**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ**

**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук

Департамент программной инженерии

|  |  |
| --- | --- |
| СОГЛАСОВАНО  Научный руководитель  доцент департамента  программной инженерии  факультета компьютерных наук,  канд. техн. наук  Родригес Залепинос Р.А.  **«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  2018 г. | УТВЕРЖДЕНО  Академический руководитель  образовательной программы  «Программная инженерия»  профессор департамента программной инженерии, канд. техн. наук  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** В.В. Шилов  **«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  2018 г. |

|  |  |
| --- | --- |
| **Подп. и дата** |  |
| **Инв. № дубл.** |  |
| **Взам. инв. №** |  |
| **Подп. и дата** |  |
| **Инв. № подл** |  |

**ПРОГРАММА ОБНАРУЖЕНИЯ ИЗМЕНЕНИЙ ЗЕМЛЕПОЛЬЗОВАНИЯ ПО МУЛЬТИСЕНОСОРНЫМ СПУТНИКОВЫМ ДАННЫМ**

**Пояснительная записка**

**ЛИСТ УТВЕРЖДЕНИЯ**

**RU.17701729.04.16 81 01-1-ЛУ**

|  |  |
| --- | --- |
|  | Исполнитель:  студент группы БПИ152  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** /А.А. Лукин/  **«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  2018 г. |

УТВЕРЖДЕН

RU.17701729.04.14 81 01-1-ЛУ

**ПРОГРАММА ОБНАРУЖЕНИЯ ИЗМЕНЕНИЙ ЗЕМЛЕПОЛЬЗОВАНИЯ ПО МУЛЬТИСЕНСОРНЫМ СПУТНИКОВЫМ ДАННЫМ**

|  |  |
| --- | --- |
| **Подп. и дата** |  |
| **Инв. № дубл.** |  |
| **Взам. инв. №** |  |
| **Подп. и дата** |  |
| **Инв. № подл** |  |

**Пояснительная записка**

**RU.17701729.04.16 81 01-1**

**Листов 29**

# **АННОТАЦИЯ**

В данном документе приведена пояснительная записка для «Программы обнаружения изменений землепользования по мультисенсорным спутниковым данным». Данная программа предназначена для обнаружения изменений землепользования по спутниковым снимкам.

Оформление программного документа «Пояснительная записка» произведено по требованиям ГОСТ 19.404-79 «Пояснительная записка. Требования к содержанию и оформлению» [1].

**Содержание**

[**АННОТАЦИЯ** 2](#_Toc514032378)

[**1.** **Введение** 4](#_Toc514032379)

[**1.1. Наименование разработки** 4](#_Toc514032380)

[**2.** **Назначение и область применения** 5](#_Toc514032381)

[**2.1. Назначение программного модуля** 5](#_Toc514032382)

[**2.2. Область применения программы** 5](#_Toc514032383)

[**3.** **Технические характеристики** 6](#_Toc514032384)

[**3.1. Постановка задачи для разработки программы** 6](#_Toc514032385)

[**3.2. Описание алгоритмов и функционирования программы** 6](#_Toc514032386)

[**3.2.1. Поиск и загрузка данных Sentinel 2** 6](#_Toc514032387)

[**3.2.2. Преобразование данных Sentinel 2 Level 1C к Sentinel 2 Level 2A** 8](#_Toc514032388)

[**3.2.3. Работа с мультисенсорными сценами Sentinel-2 Level 2A** 9](#_Toc514032389)

[**3.2.3. Модуль обучения классификатора SVM** 11](#_Toc514032390)

[**3.2.4. Определение изменений землепользования по мультисенсорным спутниковым данным Sentinel-2 с использованием классификатора SVM** 15](#_Toc514032391)

[**3.3. Описание и обоснование выбора метода организации входных и выходных данных** 19](#_Toc514032392)

[**3.3.1. Описание метода организации входных и выходных данных** 19](#_Toc514032393)

[**3.4. Описание и обоснование выбора состава технических и программных средств** 20](#_Toc514032394)

[**3.4.1. Состав технических и программных средств** 20](#_Toc514032395)

[**3.4.2.** **Обоснование выбора технических и программных средств** 20](#_Toc514032396)

[**4.** **Ожидаемые технико-экономические показатели** 22](#_Toc514032397)

[**4.1. Предполагаемая потребность** 22](#_Toc514032398)

[**4.2. Ориентировочная экономическая эффективность** 22](#_Toc514032399)

[**5.** **Список использованной литературы** 23](#_Toc514032400)

[**Приложение 1 Терминология** 25](#_Toc514032401)

[**Приложение 2 Описание и функциональное назначение файлов** 26](#_Toc514032402)

# **Введение**

## **1.1. Наименование разработки**

Наименование программного продукта – «Программа обнаружения изменений землепользования по мультисенсорным спутниковым данным».

Наименование программного продукта на английском языке – «Program for Land-Use Change Detection using Multisensor Satellite Data».

**1.2. Документы, на основании которых ведется разработка**

Разработка программы ведется на основании приказа Национального исследовательского университета «Высшая Школа Экономики» № 2.3-02/1903-01 «Об изменении тем, руководителей курсовых работ студентов образовательной программы Программная инженерия факультета компьютерных наук» от 19.03.2017.

# **Назначение и область применения**

## **2.1. Назначение программного модуля**

Программа предназначена для детектирования изменений землепользования на основе мультисенсорных спутниковых данных, предоставляемых спутниками программы Sentinel 2 Европейского космического агентства.

Виды землепользования должны быть представлены следующими классами:

1. Водные ресурсы.
2. Пахотные земли и пастбища.
3. Застроенные области.
4. Лесные массивы.

