



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ (ИУ6)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.03.01 Информатика и вычислительная техника

БАКАЛАВРСКАЯ ПРОГРАММА 09.03.01/03 Вычислительные машины, комплексы,
системы и сети

ОТЧЕТ ПО ПРОИЗВОДСТВЕННОЙ ПРАКТИКЕ

Тип практики Преддипломная практика


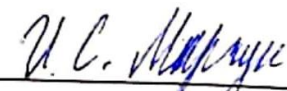
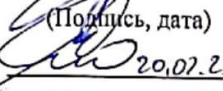
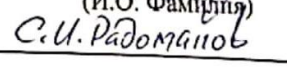
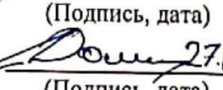
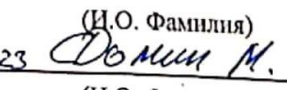
Название

предприятия ФГБУ «27 ЦНИИ» Минобороны России

Студент ИУ6-82Б

Руководитель практики
от предприятия

Руководитель практики
от МГТУ им. Н.Э.Баумана

| | |
|--|---|
|  (Подпись, дата) |  (И.О. Фамилия) |
|  (Подпись, дата) |  (И.О. Фамилия) |
|  (Подпись, дата) |  (И.О. Фамилия) |

Оценка


27.02.2023

2023 г.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ЗАДАНИЕ на производственную практику

по теме предприятия практики

Студент группы ИУ6-82Б

Марчук Иван Сергеевич

(Фамилия, имя, отчество)

Направление подготовки 09.03.01 Информатика и вычислительная техника

Бакалаврская программа 09.03.01/03 Вычислительные машины, комплексы, системы и сети

Тип практики Преддипломная практика

Название предприятия «ФГБУ «27 ЦНИИ» Минобороны России»

Техническое задание Изучение базовых производственных процессов разработки программно-аппаратных средств вычислительной техники на предприятии практики. Сбор материалов для подготовки ВКР.

Оформление отчета по практике:

Отчет на 15-25 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Нет

Дата выдачи задания « 07 » февраля 2023 г.

Руководитель практики
от МГТУ им. Н.Э.Баумана

Студент


(Подпись, дата) 07.02.2023

(И.О. Фамилия)

(Подпись, дата) 20.02.2023
(И.О. Фамилия) Марчук ИС

Примечание: Задание оформляется в двух экземплярах.

З А Д А Н И Е

на производственную практику от предприятия

по теме _____

Студент группы ИУ6-82Б

Марчук Иван Сергеевич

(Фамилия, имя, отчество)

Направление подготовки 09.03.01 Информатика и вычислительная техника

Бакалаврская программа 09.03.01/03 Вычислительные машины, комплексы, системы и сети

Тип практики Преддипломная практика

Название предприятия «ФГБУ «27 ЦНИИ» Минобороны России»

Техническое задание Анализ существующих методов обработки рукописных текстов с применением технологии искусственного интеллекта

Оформление отчета по практике:

Отчет на 15-25 листах формата А4.

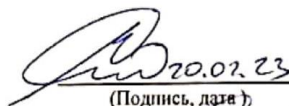
Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Нет

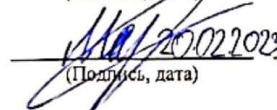
Дата выдачи задания « 07 » февраля 2023 г.

**Руководитель практики
от предприятия**

Студент


(Подпись, дата)

С.И. Радоманов
(И.О. Фамилия)


(Подпись, дата)

И.С. Марчук
(И.О. Фамилия)

Примечание: Задание оформляется в двух экземплярах.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

СПИСОК-ВЕДОМОСТЬ ПРОИЗВОДСТВЕННОЙ ПРАКТИКИ

Факультет: ИУ, группа: ИУ6-82Б.

Название практики: Преддипломная практика.

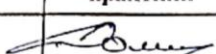
Сроки практики: с 07.02.2023 по 20.02.2023.

Место проведения практики: г. Москва.

Название предприятия: ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
"27 ЦЕНТРАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ" МИНИСТЕРСТВА
ОБОРОНЫ РОССИЙСКОЙ ФЕДЕРАЦИИ.

