

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра математического обеспечения и применения ЭВМ

ОТЧЕТ
по научно-исследовательской работе
Тема: Анализ методов повышения разрешения изображения
для проведения экспериментов

Студентка гр. 3303

Саклакова В.А.

Руководители

Шолохов А.В.

Кринкин К.В.

Санкт-Петербург

2018

ЗАДАНИЕ НА НАУЧНО-ИССЛЕДОВАТЕЛЬСКУЮ РАБОТУ

Студентка Саклакова В.А.

Группа 3303

Тема НИР: Анализ методов повышения разрешения изображения
для проведения экспериментов

Задание на НИР:

В ходе выполнения данной научной исследовательской работы предстоит проанализировать различные методы и варианты построения архитектур в рамках задачи повышения разрешения изображения для проведения экспериментов и оценки производительности. Были определены следующие задачи научной исследовательской работы:

- Изучение основ машинного обучения.
- Изучение методов повышения разрешения изображения.
- Проведение сравнительного анализа методов по результатам экспериментов.

Сроки выполнения НИР: 13.02.2018 – 24.05.2018

Дата сдачи отчета: 24.05.2018

Дата защиты отчета: 24.05.2018

Студентка

Саклакова В.А.

Руководители

Шолохов А.В.

Кринкин К.В.

АННОТАЦИЯ

Целью научно-исследовательской работы является изучение основ машинного обучения, исследование методов повышения разрешения изображения и проведение экспериментов на основе существующих методов. В ходе данного исследования был произведен обзор предметной области, в частности были рассмотрены такие понятия, как машинное обучение, нейронные сети, были рассмотрены различные методы повышения разрешения изображения, и проведён сравнительный анализ по результатам экспериментов.

SUMMARY

The purpose of the research work is to study of fundamentals of machine learning, to research methods of image super-resolution and carrying out experiments on the basis of the existing methods. During this research the review of data domain was made, in particular such concepts as machine learning, neural networks, were considered, different methods of image super-resolution were considered, and the comparative analysis by results of experiments is carried out.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1. Обзор предметной области	6
1.1. Машинное обучение	6
1.2. Нейронные сети	8
2. Описание методов повышения разрешения изображения	10
2.1. Постановка задачи	10
2.2. Методы повышения разрешения изображения	10
3. Сравнительный анализ подходов к решению проблемы	16
3.1. Проведение экспериментов	16
3.2. Сравнительный анализ	19
3.3. Актуальность работы	19
ЗАКЛЮЧЕНИЕ	21
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	22

ВВЕДЕНИЕ

В ходе выполнения данной научной исследовательской работы предстоит подробно изучить предметную область, проанализировать различные методы повышения разрешения изображения и провести эксперименты для сравнительного анализа.

Были определены следующие задачи научной исследовательской работы:

- Изучить предметную область: виды машинного обучения, нейронные сети.
- Изучить методы повышения разрешения изображения: линейные, нелинейные, нейросетевые.
- Провести сравнительный анализ методов по результатам экспериментов.

1. Обзор предметной области

1.1. Машинное обучение

Машинное обучение (Machine Learning) — обширный подраздел искусственного интеллекта, изучающий методы построения алгоритмов, способных обучаться. Различают два типа обучения. Обучение по прецедентам, или индуктивное обучение, основано на выявлении общих закономерностей по частным эмпирическим данным. Дедуктивное обучение предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний. Дедуктивное обучение принято относить к области экспертных систем, поэтому термины машинное обучение и обучение по прецедентам можно считать синонимами.

Машинное обучение находится на стыке математической статистики, методов оптимизации и классических математических дисциплин, но имеет также и собственную специфику, связанную с проблемами вычислительной эффективности и переобучения. Многие методы индуктивного обучения разрабатывались как альтернатива классическим статистическим подходам. Многие методы тесно связаны с извлечением информации и интеллектуальным анализом данных (Data Mining).

