

Исследование генеративно-состязательных нейронных сетей для решения задачи реконструкции изображений в дифракционно-оптических системах

В.В. Евдокимова^{1,2}, М.В. Петров^{1,2}, М.А. Ключева², А.П. Алексеев², С.А. Бибиков^{1,2},
Р.В. Скиданов^{1,2}, А.В. Никоноров^{1,2}

¹Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

²Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В данной работе проводится исследование причин возникновения артефактов, возникающих в результате цветовой коррекции изображений на основе генеративно-состязательной нейронной сети. Исходное изображение получено с помощью многоуровневой дифракционной линзы. В данной работе показано, что на наличие артефактов влияют переэкспонирование фрагментов исходного изображения и наличие маркеров на изображениях обучающей выборки. В работе предлагается архитектура генеративно-состязательной сети с масштабирующим слоем, обеспечивающим увеличение разрешения восстановленного изображения в два раза.

1. Введение

В настоящее время все больше исследований проводится в области реконструкции изображений, полученных на основе плоских оптических элементов [1-6]. Использование плоской оптики позволяет значительно сократить вес и упростить конструкцию оптической системы, что является большим преимуществом перед классическими рефракционными аналогами. Наиболее популярным подходом к проектированию плоской оптики является использование т.н. *металинз* [1], которые позволяют создать оптическую систему с субволновой толщиной. Другой пример плоской оптики – многоуровневые дифракционные линзы (МДЛ), которые стали использоваться для получения цветных изображений высокого разрешения только с 2015 года [2]. Подход, основанный на использовании метаповерхностей, обеспечивает меньшую толщину линзы, однако МДЛ проще в изготовлении и позволяют получить изображения лучшего качества.

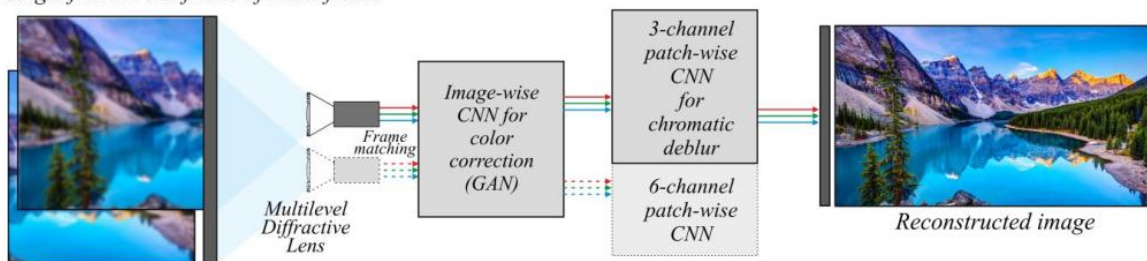
Несмотря на все достоинства для изображающих систем на основе плоской оптики характерны хроматические aberrации, которые являются причиной существенных искажений на регистрируемых изображениях. В работах [2-6] описываются различные методы компенсации возникающих искажений в системах на основе МДЛ. В работе [5] предлагается сквозная нейросетевая технология реконструкции изображений и впервые используется генеративно-состязательная сеть (GAN) для решения задачи цветовой коррекции. Предложенный пайплайн реконструкции обучен на наборе данных, который сформирован путем съемки изображений, выведенных на экран монитора [5], и обеспечивает на тестовой

выборке 26 Дб по критерию PSNR. Однако на изображениях реальных сцен после цветовой коррекции на основе GAN наблюдались артефакты, природа которых не была исследована. В данной работе проводится исследование причин появления артефактов на изображениях реальных сцен после GAN обработки, а также методов их устранения.

2. Сквозная нейросетевая технология реконструкции изображений

На рисунке 1 представлена схема сквозной нейросетевой технологии реконструкции изображений, полученных в монокулярной и бинокулярной системах на основе МДЛ [5]. На изображениях, полученных с помощью МДЛ выделяются два типа искажений: локальные искажения, вызванные хроматической расфокусировкой (хроматическое размытие), и глобальные цветовые искажения, вызванные перераспределением энергии между дифракционными порядками (контекстно-зависимый цветовой сдвиг). Разработанная в [5] технология реконструкции позволяет устранить описанные искажения благодаря комбинации двух сверточных нейронных сетей (СНС) глобально (image-wise) и локально (patch-wise) действующих на изображение.

