МЕТОДЫ И СРЕДСТВА ОБРАБОТКИ И ИНТЕРПРЕТАЦИИ КОСМИЧЕСКОЙ ИНФОРМАЦИИ

## УТОЧНЕНИЕ ПАРАМЕТРОВ КЛАССИФИКАЦИИ МЕТОДОМ ОПОРНЫХ ВЕКТОРОВ (SVM) ПРИ КРУПНОМАСШТАБНОМ КАРТОГРАФИРОВАНИИ АРКТИЧЕСКИХ ЛАНДШАФТОВ (НА ПРИМЕРЕ ОСТРОВА БЕЛЫЙ, КАРСКОЕ МОРЕ)

© 2020 г. О. С. Сизов<sup>а, b, c, \*</sup>, И. Р. Идрисов<sup>d</sup>, А. А. Юртаев<sup>d</sup>

<sup>a</sup>AO "Российские космические системы", Москва, Россия <sup>b</sup>Институт проблем нефти и газа РАН, Москва, Россия <sup>c</sup>Российский государственный университет нефти и газа (национальный исследовательский университет) им. И.М. Губкина, Москва, Россия <sup>d</sup>Тюменский государственный университет, Тюмень, Россия \*e-mail: sizov\_os@ntsomz.ru Поступила в редакцию 03.12.2019 г.

На примере острова Белый, расположенного в Карском море, изучаются вопросы автоматизации ландшафтного картографирования. Выбор острова в качестве тестового объекта обусловлен решением правительства Ямало-Ненецкого округа о создании научного стационара для планомерных фундаментальных исследований последствий антропогенного воздействия на арктические экосистемы. На примере небольшого тестового участка, полевое изучение которого выполнено летом 2016 г., рассматриваются актуальные методические подходы ландшафтного картографирования на основе объектно-ориентированного анализа изображений (OBIA) и классификации методом опорных векторов (SVM). Предлагается комбинированный способ, основанный на независимой обработке исходных данных с последующей статистической обработкой полученной информации. Результаты исследования показали, что подобный способ позволяет более детально подходить к оценке ландшафтной дифференциации территории исследований. Впервые для территории острова Белый используются высокодетальные снимки со спутника "Ресурс-П" и цифровые модели рельефа ArcticDEM. Использование объектно-ориентированного дешифрирования позволит в дальнейшем расширить набор исходных данных за счет материалов съемки с БПЛА, а результаты классификации SVM могут быть улучшены за счет дополнительных тематических слоев морфометрического анализа цифровых моделей рельефа. Предложенная технология применима для выполнения крупномасштабного ландшафтного картографирования всей территории острова Белый.

*Ключевые слова:* Остров Белый, классификация SVM, данные Д33, Ресурс-П, ландшафтное картографирование, сегментация, объектно-ориентированное дешифрирование **DOI:** 10.31857/S0205961420020050

#### **ВВЕДЕНИЕ**

Многочисленные исследования (https://www.amap. no/documents/doc/arctic-arctic-climate-impactassessment/796) показывают, что в настоящее время в арктических регионах под влиянием глобальных климатических колебаний и антропогенного воздействия, связанного с интенсификацией добычи углеводородов и традиционного природопользования (оленеводства) наблюдаются масштабные изменения природных условий. В связи с этим на севере Западной Сибири по инициативе правительства Ямало-Ненецкого автономного округа (ЯНАО) с 2014 г. организованы планомерные научные исследования происходящих изменений (http://docs.cntd.ru/document/460282983). В качестве одной из фоновых площадок мониторинга выбран остров Белый, что определяет необходимость актуализации данных о современном состоянии его природных условий и, в частности, о структуре и разнообразии ландшафтного покрова.

Наличие возможности использования в качестве фактической основы для составления ландшафтной карты различных высокодетальных мультиспектральных данных дистанционного зондирования Земли (ДЗЗ) и цифровых моделей рельефа позволяет применить современные алгоритмы автоматического дешифрирования, способные обеспечить максимальную достоверность и объективность результатов. Анализ актуальных работ в данной области (Li et al., 2014; Mountrakis et al., 2011; Blaschke, 2010; Waske, Benediktsson, 2007) показал, что одним из наиболее эффективных и доступных методических подходов является объектно-ориентированное дешифрирование (object based image analysis, OBIA) на основе метода опорных векторов (support vector machines, SVM).

В условиях большого разнообразия научных подходов и программных средств к выполнению OBIA целью данной работы является разработка оптимальной методики крупномасштабного ландшафтного дешифрировании (1 : 10000), основанной на использовании открытых исходных данных и общедоступного программного обеспечения. При этом основной исследовательской задачей является определение интервалов значений базовых параметров метода SVM, которые непосредственно определяют точность и достоверность результатов.

## РАЙОН И МАТЕРИАЛЫ ИССЛЕДОВАНИЙ

Остров Белый – небольшой (58 × 43 км) остров, расположенный в Карском море, отделятся от полуострова Ямал проливом Малыгина шириной до 9 км. Широта самой северной точки острова – мыса Белый – 73°30' с.ш., максимальная высота над у. м. – 12 м. Средняя многолетняя температура воздуха составляет -0.3°C в мае. +4.1°C в июле, +5.3°С в августе, +1.9°С в сентябре. Влажность воздуха около 90%. Преобладают ветра северного, северо-восточного и юго-западного направления. Глубина залегания многолетнемерзлых пород (ММП) составляет в среднем 50-60 см, глубина сезонноталого слоя (СТС) изменяется от нескольких дециметров на водораздельных равнинах до 1 м и более на морском пляже (Дружинин и др., 2015).

Остров является типичным примером западноносибирских арктических островных тундр. Важным ландшафтообразующим фактором являются приливы и отливы, сочетающиеся со сгонно-нагонными процессами и штормами. Так, в 2005 г. шторм привел к полъему уровня волы на острове на 6 м. Вдоль побережья расположены места массового гнездования, линьки и остановок во время миграций белолобого гуся и черной казарки, линьки гуменника. На острове постоянно обитает дикий северный олень, белый медведь, на песчаных отмелях встречаются атлантический морж, нерпа, морской заяц и белуха (Васильчук, Васильчук 2015). Течения Карского моря выносят на остров остатки древесины (плавник, которым нередко заполнена овражно-балочная сеть острова) и различные виды мусора.

В ноябре 1933 г. у протоки Рагозина в северозападной части острова открыта морская гидрометеорологическая (полярная) станция. В феврале 1972 г. станции присвоено имя бывшего начальника Амдерминского РМЦ Михаила Владимировича Попова. Из-за разрушения берега протоки метеорологическая площадка несколько раз переносилась: в 1950-х гг. и в сентябре 1966 г. дважды на 50 м на юго-юго-восток, в сентябре 1973 г. на 180 м на северо-восток. В советское время неподалеку от станции располагалась военная часть, а также проводились геологоразведочные работы на нефть и газ — в северной и северо-западной частях острова пробурены 4 скважины глубиной 2.5 км (Дружинин и др., 2015). С 2012 по 2016 гг. при поддержке Администрации ЯНАО на острове проводились ежегодные экологические работы по сбору и вывозу металлолома и рекультивации нарушенных участков (Галиулин и др., 2016).

В качестве фактической основе в рамках данной работы использовались два основных источника пространственных данных:

— мультиспектральные снимки со спутника "Ресурс-П" (аппаратура Геотон) за 2016 г., обеспечивающие полное покрытие территории острова безоблачной съемкой с предельным пространственным разрешением порядка 1.6 м/пикс. (каналы R, G, B, NIR). Материалы съемки были получены на безвозмездной основе от Научного центра оперативного мониторинга Земли АО "Российские космические системы" в рамках деятельности Госкорпорации "Роскосмос" по обеспечению российских государственных потребителей, включая научные организации РАН и высшие учебные заведения, что делает их использование предпочтительным для научных исследований в условиях высокой стоимости зарубежных аналогов.

