

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОННОЙ ТЕХНИКИ»

Д Н Е В Н И К П Р А К Т И К И

Обучающийся Ильютченко Павел Сергеевич
(ФИО)

Форма обучения: очная

Кафедра ВМ-1 Группа ВМ-21
(сокращенное название)

Направление подготовки 01.04.04 Прикладная математика

Программа «Цифровая обработка сигналов и изображений»

Вид практики Производственная (Научно-исследовательская работа)

Место практики Кафедра Высшей математики №1
(организация, подразделение)

Начало «1» сентября 2018 г.

Окончание «31» декабря 2018 г.

Документы сопровождения практики в отдел практики и трудоустройства студентов
(ауд.4354) обучающимся предоставлены:

« _____ » _____ Г. _____ (_____)
(подпись сотрудника ОПТС) (ФИО сотрудника ОПТС)

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОННОЙ ТЕХНИКИ»

Д Н Е В Н И К П Р А К Т И К И

Обучающийся Ильютченко Павел Сергеевич
(ФИО)

Форма обучения: очная

Кафедра ВМ-1 Группа ВМ-21
(сокращенное название)

Направление подготовки 01.04.04 Прикладная математика

Программа «Цифровая обработка сигналов и изображений»

Вид практики Производственная (Научно-исследовательская работа)

Место практики Кафедра Высшей математики №1
(организация, подразделение)

Начало «1» сентября 2018 г.

Окончание «15» января 2019 г.

Документы сопровождения практики в отдел практики и трудоустройства студентов
(ауд.4354) обучающимся предоставлены:

« _____ » _____ Г. _____ (_____)
(подпись сотрудника ОПТС) (ФИО сотрудника ОПТС)

ПАМЯТКА

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ПО ПРОХОЖДЕНИЮ ПРАКТИКИ

1.1. До начала практики руководитель от кафедры, проводит инструктаж обучающихся и выдает индивидуальные задания по практике.

1.2. По прибытии на место практики обучающийся должен представить руководителю от организации Дневник практики и ознакомить его с содержанием индивидуальных заданий, пройти инструктаж по технике безопасности и противопожарной профилактике, ознакомиться с рабочим местом, правилами технической эксплуатации оборудования и уточнить план прохождения практики.

1.3. Обучающийся во время практики обязан строго соблюдать правила внутреннего распорядка организации. О временном отсутствии на своем месте необходимо ставить в известность руководителя практики от организации.

1.4. Зачет по практике приравнивается к зачетам по теоретическому обучению и учитывается при подведении итогов общей успеваемости обучающихся.

Результаты практики оцениваются путем проведения промежуточной аттестации с выставлением оценок: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

1.5. Обучающиеся, не прошедшие практику какого либо вида по уважительной причине, проходят практику по индивидуальному плану.

Обучающийся, не прошедший практику какого-либо вида по неуважительной причине или не получивший зачета по итогам ее прохождения, признается имеющим академическую задолженность.

2. ПРАВИЛА ВЕДЕНИЯ ДНЕВНИКА ПО ПРАКТИКЕ

Дневник по практике обучающихся имеет единую форму для всех видов практик и является основным документом для текущего и итогового контроля выполнения заданий.

Обучающемуся необходимо:

2.1. Заполнить титульный лист.

2.2. Получить индивидуальное задание на практику.

2.3. Регулярно вести Табель прохождения практики, записывая все, что сделано за весь день по выполнению индивидуального задания.

2.4. По завершению практики составить отчет в соответствии с индивидуальным заданием. Отчет о практике должен содержать сведения о конкретно выполненной работе, а также краткое описание деятельности, выводы и предложения.

2.5. Получить отзыв руководителя практики от организации.

2.6. В установленный срок, обучающийся должен предоставить на кафедру Дневник практики. При отсутствии правильно заполненного Дневника практика не засчитывается.

ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ НА НАУЧНО-ИССЛЕДОВАТЕЛЬСКУЮ РАБОТУ

Тема практики: Увеличение разрешения изображения с помощью свёрточных нейронных сетей

Задание:

1. Теоретическая разработка метода для решения задачи, поставленной для исследования, и его программная или программно-аппаратная реализация.
2. Проведение экспериментов и тестов с целью определения характеристик и уточнения эмпирических параметров разработанного метода и реализующего его алгоритма.
3. Отладка и программная оптимизация разработанного ПО, оформление программного кода.
4. Анализ результатов применения разработанного метода.
5. Подготовка письменного отчета по проведенному исследованию.

