

Структурно-параметрический синтез тематического слоя объектов транспортной инфраструктуры на основе гиперспектральной информации

Т.И. Михеева^а, О.К. Головнин^а, А.А. Федосеев^б

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

^б Ракетно-космический центр «Прогресс», 443009, ул. Земеца, 18, Самара, Россия

Аннотация

Описан метод синтеза тематического слоя в интеллектуальной транспортной геоинформационной системе с отображением имманентных свойств элементов транспортной сети. Структурно-параметрический синтез тематического слоя формируется на основе гиперспектральной информации определения состояния дорожного покрытия. Тематический слой представляет собой пространственное распределение значений атрибутов объектов класса «Проезжая_часть» в границах, соответствующих размерам исходного изображения.

Ключевые слова: тематический слой; транспортная инфраструктура; гиперспектральная информация; дорожное покрытие; геоинформационная система

1. Введение

Увеличение протяжённости и качества автомобильных дорог относится к ряду приоритетных мероприятий по модернизации транспортных систем регионов. При этом актуальной становится задача оперативного выявления изменений характеристик и состояния объектов транспортной инфраструктуры [1]. Значительная площадь территории и большое число субъектов РФ диктуют необходимость разработки и внедрения специальных подходов к решению указанной задачи на основе использования данных дистанционного зондирования Земли (ДЗЗ). Современная стратегия развития интеллектуальных транспортных систем предполагает использование беспилотных летательных аппаратов (БПЛА). Перечень полезной нагрузки БПЛА включает в себя линейку инструментов, в том числе, маломассогабаритные гиперспектральные сенсоры, позволяющие получать информацию об объектах ТрИ в виде «гиперкуба», включающего в себя как пространственную информацию об объекте, так и спектральную информацию по каждой пространственной координате. Важным аспектом использования гиперспектральной информации является разработка специализированных подходов к её интерпретации с целью высокоточного распознавания объектов интереса, присутствующих в рассматриваемой сцене, и последующего построения тематического слоя объектов транспортной инфраструктуры (ТрИ).

Рассмотрим направления использования гиперспектральной информации в обеспечение синтеза тематического слоя объектов транспортной инфраструктуры:

- определение состояния дорожного покрытия;
- распознавание и классификация на снимке объектов ТрИ, соответствующих заранее определённым классам.

В качестве объектов ТрИ, подлежащих распознаванию и классификации на основе гиперспектральной информации, определим объекты следующих классов:

- «Проезжая_часть» $\tilde{t}_i^d \in \tilde{T}^D$;
- «Обочина» $\tilde{t}_i^w \in \tilde{T}^W$;
- «Тротуар» $\tilde{t}_i^f \in \tilde{T}^F$;
- «Велосипедная_дорожка» $\tilde{t}_i^c \in \tilde{T}^C$;
- «Разделительная_полоса» $\tilde{t}_i^s \in \tilde{T}^S$;
- «Разметка» $\tilde{t}_i^m \in \tilde{T}^M$;
- «Искусственное_сооружение» $\tilde{t}_i^b \in \tilde{T}^B$;
- «Природный_объект» $\tilde{t}_i^g \in \tilde{T}^G$.

2. Синтез тематического слоя состояния дорожного покрытия

Подход к определению состояния дорожного покрытия предполагает реализацию следующих задач:

- распознавание объектов класса «Проезжая_часть»;

- расчёт эмпирических соотношений *ARVNIRI* (*Asphalt Road Visible Near InfraRed Index*) (см. формулу (1)) и *ARSWIRI* (*Asphalt Road Short Wave InfraRed Index*) (см. формулу (2)), описывающих состояние покрытия [2,3].

$$ARVNIRI = \rho^{830} - \rho^{490}, \quad (1)$$

где ρ^{830} – значение отражательной способности на длине волны $\lambda=830$ нм, ρ^{490} – значение отражательной способности на длине волны $\lambda=490$ нм.

$$ARSWIRI = \rho^{2120} - \rho^{2340}, \quad (2)$$

где ρ^{2120} – значение отражательной способности на длине волны $\lambda=2120$ нм, ρ^{2340} – значение отражательной способности на длине волны $\lambda=2340$ нм.