– цифровая модель поверхности (ЦМП) Arctic-DEM (https://www.pgc.umn.edu/data/arcticdem/), доступная для всей территории Земли севернее 60 град. с.ш. с разрешением 2 м/пикс. ЦМП ArcticDEM обладает наилучшей детальностью по сравнению с другими глобальными ЦМП (AW3D30, AsterGDEM2 и др.). В условиях отсутствия растительности на острове Белом ArcticDEM может использоваться как полноценная цифровая модель рельефа (ЦМР), отражающая все особенности мезо- и микрорельефа района работ при небольших относительных высотных перепадах.

Непосредственно для проведения тестового картографирования был выбран небольшой участок размером 3.6 × 2.3 км, распложенный в центральной части острова и примечательный тем, что здесь проходит сквозная долина рек Пахаяха и Нябилахаяха (рис. 1). Максимальные высотные отметки достигают 11 м над у. м., минимальные – 2 м. Участок полностью покрывается фрагментом снимка с КА "Ресурс-П" от 7 июля 2016 г. (разрешение 1.6 м/пикс., маршрут 01772\_02) и фрагментом ЦМП ArcticDEM, созданном на основе съемки с KA WorldView-1,2 (разрешение 2 м/пикс., тайл мозаики № 48\_57\_2\_2).



Рис. 1. Обзорная карта расположения тестового участка (снимок Sentinel-2, синтез CIR, 2016).

Описание эталонов дешифрирования выполнено в ходе экспедиций сотрудниками Тюменского госуниверситета в 2014-2016 гг. (Дружинин и др., 2015, Юртаев, 2016, Юртаев, Сулкарнаев, 2018). В полевых условиях изучены формы рельефа, почвенный и растительный покров, глубина сезонного протаивания грунтов, установлена приуроченность ПТК к той или иной форме рельефа (пойме, лайде, водораздельной равнине и др.). Особое внимание уделялось формам мезорельефа как важным факторам ландшафтной дифференциации на уровне типов урочищ. Всего охарактеризованы 12 ландшафтных профилей с шагом наблюдений от 10 до 20 м и протяженностью до 1 км, выполнены детальные комплексные описания на 70 ключевых участках и репрезентативных точках (из них 5 точек находятся в пределах тестового участка).

## МЕТОДИКА ДЕШИФРИРОВАНИЯ

Согласно принятым научно-методическим подходам (Козин, 2007; Мильков, 1986) базовой единицей крупномасштабного ландшафтного картографирования является вид урочища, представляющий собой закономерный комплекс фаций, достаточно хорошо обособленный в природе в связи с неровностями рельефа и неоднородным составом почв и верхней части толщи четвертичных отложений. Ведущими факторами в дифференциации видов урочищ являются растительность и микрорельеф при относительно однородном литолого-фациальном комплексе, что определяет необходимость использования при классификации урочищ дополнительно к мультиспектральной информации тематических слоев ЦМР и вегетационных индексов.

Выделение контуров и определение видов конкретных природных комплексов ландшафтной структуры (урочищ и типов местности) традиционно выполняется на основе визуального дешифрирования экспертом-ландшафтоведом и является наиболее трудоемким этапом при создании ландшафтной карты (Идрисов и др., 2017; Козин, 2008). Для автоматизации этой процедуры в рамках подхода OBIA выполнялись следующие действия:

1. Создание на тестовый участок единого 6-канального изображения, включающего 4 мультиспектральных канала снимка с КА "Ресурс-П" (рис. 2, *a*), а также слои ЦМП ArcticDEM и вегетационного индекса NDVI, рассчитанного на основе данных с КА "Ресурс-П". Для объединения каналов использовался модуль Semi-Automatic Classification Plugin (https://www.researchgate.net/ publication/265031337\_Semi-Automatic\_Classification\_Plugin\_User\_Manual) в открытом программном обеспечении (ПО) QGIS 3.6 (https://qgis.org/ ru/site/).



**Рис.** 2. *а* – фрагмент исходного мультиспектрального снимка "Ресурс-П", разрешение 1.6 м/пикс., *б* – результат сегментации исходных данных, *в* – результат растеризации сегментов со средними значениями по каждому.

2. Сегментация полученного изображения с помощью инструмента Large-Scale MeanShift (LSMS) в открытом ПО OrfeoToolbox 6.6.0 (https://www.orfeotoolbox.org/CookBook/Applications/app LargeScale-MeanShift.html). Инструмент основан на алгоритме среднего смещения (MeanShift algorithm), который чувствителен к пространственному разрешению и адаптирован для обработки растров большого объема (Comaniciu, Meer, 2002). Настройка инструмента ограничивается выбором параметров минимальной площади сегмента и размера используемой виртуальной памяти ПК. Полученный векторный слой сегментов (рис. 2, б) в таблице атрибутов содержит средние значения яркости в каждом спектральном канале, что позволяет провести растеризацию с помощью QGIS и получить исходное сегментированное изображение для классификации (рис. 2, в).

3. Классификация с обучением методом SVM с помощью инструмента SVM Classification в открытом ПО SAGA 7.3.0 (System for Automated Geoscientific Analyses) (Conrad et al., 2015). Инструмент основан на широко распространенной библиотеке libSVM (Chang, Lin, 2011), которая используется во многих приложениях (EnMAP box, ENVI, eCognition, PCI Geomaica и др.) и позволяет задавать все необходимые параметры SVM для повышения качества классификации. 4. Оценка качества классификации проводилась с помощью инструмента Confusion Matrix (Polygons/Grid) в ПО SAGA 7.3.0, который рассчитывает все стандартные характеристики – матрицу ошибок по классам (Confusion Matrix), точность разработчика (Accuracy Producer), точность пользователя (Accuracy User), коэффициент Каппа (Карра) и общую точность (Overall Accuracy).

Таким образом, все используемые инструменты относятся к классу открытого ПО, имеющего полноценный функционал. Наиболее сложным и неопределенным моментом данной методики является выбор параметров классификации SVM, который определяется особенностями данного метода.

### ХАРАКТЕРИСТИКА МЕТОДА SVM

SVM представляет собой набор контролируемых методов обучения, используемых для классификации, регрессии и обнаружения статистических выбросов. В случае линейной разделимости исходных данных метод основан на построении гиперплоскости, разделяющей объекты выборки наиболее оптимальным способом (оптимальной разделяющей гиперплоскости, ОРГ) (Вапник, 1974). Алгоритм работает в предположении, что чем больше расстояние (зазор) между ОРГ и объектами разделяемых классов, тем меньше будет средняя ошибка классификатора (рис. 3).

Пусть имеется обучающий набор точек вида  $\{(x_1, c_1), (x_2, c_2), ..., (xn, c_n)\}$ , где  $x_i$  – это *n*-мерный вещественный вектор, а с, принимает значение 1 или -1 в зависимости от того, какому классу принадлежит точка х,. Для классификации необхолимо построить разделяющую гиперплоскость следующего вида: wx + b = 0, где w – перпендикуляр к разделяющей гиперплоскости, b определяется кратчайшим расстоянием между гиперплоскостью и началом координат. Так как мы определяем оптимальное разделение, то нас интересуют ближайшие к опорным векторам детерминируемых классов гиперплоскости, которые могут быть описаны следующими уравнениями: wx + b = 1 и wx + b = -1, ширина полосы между которыми равна 2/w. Таким образом, задача сводится к оптимизации при условии, что  $c_i(wx_i + b) \ge 1, 1 \le i \le n$ , которая является стандартной задачей квадратичного программирования и решается с помощью множителей Лагранжа (Романов, Рубанов, 2012). В итоге алгоритм классификации может быть записан в виде:

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{n} \lambda_i c_i x_i x + b\right), \tag{1}$$

где [ $\lambda = (\lambda_1, ..., \lambda_n)$ ] – вектор множителей Лагранжа.