Single frame/First frame of multi-frame



Second frame of multi-frame

Рисунок 1. Пайплайн реконструкция изображений, описанный в работе [5].



Рисунок 2. Пример изображений из тестовой выборки и результаты реконструкции.



Рисунок 3. Примеры изображений реальных сцен и результаты их цветовой коррекции.

Для устранения контекстно-зависимых цветовых искажений в [5] впервые предлагается подход на основе GAN, где используется архитектура Pix2Pix [7]. Для компенсации

хроматического размытия использовались 3-х канальная модификация VDSR [5] и DRRN [8] в случае монокулярной системы и 6-ти канальная модификация VDSR для бинокулярной системы.

Описанный пайплайн позволил достичь 26 Дб по критерию PSNR на тестовой выборке, а также обеспечил хорошее визуальное качество реконструкции (рисунок 2). Однако на обработанных снимках реальных сцен наблюдаются значительные артефакты, которые возникли в результате цветовой коррекции на основе GAN (рисунок 3).

3. Методы исследования причин возникновения артефактов

3.1. Коррекция экспозиции входного изображения

Возможной причиной возникновения артефактов является переэкспонирование снимка (отсутствие деталей в светлых частях изображения), которое чаще всего вызвано неверной экспозицией. На рисунке 3 наблюдаются пересветы в тех областях изображений, где после GAN обработки возникают артефакты.

В работе предлагается скорректировать значение экспозиции с некоторым шагом и визуально сравнить результаты GAN-обработки изображений с скорректированной экспозицией. Также предлагается выполнить выравнивание гистограммы всех изображений (обучающих, тестовых, реальных) средствами OpenCV и переобучить GAN.

3.2. Влияние маркеров изображений обучающей выборки на возникновение артефактов

Предполагается, что наличие крестообразных маркеров по углам (рисунок 2) изображений обучающей выборки также приводит к возникновению артефактов. Данные маркеры были добавлены для сопоставления изображений, снятых на основе МДЛ, и оригинальных изображений [5]. Ранее в работах [2-6] влияние маркеров на качество реконструкции не исследовалось. В данной работе предлагается сформировать новый обучающий набор, путем обрезания маркеров на изображениях.

4. Масштабирование изображений на основе GAN

В работе предлагается модификация используемой в работе [5] Pix2Pix, которая позволяет увеличивать разрешение цветокорректированного изображения. Pix2Pix состоит из U-Net в качестве генератора и PatchGAN в качестве дискриминатора. Общая архитектура GAN представлена на рисунке 4.

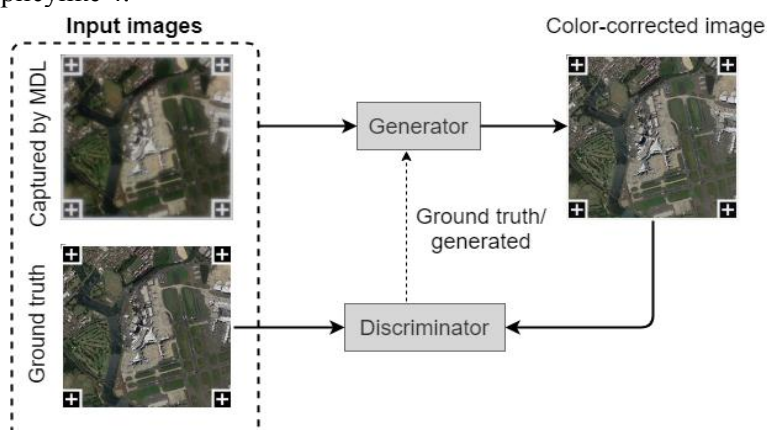


Рисунок 4. Цветовая коррекция на основе GAN.

