

Изучение эффективности восстановления изображений в методе однопиксельной визуализации с помощью генеративных состязательных сетей

Д.В. Бабухин¹, А.А. Реутов¹, Д.В. Сыч¹

¹ Физический институт им. П.Н. Лебедева Российской академии наук,
119991, Россия, Москва, Ленинский пр-т, д. 53

Аннотация

Однопиксельная визуализация является перспективным методом получения изображений, представляющим альтернативу традиционным методам визуализации с помощью многопиксельных матриц. Однако алгоритмическое восстановление изображения из измерений однопиксельной камеры является нетривиальной вычислительной задачей, для решения которой недавно стали применяться методы машинного обучения. В данной работе исследовалась возможность восстановления изображения в методе однопиксельной визуализации с помощью генеративных состязательных нейронных сетей. Для этого с использованием компьютерного моделирования однопиксельной камеры оценивалась эффективность восстановления изображений при помощи двух архитектур генеративных сетей – глубокой сверточной генеративной состязательной сети и генеративной состязательной сети наименьших квадратов. Было установлено, что генеративная состязательная сеть наименьших квадратов демонстрирует лучшее качество восстановления изображения по сравнению с глубокой сверточной генеративной состязательной сетью, однако при учете оптических искажений глубокая сверточная состязательная сеть стабильнее обучается до более высокого качества по сравнению с генеративной состязательной сетью наименьших квадратов. Полученные в работе результаты могут послужить основой для создания программного обеспечения, требуемого при практическом применении однопиксельной камеры.

Ключевые слова: однопиксельная визуализация, восстановление изображений, генеративные состязательные сети, коррекция аппаратных искажений.

Цитирование: Бабухин, Д.В. Изучение эффективности восстановления изображений в методе однопиксельной визуализации с помощью генеративных состязательных сетей / Д.В. Бабухин, А.А. Реутов, Д.В. Сыч // Компьютерная оптика. – 2025. – Т. 49, № 5. – С. 818-825. – DOI: 10.18287/2412-6179-CO-1526.

Citation: Babukhin DV, Reutov AA, Sych DV. Study of image reconstruction efficiency in single-pixel imaging method using generative adversarial networks. Computer Optics 2025; 49(5): 818-825. DOI: 10.18287/2412-6179-CO-1526.

Введение

Наиболее распространённые сегодня методы формирования изображений используют многопиксельные матричные фотосенсоры, что позволяет получить требуемое пространственное разрешение целевого объекта. Такие методы лежат в основе многочисленных фото- и видеоустройств, от мобильных телефонов и цифровых камер до специализированного профессионального оборудования. Несмотря на превосходные характеристики матричных сенсоров, их использование в нестандартных режимах работы сталкивается с серьезными трудностями. Например, технологически сложно создать сенсор с гибко настраиваемым спектральным диапазоном вне диапазона чувствительности кремния или же добавить к высокому пространственному разрешению высокое временное разрешение.

Альтернативным способом получения изображений является метод однопиксельной визуализации. Идея метода заключается в освещении объекта про-

странственно-модулированным светом (световыми паттернами), измерении общего количества света, отраженного от объекта, и вычислении пространственного изображения объекта по полученной таким образом информации [1]. Пространственная модуляция света может быть реализована в широком спектральном диапазоне, например с помощью цифровых микрозеркальных модуляторов света [2, 3]. Однопиксельная визуализация открывает качественно новые возможности получения изображений. Например, можно получать изображения в диапазоне длины волны 1,5 микрона на уровне одиночных фотонов [4], мультиспектральные изображения [5], трехмерные изображения [6, 7], рентгеновские изображения [8, 9].

Вычислительная часть метода однопиксельной визуализации заключается в решении системы уравнений, связывающих измеренные сигналы однопиксельного детектора и световые паттерны. В последнее время методы машинного обучения продемонстрировали большой потенциал в решении этой задачи [10 – 12]. Нейросетевые методы оказались успешны в за-

дачах, близких задаче восстановления изображения, как например: в задаче классификации изображений [13], оптимизации световых паттернов для однопиксельной визуализации [14], определении расстояния до объекта [15].

Недавно было показано успешное применение концепции генеративных состязательных сетей в методе однопиксельной визуализации [16–18]. Однако вопрос построения оптимальной архитектуры нейронной сети для этой цели остается открытым, так как сравнения разных типов генеративных состязательных сетей и качества работы генеративных состязательных сетей при наличии аппаратных искажений в однопиксельной визуализации до настоящего времени не проводилось. В данной работе мы сравниваем качество восстановления изображения в однопиксельной визуализации, осуществленного с помощью двух видов генеративных состязательных сетей – глубокой сверточной генеративной состязательной сети и генеративной состязательной сети наименьших квадратов. Нами установлено, что генеративная состязательная сеть наименьших квадратов демонстрирует несколько лучшее качество восстановления изображения по сравнению с глубокой сверточной генеративной состязательной сетью. В то же время при наличии аппаратных искажений в виде изменения поля зрения камеры глубокая сверточная сеть стабильно обучается восстановлению изображений с более высоким качеством, чем генеративная состязательная сеть наименьших квадратов.

