Одномоментная реконструкция встроенной голограммы очень желательна как экономически эффективный и портативный способ визуализации в условиях ограниченных ресурсов. Однако артефакты двойного изображения, вызванные распространением сопряженного волнового фронта с отсутствующей фазовой информацией, загрязняют реконструкцию. Существующие методы сквозного глубокого обучения требуют огромного количества пар обучающих данных, а также стабильности окружающей среды и системы, чего очень трудно достичь. Недавно предложенный глубокий предварительный анализ изображения (DIP) интегрирует физическую модель формирования голограммы в глубокие нейронные сети без необходимости предварительного обучения. Однако процесс приведения модели к единичной измеренной голограмме приводит к появлению шума, связанного с помехами. Для преодоления этой проблемы мы применили необученную глубокую нейронную сеть с явной регуляризацией по денуазированию (RED), которая удаляет сдвоенные изображения и шум при реконструкции. В нашей работе продемонстрировано использование метода переменных направлений множителей (ADMM) для объединения DIP и RED в надежный одномоментный процесс восстановления фазы. Использование ADMM, основанного на подходе расщепления переменных, позволило подключать и использовать различные денойзеры без необходимости явного дифференцирования. Экспериментальные результаты показывают, что денойзеры, способствующие разреженности, дают лучшие результаты по сравнению с DIP в плане отношения фазового сигнала к шуму (SNR). Учитывая вычислительные сложности, мы пришли к выводу, что денойзер с полной вариацией лучше подходит для восстановления голограмм.

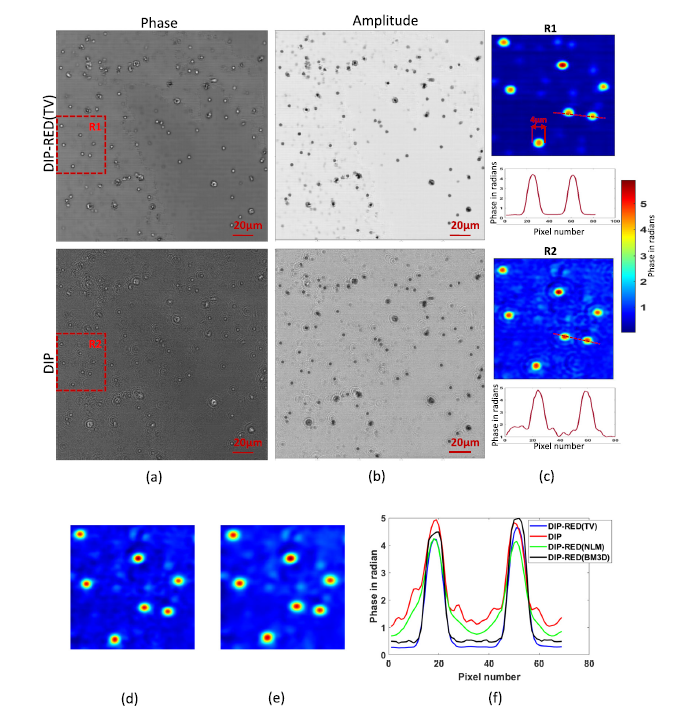
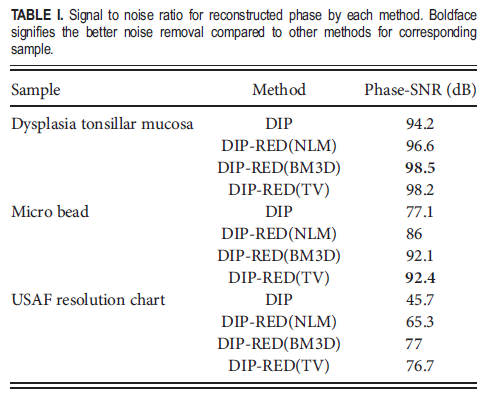
Количественная фазовая информация биологических образцов напрямую связана с толщиной и индексом рефракции образца и, следовательно, улучшает клинический анализ.1,2 Цифровая поточная голографическая микрокопия (DIHM) позволяет получать количественные изображения прозрачных образцов без меток и широко используется в медицине и физических науках. В DIHM сложная фазовая информация кодируется в интерференционную картину, известную как голограмма, где трехмерная информация объекта записывается в одно двухмерное изображение.3 Голограмма затем восстанавливается путем обратного распространения голограммы от плоскости детектора к плоскости объекта. Расширенная глубина поля (DOF) может быть достигнута в процессе извлечения фазы, чтобы получить трехмерное распределение образца.4 Из-за отсутствия фазовой информации в голограмме, реконструкция имеет нежелательный эффект, называемый артефактом двойного изображения. Для подавления этого двойного изображения в литературе было предложено несколько методов фазового восстановления. Безлинзовый DIHM (LDIHM) является простым и экономически эффективным способом получения изображений для применения в пунктах оказания медицинской помощи.5,6 Здесь объект иллюстрируется частично когерентным источником и интерференционная картина дифрагированной волны UoðÞ; и недифрагированная волна ðURÞ фиксируется на текторе, расположенном на расстоянииz(1 мм) от образца,  
Недифрагированная волна, известная как опорная, регистрируется без образца в тех же условиях записи. Эффект URjj2 устраняется путем нормализации голограммы, а самоинтерференция Uojj2 рассматривается как шум системы, e. Уравнение(1), таким образом, может быть также записано как  
Производительность Rhis сильно зависит от размера и дисперсии набора данных ST. Сбор пар входных и выходных изображений (Ik;uko) для практических приложений занимает много времени и является невыполнимой задачей. Кроме того, тестовое изображение может отличаться от изображений, полученных во время обучения, из-за механической и экологической нестабильности. Более того, глубокие сети в значительной степени зависят от морфологических характеристик образца, поэтому для каждого образца требуется отдельное обучение. Для решения этой проблемы, связанной с потребностью в данных, была разработана глубокая сеть обработки изображений (deep imageprior, DIP), которая является первой необученной сетью, реализованной для обесцвечивания, восстановления изображений, инпайнтинга и суперразрешения на основе одного зашумленного изображения.28Недавно было доказано, что необучаемые нейронные сети, учитывающие физические особенности, эффективны для восстановления изображений путем решения стандартных обратных задач, не требуя предварительного обучения.29-33 Восстановление фазы с помощью DIP достигается путем включения физической модели (прямого распространения), которая является проявлением процесса формирования изображения, нетрадиционного глубокого обучения. Фазовая реконструкция для DIHM с использованием DIP29 показала превосходную производительность по сравнению со сжатием (CS) и традиционной реконструкцией на основе распространения даже без предварительного обучения. Реконструкция на основе DIP сформулирована следующим образом

где сеть Rh генерирует объект ðuoÞ либо из вектора ﬁxednoise, z, либо из некоторого начального реконструированного объекта. Затем модель прямого распространения, представленная в уравнении (2), используется для генерации новой голограммы. Ошибка между полученной и измеренной голограммой возвращается в сеть для обновления параметров. Взаимодействие между сетью (Rh) и прямой моделью [HðuoÞ]позволяет узнать параметры сети (hÞ). Эти выученные параметры используются для восстановления амплитуды и фазы объекта, которые наиболее соответствуют входной голограмме. Поскольку фазовый поиск на основе DIP не требует большого набора данных для обучения, он подходит для фазового поиска по одному снимку. Однако параметры сети определяются с помощью модели формирования физического изображения, которая в основном зависит от пространственных частотных компонентов (fx;fyÞ, расстояния распространения (zÞ и длины волны (kÞ). Процесс приведения выходных данных модели к одной измеренной голограмме приводит к устранению шума, связанного с помехами, и затухания веса. В данном письме мы использовали DIP с регуляризацией путем денуазинга (RED), добавив явные денуазеры [qxðÞ]как приоритет.34,35 В дальнейшем мы будем называть предложенный нами метод DIP-RED. Явные денойзеры могут дополнительно усилить эффект регуляризации и уменьшить шум в восстановленном изображении

RED, добавленный в уравнении (8), является необученным и может потенциально удалить мелкие расфокусированные компоненты в пространстве признаков, что существенно улучшает результат реконструкции и определяется как a

В RED можно включить любой алгоритм денуации изображения, удовлетворяющий некоторым условиям, приведенным в дополнительном материале. Однако, Applied Physics Letters ARTICLE scitation.org/journal/aplAppl. Phys. Lett. 122, 133701 (2023); doi: 10.1063/5.0144795 122, 133701-2Опубликовано по эксклюзивной лицензии AIP Publishing Дифференцирование функции денойзера в процессе оптимизации является утомительной задачей и не рекомендуется для большинства денойзеров. К счастью, с помощью техники расщепления переменных, известной как AlternatingDirections of the Multipliers (ADMM), можно разделить терм "делимость данных" и терм регуляризации.37ADMM позволяет распараллелить выполнение денойзера и параметров сети. DIP-RED можно представить в виде

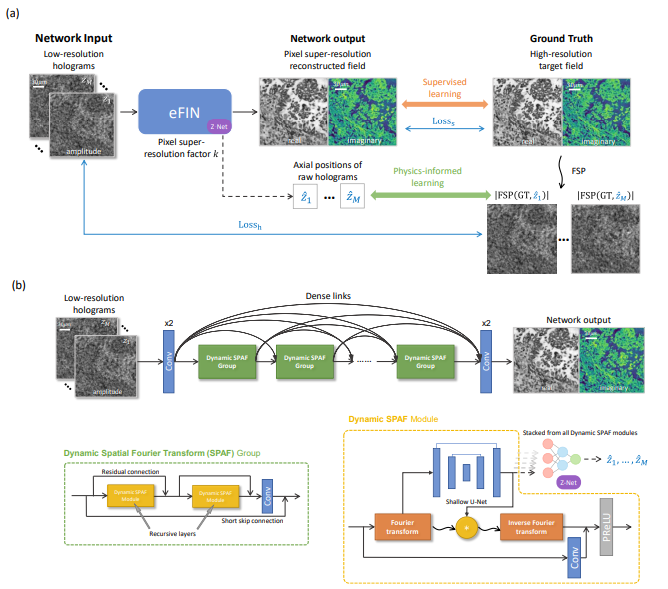
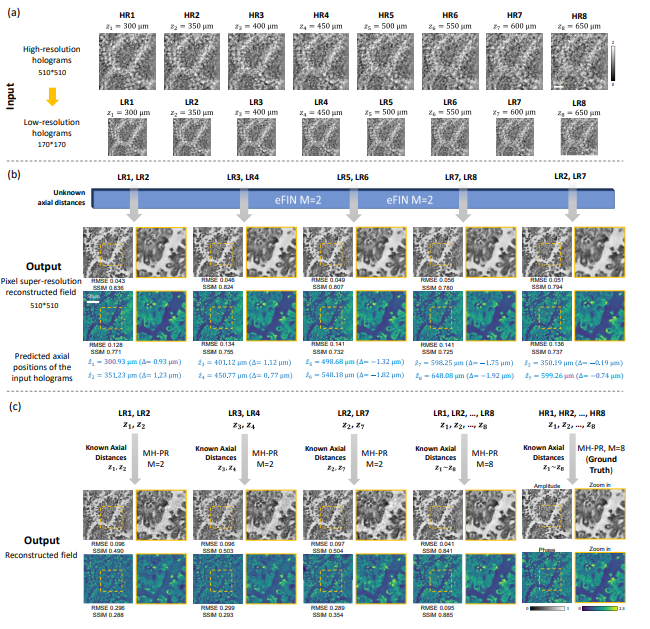
Здесь uois - реконструкция, полученная на предыдущей итерации, а ^uois - реконструкция на текущей итерации. Первый член представляет собой стандартную обратную задачу, решаемую DIP. Второй член представляет собой функцию RED, а fð:Þ - выбранную функцию денуатора. При некоторых мягких условиях на fð:Þ градиент функции RED isuofuoðÞ, что позволяет избежать излишней дифференциации денуатора.35 Третий член - регуляризация близости, которая заставляет выходной сигнал сети Rh^uoðÞ быть близким к uot. t - вектор множителей Лагранжа, a и bare - параметры регуляризации. Задача минимизации в уравнении (10) может быть решена с помощью ADMM37,38 путем последовательного обновления трех параметров, h, t иuo. Уравнение (11) вводит регуляризацию близости (второй член) в решение DIP (ﬁпервый член), что повышает стабильность и устойчивость модели.