Программа должна визуализировать изменения землепользования по всем классам в совокупности с указанием видом изменения землепользования, а также по изменению каждого класса индивидуально.

Программа предоставляет возможность сохранения результатов обработки данных в виде гео-векторных фалов.

## **2.2. Область применения программы**

Программный продукт может применяться учеными в сферах экологии и урбанистики, а также в области учета землепользования государственными органами, занимающимися контролем пользования земельными участками.

# **Технические характеристики**

## **3.1. Постановка задачи для разработки программы**

Задачей разработки программы является создание программы для определения изменений землепользования на основе мультисенсорных спутниковых данных с использованием машины опорных векторов (SVM). Виды землепользования должны быть представлены следующими классами:

1. Водные ресурсы.
2. Пахотные земли и пастбища.
3. Застроенные области.
4. Лесные массивы.

Программа должна предоставлять доступ к загрузке мультисенсорных данных Sentinel 2 Европейского и предоставлять возможность преобразования загруженных данных до уровня Level 2 [2]. Программа должна предоставлять возможность пользователю переобучить или обучить на дополнительных данных классификатор SVM для определения изменений землепользования на основе сцены Sentinel 2 и векторных файлов Open Street Map. Программа должна визуализировать результаты обнаружения изменений по мультисенсорным спутниковым данных и предоставлять возможность сохранять результаты в векторном файле ERSI Shapefile.

## **3.2. Описание алгоритмов и функционирования программы**

### **3.2.1. Поиск и загрузка данных Sentinel 2**

Поиск и загрузка мультисенсорных данных Sentinel 2 осуществляется с помощью Copernicus Open Access Hub API [3], предоставляющий поиск и загрузку данных c помощью технологии OpenSearch [4], предоставляющей результаты поиска c помощью протокола Atom [5]. Для работы с технологией OpenSearch в программе используется библиотека Apache Abdera [6].

#### **3.2.1.1. Ввод и обработка параметров поиска данных Sentinel 2**

Модуль программы для поиска и загрузки данных Sentinel 2 принимает от пользователя в качестве параметров поиска данных некоторые или все значения, указанные в списке ниже:

* Дата начала периода времени, в течении которого были зафиксированы и созданы данные спутниками ESA Sentinel 2.
* Дата завершения периода времени, в течении которого были зафиксированы и созданы данные спутниками ESA Sentinel 2.
* Область земной поверхности, на которой были зафиксированы мультисенсорные данные спутников ESA Sentinel 2.
* Максимальный процент покрытия спутниковых данных облаками.

Даты начала и завершения периода фиксирования данных, а также процент максимального покрытия мультисенсорного снимка облаками указываются пользователем в графическом интерфейсе программы с помощью предназначенных для этого контролеров. Указанные пользователем даты преобразовываются в формат ISO8601 [7].

Область земной поверхности, на котором производиться поиск данных указывается в специальном контроллере, содержащий интерактивную карту Земли, реализованной на языке JavaScript с использованием библиотеки Leaflet [8]. При начале процесса поиска данных, при указании области интереса пользователя в интерактивной карте пользователем, данных об полигоне области передаются в программу в формате WKT [9]в системе координат WGS84 [10].

#### **3.2.1.2. Создание запроса поиска данных**

После ввода необходимых пользователю параметров поиска данных Sentinel 2, программа создает URL для запроса методом GET с указанием в параметрах запроса необходимые для поиска параметры, согласно спецификации ESA Copernicus Open Access API [3]. Ниже представленные примеры создаваемых адресов для запроса данных спутников Sentic периодом фиксирования поверхности земли с 1 июля 2017 года по 31 августа 2017 года, с максимальным процентом покрытия сцены облаками 10 % и пересекающейся с областью, ограниченной областью Средиземного моря:

[*https://scihub.copernicus.eu/dhus/search?q=platformname:Sentinel-2*](https://scihub.copernicus.eu/dhus/search?q=platformname:Sentinel-2) *AND beginposition:[ 2017-07-01T00:00:00.000Z TO 2017-08-21T00:00:00.000Z] AND footprint:"Intersects(POLYGON((-4.53 29.85, 26.75 29.85, 26.75 46.80,-4.53 46.80,-4.53 29.85)))" AND cloudcoverpercentage:[0 TO 10]*

Рис.1. Пример запроса данных

Запрос поиска данных производиться с помощью OpenSearch клиента, созданного с помощью библиотеки Apache Abdera [6] и программа производить парсинг Atom для демонстрации результата пользователю.

#### **3.2.1.3. Получение, обработка результата запроса данных и загрузка данных Sentinel 2**

При успешном запросе поиска данных спутников ESA Sentinel 2 в качестве ответа с серверов ESA приход данные в формате Atom. Данные ответа сервера обрабатываются клиентом OpenSearch и преобразовываются в коллекцию элементов описания данных, из которых извлекается следующая информация:

1. Название спутниковых данных
2. Информация о дате фиксирования данных
3. Размер загружаемого файла
4. Ссылка на quick view найденных данных
5. Ссылка загрузки данных

Программа загружается изображение quick view найденных спутниковых снимков для наглядной демонстрации пользователю полученных данных, а также демонстрирует остальную информацию пользователю.

Загрузка данных ESA Sentinel 2 из списка результатов поиска происходить при переходе по ссылке загрузки данных с помощью программы загрузки файлов из Интернета, установленного на ПК пользователя..