В основном, выделяют следующие основные типы задач:

- Обучение с учителем (supervised learning) — наиболее распространённый случай. Каждый прецедент представляет собой пару «объект, ответ». Требуется найти функциональную зависимость ответов от описаний объектов и построить алгоритм, принимающий на входе описание объекта и выдающий на выходе ответ. Функционал качества обычно определяется как средняя ошибка ответов, выданных алгоритмом, по всем объектам выборки.

- Задача классификации (classification) отличается тем, что множество допустимых ответов конечно. Их называют метками классов (class label). Класс — это множество всех объектов с данным значением метки.

- Задача регрессии (regression) отличается тем, что допустимым ответом является действительное число или числовой вектор.

- Задача ранжирования (learning to rank) отличается тем, что ответы надо получить сразу на множестве объектов, после чего отсортировать их по значениям ответов. Может сводиться к задачам классификации или регрессии. Часто применяется в информационном поиске и анализе текстов.

- Задача прогнозирования (forecasting) отличается тем, что объектами являются отрезки временных рядов, обрывающиеся в тот момент, когда требуется сделать прогноз на будущее. Для решения задач прогнозирования часто удаётся приспособить методы регрессии или классификации, причём во втором случае речь идёт скорее о задачах принятия решений.

- Обучение без учителя (unsupervised learning). В этом случае ответы не задаются, и требуется искать зависимости между объектами.

- Задача кластеризации (clustering) заключается в том, чтобы сгруппировать объекты в кластеры, используя данные о попарном сходстве объектов. Функционалы качества могут определяться по-разному, например, как отношение средних межкластерных и внутрикластерных расстояний.

- Задача поиска ассоциативных правил (association rules learning). Исходные данные представляются в виде признаковых описаний. Требуется найти такие наборы признаков, и такие значения этих признаков, которые особенно часто (неслучайно часто) встречаются в признаковых описаниях объектов.

- Задача фильтрации выбросов (outliers detection) — обнаружение в обучающей выборке небольшого числа нетипичных объектов. В некоторых приложениях их поиск является самоцелью (например, обнаружение мошенничества). В других приложениях эти объекты являются следствием ошибок в данных или неточности модели, то есть шумом, мешающим настраивать модель, и должны быть удалены из выборки, см. также робастные методы и одноклассовая классификация.

- Задача построения доверительной области (quantile estimation) — области минимального объёма с достаточно гладкой границей, содержащей заданную долю выборки.

- Задача сокращения размерности (dimensionality reduction) заключается в том, чтобы по исходным признакам с помощью некоторых функций преобразования перейти к наименьшему числу новых признаков, не потеряв при этом никакой существенной информации об объектах выборки. В классе линейных преобразований наиболее известным примером является метод главных компонент.

- Задача заполнения пропущенных значений (missing values) — замена недостающих значений в матрице объекты–признаки их прогнозными значениями.

- Обучение с подкреплением (reinforcement learning). Роль объектов играют пары «ситуация, принятое решение», ответами являются значения функционала качества, характеризующего правильность принятых решений (реакцию среды). Как и в задачах прогнозирования, здесь существенную роль играет фактор времени. Примеры прикладных задач: формирование инвестиционных стратегий, автоматическое управление технологическими процессами, самообучение роботов, и т.д.

1.2. Нейронные сети

Искусственная нейронная сеть (ИНС) — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. С точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п.

У искусственных нейронов есть входы и выходы. Выходы нейронов связаны со входами других нейронов. Места, где они соединяются, называются синапсами. В синапсах проходящий сигнал может либо усилиться, либо

ослабнуть. Каждый синапс может изменять степень усиления/ослабления сигнала. Когда синапсов очень много, можно подобрать такие их параметры, что входной сигнал будет преобразовываться в правильный выходной. На рисунке 1 представлена полная внутренняя модель искусственного нейрона.