Календарный график выполнения работ

Этапы выполнения технического задания	Форма предоставления результатов	Даты отчетов
Задачи 1,2,3	Устный отчёт	8 неделя
Задача 4	Устный отчёт	12 неделя
Задача 5	Отчёт	18 неделя

Руководитель практики
от кафедры

(подпись)

Козлитин И.А.
(Ф.И.О.)

Студент

(подпись)

Ильютченко П.С.
(Ф.И.О.)

Требования к содержанию и оформлению отчета:

К окончанию практики студент должен подготовить журнал практики, включающий:

- индивидуальное задание на практику, утвержденное руководителем практики;
- табель-календарь прохождения практики, завизированный руководителем

практики.

- отчет прохождения практики с оценкой руководителя практики.

Отчет по учебной практике включает в себя:

- титульный лист;
- введение с перечислением задач, которые решались в рамках выполнения

индивидуального задания по практике;

- основную часть с описанием результатов выполнения индивидуального задания (по части 1 приводится краткое описание предлагаемого метода решения задачи; по части 2 – описание результатов проведенных экспериментов; по части 3 – обзор программных средств, использованных при проведении исследований заданной направленности, по части 4 – выводы по результатам анализа полученных результатов);

- заключение, содержащее описание возможных направлений дальнейшей работы;
- оглавление;
- список использованных источников.

Примерный объем отчета по научно-исследовательской работе – 12-18 страниц.

Шаблон титульного листа - см. ниже

Т А Б Е Л Ь
п р о х о ж д е н и я п р а к т и к и

Дата	Содержание работы	Подпись руководителя практики от организации
28.09	Теоретическая разработка метода для решения задачи, поставленной для исследования, и его программная или программно-аппаратная реализация	
19.10	Проведение экспериментов и тестов с целью определения характеристик и уточнения эмпирических параметров разработанного метода и реализующего его алгоритма	
23.11	Отладка и программная оптимизация разработанного ПО, оформление программного кода	
6.12	Анализ результатов применения разработанного метода	
28.12	Подготовка письменного отчета по проведенному исследованию	

МИНОБРНАУКИ РОССИИ
Федеральное государственное автономное образовательное учреждение высшего
образования
«Национальный исследовательский университет
«Московский институт электронной техники»

ОТЧЕТ ПО НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Направление подготовки — 01.04.04 «Прикладная математика»

Программа — «Цифровая обработка сигналов и изображений»

Выполнил студент Ильютченко Павел Сергеевич
Оценка руководителя практики от кафедры ВМ-1
Козлитин Иван Алексеевич

Группа: ВМ-21

(оценка)

(подпись)

Москва

2018

История

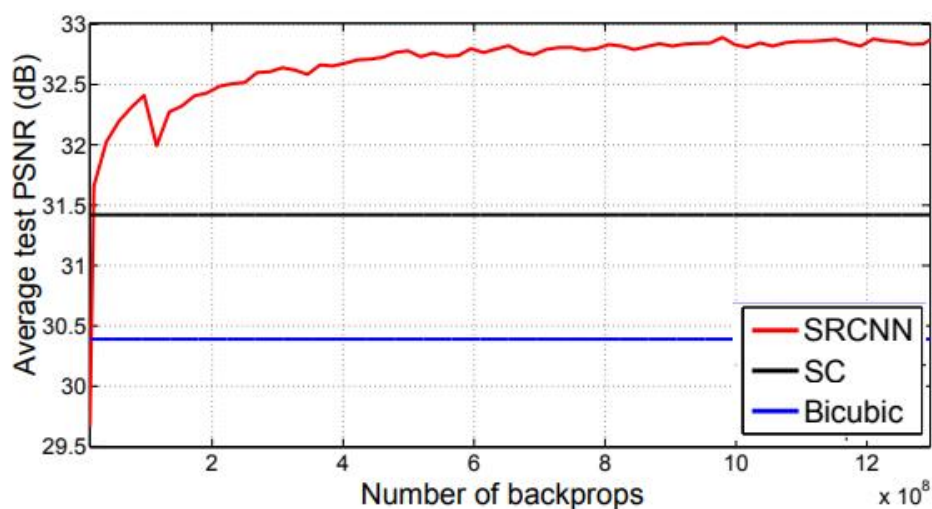
Super-resolution изображения – это восстановленное изображение более высокого разрешения изображения, чем исходное изображение. Проблема конструирования таких изображений одна из классических в компьютерном зрении. Эта проблема изначально некорректна, так как существует множество решений данной задачи. Другими словами, это недоопределенная обратная задача, решение которой не является уникальным. Эта проблема обычно смягчается путем ограничения пространства решения сильной априорной информацией. Последние современные методы в основном используют стратегию, основанную на примерах. Эти методы либо используют внутренние сходства одного и того же изображения, либо изучают функции отображения из низкого разрешения в высокое.