Объект класса «*Проезжая_часть*», представленный на гиперспектральном снимке, характеризуется следующими имманентными свойствами:

- рост отражательной способности, как функции от длины волны, хорошо прослеживается в ближнем и коротковолновом инфракрасных диапазонах при отражательной способности более 10% (на основе гиперспектральной информации) [4]:

$$(\Delta\rho^{[\lambda_1, \lambda_2]} > 0) \text{ if } (\rho^{[\lambda_1, \lambda_2]} > 10\%), \quad (3)$$

где λ_1 – значение длины волны, соответствующее левой границе ближнего инфракрасного (ИК) диапазона спектра, λ_2 – значение длины волны, соответствующее правой границе коротковолнового ИК диапазона спектра;

- поглощение излучения оксидами железа на длинах волн $\lambda=520$ нм, $\lambda=670$ нм и $\lambda=870$ нм, проявляющееся на спектральном профиле в виде «ямы» с центром, соответствующим указанным длинам волн;
- увеличение отражательной способности для старых (изношенных) дорожных покрытий (с минимальными значениями индекса состояния дорожного покрытия *PCI*, широко распространённого в международной практике) в диапазоне длин волн от 2100 до 2200 нм и уменьшение в диапазоне от 2250 до 2300 нм:

$$\{PCI \rightarrow \min\} \text{ if } \begin{cases} \Delta\rho^{[\lambda_{2100}, \lambda_{2200}]} > 0; \\ \Delta\rho^{[\lambda_{2250}, \lambda_{2300}]} < 0 \end{cases}; \quad (4)$$

- высокое значение индекса *ARVNIRI* и *ARSWIRI* для старых (изношенных) дорожных покрытий (с минимальными значениями индекса состояния дорожного покрытия *PCI*):

$$\{PCI \rightarrow \min\} \text{ if } \begin{cases} (ARVNIRI = \rho^{830} - \rho^{490}) \rightarrow \max \\ (ARSWIRI = \rho^{2120} - \rho^{2340}) \rightarrow \min \end{cases}. \quad (5)$$

Модель процесса определения состояния дорожного покрытия представляет собой математическое описание процесса построения тематического слоя \tilde{S}^C состояния дорожного покрытия объектов класса «*Проезжая_часть*» на основе исходного гиперспектрального снимка \tilde{I} :

$$M_{\tilde{S}^C} = F_{\tilde{S}^C}(\tilde{I}, \tilde{S}^C). \quad (6)$$

Тематический слой \tilde{S}^C представляет собой пространственное распределение значений атрибута *condition^{TD}* объектов класса «*Проезжая_часть*» в границах, соответствующих размерам исходного изображения.

Пространственное распределение значений атрибута *condition^{TD}* объектов класса «*Проезжая_часть*» определяется выражениями (1) и (2).

Диапазон значений изображения, полученного в результате вычислений согласно (1):

$$ARVNIRI = [0, \rho_{\max}^{ARVNIRI}], \quad (7)$$

где $\rho_{\max}^{ARVNIRI}$ – максимальное значение индекса *ARVNIRI*.

Диапазон значений изображения, полученного в результате вычислений согласно (2):

$$ARVNIRI = [0, \rho_{\max}^{ARSWIRI}], \quad (8)$$

где $\rho_{\max}^{ARSWIRI}$ – максимальное значение индекса *ARSWIRI*.

Каждый из диапазонов, определённых (7) и (8), разбивается на 4 части, каждая из которых соответствует определённому значению атрибута $condition^{TD}$.

