

Глубокое обучение в задаче реконструкции суперразрешения изображений¹

А. А. Сенов², аспирант

Санкт-Петербургский государственный университет

alexander.senov@gmail.com

Обработка изображений несколько последних десятилетий является активно развивающейся научной областью. Один из важных ее разделов — получение суперразрешения изображений. Изображения с высоким разрешением имеют лучшие характеристики для восприятия человеком и позволяют проводить более качественный последующий анализ. Основной способ повышения разрешения изображения — улучшение технических характеристик оборудования, используемого для съемки. Однако, в этом случае возникают сложности, связанные со стоимостью и техническими ограничениями. Для их преодоления был предложен подход — получение суперразрешения изображения, который состоит в получении изображения высокого разрешения из одного или нескольких изображений с низким разрешением. В этой работе приводится обзор подходов к решению задачи получения суперразрешения изображений. Особенный упор делается на подходы, основанные на методах глубокого обучения, которые в последние годы показывают конкурентные результаты. В том числе обсуждается проблема оценки параметров сверточной нейронной сети, особенно актуальной в задаче суперразрешения.

Ключевые слова: суперразрешение, глубокое обучение, оценка параметров.

1. Введение

Во многих задачах цифровой обработки, изображения с высоким разрешением предпочтительны для качественных последующих обработки и анализа. Потребность в изображениях высокого разрешения связана с двумя основными областями применения: улучшение изобразительной информации для интерпретации человека; и помощь в представлении для автоматического восприятия машины. Разрешение изображения описывает детали, содержащиеся в изображении, и чем выше разрешение, тем детальнее изображение. Высокое разрешение означает, что плотность пикселей в изображении высока, и поэтому изображение с высоким разрешением может предложить более подробную информацию, которая в

¹Работа выполнена при поддержке РНФ, проект №17-51-53053.

²©А. А. Сенов, 2017

свою очередь может иметь решающее значение в различных приложениях. Например, медицинские снимки очень полезны для врача, и их детализация может оказать решающее влияние на правильность постановки диагноза, в задачах обработки спутниковых изображений высокое изображение поможет обнаружить мелкое тело, а эффективность распознавания образов в компьютерном зрении может быть улучшена, если разрешение рассматриваемого изображения будет высоким.

Основной способ повышения разрешения изображения — улучшение технических характеристик оборудования, используемого для съемки. Однако, в этом случае возрастает стоимость, вес и габариты оборудования, что во многих случаях может оказаться неприемлемым. Так, наращивание разрешения в широко используемых камерах наблюдения и встроенных камерах сотового телефона нецелесообразно из соображений стоимости. Также, разрешение камеры наблюдения ограничено скоростью камеры и аппаратным обеспечением. В некоторых других сценариях, таких как спутниковые снимки, трудно использовать датчики высокого разрешения из-за физических ограничений. Другим способом решения этой проблемы является принятие плохого качества изображений и его последующее улучшение программными средствами для компенсации вычислительных затрат за счет стоимости оборудования. Методы, решающие эту задачу называются методами реконструкцией суперразрешения или просто методами суперразрешения (СР).

Методы суперразрешения — это методы, которые создают изображения высокого разрешения (ВР) из нескольких наблюдаемых изображений с низким разрешением (НР), тем самым увеличивая высокочастотные компоненты и устраняя ухудшения, вызванные процессом формирования изображения камеры с низким разрешением. Основная идея СР заключается в объединении не избыточной информации, содержащейся в нескольких кадрах с низким разрешением, для создания изображения с высоким разрешением. Тесно связанный с СР метод — это метод интерполяции одиночного изображения, который также можно использовать для увеличения размера изображения. Однако, поскольку в таком случае дополнительная информация отсутствует, качество интерполяции одиночного изображения очень ограничено из-за некорректного характера проблемы, а потерянные частотные компоненты не могут быть

восстановлены. В целом можно выделить две основные группы методов СР: методы суперразрешения одного и методы суперразрешения многих изображений.

В первой группе для получения изображения ВР дано лишь одно соответствующее изображение НР, во второй — несколько изображений НР, что упрощает задачу. Не избыточная информация, содержащаяся в этих изображениях НР, обычно вводится субпиксельными сдвигами между ними. Эти сдвиги могут возникать из-за неконтролируемых движений между системой формирования изображения и сценой, (например, обходами объектов) или из-за контролируемых движений. В качестве примера можно привести систему получения спутниковых изображения Земли с заранее заданной скоростью и траекторией.