Метод на основе разреженного кодирования является одним из самых репрезентативных методов Super-resolution, основанный на внешних примерах. Этот метод включает несколько шагов. Во-первых, пересекающиеся патчи плотно (с маленьким шагом) вырезаются из исходного изображения и предобрабатываются (например, вычитается среднее и нормализуются). Эти патчи затем кодируются словарём более низкого разрешения. Разреженные коэффициенты передаются в словарь с высоким разрешением для восстановления патчей более высокого разрешения. Перекрывающиеся фрагменты агрегируют (например, взвешенное среднее) для получения окончательного результата. Такая последовательность действий используются большинством методов, основанных на примерах, которые выделяют особое внимание оптимизации словарей или созданию эффективных функций отображения, однако остальные шаги таких подходов редко можно оптимизировать.

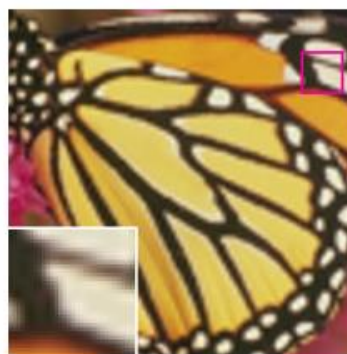
В данной работе хотелось показать, что такой сложный алгоритм из нескольких шагов эквивалентен свёрточной нейронной сети. Принимая это за факт, рассмотрим сеть, которая непосредственно изучает прямое отображение между изображениями с низким и высоким разрешениями. Этот метод принципиально отличается от существующих подходов, основанных на примерах тем, что он не явно изучает словари или многообразие для моделирования пространства патчей. Это неявно достигается через скрытые слои нейронной сети. Кроме того, извлечение и агрегация также сформулированы в этих же слоях, поэтому они и

участвуют в оптимизации. В этом методе полный алгоритм Super-resolution полностью получен с помощью обучения с пред- и постобработкой.

Авторы этого метода назвали его Super-Resolution Convolutional Neural Network (SRCNN). Предлагаемый SRCNN обладает несколькими привлекательными особенностями. Во-первых, его структура специально спроектирована с учетом простоты и обеспечения превосходной точности по сравнению с современными методами, основанными на примерах. На Рисунке 1 показано сравнение с такими методами.



Original / PSNR



Bicubic / 24.04 dB



SC / 25.58 dB



SRCNN / 27.95 dB

Рисунок 1. Сравнение SRCNN с популярными алгоритмами

Во-вторых, при умеренном количестве фильтров и слоёв в сети, этот метод обеспечивает высокую скорость для практического использования онлайн даже на CPU. Он работает быстрее чем алгоритмы, основанные на примерах, так как он полностью прямонаправленный и не нуждается в решении каких-либо оптимизационных проблем. В-третьих, примеры показывают, что восстановление сети может быть улучшено, когда доступны большие и разнообразные датасеты и используется более глубокая модель. Что напротив для других методов, основанных на примерах, может представлять проблемы. Также следует отметить, что нейронная сеть может обрабатывать сразу три цветовых канала изображения для достижения лучших результатов суперразрешения.

В целом, вклад этого подхода можно описать в трех аспектах:

1. Представлена полностью свёрточная нейронная сеть для Super-resolution изображения. Сеть непосредственно обучается отображать изображение низкого разрешения в изображение высокого разрешения с минимальной пред- и постобработкой.
2. Установлена связь между Deep Learning Super-Resolution методом и традиционными методами, основанными на разреженном кодировании. Эта связь обеспечивает руководство к проектированию структуры сети.
3. Продемонстрировано, что глубокое обучение применимо в классической проблеме компьютерного зрения суперразрешения, и показывает отличные результаты качества и скорости.