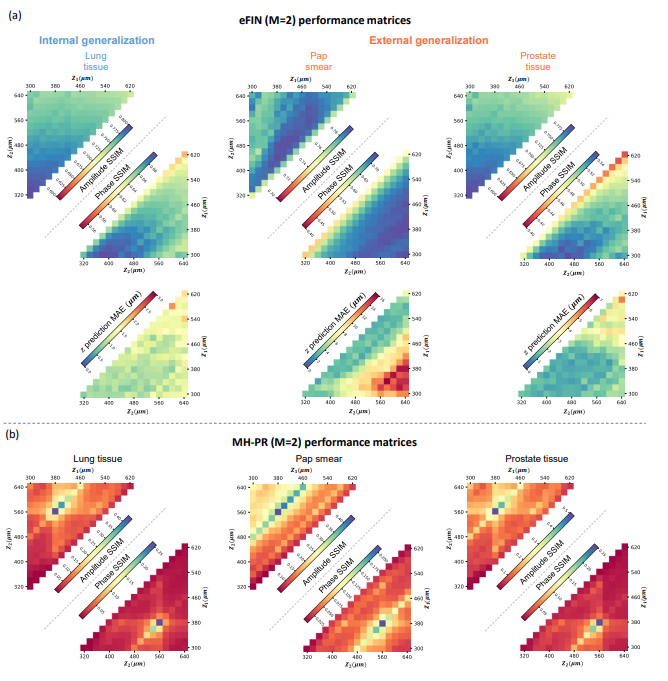
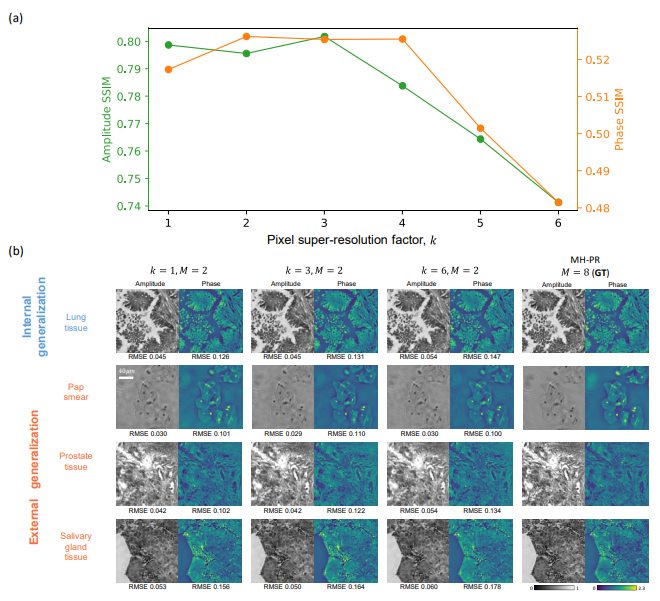
В статье [2] авторы предложили алгоритм DIP-RED восстановления фазы.   
На рисунке (а) показана восстановленная фаза, на (b) – амплитуда, (c) – увеличенная область интереса для демонстрации фазового профиля бусины. Результаты реконструкции на рисунке показывают, что, несмотря на регуляризацию DIP без сильных предварительных условий, сеть может обрабатывать шумы, связанные с помехами. Такой шум легко удаляется с помощью регуляризаторов, применяемых в DIP-RED. На рисунках 4(d) и 4(e) показана фазовая реконструкция, выполненная DIP-RED(NLM) и DIP-RED(BM3D), соответственно. Фазовый профиль, как показано на рис. 4(f), показывает, что все денойзеры могут удалять небольшие расфокусированные значения и приводят к лучшему восстановлению по сравнению с DIP. Было замечено, что все денуазеры могут удалить небольшие расфокусированные значения и привести к лучшему восстановлению по сравнению с DIP. Отношение сигнал/шум фазы (SNR)23 было рассчитано для восстановленной фазы каждым методом, и значения приведены в таблице I. DIP-RED дает значительно более высокие значения SNR фазы, что свидетельствует о лучшем восстановлении фазы при уменьшении шума.   
  
  
  
  
Применение [1] методов глубокого обучения значительно расширило возможности голографической визуализации, что привело к улучшению восстановления фазы и реконструкции изображений. Здесь мы представляем глубокую нейронную сеть, названную улучшенной сетью Фурье-изображений (eFIN), как высоко обобщаемую и надежную основу для реконструкции голограмм с суперразрешением пикселей и автофокусировкой изображения. В ходе экспериментов по голографической микроскопии с использованием срезов тканей легких, простаты, слюнных желез и мазков Папаниколау (Pap) мы продемонстрировали, что eFIN обладает превосходным качеством реконструкции изображений и демонстрирует внешнюю обобщенность для новых типов образцов, не встречавшихся на этапе обучения. Эта сеть достигает широкого осевого диапазона автофокусировки Δz ∼ 350 мкм и способна точно предсказывать осевые расстояния голограмм с помощью физически обоснованного обучения. eFIN позволяет получать изображения со сверхразрешением 3× пикселей и увеличивает пространственно-пропускное произведение реконструированных изображений в 9 раз практически без потери производительности, что позволяет значительно экономить время на этапах голографической визуализации и обработки данных. Наши результаты показывают, что eFIN расширяет границы голографической визуализации для различных приложений, например, для количественной фазовой визуализации и микроскопии без меток.

В последние годы цифровая голография привлекла к себе большое внимание как мощный инструмент микроскопической визуализации, способный реконструировать сложные оптические поля образцов без необходимости окрашивания, фиксации или маркировки образца [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. Ключевой проблемой цифровой голографии является восстановление недостающей фазовой информации из записанной голограммы только интенсивности, которая может быть решена с помощью вычислительных подходов, таких как итерационные алгоритмы, основанные на физическом распространении волн [12], [13], [14], [15], [16], [17], [18], [19], [20], [21] и, в последнее время, глубокие нейронные сети. Последние подходы, основанные на глубоком обучении, используют нейронные сети для реконструкции сложного поля образца из одной или нескольких голограмм за один шаг прямого вывода, что дает преимущества по сравнению с традиционными методами, включая более высокую скорость реконструкции, улучшенное соотношение сигнал/шум и увеличенную глубину поля [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44]. Кроме того, эти методы успешно достигли внутренней генерализации, определяемой как способность глубокой нейронной сети восстанавливать голограммы новых, невидимых объектов, принадлежащих к тому же типу образца, который использовался во время обучения. Недавно Чен и др. представили Fourier Imager Network (FIN), [45] глубокую нейронную сеть, которая использует модуль пространственного преобразования Фурье для обучения отображению между входными голограммами и сложными полями образцов через глобальное рецептивное поле. FIN успешно справилась с задачей внешнего обобщения, реконструировав голограммы новых, невидимых образцов, принадлежащих к различным типам образцов, которые не были замечены во время обучения. Однако в FIN отсутствует возможность автофокусировки, что означает, что одна обученная сеть FIN может реконструировать сложное поле образца только по голограммам, снятым на определенных, известных осевых расстояниях. В данной работе мы представляем улучшенную сеть Фурье-изображений (eFIN), надежную глубокую нейронную сеть для сквозного поиска фазы и реконструкции голографических изображений, которая достигает как сверхразрешения пикселей, так и автофокусировки - одновременно, благодаря быстрому выводу. Хотя eFIN и является производной от FIN, она значительно отличается по своей архитектуре, используя неглубокую U-сеть в своем модуле динамического пространственного преобразования Фурье (SPAF), и обеспечивает дополнительные степени свободы для интеграции возможностей попиксельного сверхразрешения и автофокусировки в одной нейронной сети. В частности, eFIN отличается способностью выполнять автофокусировку в широком осевом диапазоне (Δz) 350 мкм и точно предсказывать осевые расстояния входных голограмм с помощью физически обоснованного обучения, что избавляет от необходимости использовать истинные значения осевых расстояний. Более того, eFIN демонстрирует превосходное качество реконструкции изображений в задачах внутреннего и внешнего обобщения и способен достичь сверхразрешения 3× пикселей (PSR) при минимальном снижении производительности, что позволяет сэкономить значительное количество времени на этапах получения изображений и обработки данных. Говоря иначе, eFIN обучен (1) выполнять реконструкцию голограмм с пиксельным сверхразрешением, выявляя амплитудные и фазовые изображения входного образца с более высоким разрешением; (2) выполнять осевую автофокусировку, поскольку расстояния от образца до датчика неизвестны и меняются от входа к входу; и (3) выявлять осевые расстояния входных голограмм в дополнение к извлечению фазы, реконструкции голограмм PSR и автофокусировке.

Чтобы экспериментально подтвердить успешность eFIN, мы обучили его на голограммах образцов ткани человеческого легкого (гистопатологические слайды) и вслепую проверили его вывод, реконструкцию изображений и работу автофокусировки на мазках Папса, человеческих легких, простате и срезах тканей слюнных желез (все они не были видны на этапе обучения). Превосходные результаты реконструкции изображений с помощью PSR и автофокусировки eFIN представляют собой значительное достижение в области голографической визуализации на основе глубокого обучения, открывая возможности для значительного повышения эффективности и скорости голографической микроскопии высокого разрешения. Универсальность eFIN также позволяет использовать его в других когерентных системах формирования изображений, включая, например, внеосевую голографию и различные некогерентные режимы формирования изображений, что делает его перспективным инструментом для широкого спектра приложений визуализации и микроскопии.

