

Исследование влияния предобученных оснований нейронных сетей на качество сегментации кусков руды на фотографии

В.Ю. Устимов¹, Д.А. Полеценко²

¹*Старооскольский технологический институт им. А.А.Угарова (филиал)
Национального исследовательского технологического университета «МИСИС»,
г. Старый Оскол*

²*Филиал Национального исследовательского технологического университета
«МИСИС», г. Губкин, Старый Оскол*

Аннотация: В статье рассмотрены самые популярные семейства предобученных оснований нейронных сетей. Были рассмотрены их структуры, а также произведено сравнение результатов работы при сегментации изображения на FPN – сети. После тестирования и сравнения было получено, что самое высокое качество сегментации, равное 55,1 % по классу «куски руды» при общем качестве сегментации по трем классам составляющем 98,93 %, показала FPN – сеть на предобученном основании EfficientNetB2 после 7 эпох обучения.

Ключевые слова: сегментация, нейронная сеть, предобученное основание, EfficientNet, SEResNet.

Введение

На сегодняшний день задача распознавания кусков руды после взрыва в карьере является актуальной. От качества взрывных работ, производимых в карьере, зависит эффективность работы всех последующих этапов обработки руды.

На данный момент оценка качества взрывных работ производится с помощью программного обеспечения, которое находит контуры кусков руды на основе разности цветов. Трудозатраты человека на дорисовывание и исправление контуров после этой программы колоссальные. Необходимо уменьшить работу человека при распознавании кусков руды.

На сегодняшний день существует большое количество способов распознавания образов изображений, но это достаточно трудная задача, так как у одного и того же образа могут быть разные цветовые решения, различное освещение, наклон в сторону.

Первые методы распознавания объектов были следующими [1]:

- линейный дискриминантный анализ – метод, который минимизирует внутриклассовое расстояние и максимизирует межклассовое расстояние, где сильное влияние на результат оказывает освещенность [2];
- гибкие контурные модели объектов – анализ, который используется для построения контуров, описывающих объект. Данный метод сталкивается с проблемой определения вида контура и его центром [3];
- сравнение эталонов, в котором используется сравнение значений двух участков фотографии. Недостатком является потребление большого количества ресурсов компьютера, а также невозможность изменения условий проведения фотосъемки [4];
- скрытые Марковские модели – алгоритм обучения, который максимизирует отклик каждой модели на запрос своего класса, то есть данный метод не может распознавать объекты [5].

Данные методы не дают хороших результатов и потребляется большое количество ресурсов компьютера, поэтому, на сегодняшний день, используют способы, основанные на искусственном интеллекте.

Нейронная сеть состоит из весовых коэффициентов предобученных оснований и настраиваемых весов. В процессе обучения нейронной сети веса основания используются в качестве базовых, а настраиваемые веса изменяются в конкретности от задачи использования нейронной сети. Процесс обучения становится более легким и быстрым, потому что не требуется полностью настраивать все коэффициенты, а только небольшое количество от общего числа.

Для наиболее верного сравнения результатов качества сегментации будем сравнивать два предобученных основания EfficientNet и SEResNet на

нейронной сети FPN и при одинаковом количестве эпох обучения, а конкретнее 3, 5, 7, 10 эпох. Обучение происходило в Google Colaboratory на GPU.

Описание обучающей выборки

Обучающая выборка состоит из 67 фотографий с результатами взрывных работ в карьере. Исходный размер фотографии составляет 3808x5152 пикселей. Данный размер изображения не помещается в памяти видеокарты, что делает невозможным обучение на изображении такого большого размера. FPN сеть и предобученные основания, которые были выбраны для сравнения, наиболее эффективно работают на размере изображения 224x224 пикселей.

Для того, чтобы было возможным обучение, было принято решение разрезать исходную фотографию на маленькие изображения с присвоением нового имени, указывающего конкретное местоположение данного сегмента на большой фотографии. Пример разрезания фотографии представлен на рисунке 1.

