

## RESEARCH ON CROP CLASSIFICATION BASED ON SENTINEL-2 IMAGE USING SVM

Zhao B.<sup>1</sup>, Cao Q.<sup>2</sup>

<sup>1</sup>Zhao Bochao – graduate student;

<sup>2</sup>Cao Jiawei – master's student,

FACULTY OF GEOGRAPHY AND GEOINFORMATICS,  
BELARUSIAN STATE UNIVERSITY,  
MINSK, REPUBLIC OF BELARUS

**Abstract:** this paper uses Sentinel-2 image data, combined with ground survey data, to study the use of SVM model to classify crops in the Zhdanovich region of Belarus. First, we performed image preprocessing, including steps such as geometric correction, atmospheric correction, and pixel resampling, and then extracted image features for classification. And analyze the characteristics of crops based on the time series NDVI. In the SVM model classification, we choose the radial basis function kernel (RBF) as the kernel function, and perform parameter tuning through cross-validation. The results show that the SVM model can effectively classify Zhdanovich farmland, and the recognition accuracy is as high as 95%.

**Keywords:** sentinel-2 image; crop classification; time series; NDVI; SVM model.

## ИССЛЕДОВАНИЕ КЛАССИФИКАЦИИ КУЛЬТУР НА ОСНОВЕ ИЗОБРАЖЕНИЯ SENTINEL-2 С ИСПОЛЬЗОВАНИЕМ SVM

Чжао Б.<sup>1</sup>, Цао Ц.<sup>2</sup>

<sup>1</sup>Чжао Бочао – аспирант;

<sup>2</sup>Цао Цзявэй – магистрант,

факультет географии и геоинформатики,  
Белорусский государственный университет,  
г. Минск, Республика Беларусь

**Аннотация:** в этой статье используются данные изображений Sentinel-2 в сочетании с данными наземной съемки для изучения использования модели SVM для классификации сельскохозяйственных культур в Ждановичском районе Беларуси. Сначала мы выполнили предварительную обработку изображения, включая такие этапы, как геометрическая коррекция, атмосферная коррекция и повторная выборка пикселей, а затем извлекли признаки изображения для классификации. И анализировать характеристики сельскохозяйственных культур на основе временного ряда NDVI. В классификации модели SVM мы выбираем ядро радиальной базисной функции (RBF) в качестве функции ядра и выполняем настройку параметров посредством перекрестной проверки. Результаты показывают, что модель SVM может эффективно классифицировать сельхозугодья Ждановича, а точность распознавания достигает 95%.

**Ключевые слова:** изображение Sentinel-2; классификация культур; Временные ряды; НДВИ; Модель SVM.

В последние годы, с непрерывным развитием технологий дистанционного зондирования, изображения дистанционного зондирования широко используются в области сельского хозяйства. Классификация культур является одним из важных сценариев применения. В данной работе изучается метод классификации посевов по снимкам Sentinel-2, а в качестве объекта исследования берется посевная площадь г. Жданович, Беларусь.

1. Сбор данных и методы исследования.

1.1 Сбор и предварительная обработка данных.

Во-первых, мы собирали изображения Sentinel-2 Ждановичского района Беларуси в течение 2017 года и использовали программное обеспечение ENVI для предварительной обработки изображений, включая геометрическую коррекцию, атмосферную коррекцию и передискретизацию пикселей. И вырезаем область исследования, и вырезаем основную площадь урожая. Результаты предварительной обработки следующие:

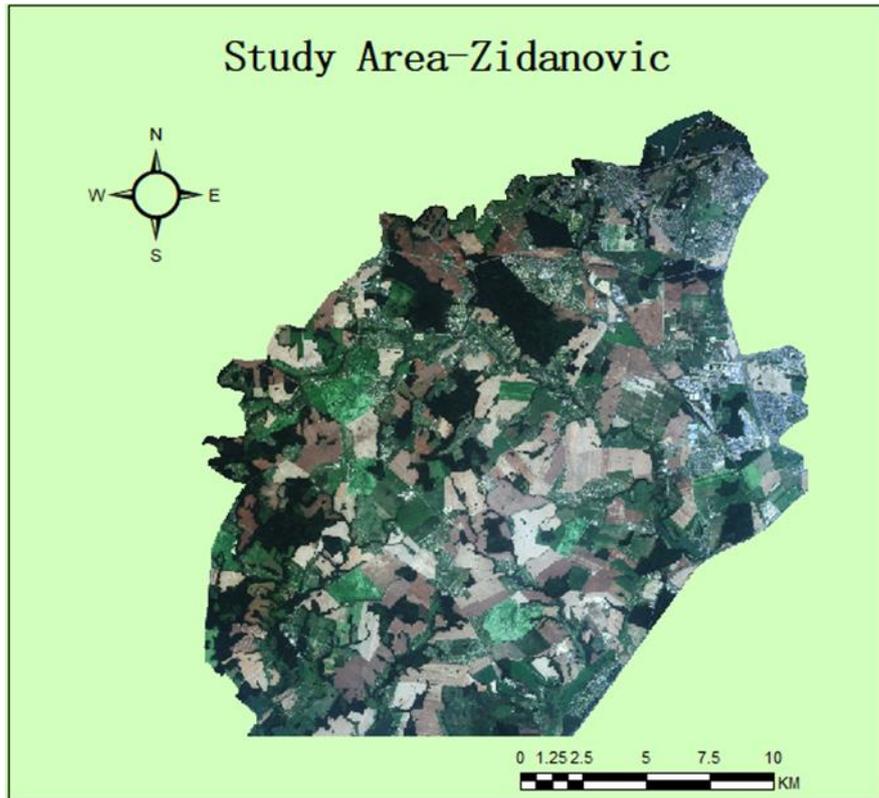


Рис. 1. Область исследования.

#### 1. NDVI на основе времени

Нормализованный разностный индекс растительности (NDVI) дает количественную оценку растительности путем измерения разницы между ближним инфракрасным светом (который сильно отражается растительностью) и красным светом (который поглощается растительностью). NDVI всегда находится в диапазоне от -1 до +1. Но четких границ для каждого типа наземного покрова нет. Когда значение отрицательное, это, скорее всего, вода. С другой стороны, если значение NDVI ближе к +1, это, скорее всего, густая зеленая листва. Однако, когда NDVI близок к нулю, зеленых листьев нет и, возможно, даже урбанизированных территорий [20].

Нормализованный разностный индекс растительности (NDVI) использует в своей формуле как NIR, так и красный канал. Как показано ниже:

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad (1.1)$$

Посредством последовательного анализа NDVI можно лучше выделить классификационные характеристики различных культур, а кривая NDVI показана на рисунке ниже:

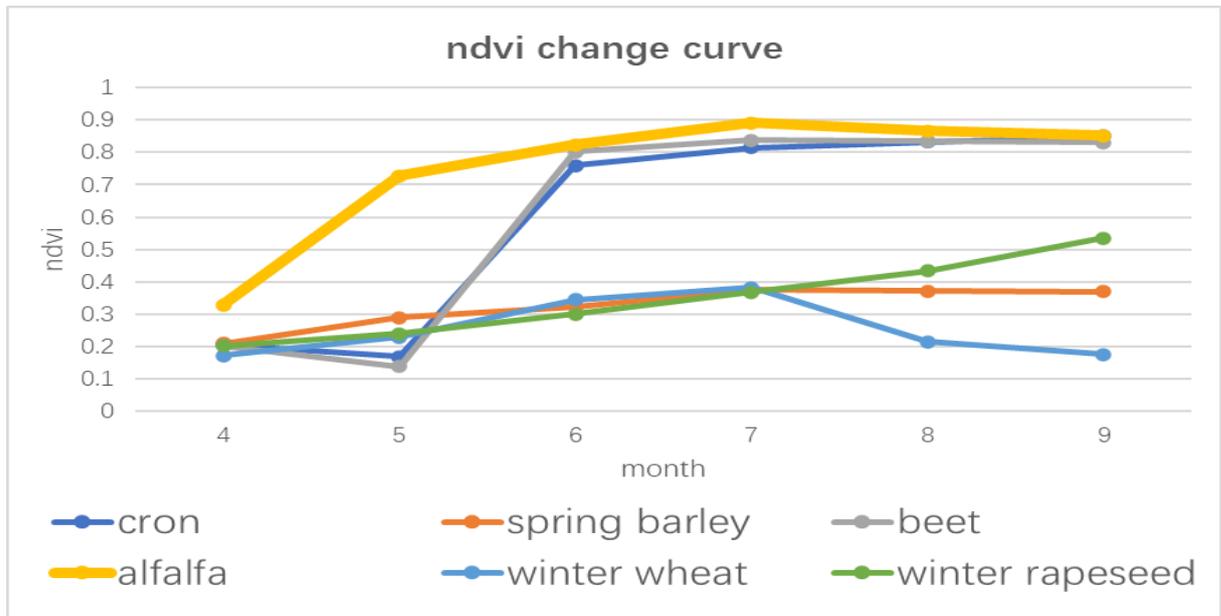


Рис. 2. График изменения NDVI.