Пространственное распределение значений атрибута $condition^{TD}$ объектов класса «Проезжая_часть» определяется следующим образом:

$$\{condition^{TD} = отличное\}if\{ARVNIRI = [0, \rho_1^{ARVNIRI}], ARSWIRI = [\rho_3^{ARSWIRI}, \rho_{max}^{ARSWIRI}]\}, \quad (9)$$

где $\rho_1^{ARVNIRI}$ – правая граница первой части, а также левая граница второй части диапазона значений $ARVNIRI$; $\rho_3^{ARSWIRI}$ – левая граница четвёртой части, а также правая граница третьей части диапазона значений $ARSWIRI$;

$$\{condition^{TD} = хорошее\}if\{ARVNIRI = [\rho_1^{ARVNIRI}, \rho_2^{ARVNIRI}], ARSWIRI = [\rho_2^{ARSWIRI}, \rho_3^{ARSWIRI}]\}, \quad (10)$$

где $\rho_2^{ARVNIRI}$ – правая граница второй части, а также левая граница третьей части диапазона значений $ARVNIRI$; $\rho_2^{ARSWIRI}$ – левая граница третьей части, а также правая граница второй части диапазона значений $ARSWIRI$;

$$\{condition^{TD} = нормальное\}if\{ARVNIRI = [\rho_2^{ARVNIRI}, \rho_3^{ARVNIRI}], ARSWIRI = [\rho_1^{ARSWIRI}, \rho_2^{ARSWIRI}]\}, \quad (11)$$

где $\rho_3^{ARVNIRI}$ – правая граница третьей части, а также левая граница четвёртой части диапазона значений $ARVNIRI$; $\rho_2^{ARSWIRI}$ – левая граница второй части, а также правая граница первой части диапазона значений $ARSWIRI$;

$$\{condition^{TD} = плохое\}if\{ARVNIRI = [\rho_3^{ARVNIRI}, \rho_{max}^{ARVNIRI}], ARSWIRI = [0, \rho_1^{ARSWIRI}]\}. \quad (12)$$

Схема определения состояния дорожного покрытия представлена на рис. 1

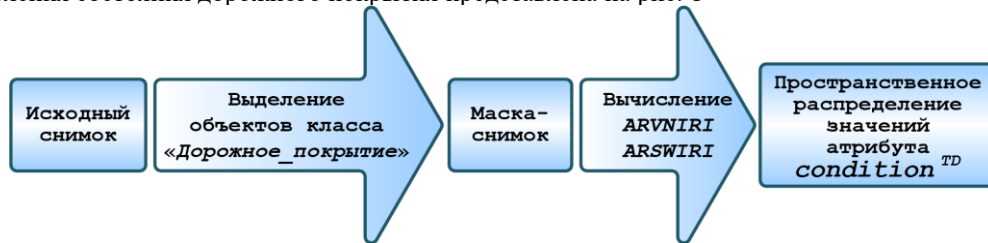


Рис. 1. Схема определения состояния дорожного покрытия.

Для оценки состояния асфальтового покрытия использовался индекс $ARVNIRI$, рассчитывающийся с учётом имеющихся исходных данных следующим образом:

$$ARVNIRI = \rho^{831} - \rho^{489} \quad (13)$$

где ρ^{831} и ρ^{489} – отражательная способность соответственно для длин волн $\lambda=831$ нм и $\lambda=489$ нм.

Анализ состояния асфальтового покрытия с использованием данного индекса производился на основе фрагмента снимка с изображением кольцевого участка автодороги. Сначала построена маска транспортной сети, затем выполнена указанная арифметическая операция, в результате чего получено цветное индексное изображение, характеризующее состояние покрытия (рис. 2). Фиолетовый цвет соответствует «свежему» асфальтовому покрытию, тёмно-красный цвет соответствует старому и изношенному покрытию.