Области применения методов СР включают, но не ограничиваются: видеонаблюдением, медицинскими изображениями (КТ, МРТ, УЗИ), конвертацией из одного формата видео/изображений в другой (например, из NTSC в HDTV, любительская фото- и видеосъемка. Исходя из различных входных данных можно сделать вывод о широком спектре входных условий в задаче СР: наличия одного или нескольких изображений, наличие видеопотока, характере сдвигов между изображениями НР, их разрешения и т.п. В зависимости от актуальности тех или иных условий зависит применимость и оптимальность методов СР. Общий обзор методов суперразрешения приведен в следующем разделе.

2. Методы суперразрешения

2.1. Модель наблюдения

Обозначим $\mathbf{x} \in \mathbb{R}^{L_1 N_1 L_2 N_2}$ — вектор изображения высокого разрешения, и $\mathbf{y} \in \mathbb{R}^{N_1 N_2}$ — вектор изображения низкого разрешения, где L_1 и L_2 — показатели понижения размерности по вертикали и горизонтали соответственно. Общепринятой является следующая модель получения изображения низкого разрешения из изображения высокого разрешения

$$\mathbf{y} = \mathbf{DHF}\mathbf{x} + \mathbf{v}, \quad (1)$$

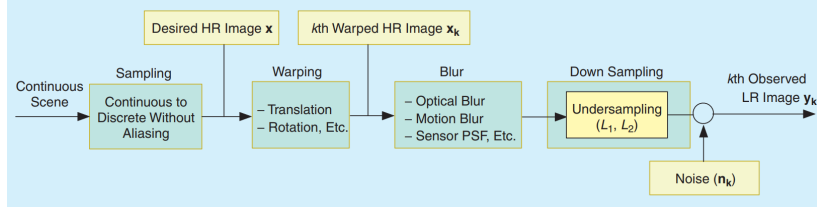
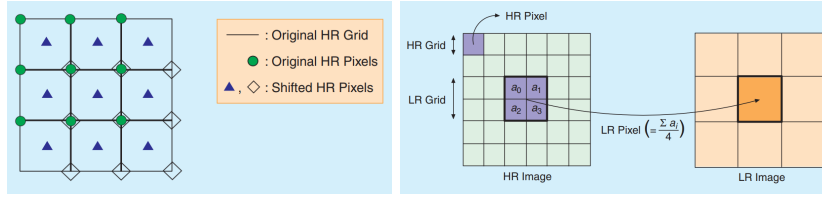


Рис. 1: Модель получения изображения ВР из изображений НР. Схема взята из работы [1], Рис. 3



(a) Интерполяция на сетке ВР. (b) Оператор пространственного усреднения. Схема взята из работы [1], Рис. 4

Рис. 2

где

- $\mathbf{F} \in \mathbb{R}^{L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2}$ — матрица движения (сдвига, поворота, и т. п.),
- $\mathbf{H} \in \mathbb{R}^{L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2}$ — матрица размытия (оптическое размытие, размытие от движения и т. д.),
- $\mathbf{D} \in \mathbb{R}^{N_1 N_2 \times L_1 N_1 L_2 N_2}$ — матрица понижения размерности,
- $\mathbf{v} \in \mathbb{R}^{N_1 N_2}$ — вектор помех.

Модель связи изображений НР и изображения ВР проиллюстрирована на Рис. 1. Движение, которое происходит во время получения изображения, представлено матрицей деформаций \mathbf{F} . Оно может содержать глобальный или локальный перевод, ротацию и пр. Вообще говоря, эта матрица неизвестна и необходимо оценить движение сцены для каждого кадра относительно других. Процесс

деформирования, выполняемый на изображении ВР фактически определяется в терминах интервалов $\{L_k\}_1^K$ изображений НР, во время его оценивания. Таким образом, этот шаг требует интерполяции, когда дробная единица движения не равна сетке изображения ВР. Пример глобального перевода показан на Рис. 2а. Здесь круг (○) Представляет собой исходный (ссылочный) образ \mathbf{x} , а треугольник (\triangle) и ромб (\diamond) представляют собой глобально сдвинутые версии. Если коэффициент понижающей выборки равен двум, ромб (\diamond) имеет субпиксельный сдвиг (0.5, 0.5) для горизонтального и вертикального направлений, а треугольник (\triangle) имеет сдвиг, который меньше (0.5, 0.5). Следовательно, как показано на Рис. 2а, ромб (\diamond) не нуждается в интерполяции, но треугольник (\triangle) должен быть интерполирован из \mathbf{x} , так как он не расположен на сетке ВР. Теоретически можно использовать идеальную интерполяцию, на практике же в основном используются простые методы, например методы нулевого порядка и билинейные методы интерполяции.

