ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ

ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОННОЙ ТЕХНИКИ»

**ДНЕВНИК ПРАКТИКИ**

**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 ВМ-1\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

(организация, подразделение)

|  |  |
| --- | --- |
| **Начало** | **«11» февраля 2019 г.** |
| **Окончание** | **«9» апреля 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.

1. **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 проделано 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** **Фотореалистичное суперразрешение изображения с помощью генеративно-состязательной сети**

**Задание:**

1. Теоретическая разработка метода для решения задачи, поставленной для исследования, и его программная реализация.
2. Проведение экспериментов и тестов с целью определения характеристик и уточнения параметров разработанного метода и алгоритма.
3. Отладка и 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 |
| Задача 1 | Устная | 5 неделя |
| Задача 2 | Устная | 7 неделя |
| Задачи 3, 4, 5 | Отчет | 9 неделя |

|  |  |  |
| --- | --- | --- |
| Руководитель практики  от кафедры |  |  |
|  | (подпись) | (Ф.И.О.) |
| Студент |  |  |
|  | (подпись) | (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

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**ТАБЕЛЬ**

**прохождения практики**

|  |  |  |
| --- | --- | --- |
| Дата | Содержание работы | Подпись руководителя практики от организации |
|
| 28.02 | Теоретическая разработка метода для решения задачи, поставленной для исследования, и его программная реализация |  |
| 8.03 | Проведение экспериментов и тестов с целью определения характеристик и уточнения параметров разработанного метода и алгоритма |  |
| 25.03 | Отладка и оптимизация разработанного ПО, оформление программного кода. Анализ результатов применения разработанного метода |  |
| 8.04 | 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 | | Группа: ВМ-21 |
| Оценка руководителя практики от кафедры ВМ-1  Козлитин Иван Алексеевич |  |  |
|  | (оценка) | (подпись) |

Москва

2019

# **1. Введение**

Очень сложная задача получения изображения с высоким разрешением (HR) из его дубликата с низким разрешением (LR), это называется суперразрешением (SR). SR получил значительное внимание со стороны сообщества исследователей компьютерного зрения и имеет широкий спектр применений.

Неправильно спозиционированная природа недоопределенной проблемы SR особенно выражена для больших коэффициентов масштабирования, для которых детали текстуры в реконструированных SR изображениях обычно отсутствуют. Цель оптимизации SR алгоритмов с учителем обычно сводится к минимизации среднеквадратичной ошибки между восстановленным HR изображением и истинным HR изображением. Это удобно, так как минимизация MSE также максимизирует пиковое отношение сигнал-шум (PSNR), которое является стандарной мерой для оценки и сравнения алгоритмов SR. Однако способность MSE (и PSNR) к улавливанию воспринимаемых различий, как высокодетализированные текстуры, очень ограничена, так как они определены на попиксельной разнице. Это показано на Рисунке 1, где самый высокий PSNR не обязательно отображает лучший результат SR по восприятию. Разница восприятия между SR изображением и оригинальным показывает, что восстановленное изображение не является фотореалистичным, как определено в работе Ferwerda ( http://cin.ufpe.br/~in1123/material/vor\_hvei03\_v20.pdf Фотореалистичное изображение производит тот же визуальный эффект как и реальная сцена ).

