# **Отчёт**

## **По обучению нейронной сети на имитированных данных для распознавания объектов реальной техники на РЛИ**

## **Постановка задачи**

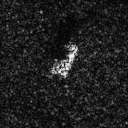
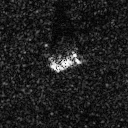
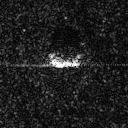
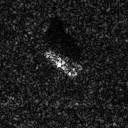
Подобрать тип и гиперпараметры свёрточной нейронной сети (СНС) для максимизации вероятности правильного распознавания (ВПР) на радиолокационных изображениях (РЛИ) объектов реальной техники при обучении СНС на имитированных данных.

## **Исходные данные**

Для обучения и проверки работы СНС использовались два типа данных: имитированные и реальные. Данные представляют собой одноканальные РЛИ размером 128х128 пикселей.

В качестве обучающих данных использовались имитированные РЛИ по трёхмерным моделям техники, полученные с помощью разработанного в ННИИРТ имитатора [1].

В качестве проверочных реальных данных использовались РЛИ техники из базы данных MSTAR [2]. В данных содержатся РЛИ 7 классов наземной техники: самоходная артиллерийская установка 2С1 (2S1), БРДМ-2 (BRDM\_2), бульдозер (D7), танк Т-62 (T62), танк Т-72 (T72), грузовик ЗИЛ-131 (ZIL131) и ЗСУ-23-4 «Шилка» (ZSU\_23\_4). Пример изображений приведён на рисунке 1.



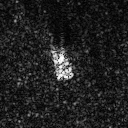
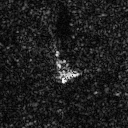
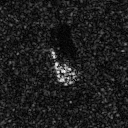


Рисунок 1. Примеры РЛИ каждого класса по порядку: 2С1 (2S1), БРДМ-2 (BRDM\_2), бульдозер (D7), танк Т-62 (T62), танк Т-72 (T72), грузовик ЗИЛ-131 (ZIL131) и ЗСУ-23-4 «Шилка» (ZSU\_23\_4)

Из базы данных MSTAR были выбраны именно эти 7 классов, поскольку только они были получены при одинаковых полётных и погодных условиях. Это важно учитывать, так как при обучении на РЛИ, полученных в разных условиях, нейронная сеть может начать классифицировать некоторые РЛИ не по участкам, содержащим части объекта, а по пикселям фона.

*Таблица 1. Количество РЛИ в реальной и имитированной выборках*

|  |  |  |
| --- | --- | --- |
| Название класса | Количество РЛИ | |
| Реальные | Имитированные |
| 2S1 | 299 | 2880 |
| BRDM\_2 | 298 | 2880 |
| D7 | 299 | 2870 |
| T62 | 299 | 2880 |
| T72 | 299 | 2880 |
| ZIL131 | 299 | 2880 |
| ZSU\_23\_4 | 299 | 2880 |

Из таблицы 1 видно, что выбранные базы данных для обучения и тестирования СНС сбалансированы.

Обучение СНС производилось на выборке, основанной на имитированных данных. Для увеличения обучающей выборки, имитированные РЛИ были дополнительно аугментированы различными способами, а именно: применялся сдвиг по яркости, сдвиг центра и зашумление.

Для сдвига по яркости ко всем пикселям исходного изображения добавлялось или вычиталось случайное значение из диапазона от 1 до 10. Полученное изображение добавлялось к исходным имитированным данным.

Для сдвига изображения по горизонтали или вертикали, выбирались два случайных числа. Далее каждый пиксель изображения смещался на эти два значения. При этом, если пиксель был граничным, то он переносился через границу на противоположную сторону.

Добавление шума проходило в два шага. В начале выбирался случайный пиксель для зашумления с вероятностью 15%, далее значение пикселя с одинаковым шансом принимало либо максимальное, либо минимальное значение. Так называемый метод «соль и перец» («salt and pepper»).

На рисунках 2-4 показаны статистические значения яркости выборок имитированных РЛИ, имитированных и аугментированных РЛИ, а также РЛИ из базы данных MSTAR.