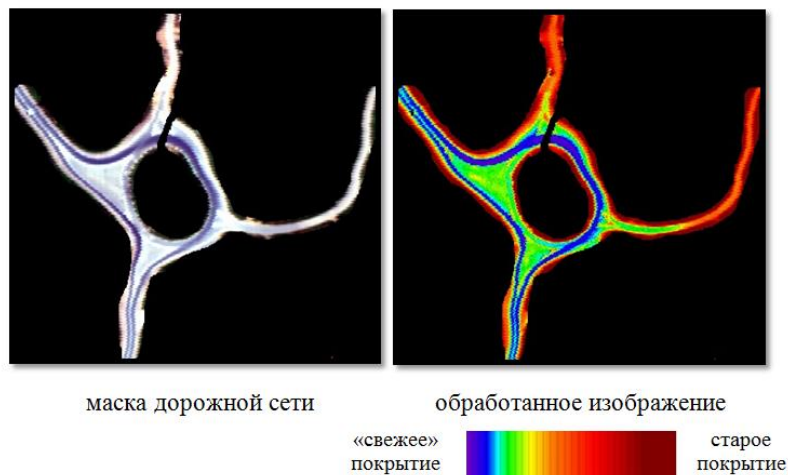


Рис. 2. Результат анализа состояния дорожного покрытия.

3. Синтез тематического слоя объектов транспортной инфраструктуры

В основе процесса синтеза тематического слоя объектов ТрИ лежит задача распознавания и классификации на гиперспектральном снимке этих объектов.

Распознаванию и классификации подлежат значения атрибутов объектов:

- для объектов класса «*Проезжая_часть*» атрибуты $material^{TD} \in \{асфальт, бетон, асфальтобетон, грунт, гравий, щебень, грунтощебень\}$ и $condition^{TD} \in \{отличное, хорошее, удовлетворительное, плохое\}$;
- для объектов класса «*Обочина*» атрибут $material^{TW} \in \{грунт, щебень, грунтощебень\}$;
- для объектов класса «*Тротуар*» атрибут $material^{TF} \in \{асфальт, бетон, асфальтобетон, грунт\}$;
- для объектов класса «*Велосипедная_дорожка*» атрибут $material^{TC} \in \{асфальт, бетон, асфальтобетон\}$;
- для объектов класса «*Разделительная_полоса*» атрибут $material^{TS} \in \{асфальт, бетон, асфальтобетон, грунт, растительность\}$;
- для объектов класса «*Природный_объект*» атрибут $type^{TG} \in \{трава, деревья, открытая_почва, вода\}$.

В качестве методического инструментария для решения задачи распознавания и классификации объектов ТрИ на гиперспектральных снимках выступает подход на основе использования свёрточных нейронных сетей, хорошо себя зарекомендовавший при обработке трёхканальных изображений [5, 6, 7], а также распознавании рукописных цифр и текста за счёт высокой точности результатов обработки при сравнительно невысоких вычислительных затратах. Использование свёрточных сетей для решения поставленной задачи выглядит более предпочтительным [8, 9], чем применение классических нейронных сетей, сопряжённое со значительными затратами вычислительных ресурсов на обучение сети и дальнейшее распознавание объектов (за счёт большой размерности обрабатываемых изображений и, следовательно, большого числа нейронов и синаптических связей в сети), а также более эффективным по сравнению с рядом наиболее распространённых методов обработки гиперспектральных изображений (например, такими как базовые методы Data Mining: метод максимального правдоподобия, метод спектрального угла, метод расстояния Махаланобиса и пр.) [1, 10, 11].

Архитектура и алгоритмы обучения свёрточных нейронных сетей подобны классическим нейронным сетям прямого распространения, для обработки больших массивов входных данных они оптимизированы. В отличие от классических нейронных сетей свёрточные сети не являются полносвязными. Входной слой в свёрточных нейронных сетях представлен в виде 3D массива данных. При прохождении через слои нейронной сети происходит трансформация обрабатываемого массива данных и в конечном счёте его сведение к одномерному массиву, который обрабатывается классической нейронной сетью прямого распространения. Подобное преобразование с сохранением высокой обучающей способности требует большого числа слоёв.