Для нелинейного разделения классов было предложено заменить скалярные произведения нелинейными функциями ядра. При этом все образцы обучающей выборки  $R^n$  переводятся в пространство более высокой размерности H с помощью специального отображения  $\varphi: R^n \to H$  так, чтобы в новом пространстве объекты были линейно разделимы (Boser et al., 1992, Cortes, Vapnik, 1995).

Существует несколько типов ядер (22 Chang et al., 2011):

- линейное:  $K(x_i, x_j) = x_i^T x_j$ .

- полиномиальное:  $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0.$ 

— радиальная базисная функция (radial basis function, RBF):  $K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2), \gamma > 0.$ 

- сигмоид:  $K(x_i, x_i) = \operatorname{th}(\gamma x_i^T x_i + r),$ 

где ү, *r*, и *d* – параметры ядра.

К преимуществу метода SVM относится возможность выбора наиболее подходящего ядра для решения конкретной задачи, включая возможность создания собственных типов ядер. Однако существует и ряд недостатков SVM, которые усложняют его практическое применение (Mountrakis et al., 2011):



Рис. 3. Метод опорных векторов (SVM).

 не существует общего подхода к автоматическому выбору ядра и его параметров в случае нелинейной разделимости классов;

 – напрямую не рассчитываются оценки вероятности, они получаются дополнительно с использованием затратной перекрестной проверки (кросс-валидации).

Для успешной классификации методом SVM должны быть определены оптимальный тип ядра и его параметры (Chang, Lin, 2011).

Для выбора типа ядра с помощью инструмента, предложенного на странице разработчиков библиотеки libSVM (https://www.csie.ntu.edu.tw/ ~cjlin/libsvm/) была создана простейшая выборка, имитирующая ландшафт тестового участка (рис. 4). Сравнение различных вариантов классификаций показало, что наилучшие результаты дает ядро RBF (табл. 1). Использование RBF рекомендовано для большинства задач, поскольку в ядре локализованные и конечные отклики наблюдаются по всему спектру действительной оси *x*.

Параметры ядра определяются моделью классификатора опорных векторов (Support Vectors Classifier, SVC) и в случае использования RBF фактически ограничиваются двумя переменными (https://chrisalbon.com/machine\_learning/support\_ vector\_machines/svc\_parameters\_using\_rbf\_kernel/, https://scikit-learn.org/stable/auto\_examples/svm/ plot rbf parameters.html):

— параметр гамма ( $\gamma$ ) — определяет как далеко распространяется влияние эталона (чем ниже значение, тем дальше влияние, но выше дисперсия и наоборот), рассматривается как обратная



**Рис. 4.** Тестовая выборка для выбора ядра типа SVM на основе.

величина радиуса влияния выборок, выбранных моделью в качестве опорных векторов.

– параметр регуляризации или штрафной параметр (С) – противопоставляет правильность классификации значению решающей (определяющей границы) функции – для больших значений параметра запас решающей функции будет минимальным для всех объектов выборки и наоборот, снижение С будет увеличивать запас функции за счет снижения точности обучения.

Программная реализация метода SVM включает алгоритм поиска оптимальных значений у и С методом кросс-валидации (cross-validation) путем сравнения точности классификации для различных сочетаний параметров по эталонам и по объектам тестовой выборки (Chang, Lin, 2011, https://scikit-learn.org/stable/modules/cross\_validation.html#cross-validation) (puc. 5). Сложность такого подхода состоит в том, что в результате статистического перебора рассчитываются оптимальные интервалы значений параметров, реальную достоверность которых из-за возникающего эффекта переобучения невозможно подтвердить математически, т.е. требуется дополнительный эксперт-

Параметр	Linear: $u' \times v$	Polynomial: (gamma $\times u' \times v +$ coef0)°	Radial basis function: exp $(-\text{gamma} \times  u - v ^2)$	Sigmoid: tanh(gamma $\times$ u' $\times$ v + coef0)		
Gamma – 0; C – 1	Change Run Clear  =0-10-g0-1	Change Run Clear - 6 0-11-g 0-c 1	Charge Run Clear 4-9-12-g-0-1	Change Run Clear 49-13-g 0-51		
Gamma — 10; C — 10	Change Run   Cear  =0.4.0-g 10-c 10	Charge Run   Clear  = 50-11-9 10-5 10	Change Run Clear 4:0-12-910-410	Charge Run Clear 1:6 0-13-g 10-5 10		
Gamma — 100; C — 10	Change Run   Clear  =0.4.0-g 100-c 100	Change Run Clear = 0-41-g-100-< 100	Change Run Clear Is 0-12-g 100 -5 100	Change Run Clear 1-5 0-13 -g 100-<100		

Таблица 1. Результаты классификации тестовой выборки при различных типах и параметрах ядер SVM



**Рис. 5.** Пример определения наилучшего сочетания параметров  $\gamma$  и C с помощью кросс-валидации на примере тестовой выборки (результат  $\gamma = 0.1$ ; C = 1.0, точность 0.97).

ный анализ исходя из особенностей исходных данных и требований к их моделированию.

В программной реализации SVM также предлагается важный методический прием, существенно повышающий точность классификации и сокращающий объем требуемых вычислений — масштабирование (scaling) исходных многопараметрических данных (Chang, Lin, 2011, https://scikit-learn.org/stable/modules/preprocessing.html). Существует два способа масштабирования, которые могут рассматриваться как взаимозаменяемые:

 нормализация (normalization) — линейное приведение значений всех атрибутов объектов классификации к общему интервалу [-1, +1] или [0, 1],

 стандартизация (standardization) – приведение значений всех атрибутов объектов к стандартному гауссовому распределению с нулевым средним значением и границами в пределах единичной дисперсии.

Наиболее распространенной моделью SVC является модель C-SVC на основе параметра регуляризации C, однако в качестве улучшенной версии предложена модель nu-SVM, ограничивающая значения C в пределах [0, 1] (Scholkopf, Smola, 2000). Результаты тестирования показали, что модель nu-SVM не показывает существенной разницы по качеству классификации с C-SVC, но более сложна в оптимизации (требует большего объема вычислений) (Chang, Lin, 2011).

Таким образом, оптимальными исходными параметрами SVM для проведения классификации являются: ядро – RBF, модель классификатора – C-SVC. В рамках исследования требуется оценить влияние масштабирования данных и различных сочетаний параметров γ и С на точность классификации на основе статистической и экспертной оценки.

### РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На основе материалов полевых описаний была составлена легенда, включающая 10 основных классов (табл. 2). Эталоны обучения и верификации выбирались визуально при соблюдении условия равномерного распределения. Всего было выбрано 125 эталонов обучающей выборки (от 3 до 29 на класс) и 34 тестовых эталона (от 2 до 8 на класс).