## 2. Метод классификации - SVM.

Поскольку основным исследовательским содержанием этой статьи являются сельскохозяйственные культуры, контролируемая классификация сельскохозяйственных культур в экспериментальной зоне устранит влияние других факторов на результаты классификации, что может значительно повысить точность классификации. В сочетании с характеристиками испытательного участка и на основе анализа кривых временных рядов NDVI для типичных культур и, наконец, с использованием метода классификации SVM точность классификации культур может быть значительно улучшена.

Машина опорных векторов (SVM) — это своего рода обобщенный линейный классификатор (обобщенный линейный классификатор), который выполняет бинарную классификацию данных в соответствии с обучением с учителем, а его граница решения — максимальный запас для решения обучающих выборок Гиперплоскость (гиперплоскость с максимальным запасом).

SVM — это дискриминативный классификатор, определяемый гиперплоскостью классификации. То есть, учитывая набор размеченных обучающих выборок, алгоритм выведет оптимальную гиперплоскость для классификации новых выборок (тестовых выборок). Как показано на рисунке ниже, видно, что существует несколько прямых линий, разделяющих два типа координат. Есть ли лучший из них? Мы можем интуитивно определить следующее правило: если сегментированная линия находится слишком близко к координатным точкам, она не оптимальна. Потому что он будет чувствителен к шуму и не сможет правильно обобщить. Поэтому наша цель — найти разделительную линию, максимально удаленную от всех точек выборки. Алгоритм SVM заключается в поиске гиперплоскости, причем расстояние от нее до ближайшей обучающей выборки должно быть наибольшим. То есть оптимальная гиперплоскость сегментации максимизирует границы обучающей выборки.

Алгоритм SVM обладает хорошей вычислительной эффективностью, надежностью и статистической стабильностью. Классификатор SVM отличается простотой, стабильностью и высокой точностью классификации и может решать такие проблемы классификации, как многомерность, небольшой размер выборки и неопределенность.

В программном обеспечении ENVI функция ядра радиального базиса SVM используется в качестве функции ядра классификации, и эффект классификации является лучшим при условии штрафного коэффициента  $c = 100$  и  $\gamma = 0,25$ .

### 2. Результаты и анализ

#### 2.1 Результаты классификации

Через NDVI на основе временных рядов алгоритм SVM используется для классификации изображения, и окончательный результат вывода выглядит следующим образом:

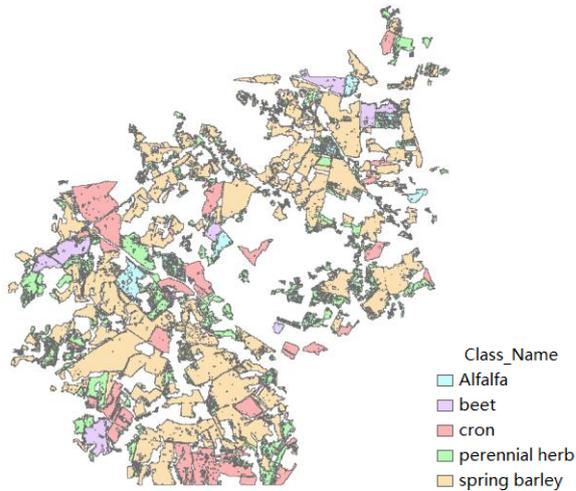


Рис. 3 Карта распределения урожая весной 2017 г.

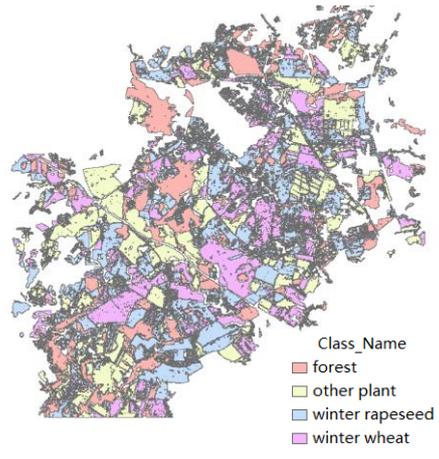


Рис. 4. Карта распределения урожая осенью 2017 г.