В работе [5] использовалась следующая архитектура генератора:

Encoder: C64-C128-C256-C512-C512-C512-C512,

Decoder: CD512- CD512- CD512-C512-C256-C128-C64,

где C_k – это комбинация сверточного слоя, пакетной нормализации и функции активации ReLU с количеством фильтров k , CD_k означает, что перед ReLU добавляется dropout-слой. Размер

окна свертки во всех случаях - 4×4 . На каждом сверточном слое кодировщика (encoder) ширина и высота изображения уменьшаются в 2 раза, сверточные слои декодировщика напротив увеличивают разрешение изображения. Используя описанную архитектуру генератора, на выходе сети всегда будет изображение такого же размера, что и входное.

В данной работе предлагается добавить дополнительную комбинацию слоев C64 для увеличения разрешения выходного изображения в два раза. Таким образом, архитектура генератора примет вид:

Encoder: C64-C128-C256-C512-C512-C512-C512-C512,
Decoder: CD512- CD512- CD512-C512-C256-C128-C64-C64.

5. Результаты экспериментов

5.1. Набор изображений

В работе использовался набор данных, сформированный в [5] для монокулярной системы на основе МДЛ. Набор данных состоит из 350 пар изображений, снятых МДЛ, и оригинальных изображений размера 1024×1024 . Примеры изображений представлены на рисунке 2.

Кроме того, на основе данного набора был сформирован новый набор данных, путем обрезания крестообразных маркеров на изображении. Таким образом, новый набор составил 350 пар изображений размера 512×512 . Также в работе использовались изображения реальных сцен, представленные на рисунке 3.

5.2. Исследование влияния эффекта переэкспонирования на возникновение артефактов

Коррекция экспозиции проводилась на интервале $[-2, 2]$ с шагом $h=0,25$. На рисунке 5 представлены примеры изображений с скорректированным значением экспозиции. Слева изображение с экспозицией, равной -1, в центре – исходное изображение, а справа указано изображение с экспозицией, равной 1.

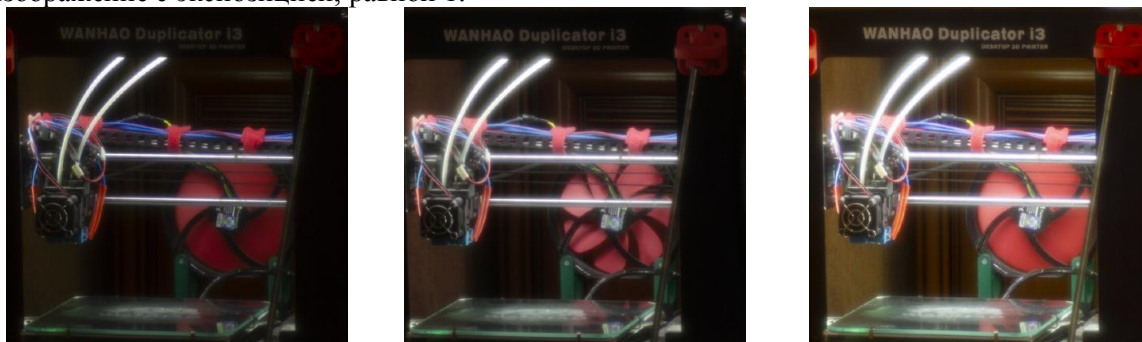


Рисунок 5. Изображение реальной сцены. Корректирование экспозиции: -1 (слева), 0 (в центре), 1 (справа).

На рисунке 6 представлены результаты обработки сетью Pix2Pix [5], обученной на исходном наборе изображений размера 1024×1024 . Видно, что изменение экспозиции влияет на количество артефактов. При значении экспозиции, равным -1, образуется немного меньше артефактов, чем в остальных случаях. Однако добиться полного устранения артефактов путем изменения экспозиции не удалось.

Также был проведён эксперимент, при котором проводилось выравнивание гистограмм всех изображений реальных сцен и исходной выборки. Изображения переводились в цветовое пространство HSV, после чего выполнялась эквализация гистограммы компоненты V средствами OpenCV. Данная процедура не привела к видимым улучшениям качества реконструкции.

5.3. Исследование влияния маркеров изображения на возникновение артефактов

Сеть Pix2Pix, используемая в работе [5], была обучена на новом сформированном наборе, состоящим из фрагментов размера 512×512 изображений исходного набора без крестообразных

маркеров. На рисунке 7 представлены результаты обработки изображений реальных сцен сетью, обученной на новом наборе. Наблюдается значительное сокращение количества артефактов.

