



Б.И. Чертков, О.П. Солдатова

СУПЕРСЕМПЛИРОВАНИЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ УЛУЧШЕННОЙ ОСТАТОЧНОЙ ГЛУБОКОЙ НЕЙРОННОЙ СЕТИ

(Самарский университет)

В последнее время в мире машинного и, в частности, глубокого обучения прослеживается четкая тенденция увеличения спроса на программные продукты, выполняющие суперсемплирование графики. То есть программы, которые, получая на вход изображение небольшого разрешения, тем или иным образом увеличивают его размер и четкость. Данная технология уже активно применяется на практике. Так, суперсемплирование внедряется в игровую индустрию для улучшения качества выводимой на экран картинки. Другой пример использования – модернизация дестояния мирового кинематографа [1]. Для решения этой задачи видеопоток последовательно прогоняется через несколько нейронных сетей, одна из которых увеличивает разрешение каждого отдельного кадра, а другая – производит колоризацию (то есть, раскрашивание) кадра. В результате исходная монохромная кинолента посредственного разрешения преобразовывается в четкий цветной фильм [2]. Стоит отметить, что принципы суперсемплирования видеопотока базируются на идеях суперсемплирования изображений.

Основу остаточных нейронных сетей составляют остаточные блоки. Ключевую роль в этих блоках играют прямые соединения. На рисунке 1 представлена схема такого блока.

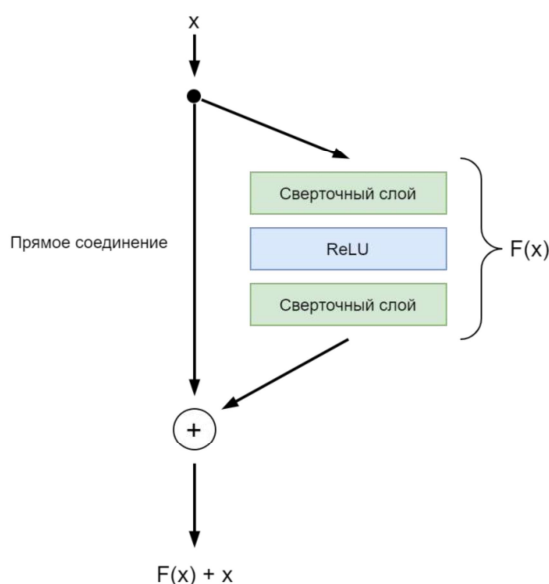


Рисунок 1 – Остаточный блок

Прямое соединение – это принцип при котором входное значение набора слоев суммируется с выходным значением (результатом работы) этого набора слоев [3]. Такой подход позволяет улучшать качество работы сети посредством



увеличения количества слоев. В то время, как качество работы сетей без прямых соединений начинает резко деградировать по достижению определенной глубины. Кроме того, с увеличением количества слоев их становится сложнее оптимизировать. Состав слоев, обозначенный на рисунке 1, как $F(x)$ может меняться от архитектуры к архитектуре. В данном случае для сети, именуемой EDSR он состоит из двух сверточных слоев и функции активации ReLU.

Предлагается два возможных варианта реализации сети: одномасштабная и мультимасштабная. В первом случае на выходе сети получается изображение, увеличенное в некоторое количество раз (чаще всего, в 4). Во втором же случае, имеем серию изображений, каждое из которых также увеличено до определенного размера. Соотношение разрешения выходного изображения к входному будем называть коэффициентом апскейлинга (коэффициентом мультисемплирования). Очевидно, что структуры для одномасштабной и мультимасштабной вариаций модели имеют некоторые различия.

Обучается такая сеть классическим алгоритмом обратного распространения ошибки с применением градиентных методов [4,5]. Для обучения EDSR с коэффициентом апскейлинга 4 в качестве начальных весов используется модель, обученная для мультисемплирования с коэффициентом 2. Такой подход в значительной степени ускоряет процесс обучения, а также улучшает объективное качество результирующей модели в сравнении со случайно инициализированными весами. Рисунок 3 демонстрирует зависимость значения метрики PSNR от количества эпох обучения для различных способов инициализации весов. Преимущество описанного способа очевидна и не нуждается в комментариях.

