ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВОССТАНОВЛЕНИЯ СПЕЦИАЛИЗИРОВАННЫХ ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ

Е.А. Степанова^{1,2}, Е.В. Родина¹, А.К. Шмаков¹

¹Омский научно-исследовательский институт приборостроения, Омск, Российская Федерация

²Институт радиофизики и функциональной электроники Омского научного центра СО РАН,
Омск, Российская Федерация

Выполнен обзор существующих современных методов восстановления искаженных изображений, представлены архитектуры искусственных нейронных сетей, эффективно применяемые для фильтрации, улучшения качества, восстановления утраченных фрагментов изображений. Рассмотрены особенности обработки и восстановления текстовых данных на графическом изображении.

Ключевые слова: искусственная нейронная сеть, интерполяция, глубокое обучение, искаженное изображение, сверточная нейронная сеть.

THE POSSIBILITY OF USING ARTIFICIAL NEURAL NETWORKS TO RESTORE SPECIALIZED DISTORTED GRAPHIC IMAGES

E.A. Stepanova^{1,2}, E.V. Rodina¹, A.K. Shmakov¹

¹Omsk Scientific-Research Institute of Instrument Engineering, Omsk, Russian Federation

²Institute of Radiophysics and Physical Electronics OSC SB RAS, Omsk, Russian Federation

A review of existing modern methods for recovering distorted images is performed, the architecture of artificial neural networks that are effectively used to filter, improve quality, and restore lost fragments of images is presented. The features of processing and restoration of text data on a graphic image are considered.

Keywords: artificial neural network, interpolation, deep learning, distorted image, convolutional neural network.

В настоящее время область применения технологий машинного обучения, в частности искусственных нейронных сетей, постоянно пополняется новыми направлениями исследований. Безусловно, это один из передовых и перспективных инструментов для решения нетривиальных, не поддающихся стандартным методам анализа задач.

Проблема восстановления искаженных, зашумленных графических изображений актуальна при обработке снимков низкого разрешения, медицинских диагностических данных, при работе со старыми архивными документами и фотографиями, в криминалистике, искусствоведении и многих других областях.

Основными, использующимися на сегодняшний день методами восстановления изображений являются:

- 1) масштабирование (upscaling) – процесс повышения разрешения и качества цифрового изображения;
- 2) метод слепой деконволюции – метод восстановления изображения без априорной информации о функции размытия точки графической системы, которая определяет «направления смазывания» [1];
- 3) применение рекуррентных нейронных сетей, таких как нейронная сеть Элмана, сеть Джордана, сеть Хопфилда, цепь Маркова и т.п.;
- 4) применение технологии глубокого обучения (deep learning).

- Помимо данных методов, существует широкий набор сопутствующих инструментов:
- фильтрация шума на графическом изображении (сглаживание кадра);
 - увеличение/ уменьшение изображения (масштабирование);
 - коррекция геометрических искажений;
 - увеличение числа пикселей на координату;
 - разбиение на фрагменты и т. д.

В настоящее время существует множество онлайн-сервисов, использующих представленные методы и инструменты. Сравнительная характеристика некоторых из них представлена в таблице.

Сравнение онлайн-сервисов по обработке графической информации

<i>Сервис</i>	<i>Метод</i>	<i>База</i>	<i>Область использования</i>
Adobe Photoshop [2]	Бикубическая интерполяция	Математический аппарат	Любительские снимки
Let's Enhance [3]	Сверточная нейронная сеть (CNN)	Обучающая выборка – фото природы, животных, зданий	Фотографии животных, природы, зданий
Improve Photo [4]	Фильтры	Математический аппарат	Пейзажи, мультфильмы
Enhance Photo [5]	Фильтры	Математический аппарат	Пейзажи, мультфильмы
PinkMirror [6]	Фильтры	Математический аппарат	Портреты людей
Movavi Photo DeNoise [7]	Фильтры, CNN	Обучающая выборка – фото природы, животных, зданий	Фотографии животных, природы, зданий
IMGonline.com.ua [8]	CNN	Обучающая выборка – фото природы, животных, зданий	Фотографии животных, природы, зданий
Webinpaint [9]	Нейронная сеть, фильтры	Совмещение рядом стоящих пикселей, мат. аппарат	Фотографии с водяными знаками
waifu2x.udp.jp [10]	CNN	Обучающая выборка – рисованные персонажи аниме	Фотографии жанра аниме
9may.mail.ru [11]	CNN	Обучающая выборка – военные фотографии	Компьютерное зрение

Некоторые продукты, такие как Adobe Photoshop [12] для работы с фото и видео, способны снизить зашумленность и улучшить качество кадра с использованием бикубической интерполяции. Результаты интерполяции в случае с работой над фото с текстом зачастую не подходят, так как «сглаживание» четких границ может только снизить читаемость информации с исходного графического изображения.

