

Визуализация шаблонов максимальной активации фильтров сверточной нейронной сети в задаче классификации радиолокационных изображений с различным разрешением

И.Ф. Купряшкин, email: ifk78@mail.ru
А. С. Мазин, email: mazinant@rambler.ru

ВУНЦ ВВС ВВА им. проф. Н.Е. Жуковского и Ю.А. Гагарина

***Аннотация.** Описан порядок подготовки обучающих, проверочных и тестовых наборов изображений, приведена структура глубокой сверточной нейронной сети и условия ее обучения. Приведены изображения шаблонов, обеспечивающие максимальную активацию фильтров сверточных слоев нейронной сети, обученной на наборе радиолокационных изображений с различным разрешением.*

***Ключевые слова:** радиолокационное изображение, глубокая сверточная нейронная сеть, классификация изображений объектов.*

Введение

В настоящее время глубокие сверточные нейронные сети (ГСНС) стали инструментом, широко применяемым в задаче классификации объектов на радиолокационных изображениях (РЛИ) [1-4]. Несмотря на значительный объем уже выполненных исследований в этой сравнительно новой области, в известных публикациях практически не отражены вопросы влияния изменения разрешающей способности РЛИ на работу ГСНС.

Цель статьи – иллюстрация базовых возможностей ГСНС в задаче классификации объектов на РЛИ с разным разрешением.

1. Подготовка обучающих, проверочных и тестовых наборов изображений

В качестве исходных данных использован набор MSTAR (Moving and Stationary Target Acquisition and Recognition) [4,5], с 90-х годов являющийся базовым для создания и изучения алгоритмов распознавания отметок объектов на РЛИ. Используемая в настоящей работе открытая часть набора MSTAR включает РЛИ десяти образцов военной техники, полученные РЛС X-диапазона. Заявленное исходное пространственное разрешение составляет около 0,3 м. Примеры оптических и радиолокационных изображений объектов набора MSTAR приведены на рис. 1.

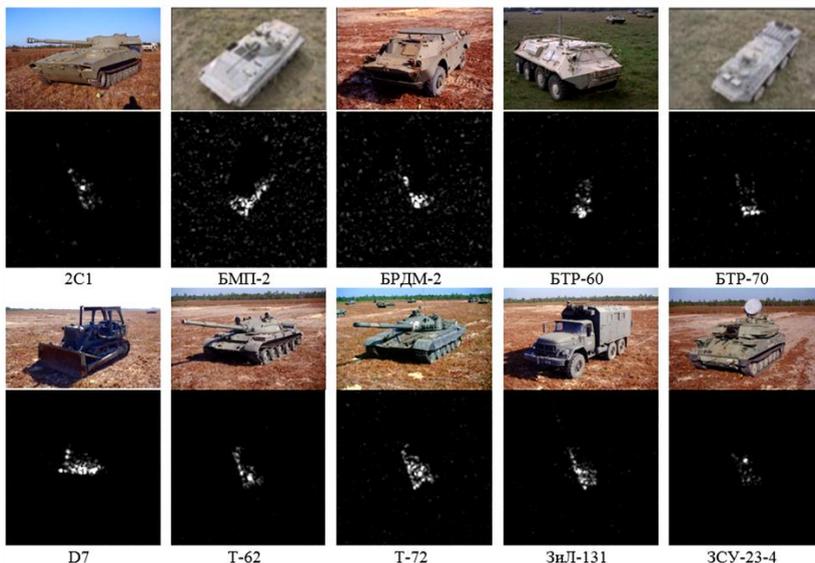


Рис. 1. Оптические и радиолокационные изображения объектов набора MSTAR

Весь имеющийся набор изображений разбит на обучающий (1923 изображения, угол визирования при съемке 17 градусов), проверочный (891 изображение, угол визирования 17 градусов) и тестовый (2503 изображения, угол визирования 15 градусов) наборы.

