

**СВЕРТОЧНЫЕ ВЕЙВЛЕТ-БЛОКИ КАК ИНСТРУМЕНТ УМЕНЬШЕНИЯ
КОЛИЧЕСТВА ПАРАМЕТРОВ МОДЕЛЕЙ**

Воробей Владислав Александрович, аспирант

Белорусский государственный университет

Varabei Uladzislau, PhD student, Belarussian State University, v.vorobey.edu@gmail.com

Аннотация. На основе дискретного вейвлет-преобразования семейства CDF-9/7 реализован сверточный вейвлет-блок, который может быть использован для дополнительного извлечения признаков из входного сигнала. Полученный блок встроен в модели ResNetV2-50 и MobileNetV2. Исходные и модифицированные модели протестированы на наборе данных Flowers. Использо-

ние реализованного блока позволило значительно сократить количество параметров у моделей с сохранением высокого уровня качества.

Ключевые слова: нейронные сети, глубокое обучение, вейвлеты, дискретное вейвлет-преобразование, классификация изображений.

Описание блоков. Для дополнительного извлечения признаков был реализован сверточный вейвлет-блок на основе семейства вейвлетов CDF-9/7. Реализация использует лифтинг-схему, которую можно найти в [1]. Блок имеет следующую архитектуру: сперва к сигналу применяется сверточный блок (свертка 1×1 , пакетная нормализация, активация LeakyReLU), затем 3 уровня вейвлет-разложения, конкатенация выходов после каждого шага разложения и наконец выходной сверточный блок для восстановления исходного количества каналов (свертка 1×1 , пакетная нормализация, активация LeakyReLU). Наглядно схема блока представлена на рис. 1.

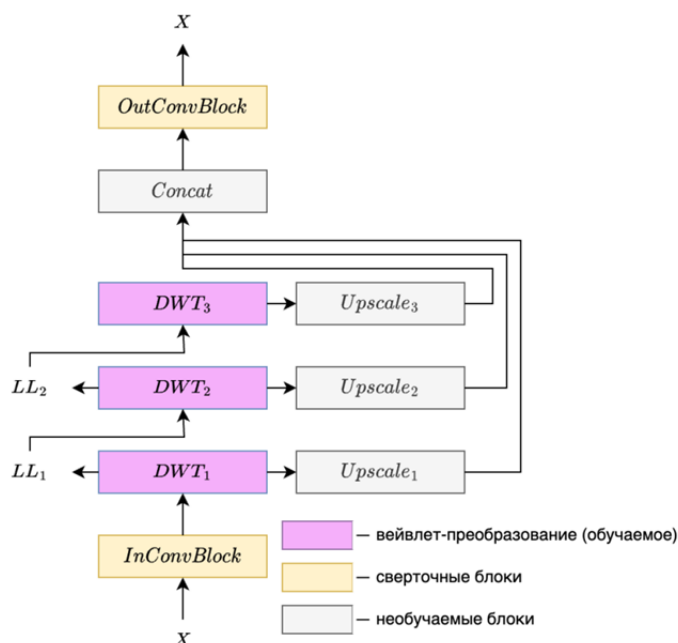


Рисунок 1. – Сверточный вейвлет-блок

Блок встраивался в существующие модели следующим образом: для каждого набора слоев, который уменьшает размерность изображения в 2 раза, входной сигнал шел параллельно как по исходным (существующим) слоям, так и по реализованному блоку, а затем полученные выходы поэлементно складывались. Зачастую таких блоков в моделях 5 (размерность входного изображения к концу сверточных операций уменьшается в 32 раза). Также входное изображение для модели раскладывалось одним шагом вейвлет-разложения, выходы конкатенировались вдоль оси каналов, а затем полученное изображение увеличивалось в 2 раза для сохранения исходного размера по ширине и высоте. Схема встраивания блока в модель показана на рис. 2.

Описание данных и параметров моделей

Для проведения экспериментов использовались данные из набора Flowers [2]. Всего датасет содержит 4317 изображений из 5 категорий: ромашка, одуванчик, роза, подсолнух и тюльпан. Из каждой категории отбиралось 20% данных для тестирования. Далее из всех тренировочных данных 15% выделялось на валидацию моделей. Все разбиения проводились с зафиксированным параметром случайности. Таким образом, для тренировки, валидации и тестирования было доступно 2934, 518 и 865 изображений соответственно. Количество изображений в классах сбалансировано, поэтому в качестве метрики использовалась ассигасу.