Статья организована следующим образом: в параграфе 1 приводятся основные сведения о методе однопиксельной визуализации, в параграфе 2 приводится описание процесса обучения анализируемых генеративных состязательных сетей, в параграфе 3 приводится описание процедуры генерации паттернов для однопиксельной камеры, в параграфах 4 и 5 приводятся результаты моделирования работы однопиксельной камеры и восстановления изображений при отсутствии (параграф 4) и при наличии (параграф 5) аппаратных искажений поля зрения камеры.

1. Однопиксельная визуализация

Получение изображения методом однопиксельной визуализации можно условно разделить на два этапа. Первый этап однопиксельной визуализации заключается в детектировании пространственно модулированного света, отраженного от объекта, с помощью однопиксельного детектора, измеряющего интегральную интенсивность света без пространственного разрешения. Допустим, мы хотим построить изображение объекта шириной W и высотой H пикселей. Свет от источника (непрерывного или импульсного) направляется на цифровое микрзеркальное устройство, которое проецирует матричный паттерн $\vec{P}^i = (p_1^i, \dots, p_{WH}^i)$ для пространственной модуляции отраженного света. Отраженный от объекта свет со-

бирается посредством линзы на детекторе, который измеряет интегральное значение интенсивности отраженного света s_i для выбранного паттерна. Эту процедуру можно провести для M различных матричных паттернов и записать уравнение

$$S = PI, \quad (1)$$

где в векторной форме представлены изображение целевого объекта $I = (i_1, \dots, i_{WH})$, матрица паттернов $P = (\vec{P}^1, \dots, \vec{P}^M)^T$ и набор интенсивностей, измеренных для M паттернов $S = (s_1, \dots, s_M)$. Отношение количества использованных паттернов M к размеру матрицы паттерна WH называется частотой выборки $SR = M / WH$ (sampling rate).

Вторым этапом однопиксельной визуализации является восстановление целевого изображения I по значениям интенсивностей S и информации об использованных паттернах P . В случае, когда для получения изображения используется столько паттернов, сколько пикселей в изображении ($M = WH$, $SR = 1$), результирующее изображение I является решением системы линейных уравнений

$$I = P^{-1}S. \quad (2)$$

На практике число паттернов может быть не равно числу пикселей. С одной стороны, наличие шума в сигнале требует увеличения числа паттернов. С другой стороны, реальные изображения имеют некоторую избыточность в своей структуре, поэтому могут быть восстановлены с меньшим числом паттернов методами сжатого сэмпирования или машинного обучения.

2. Генеративные состязательные сети

Генеративная состязательная сеть представляет собой нейронную сеть, состоящую из двух компонент – сети-генератора $G(z, \theta_G)$ и сети-дискриминатора $D(x, \theta_D)$. Здесь z – входной вектор (в нашем случае – вектор интенсивностей, полученный в результате работы однопиксельной камеры), из которого генератор создаёт изображение, x – изображение (реальное или сгенерированное), которое классифицирует дискриминатор, а $\theta_{D(G)}$ – набор настраиваемых параметров дискриминатора (генератора). Обе нейросети в нашем случае представляют собой сверточные нейросети с архитектурой, соответствующей требуемым выходным данным: на выходе генератора получается изображение размером WH , а на выходе дискриминатора получается вещественное число в диапазоне между 0 и 1, характеризующее вероятность того, что поданное на вход изображение не было сгенерировано сетью-генератором.

Схематично процесс обучения нейросетей имеет следующий вид. Пусть имеется обучающая выборка изображений («истинные» изображения). Для каждого изображения мы получаем вектор отсчётов интенсивностей посредством симуляции работы однопиксельной

камеры (см. параграф 1), и для каждого такого вектора сеть-генератор восстанавливает изображение объекта («сгенерированные» изображения). В процессе обучения сети-дискриминатора на вход дискриминатора попеременно подаются истинные и сгенерированные изображения с соответствующей меткой (1 для истинных и 0 для сгенерированных изображений) и дискри-

минатор обучается правильной разметке подаваемых изображений. В процессе обучения сети-генератора сгенерированные сетью-генератором изображения подаются на вход сети-дискриминатора. Сеть-генератор настраивается в сторону увеличения вероятности присвоения метки 1 («истинное» изображение) сгенерированному изображению.