Данная сеть не является полносвязной. Преимуществом распознавания гиперспектральных изображений данной сетью является учёт как спектральных признаков объектов, присутствующих на снимке, так и пространственных. Свёрточная нейронная сеть для обработки гиперспектральных изображений состоит из следующих типов промежуточных слоёв (рис. 3):

- свёрточный слой;
- обобщающий слой;
- слой с линейной функцией активации;
- слой со случайным отключением связи нейронов;
- полносвязный слой.

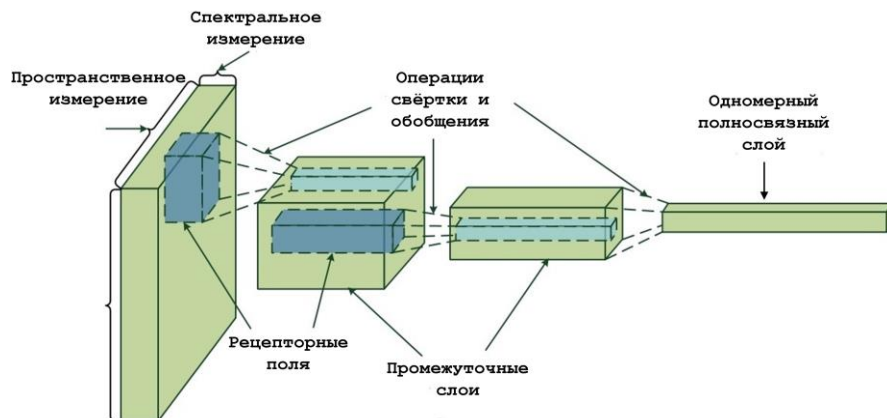


Рис. 3. Свёрточная нейронная сеть.

Свёрточный слой служит для выявления характерных признаков (имманентных свойств) объектов в соответствии с заранее обученными паттернами. Каждый следующий промежуточный слой сети связан с небольшим количеством нейронов предыдущего слоя, которые объединяет нахождение в некоторой небольшой локальной области – рецепторном поле. Использование одних и тех же весовых коэффициентов для всех рецепторных полей одного свёрточного слоя ускоряет обучение и работу нейронной сети. При проектировании свёрточного слоя важно правильно подобрать параметры глубины выходного массива, шага смещения рецепторного поля и дополнения нулями. Вариацией глубины выходного массива контролируется число распознаваемых слоев признаков. Величина шага отвечает за перекрытие рецепторных полей. Чтобы не «перешагнуть» искомый признак при смещении рецепторного поля для проектируемой сети задано значение шага $s=1$. Дополнение нулями не используется, так как предназначено для сохранения оригинального размера изображения на входе слоя сети.

Обобщающий слой позволяет выбрать наиболее сильный сигнал из рассматриваемой области и редуцировать размер массива данных. Обобщающий слой выполняет уменьшение размерности сформированных карт признаков. Нейроны этого слоя не используют параметры и, следовательно, не нуждаются в обучении. На практике широко используется два типа обобщения: выбор максимального значения и вычисление среднего значения. В проектируемой сети используется первый тип обобщения, обеспечивающий определение доминантных признаков путём выбора самого сильного сигнала из рецепторного поля.

Слой с линейной функцией активации позволяет внести элемент нелинейности в распространяемый по сети сигнал. Операции свёртки и обобщения являются линейными и для получения более сложных закономерностей необходимо включение функции активации в цепь обработки сигнала. Обычно в нейронных сетях для этого используются функции сигмоида или гиперболического тангенса. Эти функции являются нелинейными, однако, приводят к проблемам с затуханием или увеличением градиентов. Для решения данной проблемы в работе предложено использовать выпрямленную линейную функцию активации *ReLU*, показавшую свою эффективность в ходе проведённых исследований. Производная данной функции равна единице либо нулю и поэтому не может произойти возрастания или затухания градиентов. Работа данного слоя не нуждается в настройке весов, а вычисление функции сводится к вычислению максимума между нулём и значением сигнала.