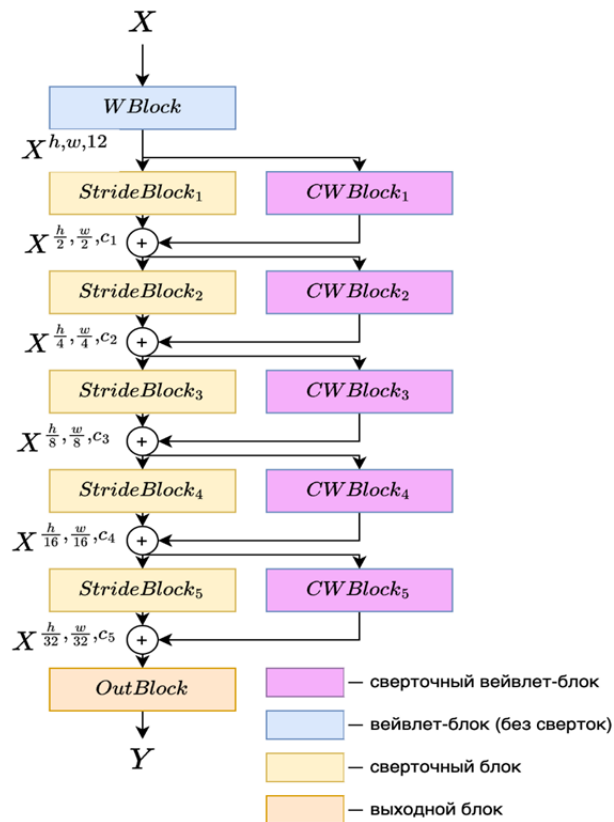


Рисунок 2. – Модель на основе сверточного-вейвлет блока

Моделям на вход поступали изображения размером 256×256 . Для валидации и тестирования изображения масштабировались с сохранением соотношения сторон так, чтобы новое изображение было в центре прямоугольника требуемого размера (для меньшей стороны с обеих краев недостающие пиксели заполнялись 0). Для тренировочных изображений при масштабировании использовался размер 256×352 , и затем из полученного изображения извлекался случайный патч размером 256×256 . К тренировочным изображениям также применялись следующие аугментации: отражение относительно вертикальной оси ($p = 0.5$), вырезание патча, в котором ширина и высота выбираются случайным образом в диапазоне $[0.75, 1.0]$ от исходного размера, и последующее его приведение к требуемому размеру методом билинейной интерполяции.

Примеры полученных тестовых изображений показаны на рис. 3.

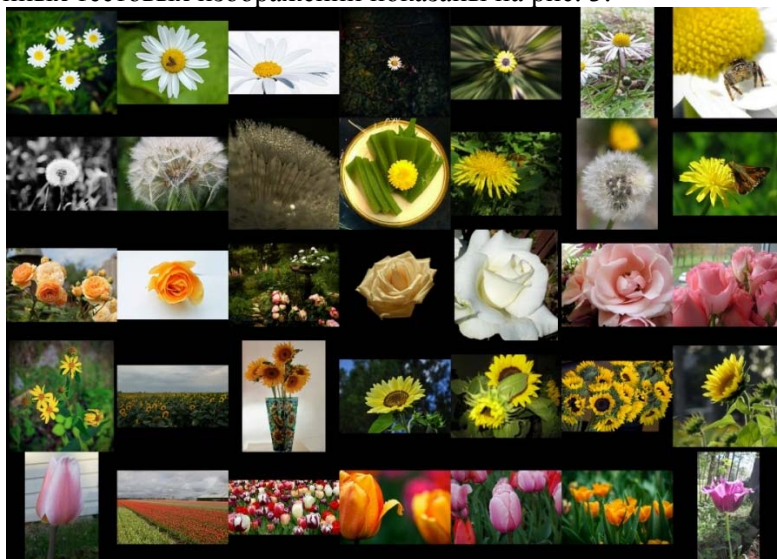


Рисунок 3. – Примеры тестовых изображений

Для проведения экспериментов использовались модели ResNetV2-50 [3] и MobileNetV2 [4]. Для уменьшения общего числа параметров моделей количество фильтров в сверточных слоях масштабировалось при помощи параметра α . В ходе экспериментов использовалось фиксированное значение $\alpha = 0.75$ для моделей со сверточными вейвлет-блоками. Для модели на основе ResNetV2-50 5-й блок уменьшения размерности не подвергался изменениям, поскольку в нем содержится очень большое количество параметров относительно их общего числа в модели.

Результаты.

В таблице показаны значения метрики на тестовой выборке, полученные для исходных и обновленных моделей. Для исключения влияния случайности все варианты архитектуры просчитывались по 5 раз (в таблице указано среднее значение).

Таблица – Сравнение результатов

Модель	Кол-во общих пар- ов	Кол-во DWT пар-ов	Доля DWT пар-ов	Размер (Мб)	Метрика (accuracy)
ResNetV2-50_base	23.56 М	0	0	90.1	0.7732
ResNetV2-50_a0.75_ per-block_no-b5	20.08 М	786 К	0.039	76.1 (-15.4%)	0.7799 (+0.0067)
MobileNetV2_base	2.26 М	0	0	8.9	0.7792
MobileNetV2_a0.75_ per-block	1.49 М	94 К	0.063	5.9 (-33.7%)	0.7665 (- 0.0127)

Можно видеть, что для обеих моделей удалось значительно сократить количество параметров, практически не потеряв при этом в качестве. Модель на основе MobileNetV2 показала себя лучше: при использовании реализованных вейвлет-блоков количество параметров уменьшилось на 33%, а качество упало лишь на 1.27%.

Список использованных источников

1. Wavelet CDF 9/7 Implementation [Electronic resource] // Getreuer: On Wavelet Implementation, 1997. – Mode of access: <https://getreuer.info/posts/waveletcdf97/index.html>. – Date of access: 21.10.2023.
2. TF Flowers dataset [Electronic resource] // The TensorFlow Team. – Mode of access: https://www.tensorflow.org/datasets/catalog/tf_flowers. – Date of access: 19.04.2024.
3. Identity Mappings in Deep Residual Networks / K. He [et al.] // arXiv.org. – Mode of access: <https://arxiv.org/abs/1603.05027>. – Date of access: 19.04.2024.
4. MobileNetV2: Interested Residuals and Linear Bottlenecks / M. Sandler [et al.] // arXiv.org. – Mode of access: <https://arxiv.org/abs/1801.04381>. – Date of access: 19.04.2024.