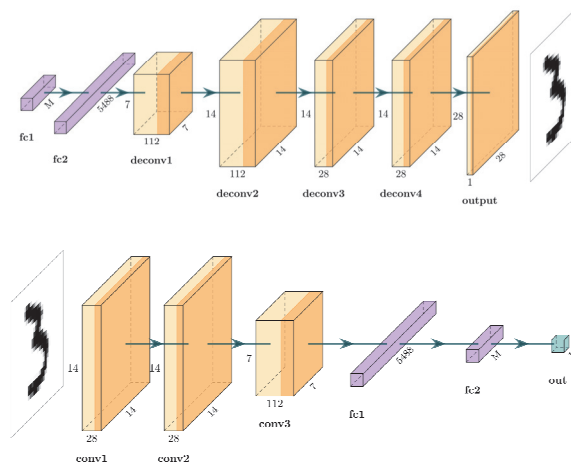


Рис. 1. Схема архитектуры нейронной сети: сеть-генератор (сверху) и сеть-дискриминатор (снизу). На вход сети-генератора подаётся вектор отсчётов интенсивностей размером M , полученный в результате работы однопиксельной камеры. На выходе сети-генератора появляется восстановленное изображение объекта. На вход сети-дискриминатора подаётся изображение объекта – сгенерированное сетью-генератором или истинное. На выходе сети-дискриминатора появляется число от 0 до 1, характеризующее вероятностью поданного на вход изображения быть истинным (не сгенерированным сетью-генератором)

Процесс обучения генеративной состязательной сети заключается в попеременной оптимизации параметров двух нейросетей с соответствующими функциями ошибок. Первая стадия обучения – обучение дискриминатора D , при котором дискриминатору демонстрируются изображения из обучающей выборки («истинные» изображения), а также сгенерированные сетью G изображения. Задачей дискриминатора является разметка изображений на истинные и сгенерированные. Функцией ошибки дискриминатора для глубокой сверточной генеративной состязательной сети является бинарная кросс-энтропия [19]

$$L_D^{DC}(D) = \mathbb{E}_x[\log(D(x))] + \mathbb{E}_z[\log(1 - D(G(z)))], \quad (3)$$

а для генеративной состязательной сети наименьших квадратов [20]

$$L_D^{LS}(D) = \frac{1}{2} \mathbb{E}_x[(D(x) - 1)^2] + \frac{1}{2} \mathbb{E}_z[(D(G(z)))^2], \quad (4)$$

где усреднение производится по подвыборке истинных изображений x или по подвыборке векторов отсчётов z . Вторая стадия обучения – обучение генератора G . На этой стадии сгенерированные генератором изображения подаются на вход дискриминатору попеременно с реальными изображениями, а дискриминатор (с фиксированными на этой стадии параметрами) размечает изображения на реальные и сгенериро-

ванные. Задачей генератора является научиться генерировать изображения таким образом, чтобы генератор не мог достоверно отличить сгенерированные изображения от реальных. Функция ошибки для генератора в глубокой сверточной генеративной состязательной сети [19] имеет вид

$$L_G^{DC}(G) = \mathbb{E}_z[\log(D(G(z)))], \quad (5)$$

а для генеративной состязательной сети наименьших квадратов [20]

$$L_G^{LS}(G) = \frac{1}{2} \mathbb{E}_z[(D(G(z)) - 1)^2]. \quad (6)$$

На первой стадии обучения минимизируется функция $L_D^{DC(LS)}(D)$ и настраиваются параметры сети-дискриминатора D , тогда как на второй стадии обучения минимизируется функция $L_G^{DC(LS)}(G)$ и настраиваются параметры сети-генератора G . В результате обучения получается сеть-генератор, способная восстанавливать изображение из вектора интенсивностей, полученного с помощью съемки целевого объекта посредством однопиксельной камеры. Обученная сеть способна восстанавливать изображения объектов, не присутствующих в обучающей выборке.

Помимо конкурентной составляющей процесса обучения дискриминатора и генератора в функции ошибок генератора входят дополнительные компо-

ненты, позволяющие генератору создавать более качественные изображения. Такие дополнительные компоненты функции ошибок называются в литературе функциями ошибок содержания (content loss). В данной работе используются две дополнительные функции ошибок: функция ошибок L_1

$$L_1 = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H |x_{i,j} - G(z)_{i,j}|, \quad (7)$$

где $x_{i,j}$ – компоненты истинного изображения и $G(z)_{i,j}$ – компоненты восстановленного генератором изображения, и функция ошибок, основанная на нейросети VGG19 (perceptual similarity loss) [21]:

$$L_{VGG} = \frac{1}{W_k H_k} \sum_{i=1}^{W_k} \sum_{j=1}^{H_k} (VGG_k(x)_{i,j} - VGG_k(G(z))_{i,j})^2, \quad (8)$$

где VGG_k – k -й слой нейронной сети VGG19, и $VGG_k(x)_{i,j}$ – выход k -го слоя VGG19 для поданного на вход изображения x , W_k и H_k – ширина и высота результата на выходе k -го слоя VGG19.

3. Генерация паттернов

Важным вопросом в задаче однопиксельной визуализации является выбор паттернов, поскольку он напрямую влияет на её быстродействие и качество. С точки зрения скорости визуализации выбор бинарных паттернов представляется естественным, поскольку позволяет использовать каждое микрозеркальце цифрового микрозеркального устройства как один пиксель паттерна. В то же время выбор бинарных матриц, обеспечивающих наилучшее качество визуализации, неоднозначен. Например, использование случайных бинарных паттернов с фиксированным соотношением 0 и 1 вместе со сжатым сэмплированием требует большого (не менее 300 для изображения с 784 пикселями) количества паттернов для достижения приемлемого качества восстановленного изображения [22]. В то же время использование таких паттернов вместе со свёрточной нейронной сетью для классификации изображений демонстрирует хорошие результаты [23]. Таким образом, один и тот же способ генерации паттернов может показывать разную эффективность в разных задачах.

