. Тщательно перемешав смесь, пробирки были помещены на водяную баню на 1,5 ч. После чего в каждую пробирку были добавлены 0,06 г мочевины, 0,16 г 10 % формальдегида, 0,097 г КОН 10 %. После фиксации в течение 24 часов пробирки были помещены в центрифугу при 3000 об./мин на 15 мин, надсадочная жидкость сливалась, после чего пробирки доливались водой до 15 мл и снова центрифугировались. Операция была повторена 3 раза. После чего содержимое пробирок было помещено в отдельные чашки Петри и накрыто фильтровальной бумагой до полного высыхания.

После высыхания 100 мг окрашенных спор добавлялось в 200 мл дистиллированной воды и перемешивалось магнитной мешалкой в течение 2 часов. Полученная суспензия раскапывалась в 100 пластиковых пробирок объемом 2 мл по 200 мкл суспензии в каждую.

Из 30 пробирок было взято по 30 капель по 3 мкл и помещены на предметные стекла, после высыхания были получены микрофотографии. Всего получено 900 фотографий, по которым подсчитывается количество спор в пробирке (рис. 1).

Для подсчета количества спор в каждой капле (3 мкл) была обучена нейросеть на основании YOLOv8 (Jocher et al., 2023). Для обучения модели использовалось 100 изображений и 20 для валидации. Искусственного увеличения тренировочного набора данных за счет поворота изображений, отражения по вертикали и т. д. не

потребовалось. Обучение модели проводилось на суперкомпьютере «Уран» института математики и механики им. Н. Н. Красовского УрО РАН, г. Екатеринбург. После обучения модели были проанализированы оставшиеся 780 изображений.

### Результаты и их обсуждение

Окрашенные споры представляют собой хорошо различимые частицы разного цвета со стойким окрашиванием, не слипающиеся в комки (рис. 2).

Обучение модели проводилось на 120 изображениях, полученных при фотографировании окрашенных спор. Для оценки точности модели использовалась метрика mAP50 (Pascal VOC Challenge, 2007). Эта метрика применяется для оценки точности моделей обнаружения объектов путем расчета точности и полноты предсказанных ограничивающих рамок (объектов) по сравнению с реальными данными. Значения mAP50 близки к 1, что указывает на то, что модель стабильна и при различных пороговых значениях достоверности и имеет высокую точность предсказаний на различных изображениях (рис. 3). Весь цикл тренировки модели составил 150 итераций обучения (эпох) (Jocher et al., 2023).

Для валидации обученной модели проводился подсчет числа спор на изображениях, не

вошедших в обучающий набор данных, модель продемонстрировала высокую способность к верному определению числа спор (рис. 4). При сопоставлении результатов определения числа спор моделью на 30 изображениях с этими же изображениями, на которых количество спор подсчитывалось вручную, была установлена средняя ошибка модели – 0,9 споры на изображение.

После чего был проведен анализ с помощью нейросети 900 изображений, представляющих 30 пробирок, содержащих споры (табл. 1). Среднее значение ( $S$ ) = 4400 шт., стандартное отклонение ( $SD$ ) = 429, ошибка среднего ( $SE$ ) = 78, коэффициент вариации ( $CV$ ) = 9,88 %.

Результаты проверки распределения на нормальность показали, что изученная выборка из нормально распределенной генеральной совокупности (рис. 5).

В результате была получена прототипная партия проб, для каждой установлено число спор в количестве  $4400 \pm 78$  шт. Пробы представляют собой пробирки с бесцветной жидкостью 0,2 мл, готовые к добавлению на любой стадии пробоподготовки на спорово-пыльцевой анализ. А также отработан подход для получения любого количества проб с любым известным количеством спор.



Рис. 1. Пример фотографии для оценки количества спор в препарате, на которой зафиксированы подготовленные споры *Lycopodium* sp. Одно изображение позволяет описать 3 мкл суспензии.

