

УДК 004.932

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ИДЕНТИФИКАЦИИ СТРУКТУРНО-ОДНОРОДНЫХ ОБЛАСТЕЙ НА КОСМИЧЕСКИХ СНИМКАХ ЗЕМЛИ

Еремеев В.А., НИИ «Фотон» РГРТУ, инженер, e-mail: foton@rsreu.ru

Макаренков А.А., НИИ «Фотон» РГРТУ, с.н.с., к.т.н., e-mail: foton@rsreu.ru

APPLICATION OF CONVOLUTION NEURAL NETWORKS FOR IDENTIFICATION OF STRUCTURALLY HOMOGENEOUS AREAS ON EARTH REMOTE SENSING IMAGES

Eremeev V.A., Makarenkov A.A.

The analysis of different approaches to automatic identification of objects with uniform brightness and structure on Earth remote sensing images is presented. Application of the convolutional neural network U-Net with textural processing of images for homogeneous areas identification is proposed. Efficiency evaluation of the proposed method using real images from Russian remote sensing system «Resurs-P» is performed.

Key words: remote sensing, identification of homogeneous areas, texture features, convolutional neural network U-Net.

Ключевые слова: дистанционное зондирование Земли, идентификация однородных областей, текстурный признак, сверточная нейронная сеть U-Net.

Введение

Идентификация объектов, зафиксированных на спутниковых изображениях – это трудно формализуемая задача, до сих пор не имеющая определённого решения. Сегодня приборы космического наблюдения Земли позволяют выполнять съёмку в видимом, тепловом, радиолокационном диапазоне спектра; с геометрическим разрешением до 0,3 м; с полосой обзора от нескольких километров до глобального наблюдения Земли на фоне окружающего космоса. Главной же задачей систем ДЗЗ является получение новых оперативных сведений по конкретным объектам интереса (поверхность Земли и её атмосфера, регионы, малоразмерные объекты и т.п.). Решение этой задачи во многом зависит от уровня исследований проблемы автоматической идентификации на спутниковых снимках объектов с заданными свойствами. Наиболее простыми для определения казалось-бы являются текстурно-однородные по характеристикам объекты: водная поверхность, песчаные пустыни, облака и др. Но и при их идентификации возникают большие трудности. Здесь необходимо выделить два подхода.

Первый подход – это строгое (адекватное) математическое описание характеристик структурно-однородных объектов и создание алгоритмов их идентификации. В данном случае задача нахождения объектов на снимках может качественно и быстро решаться в автоматическом режиме и необходимость в использовании второго подхода отпадает.

Второй подход – это наличие знаний лишь для создания приближенных (грубых) математических моделей идентификации однородных объектов. В этом случае «напрашивается» решение о привлечении искусственных нейронных сетей для получения более каче-

Выполнен анализ подходов к автоматической идентификации однородных по яркости и структуре объектов по данным ДЗЗ. Обосновано использование для решения этой задачи свёрточной нейронной сети U-Net с привлечением текстурных признаков изображений. Представлены результаты экспериментальных исследований предложенного подхода на примерах использования натурной информации от российской системы дистанционного зондирования Земли (ДЗЗ) «Ресурс-П».

ственных результатов идентификации. В рамках второго подхода рассматриваются некоторые вопросы, поднятые в настоящей статье.

Автоматизация процесса идентификации структурно-однородных областей на космических снимках Земли востребована во многих важных практических приложениях. Укажем лишь две диаметрально противоположные области использования. Значительную часть сюжета космических изображений составляют водные объекты, песчаные пустыни, снежные и ледовые образования, облака и др. Такие участки являются не нужными для многих потребителей, например для картографов. В то же время для их передачи на наземные средства приёма и обработки неоправданно загружаются космические каналы связи, средства обработки, хранения и доставки видеoinформации потребителям. После идентификации однородных участков они могут быть исключены из ресурсно затратных процедур обработки данных ДЗЗ, подвергнуты сжатию с потерями и т.д. Другая область использования структурно-однородных участков связана с радиометрической калибровкой бортовых средств наблюдения Земли при которой однородные участки используются для определения параметров радиометрической коррекции структурных искажений.