Описание метода

Рассматривая одно изображение низкого разрешения, во-первых, увеличиваем разрешение до желаемого размера используя бикубическую интерполяцию, это единственная предобработка. Назовём интерполированное изображение как Y . Задача состоит в том, чтобы из изображения Y восстановить изображение $F(Y)$, которое максимально похоже на изображение X с высоким разрешением. Для простоты Y все равно воспринимается как изображение низкого разрешения, хотя его разрешение такое же, как и у X . Метод сводится к обучению отображения F , которое концептуально состоит из трех операций:

1. **Извлечение и представление патчей:** эта операция извлекает пересекающиеся патчи из изображения Y низкого разрешения и представляет каждый патч как вектор высокой размерности. Эти векторы составляют набор карт признаков, число которых равно размерности векторов.
2. **Нелинейное отображение:** эта операция нелинейно отображает каждый многомерный вектор в другой многомерный вектор. Каждый отображенный вектор является представлением патча высокого разрешения. Эти векторы составляют другой набор признаков.
3. **Реконструкция:** эта операция агрегирует вышеприведенные представления патчей высокого разрешения для генерации окончательного изображения высокого разрешения.

Все эти операции образуют одну свёрточную нейронную сеть. Обзор сети изображен на Рисунке 2. Далее более подробно разберем каждую операцию.

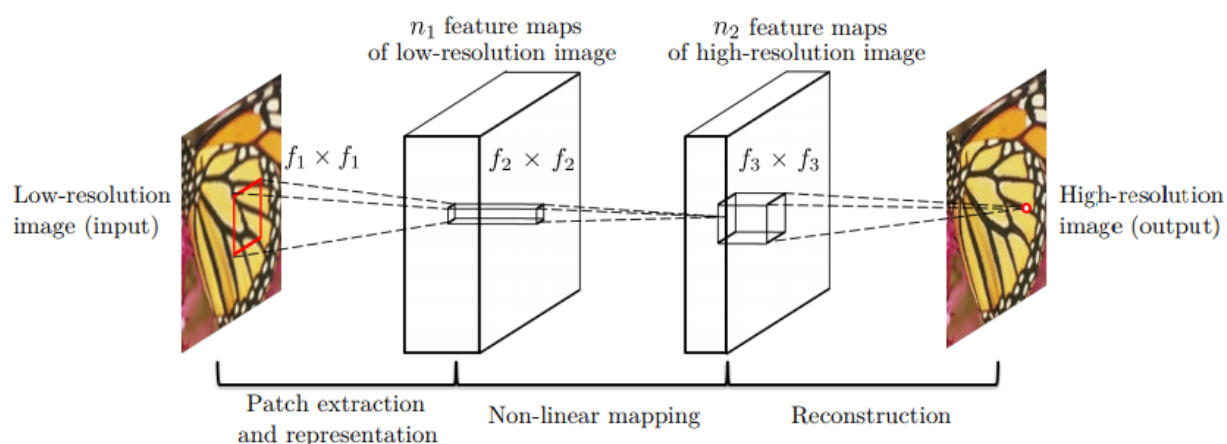


Рисунок 2. Структура SRCNN

Извлечение и представление патчей

Это популярная стратегия в восстановлении изображений состоит в том, чтобы плотно извлекать патчи и затем представлять их с помощью предварительно обученного базиса, как PCA, DCT, Haar и т.д. Это эквивалентно свёртке изображения с множеством фильтров, каждый из которых является базисом. В этой формулировке в оптимизацию сети включается оптимизация этих базисов. Формально этот слой можно описать одной формулой:

$$F_1(\mathbf{Y}) = \max(0, W_1 * \mathbf{Y} + B_1)$$

где W_1 и B_1 представляют собой фильтры и смещения соответственно, а $*$ означает операцию свёртки. Здесь, W_1 соответствует n_1 количеству фильтров размером (c, f_1, f_1) , где c – это количества каналов входного изображения, а f_1 – пространственный размер фильтра. Отсюда понятно, что W_1 применяет n_1 свёрток к изображению, и каждая свёртка имеет ядро (c, f_1, f_1) . Выход состоит из n_1 карты признаков. Смещение B_1 – это вектор размерности n_1 , где каждый элемент связан с фильтром. И к результату свёртки применяется функция активации ReLU (Rectified Linear Unit).