### **3.2.2. Преобразование данных Sentinel 2 Level 1C к Sentinel 2 Level 2A**

Согласно спецификации данных Sentinel 2 «Sentinel 2 Product Specification» [2] мультиспектральные спутниковые снимки, создаваемые спутниками Sentinel 2, имеют несколько типов данных, характеристики которых представлены в таблице № 1.

|  |  |  |  |
| --- | --- | --- | --- |
| **Тип** | **Уровень обработки** | **Описание** | **Целевые пользователи** |
| S2HKTM | N/A | Телеметрия космического спутника Sentinel 2 в формате Transfer Frame. | Полетные операции Sentinel 2. |
| S2MSI0 | 0 | Сжатые необработанные мультиспектральные (MSI) данные в формате ISP. | Эксперты MSI и внутренние пользователи ESA. |
| S2MSI1A | 1A | Несжатые необработанные мультиспектральные (MSI) данные с файлами спектральных полос. | Не распространяется среди внешних пользователей. |
| S2MSI1B | 1B | Радиометрически скорректированные (откалиброванные) данные изображения MSI с файлами спектральными полосами. | Конечные пользователи-эксперты. |
| S2MSI1C | 1C | Ортофототрансформированная и UTM-геокодированная отражательная способность верхней атмосферы с субпиксельной мультиспектральной и многорежимной регистрацией. | Общие конечные пользователи. |
| S2MSI2A | 2A | Ортофототрансформированная и UTM-геокодированная мультиспектральная отражательная способность нижней части атмосферы. Дополнительные выходы - карта аэрозольной оптической толщины (AOT), карта водяного пара (WV), карта классификации облаков и снега вместе с данными показателей качества. |

Таблица №1. Уровни данных спутников ESA Sentinel-2

В целях определения изменения землепользования программа нуждается в данных отражения поверхности Земле, то есть отражения нижних областей атмосферы планеты (Bottom Of Atmosphere – BOA) [11] с соответствующей атмосферной коррекцией, поэтому программа должна работать согласно таблице №1 с данными S2MSI2A или Level 2. Однако данные уровня Level 2A не регулярно публикуются и доступны с помощью Copernicus Open Access Hub. [3] связи с чем возникает необходимость предоставлять пользователю возможность преобразовывать данные Level 1C, полученные с помощью Copernicus Open Access Hub API, до уровня Level 2A.

ESA рекомендует [12] преобразовывать данные Sentinel Level 1C к данным Level 2A c помощью утилиты Sen2Cor [12] для программной среды Anaconda Python 2.7 [13] , которая имеет интерфейс командой строки. Разработанная программа преобразовывает данные c уровня Level 1С к уровню Level 2A при помощи запуска утилиты Sen2Cor с указанием выбранных пользователем параметров преобразования, с отображением состояния преобразования, а также возможностью прервать процесс преображения данных. Модель предоставляет пользователю возможностью выбора следующих параметров преобразования данных:

1. Преобразование спутниковых данных с резолюцией 10 х 10 метров.
2. Преобразование спутниковых данных с резолюцией 20 х 20 метров.
3. Преобразование спутниковых данных с резолюцией 60 х 60 метров.
4. Преобразование спутниковых данных с резолюциями 10 х 10 метров, 20 х 20 метров и 60 х 60 метров.

Таким образом, модуль преобразования мультиспектральных спутниковых данных Sentinel 2 Level 1C к данным c представлением отражения света Sentinel 2 Level 2A представляет собой обертку графического интерфейса Java для управления утилитой Sen2Cor [12], выполняющаяся в среде Anaconda [13].

### **3.2.3. Работа с мультисенсорными сценами Sentinel-2 Level 2A**

#### **3.2.3.1. Мультисенсорные данные Sentinel-2 Level-2A**

Согласно спецификации данных Sentinel-2 [11] Загруженная сцена представляет собой одну или несколько гранул в разрешениях 60 x 60 метров, 20 х 20 метров, 10 х 10 метров, содержащие в себе файлы формате JP2000 [14] с геопривязкой, каждый из которых содержит один канал со значениями отражения света определенного диапазона света земной поверхностью. На Рис.2 представлена диаграмма соответствия названия канала с фиксированным диапазоном данных и максимальную резолюцию данного файла.

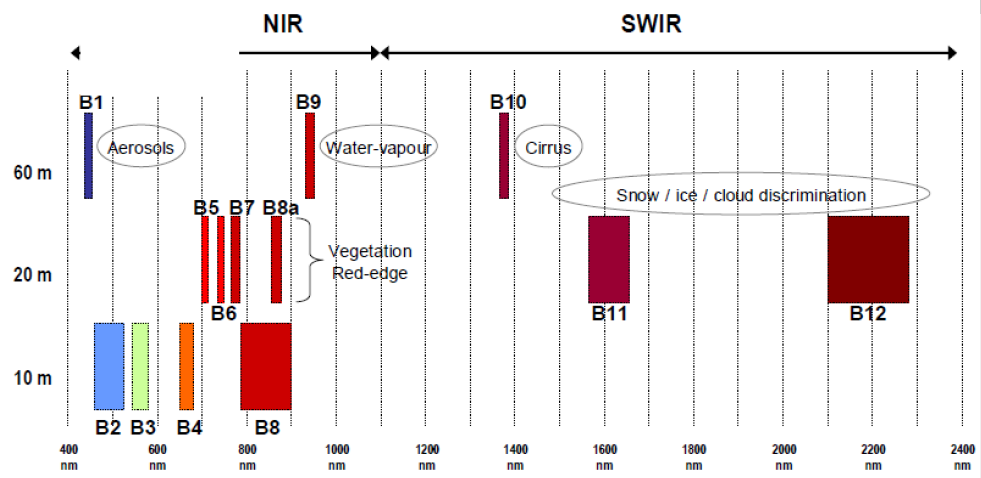


Рис. 2 Диаграмма соответствия названия канала с фиксированным диапазоном данных и максимальную резолюцию данного файла [11]

Также вместе с данными каналами предоставляются следующие маски:

* Маска вероятности облаков
* Маска вероятности облаков
* Маска первичной классификации земной поверхности
* Маска верхнего слоя атмосферы
* Маска плохих пикселей
* Маска первичной классификации водной поверхности

Перечисленные выше маски предоставляются исключительно с величиной резолюции 60х60 и 20х20 метров в связи с тем, что данный резолюции 10х10 метров содержит не полный набор полос, необходимых для обнаружения изменений и следовательно поддерживаются программой.