Ha основе рекомендаций (Chang, Lin, 2011, https://scikit-learn.org/stable/auto\_examples/svm/ plot\_rbf\_parameters.html) и опыта подобных исследований (Qian et al., 2014) были выбраны пары параметров ү и С, для которых выполнена классификация с двумя вариантами масштабирования и

#### Таблица 2. Характеристика классов ландшафтной карты

N⁰	Описание урочища							
Водораздельно-тундровый тип местности								
1	Плоские недренированные поверхности первой морской террасы с мезотрофными местами обводнен- ными осоково- сфагновыми и пушицево-травяно-моховыми болотами в сочетании с травяно-моховыми влажными тундрами и прогрессивно заторфовывающимися хасыреями молодых генераций на глееземах перегнойных тундровых, глееземах торфянистых и торфяных болотных							
2	Плоские слабодреннированные слабонаклонные с мелкокочковатым микрорельефом поверхности пер- вой морской террасы с травяно-моховыми влажными тундрами на глееземах арктических и глееземах перегнойных тундровых							
3	Слабонаклонные с сетью вытянутых ложбин поверхности первой морской террасы с травяно-моховыми влажными тундрами по основной поверхности и ивнячково-травяно-моховыми болотами по ложбинам на глееземах арктических и глееземах торфянистых в сочетании с полигональными тундрами							
	Склоновый тип местности							
4	II. Пологие с мелкокочковатым и местами мелкобугристым микрорельефом склоны долины с травяно- моховыми влажными и лишайниково-мохово-травяными тундрами на арктотундровых слабооглеенных гумусовых почвах и почвах пятен							
Мелк	одолинный тип местности							
5	VIII. 6. Плоские слабодренированные долины рек малых порядков с травяно-моховыми и пушицево- злаковыми влажными тундрами и низинными травяно-моховыми болотами на глееземах торфянистых в сочетании с хасыреями и приозерными понижениями							
Пойм	енный тип местности							
6	VII. 5. Плоские мелкокочковатые по межбугорным понижениям обводненные поверхности высокой поймы с низинными мохово-травяными болотами на торфяных болотных почвах							
7	VII. 4 . Редко затапливаемые поверхности высокой поймы с травяно-моховыми тундрами и низинными осоково-сфагновыми болотами по ложбинам стока							
Проч	ие территории							
8	Песчаные раздувы на дренированных водоразделах первой морской террасы							
9	Водная поверхность, тип 1							
10	Водная поверхность, тип 2							

проведена статистическая оценка ее точности. Для сравнения на примере отдельных сочетаний параметров  $\gamma$  и С, показавших наибольшую достоверность, была выполнена классификация и оценка точности без масштабирования. Кроме того, для ряда сочетаний была выполнена классификация и оценка точности исходного космического снимка (с использованием нормализации). Всего выполнено 142 тестовых классификации, результаты оценки общей точности (Overall Accuгасу) по 96 классификациям представлены в табл. 3. Максимальные показатели составили 0.9547–0.9585.

#### Итоги экспериментов показали следующее:

— наилучшие результаты при стандартизации данных дают значения  $\gamma = 0.01, C = 50 (0.9585)$  и

 $\gamma = 1, C = 0.1$  (0.9554). Увеличение  $\gamma$  более 1 существенно снижает качество классификации для любых значений C;

— наилучшие результаты при нормализации данных дают значения  $\gamma = 0.1$ , C = 100 (0.9547) и  $\gamma = 50$ , C = 1 (0.9585);

 – без нормализации данных точность классификации не превышает 0.78 для любых сочетаний параметров ү и С;

 – классификация исходного несегментированного снимка при условии нормализации данных показывает худшую точность по сравнению с сегментированным изображением для любых со-

Таблица 3. Оценка общей точности (Overall Accuracy) классификации для различных параметров ядра RBF (значения более 0.95 выделены красным; 0.94–0.949 – зеленым; 0.93–0.939 – синим; 0.9–0.929 – черным; менее 0.9 – серым)

Стандартизация		γ									
		0.01	0.1	1	10	50	100				
С	0.1	0.9141	0.9387	0.9554	0.8995	0.7882	0.7463				
	1	0.9395	0.9511	0.9546	0.9021	0.8344	0.7556				
	10	0.9508	0.9431	0.9447	0.9020	0.7980	0.7556				
	50	0.9585	0.9431	0.9416	0.9096	0.7980	0.7556				
	100	0.9502	0.9447	0.9450	0.9020	0.7980	0.7556				
Нормализация		γ									
		0.01	0.1	1	10	50	100				
	0.1	0.8330	0.9192	0.9293	0.9413	0.9584	0.9517				
	1	0.9192	0.9260	0.9442	0.9480	0.9585	0.9446				
С	10	0.9260	0.9526	0.9566	0.9370	0.9408	0.9420				
	50	0.9446	0.9541	0.9456	0.9385	0.9407	0.9412				
	100	0.9531	0.9547	0.9531	0.9366	0.9356	0.9350				
Без масштабирования		γ									
		0.01	0.1	1	10	50	100				
	0.1	0.7584	_	0.7278	—	0.7278	_				
	1	0.7736	_	0.7355	—	0.7355	_				
С	10	0.7754	_	0.7355	—	0.7355	_				
	50	0.7754	_	0.7355	0.7001	0.7355	_				
	100	0.7754	_	0.7355	_	0.7355	_				
Нормализация (снимок)		γ									
					10	50	100				
	мок)	0.01	0.1		10	50	100				
	мок) 0.1	0.01	0.1	0.8583	10	0.8868	-				
	мок) 0.1 1	0.01 0.8123 0.8399	0.1 0.8400 0.8581	1 0.8583 0.8832		0.8868					
С	мок) 0.1 1 10	0.01 0.8123 0.8399 0.8581	0.1 0.8400 0.8581 0.8837	1 0.8583 0.8832 0.8873		0.8868 0.8891 0.8890					
С	мок) 0.1 1 10 50	0.01 0.8123 0.8399 0.8581 0.8812	0.1 0.8400 0.8581 0.8837 0.8874	1 0.8583 0.8832 0.8873 0.8872		30   0.8868   0.8891   0.8890   0.8879					

четаний параметров  $\gamma$  и C (максимальная точность составила 0.8891 при  $\gamma = 50$ , C = 1).

Наглядно видно, что для обоих типов масштабирования существуют интервалы наиболее эффективных значений параметров  $\gamma$  и C, формирующие две обособленные группы. Поскольку не существует математических способов выбора единственного наилучшего значения, для окончательного решения использовался визуальный анализ путем попарного сравнения с исходным снимком. Сравнение показало, что при незначительных отличиях, связанных с точностью выделения долин мелких водотоков и небольших озер, лучший результат достигается при стандартизации с параметрами  $\gamma = 0.01$  и C = 50 (коэффициент Каппа – 0.94623, общая точность – 0.95845). Матрица ошибок по отдельным классам представлена в табл. 4. Наглядно результат классификации показан на рис. 6.

### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

SVM зарекомендовал себя как эффективный метод классификации с обучением, позволяющий получать достаточно высокие результаты при относительно небольшом объеме обучающей выборки (Романов, Рубанов, 2012; Foody, 2004). При этом он менее чувствителен к явлению Хьюза, которое состоит в том, что при ограниченной обучающей выборке точность классификации имеет тенденцию уменьшаться по мере увеличения пространства признаков (Hughes, 1968; Plaza et al., 2009). Это свойство делает его предпочтительным при классификации гиперспектральных данных, а также мультиспектральных данных с

Класс	1	2	9	4	5	10	3	7	8	6	Точность пользователя, %	Точность разработчика, %
1	3065	274	0	460	45	0	24	0	0	0	79	55
2	0	5736	0	0	0	0	0	0	0	0	100	94
9	0	0	39439	0	0	0	0	0	0	0	100	100
4	92	0	0	14084	0	0	13	0	2	0	99	97
5	181	112	0	0	2343	0	0	0	0	0	89	98
10	0	0	0	0	0	2791	0	0	0	0	100	100
3	2203	0	0	0	2	0	19561	0	0	0	90	100
7	0	0	0	0	0	0	0	2486	0	0	100	76
8	0	0	0	0	0	0	0	0	4442	0	100	100
6	0	0	0	0	0	0	0	766	0	2347	75	100
Итого	5541	6122	39439	14544	2390	2791	19598	3252	4444	2347		
Точность, %	55	94	100	97	98	100	100	76	100	100		

**Таблица 4.** Матрица ошибок для классификации с параметрами ядра  $\gamma = 0.01$  и C = 50 (стандартизация), номера классов – см. табл. 2

привлечением дополнительных тематических слоев (ЦМР, NDVI и др.) (Watanachaturaporn et al., 2008, Борзов, Потатуркин, 2017). Также можно отметить, что процесс настройки SVC существенно проще настройки и обучения нейронной сети.