Использование нейронных сетей в однопиксельной визуализации позволяет (помимо восстановления изображений) генерировать оптимальные (для конкретной задачи) паттерны. Так, в работе [14] свёрточный автокодировщик использовался для одновременного восстановления изображений из значений интенсивностей, измеренных однопиксельной камерой, и для генерации паттернов, оптимальных для работы камеры.

В данной работе мы используем набор паттернов, сгенерированных свёрточным автокодировщиком, обученным восстановлению изображений из набора MNIST. Кодировочная часть свёрточного автокоди-

ровщика состоит из одного полносвязного слоя, переводящего N (784 в нашем случае) пикселей изображения в M (64 в нашем случае) переменных скрытого слоя; такая архитектура была выбрана для того, чтобы матрица весов кодировщика по размеру соответствовала матрице паттернов P .

Процесс обучения свёрточного автокодировщика состоял из трёх этапов:

- на первом этапе кодировщик и декодировщик обучались совместно без каких-либо ограничений. Обучение было прервано на моменте, когда значения функции потерь перестали значимо изменяться для обучающего и тестового набора;
- на втором этапе был введен дополнительный член функции потерь вида:

$$L_{\text{norm}} = R * \sum_m \sum_n w_{mn}^2 * (1 - w_{mn})^2, \quad (9)$$

где R – константа, w_{mn} – веса кодировщика. Этот член «штрафует» нейронную сеть, если значения весов кодировщика слишком сильно отличаются от 0 или 1. В процессе дообучения свёрточного автокодировщика константа R увеличивалась несколько раз для достижения различия весов кодировщика w_{mn} от нуля (либо единицы) не более чем 10^{-3} ;

- на третьем этапе веса кодировщика были зафиксированы, округлены до целого значения и использованы как матрица паттернов P .

Таким образом, с помощью весов кодировщика была получена бинарная матрица паттернов P для эффективного кодирования изображений в сигналы S .

4. Восстановление изображений при отсутствии аппаратных искажений

Обучение генеративных состязательных сетей осуществлялось на наборе изображений MNIST, состоящем из 60000 изображений в обучающей выборке и 10000 изображений в тестовой выборке. Задачей генератора являлось восстановление изображения цифры от 0 до 9 из симулированных результатов работы однопиксельной камеры, а задачей дискриминатора – разметка изображений на реальные и сгенерированные. Генератор состоял из 2 входных линейных полносвязных слоёв и 4 слоёв обратной свертки с функциями активации $\text{ReLU } f(x) = \max(0, x)$ на 3 внутренних слоях и с сигмоидной выходной функцией $f(x) = 1 / (1 + e^{-x})$ на последнем. Дискриминатор состоял из 3 свёрточных слоёв с функциями активации $\text{Leaky ReLU } f(x) = \max(0, 0.1, x)$ и 2 выходных линейных полносвязных слоёв с сигмоидными функциями активации. Минимизация функций ошибок осуществлялась оптимизатором RMSProp с начальной скоростью обучения $lr = 4 \times 10^{-4}$ для генератора и дискриминатора, а также параметром $\alpha = 0,99$. Обучение длилось в течение 50 эпох, при этом каждые 5 эпох скорость обучения для обеих компонент сети падает в

2 раза. Обучающая выборка разбивалась на подвыборки (батчи) по 30 изображений, и каждые 10 эпох размер подвыборки увеличивался вдвое. Для оценивания качества восстановления изображений сетью-генератором использовались метрики Peak Signal-to-Noise Ratio (PSNR)

$$PSNR(x, \hat{x}) = 10 \log_{10} \left(\frac{Max(x)}{MSE(x, \hat{x})} \right) \quad (10)$$

и Structural Similarity Index Measure (SSIM)

$$SSIM(x, \hat{x}) = \frac{(2\mu_x \mu_{\hat{x}} + c_1)(2\sigma_{xx} + c_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2)}, \quad (11)$$

где x – истинное изображение, $\hat{x} = G(z)$ – изображение, восстановленное сетью-генератором из вектора интенсивностей z , $\mu_{x(x)}$ – среднее значение матричных элементов в изображении $x(x)$, $\sigma_{x(x)}$ – дисперсия значений матричных элементов в изображении $x(x)$, а σ_{xx} – коэффициент корреляции между изображениями x и $\hat{x}$.

