

## ВИЗУАЛИЗАЦИЯ ШАБЛОНОВ МАКСИМАЛЬНОЙ АКТИВАЦИИ ФИЛЬТРОВ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ В ЗАДАЧЕ КЛАССИФИКАЦИИ ЗАШУМЛЕННЫХ РАДИОЛОКАЦИОННЫХ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ

**Купряшкин И. Ф.**, д.т.н., доцент, начальник кафедры «Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), e-mail: ifk78@mail.ru;

**Мазин А. С.**, адъюнкт кафедры «Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), e-mail: mazinant@rambler.ru

## VISUALIZATION OF CONVOLUTIONAL NEURAL NETWORK PATTERNS IN THE NOISY RADAR IMAGES CLASSIFICATION PROBLEM

*Kupryashkin I.F., Mazin A.S.*

*The procedure for preparing training and test data is described, the structure of a deep convolutional neural network and the results of object classification on MSTAR dataset radar images with different noise-to-signal ratios are given. The images of the patterns providing the maximum activation of the convolutional layers filters in the presence and absence of noise are shown. A conclusion is formulated about the validity of radar images of the general property of convolutional neural networks, which consists in the complication of the textures of the filters of the higher layers. It is shown that the textural features of the classification become less pronounced under jamming conditions.*

**Key words:** synthetic aperture radar, deep convolutional neural network, object classification.

**Ключевые слова:** РЛС с синтезированной апертурой антенны, глубокая сверточная нейронная сеть, классификация изображений объектов.

### Введение

Вопросы классификации отметок объектов на радиолокационных изображениях (РЛИ) с использованием глубоких сверточных нейронных сетей (ГСНС) на сегодняшний день достаточно подробно освещены в ряде зарубежных публикаций [1-8]. Как правило, ГСНС демонстрируют более высокие точности распознавания объектов по сравнению с алгоритмами, рассматриваемыми как традиционные (K-Neighbor, SVM, AdaBoost, Random Forest, CART и др.) [9]. Высокая эффективность распараллеливания вычислительного процесса, реализующего алгоритмы работы ГСНС, стимулирует их аппаратную реализацию на ПЛИС [10, 11] и других платформах цифровой обработки сигналов.

Однако исследованиям вопросов эффективности ГСНС в условиях, когда классифицируемые РЛИ сформированы в условиях преднамеренных шумовых помех, на сегодняшний день уделяется недостаточное внимание. В связи с этим для иллюстрации базовых возможностей ГСНС в задаче классификации объектов на зашумленных РЛИ в настоящей работе приведено краткое описание процедуры подготовки данных, архитектуры сети VGG-типа, условий ее обучения и результатов работы. Кроме того, с целью наглядной визуализации представлений, формируемых слоями ГСНС в процессе обучения, показаны примеры входных изображений (шаблонов), обеспечивающих максимальную активацию их фильтров. Эти шаблоны характеризуют классифика-

*Описан порядок подготовки обучающих и тестовых данных, приведена структура глубокой сверточной нейронной сети и результаты классификации отметок объектов на радиолокационных изображениях набора MSTAR при различном отношении помех/сигнал. Приведены изображения шаблонов, обеспечивающие максимальную активацию фильтров сверточных слоев при наличии и отсутствии помех. Сформулирован вывод о справедливости и для радиолокационных изображений общего свойства сверточных нейронных сетей, состоящего в усложнении текстур фильтров более высоких слоев. Показано, что текстурные признаки классификации в условиях помех становятся менее выраженными.*

ционные признаки изображений, и дают некоторое представление о том, в виде комбинации каких составляющих описывается представление объекта в слоях обученной ГСНС.

### Подготовка обучающих и тестовых наборов данных

В качестве исходных данных использована открытая часть известного набора MSTAR [12], включающего радиолокационные изображения десяти образцов военной техники с разрешением 0,3 м (рис. 1), полученные по результатам съемки авиационной РЛС X-диапазона.