Результаты и обсуждения A. Архитектура eFIN eFIN представляет собой глубокую нейронную сеть, которая выполняет сквозное восстановление фазы и реконструкцию голографического изображения с возможностью автофокусировки в широком осевом диапазоне и попиксельной реконструкции изображений со сверхразрешением, как показано на рис. 1. Кроме того, подсеть eFIN, названная Z-Net, обучена слепому прогнозированию осевых расстояний от образца до датчика для входных голограмм без какой-либо априорной информации.   
Рис. 1. Реконструкция голограмм и автофокусировка в системе eFIN. (a) Входные голограммы низкого разрешения снимаются на различных расстояниях от образца до датчика zi, i = 1,...,M. Losss - контролируемая функция потерь, полученная из целевого поля, в то время как Lossh генерируется в результате обучения с учетом физики и применяется только к сети Z-Net. (b) Архитектура сети eFIN, где Z-Net является подсетью eFIN. Выходы U-Net всех модулей Dynamic SPAF собираются и поступают в Z-Net для предсказания осевых расстояний (z^) входных голограмм. Показать все   
На вход сети подается последовательность линейных голограмм низкого разрешения только интенсивности [46], снятых на M различных расстояниях от образца до датчика, т. е. z1...,zM (см. раздел "Методы"). Выходными данными являются действительная и мнимая части пиксельного сверхразрешенного реконструированного комплексного поля образца, а также расстояния от образца до датчика соответствующих входных голограмм. Истинные комплексные изображения высокого разрешения (GT) для контролируемой части обучения получены с помощью итеративного алгоритма многовысотного фазового поиска (MH-PR) [17], [18] с использованием пиксельных сверхразрешенных голограмм, полученных на восьми различных осевых расстояниях (M = 8) (см. раздел "Методы").

  
B. Реконструкция голограмм с пиксельным сверхразрешением и эффективность автофокусировки eFIN Чтобы продемонстрировать превосходную производительность eFIN, мы обучили сеть на образцах легочной ткани человека (см. раздел "Методы"). Во время каждой итерации обучения в сеть подавались две голограммы низкого разрешения (т. е. M = 2, k = 3), снятые на случайно выбранных и неизвестных осевых расстояниях между образцом и датчиком, z, в диапазоне от 300 мкм до 650 мкм, в порядке возрастания, где k = 3 означает целевой коэффициент сверхразрешения пикселя, как показано на рис. 2(a). Говоря иначе, каждая поточная голограмма низкого разрешения имеет в k2 раз меньшее количество пикселей, что приводит к потере разрешения из-за недодискретизации; для голографического микроскопа на кристалле с единичным увеличением [47] это эквивалентно использованию оптоэлектронного датчика изображения с k-кратно большей шириной/периодом пикселя, что приводит к k2-кратному уменьшению количества пикселей (например, k2 = 9 на рис. 2). Рис. 2. Восстановление пиксельной сверхразрешенной голограммы и производительность автофокусировки в eFIN. (a) Входные голограммы низкого разрешения были получены с помощью понижающей дискретизации 3× (путем бинирования пикселей). (b) Выходные комплексно-значимые поля были сгенерированы с помощью обученной сети eFIN (M = 2) с пиксельным коэффициентом сверхразрешения k = 3. (c) Выходные комплексно-значимые поля алгоритма MH-PR, использующего те же входные голограммы низкого разрешения (M = 2), приводят к более низким по качеству реконструкциям, даже если он использует известные аксиальные расстояния для входных голограмм.  
  
Рис. 3. Эффективность реконструкции изображений с помощью eFIN и MH-PR. (a) Амплитудные и фазовые значения SSIM реконструированных комплексно-значных полей и значения MAE предсказанных осевых расстояний с использованием сети eFIN (M = 2, k = 3). (b) Матрицы эффективности реконструкции изображений алгоритмом MH-PR (M = 2) при использовании различных комбинаций расстояний z в качестве входных данных, при этом входные голограммы низкого разрешения были сняты при фиксированных z1 = 380 мкм и z2 = 560 мкм. Показать все Рис. 4. Эффективность реконструкции голограмм с помощью eFIN в зависимости от коэффициента PSR (k). Независимая модель eFIN (M = 2) была обучена на образцах легочной ткани человека для каждого фактора PSR k. (a) Средние значения SSIM амплитудной и фазовой частей восстановленных полей образцов для различных типов образцов, снятых при различных комбинациях осевых расстояний. (b) Визуализация реконструированных eFIN полей образцов. Истина для каждого образца получена с помощью алгоритма MH-PR, который использовал M = 8 пиксельных сверхразрешенных голограмм, снятых на 8 различных расстояниях от образца до датчика. Показать все Для дальнейшего сравнения эффективности eFIN с традиционными методами реконструкции голограмм мы также оценили качество реконструкции алгоритма MH-PR. В данном случае, чтобы обеспечить справедливое сравнение с eFIN, мы не предоставляли точные осевые расстояния (значения z) входных голограмм для MH-PR. Вместо этого в алгоритм MH-PR были поданы две фиксированные голограммы, снятые при z1 = 380 мкм и z2 = 560 мкм, а также различные комбинации изменяющихся осевых расстояний (z1 и z2), чтобы смоделировать сценарий, когда точные осевые расстояния входных/полученных голограмм недоступны (что часто случается, особенно в относительно недорогом оборудовании для голографической визуализации). Как показано на рис. 3(b), классический алгоритм MH-PR не может точно восстановить поле образца, используя неточные осевые расстояния голограмм; для сравнения, результаты автофокусировки eFIN показаны на рис. 3(a).

  
  
C. Эффективность реконструкции eFIN в зависимости от коэффициента PSR (k) Для дальнейшего изучения эффективности реконструкции голограмм с помощью eFIN мы индивидуально обучили шесть различных сетей eFIN с коэффициентом PSR k в диапазоне от 1 до 6. Все модели eFIN были обучены только на образцах легочной ткани человека с двумя входными голограммами (M = 2), полученными на случайно выбранных и неизвестных расстояниях от образца до датчика от 300 мкм до 650 мкм. Для коэффициента PSR k = 3 мы использовали те же веса модели, что и на рис. 2 и 3. Эти шесть различных сетей eFIN были протестированы с различными комбинациями z1 и z2 в диапазоне 300 мкм, 400 мкм, 500 мкм и 600 мкм и слепо оценены на срезах легочной ткани человека (внутреннее обобщение) и образцах мазков, тонких тканей простаты и слюнных желез человека (внешнее обобщение). Средние значения SSIM в зависимости от целевых коэффициентов PSR (k) представлены на рис. 4(a), а восстановленные поля образцов и соответствующие им метрики изображений - на рис. 4(b). Эти результаты подтверждают выдающиеся PSR-характеристики eFIN, показывая, что качество реконструкции не страдает даже при использовании входных голограмм низкого разрешения с k2= 9 или k2= 16 раз меньшим количеством пикселей по сравнению со случаем k = 1 (см. рис. 4).  
Не изменяя архитектуру eFIN, мы также обучили с нуля новую модель eFIN (M = 2), которая адаптируется к различным коэффициентам PSR от k = 1 до 6; иначе говоря, эта новая модель eFIN была обучена восстанавливать голограммы низкого разрешения, снятые при различных размерах пикселей, с целью достижения различных коэффициентов PSR (k) с помощью одной и той же сети. Мы использовали тот же набор данных для обучения человеческой легочной ткани, что и ранее, и вслепую протестировали полученную сеть на новых срезах легочной ткани (для внутреннего обобщения) и образцах мазка Папса, простаты и слюнных желез человека (для внешнего обобщения). Во время обучения этой модели eFIN осевые расстояния входных поточных голограмм выбирались случайным образом из диапазона от 300 мкм до 650 мкм, а тестовые расстояния между образцом и датчиком представляли собой комбинации двух значений z, выбранных из 300 мкм, 400 мкм, 500 мкм и 600 мкм. Средние значения SSIM выходных изображений унифицированной модели eFIN для этих тестовых образцов представлены на рис. 5(a) в зависимости от коэффициента PSR (k). На этом же рисунке сравниваются показатели различных моделей eFIN, которые были обучены по отдельности для каждого значения k (см. пунктирные линии на рис. 5(a)). Это сравнение на рис. 5(a) показывает, что единая модель eFIN (т. е. объединенная сеть) может выполнять реконструкцию голограммы в большом диапазоне PSR (k = 1-6), соответствуя производительности отдельно обученных моделей eFIN, которые были специализированы/обучены для данного k. Результаты реконструкции голограммы этой объединенной модели eFIN PSR также показаны на рис. 5(b). Эти результаты показывают, что единая модель eFIN может быть обучена обрабатывать/реконструировать голограммы с различными коэффициентами PSR через одну общую сеть. Рис. 5. Результаты реконструкции голограмм с помощью единой модели PSR eFIN. Одна модель eFIN (M=2) была обучена с нуля на образцах легочной ткани человека с различными коэффициентами PSR (k) в диапазоне от 1 до 6. (a) Средние значения SSIM амплитудной и фазовой частей восстановленных изображений для различных типов образцов; голограммы низкого разрешения в линию были сняты при различных комбинациях осевых расстояний (от образца до датчика). Каждый кружок представляет собой индивидуально обученную модель eFIN для определенного фактора PSR, которая сравнивается с результатами реконструкции голограмм с помощью унифицированной модели PSR eFIN. (b) Примеры полей образцов, реконструированных по унифицированной модели PSR eFIN. Истина для каждого образца получена с помощью алгоритма MH-PR, который использовал M=8 пиксельных сверхразрешенных голограмм, снятых на 8 различных расстояниях от образца до датчика.