Нелинейное отображение

Первый слой извлёк признаки размерности n_1 для каждого патча. Во второй операции каждый вектор признаков размерности n_1 отображается в вектор признаков размерности n_2 . Это эквивалентно применению n_2 фильтров, которые имеют пространственные размеры 1×1 . Эта интерпретация справедлива только для фильтров размерности 1×1 . Но это можно легко обобщить на фильтры больших размеров, например, 3×3 или 5×5 . В этом случае нелинейное отображение находится не на патче из входного изображения, а на 3×3 или 5×5 патче карты признаков. Операция второго слоя записывается как:

$$F_2(\mathbf{Y}) = \max(0, W_2 * F_1(\mathbf{Y}) + B_2)$$

Здесь W_2 состоит из n_2 фильтров размерности (n_1, f_2, f_2) и смещения B_2 вектора размерности n_2 . Каждый вектор размерности n_2 является представлением патча высокого разрешения, который будет использован для реконструкции. Также можно добавить больше свёрточных слоёв для усиления нелинейности, но это может увеличить сложность модели и понадобится больше времени для обучения.

Реконструкция

В обычных методах полученные перекрывающиеся патчи высокого разрешения часто усредняются для получения окончательного результирующего

изображения. Усреднение может быть рассмотрено как предопределенный фильтр на множестве карт признаков, где каждая позиция “плоской” векторной формы патча высокого разрешения. Основываясь на этом можно определить свёрточный слой для получения конечного изображения высокого разрешения:

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3$$

Здесь W_3 состоит из s фильтров размером (n_2, f_3, f_3) и смещения B_3 размерности s . Если представления патчей высокого разрешения в пространстве изображения (то есть можно просто изменить размер каждого представления для формирования патча) можно ожидать, что фильтр ведет себя как усредняющий. Если представления находятся в других пространствах (например, коэффициенты в терминах некоторого базиса) можно ожидать, что W_3 ведет себя как проектирующие отображение на пространство изображения, а затем усреднение. В любом случае W_3 – это набор линейных фильтров.

Интересно получается, что все три операции мотивированными разными представлениями, однако они приводят к одной и той же форме, что и свёрточные слои. Объединим все три операции вместе в одну свёрточную нейронную сеть (как на Рисунке 2). В этой модели все веса фильтров и смещения должны быть оптимизированны. Несмотря на сжатость всей структуры, такая модель SRCNN тщательно разработана с учетом накопленного большого опыта авторов.

Обучение

Обучение полной отображающей функции F требует оценки параметров сети $\{W_1, W_2, W_3, B_1, B_2, B_3\}$. Этого можно достичь с помощью минимизации ошибки между реконструированным изображением и соответствующему реальному изображению высокого разрешения. Используя созданный набор изображений высокого разрешения и соответствующих им изображений низкого разрешения с среднеквадратичной функцией (MSE) ошибки, задачу минимизации можно записать:

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F(\mathbf{Y}_i; \Theta) - \mathbf{X}_i\|^2$$

где n – это количество образцов в выборке. Использование MSE способствует высокому значению PSNR. PSNR – это широко используемый показатель для качественной оценки восстановленного изображения, и хотя бы частично связан с качеством восприятия. Несмотря на то, что обучение направлено на увеличение PSNR, также наблюдается улучшение и в других метриках, как SSIM, MSSIM и др.

Ошибка минимизируется с помощью стохастического градиентного спуска со стандартным алгоритмом обратного распространения. Веса фильтров изменяются соответственно формулам:

$$\Delta_{i+1} = 0.9 \cdot \Delta_i - \eta \cdot \frac{\partial L}{\partial W_i^\ell}, \quad W_{i+1}^\ell = W_i^\ell + \Delta_{i+1}$$

На этапе обучения из исходных изображений обучающей выборки вырезался случайный кусок размером (f_sub, f_sub, c) изображения с высоким разрешением. Для создания образца низкого разрешения размывается изображение Гауссовым ядром и уменьшается разрешение на порядок, а потом увеличивается на этот же порядок с помощью бикубической интерполяции.

