

Ревью

Акимова Екатерина

November 2021

В настоящий момент тема обработки изображений крайне актуальна, потому что наступила информационная эра. Огромные объемы информации хранятся в виде изображений, однако не все объекты, присутствующие на изображениях, являются нужными, поэтому возникает задача по их устранению таким образом, чтобы изображение не оказалось безнадежно испорченным. Именно по темам, связанным с этой проблемой, будут рассмотрены статьи.

Научная статья О.А. Строя и В.В. Буряченко посвящена удалению нежелательных объектов с аэроснимкой. В ней они рассматривают различные способы обработки изображения, одним из которых является нейронная сеть ResNet, которую легче оптимизировать и точность классификации которой повышается за счёт значительного увеличения глубины, при этом обучать её легче. Схема архитектуры данной сети представлена на рисунке 1. [1]

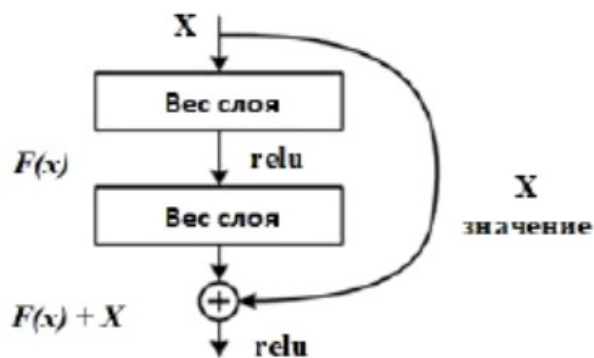


Figure 1: Архитектура ResNet

В следующей статье рассматривается метод реконструкции изображений на основе поиска подобных областей состоит из нескольких этапов. Данный метод состоит из двух шагов. На первом шаге автоматизировано определяются поврежденные участки изображения или пользователем формируется маска дефектных областей. На втором шаге изображение с потерянными областями

восстанавливается с помощью метода ЕВМ (exemplarbasedmethod). Данный метод позволяет корректно восстанавливать потерянные области с помощью созданной маски, правильно реконструируя значения пикселей. Метод основывается на вычислении приоритета для каждого пикселя границы с последующим поиском похожего квадратного блока в области доступных пикселей и его копировании в область отсутствующих пикселей. [2]

Ю.В. Алексеенко в своей статье говорит о решении задачи по идентификации изображений при помощи свёрточной нейронной сети с использованием различных видов обучения: полное, с использованием bottleneck признаков уже обученной сети, с поднастройкой последних слоёв обученной сети. Более подробно о каждом методе приведено в таблице 1. В результате проведенных экспериментов лучшим был выделен третий метод. [3]

Полное обучение	В данном случае необходимо помнить, что обучение на малом объеме данных может привести к проблеме переобучения. Один из вариантов решения проблемы переобучения - это увеличение количества данных для обучения. Но во многих случаях этого недостаточно, так как данные могут быть сильно коррелированы. Дропаут (Dropout) также помогает бороться с переобучением, не позволяя слою увидеть дважды один и тот же образ, таким образом, действуя аналогично аугментации (ugmentation). Таким образом, можно сделать вывод, что аугментация и дропаут помогают нарушить случайные корреляции, возникающие в данных. На практике метод полного обучения сети реализуется крайне редко, так как проблематично найти набор данных достаточного размера для обучения сети. Вместо этого обычно принято брать уже обученную сверточную нейронную сеть и использовать ее для решения поставленной задачи.
Обучение с использованием bottleneck признаков уже обученной сети	Использование уже обученной нейронной сети на наборе данных, предварительно обученную на наборе данных, затем убрать последний полносвязный слой. После чего запустить один раз полученную модель на имеющемся обучающем наборе данных для тестирования, записывая при этом признаки в массив. признаками называются те, которые извлекаются из малоразмерного скрытого слоя, расположенного уже ближе к последним скрытым слоям нейронной сети. В завершении, и, необходимо обучить полносвязный слой с учетом этих сохраненных признаков.
Обучение с поднастройкой последних слоёв обученной сети	Работая уже с обученной нейронной сетью, необходимо лишь переучить ее на новом наборе данных, используя при этом очень небольшие обновления весов. Другими вариантами в рамках этого подхода являются также: подстройка всех слоев обученной сети или использование предыдущих слоев и подстройка слоев, находящихся ближе к выходу сети.