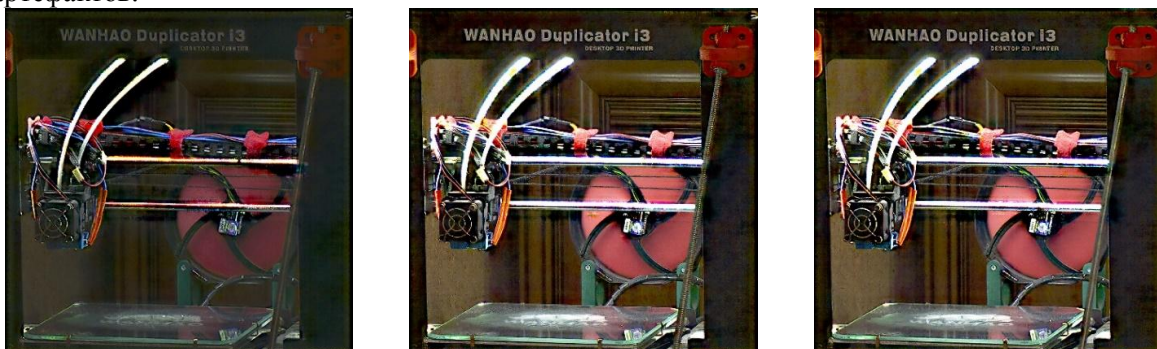


Рисунок 6. Результат Pix2Pix обработки. Экспозиция: -1 (слева), 0 (в центре), 1 (справа).



Рисунок 7. Эффект предобработки обучающей выборки.

На рисунке 8 представлены примеры обработки данной сетью изображений с разными значениями экспозиции. Видно, что значение экспозиции в данном случае также влияет на количество артефактов.

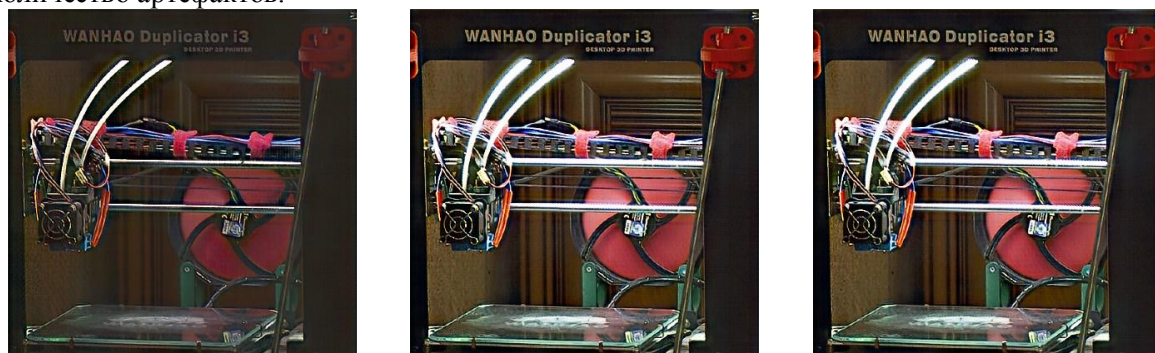


Рисунок 8. Результат Pix2Pix обработки. Экспозиция: -1 (слева), 0 (в центре), 1 (справа).

5.4. Оценка качества реконструкции Pix2Pix с масштабирующим слоем

Использование обучающей выборки изображений без крестообразных маркеров привело к тому, что на выходе Pix2Pix разрешение изображения уменьшилось в два раза. В данной работе предлагается использовать Pix2Pix с масштабирующим слоем для получения изображений исходного разрешения 1024×1024 . На рисунках 7 и 8 обработка снимков реальных сцен проводилась на основе сети с масштабирующим слоем. Среднее значение PSNR на тестовой выборке при размерностях входного и выходного изображений 512×512 и 1024×1024 соответственно составило 22 Дб. Для размерностей 256×256 (на входе) и 512×256 (на выходе) среднее значение PSNR составило 21,8 Дб.