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  $\hat{I}$ , размерность которого для объектов разных классов является различной (минимальная составляет  $128 \times 128$  отсчетов). На начальном этапе подготовки данных осуществляется подавление фона местности на каждом изображении для исключе-

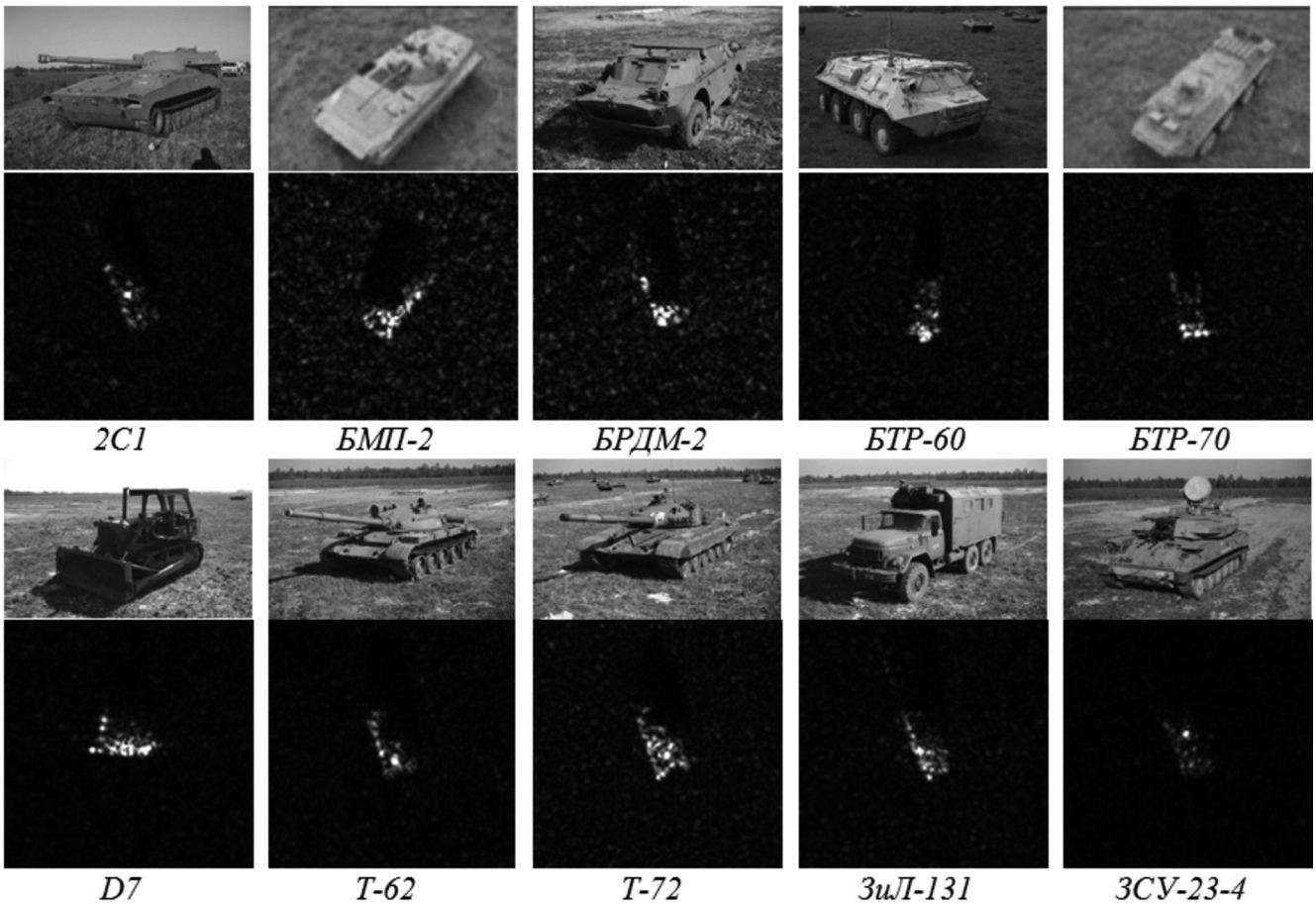


Рис. 1. Оптические и радиолокационные изображения объектов набора MSTAR

ния 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).

Результат воздействия шумовой помехи представляется как помеховое изображение в виде матрицы комплексных некоррелированных значений  $\dot{\eta} = \eta_c + j\eta_s$  размерностью 44×44, при этом элементы матриц действительной  $\eta_c$  и мнимой  $\eta_s$  частей распределены по гауссовскому закону с нулевым средним и единичной дисперсией.