Предлагаемый подход к идентификации структурно-однородных областей

На сегодняшний день свёрточные нейронные сети показывают наилучшие результаты в задачах обнаружения и распознавания объектов. Для реализации постав-

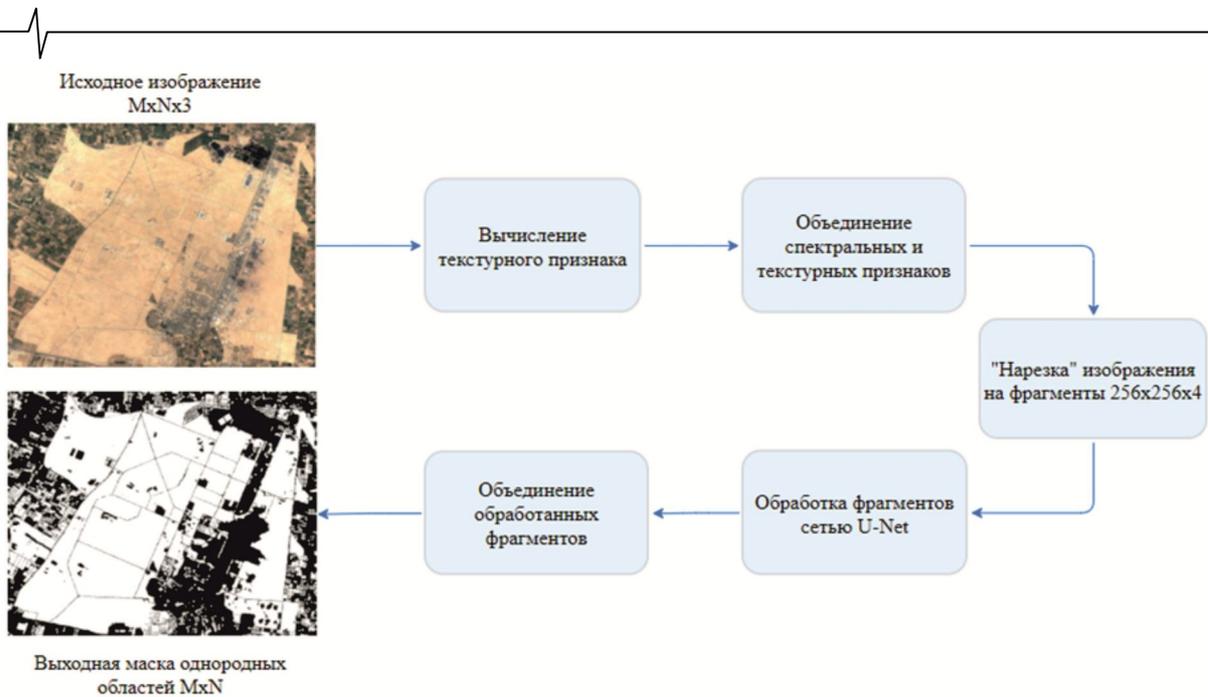


Рис. 1. Схема идентификации структурно-однородных областей на основе сети U-Net и текстурного признака

ленной задачи выбрана популярная архитектура U-Net [1]. Архитектура U-Net была разработана в 2015 году и успешно применяется, в том числе, для автоматической сегментации данных ДЗЗ. Существуют и другие архитектуры свёрточных нейронных сетей, широко используемые для сегментации спутниковых изображений, например FCN (Fully Convolutional Network), ResNet (Residual Network), VGGNet (Very Deep Convolutional Network), однако выбрана архитектура U-Net из-за низких требований к объёму обучающей выборки и достаточно высокой точности идентификации. Сеть U-Net эффективно использует идею сквозных соединений между «кодировщиком» и «декодировщиком», благодаря чему имеет сравнительно небольшое количество обучаемых параметров и, соответственно, не требует большого количества изображений для обучения.

Использование только спектральных признаков изображения (яркостей пикселей в спектральных каналах) не позволяет нейронной сети с высокой вероятностью отличать однородную область от неоднородной, в связи с чем, возникает необходимость в привлечении дополнительных признаков. Текстуальный анализ позволяет с высокой точностью получать информацию о структуре объектов на изображении. В ряде случаев, текстурные признаки могут быть более информативны, чем физические характеристики, полученные по данным ДЗЗ. Исходя из этого, поставлена задача дополнить спектральную информацию текстурным признаком при обучении свёрточной нейронной сети U-Net.

На рис. 1 представлена схема идентификации структурно-однородных областей. На первом этапе исходное трехканальное изображение преобразуется в одноканальное путем усреднения спектральных каналов и вычисляется текстурный признак «Sum of square» Харалика. В результате формируется одноканальное изображение, каждый элемент которого является значение текстурного признака. После этого изображение тек-

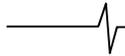
стурного признака объединяется с трехканальным исходным снимком, в результате формируется четырехканальное изображение. Далее это изображение разбивается на фрагменты размером $256 \times 256 \times 4$, после чего каждый фрагмент последовательно обрабатывается сетью U-Net. Затем обработанные фрагменты воссоединяются, и формируется выходная маска однородных областей.