Слой со случайным отключением связей *Dropout* служит для повышения обобщающих свойств сети. При продолжительном обучении нейронной сети по обучающей выборке существует большая вероятность возникновения проблемы переобучения. Сеть при этом может давать хороший результат на обучающей выборке, но плохо распознавать объекты на новых данных. Для предотвращения проблемы переобучения нейронной сети, в работе предложено использовать метод случайного отключения связей нейронов с заданием вероятности отключения. Таким образом, в слое рассматриваемой нейронной сети половина сигналов отбрасывается при обучении, что, в итоге, приводит к настройке параметров сети таким образом, что она хорошо распознает объекты на гиперспектральных изображениях, не использованных при обучении.

На конечном этапе обработки данных используется полносвязный слой, который определяет принадлежность объектов ТрИ, описываемых входным массивом данных, к определённому классу.

Обучение свёрточной нейронной сети проводится методом обратного распространения ошибки с использованием алгоритма *Adam* [12].

Для настройки нейросети использовался тестовый набор гиперспектральных данных «*Pavia University*», имеющий разметку классов объектов (рис. 4а – исходное изображение (синтез в естественных цветах), рис. 4б – эталонная классификация).

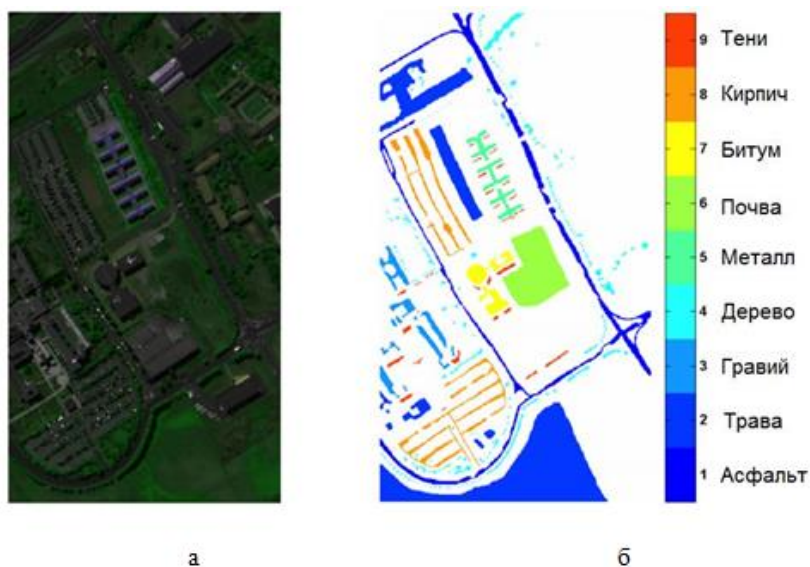


Рис. 4. Тестовое изображение «*Pavia University*».

Настройка архитектуры сети производилась путём подбора оптимальных значений следующих параметров:

- размер рецепторного поля;
- значение шага смещения фильтра;
- число карт признаков;
- тип слоя подвыборки (*max/general*);
- тип функции активации;
- алгоритм обучения.

По результатам настройки выбрана оптимальная архитектура сети: размер рецепторного поля 5×5 , значение смещения фильтра $M = 5$, слой подвыборки – *maxpooling*. В качестве алгоритма обучения выбран алгоритм *Adam*, показавший наилучшие характеристики по сравнению с рядом других алгоритмов (рис. 5). В качестве функции активации выбрана выпрямленная линейная функция *ReLU*, использование которой позволило преодолеть проблему затухания градиентов в процессе обучения.