В статье "Разработка алгоритма сегментации рукописного текста с использованием

маски" освещается проблема выделения рукописного текста на изображении для последующего распознавания. Приводится описание широко распространенных методов сегментации, позволяющих решить данную проблему: метод цветового анализа изображений, бинарная сегментация изображений методом фиксации уровня, а также метод выделения и анализа контуров. Предлагается алгоритм сегментации текста, основанный на использовании маски исходного изображения, позволяющий значительно упростить сегментацию текста на изображениях с немонокотным фоном. Также производится сравнительный анализ методов. Поиск различий маски и изображения проводился по формуле: [4]

$$Diff(i, j) = Image(i, j) - Mask(i, j)$$

В статье "Использование технологии сверточных нейронных сетей в сегментации объектов изображения" речь идёт о компьютерном зрении, рассматривается применение нейронных сетей для выделения конкретных объектов на изображении, дана характеристика задачи сегментации и основных принципов компьютерного зрения. [5]

В статье "An Analysis Of Convolutional Neural Networks For Image Classification" представлен эмпирический анализ эффективности популярных сверточных нейронных сетей (CNN) для идентификации объектов в видеопотоках в реальном времени. Наиболее популярными сверточными нейронными сетями для обнаружения объектов и классификации категорий объектов на изображениях являются Alex Nets, GoogLeNet и ResNet50. Для тестирования производительности различных типов CNN существует множество наборов данных изображений. Общепринятыми эталонными наборами данных для оценки производительности сверточной нейронной сети являются набор данных ImageNet, а также наборы данных изображений CIFAR10, CIFAR100 и MNIST. Данное исследование сосредоточено на анализе производительности трех популярных сетей: Alex Net, GoogLeNet и ResNet50. Для исследования были взяты три наиболее популярных набора данных: ImageNet, CIFAR10 и CIFAR100, поскольку тестирование работы сети на одном наборе данных не позволяет выявить ее истинные возможности и ограничения. Следует отметить, что видео не используются в качестве обучающего набора данных, они используются в качестве тестового набора данных. Анализ показал, что GoogLeNet и ResNet50 способны распознавать объекты с большей точностью по сравнению с Alex Net. [6]

В статье "DOG: A new background removal for object recognition from images" говорят о том, что иногда фон изображений влияет на распознавание нейронных сетей, и его удаление может улучшить их производительность распознавания объектов. Поэтому авторы статьи разработали модель под названием DOG на основе свёрточной нейронной сети для удаления фона изображений с целью повышения эффективности распознавания объектов. Из-за нехватки образцов DOG интегрировали с DCGAN для дальнейшего повышения эффективности распознавания объектов. Экспериментальные результаты показали эффективность данного решения. [7]

В данной работе разработана система обработки изображений, объединяющая выделение и сопоставление признаков с помощью сверточной нейронной

сети (CNN), вместо того, чтобы полагаться на простой метод выделения и сопоставления признаков по отдельности при обработке изображений в обычной системе распознавания образов. Для ее реализации предлагаемая система позволяет использовать CNN для работы и анализа производительности обычной системы обработки изображений. Эта система извлекает особенности изображения с помощью CNN, а затем обучает их с помощью нейронной сети. Предложенная система показала 84% точность распознавания. Предложенная система представляет собой модель распознавания изученных изображений с помощью глубокого обучения. Поэтому она может запускаться в пакетном режиме и легко работать под любой платформой (в том числе встроенной), которая может читать все виды файлов в любое время. Кроме того, она не требует реализации алгоритма извлечения признаков и алгоритма сопоставления, что позволяет экономить время и является эффективной. В результате она может широко использоваться в качестве программы распознавания изображений. [8]

Авторы статьи "Deep features based convolutional neural network model for text and non-text region segmentation from document images" представляют модель глубокой сверточной нейронной сети, которая использует функции глубокого обучения для сегментации текстовых и нетекстовых областей на изображениях документов. Основная задача заключается в извлечении текстовых областей из изображений документов сложной компоновки без каких-либо предварительных знаний о сегментации. В реальном мире изображения документов или журналов содержат различную текстовую информацию наряду с нетекстовыми областями, такими как символы, логотипы, изображения и графика. Извлечение текстовых областей из нетекстовых является сложной задачей. Для решения этих проблем в данной работе предложена эффективная и надежная техника сегментации. Реализация предложенной модели разделена на три этапа: метод предварительной обработки изображений документов с использованием различных размеров патчей используется для обработки ситуаций, связанных с вариантами шрифтов и размеров текста на изображении; модель глубокой сверточной нейронной сети предлагается для предсказания текста, нетекста или неоднозначной области на изображении; метод постобработки изображения документа предлагается для обработки ситуации, когда изображение имеет сложные неоднозначные области, используя рекурсивное разбиение этих областей на соответствующие классы (т.е. текст или не текст), затем система накапливает ответы этих предиктивных патчей с различным разрешением для обработки ситуации с вариациями текстовых шрифтов в изображении. Было проведено обширное компьютерное моделирование с использованием коллекции изображений журналов со сложной версткой с сайтов Google и базы данных ICDAR 2015. Получены результаты и проведено сравнение с современными методами. Выяснилось, что предложенная модель является надежной и более эффективной по сравнению с современными методами. [9]

В этой статье мы представляем новый метод извлечения текстовых строк для неограниченных рукописных документов, не зависящий от языка, который справляется с такими сложностями, как соприкасающиеся и разнонаправленные