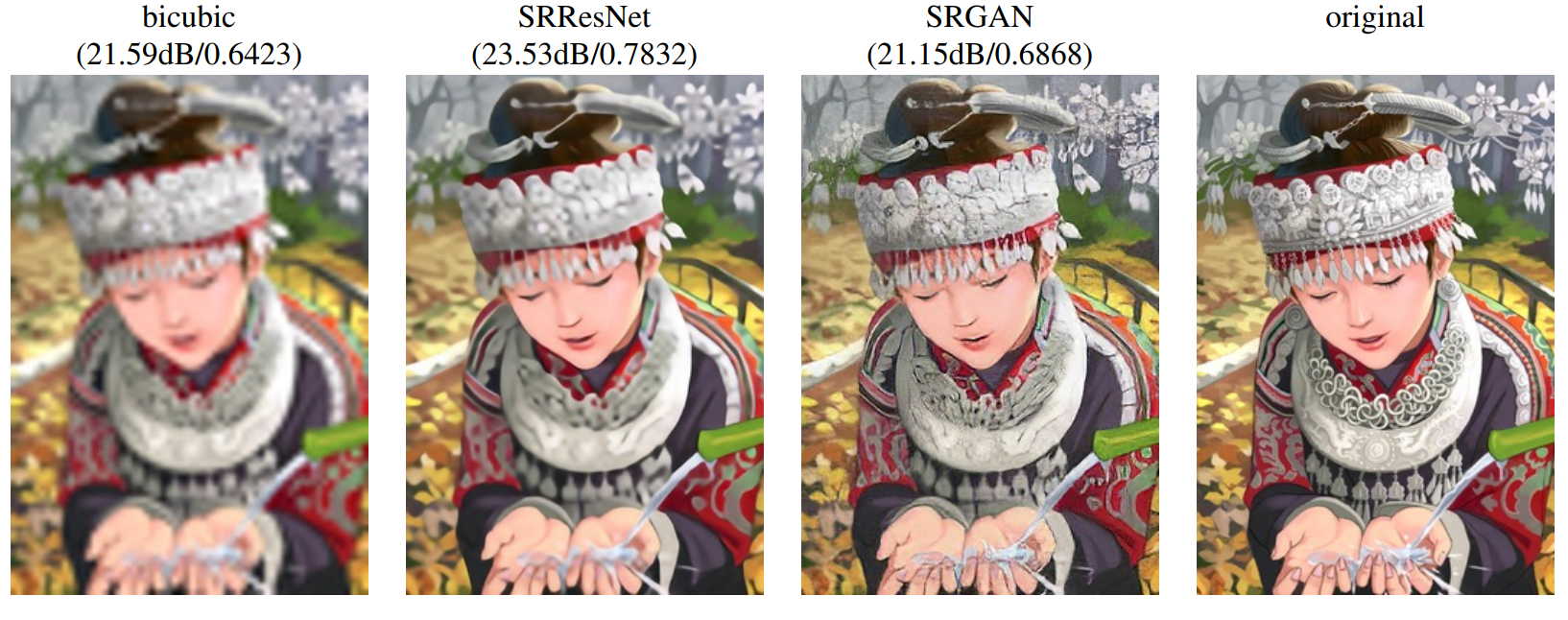


Рисунок 1. Слева направо: бикубическая интерполяция, глубокая остаточная сеть оптимизированная с МНК, глубокая остаточная генеративно-состязательная сеть оптимизированная для потери более чувствительной с человеческому восприятию, оригинальное HR изображение. Соответствующие PSNR и SSIM показаны в скобках. [4x масштабирование]

В этой работе мы предлагаем генеративно состязательную сеть суперразрешения (SRGAN), в которой мы используем глубокую свёрточную сеть с residual связями (ResNet), c пропущенными соединениями и с целью оптимизации отличающейся от MSE. В отличие от предыдущих работ, мы определяем новую потерю восприятия, использующую карты особенностей высокого уровня сети VGG в сочетании с дискриминатором, который поощряет результаты, которые трудно отличить от эталонных изображений HR. Пример фотореалистичного изображения, которое было получено масштабированием в 4 раза показано на Рисунке 2.

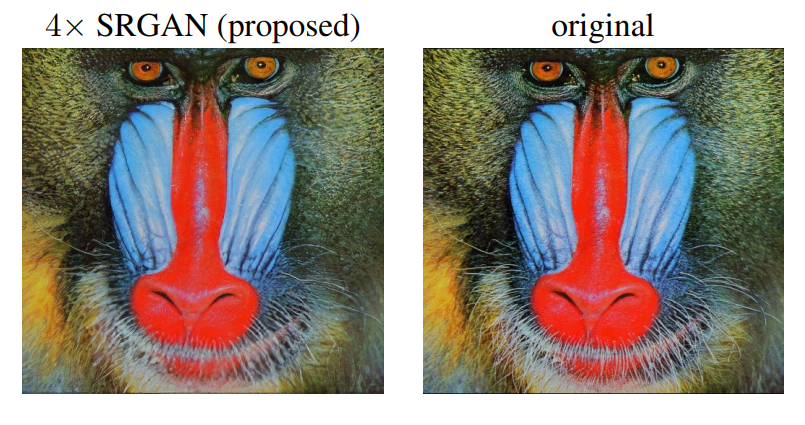
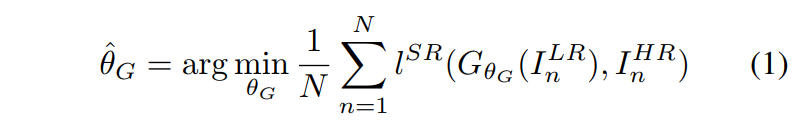


Рисунок 2. Суперразрешение изображения (слева) почти неотличимо от оригинального (справа). [4x мастабирование]

# **2. Метод**

Сконцентрируемся на суперразрешении одного изображения (SISR) и не будем затрагивать темы, покрывающие HR изображения по нескольким изображениям. В SISR целью является оценка высокого разрешения, суперразрешимое изображение из изображения низкого разрешения . Здесь дубликат с низким разрешением версии . Изображения высокого разрешения доступны только во время обучения. В течение обучения, получаются применением фильтра Гаусса перед последующим применением операции понижения дискретизации с коэффициентом . Для изображения с цветовыми каналами, мы опишем как вещественный тензор размером и , как соответственно.