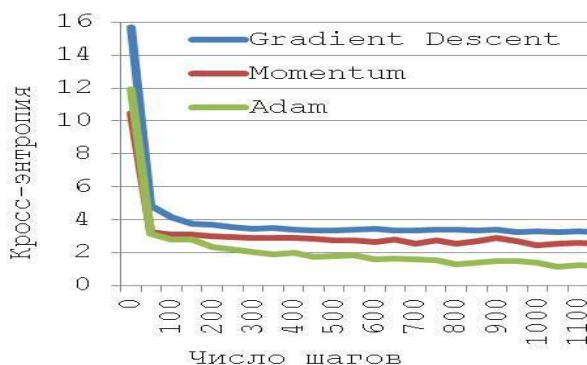


Рис. 5. Сравнение результатов применения алгоритмов обучения.

Для оценки точности классификации использовалась матрица ошибок (рис. 6), представляющая собой инструмент, использующий кросс-табуляцию для показа того, как соотносятся значения совпадающих классов, полученные из различных источников.

	1	2	3	4	5	6	7	8	9
1	54	0	0	0	4	1	0	0	0
2	2	17	6	8	4	1	0	2	9
3	0	3	30	0	12	0	0	0	0
4	0	0	0	40	0	0	0	1	3
5	0	1	6	0	36	1	0	0	0
6	0	2	0	0	0	15	8	2	3
7	0	0	0	0	0	0	24	0	0
8	0	2	0	0	0	3	8	20	3
9	0	2	0	0	0	2	0	2	36

Рис. 6. Матрица ошибок.

Главная диагональ матрицы показывает случаи, где расчётные классы и реальные данные совпадают (правильная классификация). По одной оси матрицы записаны названия классов легенды классификации проверяемого набора данных, по второй – классы легенды данных, используемых для проверки, – разметка.

Для оценки качества классификации использовался подход на основе *ROC*-анализа (*receiver operating characteristic*, рабочая характеристика приёмника). *ROC*-кривая представляет собой график (рис. 7), отображающий соотношение между долей объектов от общего количества носителей признака, верно классифицированных, как несущих признак, называемой чувствительностью алгоритма классификации) и долей объектов от общего количества объектов, не несущих признака, ошибочно классифицированных, как несущих признак при варьировании порога решающего правила.

Количественную интерпретацию *ROC* даёт показатель *AUC* (*area under ROC curve* – площадь под *ROC*-кривой) – площадь, ограниченная *ROC*-кривой и осью доли ложных положительных классификаций. Чем выше показатель *AUC*, тем качественнее классификатор, при этом значение 0,5 демонстрирует непригодность выбранного метода классификации (соответствует случайному гаданию). Значение менее 0,5 означает, что классификатор действует с точностью до наоборот: если положительные назвать отрицательными и наоборот, классификатор будет работать лучше. Для построенного классификатора на основе свёрточной нейронной сети площадь под *ROC*-кривой составляет $S = 0,95$.

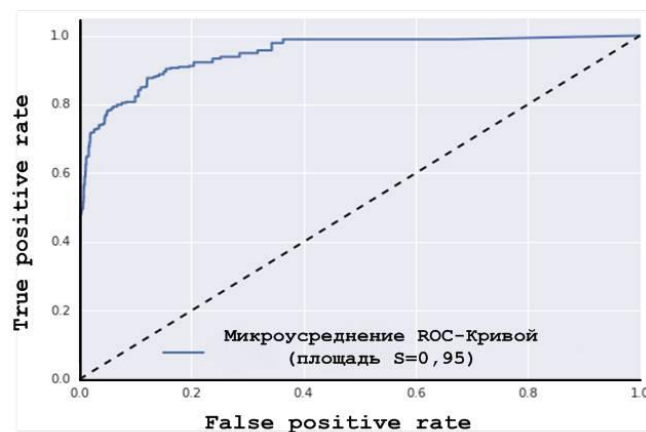


Рис. 7. ROC-кривая для построенного классификатора.