**Таблица 1.** Результат подсчета количества спор в 30 случайно выбранных пробирках из изготовленной партии

Номер пробирки	Групповое среднее, шт.	Номер пробирки	Групповое среднее, шт.
1	3762	16	4684
2	4242	17	4025
3	3510	18	5226
4	4095	19	4621
5	3980	20	4700
6	3827	21	4640
7	4676	22	4415
8	3810	23	4553
9	3977	24	4714
10	3850	25	4703
11	4264	26	4420
12	4471	27	4458
13	4331	28	4935
14	3864	29	5162
15	4286	30	3949

Для проверки возможности использования полученного аналитического стандарта в спорово-пыльцевом анализе одна из изготовленных проб была включена в стандартный процесс пробоподготовки (Nakagawa, 1988). К навеске почвы 5 г добавлялось 200 мкл суспензии со спорами, после чего к навеске добавлялся 10%-й раствор HCl до момента прекращения газообразования, затем навеска заливалась 10%-м раствором KOH

в двукратном объеме и прогревалась на водяной бане при 95 °C в течение 10 мин. Далее суспензия проливалась через сито 0,25 мм и отстаивалась в течении 8 часов, верхняя часть осадка сливалась, и оставшаяся часть концентрировалась центрифугированием. Отделение минерального компонента осадка проводилось сепарационным методом Гричука с использованием тяжелой жидкости ГПС-В (Grichuk, Zaklinskaya, 1948; Campbell et al., 2016). После чего были сделаны микрофотографии на увеличении  $\times 400$  и  $\times 200$ , на которых присутствуют пыльцевые зерна и споры. Окрашенные споры хорошо отличаются от естественных пыльцы и спор, и сохраняют окраску после пробоподготовки (рис. 6).

Существующие на рынке продукты «Pollen tablets» университета Лунд (Stockmarr, 1971) и «Palynospheres» компании Palynotech имеют следующие параметры: таблетки со спорами, подготавливаются с расчетом  $\sim 18\,000$  (зависит от партии) спор с коэффициентом вариации 3,66 %, что ниже, чем в полученном нами подходе; суспензии керамических сфер (Kitaba, Nakagawa, 2017) поставляются в вариантах 20 000 и 4 500 сфер при коэффициенте вариации 10 %, что аналогично значениям, полученным в этой работе. Отсюда можно сделать вывод, что предлагаемый аналитический стандарт обладает достаточными показателями точности для использования в палинологической практике.



**Рис. 2.** Окрашенные споры *Lycopodium* sp.

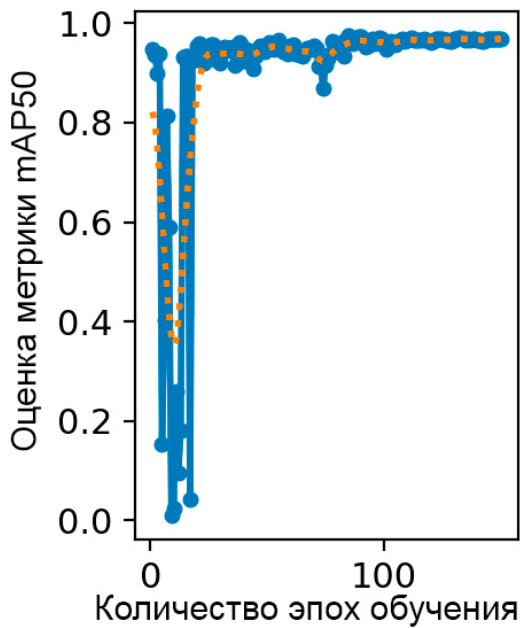


Рис. 3. Изменение метрики mAP50 обучения модели YOLOv8 для 150 итераций обучения: значения  $> 0,8$  – высокая точность; значения  $< 0,8$  – низкая точность.

Предложенный аналитический стандарт благодаря окрашиванию однозначно является экзотическим маркером вне зависимости, для какой экосистемы идет реконструкция. Удобная фасовка в жидкости позволяет пропустить этап с HCl, который необходим для растворения таблеток. Следующим этапом исследования будет повышение однородности распределения числа спор в каждой пробе, чтобы довести разброс до минимальных значений.