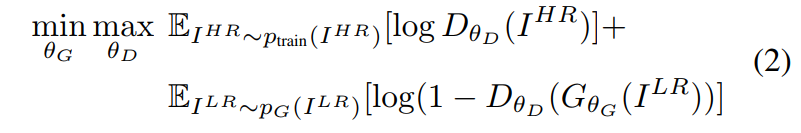
Наша конечная цель - это обучить генерирующую функцию G, которая создает по полученному LR входному изображению его соответствующий HR аналог. Для достижения этого, мы обучаем генеративную сеть как прямонаправленную свёрточную нейронную сеть , зависящую от параметров . Здесь обозначает весовые коэффициенты и смещения сети L-го уровня глубогой сети и получаются они путём оптимизации специальной для SR функции потерь . С тренировачными изображениями с соответствующими , мы решаем задачу оптимизации:



В этой работе мы специально спроектируем функцию потери восприятия как взвешенную комбинацию из нескольких компонент, которые моделируют различные желаемые характеристики восстановленного SR изображения.

**2.1. Архитектура состязательной сети**

Следуя определению Goodfellow мы также определим сеть дискриминатор , которую мы оптимизируем поочередно наряду с для решения состязательной мин-макс проблемы:



Основная идея заключенная в этой формуле - это позволить тренироваться генеративной модели G с основной целью обмануть дискриминатор D, который тренируется распознавать суперразрешенные изображения от реальных изображений. С этим подходом наш генератор может научиться создавать решения, которые очень похожи на реальные изображения и поэтому их сложно классифицировать дискриминатору D. Это поощряет решения, которые лучше воспринимаются визуально (перцептивно), находящиеся в подпространстве, многообразии естественных изображений. Это контрастирует с SR решениями прибегающими к минимизации измерений попиксельной ошибки, как MSE.

В центре нашей очень глубокой генеративной сети G, которая отображена на Рисунке 3, в которой B количество residual блоков с идентичным строением. В особенности, мы используем два свёрточных слоя с маленьким ядром 3х3 и с 64-мя картами особенностей, отправляющиеся затем в batch-normalization слои и ParametricReLU, как функцию активации. Мы увеличиваем разрешение входного изображения с помощью двух обученных субпиксельных свёрточных слоёв, предложенных Shi et al. ( https://arxiv.org/pdf/1609.05158.pdf )

Чтобы определить настоящие HR изображения от сгенерированных SR образцов, мы обучаем сеть дискриминатор. Архитектура показана на Рисунке 3. Мы следуем архитектурным рекомендациям резюмированным Radford et al. ( https://arxiv.org/pdf/1511.06434.pdf ), и используем LeakyReLU активацию (alpha = 0.2), и избегаем максимального пулинга во всей сети. Сеть дискриминатор обучена решать задачу максимизации в Выражении 2. Она содержит 8 свёрточных слоёв с увеличивающимся размером ядер от 3х3 с фактором 2, от 64 до 512 фильтров в ядре, как в сети VGG. В свёртке используются strides для уменьшения разрешения изображения, каждый раз количество особенностей удваивается. В результате 512 карт особенностей проходят через два полносвязных слоя и в конце сигмоидная функция активации для получения вероятностей классификации.