DarkFocus [4]  
В работе [4] авторы предлагают новую схему для оценки фокальной плоскости в цифровой поочной голографической микроскопии, которая обладает следующими приемуществами: (1) точность, (2) униполярность, (3) универсальность, (4) надежность, (5) высокая аксиальная точность, позволяющая определять несколько фокальных плоскостей, особенно в условиях шума и пространственно изменяющейся плотности образца, и (6) возможность качественного определения относительного уровня слияния долей объекта в фокусе в режиме нескольких фокальных плоскостей (для сравнения плотности образца в разных фокальных плоскостях). Для сравнения выбрали две очень эффективные эталонные метрики: комплексную амплитуду с фильтрацией высоких частот, называемую методом Дюбуа, и метод градиента Тамура, называемый методом ToG. Исследовали полистироловые микробусины - стандартные монодисперсные полистироловые микросферы, погруженные в водную среду - в качестве калибровочных образцов в различных экспериментальных условиях, чтобы подчеркнуть широкую применимость предлагаемой автофокусировки DIHM. Микробусины выбрали в качестве смешанных фазово-амплитудных образцов, поскольку они в основном преломляют свет как фазовые объекты, но благодаря высокой числовой апертуре и локальной неясности поля они также ведут себя как амплитудные образцы. Для проверки подхода мы рассмотрели два различных размера статических микробусин (Polybead Microspheres, диаметром 45 мкм или 90 мкм). Бусины вносятся в разреженном (не кластерном) режиме в счетную камеру толщиной 100 мкм с помощью микропипетки.  
Предлагаемая новая методика численной автофокусировки - DarkFocus - содержит следующую структуру алгоритмической процедуры:  
(A) фильтрация голограммы, (B) распространение отфильтрованной голограммы вдоль оптической оси на заданное расстояние z и вычисление амплитуды распространяющегося комплексного поля, (C) вычисление метрики фокусировки DarkFocus(z) как дисперсии градиента амплитуды - обычно чем выше значение метрики, тем лучше фокусировка, (D) повторение шагов (A)-(C) для всех заданных расстояний распространения z с получением кривой автоматической фокусировки.  
  
Записанная линейная голограмма – картина интенсивности, состоящая из некогерентного фона, когерентных интерференционных полос и шума – адаптивно фильтруется с помощью разложения области изображения. Голограмма разлагается на несколько двумерных [внутренних модовых функций](https://www.sciencedirect.com/topics/engineering/intrinsic-mode-function)(BIMF) с использованием улучшенного и быстрого алгоритма [эмпирического модового разложения](https://www.sciencedirect.com/topics/engineering/empirical-mode-decomposition) (EFEMD). Метод EFEMD основан на идее одномерного эмпирического модового разложения, первоначально предложенного Хуангом и др. как эффективный инструмент для [адаптивного анализа](https://www.sciencedirect.com/topics/engineering/adaptive-analysis) нестационарных и [нелинейных сигналов](https://www.sciencedirect.com/topics/engineering/nonlinear-signal) , позднее расширенный до 2D Нунесом и др. и упрощено с точки зрения [числа итераций](https://www.sciencedirect.com/topics/engineering/iteration-number) Бхуйяном и др.. Метод EFEMD адаптивно рассекает многомасштабные компоненты изображения, используя фильтры разложения различного размера, определяемые распределением локальных экстремумов изображения (подход, управляемый данными). Метод EFEMD применяется здесь для разложения линейных голограмм.

Процесс фильтрации голограмм демонстрируется с использованием наглядного линейного распределения интенсивности,[Рис. 2](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub" \l "fig0002)(a), записано камерой CCD после освещения микрошариков [полистирола](https://www.sciencedirect.com/topics/medicine-and-dentistry/polystyrene) синим светом. Остаток разложения (низкочастотная составляющая) и два первых BIMF представлены на[Рис 2](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0002)(b)-2(d), соответственно. Первые два BIMF добавляются для формирования адаптивно отфильтрованного свободного от фона интерференционного члена линейной голограммы и представлены в[Рис. 2](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0002)(e). Используется двухмодовая фильтрация (BIMF), поэтому, поскольку авторы хотят сосредоточиться на интерференционных полосах, исходящих от оптического [поля](https://www.sciencedirect.com/topics/engineering/optical-field), рассеянного на объекте, чтобы получить темное поле при распространении отфильтрованной голограммы. На вопрос о количестве BIMF, которые должны быть извлечены и добавлены при фильтрации, можно ответить вручную, просто выбрав правильное количество мод после нескольких испытаний. Разложение представленное в[Рис. 2](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0002) происходит всего за 0,8 с (реализация Matlab на CPU 2,6 ГГц, [RAM](https://www.sciencedirect.com/topics/engineering/reliability-availability-and-maintainability-reliability-engineering) 16 ГБ), что является очень благоприятным результатом для обработки паттерна интенсивности 2040 × 2040 пикселей.

### 3.2 Численное распространение голограммы

### Записанная и отфильтрованная линейная голограмма должна быть численно обработана для достижения условий визуализации путем обратного распространения оптического поля. Существуют различные [численные методы](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/mathematical-method)цифровой реконструкции голограммы, в основном путем решения [дифракционного](https://www.sciencedirect.com/topics/medicine-and-dentistry/diffraction) интеграла распространения поля. Строгие методы распространения оптического поля основаны на интеграле Рэлея-Зоммерфельда (RS) или методе углового спектра (AS)[[82]](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub" \l "bib0082), которые эффективно численно решаются с использованием [быстрого преобразования Фурье](https://www.sciencedirect.com/topics/engineering/fast-fourier-transform) (БПФ). В исследовании использовали процедуру распространения углового спектра. АС выгоден с точки зрения простого пути обработки, что приводит к низкой вычислительной нагрузке и короткому времени расчета, высокой эффективности, особенно в параксиальном режиме, высокой точности и хорошей [дискретизации](https://www.sciencedirect.com/topics/engineering/discretization). 3.3 . Новый расчет метрики DarkFocus

Предлагаемая схема автофокусировки основана на вычислении амплитуды распространяющегося комплексного поля, дающего вычислительное темное поле, и последующем вычислении дисперсии градиента темного поля. Предварительная фильтрация голограммы с целью генерации темного поля усиливает воздействие компонента интерференции, рассеянного объектом. Эту информацию об объекте теперь необходимо перенести в плоскость фокуса, где «резкость» максимальна. Операция градиента усиливает [высокочастотные компоненты](https://www.sciencedirect.com/topics/engineering/high-frequency-component) , связанные с резкостью объекта темного поля, и дисперсия является подходящей мерой для их количественной квантизации. В этом духе метрика DarkFocus (DF) определяется как:(2)ДФ=вар(град(абс(У(з))),где U(z) — комплексное оптическое поле, вычисленное после распространения отфильтрованной голограммы на расстояние z, abs(U(z)) — вычислительное темное поле, grad — оператор градиента, описанный ниже. Представленная схема обеспечивает максимизацию значения DF на расстоянии перефокусировки, указывая плоскость фокусировки, независимо от типа образца (фазовый, амплитудный, смешанный), поскольку каждый связанный с объектом компонент интерференции добавляется к общей метрике при фокусировке.

Таким образом, имеется доступ к двум направлениям ортогональных пространственных производных темного поля. Поскольку само темное поле является единым полностью положительным полем, а его две градиентные карты имеют как отрицательные, так и положительные значения, применяется ограничение неотрицательности в терминах вычисления суммы квадратов частных пространственных производных для получения единого объединенного полностью положительного представления [градиентного поля :](https://www.sciencedirect.com/topics/engineering/gradient-field)(4)град(абс(У(з))=(гх.2+гы.2)

Процедура завершается вычислением дисперсии для определения метрики DarkFocus. Возведение в квадрат является нелинейной операцией, и этот факт дополнительно увеличивает чувствительность резкости края DarkFocus и вычисленную надежность дисперсии. Можно также реализовать норму L1 в терминах суммирования абсолютных [значений двух ортогональных пространственных производных или норму L2 в терминах](https://www.sciencedirect.com/topics/engineering/absolute-value)[вычисления модуля](https://www.sciencedirect.com/topics/engineering/modulus-calculation) градиента .

Расчет темного поля как часть метрического конвейера автофокусировки был предложен Дюбуа и др, однако фильтрация была выполнена с использованием произвольно выбранного комплексного высокочастотного [гауссовского](https://www.sciencedirect.com/topics/physics-and-astronomy/gaussian-distribution)ядра, примененного к распространенной [комплексной амплитуде](https://www.sciencedirect.com/topics/engineering/complex-amplitude) . Комплексные амплитуды требуют предварительной адекватной обработки для выравнивания фазового фона и удаления постоянных дефектов в поле [зрения](https://www.sciencedirect.com/topics/physics-and-astronomy/field-of-view) . Эта метрика, называемая Дюбуа, предназначена для всех амплитудных, фазовых и смешанных выборок, всегда дает минимальное значение кривой фокусировкию По сравнению с метрикой Дюбуа, предлагаемый метод DarkFocus работает в области голограммы и устойчив к неплоскому фазовому фону и дефектам, присутствующим в поле зрения. Не нужно применять никакую дополнительную расширенную процедуру удаления дефектов. Методика предполагает более высокую надежность, универсальность и точность представленной метрики DarkFocus по сравнению с метрикой Дюбуа.

Расчет дисперсии амплитудного члена комплексного поля был проведен для целей автофокусировки. Цель состояла в том, чтобы минимизировать его в фазовой объектной автоматической фокусировке и максимизировать его в амплитудной фокусировке объекта. Недавно коэффициент Тамуры, основанный на дисперсии и среднем значении, был предложен в качестве меры [разреженности](https://www.sciencedirect.com/topics/engineering/sparsity) и применен в автофокусировке. Для повышения производительности он был объединен с вычислением градиента до оценки коэффициента Тамуры, что дало так называемую метрику автофокусировки градиента Тамуры(ToG).[[56]](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub" \l "bib0056). Следует подчеркнуть, что ToG основан на амплитуде сложного оптического поля (голографическая фильтрация не применяется, поэтому исследуется режим светлого поля) и по своей сути чувствителен к острым краям, предполагая их разреженность. Поэтому осевое разрешение часто может быть затруднено в случае плотных объектов, где разреженность ограничена. По сравнению с метрикой ToG, наш предлагаемый DarkFocus применяет голографическую фильтрацию для вычислительного создания распространенного темного поля и не делает никаких предположений относительно разреженности сфокусированных объектов. Предполагаются проблемы ToG, связанные со слиянием. Это предполагает более высокую надежность, универсальность и точность представленного DarkFocus по сравнению с ToG, которые будут проверены в следующих разделах и рисунках.

