

Сравнительный анализ данных Sentinel и MODIS в задаче классификации сельскохозяйственных культур

А.Ю. Денисова¹, А.В. Чернов²

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34а, Самара, Россия, 443086

Аннотация

Настоящая работа посвящена вопросу использования данных различного разрешения Sentinel (10 м) и MODIS (250 м) для классификации сельскохозяйственных культур. На примере классификации временных рядов вегетационного индекса NDVI методом случайного леса (Random Forest) рассматривается влияние источника данных на результат классификации для полей различной площади. В итоге экспериментально установлено, что данные Sentinel дают сравнимую с данными MODIS вероятность верной классификации для любых размеров полей, причем для полей размера менее 50 Га классификация по данным Sentinel чуть лучше классификации по данным MODIS. Результаты говорят о том, что высокая частота получения снимков MODIS позволяет получить схожую вероятность классификации при меньшем пространственном разрешении. Тем не менее для полей малых размеров остается предпочтительным применение классификатора на основе данных Sentinel.

Ключевые слова

Классификация культур, Sentinel, MODIS, Random Forest

1. Введение

Данные дистанционного зондирования Земли (ДЗЗ) MODIS [1] и Sentinel [2] являются наиболее широко используемыми в силу своей открытости и доступности. Снимки MODIS обладают низким пространственным разрешением 250 м и высокой частотой получения, тогда как данные Sentinel поступают в два раза реже, но имеют значительно более высокое пространственное разрешение 10 м. В связи с чем возникает закономерный вопрос как соотносятся данные тематической обработки, получаемые по обоим источникам данных, в тех или иных задачах и целесообразна ли замена данных низкого разрешения данными более высокого разрешения, которые сложнее в обработке.

В рамках настоящего исследования на примере задачи классификации культур рассматривается вопрос о влиянии различного пространственного и временного разрешения исходных данных на результаты классификации.

2. Метод и результаты исследования

Для исследования был использован объектно-ориентированный подход на основе классификации методом Random Forest (RF) [3] временных рядов вегетационного индекса NDVI [4]. Объектный подход заключался в расчёте среднего NDVI для поля в границах векторной маски границ поля. Расчитанные в период с 1 апреля по 1 сентября 2020 г. значения NDVI для каждого поля объединялись во временной ряд путём линейной интерполяции пропущенных значений. Полученные временные ряды классифицировались алгоритмом RF.

Для расчёта NDVI использовались данные продукта MOD09GQ [5] с пространственным разрешением 250 м, и данные NDVI, с портала Sentinel-Hub [6], с пространственным разрешением 10 м.

Исследование производилось на выборке полей, задекларированных крупнейшими субсидируемыми сельхозтоваропроизводителями Самарской области, по данным за 2020 год

из системы ГИС АПК. Выборка была разделена на две части: малые поля площадью от 30 до 50 Га – 116 полей, и большие поля площадью более 50 Га – 1280 полей. Классификация производилась на четыре класса: озимые, пар, поздние яровые, ранние яровые. Априорные вероятности каждого из классов составили 0,26, 0,18, 0,33, 0,23 соответственно.

На выборке больших полей методом кросс-валидации с количеством итераций 10 были подобраны оптимальные параметры классификатора RF. Для экспериментов использовалась реализация классификатора в пакете scikit-image [7]. Оптимальная глубина деревьев классификации составила 8, а оптимальное количество деревьев составило 100 и 50 для Sentinel и MODIS соответственно. В итоге в экспериментах количество деревьев было выбрано как среднее от оптимальных значений и было равно 75. Для описанных параметров классификатора и выборок полей было произведено по 10 классификаций с расчетом средней вероятности верной классификации по всем классам. Результаты приведены в таблице 1.

Таблица 1

Вероятность верной классификации по данным Sentinel и MODIS

Данные	Малые поля	Большие поля
Sentinel	0,855	0,908
MODIS	0,841	0,936

Видно, что для больших полей вероятность верной классификации выше для данных MODIS, а для малых полей – для данных Sentinel. Следует отметить, что выбранное нижнее ограничение по площади малых полей в 30 Га соответствует 4-5 пикселям по данным MODIS, следовательно, поля меньшего размера не могут быть оценены по данным низкого разрешения. Тогда как данные Sentinel позволяют производить расчёты и для полей площадью менее чем 30 Га.

3. Заключение

Проведенное исследование показало, что для больших полей, площадью более 50 Га, вероятность верной классификации чуть выше для данных MODIS, а для малых полей – для данных Sentinel. Причем данные Sentinel позволяют произвести расчёты и для полей площадью менее 30 Га. Таким образом, для полей малых размеров остается предпочтительным применение классификатора на основе данных Sentinel.

4. Литература

- [1] Sakamoto, T. A crop phenology detection method using time-series MODIS data / T. Sakamoto, M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, H. Ohno // Remote sensing of environment. – 2005. – Vol. 96(3-4). – P. 366-374.
- [2] Veloso, A. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications / A. Veloso, S. Mermoz, A. Bouvet, T. Le Toan, M. Planells, J.F. Dejoux, E. Ceschia // Remote sensing of environment. – 2017. – Vol. 199. – P. 415-426.
- [3] Pal, M. Random forest classifier for remote sensing classification / M. Pal // International journal of remote sensing. – 2005. – Vol. 26(1). – P. 217-222.
- [4] Carlson, T.N. On the relation between NDVI, fractional vegetation cover, and leaf area index / T.N. Carlson, D.A. Ripley // Remote sensing of Environment. – 1997. – Vol. 62(3). – P. 241-252.
- [5] MODOGQ v006 [Электронный ресурс]. – Режим доступа: <https://lpdaac.usgs.gov/products/mod09gqv006/> (19.01.2020).
- [6] Cloud API for satellite imagery [Электронный ресурс]. – Режим доступа: <https://www.sentinel-hub.com/> (19.01.2020).
- [7] Scikit-image. Image processing in python [Электронный ресурс]. – Режим доступа: <https://scikit-image.org/> (19.01.2020).