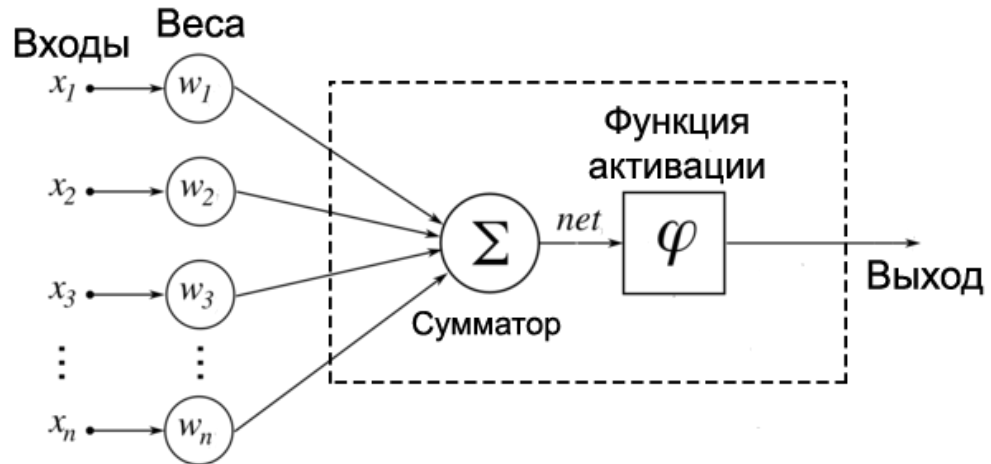


Рисунок 1 – Модель искусственного нейрона.

У каждого нейрона, в том числе и у искусственного, должны быть входы, через которые он принимает сигнал. Поступившие на входы сигналы умножаются на свои веса. Сигнал первого входа x_1 умножается на соответствующий этому входу вес w_1 . В итоге получаем $x_1 w_1$. И так до n -ого входа. В итоге на последнем входе получаем $x_n w_n$. Далее все произведения передаются в сумматор, который суммирует все входные сигналы, умноженные на соответствующие веса:

$$x_1 w_1 + x_2 w_2 + \dots + x_n w_n = \sum_{i=1}^n x_i w_i$$

Результатом работы сумматора является число, называемое взвешенной суммой. Просто так подавать взвешенную сумму на выход бессмысленно. Нейрон должен как-то обработать ее и сформировать адекватный выходной сигнал. Именно для этих целей и используют функцию активации, которая преобразует взвешенную сумму в такое число, которое и является выходом нейрона.

2. Описание методов повышения разрешения изображения

2.1. Постановка задачи

Цифровое изображение имеет конечное количество пикселей. Количество этих пикселей, деленное на единицу площади, называется разрешением цифрового изображения.

В обработке изображений существует класс методов повышения разрешения изображения, которые позволяют качественно увеличить разрешение исходного изображения, при этом происходит преодоление оптического предела объектива и/или физического разрешения цифрового сенсора, который записал изображение [1].

Методы повышения разрешения изображения часто используются для того, чтобы рассмотреть подробности небольшого фрагмента снимка. Их суть заключается в увеличении количества пикселей на некоторую единицу длины (например, на сантиметр или на дюйм). На сегодняшний день существует множество подходов к решению этой задачи, однако наибольшей популярностью пользуются алгоритмы интерполяции (в частности, метод ближайшего соседа, билинейная и бикубическая интерполяция) [1]. Принцип их работы довольно прост: каждый новый пиксель достраивается на основе некоторого набора окружающих пикселей в соответствии с выбранным фильтром. Такие методы позволяют увеличить разрешение изображения довольно быстро, однако они имеют существенный недостаток — небольшие фрагменты получаются слишком гладкими и мелкие детали картинки теряются. Использование нейросетевых подходов к решению поставленной задачи позволяют обойти эти недостатки.

2.2. Методы повышения разрешения изображения

– Линеинные методы;

Простейшим методом повышения разрешения изображений является интерполяция. При интерполяции изображение представляется в виде функции, где пиксели изображения — точки, в которых значения функции известны.

Интерполяция – нахождение промежуточных значений по дискретному набору известных значений. Точное восстановление информации путем интерполяции невозможно. Методы интерполяции основаны на использовании свертки изображения и ядер разных типов.