Из рисунка видно, что классификация культур в основном завершена, но в кромочных деталях еще много заусенцев, которые нуждаются в дальнейшей оптимизации. Обычно используемые методы постклассификационной обработки включают анализ большинства/меньшинства, кластерную обработку (сгусток) и фильтрацию (решето). Фильтр режима обычно называют фильтром режима или фильтром формы сигнала, и он хорошо влияет на устранение помех.

Сравнительная таблица до и после окончательной классификации и обработки выглядит следующим образом:

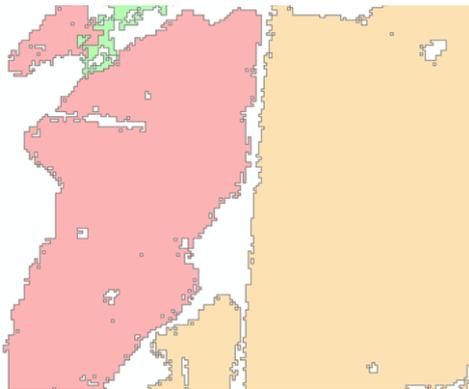


Рис. 5. До обработки.

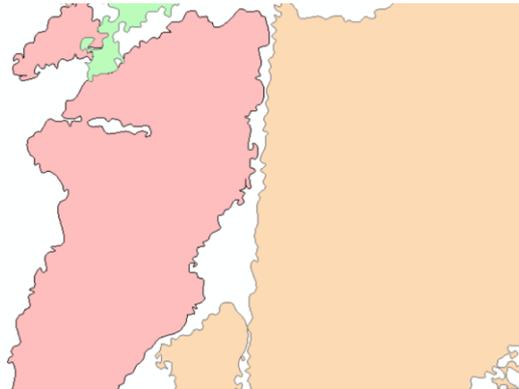


Рис.6. После обработки.

После классификации и постобработки устраняются небольшие пятна и дыры изображения, а также значительно оптимизируются свои по краям. Окончательная сгенерированная тематическая карта выглядит следующим образом:

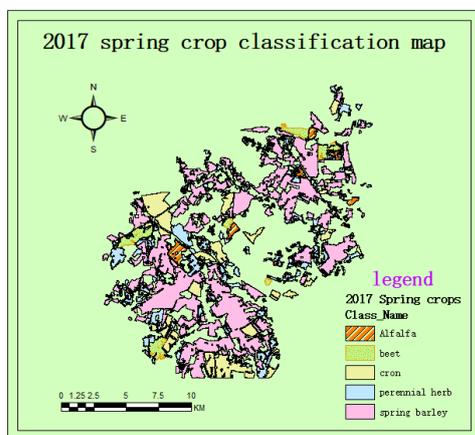


Рис. 7. Карта классификации культур весной 2017 г.

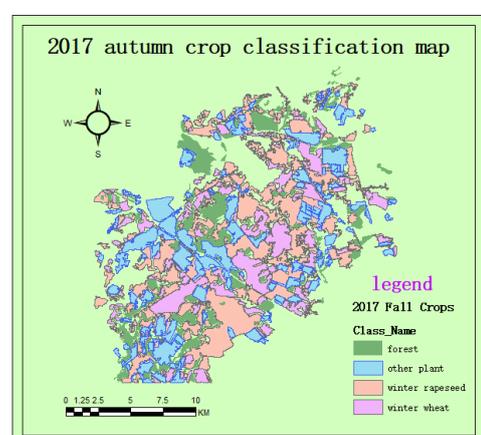


Рис. 8. Карта классификации культур осени 2017 г.

## 2.2 Анализ точности

Таблица 1. Анализ точности.