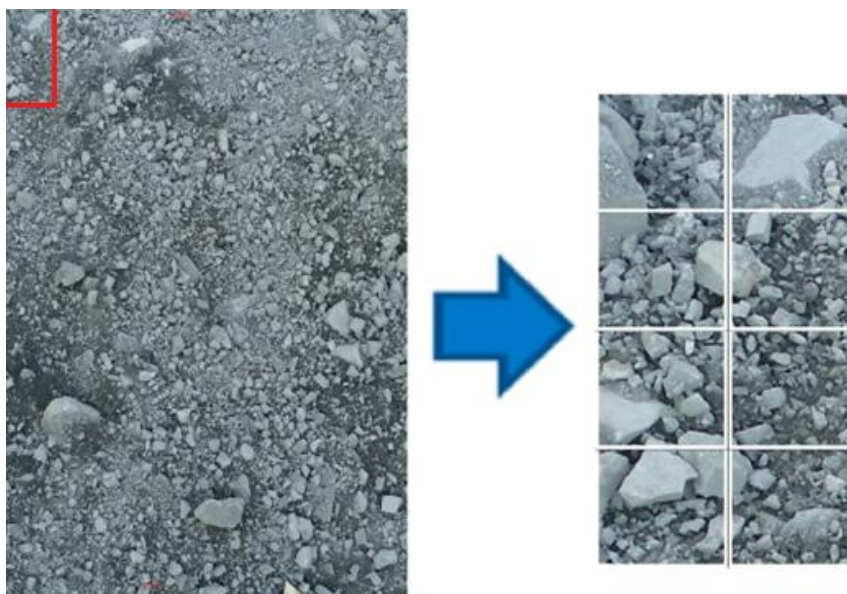


Рис. 1 – Разделение большой фотографии на маленькие изображения

К каждому входному изображению было создано сегментированное изображение, которое имело 3 категории, перекодированных с помощью one hot encoding. Классы характеризовали фон, куски руды и границы кусков руды [6].

Класс № 1 – кусок руды (интересующий нас класс). Класс № 2 – границы кусков руды. Класс № 3 – фон (самый большой класс). Пример маски представлен на рисунке 2.

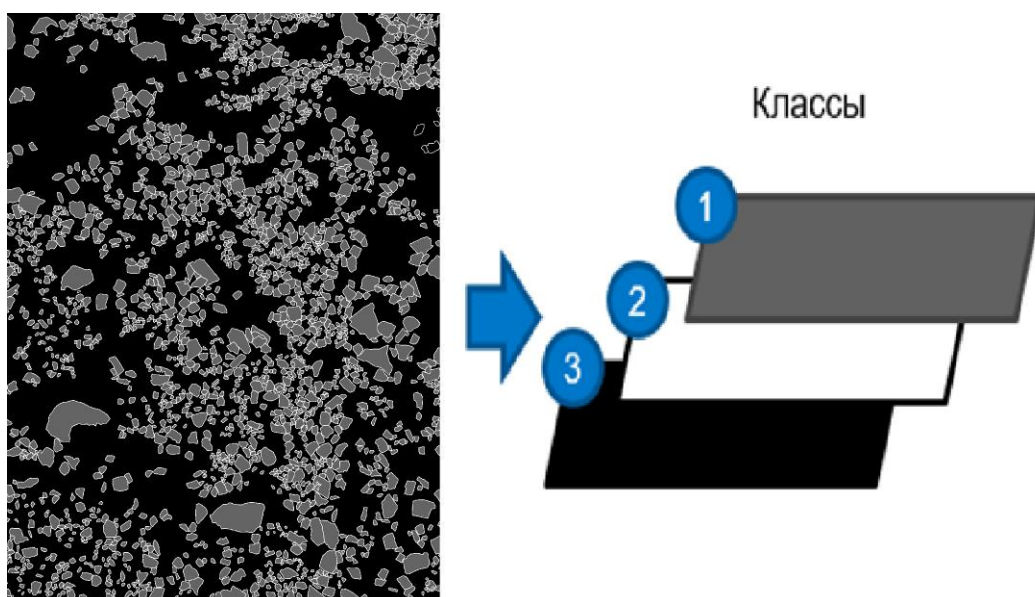


Рис. 2 – Создание сегментированного изображения

Описание предобученного семейства оснований EfficientNet

EfficientNet – это семейство предобученных оснований нейронных сетей, предназначенных для сегментации изображений. Данный вид сетей имеет быструю скорость обработки и высокую точность распознавания, а также меньший размер относительно других моделей, предназначенных для таких же целей.

