

Научная статья

УДК 004.81:58

<https://doi.org/10.35266/1999-7604-2024-4-2>



Предсказание признаков растений для моделирования экосистем: комбинированная модель на основе изображений и вспомогательных данных

Валентин Валерьевич Брыкин¹✉, Сергей Олегович Крамаров², Адриан Гросу³

^{1, 2, 3}Сургутский государственный университет, Сургут, Россия

¹valentin.brykin@mail.ru✉

²maoovo@yandex.ru

³grosu_a@surgu.ru, <https://orcid.org/0009-0004-5520-6708>

Аннотация. В данной работе представлена комбинированная модель машинного обучения для предсказания признаков растений. Модель объединяет предобученную сверточную нейронную сеть (CNN) со вспомогательными переменными, представленными в виде табличных данных (CSV). Цель исследования – повысить точность предсказания признаков растений, используя как визуальную информацию, так и контекстные данные.

Модель обучалась на наборе данных, состоящем из 9147 изображений растений и соответствующих 167 вспомогательных переменных. Изображения были предобработаны с помощью трех предварительно обученных архитектур сверточных нейронных сетей (InceptionV3, ResNet, VGG19), а дополнительные данные были объединены с помощью слоя concatenate после выпрямления выходов вышеупомянутых CNN. Модель была оптимизирована с использованием алгоритма Adam и оценена соотвествующими метриками.

Результаты показали превосходство совмещенной модели над базовой CNN в предсказании признаков растений, что подтверждает эффективность использования дополнительных данных для повышения качества работы. Исследование демонстрирует потенциал гибридных моделей машинного обучения для анализа данных о растениях, полученных от пользователей со всего мира.

Ключевые слова: комбинированная модель, извлечение признаков, изображения растений, регрессия, экосистемы

Для цитирования: Брыкин В. В., Крамаров С. О., Гросу А. Предсказание признаков растений для моделирования экосистем: комбинированная модель на основе изображений и вспомогательных данных // Вестник кибернетики. 2024. Т. 23, № 4. С. 18–24. <https://doi.org/10.35266/1999-7604-2024-4-2>.

Original article

Plant traits prediction for ecosystem modeling: Combined model based on images and auxiliary data

Valentin V. Brykin¹✉, Sergey O. Kramarov², Adrian Grosu³

^{1, 2, 3}Surgut State University, Surgut, Russia

¹valentin.brykin@mail.ru✉

²maoovo@yandex.ru

³grosu_a@surgu.ru, <https://orcid.org/0009-0004-5520-6708>

Abstract. The study presents a combined machine learning model for predicting plant traits. The model combines a pre-trained convolutional neural network (CNN) with auxiliary variables represented as comma-separated values (CSV). The aim of the study is to improve the accuracy of plant trait prediction using both visual information and contextual data.

The model is trained on a dataset consisting of 9,147 plant images and the corresponding 167 auxiliary variables. The images are preprocessed using three pre-trained convolutional neural network architectures (InceptionV3, ResNet, VGG19). The supplementary data are combined using a concatenate layer after directing the outputs of the above CNNs. The model is optimized using the Adam algorithm and evaluated with appropriate metrics.

The results show the superiority of the combined model over the baseline CNN in predicting plant traits, confirming the effectiveness of using supplementary data to improve performance. The study demonstrates the potential of hybrid machine learning models to analyze plant data from users around the world.

Keywords: combined model, trait extraction, images of plants, regression, ecosystems

For citation: Brykin V. V., Kramarov S. O., Grosu A. Plant traits prediction for ecosystem modeling: Combined model based on images and auxiliary data. *Proceedings in Cybernetics*. 2024;23(4):18–24. <https://doi.org/10.35266/1999-7604-2024-4-2>.

ВВЕДЕНИЕ

Биоразнообразие является базисом устойчивого развития нашей планеты, обеспечивающим экологическую стабильность и поддерживающим биологические ресурсы. Извлечение признаков растений, именуемых *plant traits* (например, высоты растения, плотности его стебля или площади листьев), является очень важным моментом для изучения экосистем и прогнозирования их реакции на изменения окружающей среды. Еще большую актуальность этот вопрос приобретает в контексте глобального изменения климата, оказывающего значительное влияние на функционирование экосистем. Отсутствие достаточных данных о признаках растений является серьезным ограничением для моделирования экосистем и предсказания их будущего поведения [1–4].