Каждое исходное изображение представляет собой массив комплексных отсчетов \mathbf{I} , размерность которого для объектов разных классов является различной (минимальная составляет 128×128 отсчетов). На начальном этапе подготовки данных осуществляется подавление фона местности на каждом изображении для исключения его влияния на результат работы сети. Так как при отсутствии фона исходная размерность изображений набора MSTAR от 128×128 отсчетов и более является избыточной, на следующем этапе осуществляется ее понижение до 44×44 отсчетов (именно при такой минимальной размерности отметки объектов всех классов полностью помещаются в пределах кадра изображения).

Подготовка наборов РЛИ с ухудшенным пространственным разрешением осуществлялась путем пространственной фильтрации изображений после пороговой обработки и снижения их размерности как

$$I_{ij} = \sum_{m=1}^N \sum_{n=1}^N \left| I_{mn} \right| G(m-i, n-j),$$

где $G(m, n) = \exp\left(-\frac{(m^2 + n^2)}{2k_r^2}\right)$; k_r – коэффициент, характеризующий,

во сколько раз в результате фильтрации ухудшается пространственное разрешение изображения; $i = 1, N$; $j = 1, N$; $N = 44$.

Примеры изображений объектов с исходным разрешением 0,3 м и с ухудшенным в два (0,6 м) раза показаны на рис. 2.

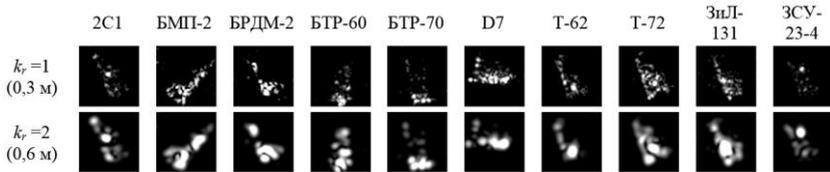


Рис. 2. Примеры изображений наборов с исходным и ухудшенным в два раза пространственным разрешением

2. Архитектура глубокой сверточной нейронной сети

Разработка и обучение сети осуществлялось с использованием фреймворка TensorFlow 2.0 и библиотеки глубокого обучения Keras [6].

Архитектура сети включает последовательности четырех пар сверточных слоев (Convolution Layers) с одинаковыми ядрами 3×3 и единичным шагом свертки. После каждой пары включены слои подвыборки (MaxPooling) 2×2 . Первая пара сверточных слоев включает по 32 фильтра, вторая – по 64, третья – по 128 и четвертая – по 256. Полносвязный классификатор включает входной слой на 4096 входов и 256 выходов, слой 50-процентного прореживания (Dropout) и выходной слой с 10 выходами по числу классов. В качестве функции активации сверточных слоев и входного слоя классификатора используется кусочно-линейная (ReLU). Размерность входного изображения составляет 128×128 . Для описанной архитектуры количество настраиваемых параметров сети составляет $2'223'082$.

При обучении использовался адаптивный алгоритм Adam, в качестве приема предотвращения переобучения применялось расширение данных (augmentation) путем смещения каждого изображения по вертикали и горизонтали на случайное число пикселей от одного до пяти и двукратной подачей обучающего набора на вход сети в течение каждой эпохи. По результатам классификации

изображений тестового набора сеть показала точность 97,91%, т.е. неверно классифицировала 52 изображения из 2503. Для набора данных с разрешением, ухудшенным в два раза ($k_r = 2$), точность составила 89,82%.

3. Шаблоны, обеспечивающие максимальную активацию фильтров сверточных слоев

На рис. 3 показаны входные изображения (шаблоны), обеспечивающие максимальную активацию фильтров сверточных слоев (по восемь для каждого слоя) сети при исходном разрешении. Для их получения использовался метод градиентного восхождения в пространстве входов, описанный в [6].

Видно, что в более высоких слоях фильтры начинают описывать все более сложные текстурные признаки, что вообще является типичным для сверточных сетей [6]. При этом отчетливо видна характерная рамка по периметру изображений, особенно для фильтров высоких слоев. Это объясняется использованием выбранного приема расширения данных (случайного смещения изображений на 1-5 пикселей по вертикали и горизонтали), приведшего к тому, что края изображений практически не определяют активацию фильтров и не являются информативными для сети.