#### **3.2.3.2. Извлечение данных гео-привязки и чтения файлов в формате JP2K**

Данный растровых файлов каналов считываются с помощью библиотеки GeoTools [15], которая требует при чтении фалов в формате .JP2K [14] фалов файла системы координат гео-привязки приложения в формате .prj и world-файла, в котором указываются граничные координаты в указанной проекции считываемого растрового изображения.

Информация о системе координат и world-файла извлекаются из фала растрового изображения с помощью библиотеки абстракции гео-пространственных данных GDAL (Geospatial Data Abstraction Library) [16] и записываются в файлы, имеющие название, соответствующее названию файла растрового изображения, и расширения .prj для системы координат растрового изображения и .j2w для world-файла, после чего данные успешно считываются с помощью библиотеки GeoTools [15].

### **3.2.3. Модуль обучения классификатора SVM**

#### **3.2.3.1. Классификация мультисенсорных спутниковых данных Sentinel-2**

Согласно базовой статье для разработки программного продукта [17], данные мультиспектральных снимков Sentinel-2 за определенную дату фиксирования земной поверхности классифицируются с помощью классификатора машины опорных векторов (Support Vector Machine) [18], классифицирующий мультиспектральные данные на следующие классы:

1. Водные ресурсы
2. Агрикультура
3. Застроенные области
4. Лесные массивы

Классификатор SVM [18] разделяет гиперпространство значений мультисенсорного снимка, принимая в качестве входных данный вектор, составленный из значений пикселей, соответствующих значениям пикселей с одинаковым географическим положением из считанных данных и величиной пикселя, соответствующей интенсивности отражения света земной поверхности в определенном диапазоне, соответствующей файлу канала данных.

В качестве классификатора SVM использовалась библиотека Statistical Machine Intelligence and Learning Engine [19], реализованной на языке Java.

В связи с тем, что миссия ESA Sentinel-2 [20] включает в себя 2 спутника: ESA Sentinel-2 Instrument A и ESA Sentinel-2 Instrument B, для классификации мультиспектральных данных для данных, зафиксированных каждым спутников, создается 2 классификатора SVM.

#### **3.2.3.2. Выбор набора каналов для классификации вида землепользования**

Для определения состава входного для классификатора SVM вектора признаков был произведено обучение классификатора на множестве всех комбинаций каналов данных Sentinel-2 Level 2A. Задачей данного эксперимента было выявление размера и состава вектора признаков для получение максимальной точности классификации вида землепользования.

В качестве элементов входного вектора признаков не использовались данные каналов AOT [11] и TCI [11] гранул Sentinel-2 Level 2A в связи с тем, что значения пикселей файла AOT (Aerosol Optical Thickness) соответствуют оптической толщине аэрозоля в атмосфере, что не связано с данными пикселей отражения света нижнем слоем атмосферы, а данные TCI (True Color Image) дублируют данные B02, B03, B04 в виде трехканального RGB-изображения в формате JP2K.

По результатам проведения измерений значений точностей классификаторов, все наборы векторов признаков длины менее 12 каналов получили точность классификации менее 80%. На рис.3 представлены значения точности классификации для наборов длины от 12 файлов.

Рис.3. Точность обучения в зависимости от множества входных каналов

Согласно рис.3 максимальная точность классификации данных Sentinel-2 Level 2A достигается на векторе входных признаков, состоящих из всех каналов данных.

#### **3.2.3.3. Обучение классификаторов SVM для данных Sentinel-2**

Классификатор SVM [18] требует обучения для классификации данных. В разработанном программной продукте представляется возможность обучения классификатора для каждого из спутников миссии ESA Sentinel-2.

Обучение классификатора SVM производиться с помощью мультиспектральных снимков Sentinel-2 Level 2A и векторных данных Open Street Map [21] за период фиксации данных Sentinel-2 в формате ERSI Shapefile. Пользователь может использовать в качестве векторных данных собственные векторные файлы, однако для успешного обучения они должны использовать типы тегов, используемые Open Street Map в качестве атрибутов, для описания класса землепользования, которые демонстрируют включенные в них векторные данные. [22]

Процесс обучения включает в себя следующие этапы:

1. Чтение, объединение и растеризация векторных данных Open Street Map [21]
2. Чтение мультиспектральных данных Sentinel-2 Level 2A и информации об спутнике, зафиксировавшего эти данные
3. Чтение данных о покрытии облаками и снегом, а также испорченных пикселей данных Sentinel-2 Level 2A и исключение данных пикселей из мультиспектральных данных для обучения классификатора SVM
4. Обучение модели SVM по предоставленным данным с использование для решетчетого поиска гиперпараметров классификатора SVM и кросс-валидацией

В следующих пунктах подробно описаны этапы обучения классификатора SVM.