В этой статье исследуется современный подход к извлечению признаков растений из изображений, сделанных пользователями со всего мира. Метод основан на использовании так называемой комбинированной модели, сочетающей сверточные нейронные сети (CNN) со вспомогательными переменными. CNN-блок используется для извлечения визуальных признаков из изображений, а вспомогательные переменные (в частности, информация о климате, почве, местоположении) предоставляют контекстную информацию. Логично, что совмещение этих двух источников информации позволит повысить точность предсказания признаков растений.

Таким образом, основная цель работы – демонстрация эффективности комбинирования моделей машинного обучения.

Работа с исходными данными и моделью машинного обучения выполнялась на языке Python с использованием соответствующих библиотек и документации.

МАТЕРИАЛЫ И МЕТОДЫ

Комбинированный алгоритм

Алгоритм основан на симбиозе двух подходов машинного обучения: извлечении признаков из фотографий сверточной нейронной сетью и нелинейной регрессии с использованием вспомогательных данных. Этот подход отражает принципы многомодального моделирования (обработка информации разной природы – изображения, переменные), способствующего улучшению качества работы. Кроме того, предобученная CNN играет роль «слабого учителя», уже обладающего знаниями о распознавании образов, тогда как гибридная модель дополнительно обучается на специфических данных о *plant traits*, что делает ее «сильным учителем», обладающим большей точностью [5].

Общий порядок действий алгоритма представлен в схеме на рис. 1.

Этапы подготовки и предварительной обработки данных

Исходные данные приведены на рис. 2 и 3 и представляют собой два набора информации разных модальностей: более 9000 цветных цифровых изображений растений разных видов.



Рис. 1. Схема алгоритма комбинирования
 Примечание: составлено авторами.

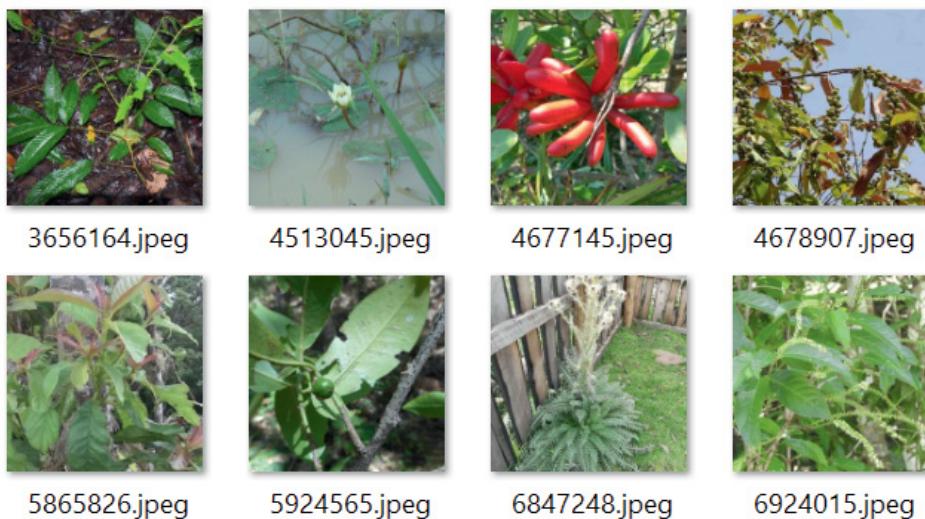


Рис. 2. Исходные данные (изображения)
 Примечание: составлено авторами по источнику [6].

Изображения были предварительно подготовлены – приведены к единому размеру 224×224 пикселя и нормализованы в диапазоне от 0 до 1 (путем деления значения яркости каждого пикселя на 255 – максимальный показатель для 8-битной глубины цвета).

Вспомогательные переменные в *CSV-файле* для каждого изображения представляют собой вектор из 167 значений, где:

- значение id – уникальный идентификатор, а также префикс имени изображения (для связи мультимодальных данных);
- 6 значений *WORLDCLIM_BIO* – климатические переменные;
- 62 значения *SOIL* – переменные почвы;
- 96 значений *MODIS/VOD* – многовременные спутниковые переменные;
- 3 значения X [*] – целевые признаки *plant traits*, которые необходимо предсказать

id	WORLDCLIM_BIO1_annual_mean_temperature	WORLDCLIM_BIO12_annual_precipitation
0	194715823	16.682079
1	194172209	16.196875
2	195910605	12.781527
3	194088633	19.231449
4	178419036	0.962713
...
9142	190581683	3.172599
9143	196623053	24.213839
9144	189605934	0.340694
9145	189425745	19.283831
9146	187853106	-7.709722

9147 rows × 167 columns

Рис. 3. Исходные данные (вспомогательные переменные в датафрейме)

Примечание: составлено авторами по источнику [6].