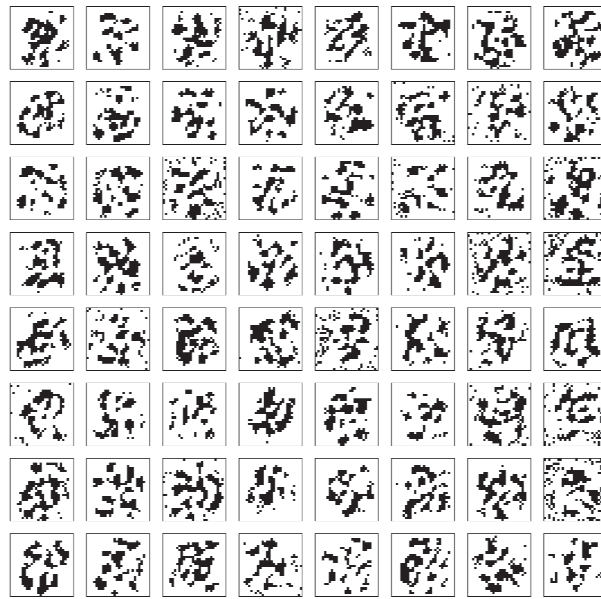


Рис. 2. Используемый в данной работе набор паттернов, который был сгенерирован сверточным автокодировщиком

Две генеративные состязательные сети были обучены на бесшумовых изображениях из набора MNIST для различных (1 %, 2 %, 4 % и 8 %) частот выборки однопиксельной визуализации. Генераторы обученных сетей восстанавливали изображения из тестовой части набора, и вычислялись средние значения метрик PSNR и SSIM на тестовых изображе-

ниях. Графики средних значений метрик от частоты выборки представлены на рис. 3. Полученные зависимости показывают, что генеративная состязательная сеть наименьших квадратов достигает несколько лучшего по сравнению с глубокой сверточной генеративной состязательной сетью качества восстановления изображений.

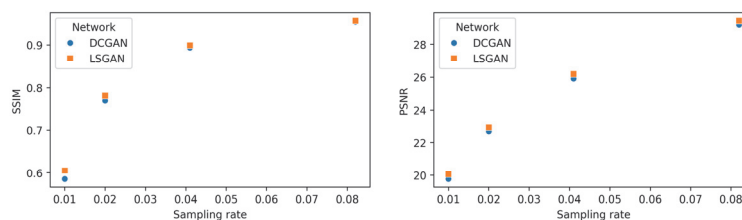


Рис. 3. Зависимости средних значений метрик PSNR и SSIM от частоты выборки однопиксельной камеры для изображений, восстановленных генераторами обученных генеративно-состязательных сетей

На рис. 4 представлены примеры восстановления изображений цифр генераторами двух нейронных сетей, а также истинное изображение цифры (до выборки с помощью однопиксельной камеры). Представлены примеры трёх цифр для демонстрации того, что качество восстановления может быть различным для различных цифр: цифра может восстанавливаться хорошо вплоть до низких частот выборки (цифра «9» на рис. 4) или же значительно искажаться при низких

частотах выборки (цифра «5» на рис. 4). В то же время при частоте выборки 8 % все цифры восстанавливаются без существенных искажений.

5. Восстановление изображений при наличии аппаратных искажений

Генеративные состязательные сети двух типов были обучены на изображениях из набора MNIST, содержащих аппаратные искажения. В качестве аппаратного ис-

кажения рассматривалась тёмная рамка вокруг изображения, соответствующая неточности в расположении визуализируемого объекта, приводящей к излишне большому полю зрения объектива камеры. Для набора статистики было обучено 10 генеративных состязатель-

ных сетей каждого вида со случайно инициализированными параметрами. Обучение каждой сети длилось в течение 10 эпох, поскольку при рассматриваемых параметрах симуляции сети достигали своего наилучшего обученного состояния в течение 10 эпох.

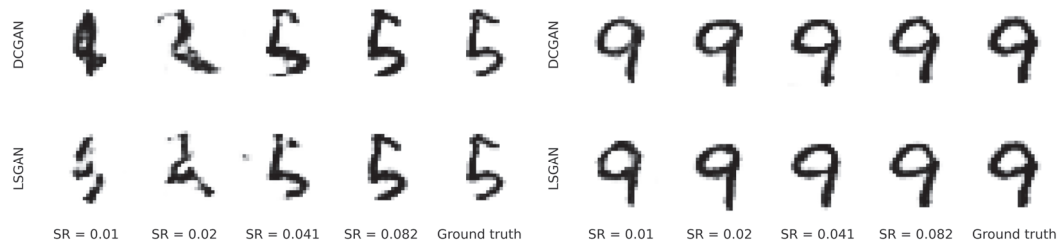


Рис. 4. Примеры цифр, восстановленных генераторами обученных генеративных состязательных сетей при различных частотах выборки, а также истинные изображения цифр