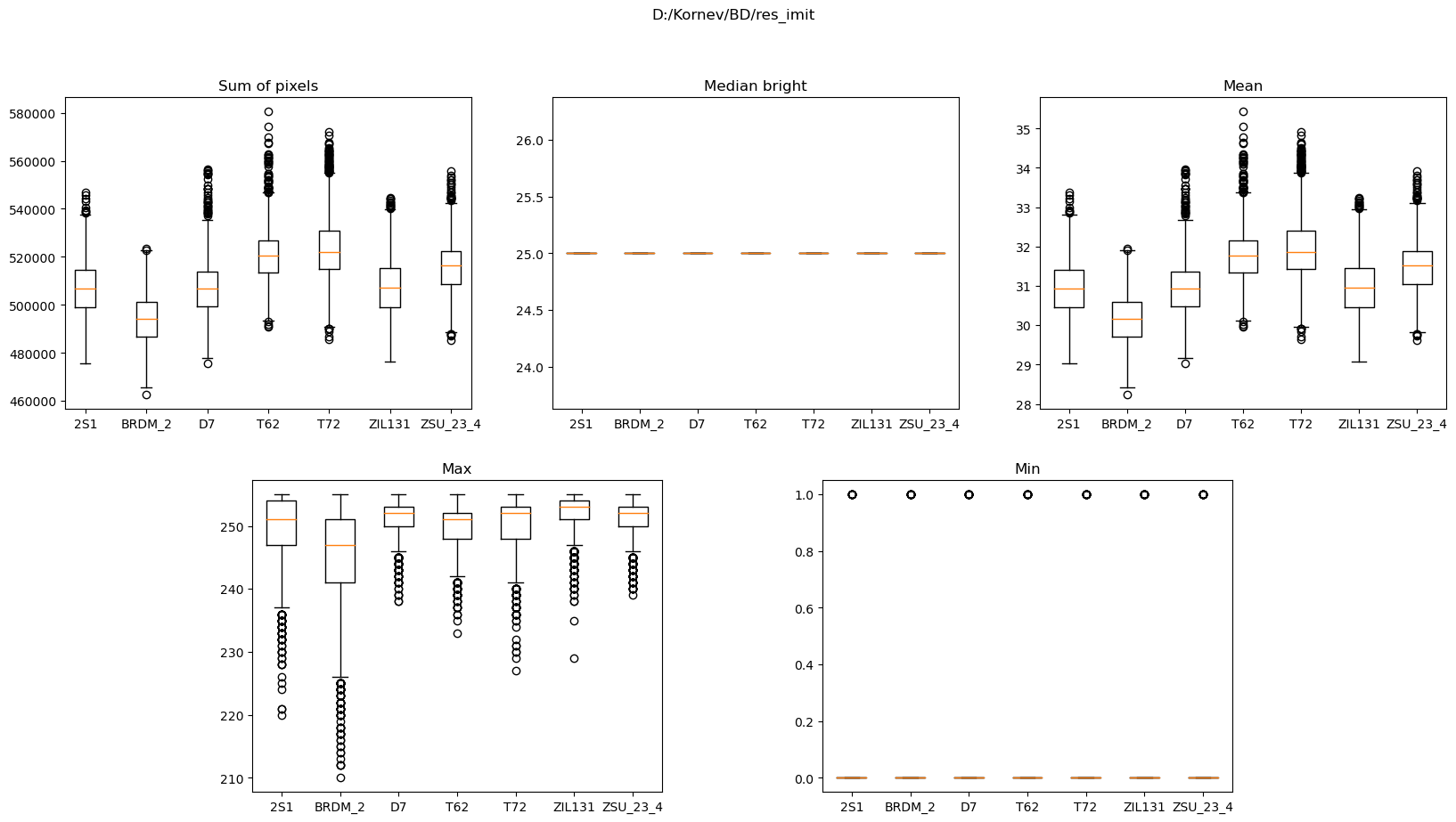


Рисунок 2. Статистика по исходным имитированным РЛИ

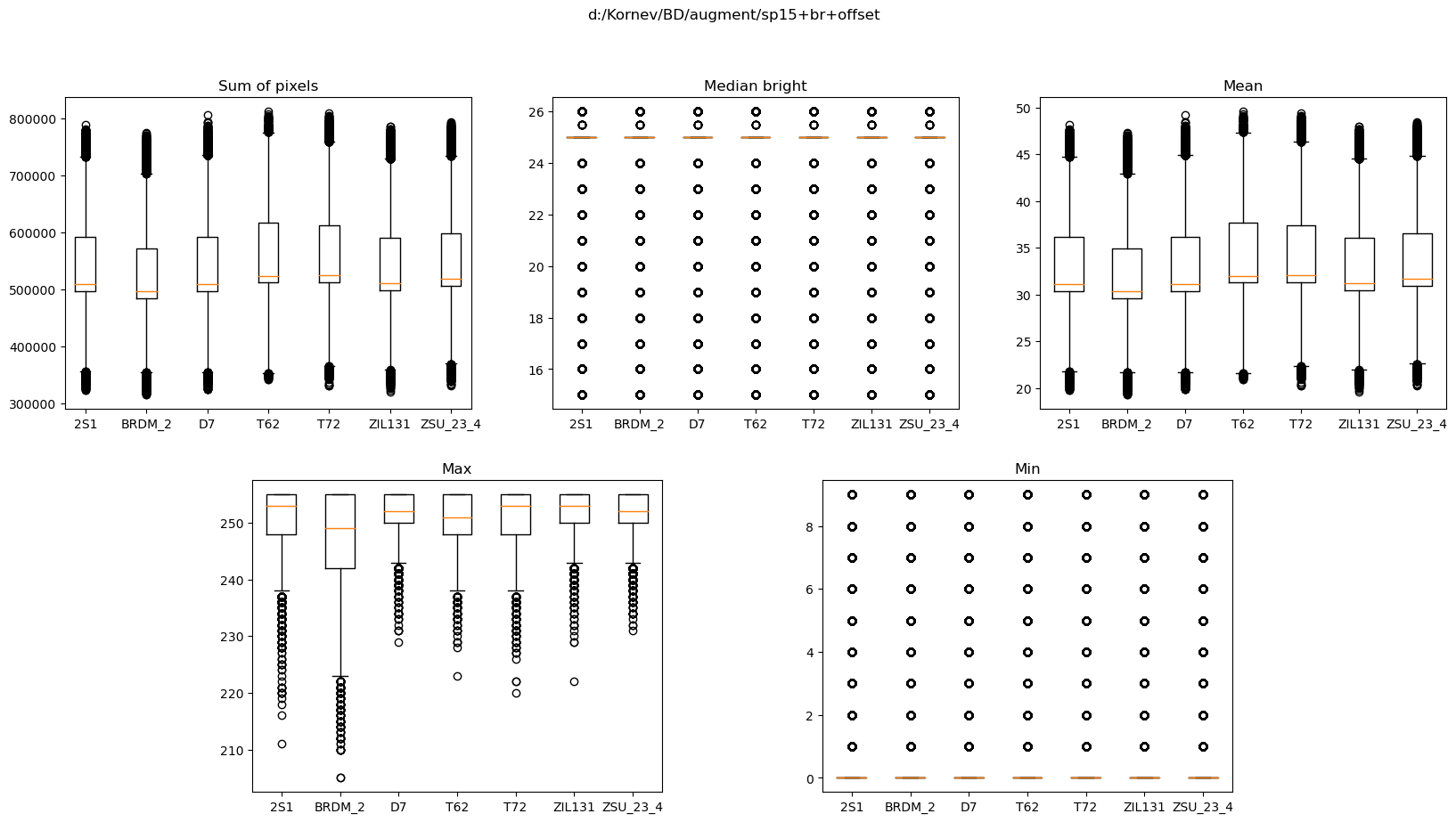


Рисунок 3. Статистика по аугментированным имитированным РЛИ

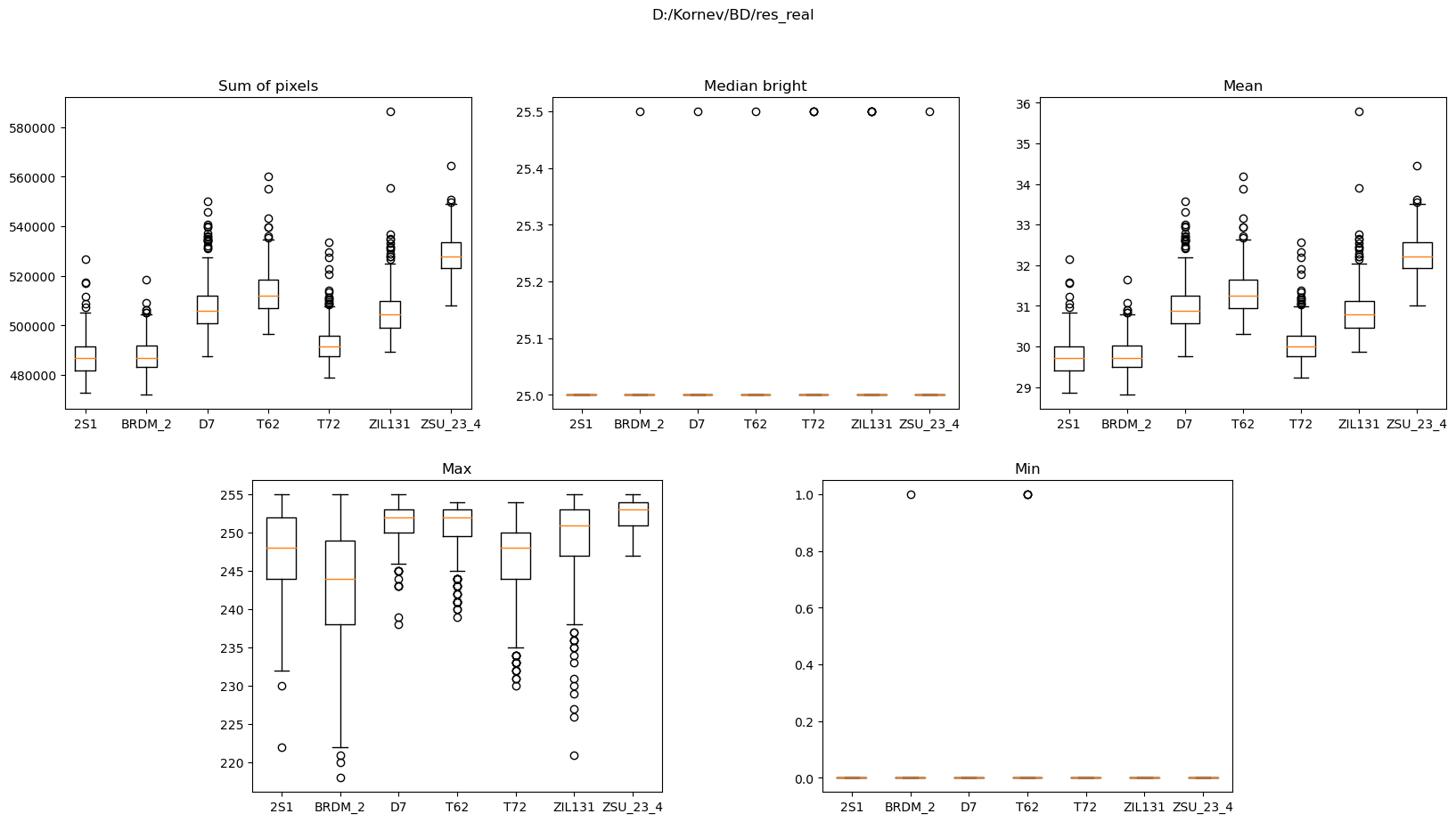


Рисунок 4. Статистика по реальным РЛИ из MSTAR

На рисунках 2-4 показаны по 5 графиков. На первом (Sum of pixels) представлена сумма пикселей РЛИ (суммарная яркость РЛИ), на втором (Median bright) – медианная яркость РЛИ, на третьем (Mean) – среднее значение яркости РЛИ, на четвёртом (Max) – максимальное значение яркости, на пятом (Min) – минимальное значение яркости. Для показа значений на графиках применялась коробчатая диаграмма, суть которой показана на рисунке 4.

Изменить графики: убрать сумму пикселей, соединить на одном изображении максимальное и минимальное значения, добавить аугментацию для вырезания квадратов и замены их пустыми участками.

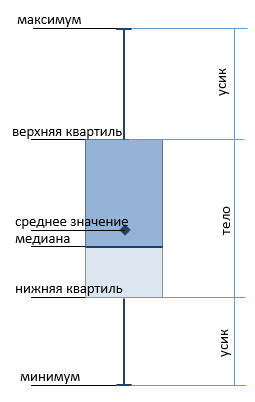


Рисунок 5. Суть коробчатой диаграммы

Аугментация имитированных РЛИ проводилась для расширения статистических характеристик так, чтобы статистика реальных изображений заведомо попадала в диапазон аугментированых РЛИ.

Для аугментации данных применялись 3 метода: смещение по яркости, смещение центра изображения и добавление шума «соль и перец» («salt and pepper»). На рисунке 6 представлены матрицы перепутывания распознавания СНС различных комбинаций методов аугментации исходных имитированных данных.

