

Разработка приложения для поддержки картографирования поверхностного покрова с использованием платформы Google Earth Engine

Фам Данг Хиен (МИИГАиК, Москва)

ВВЕДЕНИЕ

Современное ДЗЗ сталкивается с "информационным потопом", что делает традиционные локальные вычисления неэффективными. Облачные платформы, такие как Google Earth Engine (GEE), предоставляют колоссальные вычислительные мощности для анализа в планетарном масштабе.

Проблема: Точность классификации LULC сильно зависит от качества и репрезентативности обучающих данных. Ручной сбор выборок трудоемок, дорог и подвержен субъективизму эксперта, часто не охватывая всего спектрального разнообразия.

Цель: Разработать Виртуальную Исследовательскую Среду (VRE) на GEE с инновационной функцией алгоритмической поддержки процесса сбора и уточнения обучающих выборок.

Source code:

<https://code.earthengine.google.com/eb7268d3d6b99521156b242c1d581c15>



Ссылка на веб-приложение

<https://lopmaybay-gee.projects.earthengine.app/view/lulc-map>

Панель Управления:
Настройка всех параметров: выбор временных интервалов, источников данных (через поиск по каталогу GEE), каналов, алгоритмов, определение количества и характеристик классов.

Панель Отчетов:
Вывод системных сообщений, генерация подробных отчетов, включающих статистику по площадям классов и метрики точности (общая точность, матрица ошибок).

Основное Окно Карты:
Интерактивная визуализация данных, рисование и редактирование полигонов для обучающих выборок, задание границ исследуемой территории (ROI).

Настройка источников данных
↓
Определение временных параметров
↓
Задание исследуемой территории (ROI)
↓
Определение классов LULC
↓
Первичный сбор обучающих выборок
↓
Генерация вспомогательных карт
↓
Итеративное уточнение выборки
↓
Финальная классификация
↓
Анализ результатов

Легенда

Схема 1. Рабочий процесс в приложении

Рис. 1. Интерфейс приложения «Классификация землепользования»

≤5 обучающих точек/класс

30~60 обучающих точек/класс

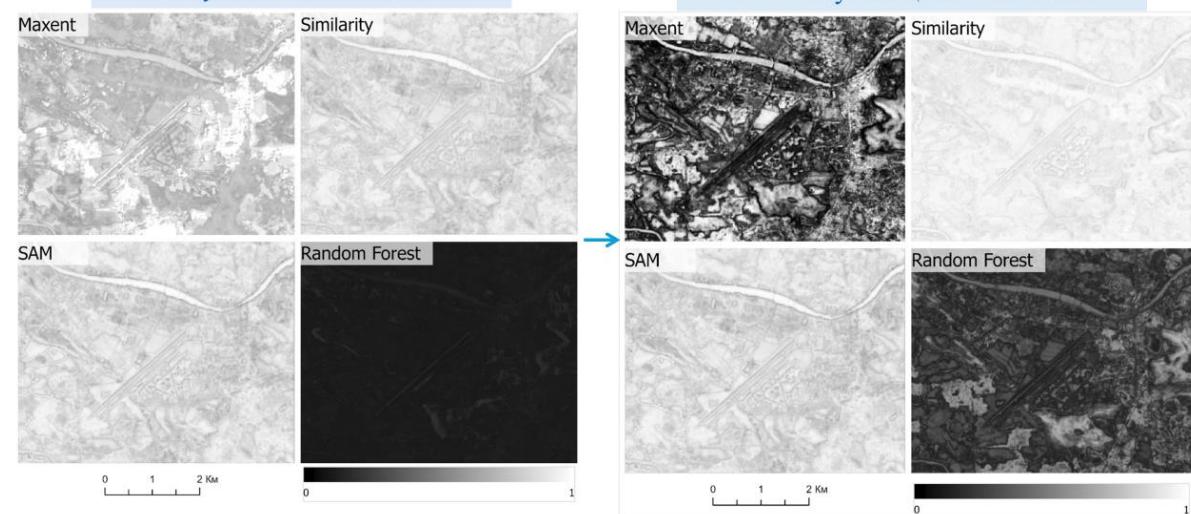


Рис. 2. Вспомогательные карты «Максимальной вероятности» 4 алгоритмов для фрагмента снимки AlphaEarth Foundations в случаях, когда каждый класс имеет 1-5 и 30-60 обучающих точек.

ЗАКЛЮЧЕНИЕ И ПЕРСПЕКТИВЫ

Приложение является ценным инструментом, который, благодаря итеративному рабочему процессу с алгоритмической поддержкой, обеспечивает значительный рост точности итоговых карт LULC.

AlphaEarth Foundations предоставляет готовый, более информативный и дискриминационный набор признаков, который превосходит традиционные подходы, особенно в условиях ограниченного количества обучающих данных.

Преодоление ограничений GEE: Для работы с большими территориями, используются стратегии масштабирования ROI (работа на репрезентативном участке) и управления параметром масштаба для предотвращения ошибок "User memory limit exceeded".

Будущее: дальнейшее развитие приложения может включать более глубокую интеграцию с **спутниковыми эмбеддингами** (satellite embeddings) фундаментальных моделей, как AlphaEarth Foundations.

Алгоритм	Особенности	Формулы
SAM (Spectral Angle Mapper)	Физически обоснованный метод. Измеряет угол между спектральными векторами, инвариантен к изменениям освещенности (яркости). Использует "средний спектральный вектор класса" для расчета и создания изображения максимального сходства.	$\alpha = \cos^{-1} \frac{\sum_{i=1}^n t_i r_i}{\sqrt{\sum_{i=1}^n t_i^2} \sqrt{\sum_{i=1}^n r_i^2}}$ t — вектор спектральных значений пикселя изображения, r — эталонный спектральный вектор, n — количество каналов.
Расчет сходства на основе скалярного произведения	Измеряет не нормализованное скалярное произведение векторов. Чувствителен к яркости. Для каждого пикселя выбирается максимальное скалярное произведение с эталонными спектрами внутри класса, что позволяет учесть внутриклассовую изменчивость. Используется для создания тематического изображения с классом наивысшего сходства.	сходство = $\sum_{i=1}^n t_i \cdot r_i$ t _i - значение i-го спектрального канала пикселя, r _i - значение i-го канала эталонного спектра, n - количество спектральных каналов.
Maxent (Максимальная энтропия)	Изначально для моделирования экологических ниш. Используется для вычисления вероятности принадлежности пикселя к классу. Предоставляет вероятностные оценки достоверности. Используется для оценки достоверности (вероятности) и поиска сложных ландшафтов.	$P(c x) = \frac{1}{Z(x)} \exp \left(\sum_{i=1}^n \lambda_{i,c} f_{i,c}(x) \right)$ x - вектор спектральных признаков пикселя $\lambda_{i,c}$ - весовые коэффициенты для i-го признака класса *c* $f_{i,c}(x)$ - функции признаков (значения спектральных каналов) $Z(x)$ - нормирующая константа, обеспечивающая $\sum P(c x) = 1$
Случайный лес (Random Forest)	Ансамблевый метод, строящий множество решающих деревьев. Вероятностный вывод позволяет идентифицировать проблемные участки со смешанной сигнатурой, то есть является более чутким к изменениям в обучающей выборке. Используется для вычисления «Максимальной вероятности классификации» для поддержки сбора выборок.	$P(c x) = \frac{1}{K} \sum_{k=1}^K I(T_k(x) = c)$ $T_k(x)$ - прогноз k-го дерева ансамбля K (равен 100 по умолчанию) - количество деревьев в ансамбле