В отношении точности SVM по сравнению с другими алгоритмами можно привести результаты обзора 220 исследований (Ma et al., 2017), из которых 64 используют классификатор NN, 55 – классификатор SVM, 44 – классификатор RF, 33 - классификатор DT, 14 - классификатор максимального правдоподобия (MLC) и 10 – другие классификаторы. Анализ показал, что метод RF обычно характеризуется максимальной общей точностью классификации (85.81%), за которым с небольшим отставанием следуют SVM (85.19%) и DT (84.15%), тогда как показатели NN и MLC в среднем ниже - 81.58 и 81.55% соответственно. Несмотря на то, что исходные данные и обучающая выборка в исследованиях отличаются, результаты оценки точности различных классификаторов в целом отражают реальные возможности алгоритмов с небольшим разбросом значений (Li et al., 2016). Наиболее перспективным подходом в настоящее время является использование нейронных сетей (сверточных нейронных сетей, глубоких сетей доверия и сложенных автоэнкодеров), которые на основе пространственно-спектральных признаков позволяют в отдельных случаях достичь точности 93-99% (Li et al., 2018).

На точность классификации в рамках любого метода оказывает воздействие выбор простран-

ства признаков. В ряде исследований (Борзов, Потатуркин, 2017; Li et al., 2018) показано, что при учете только спектральных или пространственных признаков точность на 6-14% ниже, чем при объектно-ориентированном подходе. Выбор алгоритма сегментации в данном случае является самостоятельной исследовательской задачей, но, при существующих различиях разработанных алгоритмов (Dey, 2010), общим приемом при ландшафтном дешифрировании является сокращение минимальной площади сегмента при сохранении семантического смысла и различимости границ различных фаций в пределах базового урочища. Расширение пространства признаков может происходить как на этапе сегментации, так и позже, на этапе классификации за счет любых дополнительных тематических слоев (морфометрических характеристик, вегетационных и почвенных индексов, сведений о подстилающих породах и др.). Решающую роль при этом играет методическая обоснованность выбора индикаторов с учетом минимизации эффекта слияния сходных по характеристикам классов.

Существует также ряд факторов неопределенности влияющих на точность классификации, связанный с выбором эталонов и объектов тестовых участков (Ma et al., 2017). В большинстве случаев используются алгоритмы случайного (рандомного) выбора сегментов равномерно по всему полю изображения. Тем не менее, для экономии времени в условиях мозаичной структуры ландшафтов представляется целесообразным сокращать



**Рис.** 6. *а* – исходный снимок "Ресурс-П",  $\delta$  – результат классификации при стандартизации данных и параметрах ядра RBF  $\gamma$  = 0.01 и C = 50 (коэффициент Каппа – 0.94623, общая точность – 0.95845).

количество эталонов для однородных участков. Эталоны могут выбираться на основе имеющихся сегментов, что также ускоряет подготовительный этап дешифрирования. Предельные значения объема обучающей и тестовой выборок зачастую определяются эмпирически, однако полноценных исследований на эту тему не проводилось (Li et al., 2014; Ma et al., 2017). Тем не менее отдельные работы (Qian et al., 2014) показывают, что рост числа объектов приводит к незначительному и нелинейному повышению точности. С другой стороны, можно привести пример, демонстрирующий высокую эффективность метода SVM даже в условиях ограниченной обучающей выборки (Романов, Рубанов, 2012). В рамках данного исследования общая площадь обучающей выборки составила 6.5% от общей площади участка, площадь тестовой выборки — 3%.

Необходимо отметить, что недостатком метода SVM является отсутствие учета весовых коэффициентов эталонов, поэтому для сложных классов рекомендуется введение дополнительных уточняющих подклассов с последующим объединением. Также отсутствует возможность введения весовых коэффициентов для индикационных признаков.

Метод SVM в настоящее время является основным в модулях объектно-ориентированного дешифрирования наиболее распространенного коммерческого ПО ENVI. eCognition и Geomatica (https://www.harrisgeospatial.com/docs/backgroundsvmgeneral.html, http://www.pcigeomatics.com/geomatica-help/concepts/focus c/oa classif intro svm.html), не считая пробных попыток внедрения алгоритмов глубокого обучения (deep learning). Судя по доступным руководствам пользователя, во всех случаях аналогично открытым SAGA и OrfeoToolbox используется общедоступная библиотека libSVM (Chang, Lin, 2011), поэтому выбор конкретного ПО с позиции качества результатов SVM не имеет существенного значения. Отличия программной реализации состоят в части интерфейса, быстродействия и доступности выбора отдельных параметров (например, в ENVI отсутствует выбор способа масштабирования исходных данных). Другими словами, наличие открытой библиотеки libSVM позволяет эффективно использовать возможности метода SVM без привязки к ПО, которые дополняются выполнением полготовительных и постклассификационных операций в открытом геоинформационном ПО.

В качестве перспективной альтернативы SVM может рассматриваться использование алгоритмов глубокого обучения, доступных в частности с помощью модуля *otbtf* в ПО OrfeoToolbox (https://www.orfeo-toolbox.org/a-remote-module-fordeep-learning/, https://arxiv.org/pdf/1807.06535.pdf). Фактически модуль является фрейворком открытой библиотеки TensorFlow (https://www.tensorflow.org/), куда интегрированы инструменты потоковой обработки данных ДЗЗ, реализованные в OrfeoToolbox.

Таким образом, результаты исследования показали, что для небольших территорий при оптимальном выборе параметров ядра, равномерной обучающей выборке и достаточном наборе признаков метод SVM позволяет получить точность классификации более 90%, что сопоставимо с уровнем точности нейронных сетей с современной архитектурой, требующих при этом трудоемкого обучения и значительных вычислительных ресурсов. Собственно, SVM может рассматриваться как аналог простой двухслойной нейронной сети, где число нейронов на скрытом слое определяется автоматически как число опорных векторов.

#### ЗАКЛЮЧЕНИЕ

По итогам выполненных работ можно сделать вывод, что подход OBIA для ландшафтного дешифрирования может демонстрировать высокие показатели точности при следующих условиях:

1. Наличие кондиционных исходных данных, включая высокодетальные мультиспектральные снимки и ЦМР с сопоставимым разрешением, кроме этого могут использоваться вспомогательные тематические слои, дающие значимую индикационную информацию.

2. Сегментирование многоканального изображения до уровня различимости границ различных фаций в пределах базового урочища на основе алгоритмов, учитывающих значения пространственного разрешения и эффективных для работы с большими объемами исходных данных.

3. Классификация сегментированного изображения на основе метода SVM, при этом рекомендуется использовать ядро RBF, модель классификатора C-SVC, масштабирование данных (стандартизацию или нормализацию) и параметры ядра  $\gamma$  в интервале [0.01; 1] и C в интервале [10; 100]. Наилучшее сочетание параметров ядра для тестового участка –  $\gamma = 0.01$ , C = 50, метод масштабирования – стандартизация, при этом коэффициент Каппа составил 0.94623, а общая точность – 0.95845.