<b>Общая точность = (114621/116582) 98,3179%</b>				
<b>Коэффициент Каппа = 0,9798</b>				
<b>Сорт</b>	<b>Производ.соотв. (Процент)</b>	<b>Учетная запись пользователя. (Процент)</b>	<b>Производ.соотв. (Пиксели)</b>	<b>Аккаунт пользователя (Пиксели)</b>
хрон	98.68	100.00	1274/1291	1274/1274
свекла	99.69	100.00	13682/13725	13682/13682
Люцерна	96.57	96.02	15188/15727	15188/15818
сад	97.38	97.39	9435/9689	9435/9688
многолетняя ее	95.11	95.85	17934/18857	17934/18710
яровой ячмень	99.67	99.52	26813/26902	26813/26941
лес	99.7	100	25335/25412	25335/25335
строить	99.62	96.61	4960/4979	4960/5134
озимая пшеница	95.46	99	40214/42125	40214/40622
зимний рапс	96.51	92.2	30000/31084	30000/32537

Из результатов таблицы точности можно узнать, что общая точность классификации на основе пикселей составляет более 97% с использованием данных временного ряда NDVI и алгоритма контролируемой классификации SVM, чего достаточно, чтобы доказать, что экспериментальная модель является более точной, чем традиционная контролируемая классификация в классификации культур. Выше, традиционная контролируемая классификация, общая точность классификации составляет около 80%, благодаря анализу NDVI на основе временных рядов в сочетании с характеристиками алгоритма SVM точность классификации сельскохозяйственных культур значительно улучшен. Кроме того, фактический системный эффект в основном согласуется с фактическими данными образца. Хотя на границе есть заусенцы и небольшие пятна изображения, влияние неблагоприятных факторов в основном устраняется за счет обработки после классификации. Это показывает, что данные результатов классификации на основе изображений дистанционного зондирования имеют очень высокую согласованность с данными полевых испытаний, что полностью соответствует требованиям точности. Подводя итог, в этой статье в качестве примера рассматривается сельскохозяйственный район Жданнович в Беларуси для изучения метода использования изображений Sentinel-2 для классификации сельскохозяйственных культур и использования модели SVM для классификационного анализа. Результаты показывают, что модель SVM может эффективно классифицировать различные типы культур, повышать уровень точного управления в сельскохозяйственном поле, и ожидается, что она будет популяризирована и применена в реальном производстве.

*Список литературы / References*

1. *Liu Yansui, Lu Daodao.* Основные тенденции и региональные эффекты реструктуризации сельского хозяйства в Китае // Географический журнал, 2003, № 03, с. 381-389.
2. *Wang Wensheng, Guo Leifeng.* Сельскохозяйственные большие данные и их перспективы применения // Журнал сельскохозяйственных наук Цзянсу, 2015, т. 43, № 09, с. 1-5.
3. *Chen Yangfen, Li Xiande.* Пространственные и временные паттерны и факторы изменения урожайности зерна в Китае // Журнал сельскохозяйственной инженерии, 2013, т. 29, № 20, с. 1-10.
4. *Song X., Hu Q., Lu Miao и др.* Исследование направления развития картографирования пространственного распределения сельскохозяйственных культур на основе дистанционного зондирования // Ресурсы и зонирование сельского хозяйства в Китае, 2020, т. 41, № 06, с. 57-65.
5. *Son N.T., Chen C.F., Chen C.R. и др.* Классификация многоцветных данных Sentinel-2 для мониторинга методов выращивания риса на полевом уровне в Тайване // Прогрессивные исследования космического пространства, 2019, т. 65, № 08, с. 1910-1921.
6. *Yu Wanwan, Xu Kaijian, Zhao Ping и др.* Влияние спектральных сегментов красного края изображений Sentinel-2 на определение доминирующих видов деревьев в регионах с разными периодами роста [J]. География и геоинформационная наука, 2021, т. 37, № 03, с. 42-49.
7. *Guo Yunkai, Liu Yuling, Xu Min и др.* Моделирование и анализ индекса красного края для оценки индекса площади листов растительности // Измерение и картография, 2021, т. 46, № 01, с. 93-98.

8. *Xu Qingyun, Yang Guijun, Long Huiling и др.* Выявление сельскохозяйственных культур в провинции Шаньдун на основе анализа фенологических данных многолетних временных рядов// Интеллектуальная автоматика и мягкие вычисления, 2013, т. 19, № 4, с. 513-523.
9. *Gu Xiaohu, Han Lijian, Wang Jihua и др.* Оценка площади посева кукурузы с использованием низкого и среднего разрешений метода вейвлет-слияния дистанционного зондирования// Журнал сельскохозяйственной инженерии, 2012, т. 28, № 03, с. 203-209.
10. *Mark E.J., David R.L., Jude H.K.* Идентификация культур с использованием гармонического анализа временных рядов данных NDVI AVHRR // Компьютеры и электроника в сельском хозяйстве, 2002, т. 37, с. 127-139.