6. Заключение

В данной работе проведено исследование причин возникновения артефактов на изображениях реальных сцен в результате цветовой коррекции на основе GAN. Во-первых, проведено исследование влияния эффекта переэкспонирования на появление артефактов. Было обнаружено, что изменение значения экспозиции оказывает влияние на количество артефактов, однако не обеспечивает хорошего визуального качества реконструкции. Во-вторых, проведено исследование влияния крестообразных маркеров на изображениях из обучающей выборки на появление артефактов. Обучение сети на наборе изображений без крестообразных маркеров обеспечило хорошее визуальное качество реконструкции и позволило значительно сократить количество артефактов. В-третьих, предложена модификация GAN сети для увеличения разрешения восстановленного изображения в два раза. Предложенная модификация позволила достичь 22 Дб на тестовой выборке по критерию PSNR.

7. Благодарности

Работа выполнена при поддержке грантов РФФИ (19-29-01235-мк, 16-29-11744-офи-м, № 16-29-09528-офи-м, № 17-29-03112-офи-м, № 18-07-01390-А, № 19-29-09054-мк), а также в рамках государственного задания ИСОИ РАН – филиала Федерального научно-исследовательского центра «Кристаллография и фотоника» РАН (соглашение № 007-ГЗ/Ч3363/26).

8. Литература

- [1] Genevet, P. Recent advances in planar optics: from plasmonic to dielectric metasurfaces / P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, R. Devlin // *Optica*. – 2017. – Vol. 4(1). – P. 139-152.
- [2] Nikonorov, A. Fresnel lens imaging with post-capture image processing / A. Nikonorov, R. Skidanov, V. Fursov, M. Petrov, S. Bibikov, Y. Yuzifovich // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. – 2015. – P. 33-41.
- [3] Nikonorov, A. Toward Ultralightweight Remote Sensing With Harmonic Lenses and Convolutional Neural Networks / A. Nikonorov, M. Petrov, S. Bibikov, P. Yakimov, V. Kutikova, Y. Yuzifovich, A. Morozov, R. Skidanov, N. Kazanskiy // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2018. – Vol. 11(9). – P. 3338-3348.
- [4] Nikonorov, A. Deep learning-based imaging using single-lens and multi-aperture diffractive optical systems / A. Nikonorov, V. Evdokimova, M. Petrov, P. Yakimov, S. Bibikov, Y. Yuzifovich, R. Skidanov, N. Kazanskiy // *Learning for computational imaging, ICCV*. – 2019.
- [5] Nikonorov, A. Deep learning-based image reconstruction for multi-aperture diffractive lens / A. Nikonorov, V. Evdokimova, M. Petrov, S. Bibikov, A. Alekseev, R. Skidanov, N. Kazanskiy // *Journal of Physics: Conference Series*. – 2019. – Vol. 1368.
- [6] Nikonorov, A. Towards multi-aperture imaging using diffractive lens / A. Nikonorov, R. Skidanov, V. Kutikova, M. Petrov, A. Alekseev, S. Bibikov // *Ultrafast Nonlinear Imaging and Spectroscopy VI - Proceeding of SPIE*. – 2019. – Vol. 11146.
- [7] Isola, P. Image-to-image translation with conditional adversarial networks / P. Isola, J.-Y. Zhu, T. Zhou, A. Efros // *Preprint arXiv* – 2018. – Vol. 1611.
- [8] Tai, Y. Image super-resolution via deep recursive residual network / Y. Tai, J. Yang, X. Liu // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2017)*. – 2017. – P. 3147-3155.

Study of GAN-based image reconstruction for diffractive optical systems

V.V. Evdokimova^{1,2}, M.V. Petrov^{1,2}, M.A. Klyueva², A.P. Alekseev², S.A. Bibikov^{1,2}, R.V. Skidanov^{1,2}, A.V. Nikonorov^{1,2}

¹Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

²Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. This paper studies the causes of artifacts that appear in images after color correction by a generative adversarial network. Input images are captured by diffractive optics. This paper shows overexposing of the original image and markers in training images affect the occurrence of the artifacts. The paper proposes an architecture of a generative adversarial network with a scaling layer. This architecture allows producing reconstructed images with scale factor 2.