(извлечь): $X11$ – площадь листьев, $X18$ – высота растения, $X26$ – сухая масса семян [6, 7].

Для улучшения производительности модели эти данные были также нормализованы [8].

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}. \quad (1)$$

В выражении (1):

x' – нормализованное значение признака;

x – исходное значение признака;

x_{min} – минимальное значение всех x ;

x_{max} – максимальное значение всех x .

Кроме того, строковые значения id были приведены к числовому типу *int*.

Данные на рис. 2 и 3 были загружены с общедоступного ресурса по исследованию данных Kaggle.

Этап создания комбинированной модели

В качестве основы для извлечения визуальных признаков было использовано несколько предварительно обученных архитектур CNN: InceptionV3, ResNet50, VGG19. Все они были обучены на большом наборе изображений ImageNet.

Крайний слой каждой из моделей был исключен ввиду иной выполняемой задачи – классификации. Для сохранения «знаний»,

полученных при предварительном обучении сети, их оставшиеся слои были подвергнуты «заморозке». Это предотвратит изменение весов в слоях CNN при обучении объединенных моделей задаче регрессии [9].

Данные на выходе CNN-моделей имеют вид 4-мерного тензора в формате

$$V_4d = \{b \times m \times n \times k\}, \quad (2)$$

где:

b – размер батча;

m – высота тензора;

n – глубина тензора;

k – количество каналов.

Для успешного комбинирования модели на вход регрессионного блока должен поступать одномерный тензор формата:

$$V_1d = \{1 \times (b \times m \times n \times k + f)\}, \quad (3)$$

где f – число признаков после обработки вспомогательных переменных.

Следовательно, возникает необходимость конвертирования тензора (2) в тензор (3). Выполняется данное действие благодаря применению Flatten-слоя выпрямления.

Из выражения (3) понятно, что после выпрямления тензора (2) выполняется операция конкатенации по числу признаков.

Таким образом, ветвь изображений представляет собой замороженную предобученную сверточную модель со скорректированным внешним слоем.

Принцип работы комбинированной (сочувственной, гибридной) модели представлен на рис. 4.

Параметры обучения моделей приведены в табл. 1.

Метрики, использованные для оценки работы моделей:

Средняя абсолютная ошибка:

$$MAE = \frac{1}{N \cdot M} \sum_{i=1}^N \sum_{j=1}^M |y_{i,j} - \hat{y}_{i,j}|. \quad (4)$$

Среднеквадратичная ошибка:

$$MSE = \frac{1}{N \cdot M} \sum_{i=1}^N \sum_{j=1}^M (y_{i,j} - \hat{y}_{i,j})^2. \quad (5)$$

Коэффициент детерминации:

$$R^2 = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^M (y_{i,j} - \hat{y}_{i,j})^2}{\sum_{i=1}^N \sum_{j=1}^M (y_{i,j} - \bar{y}_j)^2}. \quad (6)$$

В выражениях (4), (5), (6):

N – количество наблюдений (по числу изображений $N = 9147$);

M – количество целевых переменных (в данной работе $M = 3$);

– фактическое значение j -й целевой переменной для i -го наблюдения;

– предсказанное значение j -й целевой переменной для i -го наблюдения;

– среднее значение j -й целевой переменной.

Значение коэффициента детерминации не может превышать 1. Нижний порог неограничен, но отрицательное значение свидетельствует о неработоспособности модели в контексте решаемой задачи [10].

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Для доказательства эффективности сочетания сверточной и регрессионной частей результаты извлечения признаков приведены в двух таблицах. В табл. 2 – метрики оценки, обученных задаче регрессии сверточных сетей. В табл. 3 – метрики оценки гибридных моделей с многослойным персепtronом в роли нелинейной регрессии.