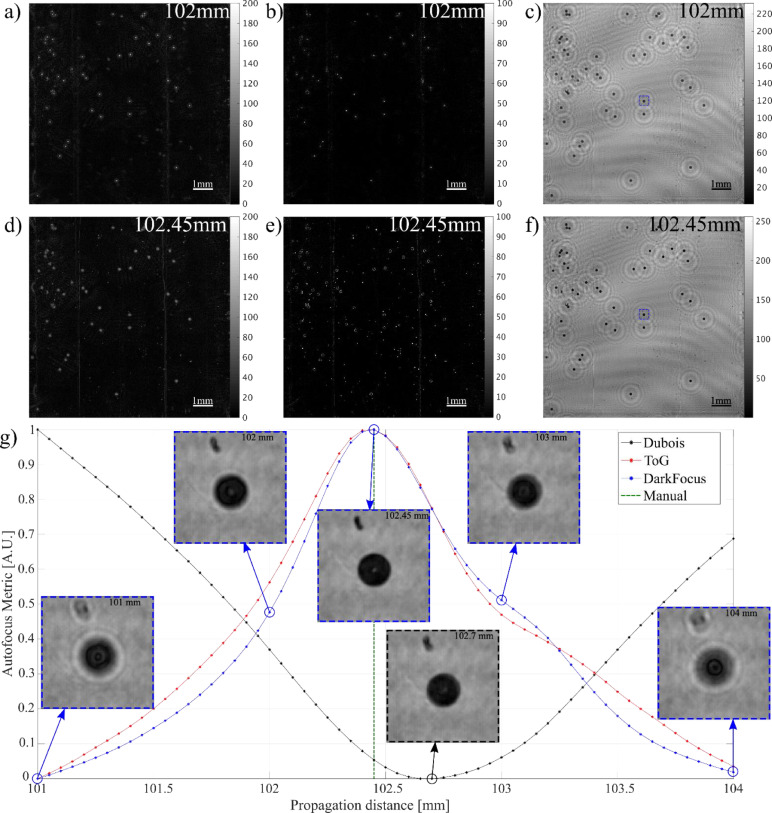
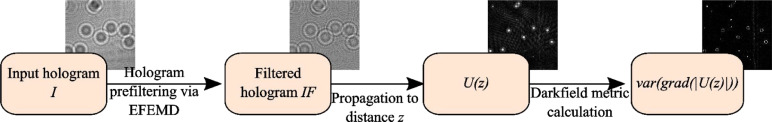
Пример распределения темного поля для голограммы ([Рис. 2](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0002)) представлены два расстояния распространения, z = 102 мм и 102,45 мм.[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub" \l "fig0003)(а) и 3(d), тогда как их соответствующие градиенты показаны на[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003)(b) и 3(e). Z = 102,45 мм выбрано вручную в качестве правильного расстояния перефокусировки. Можно заметить, анализируя оба распространяемых темных поля, что второе дает более высокую общую видимую «резкость». Это подтверждается расчетом метрики DarkFocus: DF(102) = 0,92, DF(102,45) = 1. Авторы нормализуют метрику DarkFocus и другие используемые метрики, чтобы обеспечить их четкое сравнение. Окончательная визуальная оценка выполняется путем распространения исходной голограммы на оба расстояния, см.[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003)(c) и 3(f) для распространения 102 мм и 102,45 мм соответственно.  


Рис. 3. Процедура DarkFocus пошагово: (a) темное поле, распространенное до 102 мм, (b) градиент темного поля 102 мм, (c) нефильтрованная голограмма, распространенная до 102 мм; (d) темное поле, распространенное до 102,45 мм, (e) градиент темного поля 102,45 мм, (f) нефильтрованная голограмма, распространенная до 102,45 мм, (g) график, иллюстрирующий три метрические кривые автофокусировки (Dubois – черным, ToG – красным и DarkFocus – синим). Вставки соответствуют области, отмеченной синим пунктирным квадратом. Визуализации 1, 2 и 3 представляют численную перефокусировку выбранной области в амплитудном режиме, режиме темного поля и градиентном режиме темного поля (DarkFocus) соответственно, демонстрируя перефокусировку микросфер и частиц пыли. Кроме того, включена Visualization4 для представления автоматической числовой перефокусировки другой области интереса, выделяя несколько бусин и частиц пыли.

Визуальная проверка положительно подтвердила численные автоматические расчеты. График, иллюстрирующий три метрические кривые автофокусировки (Dubois – черным, ToG – красным и DarkFocus – синим), представлен на[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003)(g) с дополнительными вставками, облегчающими качественную оценку точности числовой фокусировки. Область интереса для вставок была выбрана так, чтобы охватить одиночную бусину и одну частицу пыли. Обратите внимание на ошибочный результат метрики Дюбуа, четко указанный на не сфокусированной вставке на расстоянии распространения 102,7 мм. Метрика Дюбуа имеет проблемы с паразитными полосами (созданными многократным отражением на основе ложной интерференции) и большим полем зрения, проанализированным в[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003) [инкапсуляция](https://www.sciencedirect.com/topics/physics-and-astronomy/encapsulating)большого количества [микросфер](https://www.sciencedirect.com/topics/engineering/microsphere) и частиц пыли. Предлагаемые метрики DarkFocus и эталонного ToG демонстрируют устойчивость к упомянутым источникам ошибок, обеспечивая высокую точность.

Визуализации 1, 2 и 3 добавлены к[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003)представляют динамические последовательности изображений числовой перефокусировки выбранной области интереса в амплитудном режиме, режиме темного поля и градиентном режиме темного поля (DarkFocus) соответственно, демонстрируя микросферы и перефокусировку частиц пыли. Кроме того, включена Visualization4 для представления различной автоматической числовой перефокусировки области интереса, выделяя несколько бусин и частиц пыли.

Все данные представлены в[Рис. 3](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0003)подтверждают метрики ToG и DarkFocus и подчеркивают неотъемлемые особенности недавно предложенной схемы оценки автоматической фокусировки. Однако метод Дюбуа оказался менее точным.   
[Рис. 4](https://www.sciencedirect.com/science/article/pii/S0143816620300762?via%3Dihub#fig0004)блок [-схема](https://www.sciencedirect.com/topics/engineering/block-diagrams) для расчета метрики DarkFocus для заданного расстояния распространения z. Обратите внимание, что процедура, описанная на схеме, повторяется для всех предварительно выбранных расстояний z. Однако фильтрация голограммы выполняется только один раз.  


## Трехмерная визуализация

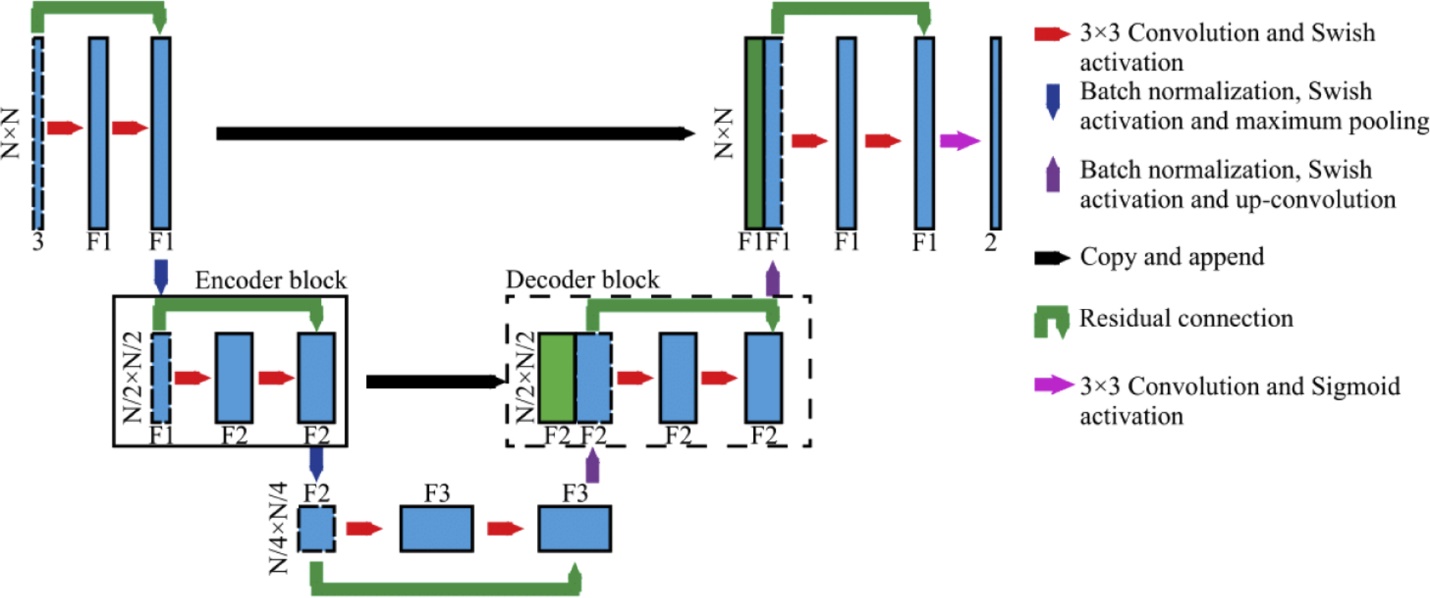
В статье [3] предлагается новый подход на основе обучения для 3D-визуализации поля частиц с использованием голографии. Предложенный подход подход использует архитектуру U-net, включающую остаточные связи, функцию активации Swish, предварительную обработку голограмм и перенос обучения, чтобы справиться с проблемами, возникающими в голограммах частиц, где точное измерение отдельных частиц имеет решающее значение. Оценки как синтетических, так и экспериментальных голограмм демонстрируют значительное улучшение скорости извлечения частиц, точности локализации и скорости по сравнению с предыдущими методами в широком диапазоне концентраций частиц, включая высокоплотные концентрации, где другие методы не подходят.  
В последнее время машинное обучение с использованием глубоких нейронных сетей (DNN) стало преобладающим инструментом для различных задач анализа изображений. Внедрение DNN значительно увеличило скорость обработки и дало более точные результаты, чем обычные обратные подходы для некоторых приложений [ [32](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref32) ] . Однако по сравнению с другими областями вычислительной визуализации машинное обучение недостаточно использовалось в DH . Машинное обучение в DH было принято для преобразования голографических реконструкций в микроскопические изображения [,](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) аналогичные тем, которые обычно используются в биологических и медицинских исследованиях, и классификации видов частиц, захваченных в голограмме.

