

КЛАССИФИКАЦИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ, ПОЛУЧЕННЫХ С ПОМОЩЬЮ КОСМИЧЕСКИХ АППАРАТОВ И БПЛА

**В.В. Ганченко¹, А.А. Дудкин¹, А.В. Инютин¹, Е.Е. Марушко¹,
В.Г. Писаренко²**

¹*Объединенный институт проблем информатики НАН
Беларуси, Минск, Беларусь
e-mail: avin@lsi.bas-net.by*

²*Институт кибернетики имени В. М. Глушкова НАН Украины, Киев, Украина*

Для идентификации и классификации объектов на изображениях, полученных с помощью съемочных средств БПЛА и орбитального базирования, предложена нейросетевая модель классификации, основанная на использовании автоэнкодера и построенная по архитектуре ансамбля многослойных перцептронов. Использование предложенной модели для классификации на четыре класса: «Пожар», «Задымление», «Расительность» и «Строения», позволяет достичь точности классификации выше 99 %.

Ключевые слова: автоэнкодер; ансамбль многослойных перцептронов; классификация.

CLASSIFICATION OF OBJECTS IN IMAGES OBTAINED USING SPACECRAFT AND UAV

V.Ganchenko^a, A.Doudkin^a, A.Inyutin^a, Ya.Marushko^a, V.Pisarenko^b

^a*United Institute of Informatics Problems of NAS of Belarus, Minsk, Belarus*

^b*V.M. Glushkov Institute of Cybernetics of the National
Academy of Sciences of Ukraine, Kiev, Ukraine*

This report presents a neural network model for identifying and classifying objects in images obtained using UAV and orbital-based imaging equipment. The model is based on the use of an autoencoder and is built on the architecture of an ensemble of multilayer perceptrons. The model has an accuracy above 99% when classified into four classes: "Fire", "Smoke", "Vegetation" and "Building".

Keywords: autoencoder; an ensemble of multilayer perceptrons; classification.

Введение

Мониторинг земной поверхности с целью обнаружения стихийных бедствий и оценки их последствий, контроля природопользования и экологической безопасности является актуальной задачей. Ряд стихийных бедствий, таких как паводковые наводнения, лесные пожары и засухи, аварии на техногенных объектах, затрагивают территории приграничных стран.

Для своевременного обнаружения на конкретном участке территории фактов возникновения и развития неблагоприятных явлений, ситуаций и стихийных бедствий необходимо с целью скорейшей ликвидации этих последствий использовать информацию от разных источников: космические снимки, данные аэросъемки, включая полученные с беспилотных летательных аппаратов (БПЛА), а также специальные данные от наземных средств мониторинга, включая мобильные.

Задачей, решаемой представленными нейросетевыми моделями, является выделение параметров объектов на изображениях для их идентификации и классификации. Идентификация и классификация объектов осуществляются для участков исходных цветных изображений размером 128 x 128 пикселей. Идентификация и классификация участка изображения разделяются на два этапа:

- 1) выделение информативных признаков, позволяющих идентифицировать содержимое изображения;
- 2) выбор класса, которому согласно выделенным признакам максимально соответствует анализируемый участок изображения.

1. Нейросетевая модель АВТОЭНКОДЕРА

Основой рассматриваемого подхода к выделению признаков является использование нейросетевой модели автоэнкодера, который формирует дескриптор изображения в виде массива из 512 элементов (числа с плавающей запятой). Получаемый массив нормируется по L2-норме.

Автоэнкодеры состоят из двух частей: энкодера g и декодера f . Энкодер переводит входной сигнал в его представление (дескриптор) $h = g(x)$, а декодер восстанавливает сигнал по его коду $x = f(h)$. Автоэнкодер, изменяя g и f , стремится выучить тождественную функцию $x = f(g(x))$, минимизируя функционал ошибки $L(x, f(g(x)))$. При этом семейства функций энкодера g и декодера f ограничены так, чтобы автоэнкодер был вынужден отбирать наиболее важные свойства сигнала.

Кодировщик представляет собой набор из трех сверточных блоков с понижением размерности и одного полносвязного слоя. Каждый сверточный блок состоит из двух сверточных слоев (например, Conv2D_1.1 и Conv2D_1.2 для первого блока) и слоя понижения пространственной размерности (например, MaxPooling2D_1 для первого блока). Сверточные слои выполняют функцию выделения признаков заданного размера и формирования карт признаков, по одной на каждый фильтр. Полносвязный слой выполняет функцию преобразования полученного набора карт признаков в одномерный вектор чисел с плавающей запятой.

Декодировщик также состоит из одного полносвязного слоя и трех сверточных блоков с повышением размерности. Каждый сверточный блок состоит из слоя повышения пространственной размерности (например, UpSampling2D_2 для первого блока) и двух сверточных слоев (например, Conv2D_4.1 и Conv2D_4.2 для первого блока).