Для избежания краевых эффектов в течение обучения все свёрточные слои не используют отступы (падинги), и сеть производит меньший выход размером $((f_sub - f1 - f2 - f3 + 3)^2, c)$. Функция потерь MSE оценивается только по разнице между центральными пикселями исходного изображения и выходом сети. Хотя обучение происходит на изображениях фиксированного размера, свёрточная сеть может быть применима к изображениям любого размера.

Эксперименты

Для обучения использовались датасеты T91, BSDS200, General100 (<http://vllab.ucmerced.edu/wlai24/LapSRN/>). Из этих датасетов на каждом изображении вырезались куски размером 127×127 – это оригинальные изображения высокого разрешения, потом этот кусок уменьшался до 63×63 , а затем увеличивали разрешение обратно до 127×127 с помощью бикубической интерполяции. Получили датасет состоящий из пар изображений $\{X, Y\}$, где X – изображение, полученное бикубической интерполяцией, а Y – изображение высокого разрешения, которое нужно получить с помощью нейронной сети, то есть получить функцию F ,

которая будет описывать $Y = F(X)$. Из этих датасетов получилось ~ 11 т. пар изображений.

В структуру сети внесены небольшие изменения, на последний слой добавлена функция активации \tanh . А также все изображения предобрабатываются так, чтобы значения каналов лежали в области $[-1, 1]$. А размеры фильтров выбраны из исследования авторов исходной статьи (см. Рисунок 3): $f1=9$, $f2=3$, $f3=5$, $n1=64$, $n2=32$.

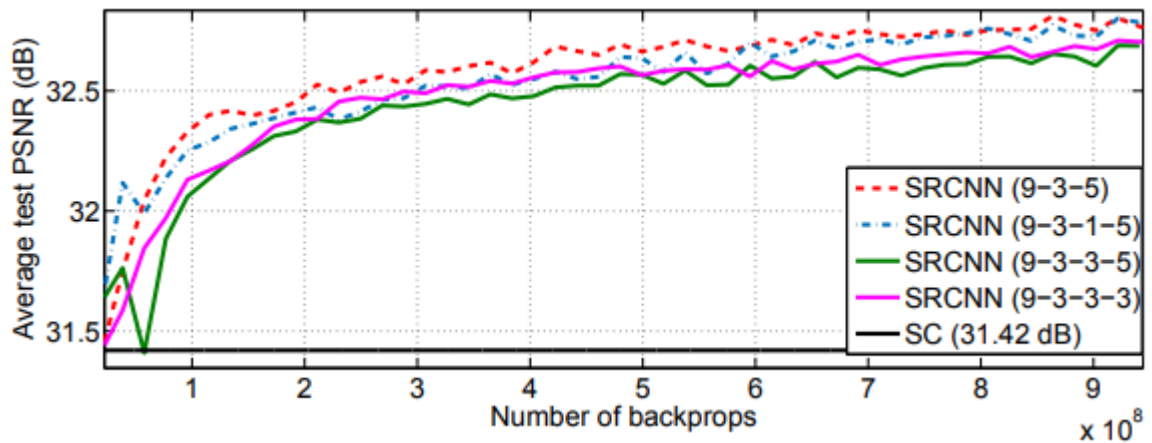


Рисунок 3. Исследование размеров сети

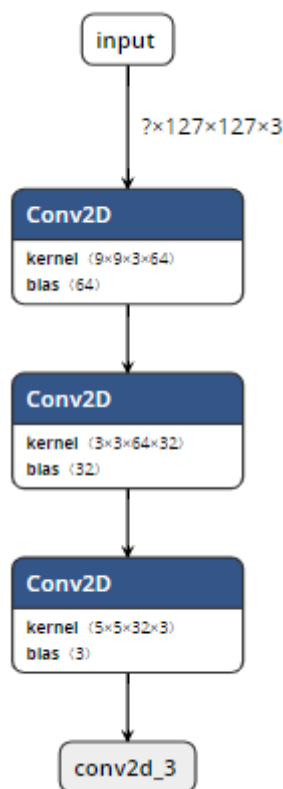


Рисунок 4. Структура сети

Также рассмотрим сравнение авторов SRCNN с основными методами суперразрешения:

- SC – Sparse Coding-based method,
- NE+LLE – Neighbor Embedding + Locally Linear Embedding method,
- ANR – Anchored Neighborhood Regression method,
- A+ - Adjusted Anchored Neighborhood Regression method,
- KK – method described in Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior.