Сеть EfficientNet настроена на максимальное повышение процента точности детектирования, но одновременно с этим подвергается наказанию, если потребляет много мощностей компьютера в единицу времени на

вычисления, а также штрафуются за долгое время вывода, когда сеть затрачивает большое количество времени на прогнозирование результата.

Нейронные сети характеризуются тремя параметрами [7]:

- глубина – это общее количество слоёв, содержащихся в сети;
- ширина – максимальное количество каналов свертки;
- разрешение – размер картинка, подаваемой нейронной сети на вход.

Семейство EfficientNet содержит 8 сетей, начиная от B0 до B7. Сложность архитектуры и количество параметров увеличивается по возрастанию, начиная от B0 до B7.

EfficientNet B0 – самая маленькая сеть из семейства предобученных оснований. Данная сеть содержит 237 слоев.

Далее все сети масштабируют исходную, начиная с B1 до B7. Все модели количество этапов работы и типы блоков остаются постоянными, в них меняются три параметра сети: разрешение, глубина, ширина сети. С развитием модели параметры только увеличиваются.

Такие модели CNN применяются для идентификации и сегментации изображений. Наиболее часто такие сети применяются при быстро меняющейся информации, которую необходимо обрабатывать.

Структура сети EfficientNet представлена на рисунке 5.

Рисунку 5 соответствуют такие обозначения, как:

- input image – входное изображение, подаваемое для обработки;
- conv 3x3 – слой свертки с размером ядра, равным 3x3;
- MBConv 1, 3x3 - инвертированный остаточный блок размером ядра свертки, равным 3x3, и количеством карт признаков, равным 1;
- MBConv 6, 3x3 - инвертированный остаточный блок размером ядра свертки, равным 3x3, и количеством карт признаков, равным 6;

- MBConv 6, 5x5 - инвертированный остаточный блок размером ядра свертки, равным 5x5, и количеством карт признаков, равным 6;
- output image – итоговое изображение, обработанное нейронной сетью.