При решении задачи повышения разрешения цифровых изображений требуется не построение непрерывного изображения, а переход с более грубой сетки на более мелкую сетку. Такой процесс называют ресамплингом изображений, а коэффициент отношения шага крупной сетки к шагу мелкой – коэффициентом повышения разрешения изображения.

Самым простым и самым быстрым линейным методом интерполяции является метод «ближайшего соседа», который заключается в растяжении пикселей. В случае ресамплинга с целым коэффициентом увеличения он представляет собой простое повторение пикселей изображения. Недостатком применения данного метода является ярко выраженный эффект ступенчатости. Восстановленная функция становится более гладкой, если вместо ступенек соединять точки прямыми. Так осуществляется линейная интерполяция первого порядка. Билинейной называется линейная интерполяция первого порядка в двумерном случае. В результате применения такого метода интерполяции появляется неярко выраженные эффекты ступенчатости и размытия изображения.

Наиболее распространенные негативные эффекты, возникающие при повышении разрешения изображений, связаны с искажением высокочастотной информации [1]. Такими эффектами являются размытие, алиасинг (ступенчатость контуров) и эффект Гиббса. В задачах обработки изображений эффект Гиббса проявляется как эффект ложного оконтуривания, проявляющийся в виде ореолов возле резких контуров.

– Нелинейные методы;

Избавиться от недостатков линейных методов интерполяции можно с помощью адаптивных нелинейных методов, для которых ядро задается отдельно для каждого интерполируемого пикселя и зависит от значений

пикселей интерполируемого изображения. Примером нелинейных методов является класс градиентных методов. Градиентные алгоритмы позволяют достичь лучших результатов при интерполяции вдоль контуров деталей изображения путем использования разных ядер вдоль и поперек границ (например, использование функции Гаусса с переменными радиусами по разным направлениям) [3]. В целом результат получается близким к результату бикубической интерполяции, но эффект алиасинга оказывается практически полностью подавленным.

– Нейросетевые методы.

Одним из преимуществ использования искусственной нейронной сети является её способность к обучению, которая может реализовываться в виде изменения внутренних свойств нейронной сети для выполнения требуемых функций. Данный подход предпочтителен тем, что мы можем генерировать почти неограниченную выборку низкоразрешенных изображений, причем как шум, так и нелинейная функция, в отличие от предыдущих методов, могут быть практически произвольными. Подобные подходы активно применяются для решения задачи удаления шума с изображений или устранения размытости в изображениях.

В последние годы наилучшие результаты во многих задачах компьютерного зрения были достигнуты с помощью использования глубоких свёрточных нейронных сетей [4].

Свёрточная нейронная сеть – специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году и нацеленная на эффективное распознавание изображений, входит в состав технологий глубокого обучения. Использует некоторые особенности зрительной коры, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток. Таким образом, идея свёрточных нейронных сетей заключается в чередовании свёрточных слоев и слоёв подвыборки. Структура сети — однонаправленная (без обратных

связей), принципиально многослойная. Для обучения используются стандартные методы, чаще всего метод обратного распространения ошибки. Функция активации нейронов (передаточная функция) — любая, по выбору исследователя.

Название архитектура сети получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.

На рисунке 2 изображена архитектура свёрточной нейронной сети.