текстовые строки, перекрывающиеся символы и неравномерные межстрочные интервалы. Наш метод осуществляет предварительную обработку страниц документа с помощью метода преобразования расстояний и использует новый алгоритм обнаружения контура для выделения отдельных текстовых строк. Предложенный метод был протестирован на шести стандартных наборах данных, которые находятся в открытом доступе, и результаты экспериментов показывают, что наш метод достигает многообещающей точности по сравнению с современными методами TLE. [10]

Список используемой литературы

1. Строй О. А., Буряченко В. В. МЕТОДЫ УДАЛЕНИЯ НЕЖЕЛАТЕЛЬНЫХ ОБЪЕКТОВ С ИЗОБРАЖЕНИЙ АЭРОФОТОСЪЕМКИ С ИСПОЛЬЗОВАНИЕМ ИТЕРАЦИОННОГО ПОДХОДА // Сибирский аэрокосмический журнал. 2021. №3. URL: <https://cyberleninka.ru/article/n/metody-udaleniya-nezhelatelnih-obektov-s-izobrazheniy-aerofotosemki-s-ispolzovaniem-iteratsionnogo-podhoda> (дата обращения: 24.11.2021).
2. Ибадов Рагим Рауфевич, Ибадов Самир Рауфевич, Воронин Вячеслав Владимирович, Федосов Валентин Петрович Модифицированный метод реконструкции изображений на основе поиска подобных областей // Известия ЮФУ. Технические науки. 2017. №6 (191). URL: <https://cyberleninka.ru/article/n/modifitsirovanny-metod-rekonstruktsii-izobrazheniy-na-osnove-poiska-podobnyh-oblastey> (дата обращения: 24.11.2021).
3. Алексеенко Юлия Вячеславовна Разработка системы распознавания изображений на основе сверточных нейронных сетей // Евразийский Союз Ученых. 2017. №7-1 (40). URL: <https://cyberleninka.ru/article/n/razrabotka-sistemy-raspoznavaniya-izobrazheniy-na-osnove-svertochnyh-neyronnyh-setey> (дата обращения: 24.11.2021).
4. Кобенко Вадим Юрьевич, Фролов Святослав Олегович, Талалаев Владимир Ильич РАЗРАБОТКА АЛГОРИТМА СЕГМЕНТАЦИИ РУКОПИСНОГО ТЕКСТА С ИСПОЛЬЗОВАНИЕМ МАСКИ // ОНВ. 2020. №4 (172). URL: [https://www.omgtu.ru/general_information/media_omgtu/journal_of_oms_k_research_journal/files/arhiv/2020/4%20\(172\)/47-52%20Кобенко%20В.%20Ю.,%20Фролов%20С.%20О.,%20Талалаев%20В.%20И..pdf](https://www.omgtu.ru/general_information/media_omgtu/journal_of_oms_k_research_journal/files/arhiv/2020/4%20(172)/47-52%20Кобенко%20В.%20Ю.,%20Фролов%20С.%20О.,%20Талалаев%20В.%20И..pdf) (дата обращения: 24.11.2021).
5. Клехо Дмитрий Юрьевич, Карелина Екатерина Борисовна, Батырев Юрий Павлович ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ В СЕГМЕНТАЦИИ ОБЪЕКТОВ ИЗОБРАЖЕНИЯ // Вестник МГУЛ – Лесной вестник. 2021. №1. URL: <https://cyberleninka.ru/article/n/ispolzovanie-tehnologii-svertochnyh-neyronnyh-setey-v-segmentatsii-obektov-izobrazheniya> (дата обращения: 24.11.2021).
6. Neha Sharma, Vibhor Jain, Anju Mishra, An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science,

Volume 132, 2018, Pages 377-384, ISSN 1877-0509, URL: <https://doi.org/10.1016/j.procs.2018.05.198>.
(дата обращения: 24.11.2021).

7. Wei Fang, Yewen Ding, Feihong Zhang, Victor S. Sheng, DOG: A new background removal for object recognition from images, Neurocomputing, Volume 361, 2019, Pages 85-91, ISSN 0925-2312, URL: <https://doi.org/10.1016/j.neucom.2019.05.095> (дата обращения: 24.11.2021).
8. Hankil Kim, Jinyoung Kim, Hoekyung Jung, Convolutional Neural Network Based Image Processing System, J. Inf. Commun. Converg. Eng. 16(3): 160-165, Sep. 2018/ URL: <https://www.koreascience.or.kr/article/JAKO201831854885756.pdf> (дата обращения: 24.11.2021).
9. aiyed Umer, Ranjan Mondal, Hari Mohan Pandey, Ranjeet Kumar Rout, Deep features based convolutional neural network model for text and non-text region segmentation from document images, Applied Soft Computing, Volume 113, Part A, 2021, 107917, URL: <https://doi.org/10.1016/j.asoc.2021.107917> (дата обращения: 24.11.2021).
10. Suman Kumar Bera, Soumyadeep Kundu, Neeraj Kumar, Ram Sarkar, Distance transform based text-line extraction from unconstrained handwritten document images, Expert Systems with Applications, Volume 186, 2021, 115666, URL: <https://doi.org/10.1016/j.eswa.2021.115666> (дата обращения: 24.11.2021).