Руководитель практики от МГТУ имени Н.Э. Баумана:

Старший преподаватель, Фомин Михаил Михайлович.

| № | Фамилия, имя, отчество | № зачётки | Группа | Дифференцированная оценка | Подпись руководителя практики |
|---|------------------------|-----------|---------|---------------------------|---|
| 1 | Марчук Иван Сергеевич | 18У373 | ИУ6-82Б | от 1 |  |

Зам. декана факультета

Управление образовательных технологий

СПРАВКА

Студенты в количестве 1, согласно списку-ведомости производственной практики

Выбыли из вуза

_____ 20__ г.

Зам. декана факультета

подпись(с расшифровкой), печать

Выбыли с базы П.П.

20 фев. 2023 г.

С. Радомянов

должность, подпись(с расшифровкой), печать

Прибыли на базу П.П.

07. фев 2023 г.

С. Радомянов

должность, подпись(с расшифровкой), печать

Прибыли в вуз

_____ 20__ г.

Зам. декана факультета

подпись(с расшифровкой), печать

Приложение 2 / _____
к договору № 05.09-00-14/23-40
от « 31 » января 2023 г.
ФГБУ «27 ЦНИИ» Минобороны России
название профильной организации

Совместный рабочий график (план) проведения практики



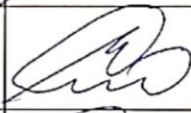
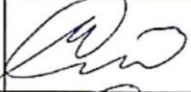
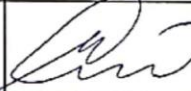

Сроки проведения практики: 07.02.2023 – 20.02.2023

Вид практики: производственная

Тип практики: преддипломная

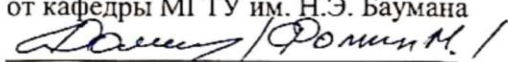
Студент Марчук Иван Сергеевич группа ИУ6-82Б

Направление 09.03.01

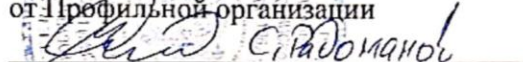
| Мероприятия * | Дата | Место проведения | Ответственное лицо |
|------------------------------------|------------------------|----------------------------------|---|
| Организационное собрание | 07.02.2023 | кафедра |  |
| Экскурсия обзорная | 07.02.2023 | ФГБУ «27 ЦНИИ» Минобороны России |  |
| Выполнение индивидуального задания | 07.02.2023-20.02.2023 | ФГБУ «27 ЦНИИ» Минобороны России |  |
| Лекции (по необходимости) * | 09.02.2023, 17.02.2023 | ФГБУ «27 ЦНИИ» Минобороны России |  |
| Консультации | 09.02.2023, 17.02.2023 | ФГБУ «27 ЦНИИ» Минобороны России |  |
| Итоговое собрание | 20.02.2023 | кафедра |  |

*-тип мероприятия устанавливаются на усмотрение руководителей

Подписи сторон:

Руководитель практики
от кафедры МГТУ им. Н.Э. Баумана


27.02.2023

Руководитель практики
от Профильной организации


Оглавление

| | |
|---|----|
| ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ | 7 |
| Введение | 8 |
| Принципы обучения нейронных сетей | 9 |
| Генеративно-сопоставительная модель | 11 |
| Обзор существующих решений GAN с открытым исходным кодом | 12 |
| Нейросеть ScrabbleGAN | 12 |
| Нейросеть GanWriting | 13 |
| Свёрточные нейронные сети | 15 |
| Дообученная нейросеть microsoft/trocr-small-handwritten..... | 16 |
| Нейросеть HTR | 17 |
| Заключение | 18 |
| Список используемых источников | 19 |

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

GAN – (Generative adversarial network) Генеративно-состязательная модель. Алгоритм машинного обучения без учителя, построенный на комбинации из двух нейронных сетей, одна из которых генерирует образцы , а другая старается отличить правильные образцы от неправильных.

OCR – (Optical character recognition) Технология оптического символьного распознавания текста

Датасет – (анг. Data Set) Это механизм хранения информации, который предоставляет быстрый доступ к большим объемам данных.

One-hot – Тип кодирования категориального признака, основывается на создании бинарных признаков, которые показывают принадлежность к уникальному значению. Под категориальными данными понимаются данные, которые не имеют численного представления, они могут иметь, как и два уникальных значения (бинарные признаки), так и более.