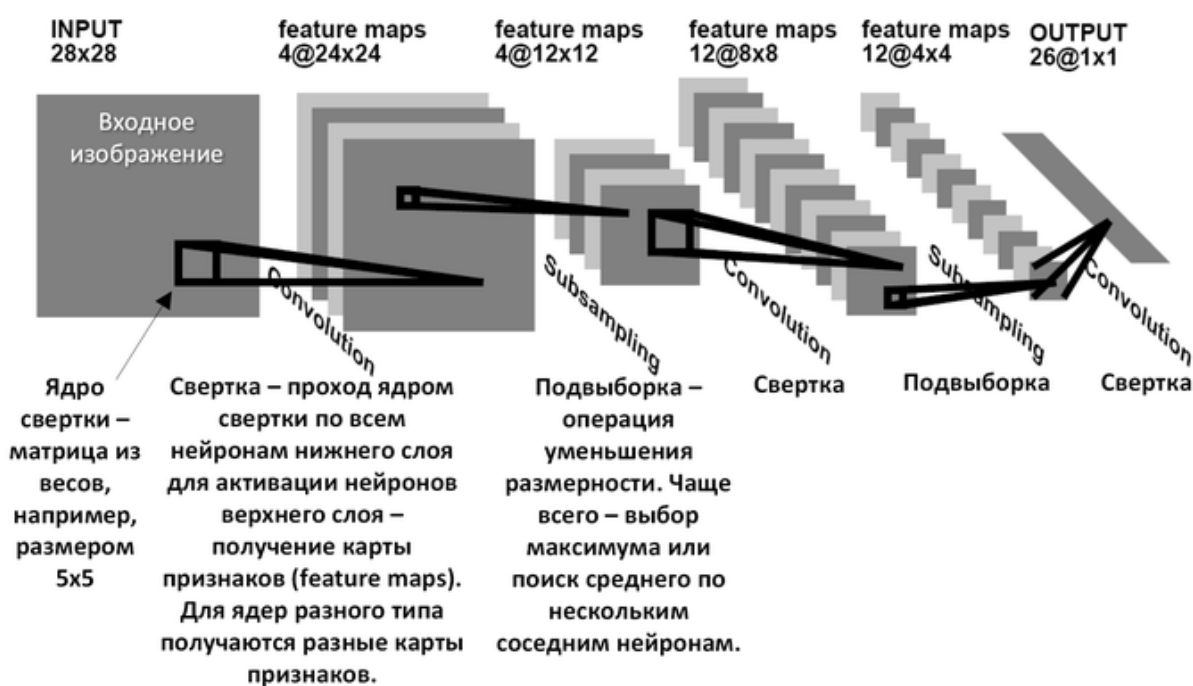


Рисунок 2 – Архитектура свёрточной нейронной сети.

Большинство подходов к решению задачи генерации фотореалистичных изображений используют для задания вероятностного распределения графические модели. Подобные модели зачастую оказываются неэффективными из-за необходимости проведения приближенных процедур вывода, а также использования Марковских цепей для генерации выборки. Недавно предложенная модель конкурирующих сетей [5] позволяет выучивать сложные распределения, а также генерировать выборки произвольного размера

используя только вычислительно эффективный алгоритм обратного распространения ошибки для нейронных сетей.

Генеративно-сопоставительная сеть (Generative adversarial network) – алгоритм машинного обучения без учителя, построенный на комбинации из двух нейронных сетей, одна из которых (сеть G) генерирует образцы, а другая (сеть D) старается отличить правильные («подлинные») образцы от неправильных. Так как сети G и D имеют противоположные цели — создать образцы и отбраковать образцы — между ними возникает Антагонистическая игра. Генеративно-сопоставительную сеть описал Ян Гудфеллоу из компании Google в 2014 г.

Использование этой техники позволяет, в частности, генерировать фотографии, которые человеческим глазом воспринимаются как натуральные изображения. Кроме того, GAN может использоваться для улучшения качества нечётких или частично испорченных фотографий [6].

На рисунке 3 изображена модель GAN сети.