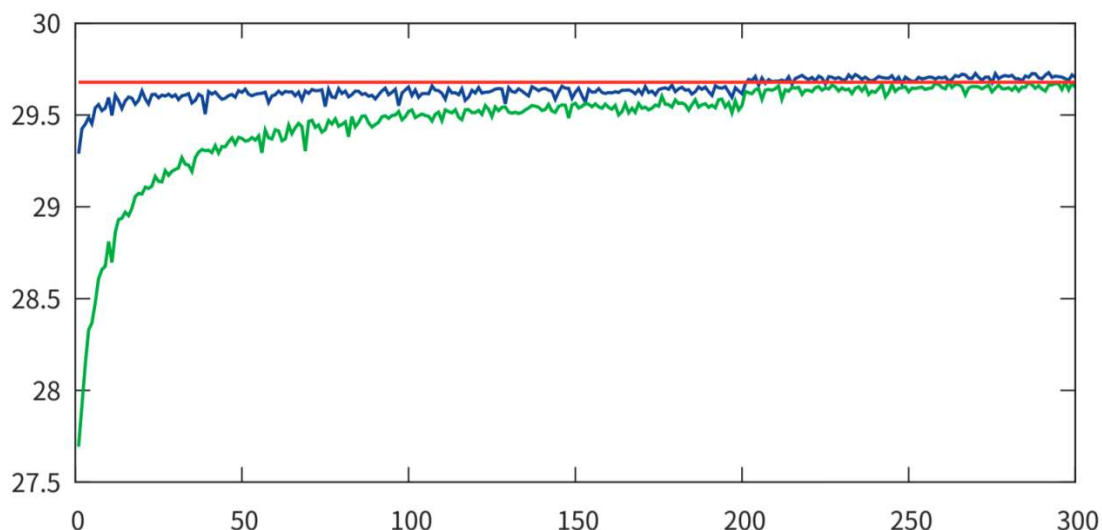


Рисунок 2 – Зависимость PSNR (дБ) от количества эпох. Зеленая линия – случайно инициализированные веса; красная линия – лучшее значение зеленого графика; синяя линия – веса сети, обученной для апскейлинга с коэффициентом 2



Результатом работы стала реализация остаточной нейронной сети, производящей одномасштабное суперсемплирование подаваемых ей на вход изображений с коэффициентом 4. Сеть разработана на языке программирования Python 3.7 с использованием библиотек машинного обучения Tensorflow 2.3 и Keras 2.4. Данный набор технологий уже стал стандартом де-факто в сфере машинного обучения. Он сочетает в себе удобство использования и высокую производительность, обеспеченную возможностями современных графических ускорителей.

Для обучения модели использовался датасет DIV2K, содержащий в общей сумме около тысячи изображений (800 – для обучения и по 100 – на проверку и тестирование) [6].

На рисунке 3 представлено сравнение изображений, полученных на выходе двух различных сетей. Изображение стороннее, то есть взято не из датасета DIV2K. Стоит заметить, что генеративно-состязательная сеть SRGAN (которая также выполняет четырехкратное суперсемплирование) в своей архитектуре тоже использует остаточные блоки [7].