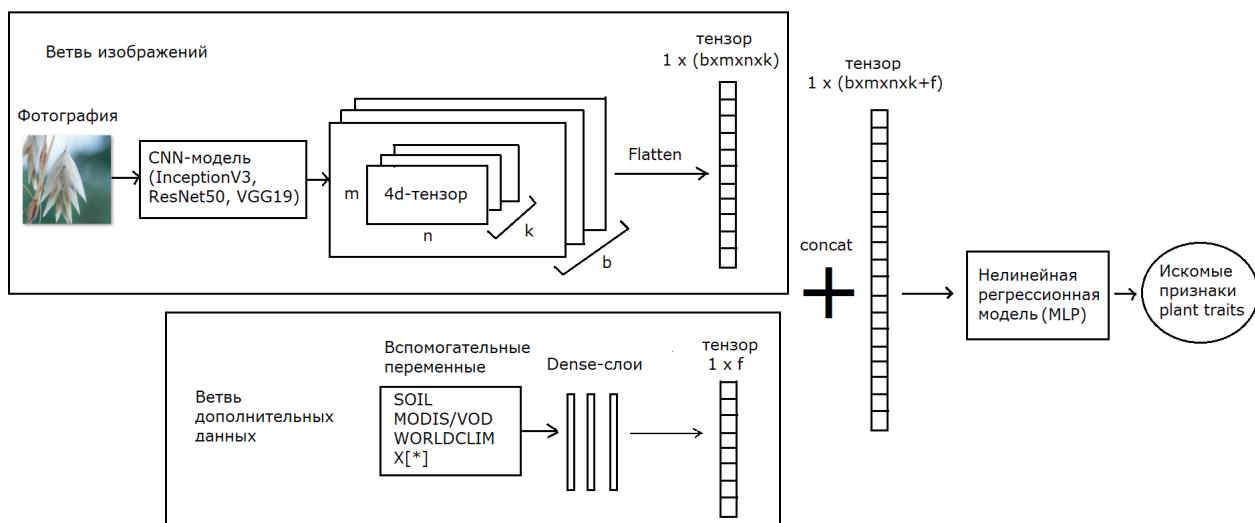


Рис. 4. Функционирование комбинированной модели

Примечание: составлено авторами.

Таблица 1

Характеристики обучения моделей

Оптимизатор/скорость обучения	Adam/0.001
Функция потерь	MSE
Количество эпох	10
Размер батча	32

Примечание: составлено авторами.

Таблица 2

Оценка мономодальных сверточных моделей

Архитектура/Метрика	InceptionV3	ResNet50	VGG19
MAE	157,5	106,9	677,1
MSE	22301,2	14301,4	214982
R ²	0,1	0,14	-2,67

Примечание: составлено авторами.

Таблица 3

Оценка многомодальных комбинированных моделей

Архитектура/Метрика	InceptionV3 + MLP	ResNet50 + MLP	VGG19 + MLP
MAE	98,7	69,9	332,2
MSE	10323,5	5770,1	94223,4
R ²	0,29	0,39	-0,95

Примечание: составлено авторами.

Из табл. 2 и 3 видно, что модель VGG-19 не подходит для исходного набора данных, скорее всего, ввиду архитектурной простоты – коэффициент детерминации отрицателен. В целом невысокое значение R^2 в таблицах объясняется большой сложностью набора исходных данных. Тем не менее эффективность рассматриваемого подхода гибридизации очевидна.

ЗАКЛЮЧЕНИЕ

В работе рассмотрен перспективный подход совмещения мультимодальных данных на примере сочетания моделей машинного обучения разной природы. Подтверждена его эффективность в решении задачи определения

признаков растений plant traits: результаты показали, что объединение преимуществ сверточных сетей (CNN) для обработки изображений и регрессионных моделей для интеграции вспомогательных данных позволяет улучшить показатели качества обучения – уменьшить величины MAE, MSE и увеличить.

Дальнейшие исследования будут направлены на совершенствование данного метода и могут включать оптимизацию гиперпараметров каждой из составляющих комбинированной модели, борьбу с переобучением, изучение исходных данных на предмет корреляции признаков и целевых переменных, а также ансамблирование.