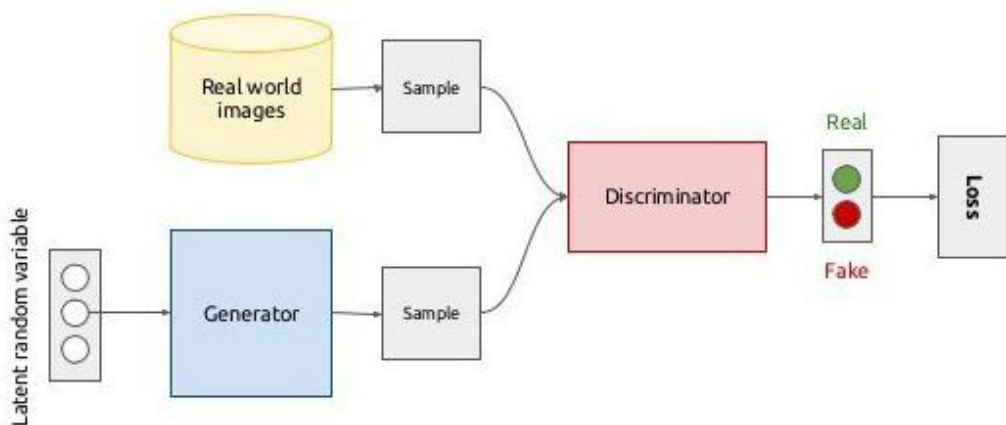


Рисунок 3 – Модель GAN сети.

Принцип состязательности в сети GAN нередко описывается через метафоры. Например, генеративная сеть уподобляется фальшивомонетчику или подделывателю картин, а дискриминативная — эксперту который стремится распознать подделку. Другой пример — образ двух боксёров, один из которых учился у мастера, а второй вынужден подражать ученику. В популярном приложении генерации человеческих лиц в качестве подлинных данных

выступают реальные фотографии, а генеративная сеть пытается создать искусственные лица, варьируя комбинации таких латентных параметров, как цвет волос, пропорции лица, разрез глаз, форма носа, размер ушей, наличие бороды и усов и т.д.

Формально подобную модель можно записать в виде следующей минимакс игры:

$$G, D \leftarrow \min_G \max_D (\mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))])$$

Здесь z — это некоторые модельные данные (случайный шум), $p_z(z)$ — это модельное распределение, выборки из которого подаются на вход сети G . x — объекты из распределения данных $p_{data}(x)$, из которого сеть G учится генерировать выборки. В качестве модельного распределения p_z обычно используется равномерное распределение $U[-1,1]$.

Модель конкурирующих сетей на сегодняшний день является одной из лучших моделей для генерации фотореалистичных изображений.

3. Сравнительный анализ подходов к решению проблемы

3.1. Проведение экспериментов

В качестве примера линейного метода была рассмотрена бикубическая интерполяция. Бикубическая интерполяция обладает одним из лучших сочетаний качества и скорости. Бикубическая интерполяция является обобщением интерполяции кубическими сплайнами на двумерный случай. Алгоритм кубической интерполяции реализован во многих редакторах цифровых изображений. По сравнению результатов бикубической интерполяции и предложенного метода можно судить том, улучшает алгоритм изображение, или нет.

Бикубическая интерполяция – в вычислительной математике расширение кубической интерполяции на случай функции двух переменных, значения которой заданы на двумерной регулярной сетке. Поверхность, полученная в результате бикубической интерполяции, является гладкой функцией, в отличие от поверхностей, полученных в результате билинейной интерполяции или интерполяции методом ближайшего соседа. Так же бикубическая интерполяция часто используется в обработке изображений, давая более качественное изображение по сравнению с билинейной интерполяцией. В случае бикубической интерполяции значение функции в искомой точке вычисляется через её значения в 16-ти соседних точках. Бикубическая интерполяция идёт на один шаг дальше билинейной, рассматривая массив из 4x4 окружающих пикселей — всего 16. Поскольку они находятся на разных расстояниях от неизвестного пикселя, ближайшие пиксели получают при расчёте больший вес. Бикубическая интерполяция производит значительно более резкие изображения, чем предыдущие два метода.

На рисунке 4 представлен результат работы метода бикубической интерполяции при увеличении масштаба изображения в 4 раза.



Рисунок 4 – Метод бикубической интерполяции.