Введение

Практика проводилась на базе федерального государственного бюджетного предприятия «27 Центральный научно-исследовательский институт» Министерства обороны Российской Федерации, Управлении № 5.

Цели изучения дисциплины:

- Закрепить и расширить теоретические знания и практические навыки, полученные во время обучения.
- Ознакомиться с содержанием и видами основных работ и исследований, выполняемых на предприятии по месту прохождения практики.
- Освоение инструментальных средств и технологических цепочек, применяемых при промышленной разработке, отладке и тестировании программных систем на примере предприятия (27 ЦНИИ МО РФ), приобретение навыков работы в коллективе.

В процессе прохождения практики в 27-м институте мне была поручена задача проанализировать технологию нейросетевого обучения, для распознавания рукописных текстов.

Поскольку это преддипломная практика, необходимо было подобрать задачу, связанную с моей дипломной работой «Андроид приложение Помощник учителя». Упор в тематику распознавания рукописных текстов делался по тому, что распознавание текста могло бы помочь учителям в проверке письменных работ. А институту, в котором я проходил практику, распознавание рукописных текстов пригодилось бы, в том числе, для оцифровки или сортировки заполненных вручную документов.

Принципы обучения нейронных сетей

Для обучения нейросети могут быть использованы различные подходы. И исходя из имеющихся ресурсов необходимо выбрать принцип, по которому будет производиться обучение.

Обучение с учителем — нейронная сеть обучается на размеченном наборе данных и предсказывает ответы, которые используются для оценки точности алгоритма на обучающих данных.

Размеченный набор – это набор заранее подготовленных данных (картинок, аудиофайлов, и т.д.), для каждого из которых человеком был заранее указан правильный ответ, который должна выдать нейросеть.

То есть алгоритм обучения с учителем является любой алгоритм машинного обучения, где мы выдаем алгоритму как ему нужно с нашей точки зрения правильно поступить.

Обучение без учителя — В этом случае, испытываемая система спонтанно обучается выполнять поставленную задачу без вмешательства со стороны человека.

Однако возникает вопрос, если мы не даем алгоритму явного указания как поступать правильно, то чему же алгоритм научиться? Как алгоритм будет обучаться, не зная куда ему корректировать свои внутренние параметры, чтобы поступать правильно и в конечном итоге решить задачу так как нам бы хотелось?

В этом случае мы не передаем программе заранее размеченные правильные ответы, однако мы даем нейросети посчитать результат с теми настройками что у неё уже есть и затем говорим правильное это решение или нет. В таком случае нейросеть обучается сама, а человек просто оценивает её действия.

Принцип «поощрения» нейросети

Однако в таком виде способы не сильно отличаются, так как и в том и в том случае человеку необходимо самому анализировать входные обучающие данные. Поэтому чтобы облегчить программисту работу был придуман еще

один принцип, по которому среда, в которой обучается нейросеть может «поощрять» нейросеть в случае правильного ответа, получая таким образом обратную связь. К примеру, нейросеть управляющую игрушечной машинкой можно поощрять за успешно пройденное расстояние, без касаний границ дороги.



Рисунок 1 – Синий кружок, это обучаемая нейросеть, чёрная стрелка это выдаваемый ею результат, красная стрелка – награда или поощрение.

Подобный способ действительно позволяет полностью убрать необходимость размечать наборы входных данных, однако применим он далеко не всегда, например в том же распознавании текстов нет алгоритма, дающего всегда правильный ответ - что написано на отсканированном документе.

Однако если бы можно было заранее знать, что будет написано на следующей сканированной рукописи то можно по ответу нейросети скорректировать её параметры. Узнать это можно только в том случае, если картинки, передаваемые на вход нейросети тоже были бы сгенерированы каким-нибудь алгоритмом. Допустим можно взять вторую нейросеть, которую необходимо обучить генерировать рукописные тексты, но тогда возникнут те же самые проблемы что и с первой нейросетью. Тогда почему бы не обучать обе нейросети одновременно на результатах друг друга? Получится что нейросети как-бы состязаются друг с другом генерируя рукописный текст и распознавая его. Таким образом и получается генеративно-состязательная модель (GAN).

Генеративно-сопоставительная модель

Это класс фреймворков глубокого обучения со структурой генеративной модели. Их задача — генерировать новые, сложные (выходные) данные: например, изображения или аудиофайлы, которых до генерации не существовало.