Размытие может быть вызвано оптической системой (например, расфокусировкой, дифракционным пределом, аберрацией и т. д.), относительным движением между системой формирования изображения и исходной сценой и функцией рассеяния точки датчика НР. Ее эффект представлен матрицей \mathbf{H} . При СР одного изображения обычно рассматривается оптическое размытие или размытие в результате движения. Однако при реконструкции изображения НР конечность физического разрешения в датчиках НР является важным фактором размытия. Обычно этот датчик моделируется оператором пространственного усреднения (смотри Рис. 2б). При использовании методов реконструкции СР характеристики размытия зачастую считаются известными, но могут быть и оценены в процессе реконструкции.

Матрица понижения размерности \mathbf{D} генерирует изображение НР из искаженного и размытого изображения ВР путем наложения (алиасинга). Хотя размытие действует более или менее как фильтр сглаживания (анти-алиасинга), при реконструкции изображения ВР предполагается, что наложение всегда присутствует в изображениях НР. Стоит отметить, что в случае СР многих изображений разрешения изображений НР и, следовательно размеры матриц \mathbf{D}_k могут различаться.

Модель (1) допускает упрощенную запись вида:

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{v}.$$

Матрицы \mathbf{D} , \mathbf{H} и \mathbf{F} зачастую являются разреженными, из-за чего уравнение (2.1.) является некорректно поставленным. Более того, эти матрицы могут быть неизвестны. В обоих случаях, для успешного решения задачи суперразрешения необходимо априорное знание о матрицах \mathbf{D} , \mathbf{H} , \mathbf{F} помехах \mathbf{v} и векторе высокого разрешения \mathbf{x} выраженное в виде регуляризации и дополнительных ограничениях. В последующих разделах мы рассмотрим различные методы СР базирующиеся на различных допущениях и ограничениях.

2.2. Методы суперразрешения многих изображений

В задаче суперразрешения многих изображений несколько изображений низкого разрешения получаются согласно модели (1):

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{H}_k \mathbf{F}_k \mathbf{x} + \mathbf{v}_k, \quad k = 1..K, \quad (2)$$

где k — индекс изображения НР. Возможна и упрощенная запись

$$\mathbf{Y} = \mathcal{M}\mathbf{x} + \mathbf{V}, \quad (3)$$

где строки матрицы \mathbf{Y} — транспонированные векторы \mathbf{y}_k , а матрица \mathcal{M} — результат конкатенации матриц $\mathbf{M}_k = \mathbf{D}_k \mathbf{H}_k \mathbf{F}_k$ по вертикали.

Далее в этом разделе приведены некоторые методы суперразрешения. Более полный список методов приведен, например, в работах [1] и [2].

Частотные методы

В основополагающей работе [3] предложена модель связи изображения ВР и изображений НР в частотном спектре, основывающаяся на свойствах преобразований Фурье относительно сдвига и наложения спектров. Рассмотрим поясняющий пример: пусть $x(t_1, t_2)$ — непрерывная сцена (изображение ВР), $x_k(t_1, t_2) = x(t_1 + \Delta_{k1}, t_2 + \Delta_{k2})$ — сдвиги, соответствующие K изображениям НР,

$y_k[n_1, n_2] = x_k(t_1 T_1, t_2 T_2)$, $0 \leq n_1 \leq N_1 - 1$, $0 \leq n_2 \leq N_2 - 1$ — дискретные изображения НР, а $\chi(u_1, u_2)$ и $\chi_k(u_1, u_2)$ — результат применения преобразования Фурье к x и x_k соответственно. По свойству преобразования Фурье относительно сдвига:

$$\chi_k(u_1, u_2) = \exp(i2\pi(u_1 \Delta_{k_1} + u_2 \Delta_{k_2})) \chi(u_1, u_2).$$

Далее, пусть $\gamma_k[r_1, r_2]$ — результат применения дискретного преобразования Фурье к y . Тогда $\gamma_k[r_1, r_2] =$

$$= \frac{1}{T_1, T_2} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \chi_k \left(\frac{2\pi}{T_1} \left(\frac{r_1}{N_1} - m_1 \right), \frac{2\pi}{T_2} \left(\frac{r_2}{N_2} - m_2 \right) \right).$$

Учитывая, что функция χ финитна, получаем

$$\gamma = \Phi \chi,$$

где $\gamma \in \mathbb{C}^K$, $\chi \in \mathbb{C}^{N_1 N_2}$.

Приведенный подход может быть использован лишь когда параметры сдвига и наложения (Δ_k и T_k) известны, а помехи отсутствуют. Не смотря на то, что существуют обобщения данного подхода как на случай неизвестных параметров движения [5], так и на случай наличия шумов [4], он трудно обобщаем на более сложные случаи. На данный момент большинство исследований сконцентрированы на методах, работающих в пространственной, а не частотной области.