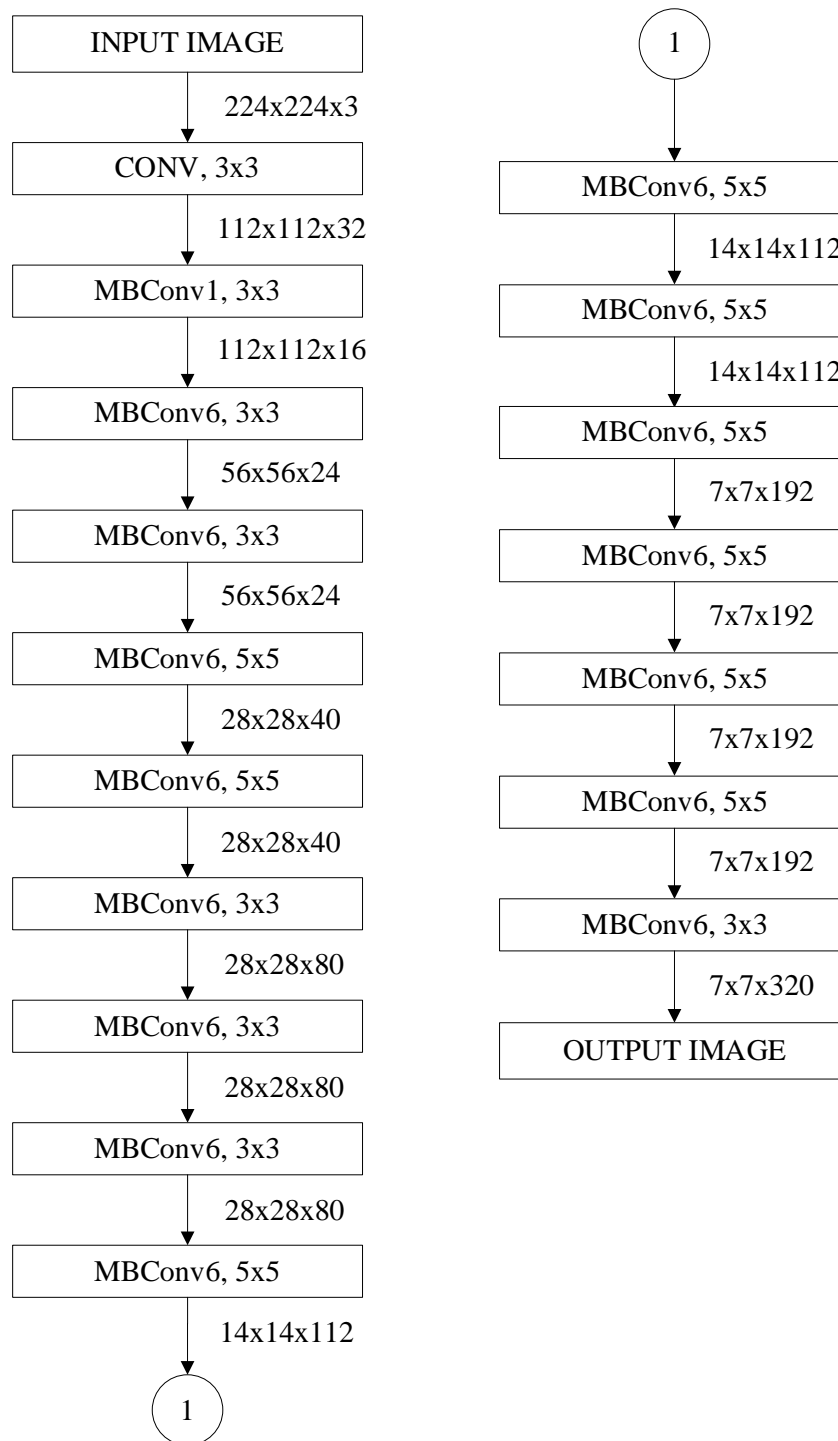


Рис. 5 – Структура сети EfficientNet [8]

Тестирование предобученного семейства оснований EfficientNet

На рисунке 6 представлен график обучения FPN–сети с предобученным основанием EfficientNet B7, остальные графики имеют похожую структуру.

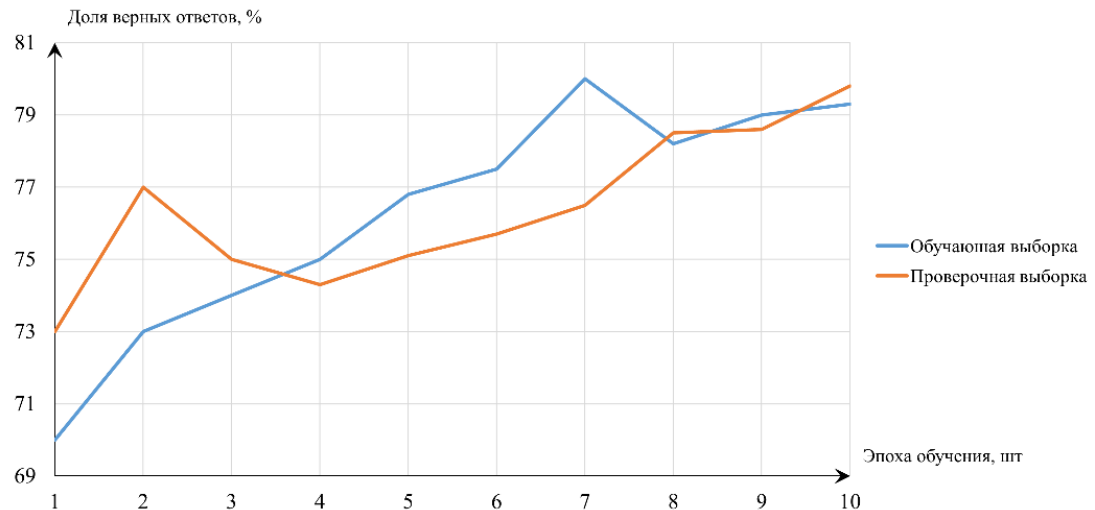


Рис. 6. – График обучения сети на основании EfficientNet B7

Среднее время обучения нейронной сети находится в широком диапазоне. На предобученном основании EfficientNetB0 составляет 50 минут, а на предобученном основании EfficientNetB7 составляет 1 час 40 минут.