#### **3.2.3.3.1. Выбор областей для обучения и векторизация данных Open Street Map**

Классификатор SVM требует обучения для классификации данных. В разработанном программной продукте представляется возможность обучения классификатора для каждого из спутников миссии ESA Sentinel-2. Для выбора областей землепользования и соответствующего ему класса землепользования используются векторные данные в формате ERSI Shapefile [22] c использованием тегов Open Street Map Tags [23]. В таблице №2 представлены значения тегов данных для классов землепользования:

|  |  |
| --- | --- |
| **Класс землепользования** | **Тег OpenStreetMap** |
| Водные ресурсы | * “water” |
| Агрикультура | * "farmland", * "meadow", * "orchard", * "plant\_nursery", * "vineyard", * "farm", * "allotments", * "farmyard", * "greenhouse\_horticulture" |
| Застроенные области | * "commercial", * "garages", * "industrial", * "religious", * "residential", * "retail", * "school", * "brownfield", * "construction", * "landfill", * "quarry", * "salt\_pond" |
| Лесные массивы | * "forest", * "wood" |

Табл. 2 Соответствие тегов OpenStreetMap классам землепользования

После выборки полигонов землепользования, данная коллекция геометрий преобразовывается в растровое изображение с разрешением с использованием библиотеки Java Advanced Imaging [22]. В качестве значения пикселя полученного растрового изображения используется двоичная маска, где i-й блок соответствует номеру классу землепользования. Например, значение 2 в двоичной записи 0010 представляет пиксель класса агрикультуры.

#### **3.2.3.3.2. Извлечение из спутниковых данных пикселей, покрытых облаками или снегом и испорченных пикселей**

Данные Sentinel-2 Level 2A включают в себя растровые файлы в формате JP2000 [14] с масками вероятностей облаков, снега и испорченных пикселей. Пиксели масок снега и облаков со значением вероятности более 90% извлекаются каждого файла полос мультисенсорных данных с помощью библиотеки Java Advanced Imaging [24] с установкой значений пикселей в файлах на значение -1, для того, чтобы не обучать классификатор землепользования SVM на пикселях со значениями интенсивности отражения света облаками в различных диапазонах.

#### **3.2.3.3.3. Подбор параметров по сетке и кросс-валидация при обучении классификаторов SVM**

При обучении классификатора SVM программа использует побор параметров классификатора SVM с помощью подбора параметра по сетке. Классификатор SVM использует для классификации [25], которая в свою очередь требует значения и стандартного отклонения и гамма-параметр [25] для разделения гиперпространства для классификации. Суть поиска заключается в переборе каждого параметра в значения от 0.5 до 213 для нахождения подходящих параметров для классификатора, при нахождении которого фиксируются параметры.

При обучении классификатора используется кросс-валидация модели на данных для обучения. Исходный набор векторов, полученных из данных Sentinel-2 для обучения, делиться случайно на 5 наборов векторов. Классификатор поочередно обучается на всех множествах длины 4 из данного набора, и проверяется на 5-м оставшемся. Из данного набора обученных классификаторов выбирается обученный классификатор с максимальным значении валидации обучения.

#### **3.2.3.3. Импорт и экспорт классификаторов SVM**

После успешного обучения объект классификатора SVM сериализуется в файл с расширением .svm, соответствующий виду спутнику, зафиксировавшего данные. В связи с тем, что значение пикселей каналов данных для различных классов могут различаться в связи с различными периодами года, климатическими и культурными особенностями в различных регионах планеты, пользователю предоставляется возможность запуска обучения классификаторов, а также их экспорт и импорт в программу определения изменений землепользования на основе мультисенсорных спутниковых данных.

### **3.2.4. Определение изменений землепользования по мультисенсорным спутниковым данным Sentinel-2 с использованием классификатора SVM**

Определение изменение землепользования производиться на основе двух гранул данных Sentinel-2 Level 2A с различной датой создания снимков спутниками. Рекомендуются использоваться данные, соответствующие одинаковым сезоном создания снимков спутниками Sentinel-2, и классификатором SVM, обученном на территории с не сильно различающимися климатическими особенностями и датой фиксации, соответствующей данным, на основе которых производиться определение изменений землепользования.

Процесс детектирования изменений землепользования на основе мультисенсорных спутниковых данных включает в себя следующие этапы:

1. Считывание и обрезка данных Sentinel-2 Level 2A
2. Обрезка считанных данных на основе области интереса пользователя или пересечении покрытия территорий выбранными спутниковыми данными
3. Извлечение из считанных данных пикселей облаков, снега и отсутствующих значений
4. Классификация считанных данных классификатором SVM, соответствующего спутника, создавшего снимок
5. Проверка пикселей классифицированных растровых данных на нахождение отдельных пикселей в кластер пикселей отличного класса
6. Векторизация классифицированных растровых изображений
7. Определение территорий с изменившимся видом землепользования
8. Расчет площадей изменений землепользования по классам
9. Демонстрация в виде векторных данных и данных в формате WKT [9]

#### **3.2.4.1. Выбор области интереса для классификации снимков**

Пользователь может выбрать область интереса для классификации данных, в таком случае входные изображения Sentinel-2 Level-2A обрезаются по указанной в векторных файлах геометриях с помощью библиотеки Java Advanced Imaging [24].

#### **3.2.4.2. Проверка пикселей классифицированных растровых данных на нахождение отдельных пикселей в кластер пикселей отличного класса**

После успешной классификации входных данных, программа проверяет растровое изображение на пиксели одного класса, окруженного пикселями одного другого класса. В таком случае, окруженному пикселю присваивается значение класса соседних пикселей.

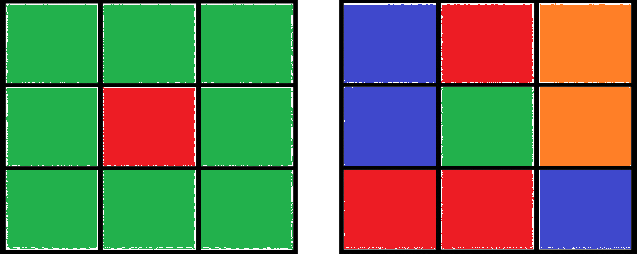


Рис 4. Примеры пикселя для замены и валидного пикселя

#### **3.2.4.2. Векторизация классифицированных данных**

После успешной классификации и проверки пикселей, данные растровые изображения переводятся в векторный формат, в котором объединяются в полигоны пиксели, имеющие одинаковые классы. Векторизация данных производиться с помощью библиотеки GeoTools [15].