Многозональное изображение может быть представлено в виде трехмерного куба, две оси которого имеют пространственную, а третья – спектральную размерность (непосредственно связанную с излучательной и отражательной характеристиками объекта).

Для получения дескриптора участка исходного цветного изображения размером 128×128 пикселей достаточно подать это изображение на вход автоэнкодера и получить выходные значения указанного скрытого слоя.

Т.к. вектор размерностью 512 отобразить не представляется возможным, то для вывода данных был использован алгоритм понижения размерности – метод главных компонент. С помощью указанного метода размерность дескрипторов была уменьшена с 512 до двух, что уже может быть отображено на плоскости.

2. Вариативные информативные признаки

Текстурные характеристики выражают разность значений интенсивности соседних пикселей изображения, при этом значения оценок этой разности отличаются для различных типов подстилающей поверхности. Примеры вычисленных текстурных характеристик приведены в таблице 1. Для оценки текстуры используются текстурные характеристики Харалика.

Таблица 1 – Примеры вычисленных текстурных характеристик

<i>Объект</i>	<i>Характеристика</i>	
	<i>ASM</i>	<i>Contrast</i>
Дым	$5.262 \cdot 10^{-3}$	4.249
Лес	$2.215 \cdot 10^{-4}$	$2.681 \cdot 10^2$
Строения	$3.915 \cdot 10^{-4}$	$1.560 \cdot 10^2$
Очаг возгорания	$2.594 \cdot 10^{-3}$	$1.656 \cdot 10^2$

Построение дескрипторов объектов основано на использовании комбинированных информативных признаков многоспектральных изображений. Важной характеристикой признаков является вариативность, которая оценивается их текстурными характеристиками. Значения текстурных характеристик собираются в матрицы, которые преобразовываются к целочисленным значениям в диапазоне от 0 до 255. Полученная матрица пред-

ставляет собой полутоновое изображение, которое позволяет сравнить значения оценок текстурных характеристик для различных участков изображений.

При выделении информативных признаков дополнительно добавляется цветовая информация, инвариантная к масштабу и поворотам изображения. Для этого для каждого обрабатываемого участка изображения строится его поканальная гистограмма. При этом для сокращения объема данных применяется квантование значений яркости не на 256 уровней, а на 16. Далее полученные гистограммы объединяются в один массив размером 48 элементов: [16 значений синего канала, 16 значений зеленого канала, 16 значений красного канала]. Полученный массив также нормируется по L2-норме.

Таким образом, результатом идентификации является массив из 560 элементов, который и используется для классификации.

3. Обучение моделей

Для классификации разработана нейросетевая модель, состоящая из двух многослойных персептронов, объединенных в виде нейросетевого ансамбля. Каждый персептрон в отдельности анализирует входной массив данных, после чего выходные слои формируют взвешенное общее решение по классу входных данных. Сама модель реализована на языке Python с использованием библиотеки Keras [1].

Для обучения модели классификации использовалась функция потерь (loss) – categorical cross entropy; обучающая и валидационная выборки – 4000 и 800 изображений (не включая аугментацию); обучающий алгоритм Adam [2] с параметрами $lr = 0,0001$, $beta_1 = 0,9$, $beta_2 = 0,999$, $epsilon = 1 \cdot 10^{-8}$, $decay = 0,0$. Результирующая точность на валидационной выборке – 99,19 %.

4. Тестирование

Для тестирования использовалась видеосъемка с БПЛА пожара на нефтебазе. Пример кадра приведен на рис. 1. Также использовались данные Sentinel Hub: съемки спутника Landsat 8 в диапазоне, полученном комбинацией красного, синего, зеленого и панхроматического каналов, и в тепловом инфракрасном диапазоне [3]. Разрешение цветных каналов составляет 30 м на пиксел, разрешение панхроматического канала – 15 м на пиксел, разрешение теплового инфракрасного канала – 100 м на пиксел.

Классификация осуществляется на четыре класса: «Пожар», «Задымление», «Растительность» и «Строения». Пример изображений для каждо-

го класса приведен на рис. 2. На рис. 3 показан пример отображения 16 дескрипторов (для четырех различных групп изображений, включающих различные типы поверхностей).

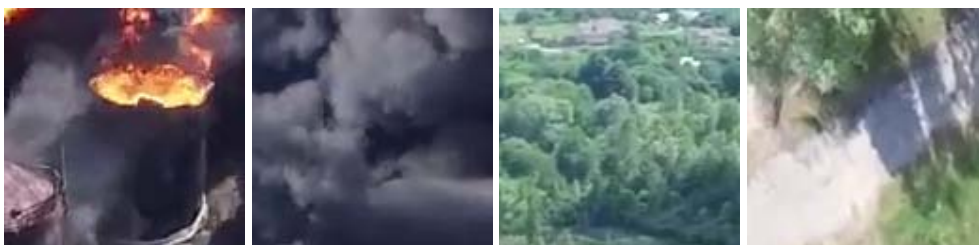
Для расширения набора данных, используемых для обучения моделей, были использованы следующие аугментации (в случайных комбинациях):

- поворот изображения на углы 0, 90, 180 и 270°;
- отражение по вертикали и горизонтали;
- масштабирование на случайную величину из диапазона [0,9; 2].