Сравнение производилось на датасете Set5 (5 изображений). Метрикой для сравнения была использована PSNR. Сравнение отображено на Рисунке 4. Своё преимущество SRCNN показывает, когда сеть достаточно обучилась, уже пройдя $3 \cdot 10^8$ шагов обучения, сеть превосходит результаты всех основных методов суперразрешения.

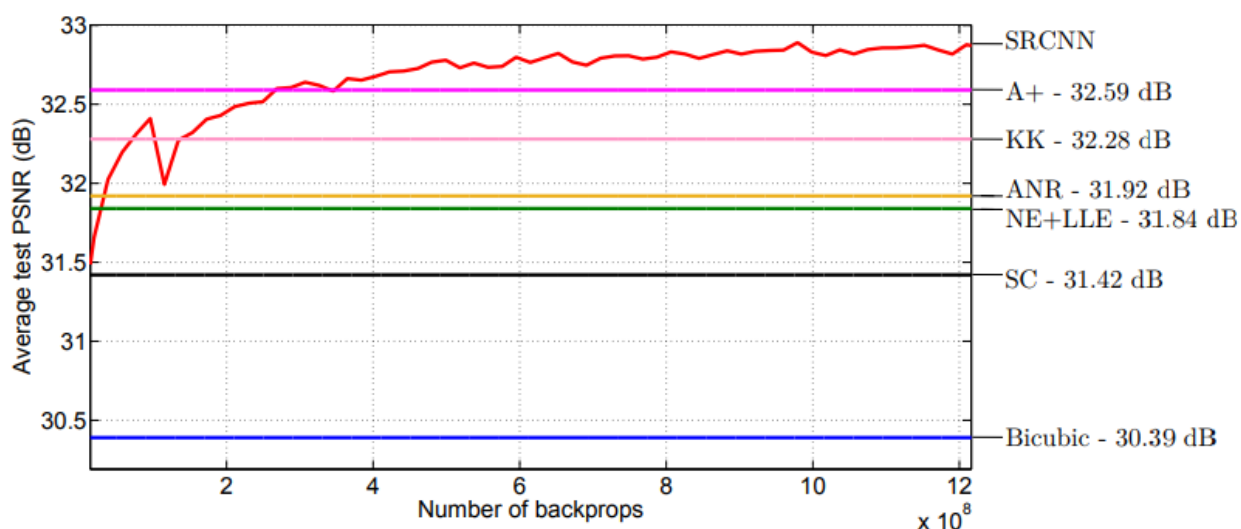


Рисунок 5. Сравнение SRCNN с другими популярными методами SR

Обучение заняло около одного часа на GTX 1050Ti. Размер обучающей выборки 10338 изображений, а тестовой 1149. В конце обучения получились ошибки на обучающем наборе 0.0077 и на тестовом наборе 0.0076 (Рисунок 5). Эти ошибки достаточно близки, что означает хорошие результаты обучения.