#### **3.2.4.3. Определение территорий с изменением вида землепользования**

Векторизованные данные проверяются на изменение вида землепользования путем перебора геометрий векторов файла «до» с определение пересечений с геометриями второго векторной коллекции. На основе этих пересечений создается новая коллекция векторных данных, где в качестве геометрии указывается область пересечения, и качестве атрибутов класс землепользования «до» и «после».

#### **3.2.4.4. Расчет площади изменений**

Итоговая векторная коллекция используется для определения площади изменений землепользования. Данный геометрии векторов и классов землепользования «до» и «после» отправляются в таблицу базы данных PostgreSQL [26] c установленным расширением PostGIS [27]. Данные агитируются по классам «до» и «после», после чего для агрегированных данных возвращаются значения площадей для измененный и нет классов землепользования. После расчета, данные удаляются из базы данных PostGIS[27].

Для расчета площадей используются следующая таблица PostGIS:

*TABLE landUses(*

*id SERIAL PRIMARY KEY,*

*the\_geom geography NOT NULL,*

*before INTEGER NOT NULL,*

*after INTEGER NOT NULL*

*Рис. 6. Таблица для определения площадей*

Запрос для расчета площадей выглядит следующим образом:

SELECT sum(st\_area(the\_geom)), before, after FROM landuses GROUP BY before, after

#### **3.2.4.4. Демонстрация в виде векторных данных и данных в формате WKT**

После расчета площадей изменений, результат детектирования отображается в интерфейсе пользователя в виде полигонов на карте LeafletJS [8], в виде записи геометрий в формате WKT, а также отображаются значения площадей классов землепользования.

#### **3.2.4.4. Сохранение и повторное открытие файла результата определения изменения землепользования**

Пользователь может сохранить и открыть позже результат определения изменения землепользования. При сохранении результата, на компьютере пользователя сериализуется файла класса определения землепользования, после чего пользователь может открыть данный файл в программе и посмотреть результат изменения землепользования.

#### **3.2.4.4. Сохранение результата в формате ERSI Shapefile**

Пользователь может сохранить файл результата классификации в формате ERSI Shapefile [22], где в качестве атрибутов указываются классы землепользования «до» и «после», для последующей работы с экспортированными данными в GIS-средствах. Создание файлов ERSI Shapefile происходить с помощью библиотеки GeoTools [15].

## **3.3. Описание и обоснование выбора метода организации входных и выходных данных**

### **3.3.1. Описание метода организации входных и выходных данных**

#### **3.3.1.1. Модуль загрузки данных Sentinel-2**

Входные данные вводятся в соответствующие элементы ввода формы:

1. Логин и пароль.
2. Выбор дат фиксирования данных вводятся с помощью контролера календаря.
3. Выбор максимальной облачности вводиться в виде числа в соответствующий контроллер.
4. Выбор области производится путем рисования области на карте LeafletJS [8]

Программа демонстрирует пользователю информацию о найденных данных и начинает загрузку по приходу по ссылке загрузки данных.

#### **3.3.1.2. Модуль преобразования данных**

Для преобразования данных с уровня Level 1C к уровню Level 2A пользователь выбирает необходимый файл загруженных данных и величину резолюции итоговых данных.

Преобразованные данные создаются в той-же директории, где находятся данные Level 1C.

#### **3.3.1.3. Модуль обучения классификаторов SVM**

Для обучения модели SVM пользователю нужно выбрать следующие файлы для обучения:

1. Папку с данными Sentinel-2 Level 2A, по которым будет происходить обучения
2. Выбрать необходимую гранулу и резолюцию данных
3. Папку с векторными файлами OpenStreetMap в формате ERSI Shapefile

Послу успешного обучения модели SVM, соответствующего спутнику, создавшего данные сериализуется в советующий файл и доступны для экспорта и импорта.

#### **3.3.1.4. Модуль определение изменения землепользования**

Программа использует в качестве мультисенсорных спутниковых данных данные Sentinel 2, т.к. данные снимки, по сравнению с аналогом данных NASA Landsat [28], обладают более высоким разрешением снимков и является более молодой миссией, в связи с чем данные спутников будут обновляться дольше, чем данные NASA.

Для определения изменения землепользования по мультисенсорным спутниковым данным пользователь должен указать путь до 2-х данных Sentinel-2 Level 2A, необходимые гранулы и резолюцию. Опционально пользователь может указать путь на файлы в формате ERSI Shapefile для того, чтобы не производить классификацию данных и определение изменения землепользования за пределами геометрий, указанных в данном файле.

Модуль предоставляет пользователю сохранить результат определения изменения землепользования в виде сериализованного объекта для последующего чтения результата в программном продукте.

Также в программе предусмотрена возможность сохранения результата в формате ERSI Shapefile для последующей работы с ним в GIS-средствах, т.е. формат ERSI Shapefile является одним из самых популярных форматов хранения векторных гео-данных.