Для обучения свёрточной нейронной сети сформирован набор данных с использованием мультиспектральных снимков (3 спектральных канала) от космического аппарата «Ресурс-П». Пиксели исходных изображений вручную помечены и разделены на два класса: однородный и неоднородный. Сформированный набор данных «разбит» на фрагменты размером 256×256 пикселей и разделён на наборы для обучения и тестирования. Для увеличения объёма обучающей выборки применена аугментация данных, включающая повороты исходных фрагментов на углы кратные 90° и поворот на 45° с отражением недостающих частей изображения.

Выполнен анализ известных текстурных признаков. Выбор пал на использование одного из наиболее популярных алгоритмов [2] (Харалик, 1979). Алгоритм основан на вычислении матрицы смежности уровней серого, которая учитывает как уровни яркости пикселей, так и относительное расположение пикселей с определённой яркостью на изображении. Размерность матрицы определяется числом градаций яркости изображения. Элемент матрицы смежности уровней серого G с координатами (i, j) вычисляется как [2]

$$G_{\Delta y, \Delta x}(i, j) = \sum_{y=1}^{N-\Delta y} \sum_{x=1}^{M-\Delta x} \begin{cases} 1, & \text{если } I(y, x) = i \vee I(y + \Delta y, x + \Delta x) = j, \\ 0 & \text{иначе} \end{cases}$$

где $I(y, x)$ – яркость пикселя с координатами (y, x) , M и N – размеры изображения, $[\Delta y, \Delta x]$ – параметры



сдвига, которые равны $[\Delta x, \Delta y] \in \{[0,1]; [-1,1]; [-1,0]; [-1,-1]\}$ для углов $\{0, 45, 90, 135\}$ соответственно. При помощи сформированной матрицы смежности уровней серого могут быть вычислены все 14 известных текстурных признаков Харалика, однако в ходе экспериментов выявлено, что для идентификации однородных областей достаточно наиболее информативного для этой задачи признака «Sum of square». Тектурный признак «Sum of square» вычисляется как

$$S = \sum_{i=1}^N \sum_{j=1}^N p_{i,j} (i - \mu)^2,$$

где S – значение текстурного признака, $p_{i,j}$ – элемент матрицы G с координатами (i, j) , N – размер матрицы G . Параметр μ в большинстве публикаций задаётся как

$$\mu = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N p_{i,j}.$$

На рис. 2 представлен один из фрагментов в наборе данных. На рисунке: слева – исходное изображение; по центру – его эталонная маска, сформированная оператором; справа – текстурный признак «Sum of square». На эталонной маске серым цветом выделены однородные области, а черным неоднородные.

Моделирование и результаты

Для оценки эффективности предложенного подхода использовались метрики качества сегментации, такие как точность, полнота, F-мера [3]. В табл. 1 приведено описание используемых метрик.

Для оценки эффективности использования текстурных признаков дополнительно обучена ещё одна сеть U-Net. Набор данных и процесс обучения оставались неизменны за исключением того, что сеть обучалась только с привлечением спектрозональных изображений без использования текстурных признаков. На рис. 3-4

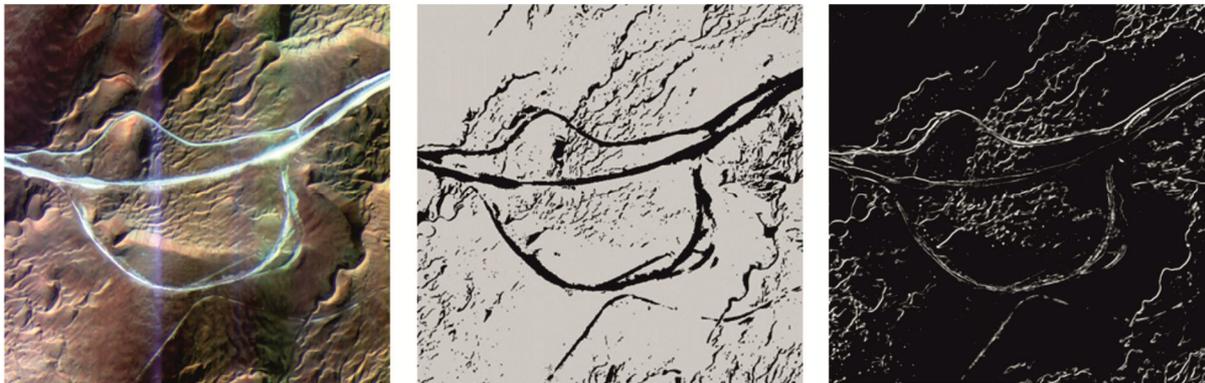


Рис. 2. Фрагмент набора данных (слева – исходное изображение, по центру – эталонная маска, справа – текстурный признак «Sum of square»)