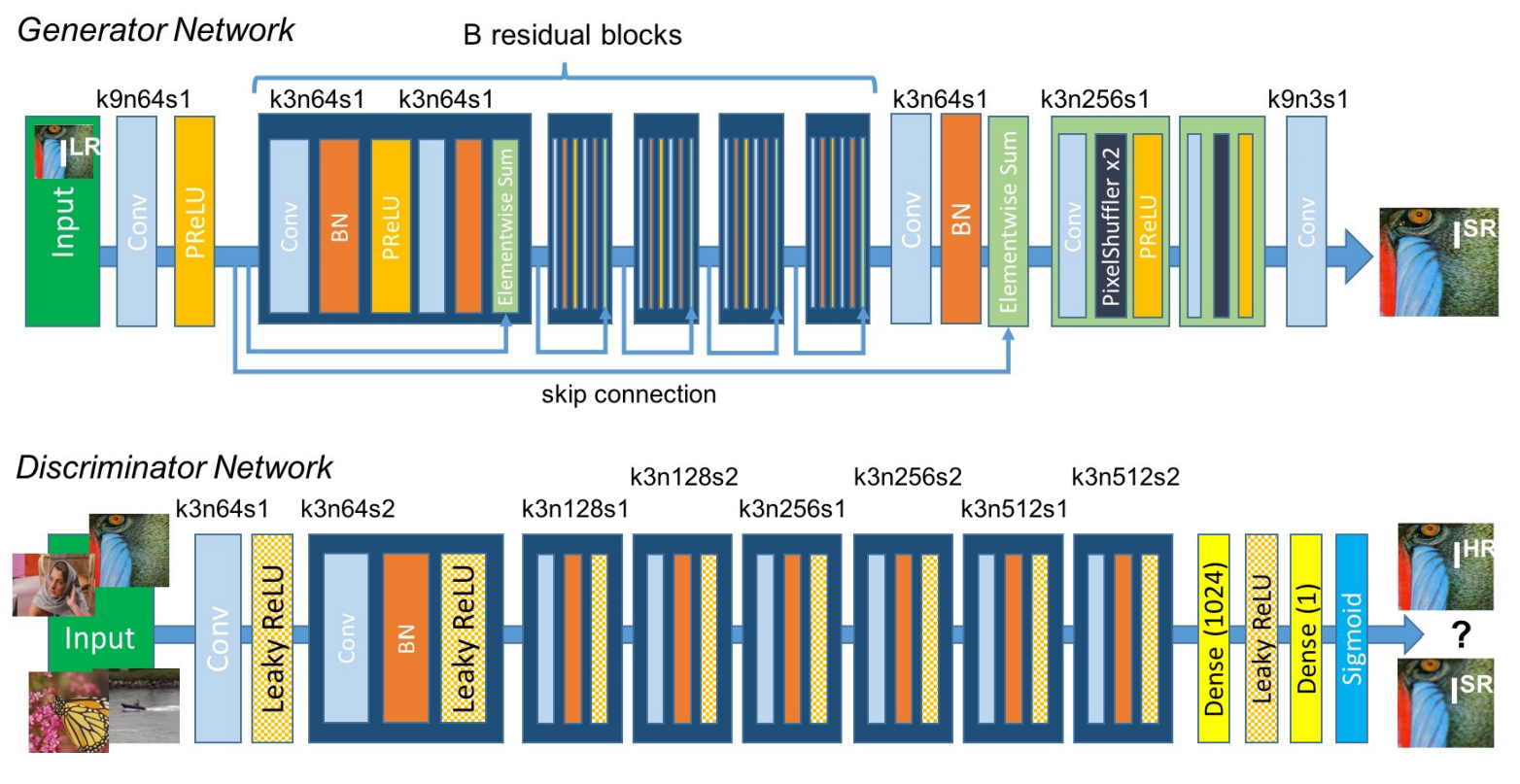
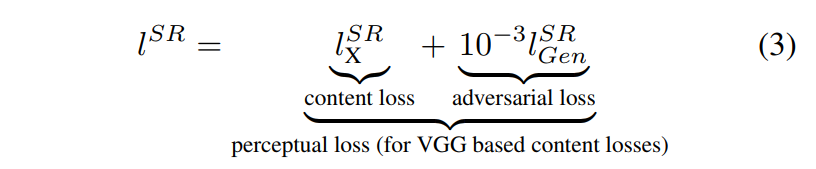


Рисунок 3. Архитектура Генерирующей и Дискриминирующей сети с соответствующими размерами ядер (k), количествами карт особенностей (n) и шагами (s), которые показаны для каждого свёрточного слоя.

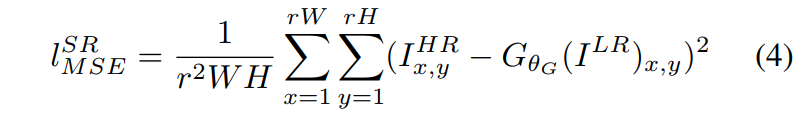
**2.2 Функция потери восприятия**

Предназначение нашей функции потери восприятия имеет решающее значение для производительности нашей генерирующей сети. Пока обычно смоделированно на основе MSE, мы улучшаем ее с помощью предложений из Johnson et al. ( https://arxiv.org/pdf/1603.08155.pdf ) и Bruna et al. ( https://arxiv.org/pdf/1511.05666.pdf ) и конструируем функцию потерь, которая оценивает решение с точки зрения воспринимательных характеристик. Мы сформулируем потери восприятия, как взвешенная сумма потери содержания и потери состязательной компоненты , как:

Далее мы опишем возможные реализации для потери содержания и состязательной потери.