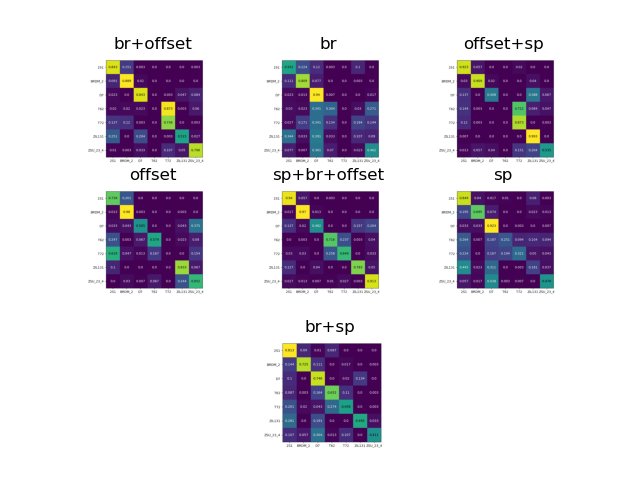


Рисунок 6. Матрицы перепутывания для различных аугментаций:  
br – сдвиг изображений по яркости; offset – сдвиг центра изображения на небольшую величину; sp - добавление на изображения мультипликативного шума «соль и перец»

Такой способ аугментации был выбран исходя из результата исследования влияния различных аугментаций на обучение нейронной сети (см. рисунок 5). Для этого одну и ту же нейронную сеть ResNet-18 одинаковым образом обучали на 7-ми разных наборах данных. Каждый из этих наборов был аугментирован так, чтобы в результате получить всевозможные комбинации аугментаций исходных изображений. На рисунке 6 представлены матрицы перепутывания для наборов данных с различными комбинациями способов аугментации.

Провести такое исследования на всех архитектурах, которые буду использоваться в работе

Из полученных матриц перепутывания видно, что отдельные аугментации данных не дают улучшения точности СНС, поскольку некоторые классы при распознавании перепутываются почти полностью. Также можно сделать вывод, что наибольший вклад в увеличение точности распознавания вносит добавление мультипликативного шума в обучающую выборку, так как после его применения уменьшается перепутывание классов и матрица принимает необходимый диагональный вид. Наилучший результат был получен при применении всех трёх аугментаций.

## **Процесс поиска типа и подбора гиперпараметров свёрточной нейронной сети**

Для решения задачи были протестированы несколько архитектур нейронных сетей, такие как: ResNet-18, ResNet-37, ResNet-52, DenseNet, ViT. В результате обучения ResNet-37 показала максимальную ВПР в 81%. В процессе подбора архитектуры были испробованы разные функции активации, например, tanhexp, swish, mish, GELU. Наилучший результат в решении задачи распознавания объектов на РЛИ показала функция активации ReLU.

Добавить новые архитектуры, сохранить их результат обучения в файлы

Также были протестированы различные гиперпараметры для модели СНС: разная конфигурация слоёв, различные размеры ядер свёртки и т.д. График обучения моделей представлен на рисунке 6.

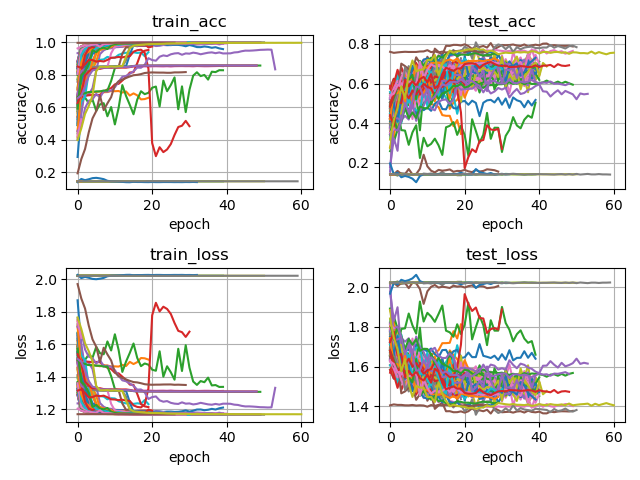


Рисунок 7. Графики точности и ошибки моделей в процессе обучения

Из графиков на рисунке 7 видно, что большинство моделей не превышает ВПР 80%. И только при определённых гиперпараметрах было достигнуто превышение этого значения ВПР.