Список источников

1. Демидчик В. В., Шашко А. Ю., Бондаренко В. Ю. и др. Феномика растений: фундаментальные основы, программно-аппаратные платформы и методы машинного обучения // Физиология растений. 2020. Т. 67, № 3. С. 227–245.
2. Гончаров С. В. Цифровое фенотипирование растений в ботанических садах // Наследие академика Н. В. Цицина: ботанические сады. Отдаленная гибридизация растений и животных : материалы Всерос. науч. конф. 03–07 июля 2023 г. г. Москва. Воронеж : Воронежский ГАУ имени Императора Петра I, 2023. С. 62–64.
3. Покровская Я. К., Каюгина С. М. Перспективы использования искусственного интеллекта в сельском хозяйстве // Достижения молодежной науки для агропромышленного комплекса : сб. материалов LVI науч.-практ. конф. студентов, аспирантов и молодых ученых. 14–18 марта 2022 г. г. Тюмень. Тюмень : ГАУ Северного Зауралья, 2022. Т. 1, ч. 4. С. 350–358.

References

1. Demidchik V. V., Shashko A. Y., Bandarenka U. Y. et al. Plant phenomics: fundamental bases, software and hardware platforms, and machine learning. *Russian Journal of Plant Physiology*. 2020;67(3):227–245. (In Russ.).
2. Goncharov S. V. Digital phenotyping of plants in botanic gardens. In: *Proceedings of All-Russian Scientific Conference with International Participation “Nasledie akademika N. V. Tsitsina: botanicheskie sady. Otdalennaia gibrizatsiia rastenii i zhivotnykh”*, July 3–7, 2023, Moscow. Voronezh: Voronezh State Agrarian University named after Emperor Peter the Great; 2023. p. 62–64. (In Russ.).
3. Pokrovskaya Ya. K., Kayugina S. M. Perspektivnye ispolzovaniya iskusstvennogo intellekta v selskom khozyaystve. In: *Proceedings of 56th Research-to-practice conference of students. postgraduates and young scientists “Dostizheniya molodezhi nauki dlya agropromyshlennogo kompleksa”*, March 14–18, 2022, Tyumen. Tyumen: Northern Trans-Ural State

4. Мишурев Н. П., Чавыкин Ю. И., Моторин О. А. Цели и задачи искусственного интеллекта в сельском хозяйстве // Управление рисками в АПК. 2021. № 3. С. 39–49.
5. Хайкин С. Нейронные сети: полный курс. 2-е изд. / пер. с англ. Н. Н. Куссуль. М. : Изд. дом «Вильямс», 2016. 1104 с.
6. PlantTraits2024 – FGVC11. URL: <https://www.kaggle.com/competitions/plantraits2024> (дата обращения: 13.09.2024).
7. Ковалев А. В., Исаева А. С. Оценка качества семян пшеницы с использованием сверточной нейронной сети // Инженерный вестник Дона. 2021. № 12. С. 206–215.
8. Предварительная обработка данных. URL: <https://scikit-learn.ru/> (дата обращения: 25.09.2024).
9. Сабитов Б. Р., Сейтказиева Н. С. Методы компьютерного зрения в задачах прогнозирования болезней растений с использованием трансферного обучения // Проблемы автоматики и управления. 2022. № 3. С. 135–144.
10. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag, 2009. 746 p.
4. Mishurov N. P., Chavykin Yu. I., Motorin O. A. Goals and objectives of artificial intelligence in agribusiness. *Agribusiness Risk Management*. 2021;(3):39–49. (In Russ.).
5. Khaykin S. Neural networks: full course. 2nd ed. N. N. Kussul, trans. Moscow: Izd. dom “Vilyams”, 2016. 1104 p.
6. PlantTraits2024 – FGVC11. URL: <https://www.kaggle.com/competitions/plantraits2024> (accessed: 13.09.2024).
7. Kovalev A. V., Isaeva A. S. Wheat seed quality assessment based on convolutional neural network algorithm. *Engineering Journal of Don.* 2021;(12):206–215. (In Russ.).
8. Predvaritelnaya obrabotka dannykh. URL: <https://scikit-learn.ru/> (accessed: 25.09.2024). (In Russ.).
9. Sabitov B. R., Seytkazieva N. S. Metody kompyuternogo zreniya v zadachakh prognozirovaniya bolezney rasteniy s ispolzovaniem transfernogo obucheniya. *Problemy avtomatiki i upravleniya*. 2022;(3):135–144. (In Russ.).
10. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag; 2009. 746 p.

Информация об авторах

В. В. Брыкин – аспирант.

С. О. Крамаров – доктор физико-математических наук, профессор.

А. Гросу – аспирант.

About the authors

V. V. Brykin – Postgraduate.

S. O. Kramarov – Doctor of Sciences (Physics and Mathematics), Professor.

A. Grosu – Postgraduate.