В качестве примера нелинейного метода был рассмотрен градиентный метод. Суть градиентного метода в том, что интерполирование производится по краям (контур) изображения и ядро интерполирования по оси абсцисс отличается от ядра интерполирования по оси ординат. На рисунке 5 представлен результат работы градиентного метода.

$$f(x, y) = \sum_{i, j=-\infty}^{+\infty} F(i, j) K_x(i - x) K_y(i - y)$$

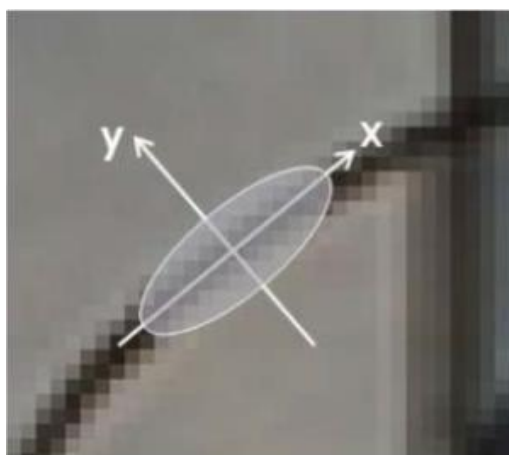


Рисунок 5 – Градиентный метод.

В качестве примера нейросетевого метода была рассмотрена модель GAN сети (кратко изложена в п.2.2). Но для сети Дискриминатора и сети Генератора в основу легли архитектуры не обычных нейронных сетей, а свёрточных. Таким образом, были совмещены генеративно-состязательные сети и свёрточные сети.

Генератор G , дискриминатор D . Есть какое-то количество оригинальных изображений X для D , и пусть на выходе у него будет число диапазоном от

нуля до единицы, чтобы оно выражало уверенность D в том, что поданные ему на вход изображения оригинальные. Еще — поскольку G нейронная сеть, ей нужны какие-то входные данные, назовем их z . В общем случае, это просто случайный шум, который модель будет стараться превратить в реальное изображение.

Тогда, очевидно, цель G — это максимизировать $D(G(z))$, то есть сделать так, чтобы D был уверен, что подделки — настоящие.

Цель D посложнее — ему нужно одновременно положительно опознавать оригиналы, и отрицательно — подделки. Запишем это как максимизацию $D(x)(1 - D(G(z)))$. Умножение можно превратить в сложение, если взять логарифм, поэтому получается:

Для дискриминатора: максимизировать $\log D(x) + \log(1 - D(G(z)))$

Для генератора: максимизировать $\log D(G(z))$

Так как в поставленной задаче на вход генератору подается не шум, а изображение с низким разрешением, то применяя общую формулу на задачу, получается данная формула:

$$\min_G \max_D V(G, D) = E_{I_{\text{generated}} \sim P_{\text{train}}(I_{\text{original}})} [\log D(I_{\text{original}})] + E_{I_{\text{generated}} \sim P_G(I_{\text{input}})} [\log(1 - D(G(I_{\text{input}})))] \quad (1)$$

которая значит, что I_{original} — множество (тренировочное), которое заранее приведено к какому-то распределению. $I_{\text{generated}}$ — множество генерируемых изображений Генератором должно быть таким же распределением, как и I_{original}

На вход такой сети подаётся изображение 16x16, которое сеть Генератора масштабирует в 4 раза (до 64x64), и пытается восстановить размытое изображение. Сеть Дискриминатора обучается на то, чтобы различать оригинальные высокоразрешенные изображения и изображения, которые являются выходными данными из сети Генератора.

На рисунке 6 представлена архитектура DCGAN сети.