4. Увеличение числа и площади эталонов обучающей выборки незначительно способствует повышению точности классификации, на тестовом участке площадь эталонов составила 6.5% от общей площади.

Все операции по подготовке данных, классификации и постклассификации (оценки точности) могут быть успешно выполнены в открытом ПО без потери качества (QGIS, SAGA, OrfeoToolbox).

В целом, разработанная на основе OBIA методика крупномасштабного ландшафтного дешифрировании (1 : 10000 и крупнее) показала на тестовом участке свою высокую эффективность, перспективы продолжения работ связаны с построением на ее основе ландшафтно-экологической карты на всю территорию острова Белый.

#### ИСТОЧНИК ФИНАНСИРОВАНИЯ

Исследование выполнено при финансовой поддержке РФФИ и Ямало-Ненецкого автономного округа в рамках научного проекта №19-45-890017.

## СПИСОК ЛИТЕРАТУРЫ

Борзов С.М., Потатуркин О.И. Исследование эффективности спектрально-пространственной классификации данных гиперспектральных наблюдений // Автометрия. 2017. Т. 53. № 1. С. 32–42. https://doi.org/10.15372/AUT20170105

*Вапник В.Н.* Теория распознавания образов // М.: Наука, 1974. 416 с.

Васильчук А.К., Васильчук Ю.К. Инженерно-геологические и геохимические условия полигональных ландшафтов острова Белый (Карское море) // Инженерная геология. 2015. № 1. С. 50–72.

Галиулин Р.В., Башкин В.Н., Алексеев А.О., Галиулина Р.А., Арабский А.К. Остров Белый: рекультивация почв // Neftegaz.RU. 2016 № 7–8. С. 96–100. https://neftegaz.ru/science/ecology/331565-ostrov-belyy-rekultivatsiya-pochv/

Дружинин А.Н., Идрисов И.Р., Маршинин А.В. Опыт картографирования ландшафтов арктической тундры на примере северо-западной части острова Белый (Карское море) // Геоинформационное картографирование в регионах России: Матер. VII Всерос. науч.практ. конф. (Воронеж, 10–12 декабря 2015 г.). Воронеж: Научная книга, 2015. С. 46–54.

Идрисов И.Р., Маршинин А.В., Марьинских Д.М. Опыт крупномасштабного картографирования арктических ландшафтов Западной Сибири // Геодезия и картография. 2017. Т. 78. № 7. С. 31–37.

Козин В.В. Ландшафтный анализ в нефтегазопромысловом регионе. // Тюмень: Изд-во ТюмГУ, 2007. 240 с.

Козин В.В., Холодилов И.В. Ландшафтно-экологический подход при комплексной оценке экологического состояния участков перспективного нефтегазового освоения на Севере Западной Сибири // Вестник Тюменского государственного университета. Социальноэкономические и правовые исследования. 2008. № 3. С. 234–240.

*Мильков Ф.Н.* Физическая география: учение о ландшафте и географическая зональность. Воронеж. 1986. 328 с.

Постановление Правительства Ямало-Ненецкого автономного округа от 25 декабря 2013 года № 1096-П "Об утверждении государственной программы Ямало-Ненецкого автономного округа "Развитие научной, научно-технической и инновационной деятельности на 2014—2021 годы". http://docs.cntd.ru/document/460282983

Романов А.А., Рубанов К.А. Эффективность непараметрических классификаторов в условиях лимитированной обучающей выборки // Журн. СФУ. Техника и технологии. 2012. № 5. С. 495–506.

*Юртаев А.А.* Комплексные исследования почвенного покрова о. Белый: первые итоги // Научный вестник Ямало-Ненецкого автономного округа. 2016. № 4(93). С. 8–11.

Юртаев А.А., Сулкарнаев Ф.Р. Комплексные исследования почвенного покрова о. Белый: промежуточные итоги // Научный вестник Ямало-Ненецкого автономного округа. 2018. № 1(98). С. 74–78.

A remote module for Deep learning [Электронный реcypc] https://www.orfeo-toolbox.org/a-remote-modulefor-deep-learning/, дата обращения 30.07.2019.

Arctic Climate Impact Assessment – Scientific Report. 2006. P. 1046. https://www.amap.no/documents/doc/arctic-arcticclimate-impact-assessment/796

*Boser B.E., Guyon I., Vapnik V.* A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual // Workshop on Computational Learning Theory. 1992. P. 144–152.

*Blaschke T.* Object based image analysis for remote sensing // ISPRS J. Photogrammetry and Remote Sensing. 2010. No 65. P. 2–16.

https://doi.org/10.1016/j.isprsjprs.2009.06.004

Chang C.C., Lin C.J. LIBSVM: a library for support vector machines // ACM Transactions on Intelligent Systems and Technology. 2011.  $\mathbb{N}$  2/3. P. 1–27.

*Conrad O., Bechtel B., Bock M., Dietrich H., Fischer E., Gerlitz L., Webberg J., Wichmann V., Boehner J.* System for Automated Geoscientific Analyses (SAGA) v.2.1.4 // Geosci. Model. 2015. Dev., 8, 1991–2007.

https://doi.org/10.5194/gmd-8-1991-2015

*Comaniciu D., Meer P.* Mean shift: a robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002 № 24(5). P. 603–619. https://doi.org/10.1109/34.1000236

Cortes C., Vapnik V. Support-vector network // Machine Learning. 1995. № 20. P. 273–297.

*Cresson R*. A framework for remote sensing images processing using deep learning techniques [Электронный ресурс] https://arxiv.org/pdf/1807.06535.pdf, дата обращения 30.07.2019.

Cross-validation: evaluating estimator performance [Электронный pecypc] https://scikit-learn.org/stable/modules/cross\_validation.html#cross-validation, дата обращения 30.07.2019.

*Dey V., Zhang Y., Zhong M.* A review on image segmentation techniques with remote sensing perspective // ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS, V. 38. № 7A. P. 31–42.

*Foody G.M.* Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification // Remote Sensing of Environment. 2004.  $\mathbb{N}^{\circ}$  1(93). P. 107–117.

Hughes G.F. On the mean accuracy of statistical pattern recognizers / G.F. Hughes // IEEE Transactions on Information Theory. 1968. № 1(14). P. 55–63.

LIBSVM — A Library for Support Vector Machines [Электронный ресурс] https://www.csie.ntu.edu.tw/~cjlin/libsvm/, дата обращения 30.07.2019.

*Li M., Zang S., Zhang B., Li S., Wu C.* A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information // European J. Remote Sensing. 2014. № 47(1). P. 389–411.

https://doi.org/10.5721/EuJRS20144723

Li M., Ma L., Blaschke T., Cheng L., Tiede. D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments // International J. Applied Earth Observation and Geoinformation. 2016. № 49. P. 87–98. https://doi.org/10.1016/j.jag.2016.01.011

Li Y., Zhang H., Xu, X., Jiang Y., Shen Q. Deep learning for remote sensing image classification: A survey // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018. № 1264.

https://doi.org/10.1002/widm.1264

*Ma L., Li M., Ma X., Cheng L., Du P., Liu Y.* A review of supervised object-based land-cover image classification // ISPRS J. Photogrammetry and Remote Sensing. 2017. V. 130. P. 277–293.

https://doi.org/10.1016/j.isprsjprs.2017.06.001

*Mountrakis G., Im J., Ogole C.* Support vector machines in remote sensing: A review // ISPRS Journal of Photogrammetry and Remote Sensing. 2011. № 66(3). P. 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001 *Plaza A.* Recent advances in techniques for hyperspectral images / Plaza A, Benediktsson J.A., Boardman J.W., Brazile J., Bruzzone L. et al. // Remote Sensing of Environment. 2009. № 10(113). P. 110–122.