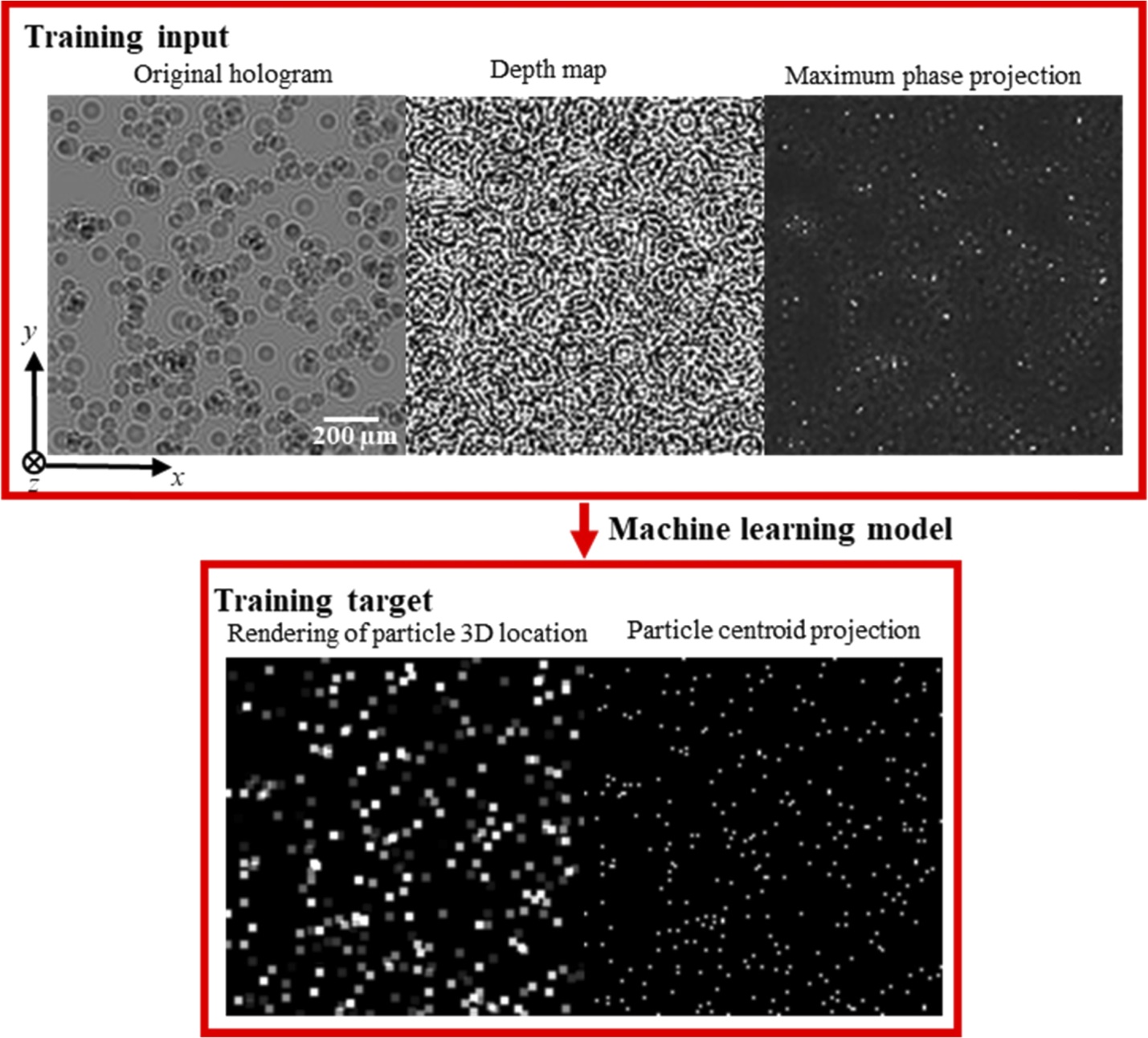
 В частности, Ren [et](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137" \l "ref43)*al* . [ [41](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref41) ] продемонстрировали, что сверточная нейронная сеть (CNN) дает более точную глубину частиц, чем обычные методы реконструкции и другие подходы к машинному обучению. На сегодняшний день только Шимобаба *и др.* [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ] применили машинное обучение для реконструкции многообъектного поля частиц из голограмм с помощью подхода сегментации машинного обучения. Они использовали архитектуру U-net CNN с функцией потерь, регуляризованной с помощью *L* 1, и обучались на синтетических голограммах с концентрацией частиц, варьирующейся от 4,7×10−5 частиц /пиксель (ppp) до 1,9×10−4 ppp . Их алгоритм продемонстрировал хорошие результаты реконструкции для синтетических голограмм с низкой концентрацией в присутствии гауссовского шума с быстрым снижением производительности с увеличением концентрации частиц. Такое увеличение концентрации обычно требуется для многих практических измерительных приложений. Кроме того, метод регуляризации, используемый в их подходе, имеет тенденцию быть нестабильным, что влияет на сходимость решения [ [45](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref45) ].

На основе обзора литературы и сравнения с другими задачами обработки изображений на основе обучения с учителем авторы выявили три уникальные проблемы, связанные с 3D-визуализацией частиц с использованием DH. Во-первых, в то время как сигнал отдельного объекта может распространяться на большую область голограммы, реконструированное поле частиц обычно состоит из группы разреженных объектов. Когда для замены реконструкции используется подход на основе обучения, эта разреженность приводит к тому, что процесс обучения становится крайне нестабильным и дает неверные результаты. Во-вторых, реконструкция 3D-поля частиц требует очень точных измерений для каждой частицы, что отличается от многих обычных задач визуализации на основе обучения, таких как классификация или глобальная регрессия изображения. Наконец, желаемые метрики, параметры записи и внешний вид голограммы связаны, что ограничивает обобщаемость модели, обученной на определенном наборе данных. Стоит отметить, что эти проблемы могут также возникать при получении изображений в световом поле [ [6](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref6) , [46](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref46) , [47](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref47) ], при получении изображений через диффузные среды [ [48](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref48) ], при получении изображений с дефокусировкой [ [49](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref49) , [50](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref50) ] и других методах [ [8](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref8) ].

Для решения вышеупомянутых проблем авторы представляют специально разработанный подход машинного обучения для реконструкции трехмерного поля частиц в DH, который также может быть использован в других задачах вычислительной визуализации, имеющих схожие черты.

Подход машинного обучения для реконструкции поля частиц из голограммы использует специально разработанную архитектуру U-net, которая принимает голограммы в качестве входных данных и вычисляет трехмерное местоположение частиц в качестве выходных данных. U-net — это тип CNN, разработанный для сегментации медицинских и биологических изображений [ [51](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref51) , [52 ]](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref52)[,](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref53) который также использовался в преобразованиях изображений на основе обучения [ [33–38](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref33) , [53](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref38) ] и многообъектной классификации из отдельных изображений [ [54](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref54) ]. U-net состоит из ряда блоков кодера и декодера, соответствующих сплошным и пунктирным черным квадратам на рис.  [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g001) соответственно. В блоке кодера два последовательных набора слоев свертки и функций активации используются для кодирования локальных признаков входных изображений в каналы. Два блока кодера соединены максимальным слоем объединения, который понижает дискретизацию карт признаков для извлечения глобальных признаков. Блок декодера аналогичен, но работает наоборот. Два последовательных сверточных слоя используются для декодирования каналов для формирования изображения, а два блока декодера соединены с помощью слоев восходящей свертки для изменения размера карт признаков. Обратите внимание, что выходная карта признаков конечного блока кодера соединена с первым блоком декодера через слой восходящей свертки. U-net также включает пропускные соединения (черные стрелки на рис.  [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g001) ), посредством чего выход кодера объединяется с блоком декодера того же размера, который объединяет локальные и глобальные признаки изображений для обучения на более глубокой стадии сети. По сравнению с простой архитектурой CNN без пропускных соединений мы предполагаем, что U-net больше подходит для реконструкции поля частиц из голограммы, поскольку пропускные соединения используют распространение информации об отдельных частицах по большой части изображения (как в локальном, так и в глобальном масштабах). В данной архитектуре U-net используется U-net с 4 кодерами и 3 декодерами, а количество выходных каналов кодера составляет 64 и 512 для первого и последнего кодера соответственно. Ключевой особенностью архитектуры U-net является то, что ее можно напрямую применять к изображениям произвольного размера (независимо от размера обучающего набора изображений), поскольку в ней нет плотно связанных слоев.

  
По сравнению с обычной архитектурой U-net, представленная архитектура U-net имеет остаточное соединение в каждом блоке кодера и декодера (зеленые стрелки на рис.  [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g001) ) и использует функцию активации Swish (Sigmoid-weighed Linear Unit) (уравнение ( [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e1) )) для всех слоев, кроме последнего. Остаточное соединение увеличивает скорость обучения и снижает вероятность того, что обучение застрянет в локальном минимуме. В декодере остаточное соединение использует каналы из предыдущего блока декодера, соединенные слоем свертки вверх. Такая конфигурация обеспечивает необходимое сокращенное соединение (т. е. пропуск одного или двух слоев свертки) для остаточной сети. Х в уравнении ( [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e1) ) соответствует выходным данным из предыдущего слоя, а *f* ( *x* ) является входными данными для следующего слоя.  
Для голограмм частиц целевые изображения (рис.  [2](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g002) ) обычно разрежены из-за малого размера центроидов частиц, что приводит к тому, что большинство значений в слоях признаков равны 0. Поэтому во время обучения параметры внутри модели имеют более высокую тенденцию быть равными 0, и впоследствии не поддаются обучению, поскольку производная ReLU равна 0, когда веса равны или меньше 0. Эта проблема вызывает существенное ухудшение для глубоких моделей CNN, использующих функцию активации ReLU. Для сравнения, функция Swish является гладкой и немонотонной вблизи 0, что увеличивает количество эффективных параметров в обучении, особенно для разреженных целей. Однако Swish может повлиять на точность прогноза из модели из-за включения отрицательных выходных значений. Чтобы решить эту проблему, используется функция активации Sigmoid в конечном блоке декодера (пурпурная стрелка на рис.  [1](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g001) ) для получения результатов в диапазоне от 0 до 1.

  
**Рис. 2.** Образец обучающего входа и обучающей цели, состоящей из 300 частиц (т.е. концентрация 0,018 ppp) с размером голограммы 128 × 128 пикселей. Голограмма сформирована с разрешением пикселей 10 мкм с длиной волны лазерного освещения 632 нм.

Обучающий вход состоит из трех каналов: исходная голограмма, соответствующие изображения проекции глубины пикселей (т. е. карта глубины) и проекция максимальной фазы (рис.  [2](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g002) ). Исходная синтетическая голограмма генерируется в соответствии с подходом в [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ] и [ [57](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref57) ]. Разрешение пикселей составляет 10 мкм с длиной волны лазерного освещения 632 нм и размером изображения 128 × 128 пикселей. Частицы внутри голограмм случайным образом распределены на расстоянии от 1 мм до 2,28 мм от датчика. Разрешение в направлении *z* также составляет 10 мкм со 128 дискретными уровнями глубины. По сравнению с [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ], карта глубины и проекции фазы являются дополнительной информацией, полученной в результате предварительной обработки голограмм. Как предложено в [ [32](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref32) ], предварительная обработка используется для включения существующих знаний о формировании голограмм в модель и снижения необходимости модели полностью изучать требуемую физику во время обучения. Кроме того, обучение с известной физикой формирования голограммы вместо того, чтобы полагаться исключительно на обучение модели, позволяет избежать ложных и нефизических выходов обученной модели. Наши первоначальные тесты показали заметное улучшение скорости извлечения частиц, особенно для случаев высокой концентрации с этими этапами предварительной обработки по сравнению с обучением непосредственно на необработанных голограммах. Используя метод углового спектра [ [58](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref58) ], трехмерное комплексное оптическое поле *u*p ( *x* , *y* , *z* ) генерируется из исходной голограммы *I(x,y)* (уравнение ( [2](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e2) )), где *λ* — длина волны, *k* — волновое число иФФ— оператор преобразования Фурье. Карта глубины генерируется путем проецирования точек *z* , где пиксели имеют максимальную интенсивность, из *u*p ( *x* , *y* , *z* ) на плоскость *xy* (уравнение ( [3](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e3) )), а максимальная фазовая проекция рассчитывается из уравнения ( [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e4) ).