Для обеспечения соответствия корреляционных характеристик сигнального и помехового изображений над матрицами  $\dot{\mathbf{I}}$  и  $\dot{\eta}$  осуществляется процедура пространственной фильтрации

$$\dot{I}_{ij} = \sum_{m=1}^N \sum_{n=1}^N \dot{I}_{mn} G(m-i, n-j);$$

$$\dot{\eta}_{ij} = \sum_{m=1}^N \sum_{n=1}^N \dot{\eta}_{mn} G(m-i, n-j), \quad (1)$$

где  $G(m, n) = \exp\left(-\frac{m^2 + n^2}{2}\right)$  – матрица весовых коэффициентов пространственного фильтра.

Полученные сигнальные и помеховые изображения поэлементно (попиксельно) суммируются с последующим вычислением амплитуд отсчетов

$$I_{ij} = |\dot{I}_{ij} + k\dot{\eta}_{ij}|, \quad (2)$$

где  $k = \frac{M\{|\dot{\mathbf{I}}| > 0\}q}{\sqrt{D\{\dot{\eta}\}}\sqrt{\pi}}$ ;  $\{|\dot{\mathbf{I}}| > 0\}$  – оценка среднего значения амплитуд ненулевых отсчетов сигнального изображения  $\dot{\mathbf{I}}$ ;  $q$  – требуемое отношение помеха/сигнал на РЛИ;  $D\{\dot{\eta}\}$  – оценка дисперсии элементов матрицы  $\dot{\eta}$  после фильтрации (1).

На заключительном этапе изображение (2) преобразуется в целочисленный беззнаковый восьмибитный формат в соответствии с правилом

$$I_{ij} = \begin{cases} \bar{I}_{ij}, & \bar{I}_{ij} \leq 2^8 \\ 2^8, & \bar{I}_{ij} > 2^8 \end{cases}, \quad (3)$$

где  $\bar{I}_{ij} = \text{int}\left\{2^9 \times \frac{I_{ij} - I_{\min}}{I_{\max}}\right\}$ ;  $I_{\max} = \max\{\mathbf{I}\}$  и

$I_{\min} = \min\{\mathbf{I}\}$  – максимальная и минимальная амплитуды отсчетов соответственно.

Сформированные изображения сохраняются в отдельных файлах. После загрузки данных из них и перед подачей на вход нейронной сети целочисленные значения диапазона [0;255] преобразуются в вещественные диапазона [0;1] путем деления на 255.