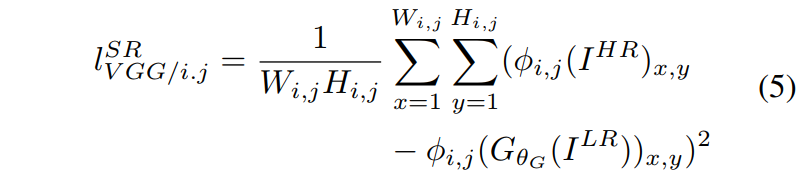
**2.2.1 Потеря содержания**

Попиксельная потеря MSE записывается как:



Это наиболее часто использующаяся задача оптимизации для SR изображения во многих современных подходах. Онако, при достижении особенно высокого PSNR, решениям на основе MSE оптимизации часто не хватает высокочастотных компонент, и это приводит к неудовлетворительному решению с точки зрения восприятия, оно содержит черезмерно гладкие текстуры (см. Рисунок 1).

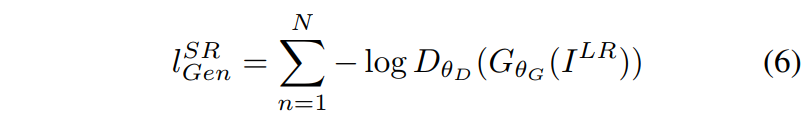
За место того чтобы использовать на попиксельные потери мы будем опираться на идеи Gatys et al. (Style-transfer), Bruna et al. и Johnson et al. и использовать функцию потерь, которая ближе к сходству восприямия. Мы определим VGG потерю основанную на слоях активации ReLU предобученной 19-ти слойной VGG модели описанной Simonyan и Zisserman. Под мы будем полагать карту особенностей, полученную на j-ом слое свёртки (после активации) перед i-ым слоем максимального пулинга в сети VGG19, которые мы считаем данными. Определим VGG потерю как евклидово расстояние между представлениями особенностей восстановленного изображения и опорноного изображения :



Здесь и описывают размерности соответствующих карт характеристик в сети VGG.

**2.2.2 Состязательная потеря**

В добавок к потерям содержания, описанным выше, мы также добавляем генеративную компоненту в наш GAN для потери восприятия. Это побуждает нашу сеть отдавать предпочтение решениям, которые лежат во множестве естественных изображений, чтобы обмануть дискриминирующую сеть. Генеративные потери определены на вероятностях дискриминатора по всем обучающим образцам:



Здесь это вероятность, что восстановленное изображение это естественное HR изображение. Для лучшего градиентного подхода мы минимизируем , заместо .

**3. Эксперименты**

**3.1. Данные и измерения сходства**

Мы производили эксперименты на трех широко используемых эталонных наборах данных Set5, Set14 и BSD100, набор тестирования BSD300. Все эксперименты производились с коэффициентом масштабирования 4х между низко- и высоко-разрешенными изображениями. Это соответствует 16х сокращению количества пикселей на изображении. Для честного сравнения, все сообщенные PSNR и SSIM измерения были высчитаны на y-канале центральной области, удалены поломы шириной 4 пикселя с каждой границы изображений, используя daala пакет ( https://github.com/xiph/daala (commit: 8d03668) ). Суперразрешенные изображения для сравнения методова включали алгоритмы ближайшего соседа, бикубический, SRCNN и SelfExSR? , были взяты из онлайн материалов, дополняющих Huang et al. ( https://github.com/jbhuang0604/SelfExSR ) и для DRCN от Kim et al. ( http://cv.snu.ac.kr/research/DRCN/ ). Результаты, полученные с помощью SRResNet (для потерь: и ) и варианты SRGAN доступны онлайн ( https://twitter.box.com/s/lcue6vlrd01ljkdtdkhmfvk7vtjhetog ). Статические тесты были произведены в виде парных двусторонних тестов Вилкоксона со знаком ранга, и значимость определялась при .