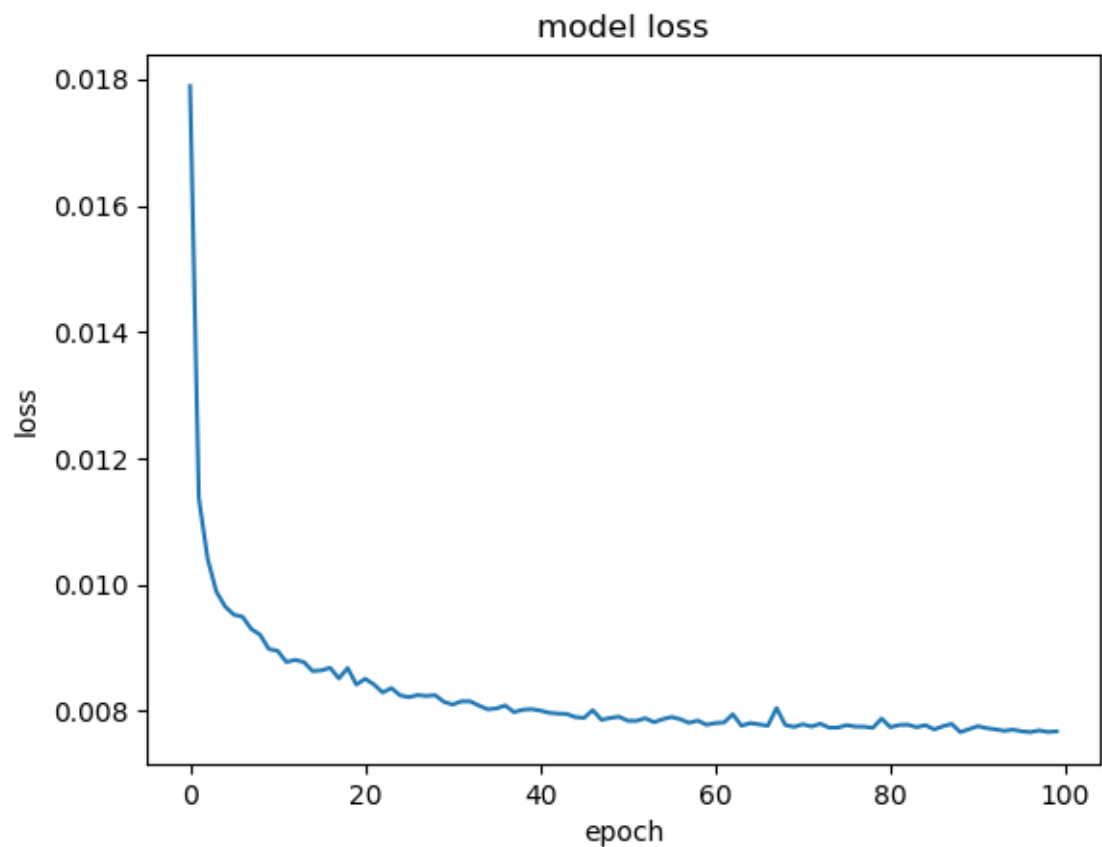














Рисунок 6. Ошибка на обучающем наборе при обучении

Рассмотрим также несколько примеров получившихся картинок:

Исходная картинка высокого разрешения	Бикубическая интерполяция	SRCNN
	 PSNR=20.28 dB	 PSNR=22.55 dB
	 PSNR=30.16 dB	 PSNR=31.27 dB

	 PSNR=24.80 dB	 PSNR=25.81 dB
	 PSNR=20.13 dB	 PSNR=21.31 dB

Заключение

Представлен новый подход глубокого обучения для одиночного суперразрешения. Показано, что традиционные методы могут быть преобразованы в глубокие свёрточные нейронные сети. Такой подход, SRCNN, обучается отображению между изображениями низкого и высокого разрешения с небольшой пред- и постобработкой, кроме оптимизации. Благодаря легкой структуре SRCNN достигает превосходной производительности, по сравнению с другими популярными методами. Можно предположить, что дополнительная производительность может быть получена путём изучения большего количества фильтров и различных стратегий обучения. Кроме того, такая структура, с ее преимуществами простоты и надёжности, может быть применена к другим низкоуровневым проблемам компьютерного зрения, как удаление размытых артефактов или одновременное суперразрешение и шумоподавление. Можно также использовать сеть, чтобы справиться с различными масштабирующими коэффициентами.

Код с архитектурой сети и обучением можно найти в репозитории на GitHub по ссылке <https://github.com/Pol22/SuperResolutionCNN>

Источники

1. Image Super-Resolution Using Deep Convolutional Networks
[<https://arxiv.org/pdf/1501.00092.pdf>]
2. Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks [<http://vllab.ucmerced.edu/wlai24/LapSRN/>]
3. Keras: The Python Deep Learning library [<https://keras.io/>]
4. Single-Image Super-Resolution: A Benchmark
[<https://pdfs.semanticscholar.org/a286/af401232dcf181af6790873d92585a85f370.pdf>]
5. Super-resolution imaging [https://en.wikipedia.org/wiki/Super-resolution_imaging]
6. Review: SRCNN (Super Resolution) [<https://medium.com/coinmonks/review-srcnn-super-resolution-3cb3a4f67a7c>]
7. Network In Network [<https://arxiv.org/pdf/1312.4400.pdf>]