Методы на основе регистрации-интерполяции-восстановления

Допустим, все матрицы размытия совпадают: $\mathbf{H}_k = \mathbf{H}$, а матрицы \mathbf{F}_k и \mathbf{H} коммутируют (например, \mathbf{F}_k состоит только из сдвигов и поворотов). Тогда формула (2) выглядит следующим образом:

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{F}_k \mathbf{H} \mathbf{x} = \mathbf{D}_k \mathbf{F}_k \mathbf{z}, \quad k = 1..K.$$

Отсюда логичным образом следуют подход, состоящий из трех этапов:

1. Регистрация изображений НР — изображения НР $\{\mathbf{y}_k\}$ выравниваются с помощью одного из методов регистрации изображений [6] до субпиксельной точности.

2. Интерполяция полученных изображений — получение \mathbf{z}
3. Устранение размытия и помех с \mathbf{z} — получение \mathbf{x} .

В рамках этого подхода предложено большое число методов (см., например, [7, 8, 9]). Однако, подобный подход не является оптимальным в теоретическом смысле, а ошибка, полученная на этапе регистрации изображений распространяется на последующие этапы.

Статистические методы

В статистических методах матрица \mathcal{M} рассматривается в параметризованном виде $\mathcal{M}(\nu, h)$, где ν — параметр движения, а h — параметр размытия. Тогда, задача восстановления вектора изображения ВР \mathbf{x} из матрицы изображений НР \mathbf{Y} формируется с помощью байесовского подхода:

$$\begin{aligned}
 \mathbf{x} &= \operatorname{argmax}_{\mathbf{x}} P(\mathbf{x}|\mathbf{Y}) \\
 &= \operatorname{argmax}_{\mathbf{x}} \int_{\nu, h} P(\mathbf{x}, \mathcal{M}(\nu, h)|\mathbf{Y}) d\nu dh \\
 &= \operatorname{argmax}_{\mathbf{x}} \int_{\nu, h} \frac{P(\mathbf{Y}|\mathbf{x}, \mathcal{M}(\nu, h))P(\mathbf{x}, \mathcal{M}(\nu, h))}{P(\mathbf{Y})} d\nu dh \\
 &= \operatorname{argmax}_{\mathbf{x}} \int_{\nu, h} P(\mathbf{Y}|\mathbf{x}, \mathcal{M}(\nu, h))P(\mathbf{x})P(\mathcal{M}(\nu, h)) d\nu dh,
 \end{aligned} \tag{4}$$

где

- вероятность $P(\mathbf{Y}|\mathbf{x}, \mathcal{M}(\nu, h))$ интерпретируется как правдоподобие данных, которая в случае аддитивного шума \mathbf{V} с стандартным распределением Гаусса принимает вид:

$$\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{Y} - \mathcal{M}(\nu, h)\mathbf{x}\|\right),$$

- вероятность $P(\mathbf{x})$ содержит априорную информацию о изображении ВР, зачастую рассматривается в виде распределения Гиббса: $\frac{1}{C_e} \exp(-\alpha A(\mathbf{x}))$.

Заметим, что если рассмотреть случай известного движения и размытия, когда $\mathcal{M}(\nu h) = \mathcal{M}$, уравнение (4) принимает значительно более простой вид:

$$\mathbf{x} = \underset{\mathbf{x}}{\operatorname{argmax}} \|\mathbf{Y} - \mathcal{M}\mathbf{x}\|^2 + \lambda A(\mathbf{x}),$$

где λ содержит в себе информацию о σ и α .

В частности, примерами реализации такого подхода могут служить: оценки максимума правдоподобия (см. [10, 11]), оценки апостериорного максимума (см. [12, 13]), совместные оценки апостериорного максимума (см. [14, 15, 16]).

2.3. Методы суперразрешения одного изображения

Не смотря на то, что наличие нескольких изображений НР позволяет извлечь больше полезной информации и, как следствие, получить изображение более высокого качества, на практике не редко приходится работать с единственным изображением НР.

Большинство методов суперразрешения одного изображения состоит из двух этапов

1. построение отображения от патчей изображения НР к патчам ВР
2. применение построенного отображения к изображению НР

Целое семейство методов строит соответствие между патчами низкого и высокого разрешений используя единственное изображение НР.

В работе [17] авторы используют марковское случайное поле для обучения соответствия между патчами низкого и высокого разрешений и выбирают патч ВР с учетом его соответствия с уже выбранными патчами слева и сверху, так как патчи высокого разрешения накладываются друг на друга, а перебор патчей идет в порядке растровой развертки.

В работе [18] соответствие между патчами низкого и высокого разрешений строится на основе метода k ближайших соседей: для

данного патча НР выбирается k наиболее похожих патчей НР из базы, и соответствующие им патчи ВР усредняются.