Тренировочная цель состоит из двух выходных каналов. Первый — это канал в оттенках серого, в котором интенсивность пикселей соответствует относительной глубине каждой частицы, а второй — это бинарное изображение центроидов *xy* частиц (рис.  [2](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g002) ). В то время как частицы кодируются как только один пиксель в бинарном канале *xy* , выполнение того же самого для канала в оттенках серого с кодированием глубины дает обученную модель, которая генерирует выходные данные с неточными интенсивностями пикселей и значительным фоновым шумом. Чтобы предотвратить это, маркированные частицы в цели в оттенках серого с кодированием глубины устанавливаются в размере 3×3 пикселя.

Из-за различий между двумя целевыми каналами каждый канал использует другой тип функции потерь. В частности, потери Хьюбера [ [59](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref59) ] оцениваются по выходному каналу, кодирующему глубину частицы. Как показано в уравнении ( [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e4) ), он использует модифицированную среднюю абсолютную ошибку (MAE) прогноза ( *Y* ) относительно истинного значения ( *X* ) в качестве потери обучения, когда MAE больше предустановленного *δ* (0,002 для синтетического набора данных), и использует среднюю квадратичную ошибку (MSE), когда MAE меньше *δ* . Потеря Хьюбера повышает надежность обучения и точность прогнозирования за счет использования MAE, когда усредненные интенсивности пикселей смещены выбросами [ [59](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref59) ]. Мы предполагаем, что параметр *δ* в уравнении ( [5](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e5) ) может быть определен на основе требований к точности измерения, при этом меньшее *δ* приводит к улучшенному разрешению глубины частицы. Однако слишком малое значение *δ* может привести к нестабильному процессу обучения и иметь множественные решения, аналогичные использованию чистой потери MAE [ [45](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref45) ].

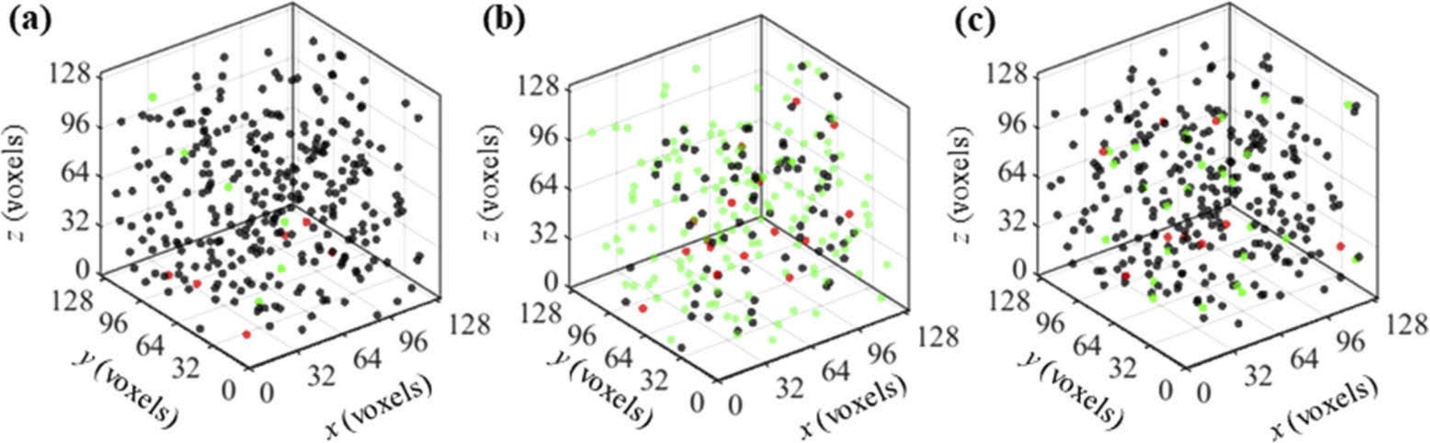
*Для канала центроида xy* (уравнение ( [6](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e6) )) используется потеря MSE, регуляризованная полной вариацией (TV) прогноза . Как показано в уравнении ( [7](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e7) ), TV представляет собой сумму градиентов первого порядка по изображению размером *N x* × *N y* .

Регуляризация TV ранее применялась в методах итеративной оптимизации для реконструкции голограмм [ [30](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref30) , [31](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref31) ]. TV устойчив к выбросам в изображениях и заставляет модель создавать гладкий фон в выходном канале *xy* -центроида. Такая регуляризация снижает вероятность того, что фоновые пиксели будут иметь ненулевые значения, что приведет к обнаружению частиц-призраков. α *в* уравнении ( [5](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e5) ) — это параметр, определяющий гладкость результатов. Мы предлагаем небольшое значение *α* (∼0,0001) для обучения, поскольку регуляризация TV действует как фильтр нижних частот, а слишком сильное сглаживание может ухудшить точность результатов.

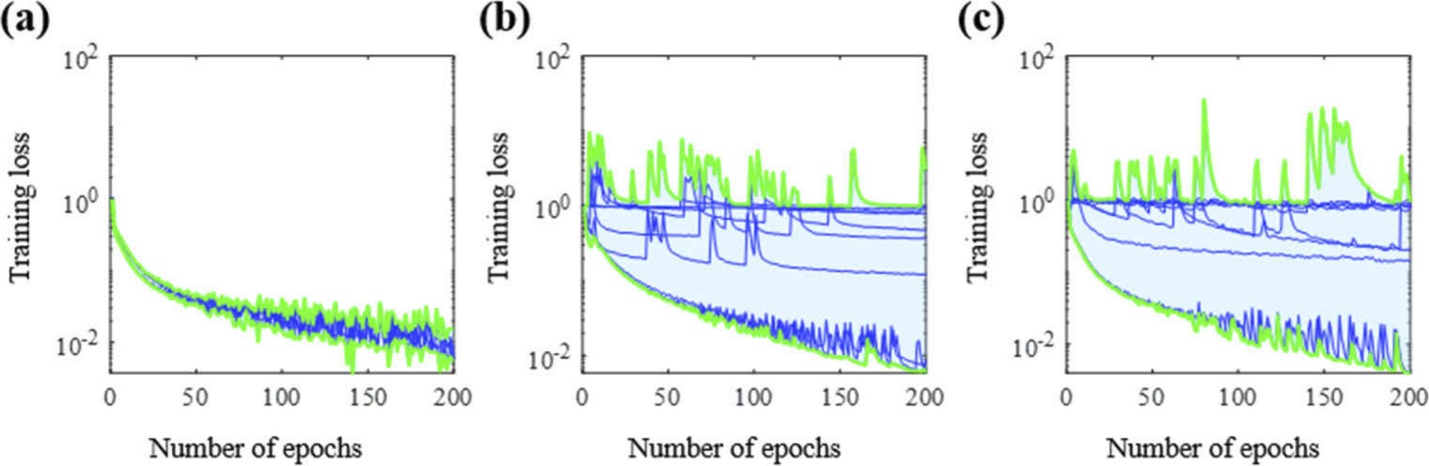
Архитектура U-net реализована с использованием Keras [ [60](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref60) ] с бэкэндом TensorFlow. Обучение проводится на графическом процессоре Nvidia RTX 2080Ti. Оптимизатор Adam [ [61](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref61) ] используется со скоростью обучения по умолчанию 0,001. Для обучения и тестирования моделей генерируются тринадцать наборов данных с концентрацией частиц от 1,9×10 −4 до 6,1×10 −2 ppp. Самая высокая концентрация частиц синтетических данных в 305 раз выше, чем самый высокий случай (1,9×10 −4 ppp) в литературе [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ]. Сначала базовая модель обучается на 2500 голограммах с концентрацией частиц 1,8×10 −2 ppp в течение 480 эпох (всего 13,5 часов). Для каждой последующей концентрации частиц набор данных из 2000 голограмм обучается в течение 120 эпох с обучением, инициализированным базовой моделью (2,7 часа для каждого случая). Этот подход к переносу обучения существенно снижает требования к обучению (т. е. размер набора данных и время обучения) для новых наборов данных голограмм [ [62](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref62) ]. Чтобы извлечь частицы из выходных данных модели, предсказанная карта центроидов *xy* частиц сначала бинаризуется с порогом 0,5 (эквивалентно максимальному правдоподобию) для извлечения центроидов *xy* частиц. Затем из выходных данных в оттенках серого с кодировкой глубины мы используем значение интенсивности соответствующих пикселей на карте глубины в качестве глубины частицы.

#### 3.1 Оценка с использованием синтетических голограмм с постоянной концентрацией частиц

Сначала мы сравниваем производительность нашего метода с реализацией подхода, предложенного Шимобаба *и др* . [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ] на тестовом наборе из 100 голограмм с той же концентрацией частиц. Поскольку синтетические голограммы не зависят от конкретных экспериментальных настроек, а разрешение дискретизации голограммы на основе вокселей в значительной степени определяет точность нашей измерительной системы с точки зрения цифровой обработки изображений, мы представляем все результаты синтетических голограмм в единицах вокселей. Обучение подхода Шимобаба проводится на наборе данных из 9000 голограмм с концентрацией частиц от 1,9×10−4 до 6,1×10−2 ppp (что совпадает с нашими синтетическими наборами данных). Сначала сравнение проводится на голограмме из 300 частиц (рис.  [3](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g003) ). Сопряжение предсказанных частиц с истинной следует методу, представленному в [ [31](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref31) ]. Как показано на рис.  [3](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g003) , наш предлагаемый подход дает значительно большее количество извлеченных частиц по сравнению с подходом Simobaba. Наш метод дает процент извлечения 98,7% (1,3% ложноотрицательных) с процентом ложноположительных (непарные частицы из прогноза) 2,3%. Для сравнения, подход Shimobaba достигает 40,4% процента извлечения (59,6% ложноотрицательных) и 17,5% ложноположительных, в то время как современная регуляризованная обратная голографическая объемная реконструкция или RIHVR [ [31](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref31) ] достигает 88,4% (11,6% ложноотрицательных) и 2,3% для показателей извлечения и ложноположительных, соответственно. Все эти три метода имеют медианную ошибку позиционирования менее 1 пикселя для *x* и *y* . Ошибка *z* для нашего метода составляет 1,48 вокселя по сравнению с 5,49 вокселями для Shimobaba и 3,50 вокселями для RIHVR.