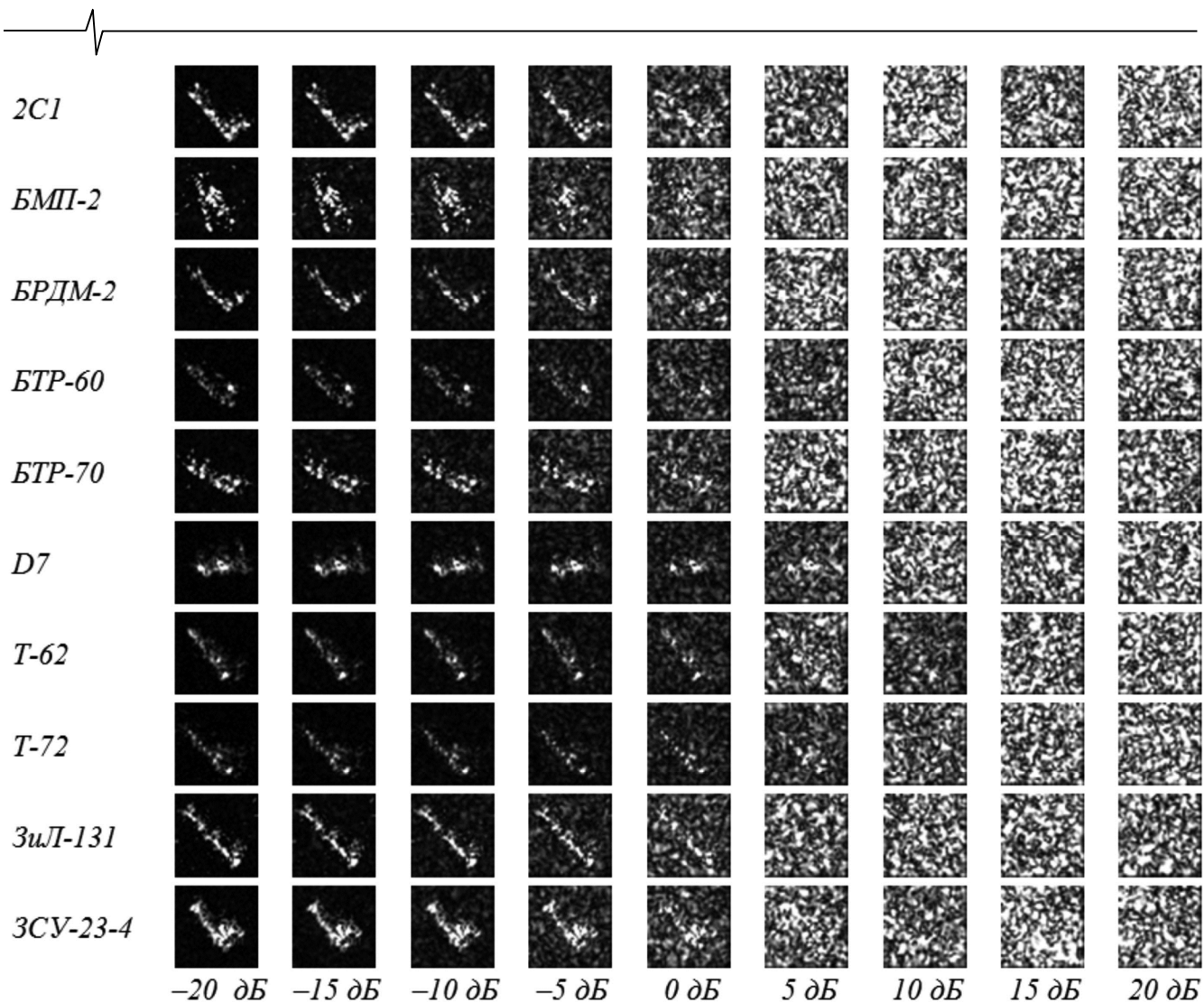


Рис. 2. Примеры изображений обучающих наборов изображений при различном отношении помеха/сигнал

Примеры изображений обучающих наборов, сформированных при  $q = -20; -15; -10; -5; 0; 5; 10; 15; 20$  дБ, приведены на рис. 2.

### Архитектура глубокой сверточной нейронной сети и результаты ее обучения

Разработка и обучение сети, для которой далее в работе принято обозначение C-CNN (Custom Convolutional Neural Network), осуществлялось с использованием фреймворка TensorFlow 2.0 и библиотеки Keras [13]. Архитектура C-CNN (рис. 3) включает четыре пары сверточных слоев (Convolution Layers) с одинаковыми ядрами  $3 \times 3$ , единичным шагом свертки и количеством фильтров 32, 64, 128 и 256 в каждой паре соответственно. После каждой пары включены слои подвыборки (MaxPooling)  $2 \times 2$ . Классификатор включает полносвязный входной слой на 4096 входов и 256 выходов, слой 50-процентного прореживания (Dropout) и выходной слой с 10 выходами по числу классов объектов. В качестве функции активации сверточных слоев и входного слоя классификатора используется кусочно-линейная (ReLU).

В беспомеховой обстановке сеть C-CNN на тестовом наборе продемонстрировала точность классификации 97,91 %, т.е. из 2503 изображений неверно распознала 52.

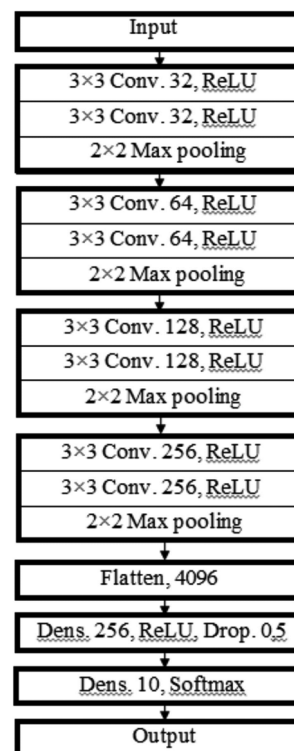


Рис. 3. Архитектура GCHC C-CNN

Точности классификации при различных отношениях помеха/сигнал на обучающем и тестовом наборах