Таблица 1. Описание метрик оценки качества сегментации [3] (TP – истинно - положительные прогнозы, FP – ложно-положительные прогнозы, FN – ложно-отрицательные прогнозы)

Метрика	Описание	Формула
Точность (P)	Доля объектов, действительно принадлежащих данному классу, относительно всех объектов, которых нейронная сеть отнесла к этому классу	$P = \frac{TP}{TP + FP}$
Полнота (R)	Доля истинно положительных прогнозов. Полнота показывает, какая доля объектов, реально относящихся к положительному классу, спрогнозирована верно	$R = \frac{TP}{TP + FN}$
F-мера (F)	Средневзвешенное значение полноты и точности	$F = \frac{2PR}{P + R}$



Рис. 3. Пример результатов идентификации однородных областей (слева – эталонная маска, сформированная оператором; в центре – U-Net; справа – U-Net + текстурный признак)

Таблица 2. Вычисленные метрики качества сегментации для фрагмента, приведённого на рис. 3. Числитель – U-Net (рис. 3, в центре), знаменатель – U-Net + текстурный признак (рис. 3, справа)

Класс/Метрика	Точность	Полнота	F-мера	Кол-во образцов в эталонной маске (пикс.)
Однородный	0,947 / 0,975 (2,9 %)	0,942 / 0,977 (3,6 %)	0,945 / 0,976 (3,2 %)	79 %
Неоднородный	0,768 / 0,906 (15,2 %)	0,786 / 0,897 (12,4 %)	0,777 / 0,902 (13,8 %)	21 %



Рисунок 4. Пример результатов идентификации однородных областей (слева – эталонная маска, сформированная оператором; по центру – U-Net; справа – U-Net + текстурный признак)

Таблица 3. Вычисленные метрики качества сегментации для фрагмента, приведённого на рис. 4. Числитель – U-Net (рис. 4, в центре), знаменатель – U-Net + текстурный признак (рис. 4, справа)

Класс/Метрика	Точность	Полнота	F-мера	Кол-во образцов в эталонной маске (%)
Однородный	0,890 / 0,914 (2,6 %)	0,767 / 0,831 (7,7 %)	0,824 / 0,870 (5,3 %)	66 %
Неоднородный	0,622 / 0,703 (8,1 %)	0,802 / 0,837 (3,5 %)	0,701 / 0,764 (6,3 %)	34 %

представлены примеры результатов идентификации однородных областей. Ниже, в табл. 2-3 представлены вычисленные метрики точности, полноты и F-меры для приведённых фрагментов.

На основе анализа рис. 3, 4 и табл. 2, 3 можно сделать следующие выводы:

– Предлагаемый подход по идентификации структурно-однородных областей изображения более эффективен по сравнению с использованием сверточной нейронной сети с архитектурой U-Net, обрабатывающей только яркости изображения.

– При обучении нейронной сети большие трудности вызывает формирование адекватных эталонов, поскольку эксперту достаточно сложно разделить изображение на два класса: однородные участки и неоднородные. В связи с этим оценка абсолютной точности выделения однородных и неоднородных участков выполнена приблизительно (с точностью ручного выделения экспертом). В целом для повышения эффективности решения поставленной задачи необходимо улучшать процесс формирования эталонов в обучающей выборке, путем, например, их синтеза.

Заключение

Результаты выполненного исследования заключаются в следующем:

– Экспериментально подтверждена целесообразность

привлечения текстурного признака для сокращения затрат при обучении сети и повышения ее эффективности для идентификации структурно-однородных областей.

– Обосновано применение свёрточной нейронной сети U-Net.

– Экспериментально получены численные оценки показателей качества идентификации структурно-однородных областей с помощью сети U-Net с привлечением текстурного признака.

– Выполнены экспериментальные исследования эффективности предложенного подхода с привлечением натурной информации от космической системы «Ресурс-П».

Литература

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, vol. 9351, 2015, pp. 234-241.
2. Haralick, R. M., Shanmugam, K. & Dinstein, I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics 3, 1973, pp. 610-621.
3. Marina Sokolova, Guy Lapalme. A systematic analysis of performance measures for classification tasks. Information Processing and Management. 2009. 45(4), pp. 427-437.