## **Проверка гипотезы центрирования объекта на РЛИ**

Для проверки гипотезы, что при центрировании объектов на РЛИ перед распознаванием повышается ВПР, было проведено исследование.

Чтобы найти центр изображения, был выбран метод вычисления центра тяжести изображения. Для этого выполнялись следующие шаги:

1. Вычислить точку центра изображения, как половину ширины и высоты.
2. Вычислить вектор смещения по формуле:

Где: – вектор до начального центра, – вектор до *i*-ого пикселя, *B*– яркость *i*-ого пикселя, *N* и *M* – ширина и высота изображения соответственно.

1. Смещаем все пиксели изображения соответственно вектору . Пиксели, которые проходят через границу, переносятся на противоположную сторону.

Такое преобразование изображения устойчиво, с точностью до шума подложки и не превышает смещения в 1 пиксель.

Далее для оценки смещений и сравнений с другими наборами данных были построены усреднённые по всем углам РЛИ по каждому классу. На рисунке 8 приведён пример такого изображения для класса 2C1 (2S1).

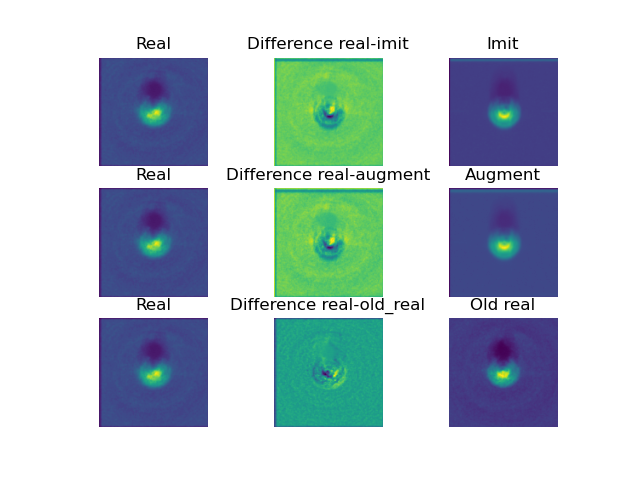


Рисунок 8. Пример усреднённых по всем углам изображения для класса 2S1 для различных наборов данных: реальных РЛИ (old real), реальных центрированных (real), имитированных (imit), аугментированных имитированных РЛИ (augment)

На центральных изображениях представлены их разности, чтобы можно было оценить различия между данными. Это помогает наглядно увидеть, какой вклад вносит центрирование изображений в набор данных и насколько сильно оно их изменяет.

В результате центрирования изображений увеличения ВПР не наблюдалось.

После аугментаций и прочего идеально будет попробовать ещё раз сравнить изображения таким способом.

## **Оптимальная СНС для распознавания объектов на РЛИ**

Наиболее высокая ВПР на проверочных реальных данных 81.4% наблюдалась при обучении на аугментированых имитированных РЛИ с вероятностью зашумления пикселя равной 15% и гиперпараметрами модели СНС ResNet-37: размер батча 16, количество эпох обучения 60, конфигурация слоёв [2, 2, 3, 2], размер ядра свёртки для первого слоя в блоке равен 7, в второй слой имеет размер ядра свёртки равный 5.

Под конфигурацией слоёв понимается количество одинаковых блоков в архитектуре модели до момента понижения размерности изображения и увеличения количества фильтров в следующем блоке.

В качестве метрики для оценки качества СНС была выбрана точность (Accuracy), для оценки ошибки выбрана метрика кросс-энтропии (CrossEntropyLoss).

График обучения оптимальной СНС представлен на рисунке 9.

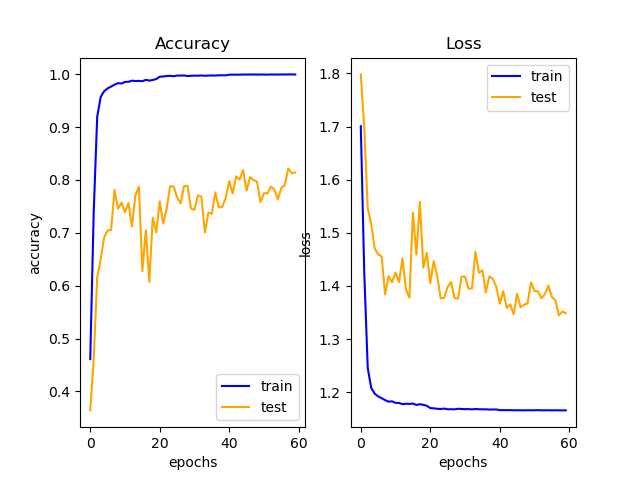


Рисунок 9. График обучения и ошибки модели

Из графика видно, что метрика Accuracy выходит на плато, следовательно, дальнейшее обучение не принесёт существенного возрастания ВПР. Это также можно пронаблюдать на метрике ошибки, где также виден выход на плато.