4. Заключение

Описаны подходы к обработке гиперспектральной информации, направленные на распознавание и классификацию объектов транспортной инфраструктуры, а также определение характеристик и состояния дорожного покрытия. Предложенные методы предназначены для информационной поддержки процесса построения модели транспортной инфраструктуры, представленной в виде тематических слоёв электронной карты, в специализированной интеллектуальной транспортной геоинформационной системе. Представленные решения могут быть встроены в контур наземной тематической обработки и цепочку создания базовых продуктов на основе информации дистанционного зондирования для создания специализированных тематических слоёв, используемых при решении практических задач в сегменте «Транспортная инфраструктура».

Литература

- [1] Михеев, С.В. Технология Data Mining в задачах прогнозирования развития транспортной инфраструктуры [Электронный ресурс] / А.А. Федосеев, С.В. Михеев, О.К. Головин // Современные проблемы науки и образования. – 2013. – № 1. – Режим доступа: www.science-education.ru/107-8153.
- [2] Федосеев, А.А. Методы обработки пространственно-распределённой спектральной информации об объектах транспортной инфраструктуры / А.А. Федосеев, А.Н. Тихонов // ИТ & Транспорт. – Самара : Интелтранс, 2015. – С 110–134.
- [3] Федосеев, А.А. Синтез тематического слоя объектов транспортной сети / А.А. Федосеев, О.К. Головин, С.В. Михеев // Известия Самарского научного центра Российской академии наук. – 2016. – Т. 18, № 4 (4). – С. 839–843.
- [4] Herold, M. Spectrometry and hyperspectral remote sensing of urban road infrastructure [Электронный ресурс] / M. Herold, M. Gardner, V. Noronha, D. Roberts. – Режим доступа : http://www.eo.uni-jena.de/~c5hema/pub/rse04_heroldetal.pdf.
- [5] Szegedy, C. Going Deeper with Convolutions [Электронный ресурс] / C. Szegedy [et al.]. – Режим доступа: http://www.cv-foundation.org/openaccess/content_cvpr_2015/app/1A_001.pdf.
- [6] Fukushima, K. Neocognitron: a hierarchical neural network capable of visual pattern recognition / K. Fukushima // Neural Networks. – 1988. – Vol. 1, No. 2. – Pp. 119–130.
- [7] LeCun, Y. Efficient BackProp / Y. LeCun [et al.] [Электронный ресурс]: – <http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf>.
- [8] Сапрыкин, О.Н. Распознавание объектов транспортной инфраструктуры на гиперспектральных снимках методом глубокого машинного обучения / О.Н. Сапрыкин, А.А. Федосеев, Т.И. Михеева, Р.Р. Мингазов // Информационные технологии интеллектуальной поддержки принятия решений (ITIDS'2016) : труды IV Международной конференции. – Уфа : Изд-во УГАТУ. – 2016. – С. 39–44.
- [9] Mikheeva, T. Recognition of urban transport infrastructure objects via hyperspectral images / O. Saprykin, A.A. Fedoseev, T. Mikheeva // VEHITS 2016 : Proceedings of the International conference on vehicle technology and intelligent transport systems. – Rome (Italy) : SCITEPRESS, 2016. – P. 203–208.
- [10] Барсегян, А.А. Технологии анализа данных : Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян М.С. Куприянов, В.В. Степаненко, И.И. Холод. – СПб. : БХВ-Петербург, 2007. – 384 с.
- [11] Han, J. Data Mining: Concepts and Techniques / J. Han, M. Kamber, J. Pei. – San Francisco : Morgan Kaufmann Publishers, 2012. – 703 p.
- [12] Ruder, S. An overview of gradient descent optimization algorithms [Электронный ресурс] / S. Ruder. – Режим доступа: <http://sebastianruder.com/optimizing-gradient-descent/index.html#momentum>.