Обучающий набор	Тестовый набор (отношение помеха/сигнал)									
	б/п	-20	-15	-10	-5	0	5	10	15	20
б/п	<b>97,91</b>	27,8	19,31	13,42	11,22	10,93	10,97	10,97	10,93	10,93
-20	97,47	<b>97,79</b>	73,59	34,85	17,74	11,81	11,01	10,97	10,93	10,93
-15	88,50	94,67	<b>96,59</b>	76,64	22,11	11,25	10,97	10,93	10,97	10,89
-10	59,61	65,10	85,01	<b>96,23</b>	49,27	17,18	10,37	08,57	07,93	07,53
-5	49,31	49,43	51,04	67,74	<b>93,18</b>	27,20	08,73	07,73	07,81	07,77
0	40,10	40,14	39,98	41,06	51,88	<b>86,13</b>	20,91	09,77	08,61	09,09
5	26,60	26,76	26,84	27,20	30,32	39,70	<b>55,01</b>	14,66	09,33	09,01
10	13,42	13,58	13,74	14,62	16,50	19,07	24,67	<b>27,44</b>	14,78	11,73
15	10,81	10,97	10,93	10,97	10,93	10,97	13,18	13,98	<b>13,18</b>	11,21
20	10,97	10,85	10,93	10,93	10,93	10,93	10,97	10,93	10,93	<b>10,93</b>

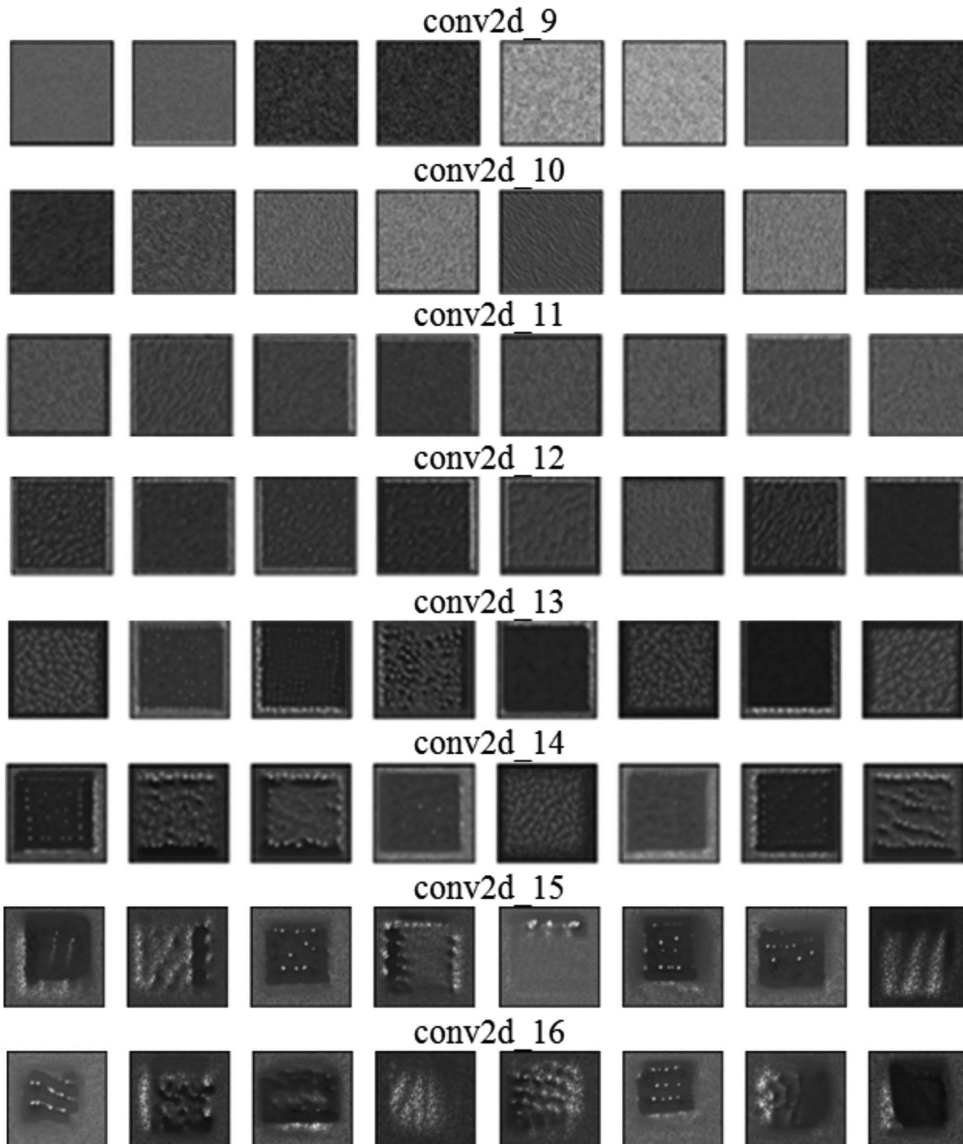


Рис. 4. Входные изображения, обеспечивающие максимальную активацию фильтров сверточных слоев

Далее сеть C-CNN обучалась на всех наборах зашумленных изображений по отдельности, для каждого полученного набора весов проводилось тестирование при всех отношениях помеха/сигнал. Достигнутые значения точности классификации приведены в таблице.