Существуют подходы, где изображение ВР для данного патча так же получается путем взвешенного усреднения известных патчей ВР, но веса подбираются не исходя из метода ближайших соседей, а на основе метода опознания со сжатием (compressive sensing) решается задача l_1 оптимизации. Так, например, в работе [19] решается задача

$$\lambda \|\omega\|_1 + \|\omega\|_2^2 \rightarrow \min_{\omega}.$$

Естественна идея, что различные типы изображений требуют различного набора базовых изображений. Так в работе [20] авторы используют в качестве базиса для патчей предобученные на изображениях человеческих лиц главные компоненты (используя анализ главных компонент [21]). В работе [22] был предложен двух этапный подход, состоящий из получения изображения среднего разрежения с помощью главных компонент, и последующего получения изображения ВР с помощью непараметрического метода, основанного на примерах.

Методы, основанные на построении соответствия между патчами низкого и высокого разрешений эксплуатируют в большинстве своем одну идею: наличия на одном или нескольких изображениях схожих или идентичных элементов. Приведенные выше методы интерпретируют эту схожесть буквально: по близости патчей в пространстве l_p . В работе [23] авторы идут дальше и понимают близость с учетом поворотов и растяжений, минимизируя расстояния между двумя патчами, домножая один из них на соответствующую матрицу.

Приведенные в этом и предыдущих разделах работы демонстрируют оригинальные подходы, положившие начало целым направлениям исследований в области суперразрешения изображений. Впоследствии эти подходы были значительно расширены и доработаны.

3. Методы на основе сверточных нейронных сетей

Сверточные нейронные сети (СНН) были успешно применены в задаче обработки изображений еще в прошлом веке [24]. Но лишь относительно недавно они обрели взрывную популярность, началом которой принято считать демонстрацию значительного превосходства над другими методами в задаче классификации изображений [25]. Значимую роль в этом прогрессе сыграли эффективная реализация обучения на современных мощных графических процессорах и легкий доступ к обилию данных для обучения моделей с большим количеством параметров.

Неудивительно, что и в задаче суперразрешения изображений методы на основе СНН вскоре продемонстрировали превосходные результаты [26], где предложен метод, состоящий из трех этапов:

1. Выделение патчей и их представление

На этом шаге происходит свертка входного изображения низкого разрешения

$$F_1(\mathbf{y}) = \max(0, \mathbf{W}_1 * \mathbf{y} + \mathbf{B}_1),$$

где $\mathbf{W}_1 \in \mathbb{R}^{c \times f_1 \times f_1 \times n_1}$ — фильтр, где c — число каналов на изображении, f_1 — размер фильтра, n_1 — число фильтров, а $\mathbf{B}_1 \in \mathbb{R}^{n_1}$ — вектор сдвига. Операция $*$ — свертка с ядром размера $c \times f_1 \times f_1$, а взятие поэлементного максимума — пространственная в СНН функция активации ReLU [27].

2. Нелинейное преобразование

Преобразует один n_1 -мерный вектор в другой n_2 -мерный вектор нелинейным образом.

$$F_2(\mathbf{y}) = \max(0, \mathbf{W}_2 * F_1(\mathbf{y}) + \mathbf{B}_2),$$

где $\mathbf{W}_2 \in \mathbb{R}^{n_1 \times 1 \times 1 \times n_2}$, $\mathbf{B}_2 \in \mathbb{R}^{n_2}$.

3. Восстановление

В котором из n_2 -мерного вектора строится искомое изображение ВР

$$F(\mathbf{y}) = \mathbf{W}_3 * F_2(\mathbf{y}) + \mathbf{B}_3,$$

где $\mathbf{W}_3 \in \mathbb{R}^{n_2 \times f_2 \times f_2 \times c}$, $\mathbf{B}_3 \in \mathbb{R}^c$.

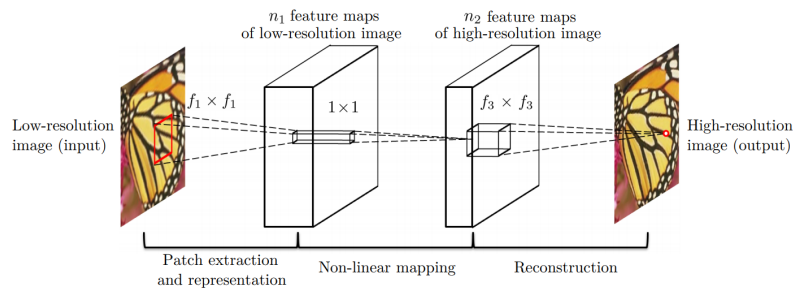


Рис. 3: Схема сверточной нейронной сети для решения задач суперразрешения из работы [26]

Легко видеть, что не смотря на различную мотивацию, все этапы — суть сверточные слои. Этот общий подход позволяет рассматривать их как единую СНН и оптимизировать все этапы одновременно с помощью алгоритма обратного распространения ошибки. На Рис. 3 приведена схема этого метода.