Поскольку в GAN используются две нейросети тренирующие друг друга, то для обучения нужен только набор данных (изображений, аудио и т. п.), которые хочется скопировать или имитировать. Сеть сама определит по какому принципу будут создаваться новые данные, которые будут выглядеть как из полученного набора. Иными словами, примеры показываются, чтобы вдохновить модель. Ей даётся полная свобода творчества.

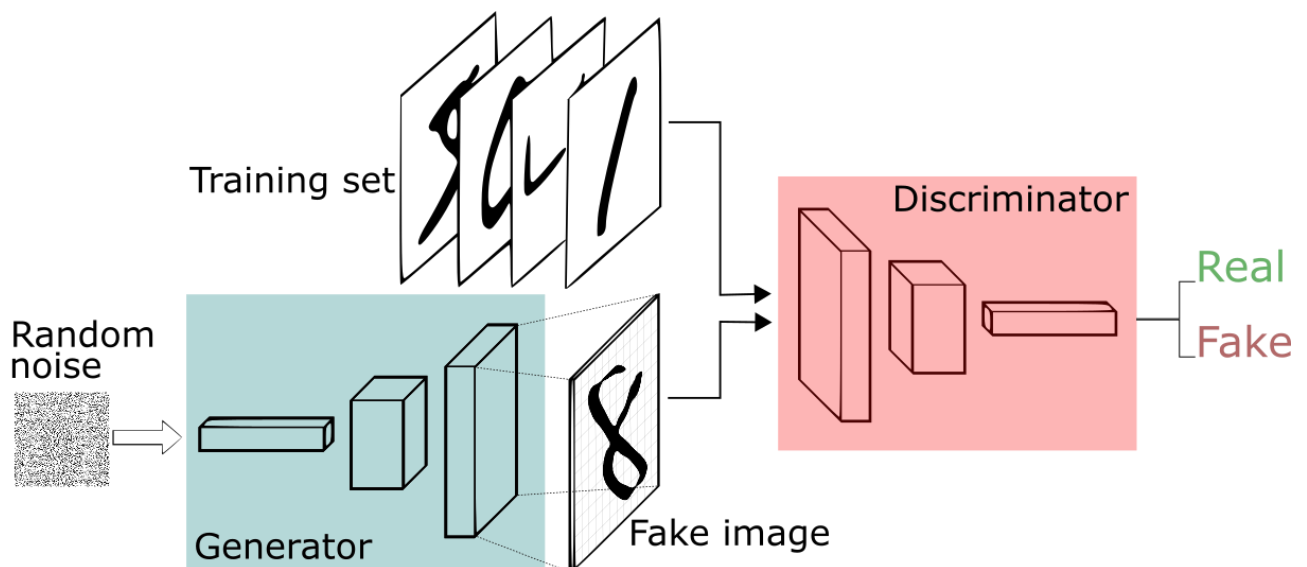


Рисунок 2 – Примерный принцип устройства GAN нейросети

Две состязающиеся между собой нейросети которые включает в себя GAN называются генеративной (G) и дискриминационной (D) или просто генератором и дискриминатором соответственно. Задача генератора — изучить функцию генерации данных, начиная со случайного шума. Дискриминатор должен определить, является ли образец данных «подлинным». При этом «подлинностью» считается принадлежность к образцам исходного набора данных. Это позволяет измерить эффективность модели и отрегулировать её параметры. Обе нейросети обучаются одновременно.

Обзор существующих решений GAN с открытым исходным кодом

Нейросеть ScrabbleGAN

Пожалуй, можно начать со ScrabbleGAN от Amazon [1] [2] – предназначена скорее для генерации текста хотя имеет внутри распознаватель, который можно использовать. Имеет довольно простую архитектуру. Есть только генератор, дискриминатор и OCR. На вход модели мы подаём текст, а генератор уже создаёт изображение. В разметке не требуется классификатор почерков или шрифтов, или же разметка bbox'ов, – нужны только картинки с текстом и аннотации для него.