Как видно из таблицы, большинство подобных сервисов используют простые методы, такие как фильтрация или интерполяция, однако для более сложных задач, таких как восстановление старых поврежденных снимков, применяются искусственные нейронные сети (ИНС).

Нейронные сети активно используются для постобработки изображений благодаря своей гибкости. ИНС находят применение в таких задачах как удаление шума и размытия с изображений, восстановление утраченных фрагментов, добавление цвета, увеличение разрешения изображений без потери качества.

Наиболее частые искажения графической информации встречаются при наложении шума, размытии снимка, затирании фрагментов, уменьшении качества изображения (сжатие). Рассмотрим эти ситуации более подробно.

Восстановление изображения с наложенным на него шумом достигается путем анализа соседних пикселей [13]. Для восстановления использовалась программа [14] на основе сверточной нейронной сети. Сверточная нейронная сеть (CNN – convolutional neural network) – это специальный вид нейронной сети для обработки данных с сеточной топологией. Свертка – особый вид линейной операции. Сверточные сети – нейронные сети, в которых

вместо общей операции умножения на матрицу, по крайней мере в одном слое, используется свертка. Для иллюстрации работы CNN было выбрано изображение размером 1080x540 пикселей формата png. На рис. 1–3 представлены результаты работы программы [14] при наложении различных видов шумов.

Рис. 1. Исходное изображение (а); изображение с наложенным шумом Гаусса (б); восстановленное изображение (в)

Рис. 2. Исходное изображение (а); изображение с наложенным шумом Гаусса с нулевым средним и переменной дисперсией (б); восстановленное изображение (в)

Рис. 3. Исходное изображение (а); изображение с наложенным шумом «соль и перец» (б); восстановленное изображение (в)

Из рис. 1 и 2 видно, что при небольшом искажении искусственная нейронная сеть восстанавливает изображение практически без потери качества, однако если искажение достаточно сильное (как на рис. 3), после восстановления на изображении остаются артефакты. Нейронные сети хорошо справляются с задачей восстановления при наложении шума.

DCNN (deep convolutional neural networks – глубокие сверточные нейронные сети) кардинально отличаются от других сетей. Они используются в основном для обработки изображений, иногда для аудио и других видов входных данных. Типичным способом применения DCNN является классификация изображений: если на вход подается изображение кошки, сеть выдаст «кошка», если картинка собаки – «собака» [15]. Такие сети обычно используют «сканер», не обрабатывающий все данные за один раз. Например, если имеется изображение 200x200 пикселей и входной слой нейронной сети состоит 40000 нейронов, то из левого верхнего угла будет считываться квадрат размером 20x20 пикселей, а затем выполнится сдвиг на 1 пиксель вправо и данные снова считаются и т.д. Потом входные данные проходят через свёрточные слои, где каждый узел соединен только со своими ближайшими соседями. Эти слои имеют свойство сжиматься, причём обычно они уменьшаются на какой-нибудь из делителей количества входных данных (например, 20 узлов в следующем слое превратятся в 10, в следующем – в 5), часто используются степени двойки. Кроме сверточных слоев, есть также так называемые слои объединения (pooling layers). Объединение – это способ уменьшить размерность получаемых данных, например, из квадрата 2x2 выбирается и передается наиболее красный пиксель. На практике к концу CNN прикрепляют сеть прямого распространения для дальнейшей обработки данных. Такие сети называются глубокими (DCNN). Архитектура сети DCNN представлена на рис. 4 а).

DCIGN (deep convolutional inverse graphics networks – глубокие сверточные обратные глубинные сети) – это вариационные автоэнкодеры со сверточными и развертывающими сетями в качестве кодирующей и декодирующей частей соответственно (рис. 4б) [16]. Такие сети представляют черты изображения в виде вероятностей и могут научиться строить изображение кошки и собаки вместе, взглянув лишь на картинки только с кошками и только с собаками. Разработчики также продемонстрировали, что DCIGN может моделировать различные сложные преобразования изображений, например, изменение источника света или поворот 3D объектов. Такие сети обычно обучают методом обратного распространения ошибки.