Общий процент качества сегментации по трем классам составляет 97,49%. В таблице 1 представлены сведения о числовом показателе IoU для интересующего нас класса сегментации «Куски руды» по всему семейству предобученных оснований EfficientNet по каждой эпохе обучения.

Таблица № 1

Показатель IoU по всему семейству EfficientNet

Кол-во эпох	Efficient Net B0	Efficient Net B1	Efficient Net B2	Efficient Net B3	Efficient Net B4	Efficient Net B5	Efficient Net B6	Efficient Net B7
3	41,5	42,9	46	47,5	50,6	50,9	47,4	42,0
5	40,4	44,8	39,4	52,0	53,6	54,6	53,6	46,2
7	49,1	51,6	55,1	53,8	42,6	54,2	54,6	21,5
10	51,4	45,1	40,6	51,5	53,3	54,1	50,2	42,1

Показатель IoU показывает количество верно распознанных пикселей по отношению к их общему количеству. Следовательно, чем больше значение критерия, тем лучше нейронная сеть сегментировала изображение.

Исходя из значений таблицы 1 видно, что наилучшее значение показала FPN – сеть на предобученном основании EfficientNetB2 при 7 эпохах обучения, её точность составила 55,1%.

После 10 эпох обучения практически для всех оснований начиналось переобучение, поэтому обучение было остановлено.

Описание предобученного семейства оснований SEResNet

Одно из самых популярных семейств предобученных оснований для сегментации изображений является SEResNet. Данное семейство является развитием семейства сетей ResNet [9].

Для понимания, как работает сеть, необходимо разобраться со структурой.

Структура сети SEResNet представлена на рисунке 7.

Рисунку 7 соответствуют такие обозначения, как:

- CONV 2D – слой свертки для двумерных данных, например, изображений;
- BATCH NORMALIZATION – слой нормализации для лучшего обучения НС;
- ReLU – функция активации для слоя;
- GLOBAL POOLING – слой объединения изученных объектов;
- DENSE – полносвязный слой, который применяется для сокращения количества входных параметров до числа классов;
- SIGMOID – функция активации для слоя.