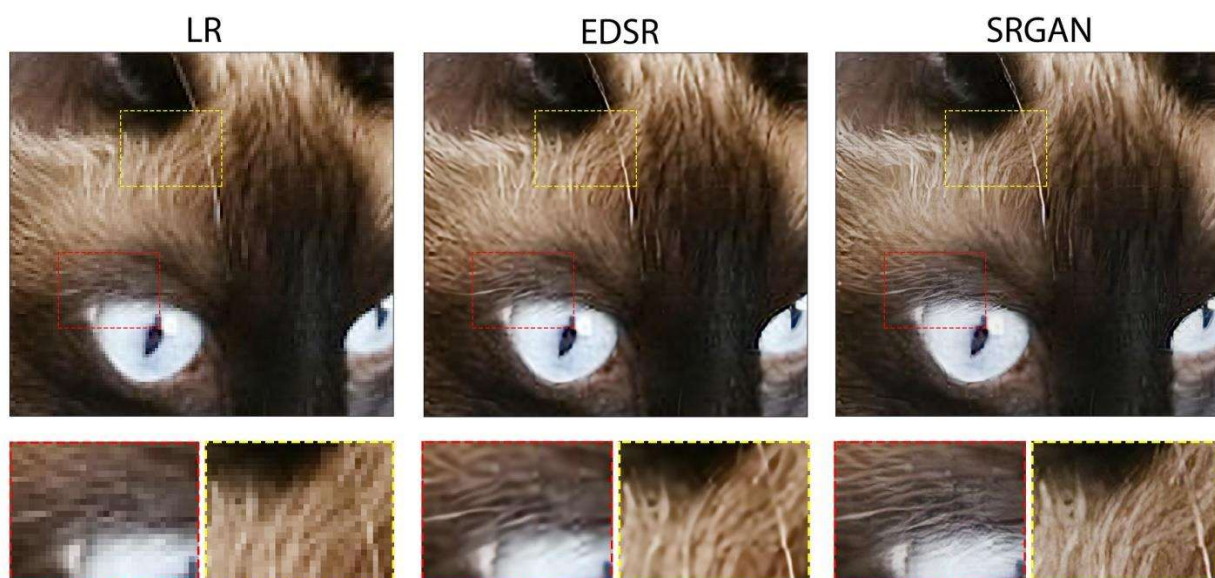


Рисунок 3 – Сравнение выходных изображений двух сетей. Слева – исходное изображение низкого качества; посередине – результат работы улучшенной остаточной сети EDSR; справа – результат работы генеративно-состязательной сети SRGAN

В целом можно заметить, что SRGAN предоставляет лучшее качество суперсемплирования, чем EDSR, однако вместе с тем является более комплексной и вычислительно более затратной, а также чаще приводит к возникновению артефактов. Решение о выборе конкретной модели для практического использования необходимо принимать на основании заданных требований к качеству выполнения поставленной задачи и уровня имеющихся вычислительных мощностей.



Литература

1. Classic 1896 short film gets upscaled to 4K using neural networks [Электронный ресурс]. URL: <https://www.theverge.com/tldr/2020/2/5/21124838/classic-1896-short-film-upscaled-4k-neural-networks-arrival-la-ciotat-station> (дата обращения: 18.04.2021).
2. Using AI to Colorize and Upscale a 109-Year-Old Video of New York City to 4K and 60fps [Электронный ресурс]. URL: <https://petapixel.com/2020/02/24/using-ai-to-colorize-and-upscale-a-109-year-old-video-of-new-york-city-to-4k-and-60fps/> (дата обращения: 18.04.2021).
3. Deep Residual Learning for Image Recognition [Электронный ресурс]. URL: <https://arxiv.org/pdf/1512.03385.pdf> (дата обращения: 18.04.2021).
4. Learning representations by back-propagating errors Recognition [Электронный ресурс]. URL: https://www.iro.umontreal.ca/~vincentp/ift3395/lectures/backprop_old.pdf (дата обращения: 18.04.2021).
5. Gradient descent [Электронный ресурс]. URL: https://en.wikipedia.org/wiki/Gradient_descent (дата обращения: 18.04.2021).
6. DIV2K dataset [Электронный ресурс]. URL: <https://data.vision.ee.ethz.ch/cvl/DIV2K> (дата обращения: 18.04.2021).
7. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [Электронный ресурс]. URL: <https://arxiv.org/pdf/1609.04802.pdf> (дата обращения: 18.04.2021).

С.В. Чурилин

ВЫЯВЛЕНИЕ И ПРЕДОТВРАЩЕНИЕ КОНФЛИКТОВ ИНТЕРЕСОВ ПРИ ВНЕДРЕНИИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ НА ПРЕДПРИЯТИИ

(Акционерное общество «Ракетно-космический центр «Прогресс»)

Интенсивное развитие информационных технологий привело к тому, что наличие современных информационных технологий и систем на всех этапах жизненного цикла (ЖЦ) изделия является необходимым условием конкурентоспособности предприятия.

Необходимо отметить, что, в соответствии со «Стратегией цифровой трансформации ракетно-космической отрасли РФ на период до 2025 г. и перспективу до 2030 г.», одним из ключевых отраслевых направлений является цифровизация производства и ЖЦ изделия.

Цифровизация должна обеспечить переход инженерного, технологического и производственного направления на полноценное использование технологий 3D моделирования и виртуализации при разработке, создании, испытании и эксплуатации изделий и систем, а также предполагает создание единого информационного пространства для конструкторских, технологических, произ-