Рис. 4. Архитектура DCNN (а); архитектура DCIGN (б)

Нейронные сети также применяют для восстановления блоков утраченной информации [17]. На рис. 5 и 6 представлены изображения, части которых были восстановлены с помощью программы [18] на основе генеративной состязательной нейронной сети (GAN).

Рис. 5. Исходное изображение (а); восстановленное изображение (б)

Генеративные нейросетевые модели создают некие синтетические данные на основе входной информации. Такие сети состоят из двух ИНС (чаще всего сети прямого распространения – FFNN и CNN), где одна часть выступает генератором и создает новые образы, а другая – дискриминатором, отличающим подлинные образы от сгенерированных. Точность определения дискриминатором источника данных служит условием для оценки ошибок генератора. То есть происходит своего рода состязание двух сетей: дискриминатор учится лучше отличать реальные данные от сгенерированных, а генератор стремится стать менее предсказуемым для дискриминатора.

Рис. 6. Исходное изображение (а); восстановленное изображение (б)

Задача обучения GAN нетривиальна: необходимо не только обучить две сети, но и найти нужный баланс между ними. Если одна из частей (генератор или дискриминатор) станет более обученной, то GAN никогда не сойдется [19].

Алгоритмы, основанные на искусственных нейронных сетях, меняют изображение целиком, поэтому хорошо справляются с дефектами большого размера. Схема нейронной сети с архитектурой GAN представлена на рис 7а.

GAN не восстанавливает утраченную информацию, а «достраивает» недостающие элементы исходя из информации, присутствующей на изображении. Стоит отметить, что нейронная сеть не сможет восстановить утраченную информацию, а сделает только предположение о том, что может находиться на утраченных блоках. В зависимости от обучающей выборки, нейронная сеть может «дорисовать» картинку настолько правдоподобно, что это не будет заметно человеческому глазу. Они используются для получения фотореалистичных изображений, восстановления утраченных фрагментов, улучшения качества изображений и т. д.

DBN (deep belief networks – глубокие сети доверия) – сети, представляющие собой композицию нескольких ограниченных машин Больцмана (RBM) или вариационных авто-энкодеров (VAE) (рис. 7б).

Такие сети показали себя эффективно обучаемыми одна за другой, когда каждая должна научиться кодировать предыдущую. Этот метод также называют «жадным обучением». Метод жадного послойного обучения состоит в принятии оптимального на данный момент решения, чтобы получить подходящий, но, возможно, неоптимальный результат [20]. DBN могут обучаться методами contrastive divergence или обратным распространением ошибки и учатся представлять данные в виде вероятностной модели, в точности как RBM или VAE. Однажды обученную и приведенную к стационарному состоянию модель можно использовать для генерации новых данных.

Рис. 7. Архитектура GAN (а); архитектура DBN (б)

Еще одним применением искусственных нейронных сетей является устранение размытия с изображений [21]. На рис. 8 и 9 представлены примеры, демонстрирующие работу программы [22] на основе масштабируемой модели рекуррентной нейронной сети (RNN) с направленной последовательностью между элементами и возможностью сохранения в памяти предыдущего состояния.

Рис. 8. Исходное изображение (а); восстановленное изображение (б)

Рис. 9. Исходное изображение (а); восстановленное изображение (б)

Преимущество данного подхода состоит в том, что с его помощью можно восстановить утраченные детали. Так, на рис. 9 видно, что удалось восстановить автомобильный номер.

Важно отметить, что различные архитектуры нейронных сетей будут давать разные результаты в одинаковых задачах по восстановлению исходного изображения.

Из всего вышесказанного можно сделать вывод о том, что для задач восстановления специализированного искаженного графического изображения наилучшим образом подходят следующие архитектуры ИНС:

- структуры GAN (generative adversarial network – генеративные состязательные сети);
- разновидности CN (convolutional network);
- разновидности DN (deep network).

Для получения лучшего результата перед подачей данных в нейронную сеть необходимо предварительно отфильтровать шумы с использованием различных интерполяционных алгоритмов, таких как бикубическая интерполяция, фильтр Ланцоша и др.

Следует отметить, что при выборе архитектуры ИНС, помимо анализа артефактов, следует понимать, какую информацию необходимо обработать: пейзаж, портрет, схемотехническую, текст, медицинские диагностические данные, цветовую гамму у изображения и т. д.

Текстовую информацию на искаженном графическом изображении зачастую довольно затруднительно прочесть (рис. 10а).