Фильтры первых четырех сверточных слоев кодируют простые направленные контуры. Фильтры 5 и 6 кодируют сочетания направленных контуров и зернистой текстуры, обусловленной спекл-эффектом, типичным для РЛИ. Фильтры 7 и 8 кодируют наиболее сложные текстурные признаки, практически неинформативные для восприятия человеком, но, по всей видимости, позволяющие представить изображения объектов в виде их суперпозиции.

На рис. 4 показаны входные изображения, обеспечивающие максимальную активацию фильтров сверточных слоев (по восемь для каждого слоя) сети, обученной на наборе изображений с двукратным ухудшением разрешения (0,6 м).

Заключение

Таким образом, применение ГСНС в задаче классификации РЛИ объектов военной техники продемонстрировало высокую, порядка 98%, точность. Такое значение является практически недостижимым при использовании традиционных (не нейросетевых) алгоритмов. Также, как и в случае ГСНС, обученных на наборах оптических изображений, более высокие слои выделяют более сложные текстурные признаки. Из сравнения шаблонов на рис. 3 и 4 видно, что хотя при ухудшении разрешения тенденция к усложнению текстурных

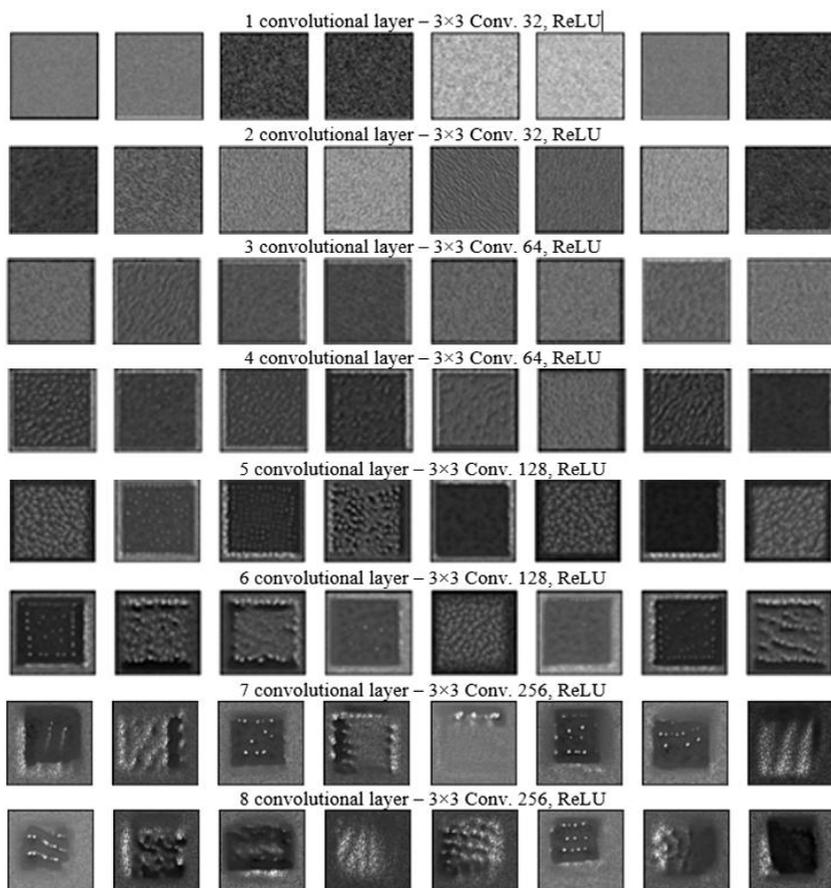


Рис. 3. Шаблоны, обеспечивающие максимальную активацию фильтров сверточных слоев при исходной разрешающей способности (0,3 м)

признаков в более высоких слоях сохранилась, вид самих текстур существенно изменился. Причем визуально это отличие можно охарактеризовать как своего рода упрощение текстур, выразившееся в их большей однородности. Уменьшение разнообразия текстурных признаков при снижении разрешения свидетельствует о меньшей выраженности классификационных признаков, и ограничении степени детализации входных изображений, которые можно представить в виде суперпозиции шаблонов фильтров.