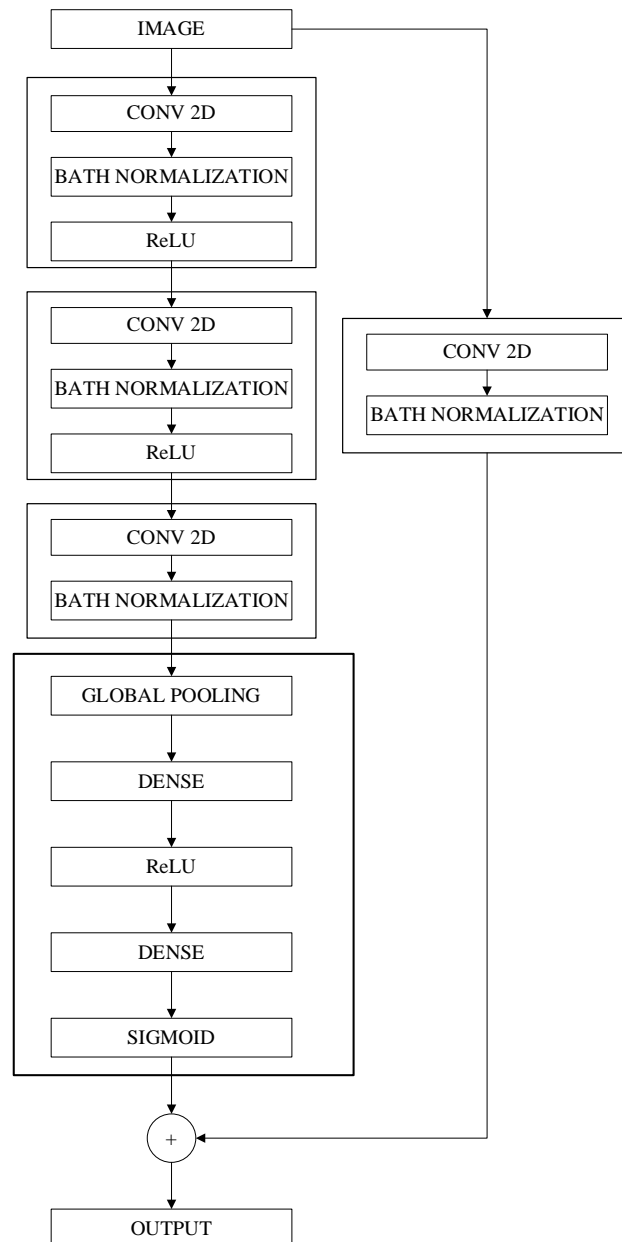


Рис. 7 – Структура SEResNet сети [10]

Тестирование предобученного семейства оснований SEResNet

Среднее время обучения нейронной сети находится в широком диапазоне. На предобученном основании SEResNET18 составляет 30 минут, а на предобученном основании SEResNET152 составляет 1 час 10 минут.

После 10 эпох обучения практически для всех оснований начиналось переобучение, поэтому обучение было остановлено.

Общий процент качества сегментации изображения составляет 98,93%. В таблице 2 представлены сведения о числовом показателе IoU для

интересующего нас класса сегментации «Куски руды» по всему семейству предобученных оснований SEResNet по каждой эпохе обучения.

Таблица 2

Показатель IoU по всему семейству SEResNet

Кол-во эпох	seresnet18	seresnet34	seresnet50	seresnet101	seresnet152
3	43,1	47,5	49,5	37,6	47,3
5	47,6	54,2	40,8	50,4	43,0
7	45,3	52,3	44,1	43,0	45,4
10	52,5	33,5	46,8	48,8	53,7

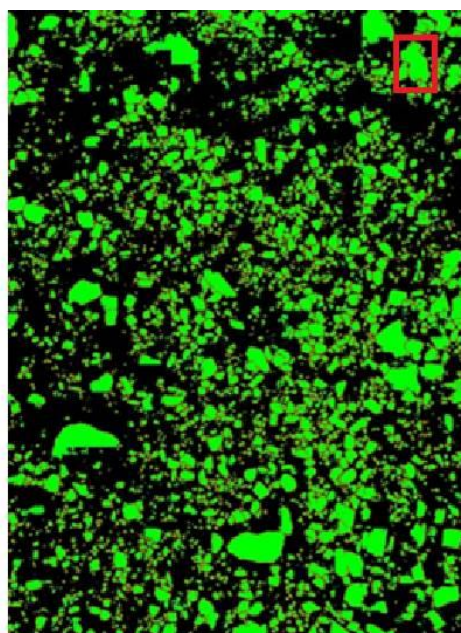
Исходя из значений таблицы 2 видно, что максимальное значение показателя IoU, равное 54,2 процента, соответствует FPN-сети на предобученном основании SEResNet34 после 5 эпох обучения.

Процесс обучения был приостановлен, так как нам необходимо сравнить значения между целыми семействами и внутри них, а многие основания уже начали терять качество сегментации, и это говорит нам о том, что наступило переобучение нейронной сети.

Результаты работы лучшей сети представлено на рисунке 8.



Исходная фотография



Результат сегментации

Рис. 8. – Результат работы лучшей нейронной сети

Заключение

Из результатов эксперимента стало очевидно, что нейронная сеть способна решать поставленную задачу. Для продолжения обучения и тестирования необходимо собрать ещё фотографии для обучающей выборки и продолжить обучение существующих нейронных сетей.

Литература

1. Романов С.А., Стоянов Ю.П. Анализ методов распознавания изображений // КиберЛенинка, 2019, №4. URL: cyberleninka.ru/article/n/analiz-metodov-raspoznavaniya-izobrazheniy/viewer.

2. Омельченко В.П., Тимошенко Е.А. Применение дискриминантного анализа для классификации ЭЭГ больных диабетической энцефалопатией // Инженерный вестник Дона, 2012, №4. URL: ivdon.ru/ru/magazine/archive/n4p1y2012/1135.