### Заключение

Предложен подход получения навесок для подсчета абсолютных концентраций пыльцевых зерен в образцах. Получена первая партия стандартов, готовых к использованию, которые содержат  $4400 \pm 78$  спор в каждой пробирке. Полученные аналитические стандарты являют собой конкурентный продукт, который может быть применен наравне с общепризнанными инструментами для определения концентраций спор.

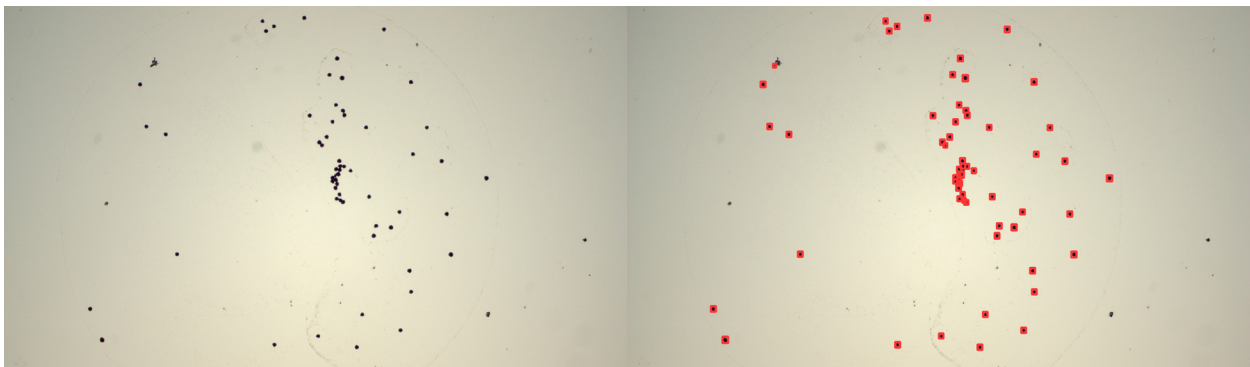


Рис. 4. Результат детектирования спор на изображении: А – изображение, предоставленное модели для обнаружения спор; В – изображение, на котором обведены объекты, определенные моделью как споры *Lycoperidium* sp.