## **3.4. Описание и обоснование выбора состава технических и программных средств**

### **3.4.1. Состав технических и программных средств**

Программой предъявляются следующие требования к составу и параметрам технических средств:

1. Процессор с тактовой частотой не менее 2.3ГГц
2. Оперативная память не менее 4Гб
3. Жесткий диск со свободным объемом не менее 20Гб
4. Монитор
5. Клавиатура
6. Мышь
7. Стабильное подключение к сети Интернет

Программой предъявляются следующие требования к информационной и программной совместимости персональных компьютеров пользователей:

1. Операционная система Windows 7, Windows 8, Windows 8.1, Windows 10 или более поздней версии.
2. Программное обеспечение Java 8 JRE или выше.
3. Дистрибутив Anaconda c версией Python 2.7.
4. Установленная база данных PostgreSQL с расширением PostGIS
5. Установления библиотека Geospatial Data Abstraction Library (GDAL/OGR) версии 2.0.1

## **3.4.2.** **Обоснование выбора технических и программных средств**

Растровые данные Sentinel-2 занимают в среднем 600Мб дискового пространства одной гранулой, в связи с чем компьютер пользователя должен обладать большим количество дискового пространства и оперативной памяти.

В качестве языка программирования был выбран язык Java, т.к. для данного языка существует довольно развитая библиотека GeoTools для работы с гео-данными и Java Advanced Imaging для работы с растровой арифметикой.

Программа требует установленного дистрибутива Anaconda, т.к. утилита Sen2Cor, требуемая для преобразования данных Sentinel-2 Level 1C к Sentinel-2 Level 2A исполняется в данной среде, а также библиотека GDAL требует Python 2.7.

Программа использует библиотеку GDAL для чтения информации о загруженные растровых данных и извлечения их системы координат и геопривязки, т.к. данная библиотека является самой широко применяемой для данных целей.

# **Ожидаемые технико-экономические показатели**

## **4.1. Предполагаемая потребность**

Программный продукт может применяться учеными в сферах экологии и урбанистики, а также в области учета землепользования государственными органами, занимающимися контролем пользования земельными участками.

## **4.2. Ориентировочная экономическая эффективность**

Аналогов данного продукта выявлено не было. Преимущество решения заключается в детектировании изменений землепользования по мультисенсорным спутниковым данным ESA Sentinel 2, которые находятся в открытом доступе для всех пользователей, а так в возможности загрузить данные .непосредственно из программы.

# **Список использованной литературы**

1. ГОСТ 19.404-79 Пояснительная записка. Требования к содержанию и оформлению. //Единая система программной документации. М.: ИПК Издательство стандартов, 2001.
2. Sentinel 2 Data Types [Электронный ресурс] // ESA: [сайт]. [2018]. URL: https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types (дата обращения: 10.02.2018).
3. Copernicus Open Access Hub API [Электронный ресурс] // ESA: [сайт]. [2018]. URL: https://scihub.copernicus.eu/twiki/do/view/SciHubWebPortal/APIHubDescription (дата обращения: 10.02.2018).
4. OpenSearch Specification [Электронный ресурс] // OpenSearch: [сайт]. [2018]. URL: http://www.opensearch.org/Home (дата обращения: 10.02.2018).
5. The Atom Syndication Format [Электронный ресурс] // The Atom Syndication Format: [сайт]. [2018]. URL: https://tools.ietf.org/html/rfc4287 (дата обращения: 10.02.2018).
6. Apache Abdera [Электронный ресурс] // Apache Abdera: [сайт]. [2018]. URL: https://abdera.apache.org/ (дата обращения: 12.02.2018).
7. ISO 8601 Date and time format [Электронный ресурс] // ISO 8601: [сайт]. [2018]. URL: https://www.iso.org/iso-8601-date-and-time-format.html (дата обращения: 10.02.2018).
8. Leaflet - an open-source JavaScript library [Электронный ресурс] // Leaflet.js: [сайт]. [2018]. URL: https://leafletjs.com/ (дата обращения: 25.01.2018).
9. Well-known text [Электронный ресурс] // WKT: [сайт]. [2018]. URL: http://www.geoapi.org/snapshot/javadoc/org/opengis/referencing/doc-files/WKT.html (дата обращения: 10.02.2018).
10. EPSG:4326 WGS84 [Электронный ресурс] // Spatial Reference: [сайт]. [2018]. URL: http://spatialreference.org/ref/epsg/wgs-84/ (дата обращения: 10.01.2018).
11. A. Gatti C.N. Sentinel-2 Products Specification Document. 144th ed. Cannes: Alenia Space France, 2018. 514 pp.
12. ESA STEP Plugins Sen2Cor [Электронный ресурс] // ESA Step: [сайт]. [2005-2018]. URL: http://step.esa.int/main/third-party-plugins-2/sen2cor/ (дата обращения: 1.03.2018).
13. Anaconda [Электронный ресурс] // Anaconda: [сайт]. [2018]. URL: https://anaconda.org/ (дата обращения: 01.03.2018).
14. JP2OpenJPEG --- JPEG2000 driver based on OpenJPEG library [Электронный ресурс] // GDAL: [сайт]. [2018]. URL: http://www.gdal.org/frmt\_jp2openjpeg.html (дата обращения: 08.03.2018).
15. GeoTools The Open Source Java GIS Toolkit [Электронный ресурс] // GeoTools: [сайт]. [2018]. URL: http://www.geotools.org/ (дата обращения: 15.02.2018).
16. GDAL - Geospatial Data Abstraction Library [Электронный ресурс] // GDAL: [сайт]. [2018]. URL: http://www.gdal.org/index\_ru.html (дата обращения: 15.02.2018).
17. Patrick Griffiths ⁎ P.H.O.G.S.V.D.L. Mapping megacity growth with multi-sensor data // Remote Sensing of Environment, No. 114, Sep 2010. P. 25.
18. Машина опорных векторов [Электронный ресурс] // MachineLearning.ru - Профессиональный информационно-аналитический ресурс, посвященный: [сайт]. [2018]. URL: http://www.machinelearning.ru/wiki/index.php?title=SVM (дата обращения: 20.02.2018).
19. Smile (Statistical Machine Intelligence and Learning Engine) [Электронный ресурс] // Statistical Machine Intelligence and Learning Engine: [сайт]. [2018]. URL: https://github.com/haifengl/smile (дата обращения: 20.02.2018).
20. SENTINEL-2 [Электронный ресурс] // ESA: [сайт]. [2018]. URL: https://sentinel.esa.int/web/sentinel/missions/sentinel-2 (дата обращения: 10.02.2018).
21. Open Street Map [Электронный ресурс] // OSM: [сайт]. [2018]. URL: https://www.openstreetmap.org/about (дата обращения: 01.03.2018).
22. Environmental Systems Research Institute, Inc. ESRI Shapefile Technical Description. New York Street. 1998. 34 pp.
23. OSM Tags [Электронный ресурс] // Open Street Map: [сайт]. [2018]. URL: https://wiki.openstreetmap.org/wiki/Tags (дата обращения: 10.02.2018).
24. Java Advanced Imaging API Home Page [Электронный ресурс] // Oracle: [сайт]. [2018]. URL: http://www.oracle.com/technetwork/java/iio-141084.html (дата обращения: 10.04.2018).
25. Chung M.K. The Gaussian kernel. 2007. Статья.
26. POSTGRESQL [Электронный ресурс] // POSTGRESQL: [сайт]. [2018]. URL: https://www.postgresql.org/ (дата обращения: 10.04.2018).
27. PostGIS - Spatial and Geographic objects for PostgreSQL [Электронный ресурс] // PostGIS: [сайт]. [2018]. URL: https://postgis.net/ (дата обращения: 10.04.2018).
28. The Landsat Program [Электронный ресурс] // NASA Landsat: [сайт]. [2018]. URL: https://landsat.gsfc.nasa.gov/ (дата обращения: 01.01.2018).
29. Sentinel 2 Level 2A Product Sentinel Online [Электронный ресурс] // The European Space Agency porta: [сайт]. [2000-2018]. URL: https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-2a (дата обращения: 18.02.2018).