3. Разработка алгоритма распознавания фонем русской речи // URL: [Studbooks.net/2042748/informatika/razrabotka_algoritma_raspoznavaniya_fonem_russkoj_rechi](https://studbooks.net/2042748/informatika/razrabotka_algoritma_raspoznavaniya_fonem_russkoj_rechi) (дата обращения: 20.07.2023).

4. Компьютерное распознавание человеческих лиц // Гид по технологиям цифровой трансформации. URL: osp.ru/os/2000/03/177945 (дата обращения: 21.07.2023).

5. Аль-Хулайди А.А. Разработка нового стохастического метода управления очередями заданий с использованием Марковских процессов для параллельных вычислений на кластере // Инженерный вестник Дона, 2011, №1. URL: ivdon.ru/magazine/archive/n1y2011/332.

6. Соловьев Р.А., Тельпухов Д.В., Кустов А.Г. Автоматическая сегментация спутниковых снимков на базе модифицированной сверточной нейронной сети UNET // Инженерный вестник Дона, 2017, №4. URL: ivdon.ru/ru/magazine/archive/n4y2017/4433.



7. ResNets - остаточные блоки и глубокое остаточное обучение // evogeeek.ru URL: evogeeek.ru/articles/175327/ (дата обращения: 01.08.2023).

8. Introduction to EfficientNet // Scaler Topics URL: scaler.com/topics/deep-learning/efficientNet/ (дата обращения: 28.07.2023).

9. Исследуем архитектуры сверточных нейронных сетей с помощью fast.ai // URL: proglib.io/p/issleduem-arhitektury-svertochnyh-neyronnyh-setey-s-pomoshchyu-fast-ai-2020-12-28 (дата обращения 27.07.2023).

10. [Image Category] Hand Tear SENET - Recount SENET (Pytorch) // ProgrammerSought. URL: programmersought.com/article/43528875588/ (дата обращения: 01.08.2023).

References

1. Romanov S.A., Stoyanov Yu.P. KiberLeninka, 2019, №4. URL: cyberleninka.ru/article/n/analiz-metodov-raspoznavaniya-izobrazheniy/viewer.

2. Omel'chenko V. P., E. A. Timoshenko E. A. Inzhenernyj vestnik Dona, 2012, №4. URL: ivdon.ru/ru/magazine/archive/n4p1y2012/1135.

3. Razrabotka algoritma raspoznavaniya fonem russkoy rechi [Development of an algorithm for recognizing phonemes of Russian speech]. URL: studbooks.net/2042748/informatika/razrabotka_algoritma_raspoznavaniya_fonem_russkoy_rechi (date accessed 20.07.2023).

4. Komp'yuternoe raspoznavanie chelovecheskikh lits [Computer recognition of human faces]. Hid po tekhnologiyam tsifrovoy transformatsii. URL: osp.ru/os/2000/03/177945 (date accessed 21.07.2023).

5. Al'-Khulaydi A.A. Inzhenernyj vestnik Dona, 2011, №1. URL: ivdon.ru/magazine/archive/n1y2011/332.

6. Solov'ev R.A., Tel'pukhov D.V, Kustov A.G. Inzhenernyj vestnik Dona, 2017, №4. URL: ivdon.ru/ru/magazine/archive/n4y2017/4433.



7. ResNets - ostatochnye bloki i glubokoe ostatochnoe obuchenie [ResNets - Residual blocks and deep residual learning]. URL: evogeeek.ru/articles/175327/ (date accessed 01.08.2023).

8. Introduction to EfficientNet. Scaler Topics. URL: scaler.com/topics/deep-learning/efficientNet/ (date accessed 28.07.2023).

9. Issleduem arkhitektury svertochnykh neyronnykh setey s pomoshch'yu fast.ai [We investigate the architecture of convolutional neural networks using fast.ai]. URL: proglib.io/p/issleduem-arhitektury-svertochnyh-neyronnyh-setey-s-pomoshchyu-fast-ai-2020-12-28 (date accessed 27.07.2023).

10. [Image Category] Hand Tear SENET - Recount SENET (Pytorch). URL: programmersought.com/article/43528875588/ (date accessed 01.08.2023).