Preprocessing data [Электронный ресурс] https://scikitlearn.org/stable/modules/preprocessing.html, дата обращения 30.07.2019.

RBF SVM parameters [Электронный pecypc] https://scikitlearn.org/stable/auto\_examples/svm/plot\_rbf\_parameters.html, дата обращения 30.07.2019.

Scholkopf B., Smola A., Williamson R. C., Bartlett P. L. New support vector algorithms. // Neural Computation. 2000. V. 12. P. 1207–1245.

Support Vector Machine Background [Электронный pecypc] https://www.harrisgeospatial.com/docs/back-groundsvmgeneral.html, дата обращения 30.07.2019.

Support Vector Machine classifier [Электронный ресурс] http://www.pcigeomatics.com/geomatica-help/con-

cepts/focus\_c/oa\_classif\_intro\_svm.html, дата обращения 30.07.2019.

SVC Parameters When Using RBF Kernel [Электронный pecypc] https://chrisalbon.com/machine\_learning/support\_vector\_machines/svc\_parameters\_using\_rbf\_kernel/, дата обращения 30.07.2019.

TensorFlow [Электронный ресурс] https://www.tensor-flow.org/, дата обращения 30.07.2019.

Trimble. Ecognition Developer 8.7 Reference Book; Trimble: Munich, Germany, 2011.

*Qian Y., Zhou W., Yan J., Li W., Han L.* Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery // Remote Sensing. 2014. V. 7 № 1. P. 153–168. https://doi.org/10.3390/rs70100153

*Waske B., Benediktsson J.A.* Fusion of support vector machines for classification fmultisensor data // IEEE Transactions in Geoscience and Remote Sensing. 2007. V. 45. P. 3858–3866.

Watanachaturaporn P., Arora M.K., Varshney P.K. Multisource classification using support vector machines: An empirical comparison with decision tree and neural network classifiers // Photogrammetric Engineering and Remote Sensing, 2008. V. 74. № 2. P. 239–246.

[Электронный pecypc] https://www.pgc.umn.edu/data/ arcticdem/, дата обращения 30.07.2019.

[Электронный pecypc] https://www.researchgate.net/ publication/265031337\_Semi-Automatic\_Classification\_ Plugin User Manual, дата обращения 30.07.2019.

[Электронный pecypc] https://qgis.org/ru/site/, дата обращения 30.07.2019.

[Электронный pecypc] https://www.orfeo-toolbox.org/ CookBook/Applications/app\_LargeScaleMeanShift.html, дата обращения 30.07.2019.

# Refinement of Classification Parameters by the Support Vector Machines (SVM) for Large-Scale Mapping of Arctic Landscapes (Using the Example of Bely Island, Kara Sea)

O. S. Sizov<sup>1, 2, 3</sup>, I. R. Idrisov<sup>4</sup>, and A. A. Yurtaev<sup>4</sup>

<sup>1</sup>Research Center for Earth Operative Monitoring of JSC "Russian Space Systems", Moscow, Russia <sup>2</sup>Institute of Oil and Gas Problems RAS, Moscow, Russia <sup>3</sup>Russian State University of Oil and Gas (National Research University) named after I.M. Gubkin, Moscow, Russia <sup>4</sup>Tyumen State University, Tyumen, Russia

On the example of the White Island, located in the Kara Sea, the issues of automation of landscape mapping are being studied. The choice of the island as a test facility is determined by the decision of the government of the Ya-malo-Nenets District to create a scientific hospital for planned fundamental research on the consequences of an-thropogenic impact on Arctic ecosystems. Based on the example of a small test site, field study of which was carried out in the summer of 2016, topical methodological approaches to landscape mapping based on object-oriented image analysis (OBIA) and SVM classification are considered. A combined method based on independent processing of initial data with subsequent statistical processing of the information obtained is proposed. The results of the study showed that such a method allows for a more detailed approach to assessing the landscape differentiation of the research area. For the first time for the territory of Bely Island, highly detailed images from the Resurs-P satellite and digital models of the ArcticDEM relief are used. The use of object-oriented interpretation will allow further expanding the set of initial data due to the materials of shooting with UAV, and SVM classification results can be improved by additional thematic layers of morphometric analysis of digital terrain models. The proposed technology is applicable for performing large-scale landscape mapping of the entire territory of White Island.

*Keywords*: Beliy Island, SVM classification, remote sensing data, Resurs-P, landscape mapping, segmentation, object-based image analysis

## REFERENCES

*Borzov S.M., Potaturkin O.I.* Issledovanie effektivnosti spektralino-prostranstvennoy klassifikatsii dannuih giper-spektral'nyh nablyudeniy [Study of the effectiveness of spectral-spatial classification of hyperspectral observation

data] // Avtometriya. 2017. T. 53. № 1. P. 32–42. doi:10.15372/AUT20170105

*Vapnik V.N.* Teoriya raspoznovaniya obrazov [Pattern recognition theory] M.: Nauka, 1974. 416 p.

Vasil'chuk A.K., Vasil'chuk Yu.K. Inzhenerno-geologicheskie I geohimicheskie usloviya poligonal'nyh lanshaftov os-

90

trova Belyi (Karskoe more) Engineering-geological and geochemical conditions of polygonal landscapes of Bely Island (Kara Sea) // Inzhenernaya geologiya. 2015.  $\mathbb{N}_{2}$  1. P. 50–72.

Galiulin R.V., Bashkin V.N., Alekseev A.O., Galiulina R.A., Arabskiy A.K. Ostrov Belui: rekul'tivatsiya pochv [White Island: soil restoration] // Neftegaz. RU. 2016 № 7–8. P. 96–100. https://neftegaz.ru/science/ecology/331565-ostrov-belyyrekultivatsiya-pochv/

*Druzhinin A.H., Idrisov I.R., Marshinin A.V.* Opyt kartografirovaniya landshaftov arkticheskoy tundry na primere severo-zapadnoy chasti ostrova [The experience of mapping landscapes of the Arctic tundra using the example of the northwestern part of Bely Island (Kara Sea] // Gioinformatsionnoe kartografirovanie v regionah Rossii: Mater. VII Vseros. nauch.-prakt. konf (Voronezh, 10–12 dekabrya 2015) Voronezh: Nauchnaya kniga, 2015 P. 46–54.

*Idrisov I.R., Marshinin A.V., Mar'inskih D.M.* Opyt krupnomasshtabnogo kartografirovaniya arkticheskih landshaftov Zapadnoy Sibiri [The experience of large-scale mapping of Arctic landscapes in Western Siberia] // Geodeziya I kartografiya. 2017. T. 78. № 7. P. 31–37.

*Kozin V.V.* Landshaftnyi analiz v neftegazopromyslovom regione[Landscape analysis in the oil and gas region]. Tyumen': Izd-vo TyumGU, 2007. 240 p.

*Kozin V.V., Holodilov I.V.* Landshaftno-ekologicheskiy podhod pri kompleksnoy otsenke ekologicheskogo sostoyaniya uchastkov perspektivnogo neftegasovogo osvoeniya na Severe Zapadnoy Sibiri [A landscape-ecological approach to a comprehensive assessment of the ecological condition of prospective oil and gas development sites in the North of Western Siberia] // Vestnik Tyumenskogo gosudarstvennogo universiteta. Sotsial'no-ekonomicheskie I pravovye issledovaniya. 2008. № 3. P. 234–240.

*Mil'kov F.N.* Fizicheskaya geografiya: uchenie o landshafte i geograficheskaya zonal'nost'. [Physical geography: the doctrine of the landscape and geographical zoning]. Voronezh, 1986. 328 p.