На рисунке 10 представлена матрица перепутывания для семи выбранных классов объектов.

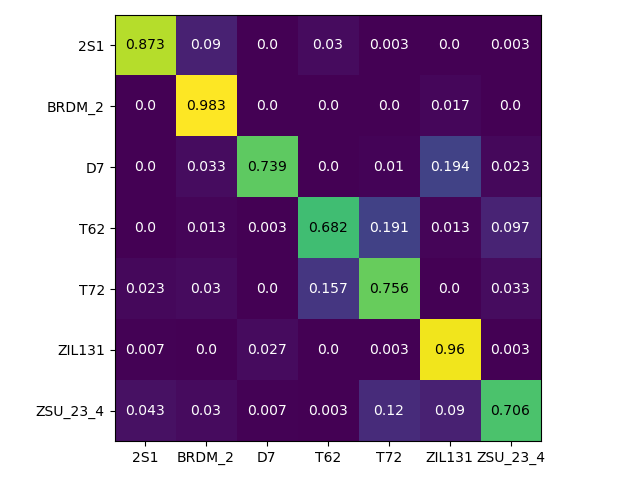


Рисунок 10. Матрица перепутывания нейронной сети

Максимально достигнутая средняя ВПР составила 81.4%, минимальная 68.2%. Однако можно учесть, что максимальное перепутывание наблюдается между классами T62 и T72, которые очень схожи между собой и их, в целом, можно объединить в один класс «танки». Тогда минимальная ВПР будет у класса ZSU\_23\_4 и составит 70.6%.

## **Выводы**

1. В результате центрирования изображений увеличения ВПР не наблюдалось.
2. Наибольший вклад по увеличению точности распознавания вносит добавление мультипликативного шума в обучающую выборку данных.
3. Из протестированных архитектур СНС наилучший результат показала ResNet-37 с функцией активации ReLU.
4. Наиболее высокая ВПР на проверочных реальных данных наблюдалась при обучении на аугментированых имитированных РЛИ с вероятностью зашумления пикселя равной 15% и гиперпараметрами модели СНС ResNet-37: размер батча 16, количество эпох обучения 60, конфигурация слоёв [2, 2, 3, 2], размер ядра свёртки для первого слоя в блоке равен 7, в второй слой имеет размер ядра свёртки равный 5. ВПР составила 81.4% на реальных данных.
5. Максимальное перепутывание наблюдается между классами T62 и T72, которые очень схожи между собой и их, в целом, можно объединить в один класс «танки».

## **Результат**

Были проанализированы реальные и имитированные наборы РЛИ и их отличия в статистических метриках. По результатам анализа были подобраны способы аугментации имитированных РЛИ. Было проведено исследование влияния различных аугментаций на наборы данных и процесс обучения моделей нейронных сетей. Получены комбинации аугментаций, оказывающие наибольшее положительное влияние на распознавание объектов на РЛИ.

Были протестированы различные архитектуры нейронных сетей, выбрана оптимальная и получены наилучшие гиперпараметры, при которых достигнута максимальная ВПР.

На основе полученных результатов можно провести исследование и разработать систему распознавания на основе СНС наземных и надводных целей для бортовой радиолокационной станции с режимом синтеза апертуры с заданными характеристиками.

**Ссылки**

1. Пат. 2835989 Российская Федерация, МПК G 06 T 15/06. Способ быстрого имитирования радиолокационных изображений / Прохоров А.Г., Ярцев И.М., Дмитриев А.В.; заявитель и патентообладатель Акционерное общество «Федеральный научно-производственный центр «Нижегородский научно-исследовательский институт радиотехники». - 2024118255; заявл. 28.06.2024; опубл. 07.05.2025.
2. T. Ross, S. Worrell, V. Velten, J. Mossing, and M. Bryant, “Standard SAR ATR evaluation experiments using the MSTAR public release data set”, Proc. SPIE, vol.3370, pp.566–573, 1998.