Читателя может также заинтересовать независимая разработка GAN-основанных решений на GitHub ( https://github.com/david-gpu/srez ). Однако это только экспериментальные результаты на ограниченном наборе лиц, что является более ограниченной и легкой задачей.

**3.2. Детали обучения и параметры**

Мы обучали все сети на NVIDIA Tesla M40 GPU, используя случайные образцы из 350 тысяч изображений из ImageNet датасета. Эти изображения отличаются от тестовых изображений. Мы получили LR изображения путем уменьшения разрешения HR изображений (BGR, ), используя бикубическое ядро с масштабом . Для каждого минибатча мы вырезаем 16 случайных HR подизображений из различных тренировачных изображений. Заметим, что мы можем применять генерирующую модель к изображениям любого размера, так как она полностью свёрточная. Мы отобразили пространство LR выходных изображений к отрезку и для HR изображений к . Потеря MSE были расчитана на изображениях с яркостью в диапазоне . Карты особенностей VGG были также масштабированны с коэффициентом , чтобы получить VGG потери сравнимые по масштабу с MSE потерями. Это эквивалентно умножению Выражению 5 с масштабирующим коэффициентом . Для оптимизации используется Adam с . SRResNet сети были обучены с шагом обучения и проведено итераций. Мы используем обученную, основанную на MSE, SRResNet сеть как первоначальный вариант для генератора, при обучении действительного GAN, чтобы избежать нежелательных локальных минимумов. Все варианты SRGAN были обучены в итераций с шагом обучения и другие в итераций с меньшим шагом . Мы чередуем обновление сети генератора и дискриминатора, что эквивалентно k = 1, как используется в алгоритме Goodfellow et al. Наша сеть генератор имеет 16 идентичных (B = 16) residual блоков. В течение времени тестов мы отключаем обновление Batch Normalization, чтобы получить выход, который детерминированно зависит только от входа.

**3.3. Тестирование средней оценки мнения (Mean opinion score = MOS)**

Мы выполнили MOS тест для количественной оценки способности различных подходов восстанавливать визуальное восприятие изображений. В частности, мы опросили 26 оценщиков присвоить интегральный балл от 1 (плохое качество) до 5 (отличное качество) для суперразрешенных изображений. Оценщики оценили 12 версий каждого изображения на датасетах Set5, Set14 и BSD100: метод ближайшего соседа, букубическая интерполяция, SRCNN, SelfExSR, DRCN, ESPCN, SRResNet-MSE, SRResNet-VGG22\* (\*не оценено на BSD100), SRGAN-MSR\*, SRGAN-VGG22\*, SRGAN-VGG54 и оригинальное HR изображение. Каждый оценщик оценил 1128 изображений, которые были представлены в случайном порядке. Оценщики были откалиброваны на NN (оценка 1) и HR (оценка 5) версиях 20 изображений из датасета BSD300. В пробном исследовании мы оценили процедуру калибровки и надежность повторного тестирования 26 оценщиков на подмножестве из 10 изображений из BSD100, добавив изображения из различных методов дважды в большой тестовый набор. Мы нашли хорошую надежность и никаких существенных различий между оценками идентичных изображений. Оценщики очень последовательно оценивали NN-интерполированные тестовые изображения как 1 и оригинальные HR изображения как 5 (см. Рисунок 4).

Экспериментальные результаты проведенных MOS-испытаний описаны в Таблице 1, Таблице 2 и на Рисунке 5.

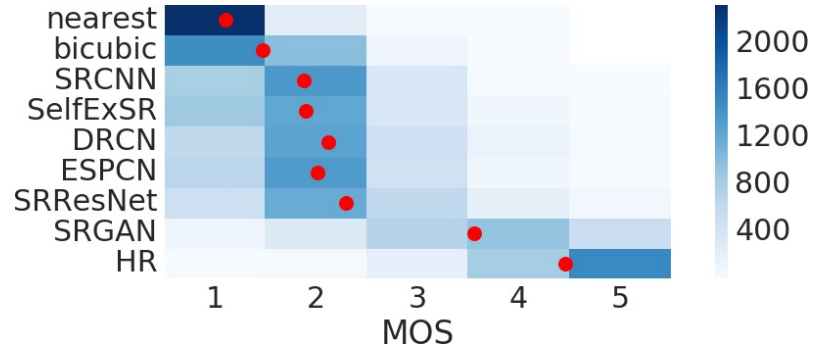


Рисунок 4. Распределение MOS баллов закодированное в цвете на BSD100. Для каждого метода 2600 образцов было оценено (100 изображений x 26 оценщиков). Среднее значение отображается красным маркером, где ячейки центрированы около значения i. [4x масштабирование]