а)

б)

Рис. 10. Искаженное комбинированное изображение (а);
восстановленное изображение с помощью ИНС (б)

Методика обработки и распознавания текста включает в себя следующие этапы:

- 1) восстановление искаженного графического изображения (предварительная фильтрация и обработка);
- 2) идентификация текста на изображении (локализация текстовой области);
- 3) восстановление текста (работа с областью локализации);
- 4) распознавание текстовой информации.

На рис. 10б представлено комбинированное изображение после обработки сверточной генеративно-состязательной сетью (convolutional generative adversarial network, CGAN), обученной на случайных выборках [23].

На рисунке видно, что с помощью нейронной сети удалось восстановить исходное изображение. В некоторых местах размытие сохранилось, однако области с текстом четкие и читаемые. Результат работы нейронной сети позволяет производить дальнейшую обработку с целью локализации текстовых областей и дальнейшего распознавания.

Учитывая все вышесказанное, можно сделать вывод о том, что при наличии разнородной информации в обрабатываемом графическом изображении, например текста и картинки, целесообразно применять разные подходы, такие как совмещение нескольких нейронных сетей: отдельно для обработки текста, отдельно – для картинок.

Работа выполнена по государственному заданию Омского научного центра СО РАН в соответствии с Программой ФНИ ГАН на 2013-2020 годы (номер госрегистрации проекта в системе ЕГИСУ НИОКТР АААА-А19-119052890058-2).

СПИСОК ЛИТЕРАТУРЫ

1. Новокишанов О.Н. Восстановление изображений, искажённых свёрткой с неизвестной функцией: автореф. дис. ... канд. тех. наук. М., 2001. 18 с.
2. Adobe, приложения Adobe Creative Cloud для настольных ПК. URL: <https://www.adobe.com/ru/creativecloud/catalog/desktop.html?promoid=PTYTQ77P&mv=other> (дата обращения: 08.08.19).
3. Let's Enhance. URL: <https://letsenhance.io/boost> (дата обращения: 07.08.19).
4. Improve Photo. URL: <http://improvephoto.net/> (дата обращения: 08.08.19).
5. Enhance.Photo.to. URL: <https://enhance.pho.to/ru/> (дата обращения: 10.08.19).
6. PinkMirror. URL: <https://pinkmirror.com/> (дата обращения: 10.08.19).
7. Movavi Photo DeNoise. URL: <https://www.movavi.ru/noise-reduction-software/> (дата обращения: 10.08.19).
8. IMGonline.com.ua. URL: <https://www.imgonline.com.ua/noise-reduction.php> (дата обращения: 10.08.19).
9. Webinpaint. URL: <https://online.theinpaint.com/> (дата обращения: 08.08.19).
10. Waifu2x image resizer. URL: <https://waifu2x.booru.pics/> (дата обращения: 08.08.19).
11. Waifu2x-multi. URL: <https://waifu2x.me/index.ru.html> (дата обращения: 10.08.19).
12. Компьютерное зрение от почты Mail.ru. URL: <https://9may.mail.ru/> (дата обращения: 09.08.19).
13. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising / K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang // Cornell University. 13 Avg. 2016. 13 p.
14. Dn-CNN-tensorflow. URL: <https://github.com/wbhu/DnCNN-tensorflow> (дата обращения: 08.08.19).
15. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // Proc. of the IEEE. November 1998. 46 p.
16. Deep convolutional inverse graphics network / Tejas D. Kulkarni, W. Whitney, P. Kohli, J.B. Tenenbaum // NeurIPS. 22 Jun 2015. 10 p.
17. Image inpainting via generative multi-column convolutional neural networks / Y. Wang, X. Tao, X. Qi, X. Shen, J. Jia // NeurIPS. 2018. 10 p.
18. Inpainting_gmcnn/ URL: https://github.com/shepnerd/inpainting_gmcnn (дата обращения: 08.08.19).
19. Generative adversarial nets / Ian J. Goodfellow, J. Pouget-Abadie, M. Mirza, Bing Xu, D. Warde-Farley, O. Aaron Courville, Y. Bengio. 10 Jun 2014. 9 p.
20. Greedy layer-wise training of deep networks / Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle // Advances in neural inf. processing systems. Jan. 2007.
21. Scale-recurrent network for deep image deblurring / X. Tao, H. Gao, X. Shen, J. Wang, J. Jia // 16 Feb. 2018. 9 p.
22. SRN-Deblur. URL: <https://github.com/jiangsutx/SRN-Deblur> (дата обращения: 11.08.19).
23. DnCNN-tensorflow. URL: <https://github.com/wbhu/DnCNN-tensorflow> (дата обращения: 12.08.19).