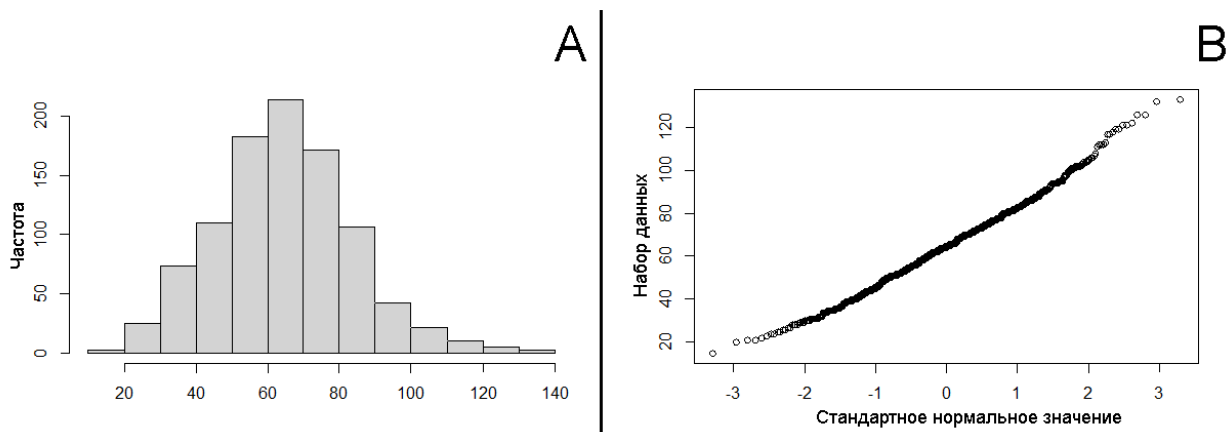
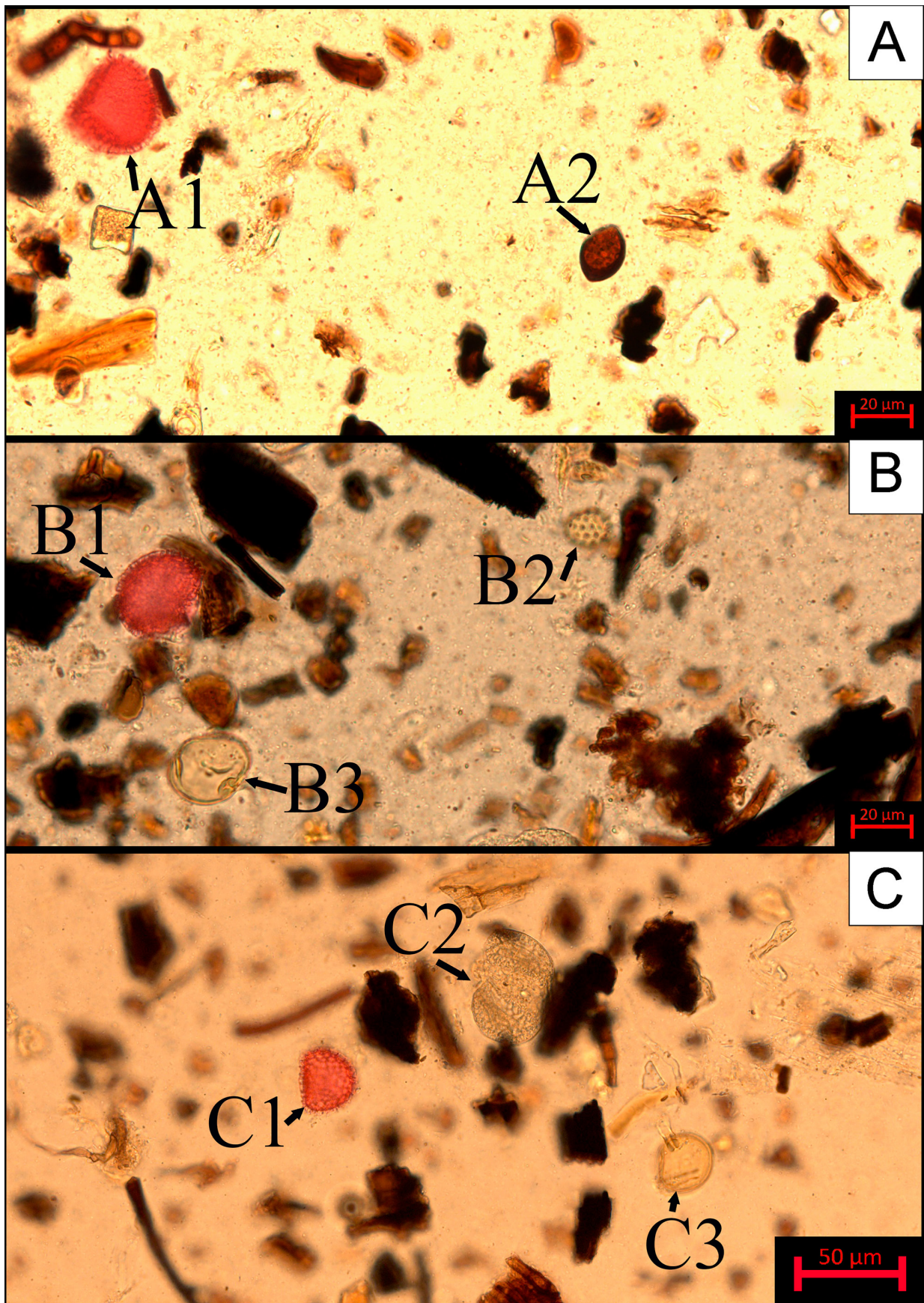


Рис. 5. Результаты графического определения нормальности распределения числа спор в пробирке (аналитическом стандарте): А – гистограмма, стремящаяся к нормальной; В – график квантилей стремящийся к нормальному.