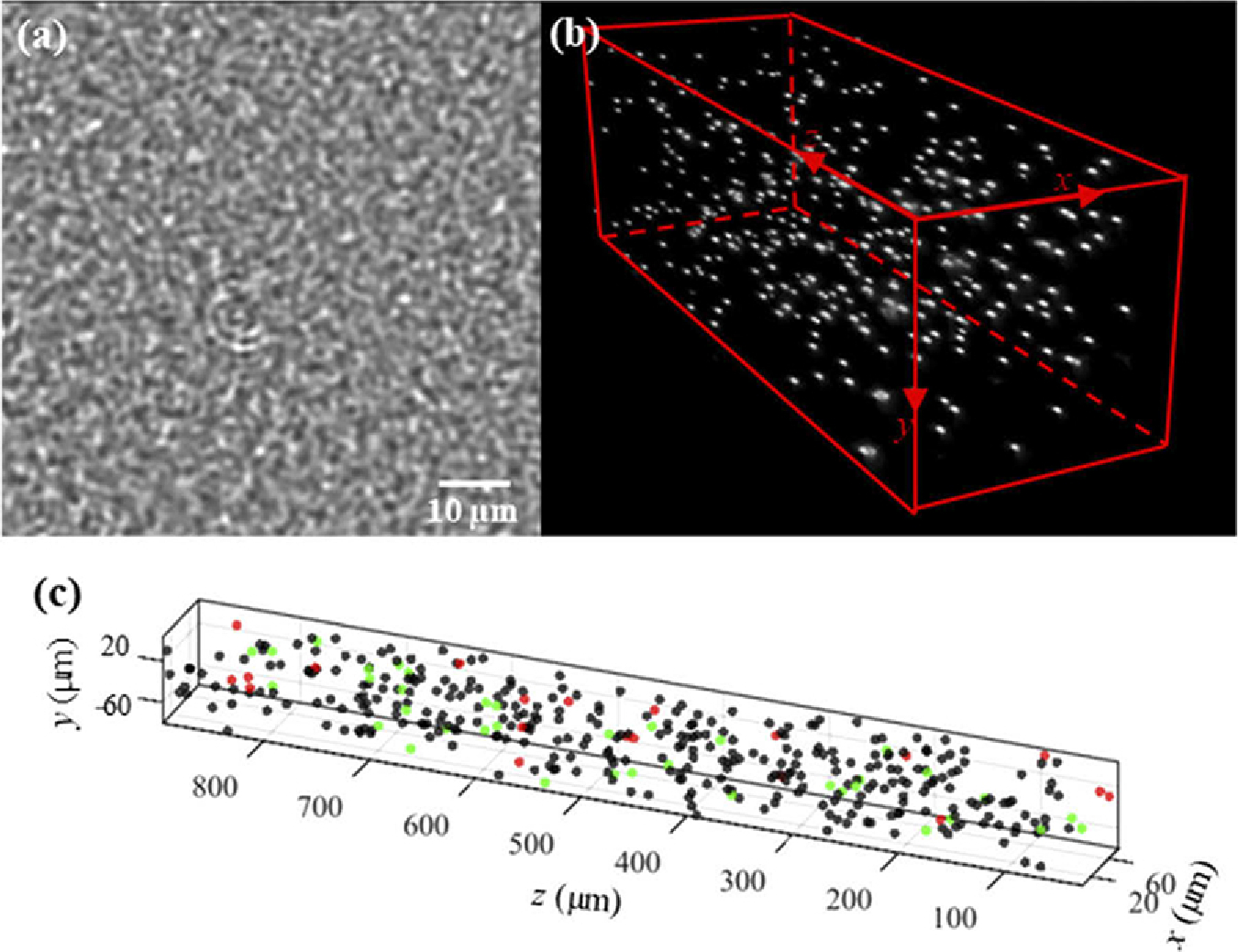
  
**Рис. 3.** Результаты прогнозирования с помощью обученной модели с использованием (a) нашей архитектуры U-net и (b) метода, представленного в Shimobaba et al. [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ] (c) и Mallery and Hong [ [31](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref31) ] для случая 0,018 ppp (голограммы из 300 частиц). Черные точки — извлеченные истинные частицы, красные точки — ложноположительные (т. е. непарные частицы из истинной наземной среды), а зеленые точки — ложноотрицательные (непарные частицы из истинной наземной среды).

Чтобы проанализировать влияние уникальных особенностей в нашей архитектуре U-net на процесс обучения, мы сравниваем спад потерь обучения для первых 200 эпох различных вариантов модели на наборе данных 1,8×10−2 ppp . Как показано на рис.  [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g004) , по сравнению с предлагаемым методом, удаление остаточных связей или использование обычной функции потерь (MSE) дестабилизируют процесс обучения. Удаление остаточных связей (рис.  [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g004) (b)) приводит к тому, что процесс обучения становится восприимчивым к локальным минимумам во время обучения, как обсуждалось Хе *и др.* [ [55](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref55) ]. Кроме того, при использовании функции потерь, восприимчивой к выбросам (такой как MSE) без какой-либо регуляризации, модель, вероятно, будет выдавать тривиальные предсказания на выходе (т. е. ложные частицы), что вызывает флуктуацию кривой потерь даже на ранних этапах обучения (рис.  [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g004) (c)). В результате модель не сходится к оптимальному решению и выдает очень неточные прогнозы интенсивности пикселей или результаты в виде белого шума для наихудших сценариев.   
Наконец, замена функций активации ReLU на Swish оптимизирует процесс обучения, избегая необучаемых параметров (т. е. мертвых нейронов) [ [56](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref56) ]. Согласно нашим тестам, случаи без Swish могут производить >6000 мертвых нейронов в последнем слое свертки в конце первой эпохи, что существенно ухудшает модель. В результате обучение, вероятно, достигнет плато в первые несколько эпох, и полученные модели будут генерировать двумерные изображения белого шума.

  
**Рис. 4.** Демонстрация влияния предлагаемых улучшений модели на процесс обучения в течение первых 200 эпох. (a) Предлагаемый подход, (b) использование архитектуры U-net без остаточной связи и (c) использование среднеквадратичной ошибки в качестве функции потерь. Потери нормализуются по их начальному значению, и каждый случай случайным образом инициализируется 10 раз, чтобы показать результирующую нестабильность обучения для случаев (b) и (c). На изображении зеленые кривые соответствуют максимальным и минимальным нормализованным потерям в каждой эпохе, синие кривые соответствуют каждой инициализации, а заштрихованная область представляет собой диапазон потерь.

#### 3.3 Оценка с использованием экспериментальных данных

Предложенный метод оценивается с использованием экспериментальных голограмм флуоресцентно меченых частиц, внедренных в твердый гель. Мы используем флуоресцентные частицы размером 2 мкм (ThermoScientific) с концентрацией ∼5000 частиц/мм 3 (2,0×10−2 ppp ), диспергированные в желатине на водной основе, помещенном между предметными стеклами. Экспериментальные голограммы (рис.  [7](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g007) (а)) регистрируются на инвертированном микроскопе Nikon Ti-Eclipse с использованием объектива Nikon 10X и Andor Zyla 5.5 sCMOS (0,65 мкм/пикс) с коллимированным 660 нм диодным лазерным освещением (QPhotonics; QLD-660-10S). Микроскоп может записывать несколько голограмм размером 2432×2048 пикселей на расстояниях от 0 до 200 мкм под объемом, все из которых используются для расчета усредненного по ансамблю фона для улучшения изображения. Измерение истинности (рис.  [7](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g007) (b)) получается путем сканирования образца в том же месте с использованием эпифлуоресцентного режима микроскопа с шагом 2,5 мкм по всей глубине при том же размере изображения и разрешении. Положения частиц для измерения истинности затем получаются с помощью ручного порогового значения и расчета центроида с весовой интенсивностью. Обучающий набор данных состоит из 7500 случайно обрезанных плиток размером 128×128 пикселей из улучшенной голограммы и соответствующих им местоположений частиц из сканируемого объема 3D флуоресценции. Здесь целевые изображения сохраняются с 16-битной точностью для кодирования большого количества плоскостей реконструкции. Потеря Хьюбера *δ* в уравнении. ( [4](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#e4) ) установлено как 0,001, что ниже, чем в синтетических случаях, поскольку интенсивность пикселей в метках кодирует более высокое разрешение по глубине ( направление *z* в экспериментальном случае). Мы используем тот же метод, что и в синтетическом случае, чтобы связать предсказанные частицы с истинными данными. Как показано на рис.  [7](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g007) (c), несмотря на шумный вход (рис.  [7](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#g007) (a)), предсказанные результаты обученной модели хорошо совпадают с истинными данными. Тест 100 случайно обрезанных плиток размером 128×128 пикселей из проверочной голограммы, отображающей другую область этого образца, дает 90%-ную скорость извлечения с ошибкой позиционирования менее 1 воксела (0,65 мкм) в направлениях *x* и *y* и 5,24 воксела (13,1 мкм) в направлении *z* . Обучение этого набора данных с использованием подхода Shimobaba [ [44](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref44) ] не может сойтись к низкому значению потерь и дает модель, генерирующую сильный фоновый шум. В этом случае трудно применить RIHVR, поскольку удаление фона, указанное в [ [31](https://opg.optica.org/oe/fulltext.cfm?uri=oe-28-3-2987&id=426137#ref31) ], зависит от движения частиц для получения точного усредненного по времени фонового изображения. Таким образом, RIHVR имеет очень высокий уровень ложных срабатываний для этого случая.

  
**Рис. 7.** (a) 128×128-пиксельная улучшенная голограмма из экспериментальных данных и соответствующее объемное изображение посредством наложения флуоресцентного сканирования в светлом поле того же образца для определения истинности (b). (c). Результаты прогнозирования в сравнении с моделью машинного обучения. Черные точки — извлеченные истинные частицы, красные точки — ложноположительные (т. е. непарные частицы из истинности), а зеленые точки — ложноотрицательные (непарные частицы из истинности).

Список литературы

1. Chen H., Huang L., Liu T., Ozcan A. eFIN: Enhanced Fourier Imager Network for Generalizable Autofocusing and Pixel Super-Resolution in Holographic Imaging // IEEE Journal of Selected Topics in Quantum Electronics. 2023. 29. 4: Biophotonics. С. 1–10.

2. Galande A. S., Thapa V., Gurram H. P. R., John R. Untrained deep network powered with explicit denoiser for phase recovery in inline holography // Applied Physics Letters. 2023. 122. № 13.

3. Shao S., Mallery K., Kumar S. S., Hong J. Machine learning holography for 3D particle field imaging // Optics Express. 2020. 28. № 3. С. 2987–2999.

4. Trusiak M., Picazo-Bueno J.-A., Zdankowski P., Micó V. DarkFocus: numerical autofocusing in digital in-line holographic microscopy using variance of computational dark-field gradient // Optics and Lasers in Engineering. 2020. 134. С. 106195.