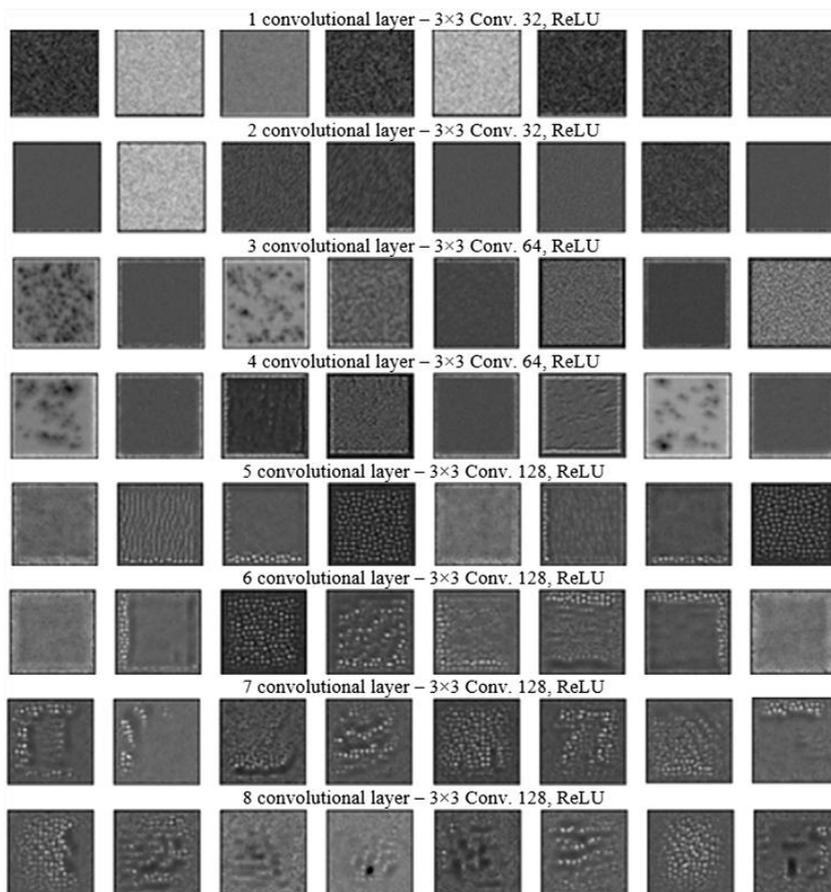


Рис. 4. Шаблоны, обеспечивающие максимальную активацию фильтров сверточных слоев при двукратном ухудшении разрешающей способности (0,6 м)

Список литературы

1. Zhu X., Montazeri S., Ali M., Hua Yu., Wang Yu., Mou L., Shi Yi., Xu F., Bamler R. Deep Learning Meets SAR. arXiv:2006.10027v2 [eess.IV] 5 Jan 2021.
2. Chen S., Wang H., Xu F., Jin Y.-Q. Target Classification Using the Deep Convolutional Networks for SAR Images. IEEE Transaction Geoscience and Remote Sensing, 2016, vol. 54, no. 8, pp. 4806-4817. doi: 10.1109/TGRS.2016.2551720.

3. Profeta A., Rodriguez A., Clouse H.S. Convolutional Neural Networks for Synthetic Aperture Radar Classification. Proc. SPIE 9843, Algorithms for Synthetic Aperture Radar Imagery XXIII, 98430M (14 May 2016). <https://doi.org/10.1117/12.2225934>.

4. Coman C., Thaens R. A Deep Learning SAR Target Classification Experiment on MSTAR Dataset. 2018 19th International Radar Symposium (IRS), 2018, pp. 1-6. doi: 10.23919/IRS.2018.8448048.

5. MSTAR [Электронный ресурс] : база данных. – Режим доступа : <https://www.sdms.afrl.af.mil/index.php?collection=mstar>.

6. Шолле, Ф. Глубокое обучение на Python / Ф. Шолле – СПб: Питер, 2018. – 400 с.