По настоящий момент область применения нейронных сетей в задаче суперразрешения активно развивается. В частности, рассматриваются гораздо более глубокие модели сверточных нейронных сетей [28], а так же используются более современные архитектуры, например, генеративные состязательные сети [29].

3.1. Оптимизация глубоких нейронных сетей

Одна из ключевых задач в тренировке глубоких нейронных сетей — оптимизация их параметров. Вследствие большого совокупного числа параметров на всех слоях нейронной сети применение методов второго порядка (таких как *BFGS*, *SR1* и других квазиньютоновских методов) крайне затруднительно, а методы нулевого порядка в большинстве случаев не позволяют найти качественное решение за приемлемое время. По этой причине для оптимизации глубоких и, в том числе сверточных нейронных сетей, наиболее распространены методы первого порядка, требующие лишь знания градиента сети как функции от ее параметров. Сперва в качестве алгоритма оптимизации глубоких нейронных сетей, в частности в упомянутых работах [25] и [26] использовался градиентный спуск.

Впоследствии было продемонстрировано, что градиентный спуск хоть и не требовательный к вычислительным ресурсам, значительно уступает многим методам первого порядка в скорости сходимости [30].

В последние годы появилось много новых методов для оптимизации глубоких нейронных сетей на основе градиентного спуска, которые хоть и не требуют вычисления Гессиана, или построения его приближения, учитывают информацию более сложную, чем значение градиента в текущей точке: метод момента, быстрый градиент Нестерова, Adam, Adagrad, Adamax, RMSprop и другие (описание и сравнительный анализ этих и других методов приведен в работе [31]). Столь большое число оптимизационных методов, появившихся в столь короткий срок, объясняется, во-первых, сложностью задачи — поверхность целевой функции глубокой нейронной сети зачастую содержит множество локальных минимумов и седловых точек [32], а во-вторых, необходимостью учитывать более сложную и подробную информацию об оптимизируемой функции, чем линейную (градиент функции).

3.2. Проекция–аппроксимация–восстановление

Исходя из приведенных выше соображений, были разработаны методы, основанные на подходе *проекции-аппроксимации-восстановления*. В частности, в работе [33] приводится модификация градиентного спуска, которая учитывает квадратичную составляющую оптимизируемой функции, полученную с помощью ее аппроксимации в пространстве меньшей размерности, а в работе [34] этот подход получает дальнейшее развитие (в частности, используется другой подход для получения матрицы проекции). Далее, в работе [35] предлагается квази-Ньютоновский метод, строящий неявную аппроксимацию Гессиана в пространстве меньшей размерности. Под подходом проекции–аппроксимации–восстановления понимается следующая процедура (далее обозначаемая акронимом *ПАВ*).

1. *Проекция*: проекция нескольких последних точек из оригинального пространства в пространство меньшей размерности посредством умножения на специальным образом построенную прямоугольную матрицу с ортонормальными строками.

2. *Аппроксимация*: построение квадратичного полинома, аппроксимирующего полученные проекции точек и соответствующие им значения целевой функции.
3. *Восстановление*: аппроксимация параметров целевой функции (например, Гессiana, точки минимума) в оригинальном пространстве на основе полученного полинома в пространстве меньшей размерности и прямоугольной матрицы из первого пункта.

В работе [35] предлагаются четыре метода, основанные на подходе ПАВ: PAR1, PAR2, PAR3 и PAR4. Их отличие заключается в различной реализации каждого из трех этапов: проекции, аппроксимации и восстановления. В Таблице 1 приведено краткое пояснение к каждому из четырех методов. Под методами 1 и 2 понимаются (неявная) оценка Гессiana согласно Утверждению 4 и явная оценка точки минимума согласно Утверждению 4 из работы [35]. В той же работе приводится сравнительный анализ указанных алгоритмов вместе с алгоритмами Бройдена–Флетчера–Гольдфарба–Шанно (*BFGS*, [36, 37]), Бройдена–Флетчера–Гольдфарба–Шанно с ограничением по памяти (*L-BFGS*, [38]) и градиентный спуском (*GD*). Результаты приведены в Таблице 2, которая содержит 25%, 50% и 75%-персентили числа итераций для каждого алгоритма. Как можно увидеть, все четыре предложенных алгоритма превосходят метод градиентного спуска в смысле числа операций, особенно хорошие показатели демонстрируют алгоритмы PAR1 и PAR3. С другой стороны, алгоритм *L-BFGS*, с которым и следует сравнивать ПАВ алгоритмы в силу схожих требований к функции и вычислительным ресурсам, превосходит их: алгоритмам PAR1 и PAR3 для сходимости требуется в два раза больше итераций, чем алгоритму *L-BFGS*. Однако, на основе представленных результатов моделирования можно сделать несколько интересных выводов. Во-первых, алгоритмы на основе ПАВ-подхода с матрицей \mathbf{P} , полученной на основе ортогонализации векторов градиентов, демонстрируют лучшее качество, нежели алгоритмы со случайной матрицей \mathbf{P} . Во-вторых, явная оценка точки минимума в алгоритмах PAR3 и PAR4 показывает несколько лучшее качество, нежели подход на основе аппроксимации Гессiana в алгоритмах PAR1 и PAR2. Более того, алгоритм PAR4 не использует информацию о градиенте функции,