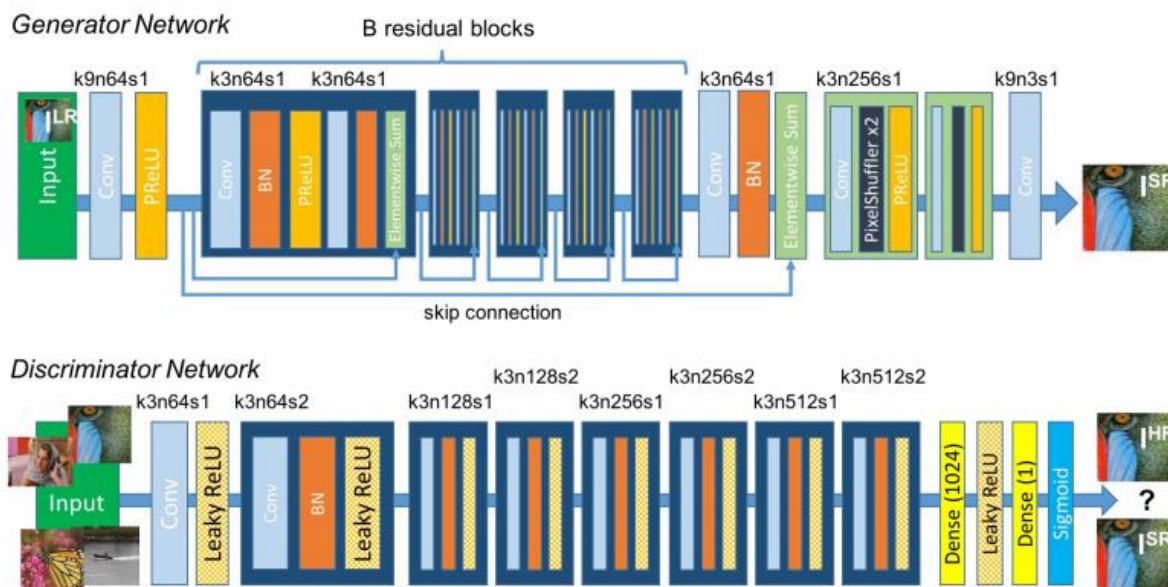


Рисунок 6 – Архитектура DCGAN сети.

3.2. Сравнительный анализ

В результате получается, что градиентный метод позволяет более качественно увеличивать разрешение изображения по сравнению с линейными методами. Изображение, полученное данным способом похоже на изображение, полученное с помощью бикубической интерполяции, но эффект ступенчатости сведен к минимуму.

Сравнивая результаты работы DCGAN сети и обработки изображения градиентным методом, было установлено, что метод DCGAN справляется лучше в отношении восприятия человеческим глазом, т.е. линии чётче и лучше прорисовываются.

3.3. Актуальность работы

Для анализа и обработки изображений в большинстве случаев используются изображения высокого разрешения, так как они позволяют увидеть детали, которые не различимы, или плохо различимы на изображениях с низким разрешением [2]. Многие задачи компьютерного зрения изначально полагаются на то, что изображение на входе в хорошем разрешении.

С первого взгляда, задача увеличения разрешения аппаратная – чтобы увеличить разрешение необходимо просто взять фотоаппарат с большей разрешающей способностью. Но часто уже имеется некоторая фотография или набор фотографий, и нет возможности повторно произвести съемку. Примером может послужить низкокачественные картинки с изображений камер наблюдения

Решение данной задачи важно для многих приложений: зуммирование фото и текста, проекция видео на большой экран и т.д. Иногда необходимо рассмотреть номер автомобиля на кадре с камеры наблюдения, приблизив фотографию до предела. И не только номер автомобиля.

ЗАКЛЮЧЕНИЕ

В результате проведения научной исследовательской работы были выполнены следующие задачи:

- Изучена предметная область: виды машинного обучения, нейронные сети.
- Изучены методы повышения разрешения изображения: линейные, нелинейные, нейросетевые.
- Проведён сравнительный анализ методов по результатам экспериментов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Маркелов К.С. Модель повышения информативности цифровых изображений на базе метода суперразрешения // Инженерный вестник. 2013, вып. (№3). С. 525-542.
2. Гитман И.А. Применение конкурирующих сетей для задач машинного обучения / МГУ им. М.В. Ломоносова, М., 2016.
3. Насонов А.В. Регуляризирующие методы повышения изображений: автореф. дисс. канд. техн. наук / МГУ им. М.В. Ломоносова, М., 2011.
4. J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” arXiv preprint arXiv:1511.04587, 2015.
5. L. Xu, J. S. Ren, C. Liu, and J. Jia, “Deep convolutional neural network for image deconvolution,” in Advances in Neural Information Processing Systems, 2014, pp. 1790–1798.
6. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.