На рис. 5 представлены значения метрик PSNR и SSIM для 10 генеративных состязательных сетей каждого вида, обученных на восстановление изображений с однопиксельной камеры при частоте выборки 8 %. Значения метрик вычислялись на 10000 тестовых изображениях цифр. Из полученных результатов видно, что после добавления аппаратных шумов глубокая сверточная генеративная состязательная сеть обучается восстановлению изображений стабильно лучше, чем генеративная состязательная сеть наименьших квадратов. Несмотря на то, что сеть наименьших квадратов способна обучиться до высокого качества (сеть LS10 на рис. 5), при других инициализациях начальных параметров процесс обучения может закончиться со значительно более низким качеством восстановления изображений (сети LS8 и LS9). Возможной причиной такого поведения является различие в свойствах сходимости конкурентной части функции ошибок двух сетей: в случае глубокой сверточной генеративной состязательной сети конкурентной частью функции ошибок является бинарная кросс-энтропия (формулы (3), (5)), а для генеративной состязательной сети наименьших квадратов – квадратичная ошибка (формулы (4), (6)). Известно, что бинарная кросс-энтропия имеет пороговое время обучения, при котором она достигает своего насыщения и перестаёт влиять на дальнейшее изменение параметров сетей, тогда как квадратичная функция ошибки даёт постепенно убывающий, но не нулевой вклад в градиенты параметров. Это может означать, что решающий вклад в качество восстановления изображений имеет именно часть функции ошибок, отвечающая за содержание изображения (content loss).

На рис. 6 представлены изображения цифр, восстановленные лучшими генераторами (по метрикам PSNR и SSIM, см. рис. 5), а также целевые изображения цифр и изображения цифр с рамкой. Видно, что оба лучших генератора дают хорошее качество восстановления цифр.

Заключение

В данной работе мы исследовали применение генеративных состязательных сетей для восстановления изображений, получаемых с однопиксельной камеры. В частности, мы сравнили работу двух видов генеративных состязательных сетей – глубокой сверточной генеративной состязательной сети и генеративной состязательной сети наименьших квадратов. Было установлено, что при отсутствии аппаратных искажений генеративная состязательная сеть наименьших квадратов (по сравнению с глубокой сверточной генеративной состязательной сетью) достигает несколько лучшего качества восстановления изображения. Однако при наличии аппаратных искажений глубокая сверточная генеративная состязательная сеть даёт стабильно лучшее качество восстановления изображения по сравнению с генеративной состязательной сетью наименьших квадратов. Мы связываем это с различными свойствами сходимости конкурентных частей функций ошибок двух видов генеративных сетей.

Применение нейронных сетей для задачи однопиксельной визуализации является перспективным и новым направлением. Хотя успешность использования нейронных сетей в однопиксельной визуализации уже продемонстрирована в работах последних лет, область содержит множество направлений для исследований. Во-первых, по мере развития алгоритмов глубокого обучения возникают всё более эффективные подходы к решению различных задач (например, генерации изображений и улучшения их качества). Поэтому использование и оценка возникающих новых нейросетевых алгоритмов для задачи однопиксельной визуализации представляет собой отдельное направление работ. Во-вторых, поскольку целью развития однопиксельной визуализации является создание новых устройств, позволяющих получать изображения объектов в условиях, недоступных традиционным многопиксельным камерам, важно учитывать требования к производительности нейросетевых алгоритмов с прикладной точки зрения. Поиск баланса

между качеством получаемых изображений и накладными расходами – временными затратами на обучение и объемом памяти, необходимым для запуска

программных приложений, обеспечивающих восстановление изображения, – представляет собой отдельное направление будущих исследований.

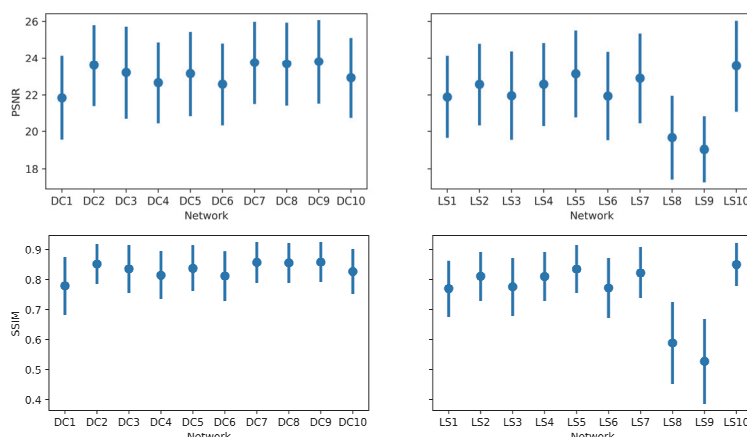


Рис. 5. Значения метрик PSNR и SSIM для 10 различных генераторов двух типов, обученных на восстановление изображений с однопиксельной камеры при частоте выборки 8%. Каждое значение вычислялось на тестовом наборе из 10000 изображений цифр. Обозначение DCN соответствует N-й глубокой сверточной генеративной состязательной сети, а обозначение LSN – N-й генеративной состязательной сети наименьших квадратов

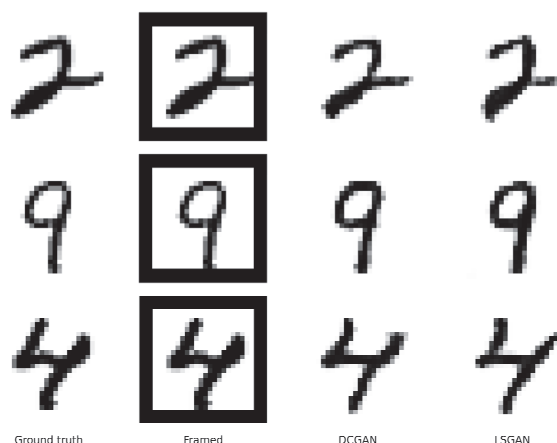


Рис. 6. Примеры изображений цифр, восстановленных с помощью лучших (по метрикам PSNR и SSIM) генераторов (см. рис. 5)

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 23-22-00381, <https://rscf.ru/project/23-22-00381/>.