Таблица 1: Классификация алгоритмов PAR1, PAR2, PAR3 и PAR4 по способу реализации ПАВ-подхода

Алгоритм	Проекция	Аппроксимация	Восстановление
<i>PAR1</i>	Ортонормализованные градиенты	Квадратичный МНК с фиксированными линейными коэффициентами	Метод 1
<i>PAR2</i>	Случайная матрица проекции	Квадратичный МНК с фиксированными линейными коэффициентами	Метод 1
<i>PAR3</i>	Ортонормализованные градиенты	Квадратичный МНК	Метод 2
<i>PAR4</i>	Случайная матрица проекции	Квадратичный МНК	Метод 2

т. о. являясь алгоритмом нулевого порядка, при этом значительно превосходит показатели градиентного спуска.

4. Заключение

Приведенный обзор методов получения суперразрешения изображений демонстрирует широкий спектр существующих задач и подходов. При условии наличия лишь одного изображения низкого разрешения наиболее перспективными являются подходы, основанные на примерах. В особенности — на основе методов глубокого обучения, для которых выбранный метод оптимизации играет ключевую роль. Предложенные в работах [33] и [35] методы демонстрируют значимое превосходство над градиентным спуском при срав-

Таблица 2: Персентиля числа итераций алгоритмов

Алгоритм	25%	50%	75%
GD	832	1000	1000
$BFGS$	34	56	84
$L - BFGS$	71	96	163
$PAR1$	96	178	391
$PAR2$	122	513	1000
$PAR3$	83	152	238
$PAR4$	126	592	1000

нительно небольших дополнительных вычислениях и линейным по размерности задачи (числу параметров) ограничении по памяти. Эти факты делают возможным успешное применение методов на основе ПАВ подхода для оптимизации глубоких сверточных сетей. В том числе, в контексте решения задачи реконструкции суперразрешения изображений.

Список литературы

- [1] *Park S. C., Park M. K., Kang M. G.* Super-resolution image reconstruction: a technical overview // IEEE signal processing magazine. 2003. Vol. 20. No. 3. P. 21–36.
- [2] *Yang J., Huang T.* Image super-resolution: Historical overview and future challenges // Super resolution imaging. 2010. P. 20–34.
- [3] *Tsai R. Y., Huang T. S.* Multipleframe image restoration and registration // In: Advances in Computer Vision and Image Processing. 1984. P. 317–339.
- [4] *Bose N. K., Kim H. C., Valenzuela H. M.* Recursive implementation of total least squares algorithm for image reconstruction from noisy,

- undersampled multiframe // In: Proc. of the IEEE Conference on Acoustics, Speech and Signal Processing. 1993. Vol. 5. P. 269–272.
- [5] *Tom B. C., Katsaggelos A. K., Galatsanos N. P.* Reconstruction of a high resolution image from registration and restoration of low resolution images // In: Proc. of IEEE International Conference on Image Processing. 1994. P. 553–557.
 - [6] *Zitova B., Flusser J.* Image registration methods: a survey // Image and Vision Computing, 2003. Vol. 21. No. 11. P. 977–1000.
 - [7] *Keren D., Peleg S., Brada R.* Image sequence enhancement using subpixel displacements // In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. 1988. P. 742–746.
 - [8] *Ur H., Gross D.* Improved resolution from subpixel shifted pictures // CVGIP: Graphical Models and Image Processing. 1992. Vol. 54. No. 2. P. 181–186.
 - [9] *Alam M. S., et al* Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames // IEEE Transactions on Instrumentation and Measurement. 2000. Vol. 49. No. 5. P. 915–923.
 - [10] *Irani M., Peleg S.* Motion analysis for image enhancement: resolution, occlusion and transparency // Journal of Visual Communications and Image Representation. 1993. Vol. 4. No. 4. P. 324–335.
 - [11] *Capel D.* Image Mosaicing and Super-resolution. Springer, 2004.
 - [12] *Capel D., Zisserman A.* Computer vision applied to superresolution // IEEE Signal Processing Magazine. 2003. Vol. 20. No. 3. P. 75–86.
 - [13] *Rudin L., Osher S., Fatemi E.* Nonlinear total variation based noise removal algorithms // Physica D: Nonlinear Phenomena. 1992. Vol. 60. No. 1–4. P. 259–268.
 - [14] *Hardie R. C., Bernard K. J., Armstrong E. E.* Joint MAP registration and high resolution image estimation using a sequence of undersampled images // IEEE Transactions on Image Processing. 1997. Vol 6. No. 12. P. 1621–1633.