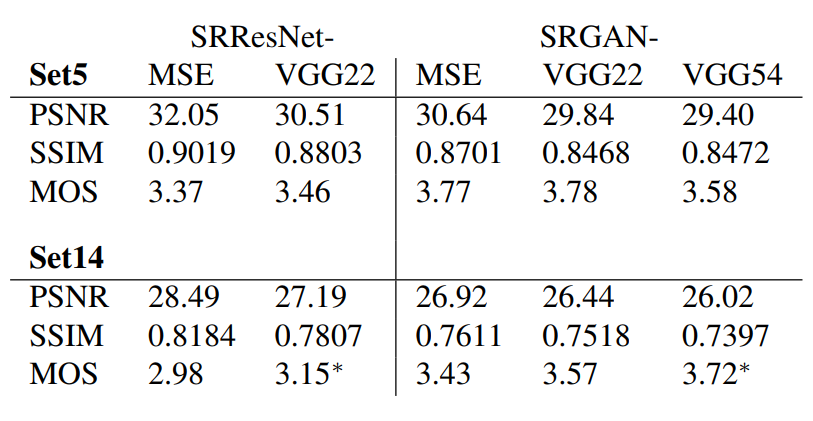


Таблица 1. Результаты различных функций потерь для SRResNet и состязательных сетей на Set5 и Set14 датасетах. MOS значительно выше (p<0.05) чем с другими потерями в этой категории [4x мастабирование]

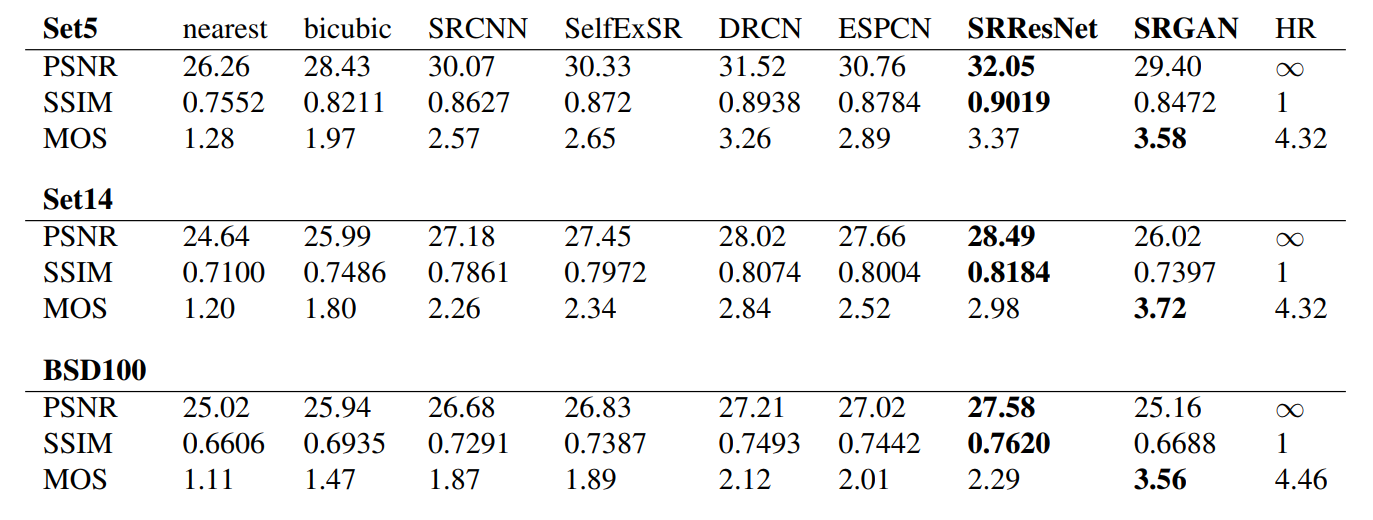


Таблица 2. Сравнение NN, бикубическая, SRCNN, SelfExSR, DRCN, ESPCN, SRResNet, SRGAN-VGG54 и оригинальное HR изображение на оценочных данных. [4x масштабирование]

**3.4. Изучение потери содержания**

Мы исследовали влияние различных вариантов потери содержания в потере восприятия для сетей на основе GAN. Конкретно мы исследуем для соответсвующей потери содержания