**Рис. 6.** Окрашенные споры *Lycoperidium* в палинологическом препарате. На микрофотографии представлены: А – окрашенная спора *Lycoperidium* (A1) и спора Ascomycota (A2); В – окрашенная спора *Lycoperidium* (B1), спора гриба рода *Glomus* (B2), пыльца Chenopodiaceae (B3); С – окрашенная спора *Lycoperidium* (C1), пыльца *Pinus sylvestris* (C2), спора гриба рода *Glomus* (C3).

Была обучена нейронная сеть, которая эффективно выполняет задачу определения количества спор на изображении ( $mAP50 = 0,98$ ). Создан размеченный массив данных, содержащий 120 изображений, который может задействован в дальнейшем обучении ИИ. Обученная модель, хоть и является более тяжеловесной, чем алгоритм, построенный на классическом машинном зрении, но позволяет получать хорошие результаты подсчета числа спор вне зависимости от качества входных изображений, которые могут включать мусорные частицы.

#### Благодарности

Работа выполнена при поддержке проекта РФФ № 22-14-00332 «Фаунистические связи Европы и Азии в четвертичном периоде на примере широко распространенных мелких млекопитающих Северной Евразии».

Автор выражает благодарность д. б. н., проф., Александру Васильевичу Бородину, за помощь при подготовке материалов к публикации.

#### REFERENCES / ЛИТЕРАТУРА

- Benninghoff W. S.** 1962. Calculation of pollen and spore density in sediments by addition of exotic pollen in known quantities. *Pollen et Spores* 4(2): 332–333.
- Bonny A. P.** 1972. A method for determining absolute pollen frequencies in lake sediments. *New Phytol.* 71: 393–405. <https://doi.org/10.1111/j.1469-8137.1972.tb04086.x>
- Campbell J. F. E., Fletcher W. J., Hughes P. D.** 2016. A comparison of pollen extraction methods confirms dense-media separation as a reliable method of pollen preparation. *J. Quat. Sci.* 31(6): 631–640. <https://doi.org/10.1002/jqs.2905>
- Davis M. B.** 1966. Determination of absolute pollen frequency. *Ecology* 47(2): 310–311.
- Faegri K., Iversen J.** 1975. *Textbook of pollen analysis*. Copenhagen: Munksgaard. 237 pp.
- Gardner G. D., Gray R. E.** 1976. Tracing subsurface flow in karst regions using artificially colored spores. *Bulletin of the Association of Engineering Geologists* 13: 177–197.
- Grichuk V. P., Zaklinskaya E. D.** 1948. *Analiz iskopaemykh pyltisy i spor i ego primeneniye v paleogeografii* [Fossil pollen and spore analysis and its application in paleogeography]. Moscow: OGIZ. 217 pp. [In Russian] (**Гричук В. П., Заклинская Е. Д.** Анализ ископаемых пыльцы и спор и его применение в палеогеографии. М.: ОГИЗ, 1948. 217 с.).
- Jocher G., Chaurasia A., Qiu J.** 2023. *Ultralytics YOLO* (Version 8.0.0). URL: <https://github.com/ultralytics/ultralytics> (Accessed 18 November 2024).
- Kitaba I., Nakagawa T.** 2017. Black ceramic spheres as marker grains for microfossil analyses, with improved chemical, physical, and optical properties. *Quat. Int.* 455: 166–169. <https://doi.org/10.1016/j.quaint.2017.08.052>
- Nakagawa T., Hayashi R., Yokoyama Y.** 1998. Dense-media separation as a more efficient pollen extraction method for use with organic sediment/deposit samples: comparison with the conventional method. *Boreas* 27(1): 15–24. <https://doi.org/10.1111/j.1502-3885.1998.tb00864.x>
- Pascal VOC Challenge* [2007]. The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results. URL: <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/> (Accessed 18 November 2024).
- Stockmarr J.** 1971. Tablets with spores used in absolute pollen analysis. *Pollen et Spores* 13: 615–621.
- Stockmarr J.** 1973. Determination of spore concentration with an electronic particle counter. *Danmarks Geol. Undersøgelse, Raekke* 15: 87–89.