- [15] *Shen H., Zhang L., Huang B., Li .P.* A MAP approach for joint motion estimation, segmentation and super resolution // IEEE Transactions on Image Processing. 2007. Vol. 16. No. 2. P. 479–490.
- [16] *Woods N. A., Galatsanos N. P., Katsaggelos A. K.* Stochastic methods for joint registration, restoration and interpolation of multiple undersampled images // IEEE Transactions on Image Processing. 2006. Vol. 15. No. 1. P. 210–213.
- [17] *Freeman W. T., Jones T. R., Pasztor E. C.* Example-based super-resolution // IEEE Computer Graphics and Applications. 2002. Vol. 22. No. 2. P. 56–65.
- [18] *Chang H., Yeung D. Y., Xiong Y.* Super-resolution through neighbor embedding // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. 2004. Vol. 1. P. 275–282.
- [19] *Yang J. et al.* Image super-resolution as sparse representation of raw image patches // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2008. P. 1–8.
- [20] *Capel D., Zisserman A.* Super-resolution from multiple views using learnt image models // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2001. Vol. 2. P. 1–8.
- [21] *Jolliffe I. T.* Principal Component Analysis. Series: Springer Series in Statistics. 2002.
- [22] *Liu C., Shum H. Y., Zhang C. S.* A two-step approach to hallucinating faces: global parametric model and local nonparametric model // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2001. Vol. 1. P. 192–198.
- [23] *Huang J. B., Singh A., Ahuja N.* Single image super-resolution from transformed self-exemplars // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 5197–5206.
- [24] *LeCun Y. et al.* Backpropagation applied to handwritten zip code recognition // Neural computation. 1989. Vol. 1. No. 4. P. 541–551.
- [25] *Krizhevsky A., Sutskever I., Hinton G. E.* Imagenet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems. 2012. P. 1097–1105.

- [26] *Dong C. et al.* Learning a deep convolutional network for image super-resolution // European Conference on Computer Vision. Springer, Cham. 2014. P. 184–199.
- [27] *Nair V., Hinton G. E.* Rectified linear units improve restricted boltzmann machines // In: Proc. of the 27th International Conference on Machine Learning. 2010. P. 807–814.
- [28] *Kim J., Kwon L. J., Mu L. K.* Accurate image super-resolution using very deep convolutional networks // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. P. 1646–1654.
- [29] *Ledig C. et al.* Photo-realistic single image super-resolution using a generative adversarial network // In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. P. 4681–4690.
- [30] *Bengio Y., Boulanger-Lewandowski N., Pascanu R.* Advances in optimizing recurrent networks // In: Proc. of IEEE Conference on Acoustics, Speech and Signal Processing. 2013. P. 8624–8628.
- [31] *Ruder S.* An overview of gradient descent optimization algorithms. ArXiv preprint arXiv:1609.04747. 2016.
- [32] *Goodfellow I., Bengio Y., Courville A.* Deep Learning. — Cambridge: MIT press, 2016.
- [33] *Senov A.* Accelerating gradient descent with projective response surface methodology // In: Proc of International Conference on Learning and Intelligent Optimization. 2017. P. 376–382.
- [34] *Senov A., Granichin O.* Projective Approximation Based Gradient Descent Modification // In Proc. of 20th World Congress of the International Federation of Automatic Control. 2017. Vol. 50. No. 1, P. 3899–3904.
- [35] *Senov A.* Projective approximation based quasi-Newton methods // In: Proc. of International Workshop on Machine Learning, Optimization, and Big Data. 2017. P. 29–40.
- [36] *Broyden C. G.* The convergence of a class of double-rank minimization algorithms. 1. General considerations // IMA Journal of Applied Mathematics. 1970. Vol. 6. No. 1. P. 76–90.

- [37] *Fletcher R.* Practical Methods of Optimization (Second ed.). New Your: John Wiley, 1987.
- [38] *Nocedal J.* Updating quasi-newton matrices with limited storage // Mathematics of Computation. 1980. Vol. 35. No. 151. P. 773–782.