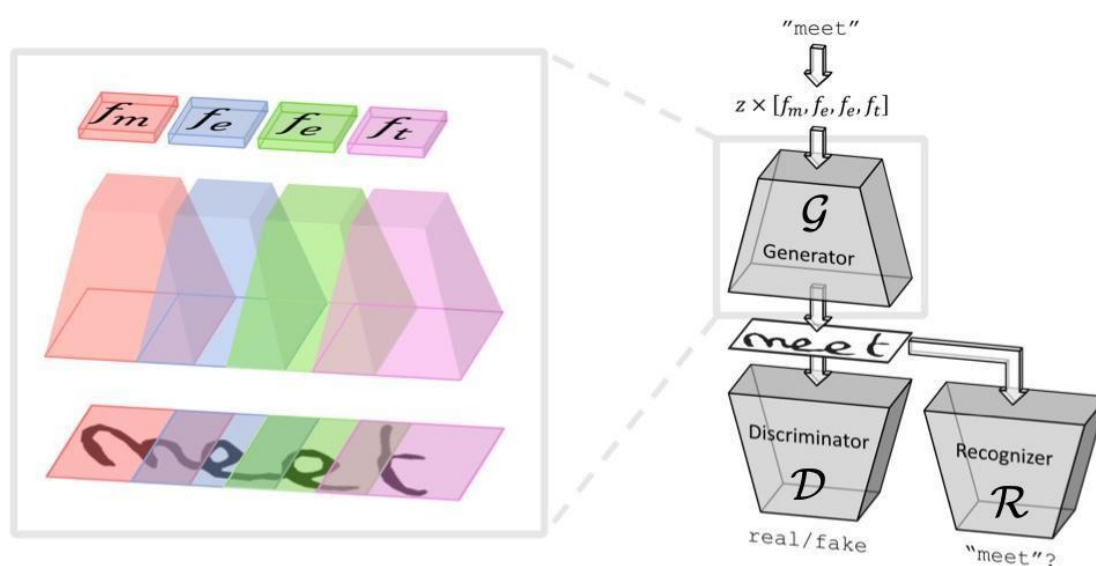


Рисунок 3 – Архитектура ScrabbleGAN

Немного об архитектуре: текст, разобранный на признаки, умножается на вектор случайного шума, проходит через linear-слои и далее повышается в качестве генератором (BigGAN-модель). Вектор шума отвечает за стиль букв: почерк/толщину/курсив и прочее. На выводе, если не менять шум, то картинки, соответственно, все будут в одном стиле. И основной недостаток ScrabbleGAN: на стиль генерации можно повлиять только с помощью вектора случайного шума. Сгенерированная картинка подаётся на вход дискриминатору, который помогает улучшить общее качество изображения, и рекогнайзеру (OCR), который делает текст читаемым.

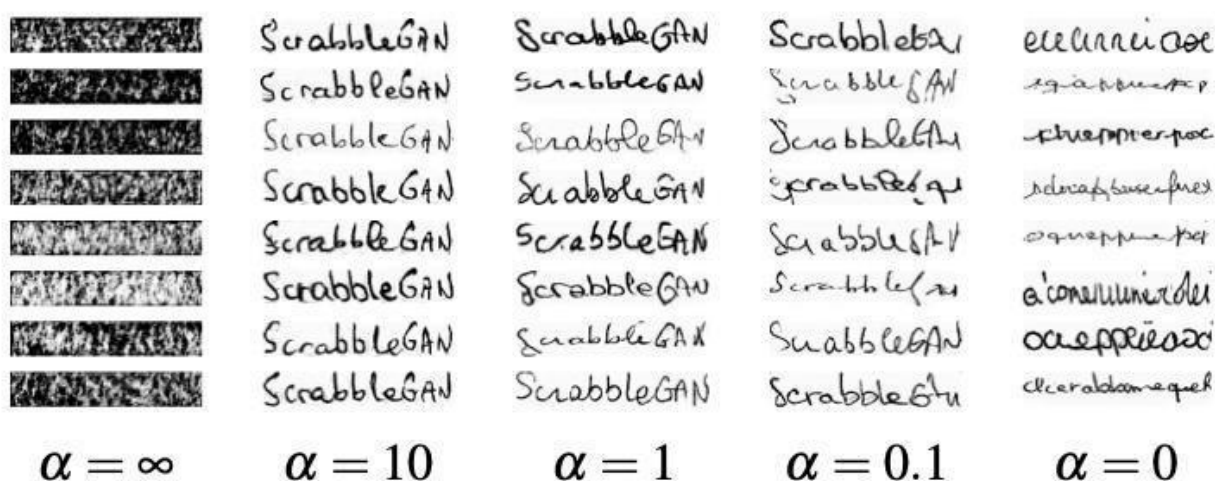


Рисунок 4 – примеры генерации рукописного текста при разных коэффициентах альфа, отвечающего за вес OCR-loss во время обучения