# **Приложение 1 Терминология**

**Землепользование (Land Use) -**  управление и изменение окружающей природной среды или дикой природы в построенных условиях, таких как поселения и полу-естественные среды обитания, такие как пахотные поля , пастбища и управляемые леса.

**Мультисенсорные (мультиспектральные) снимки -** набормоноспектральных изображений одной и той же сцены, полученных одновременно, но в разных спектральных каналах. Поочередный синтез отдельных каналов позволяет решать многочисленные тематические задачи, а также помогает при дешифрировании снимков**.**

**Гранула Sentinel-2 –** мультисенсорное изображение Sentinel-2, покрывающее территорию 100х100 км2.

**Метод опорных векторов (Support Vector Machine, SVM) -** метод решения задачи классификации и регрессии путем построения нелинейной плоскости, разделяющей решения. Благодаря особенностям природы пространства признаков, в котором строятся границы решения, метод опорных векторов обладает высокой степенью гибкости при решении задач регрессии и классификации различного уровня сложности.

# **Приложение 2 Описание и функциональное назначение файлов**

Таблица 1.1

Описание и функциональное назначение файлов

|  |  |
| --- | --- |
| **Файл** | **Функциональное назначение** |
| Data.java | Класс, содержащий статические методы создание соединения JDBC для работы c базой данных PostGIS |
| setup.sql | SQL скрипты для создания таблицы для определения площадей изменений землепользования |
| LevelUpForm.java | Класс JavaFX контроллера модуля преобразования данных Sentinel-2 |
| ProgressForm.java | Класс Java FX формы процесса детектирования |
| SearchAndDownloadForm.java | Модуль поиска и загрузки данных, влючающий JavaFX контроллер |
| TrainForm.java | Модуль обучения классификаторов SVM |
| App.java | Класс загрузчик программы JavaFX |
| ChangeDetector.java | Класс для обнаружения изменений землепользования по мультисенсорным спутниковым данным |
| Classificator.java | Класс, содержащий логику модуля обучения данных Sentinel-2 и предоставления синглтонов классификаторов для каждого из спутников |
| ClassificationEnum.java | Класс, представляющий вид спутника ESA Sentinel |
| Controller.java | Класс, предоставляющий JavaFX контроллер для модуля обнаружения изменений землепользования по мультисенсорным спутниковым данным |
| LandUseChangeDetectionResult.java | Класс, представляющий результат обнаружения изменений землепользования по мультисенсорным спутниковым данным |
| Resolution.java | Класс, представляющий доступные резолюции данных Sentinel-2 |
| SentinelData.java | Класс абстракции спутниковых данных, считывает и предоставляет данные Sentinel-2, а так-же содержит методы для обработки данных Sentinel-2 |
| SentinelLevelUpdate.java | Класс, содержащий логику модуля преобразования данных Sentinel-2 Level 1C к Sentinel-2 Level 2A |
| Utills.java | Утилитный класс программы обнаружения изменений землепользования по мультисенсорным спутниковым данным |
| AppWebForm.index.js | Файл, содержащий JavaScript код для интерактивной карты модуля обнаружения изменений |
| SaDWebFrom.index.js | Файл, содержащий JavaScript код для интерактивной карты модуля обнаружения изменений |
| Model.svm | Обученный классификатор SVM для спутника ESA Sentinel-2 A, для начального использования программой |
| ModelB.svm | Обученный классификатор SVM для спутника ESA Sentinel-2 B, для начального использования программой |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ** | | | | | | | | | |
| Изм. | Номера листов (страниц) | | | | Всего  листов  (страниц)  в докум. | № докум. | Входящий № сопроводитель-ного документа и дата | Подпись | Дата |
| изме-ненных | заме-  ненных | новых | анну-  лиро-  ванных |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |