

## СРАВНЕНИЕ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ОБРАЗОВ ДЛЯ СЕМАНТИЧЕСКОГО АНАЛИЗА ИЗОБРАЖЕНИЙ

Е.А. Васильева

ekaterina26v@gmail.com

SPIN-код: 5396-9369

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

---

### Аннотация

Работа посвящена исследованию проблемы автономного контроля сельскохозяйственного сектора в рамках ограниченной информации об окружающем пространстве. Рассмотрена задача обнаружения угнетенных участков с помощью беспилотных летательных аппаратов, оснащенных системой технического зрения. Представлены современные алгоритмы классификации объектов, приведено описание алгоритмов распознавания образов. Выполнен сравнительный анализ алгоритмов распознавания изображений на основе линейных методов и методов с использованием градиентного бустинга. Представлены результаты тестирования эффективности работы алгоритмов. В результате анализа установлено, что алгоритм *SGDClassifier* показывает наилучшее качество распознавания на основе метрик *Precision*, *Recall* и *F-меры*.

### Ключевые слова

Техническое зрение, распознавание образов, машинное обучение, классификация изображений, нейронная сеть, алгоритм SVM, алгоритм CatBoost, алгоритм LightGBM, алгоритм *SGDClassifier*

Поступила в редакцию 21.01.2019

© МГТУ им. Н.Э. Баумана, 2019

---

**Введение.** Круг задач в области сельского хозяйства достаточно велик. Дефекты при посеве, очаговые пустоты, потеря урожая и другие факторы требуют оперативного вмешательства [1]. Большинство оценок, проводимых в подобных ситуациях, выполняют на земле силами выезжающих на поля экспертных групп. В такой ситуации невозможно оценить весь масштаб происшествий. Для полноценного решения задач в рамках ведения сельского хозяйства необходим качественный сбор и анализ данных об окружающем пространстве [2]. На сегодняшний день сбор и анализ данных осуществляются с помощью беспилотных летательных аппаратов, оснащенных системой технического зрения.

Целью данного исследования является сравнение алгоритмов, применяемых для распознавания образов, с целью улучшения разработанного алгоритма системы технического зрения, основанного на методах нейронных сетей и алгоритма SVM [3].

**Сравнительный анализ.** Для решения задач в области распознавания изображений существует множество методов [4]. К числу современных методов относятся нейронные сети, алгоритмы на основе бустинга (процедура последовательного построения композиции алгоритмов машинного обучения, когда каж-

дый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов) и бэггинга технология классификации, использующая композиции алгоритмов, каждый из которых обучается независимо). Разработанный алгоритм [3] для распознавания объектов был основан на использовании сверточной нейронной сети VGG16 [5] в качестве выделения характерных признаков в изображении и алгоритма машинного обучения SVM [6] в качестве классификатора. В рамках исследования предлагается провести сравнительный анализ алгоритмов классификации с целью выявления оптимального метода с точки зрения скорости работы и качества классификации.

Для сравнения были выбраны 4 алгоритма:

- CatBoost, разработанный командой Яндекс;
- LightGBM, разработанный командой Microsoft;
- SVM, ранее примененный в работе [3] для семантического анализа изображений;
- SGDClassifier на основе стохастического градиентного бустинга.

CatBoost — открытая программная библиотека, разработанная компанией Яндекс и реализующая уникальный патентованный алгоритм построения моделей машинного обучения, в котором используется одна из оригинальных схем градиентного бустинга [7]. Кроме того, CatBoost применяется Европейской организацией по ядерным исследованиям (CERN) для объединения данных, полученных с разных частей детектора LHCb. В основе алгоритма CatBoost лежит градиентный бустинг, т. е. метод, в котором строится серия очень слабых алгоритмов (в данном случае — решающих деревьев), последовательно минимизирующих ошибку друг друга и в итоге, в комбинации, хорошо описывающих обучающую выборку данных. Алгоритм CatBoost изначально предназначался для того, чтобы наилучшим образом работать не только с числовыми, но и с категориальными признаками (признаками, не имеющие числового выражения).

LightGBM — это алгоритм машинного обучения, разработанный командой Google на основе градиентного бустинга [8]. Суть алгоритма LightGBM заключается в «выращивании» дерева вертикально, в то время как другие алгоритмы на основе бустинга деревьев «выращивают» деревья горизонтально. Это означает, что LightGBM выращивает дерево по листьям, в то время как другой алгоритм растет по уровням. Структура алгоритма LightGBM представлена на рис. 1.

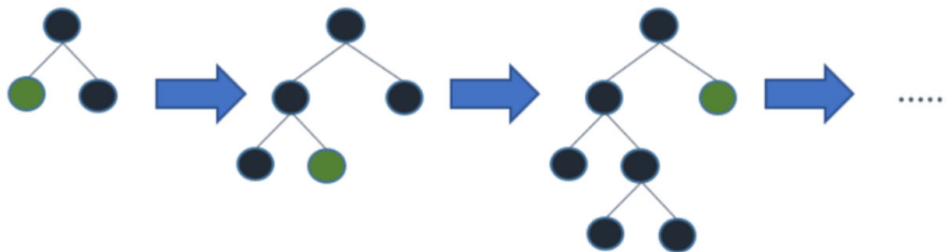


Рис. 1. Структура алгоритма LightGBM

Структура других алгоритмов на основе бустинга деревьев показана на рис. 2.

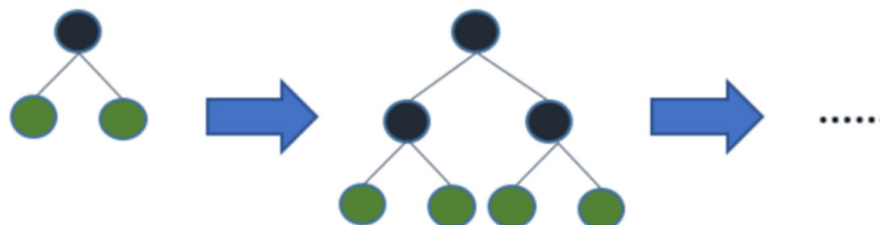


Рис. 2. Структура алгоритмов на основе градиентного бустинга

Метод опорных векторов (Support Vector Machine), или алгоритм SVM — один из методов семейства линейных классификаторов; его применяют в задачах бинарной классификации (когда объект может принадлежать одному из двух классов).

Основная идея метода — перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строят по обеим сторонам гиперплоскости, разделяющей классы. Разделяющей будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей. Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя погрешность классификатора.

Алгоритм SGDClassifier основан на алгоритме стохастического градиентного спуска [9]. Стохастический градиентный спуск относится к оптимизационным алгоритмам и нередко применяется для настройки параметров модели машинного обучения. В роли базового алгоритма может выступать алгоритм SVM или логистическая регрессия. Суть метода SGDClassifier заключается в подборе вектора весов в линейных классификаторах, т. е. в минимизации функции потерь методом стохастического градиентного спуска. На каждой итерации алгоритма вектор весов изменяется в направлении наибольшего убывания функционала (в сторону антиградиента). Таким образом, параметры модели изменяются после каждого объекта обучения. Для больших массивов данных стохастический градиентный спуск может дать значительное преимущество в скорости по сравнению со стандартным градиентным спуском. В рамках тестирования в качестве базового метода для SGDClassifier используется алгоритм SVM.

**Тестирование алгоритмов.** Для тестирования алгоритмов был подготовлен датасет, содержащий 300 изображений, 150 из них входили в обучающую выборку, 150 — в тестовую. Стояла задача классифицировать изображения по пяти классам, причем каждому классу соответствовало по 30 изображений.

Для выделения характерных признаков на изображениях используется предобученная нейронная сеть VGG16, в рамках исследования необходимо верно классифицировать полученные матрицы признаков каждого изображения. Оценками качества классификаторов выбраны метрики Precision, Recall и F-мера.

Результаты проведенного сравнительного анализа представлены в таблице.

### Результаты анализа

Алгоритм	Precision	Recall	F-мера	Время работы, с
CatBoost	0,248	0,353	0,251	119,20
LightGBM	0,941	0,940	0,940	3,86
SVM	0,948	0,949	0,947	0,36
SGDClassifier	0,967	0,967	0,968	0,07

**Выводы.** В результате проделанной работы проведен сравнительный анализ алгоритмов классификации объектов. На основе анализа алгоритм SGDClassifier с базовым методом SVM показал лучшие результаты. Данный алгоритм оказался эффективнее алгоритма обычного SVM, используемого в разработанном алгоритме для анализа семантических изображений [3], и быстрее на 0,29 с.

### Литература

- [1] Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов. М., Физматлит, 2009.
- [2] Рубцов И.В., Машков К.Ю., Наумов В.Н. 185 лет МГТУ им. Н. Э. Баумана – Состояние и перспективы развития специальной робототехники. *Сб. мат. 10 Всерос. науч.-практ. конф. Перспективные системы и задачи управления*, 2015, № 1, с. 247–255.
- [3] Васильева Е.А. Разработка алгоритма системы технического зрения для семантического анализа изображений. *Политехнический молодежный журнал*, 2018, № 9(26). DOI: 10.18698/2541-8009-2018-9-366  
URL: <http://ptsj.ru/catalog/menms/robots/366.html>
- [4] Местецкий Л.М. Математические методы распознавания образов. М., МГУ, 2002–2004.
- [5] Chen L.-C., Papandreou G., Kokkinos I., et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, vol. 40, no. 4, pp. 834–848.  
DOI: 10.1109/TPAMI.2017.2699184 URL: <https://ieeexplore.ieee.org/document/7913730>
- [6] Cristianini N., Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, 2000.
- [7] Prokhorenkova L., Gusev G., Vorobev A., et al. CatBoost: unbiased boosting with categorical features. *arxiv.org*: веб-сайт. URL: <https://arxiv.org/abs/1706.09516> (дата обращения: 01.12.2018).
- [8] Ke G., Meng Q., Finley T., et al. LightGBM: a highly efficient gradient boosting decision tree. *Proc. NIPS 2017*. URL: <https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree> (дата обращения: 01.12.2018).
- [9] Zhang T. Solving large scale linear prediction problems using stochastic gradient descent algorithms. *Proc. ICML*, 2004, p. 116. DOI: 10.1145/1015330.1015332  
URL: <https://dl.acm.org/citation.cfm?id=1015332>