References

- [1] Duarte MF, Davenport MA, Takhar D, Laska JN, Sun T, Kelly KF, Baraniuk RG. Single-pixel imaging via compressive sampling. IEEE Signal Process Mag 2008; 25(2): 83-91. DOI: 10.1109/MSP.2007.914730.
- [2] Aksenov MD, Sych DV. Optimal data acquisition methods for single-pixel imaging. J Russ Laser Res 2018; 39: 492-498. DOI: 10.1007/s10946-018-9744-y.
- [3] Sych D, Aksenov M. Computational imaging with a single-pixel detector and a consumer video projector. AIP Conf Proc 2018; 1936(1): 020016. DOI: 10.1063/1.5025454.
- [4] Shcherbatenko M, Elezov M, Manova N, Sedykh K, Korneev A, Korneeva Yu, Dryazgov M, Simonov N, Feimov A, Goltsman G, Sych D. Single-pixel camera with a large-area microstrip superconducting single photon detector on a multimode fiber. Appl Phys Lett 2021; 118(18), 181103. DOI: 10.1063/5.0046049.
- [5] Bian L, Suo J, Situ G, Li Z, Fan J, Chen F, Dai Q. Multi-spectral imaging using a single bucket detector. Sci Rep 2016; 6: 24752. DOI: 10.1038/srep24752.
- [6] Zhang Z, Liu S, Peng J, Yao M, Zheng G, Zhong J. Simultaneous spatial, spectral, and 3D compressive imaging via efficient Fourier single-pixel measurements. Optica 2018; 5(3): 315-319. DOI: 10.1364/OPTICA.5.000315.
- [7] Zhang Y, Edgar MP, Sun B, Radwell N, Gibson GM, Padgett MJ. 3D single-pixel video. J Opt 2016; 18(3): 035203. DOI: 10.1088/2040-8978/18/3/035203.
- [8] Greenberg J, Krishnamurthy K, Brady D. Compressive single-pixel snapshot x-ray diffraction imaging. Opt Lett 2013; 39(1): 111-114. DOI: 10.1364/OL.39.000111.
- [9] Olbinado MP, Paganin DM, Cheng Y, Rack A. X-ray phase-contrast ghost imaging using a single-pixel camera. Optica 2021; 8(12): 1538-1544. DOI: 10.1364/OPTICA.437481.
- [10] Lyu M, Wang W, Wang H, Wang H, Li G, Chen N, Situ G. Deep-learning-based ghost imaging. Sci Rep 2017; 7: 17865. DOI: 10.1038/s41598-017-18171-7.

- [11] He Y, Wang G, Dong G, Zhu S, Chen H, Zhang A, Xu Z. Ghost imaging based on deep learning. *Sci Rep* 2018; 8: 6469. DOI: 10.1038/s41598-018-24731-2.
- [12] Shimobaba T, Endo Y, Nishitsuji T, Takahashi T, Nagahama Y, Hasegawa S, Sano M, Hirayama R, Kakue T, Shiraki A, Ito T. Computational ghost imaging using deep learning. *Opt Commun* 2018; 413: 147-151. DOI: 10.1016/j.optcom.2017.12.041.
- [13] Zhang Z, Li X, Zheng S, Yao M, Zheng G, Zhong J. Image-free classification of fast-moving objects using “learned” structured illumination and single-pixel detection. *Opt Express* 2020; 28(9): 13269-13278. DOI: 10.1364/OE.392370.
- [14] Higham CF, Murray-Smith R, Padgett MJ, Edgar MP. Deep learning for real-time single-pixel video. *Sci Rep* 2018; 8: 2369. DOI: 10.1038/s41598-018-20521-y.
- [15] Radwell N, Johnson SD, Edgar MP, Higham CF, Murray-Smith R, Padgett MJ. Deep learning optimized single-pixel LiDAR. *Appl Phys Lett* 2019; 115(23): 231101. DOI: 10.1063/1.5128621.
- [16] Zhao M, Li F, Huo F, Tian Z. Generative adversarial network-based single-pixel imaging. *Soc for Inf Disp* 2022; 30(8): 648-656. DOI: 10.1002/jsid.1128.
- [17] Sun S, Yan Q, Zheng Y, Wei Z, Lin J, Cai Y. Single pixel imaging based on generative adversarial network optimized with multiple prior information. *IEEE Photonics J* 2022; 14(4): 8538110. DOI: 10.1109/JPHOT.2022.3184947.
- [18] Lu Y, Tan B, Ding S, Li Y. SPI-CGAN: Single-pixel imaging based on conditional generative adversarial network. *J Phys: Conf Ser* 2022; 2216: 012092. DOI: 10.1088/1742-6596/2216/1/012092.
- [19] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv Preprint*. 2015. Source: <<https://arxiv.org/abs/1511.06434>>. DOI: 10.48550/arXiv.1511.06434.
- [20] Mao X, et al. Least squares generative adversarial networks. 2017 IEEE Int Conf on Computer Vision (ICCV)2017: 2813-2821. DOI: 10.1109/ICCV.2017.304.
- [21] Ledig C, et al. Photo-realistic single image super-resolution using a generative adversarial network. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 105-114. DOI: 10.1109/CVPR.2017.19.
- [22] Sych D. Optimization of compressed sampling in single-pixel imaging. *Bull Lebedev Phys Inst*. 23024; 51: 202-205. DOI: 10.3103/S1068335624600463.
- [23] Reutov AA, Babukhin DV, Sych DV. Object classification using a single-pixel camera and neural networks. *Computer Optics* 2025; 49(3): 517-524. DOI: 10.18287/2412-6179-CO-1538.