Тестирование осуществлялось на валидационной выборке. При этом оценивалась точность как для каждого класса в отдельности, так и для всех классов в целом (таблица 2). Для оценки качества классификации были вычислены соответствующие значения точности (precision), полноты (recall) и F1-меры [4]. Общая точность составила 99,19 %. Наибольшее количество ошибок возникало на участках, соответствующих границе задымления и пожара, а также в случае горения строений.



Рисунок 1 – Пример кадра видеосъемки БПЛА



а)

б)

в)

г)

Рисунок 2 – Примеры участков изображений для идентификации и классификации на классы:

а) «Пожар»; б) «Задымление»; в) «Растительность»; г) «Строения»

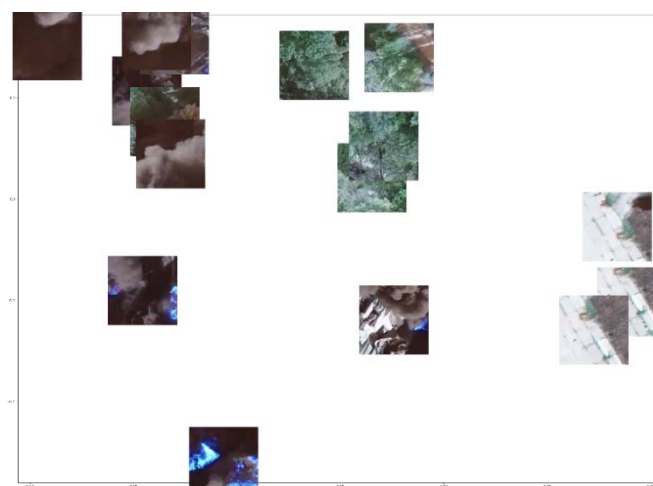


Рисунок 3 – Расположение двумерных проекций дескрипторов на плоскости в виде миниатюр

Таблица 2 – Результаты тестирования сегментаторов и значения полноты и F1-меры

<i>Классы</i>	<i>Точность</i>	<i>Полнота</i>	<i>F1-мера</i>
Задымление	0,869	1,000	0,930
Растительность	0,997	0,981	0,989
Строения	1,000	0,848	0,918
Пожар	0,989	0,994	0,991

Заключение

Для идентификации и классификации объектов на изображениях, полученных с помощью съемочных средств БПЛА и орбитального базирования, предложена нейросетевая модель классификации, основанная на использовании оригинального автоэнкодера и построенная по архитектуре ансамбля многослойных персептронов.

Дескрипторы представляют собой массив из 512 чисел с плавающей запятой. В качестве основы для построения дескрипторов использован нейросетевой автоэнкодер, архитектура которого зависит от модальности и цветности исследуемых данных. Отличительными чертами разработанных моделей является применение как многомодальных и разноцветных данных, так и дополнительных информативных признаков вариативности данных. Для оценки вариативности данных использованы текстурные признаки Харалика, построенные для локальных участков изображения. При выделении информативных признаков дополнительно добавляется

цветовая информация, инвариантная к масштабу и поворотам изображения и основанная на построении поканальных гистограмм.

Использование предложенной модели для классификации на четыре класса: «Пожар», «Задымление», «Растительность» и «Строения» – позволяет достичь точности классификации выше 99 %.

Работа выполнена при поддержке БРФФИ (договор № Ф21УКРГ-007 от 30.04.2021).

Библиографические ссылки

1. Keras optimizers: Adam [Electronic resource]. URL: <https://keras.io/api/optimizers/adam/> Date of access: 28.05.2022.
2. Kingma, D. P. Adam: A Method for Stochastic Optimization / D. P. Kingma, J. Ba // 3rd Intern. Conf. for Learning Representations, 7–9 May, 2015, San Diego. San Diego, 2015. Vol. 1412. P. 6980.
3. Sentinel Hub EO Browser [Electronic resource]. URL: https://apps.sentinel-hub.com/eo-browser/?zoom=13&lat=50.20921&lng=30.23931&themeId=DEFAULT-THEME&visualizationUrl=https%3A%2F%2Fservices.sentinel-hub.com%2Fogc%2Fwms%2F%3E35192fe-33a1-41f3-b798-b755e771c5a5&datasetId=AWS_LOTL1&fromTime=2015-06-09T00%3A00%3A00.000Z&toTime=2015-06-09T23%3A59%3A59.999Z&layerId=1_TRUE_COLOR. Date of access: 28.05.2022.
4. Sokolova, M. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation / M. Sokolova, N. Japkowicz, S. Szpakowicz // Advances in Artificial Intelligence: 19th Australian Joint Conf. on Artificial Intelligence, Hobart, 4–8 Dec., 2006. Hobart, 2006. P. 1015–1021.