На рис. 4 показаны входные изображения (шаблоны), обеспечивающие максимальную активацию фильтров сверточных слоев (по восемь для каждого слоя) сети,

обученной в беспомеховой обстановке. Для их получения использовался метод градиентного восхождения в пространстве входов, описанный в [13].

Видно, что в более высоких слоях фильтры начинают описывать все более сложные текстурные признаки, что вообще является типичным для сверточных сетей [13]. При этом отчетливо видна характерная рамка по периметру изображений, особенно для фильтров высо-

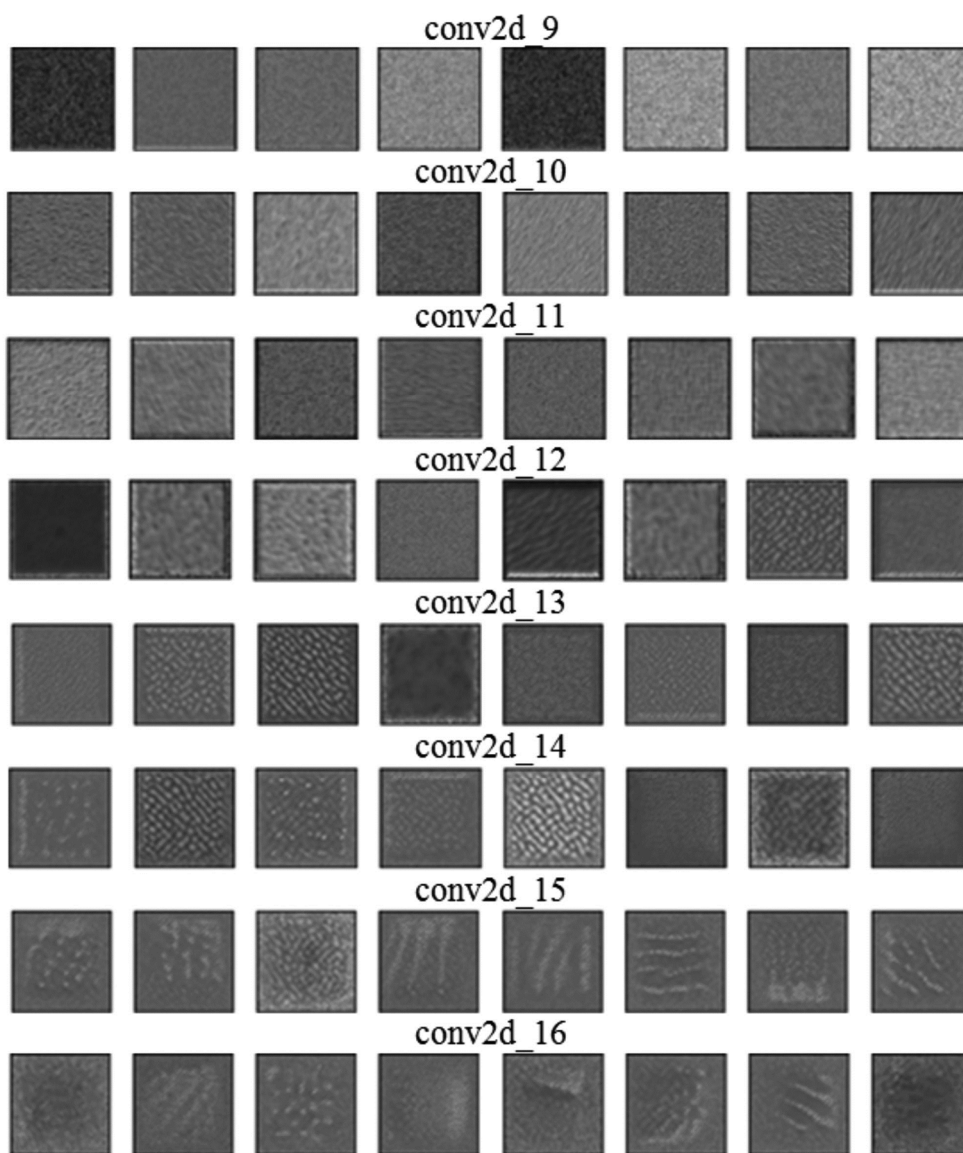


Рис. 5. Шаблоны, обеспечивающие максимальную активацию фильтров сверточных слоев при отношении помеха/сигнал 0 дБ