Postanovlenie Pravitel'stva Yamalo-Nenetskogo avtonomnogo okruga ot 25.12.2013 № 1096-P "Ob utverzhdenii gosudarstvennoy programmy Yamalo-Nenetskogo avtonomnogo okruga "Razvitie nauchnoy, nauchno-tehnicheskoy i innovatsionnoy deyatel'nosti na 2014–2021 gody". http://docs.cntd.ru/document/460282983

*Romanov A.A., Rubanov K.A.* Effektivnostj neparametricheskih klassifikatorov v usloviyah limitirovannoy obuchauschey vuborki [The effectiveness of nonparametric classifiers in the conditions of a limited training sample] // Zhurnal SFU. Tehnika i tehnologii. 2012. № 5. P. 495–506.

*Yurtaev A.A.* Kompleksnye issledovaniya pochvennogo pokrova o.Belyi: pervue itogi [Comprehensive studies of the soil cover of Bely Island: first results] // Nauchnui vestnik Yamalo-Nenetskogo avtonomnogo okruga. 2016. № 4(93). P. 8–11.

Yurtaev A.A., Sulkarnaev F.R. Kompleksnye issledovaniya pochvennogo pokrova o.Belyi: promezhutocnye itogi [Comprehensive studies of the soil cover of Bely Island: interim results] // Nauchnyi vestnik Yamalo-Nenetskogo avtonomnogo okruga 2018. № 1(98). P. 74–78.

A remote module for Deep learning [Электронный реcypc] https://www.orfeo-toolbox.org/a-remote-modulefor-deep-learning/, дата обращения 30.07.2019. Arctic Climate Impact Assessment – Scientific Report. 2006. 1046 p. https://www.amap.no/documents/doc/arc-tic-arctic-climate-impact-assessment/796

*Boser B.E., Guyon I., Vapnik V.* A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, ACM Press, 1992. P. 144–152.

*Blaschke T.* Object based image analysis for remote sensing. // ISPRS J. Photogrammetry and Remote Sensing. 2010. V. 65. P. 2–16. doi.org/10.1016/j.isprsjprs.2009.06.004

*Chang C.-C., Lin C.-J.* LIBSVM: a library for support vector machines. // ACM Transactions on Intelligent Systems and Technology. 2011. V. 2/3. P. 1–27.

*Conrad O., Bechtel B., Bock M., Dietrich H., Fischer E., Gerlitz L., Webberg J., Wichmann V., Boehner J.* System for Automated Geoscientific Analyses (SAGA) v.2.1.4. // Geosci. Model. 2015. Dev., 8, 1991-2007, doi:10.5194/gmd-8-1991-2015

*Comaniciu D., Meer P.* Mean shift: a robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. № 24(5). P. 603–619. doi:10.1109/34.1000236

Cortes C., Vapnik V. Support-vector network // Machine Learning. 1995 № 20. P. 273–297.

*Cresson R*. A framework for remote sensing images processing using deep learning techniques [Электронный ресурс] https://arxiv.org/pdf/1807.06535.pdf, дата обращения 30.07.2019.

Cross-validation: evaluating estimator performance [Электронный pecypc] https://scikit-learn.org/stable/modules/cross\_validation.html#cross-validation, дата обращения 30.07.2019.

*Dey V., Zhang Y., Zhong M.* A review on image segmentation techniques with remote sensing perspective // ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7. 2010. IAPRS. V. XXXVIII. Part 7A. P. 31–42.

*Foody G.M.* Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification // Remote Sensing of Environment. 2004.  $N^{\circ}$  1 (93). P. 107–117.

Hughes G.F. On the mean accuracy of statistical pattern recognizers / G.F. Hughes // IEEE Transactions on Information Theory.1968. № 1(14). P. 55–63.

LIBSVM — A Library for Support Vector Machines [Электронный ресурс] https://www.csie.ntu.edu.tw/~cjlin/libsvm/, дата обращения 30.07.2019.

*Li M., Zang S., Zhang B., Li S., Wu C.* A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information // European J. Remote Sensing. 2014. № 47(1). P. 389–411. doi: 10.5721/EuJRS20144723

Li M., Ma L., Blaschke T., Cheng L., Tiede D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. // International J. Applied Earth Observation and Geoinformation. 2016 № 49. P. 87–98. doi:10.1016/j.jag.2016.01.011

Li Y., Zhang H., Xue X., Jian Y., Shen Q. Deep learning for remote sensing image classification: A survey // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018. № 1264. doi: 10.1002/widm.1264

Ma L., Li M., Ma X., Cheng L., Du P., Liu Y. A review of supervised object-based land-cover image classification // IS- PRS Journal of Photogrammetry and Remote Sensing. 2017. V. 130. P. 277–293. doi: 10.1016/j.isprs-jprs.2017.06.001

*Mountrakis G., Im J., Ogole C.* Support vector machines in remote sensing: A review // ISPRS Journal of Photogrammetry and Remote Sensing. 2011. № 66(3). P. 247–259. doi: 10.1016/j.isprsjprs.2010.11.001

*Plaza A.* Recent advances in techniques for hyperspectral images / A. Plaza, J.A. Benediktsson, J.W. Boardman, J. Brazile, L. Bruzzone et al. // Remote Sensing of Environment. 2009. № 10(113). P. 110–122.

Preprocessing data [Электронный pecypc] https://scikitlearn.org/stable/modules/preprocessing.html, дата обращения 30.07.2019.

RBF SVM parameters [Электронный pecypc] https://scikitlearn.org/stable/auto\_examples/svm/plot\_rbf\_parameters.html, дата обращения 30.07.2019.

Scholkopf B., Smola A., Williamson R.C., Bartlett P.L. New support vector algorithms. // Neural Computation. 2000. № 12. P. 1207–1245.

Support Vector Machine Background [Электронный реcypc] https://www.harrisgeospatial.com/docs/backgroundsvmgeneral.html, дата обращения 30.07.2019.

Support Vector Machine classifier [Электронный ресурс] http://www.pcigeomatics.com/geomatica-help/con-

cepts/focus\_c/oa\_classif\_intro\_svm.html, дата обращения 30.07.2019.

SVC Parameters When Using RBF Kernel [Электронный pecypc] https://chrisalbon.com/machine\_learning/sup-

port\_vector\_machines/svc\_parameters\_using\_rbf\_ker-nel/, дата обращения 30.07.2019.

TensorFlow [Электронный ресурс] https://www.tensor-flow.org/, дата обращения 30.07.2019.

Trimble. Ecognition Developer 8.7 Reference Book; Trimble: Munich, Germany, 2011.

*Qian Y., Zhou W., Yan J., Li W., Han L.* Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery // Remote Sensing. 2011. № 7(1). P. 153–168. doi: 10.3390/rs70100153

Waske B., Benediktsson J.A. Fusion of support vector machines for classification fmultisensor data // IEEE Transactions in Geoscience and Remote Sensing. 2007. № 45. P. 3858–3866.

Watanachaturaporn P, Arora M.K., Varshney P.K. Multisource classification using support vector machines: An empirical comparison with decision tree and neural network classifiers // Photogrammetric Engineering and Remote Sensing. 2008. № 74(2). P. 239–246.

[Электронный pecypc] https://www.pgc.umn.edu/data/arcticdem/, дата обращения 30.07.2019.

[Электронный pecypc] https://www.researchgate.net/publication/265031337\_Semi-Automatic\_Classification\_Plugin\_ User\_Manual, дата обращения 30.07.2019.

[Электронный pecypc] https://qgis.org/ru/site/, дата обращения 30.07.2019.

[Электронный pecypc] https://www.orfeo-toolbox.org/Cook-Book/Applications/app\_LargeScaleMeanShift.html, дата обращения 30.07.2019.