Сведения об авторах

Бабухин Данила Валерьевич, 1995 года рождения, окончил Московский государственный университет имени М.В. Ломоносова, младший научный сотрудник в Физическом институте им. П.Н. Лебедева РАН. Область научных интересов: квантовая связь, однопиксельная визуализация, машинное обучение в задачах физики. E-mail: dv.babukhin@gmail.com

Реутов Алексей Алексеевич, 1995 года рождения, окончил Московский физико-технический институт, младший научный сотрудник в Физическом институте им. П.Н. Лебедева РАН. Область научных интересов: квантовая связь, однопиксельная визуализация, машинное обучение в задачах физики. E-mail: aleksey.reutov@phystech.edu

Сыч Денис Васильевич, 1980 года рождения, окончил Московский государственный университет имени М.В. Ломоносова, кандидат физико-математических наук, старший научный сотрудник Физического института им. П.Н. Лебедева РАН. Область научных интересов: квантовая информация, квантовая и классическая оптика, однопиксельная визуализация. E-mail: denis.sych@gmail.com

ГРНТИ: 28.23.15

Поступила в редакцию 30 марта 2024 г. Окончательный вариант – 06 ноября 2024 г.

Study of image reconstruction efficiency in a single-pixel imaging method using generative adversarial networks

D.V. Babukhin¹, A.A. Reutov¹, D.V. Sych¹

¹ P.N. Lebedev Physical Institute, RAS, 53 Leninskiy Prospekt, Moscow, 119991, Russia

Abstract

Single-pixel imaging is a promising image acquisition method that provides an alternative to traditional imaging methods using multi-pixel matrices. However, algorithmic image reconstruction from measurements of a single-pixel camera is a non-trivial computational task that can be solved by machine learning methods. In this work, we investigate the possibility of image reconstruction in the single-pixel imaging method using generative adversarial neural networks. Using computer simulation of a single-pixel camera, we study the efficiency of image reconstruction using two generative network architectures — a deep convolutional generative adversarial network and a generative least squares adversarial network. We find that the generative least squares adversarial network demonstrates a better image reconstruction quality compared to the deep convolutional generative adversarial network. However, when taking into account optical distortions, the deep convolutional adversarial network is more stable in learning to a higher quality compared to the generative least squares adversarial network. The results obtained in this work may serve as a basis for the development of software for the practical application of a single-pixel camera.

Keywords: single-pixel imaging, image restoration, generative adversarial networks, hardware distortion correction.

Citation: Babukhin DV, Reutov AA, Sych DV. Study of image reconstruction efficiency in single-pixel imaging method using generative adversarial networks. *Computer Optics* 2025; 49(5): 818-825. DOI: 10.18287/2412-6179-CO-1526.

Acknowledgements: This work was supported by the Russian Science Foundation under project No. 23-22-00381, <https://rscf.ru/project/23-22-00381/>.

Author's information

Danila Valerevich Babukhin, (b. 1995), studied at Lomonosov Moscow State University, junior researcher at P.N. Lebedev Physical Institute of the Russian Academy of Sciences. Research interests: quantum communication, single-pixel imaging, machine learning in physics. E-mail: dv.babukhin@gmail.com

Alexey Alexeevich Reutov, (b. 1995), studied at Moscow Institute of Physics and Technology, junior researcher at P.N. Lebedev Physical Institute of the Russian Academy of Sciences. Research interests: quantum communication, single-pixel imaging, machine learning in physics. E-mail: aleksey.reutov@phystech.edu

Denis Vasilevich Sych, (b. 1980), studied at Lomonosov Moscow State University, PhD. in Physics and Mathematics, senior researcher at P.N. Lebedev Physical Institute of the Russian Academy of Sciences. Research interests: quantum information, quantum and classical optics, single-pixel imaging. E-mail: denis.sych@gmail.com

Received March 30, 2024. The final version – November 06, 2024.