ких слоев. Это, по всей видимости, 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 слоев (conv2d\_9...conv2d\_12) кодируют простые направленные контуры. Фильтры conv2d\_13 и conv2d\_14 кодируют сочетания направленных контуров и зернистой текстуры, обусловленной спекл-эффектом, типичным для РЛИ. Фильтры conv2d\_15 и conv2d\_16 кодируют наиболее сложные текстурные признаки, практически неинформативные для восприятия человеком, но, по всей видимости, позволяющие представить изображения объектов в виде их суперпозиции.

На рис. 5 показаны входные изображения, обеспечивающие максимальную активацию фильтров сверточных слоев (по восемь для каждого слоя) сети С-CNN, обученной на наборе изображений с отношением помеха/сигнал 0 дБ.

Из сравнения шаблонов на рис. 4 и 5 видно, что в

условиях помех разнообразие текстурных признаков в более высоких слоях стало гораздо меньшим, а сами шаблоны – более однородными.

### 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. Уменьшение разнообразия текстурных признаков в условиях помех свидетельствует о снижении чувствительности сети к особенностям конкретного набора изображений.

Влияние шумовых помех вполне ожидаемо проявляется в снижении точности классификации отметок объектов на РЛИ. Ее максимальное значение в беспомеховой обстановке составило 97,91%, при уровне помехи, сопоставимом со средним уровнем отметок объектов ( $q = 0$  дБ) она остается достаточно высокой – 86,13 %,

но при дальнейшем увеличении отношения помеха/сигнал резко снижается. Так, при  $q = 5$  дБ правильная работа сети наблюдается примерно в половине случаев (55,01 %), а при  $q = 15$  дБ и более – сводится к простому угадыванию (для десятиальтернативной классификации точность составляет порядка 10 %).

Все же следует отметить, что даже при  $q = 0$  дБ, когда визуальное дешифрирование отметок представляется практически невозможным (рис. 2), сеть демонстрирует высокие результаты, что свидетельствует об эффективности применения технологий глубокого обучения в этой задаче.

### Литература

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. Wang H., Chen S., Xu F., Jin Y.-Q. Application of Deep-Learning Algorithms to MSTAR Data. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 3743-3745. DOI: 10.1109/IGARSS.2015.7326637.
3. 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.
4. Anas H., Majdoulayne H., Chaimae A., Nabil S.M. Deep Learning for SAR Image Classification. 2020. DOI: 10.1007/978-3-030-29516-5\_67.
5. Chen S., Wang H. SAR Target Recognition Based on Deep Learning. 2014 International Conference on Data Science and Advanced Analytics (DSAA), 2014, pp. 541-547. DOI: 10.1109/DSAA.2014.7058124.
6. 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.
7. Furukawa H. Deep Learning for End-to-End Automatic Target Recognition from Synthetic Aperture Radar Imagery. arXiv:1801.08558v1 [cs.CV] 25 Jan 2018.
8. 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>.
9. Бородинов А.А., Мясников В.В. Сравнение алгоритмов классификации радарных изображений при различных методах предобработки на примере базы MSTAR. IV Международная конференция и молодежная школа «Информационные технологии и нанотехнологии» (ИТНТ-2018).
10. Zhang C., Li P., Sun G., Guan Y., Xiao B., Cong J. Optimizing FPGA-based accelerator design for deep convolutional neural networks. Proceedings of the 2015 ACM/SIGDA International Symposium on Field Programmable Gate Arrays, 2015, pp. 161-170. DOI: 10.1145/2684746.2689060.
11. Зоев И.В., Марков Н.Г., Береснев А.П., Ягунов Т.А. Аппаратная реализация на ПЛИС сверточных нейронных сетей для распознавания объектов на изображениях. GraphiCon, 2018. С. 200-203.
12. <https://www.sdms.afrl.af.mil/index.php?collection=mstar>.
13. Шолле Ф. Глубокое обучение на Python. СПб: Питер, 2018. 400 с.