**Васильева Екатерина Андреевна** — студентка кафедры «Специальная робототехника и мехатроника», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

**Научный руководитель** — Рубцов Василий Иванович, кандидат технических наук, доцент кафедры «Специальная робототехника и мехатроника», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

## THE COMPARISON OF PATTERN RECOGNITION ALGORITHMS FOR SEMANTIC IMAGE ANALYSIS

E.A. Vasileva

ekaterina26v@gmail.com

SPIN-code: 5396-9369

Bauman Moscow State Technical University, Moscow, Russian Federation

---

### Abstract

The work is devoted to the study of the problem of the agricultural sector autonomous control within the scope of limited information about the surrounding area. The problem of detecting depressed areas with the help of unmanned aerial vehicles equipped with a computer vision system is considered. The modern algorithms for the object classification have been presented; a description of pattern recognition algorithms has been made. A comparative analysis of image recognition algorithms based on linear methods and methods using gradient boosting has been performed. The results of testing the algorithms performance are presented. As a result of the analysis, it was found out that the SGDClassifier algorithm showed the best recognition quality based on the Precision, Recall and F-measure metrics.

### Keywords

Computer vision, pattern recognition, machine learning, image classification, neural network, SVM algorithm, CatBoost algorithm, LightGBM algorithm, SGDClassifier algorithm

Received 21.01.2019

© Bauman Moscow State Technical University, 2019

---

### References

- [1] Kalyaev I.A., Gayduk A.R., Kapustyan S.G. Modeli i algoritmy kollektivnogo upravleniya v gruppakh robotov [Collective direction models and algorithms in robot groups]. Moscow, Fizmatlit Publ., 2009 (in Russ.).
- [2] Rubtsov I.V., Mashkov K.Yu., Naumov V.N. [185-anniversary of Bauman MSTU: current state and future prospects of special robotics development]. *Sb. mat. 10 Vseros. nauch.-prakt. konf. Perspektivnye sistemy i zadachi upravleniya* [Proc. 10<sup>th</sup> Russ. Sci.-Pract. Conf. Future systems and control problems], 2015, no. 1, pp. 247–255 (in Russ.).
- [3] Vasil'yeva E.A. Developing the algorithm of the machine vision system for the images semantic analysis. *Politekhnikheskiy molodezhnyy zhurnal* [Politekhnikheskiy molodezhnyy zhurnal [Politechnical student journal], 2018, no. 9(26). DOI: 10.18698/2541-8009-2018-9-366 URL: <http://ptsj.ru/catalog/menms/robots/366.html> (in Russ.).
- [4] Mestetskiy L.M. Matematicheskie metody raspoznavaniya obrazov [Mathematical methods of image recognition]. Moscow, MSU Publ., 2002–2004 (in Russ.).
- [5] Chen L.-C., Papandreou G., Kokkinos I., et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, vol. 40, no. 4, pp. 834–848. DOI: 10.1109/TPAMI.2017.2699184 URL: <https://ieeexplore.ieee.org/document/7913730>
- [6] Cristianini N., Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, 2000.
- [7] Prokhorenkova L., Gusev G., Vorobev A., et al. CatBoost: unbiased boosting with categorical features. *arxiv.org*: website. URL: <https://arxiv.org/abs/1706.09516> (accessed: 01.12.2018).

- [8] Ke G., Meng Q., Finley T., et al. LightGBM: a highly efficient gradient boosting decision tree. *Proc. NIPS 2017*. URL: <https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree> (accessed: 01.12.2018).
- [9] Zhang T. Solving large scale linear prediction problems using stochastic gradient descent algorithms. *Proc. ICML, 2004*, p. 116. DOI: 10.1145/1015330.1015332  
URL: <https://dl.acm.org/citation.cfm?id=1015332>

**Vasileva E.A.** — Student, Department of Robotics and Mechatronics, Bauman Moscow State Technical University, Moscow, Russian Federation.

**Scientific adviser** — V.I. Rubtsov, Cand. Sc. (Eng.), Assoc. Professor, Department of Robotics and Mechatronics, Bauman Moscow State Technical University, Moscow, Russian Federation.