* **SRGAN-MSE**: , чтобы исследовать состязательную сеть со стандартной потерей содержания MSE.
* **SRGAN-VGG22**: с , потеря определенная на картах особенностей определяющих низкоуровневые особенности.
* **SRGAN-VGG54**: с , потеря определенная на картах особенностей определяющих высокоуровневые особенности из глубоких слоёв сети с большим потенциалом сфокусированным на содержания из изображений. Дальше эту сеть будем называть **SRGAN.**

Мы также оценивает производительность генераторной сети без состязательной компоненты для двух потерь (**SRResNet**-MSE) и (SRResNet-VGG22). Мы называем SRResNet-MSE как **SRResNet**. Заметим, что при обучении SRResNet-VGG22 мы добавили дополнительную общую вариационную потерю с весом . Количественные результаты суммированы в Таблице 1 и визуальные примеры представлены на Рисунке 5. Даже в сочетании с состязательной потерей MSE обеспечивает решения с наибольшим PSNR, которые воспринимаются довольно гладкими и менее убедительными, чем результаты, достигнутые с компонентами потерь более чувствительными к визуальному восприятию. Это вызвано конкуренцией между потерей содержания на основе MSE и состязательной потерей. Кроме того, мы приписываем этим конкурирующим целям незначительные артефакты реконструкции, которые мы наблюдали в меньшенстве реконструкций на основе SRGAN-MSE. Мы не можем определить существенно лучшую функцию потерь для SRResNet или SRGAN в отношении оценки MOS на Set5. Однако SRGAN-VGG54 значительно превзошел другие SRGAN и SRResNet варианты на Set14 с точки зрения MOS. Мы наблюдали тренд, что используя высокоуровневые карты особенностей VGG дает лучшую детализацию текстуры по сравнению с (см. Рисунок 5).

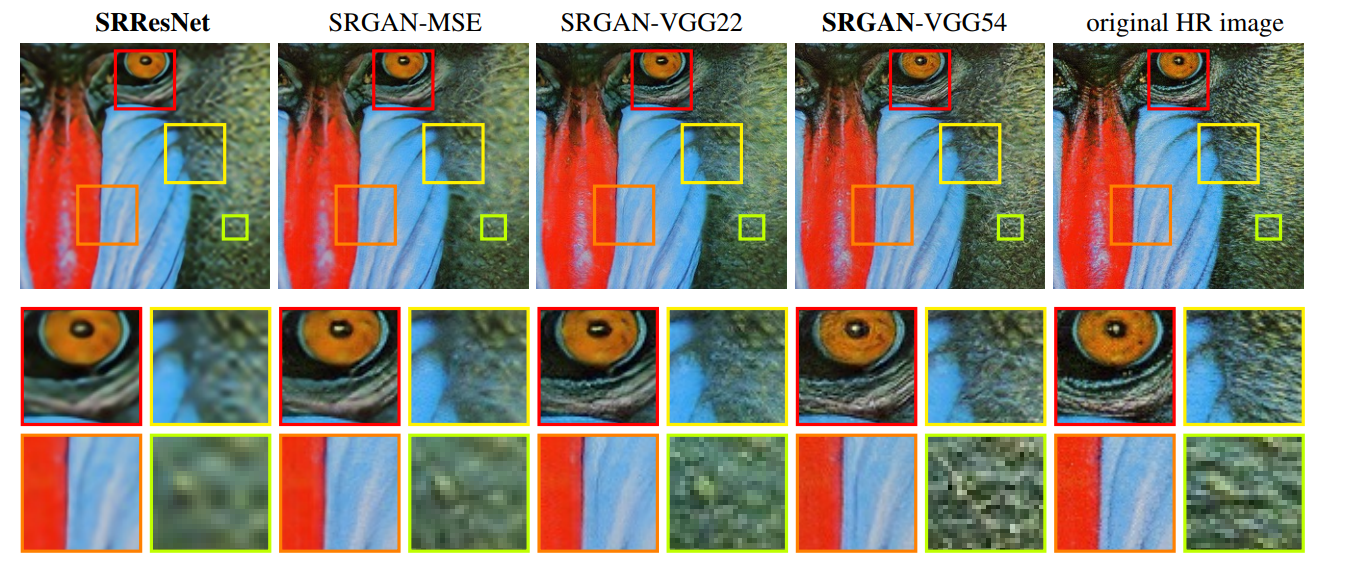


Рисунок 5. **SRResNet**, SRGAN-MSE, SRGAN-VGG22, **SRGAN**-VGG54 результаты восстановления и соответствующее HR изображение [4x масштабирование]

**3.5. Производительность финальных сетей**

**4. Обсуждение и будущая работа**