Нейросеть GanWriting

Она была опубликована в статье 2020-го года GanWriting [3] – Это также GAN по генерации рукописных картинок. На вход модели мы подаём текст, который хотим напечатать и несколько примеров текста с похожим шрифтом (авторы использовали 15). Нейросеть создает новый шрифт в усреднённом стиле/почерке из 15 входных примеров.

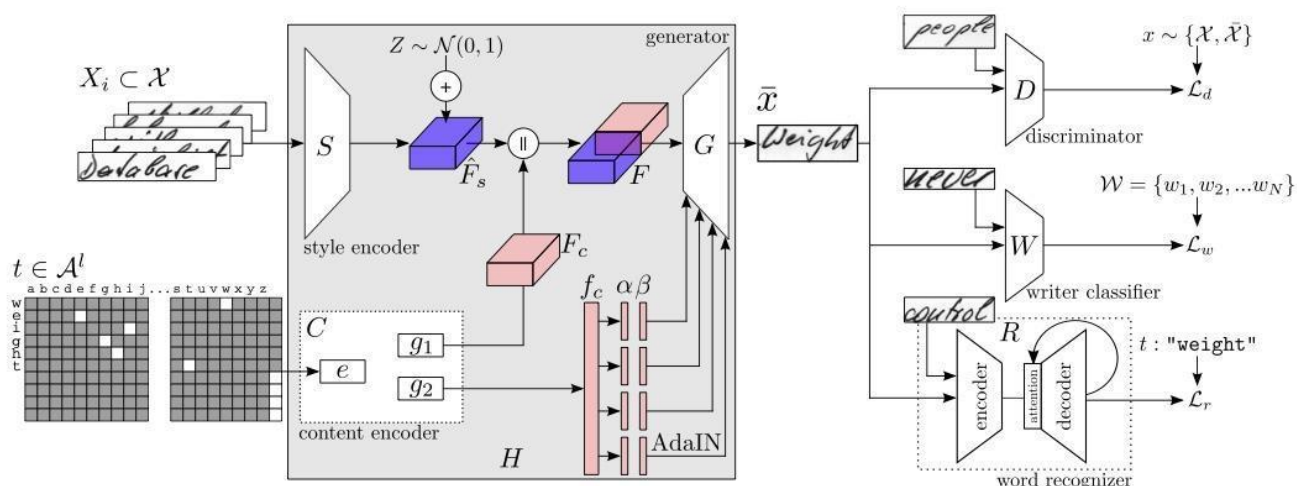


Рисунок 5 – внутренняя структура сети GanWriting

Немного об архитектуре GANwriting:

Для картинок-стилей и текста в ней есть энкодеры. Style encoder принимает на вход 15 изображений и выдаёт тензор, кодирующий стиль (авторы добавляют ещё случайный шум к тензору для искусственного создания вариативности). Content encoder принимает текст в виде one-hot матрицы (в отличие от TextStyleBrush, который рендерил текст стандартным шрифтом на белом фоне), далее энкодер делится на две головы g1 и g2. Выход g1 коннектится к выходу style encoder и такой объединённый тензор уже подаётся на вход генератору, который затем апскейлит его в результирующую картинку. Выходные контент-векторы головы g2 подаются генератору на четырёх его уровнях в AdaIN слои (тогда как в TextStyleBrush в генератор пробрасывали стиль).

В качестве loss используется дискриминатор, классификатор авторов/стилей и OCR-модель. Роль writer classifier дать генератору дополнительную информацию о типах/особенностях почерков.

Все сети модели учатся с нуля: разработчики получили лучшие результаты именно при таком подходе, чем при создании предобученных сетей какой-либо из частей архитектуры.

Также авторы приводят примеры генерации, используя в качестве входных параметров текст или изображения, которых не было во время обучения. Это очень здорово, ведь модель можно использовать для генерации синтетического датасета, имея всего несколько настоящих примеров почерка одного человека, — без необходимости дообучения.

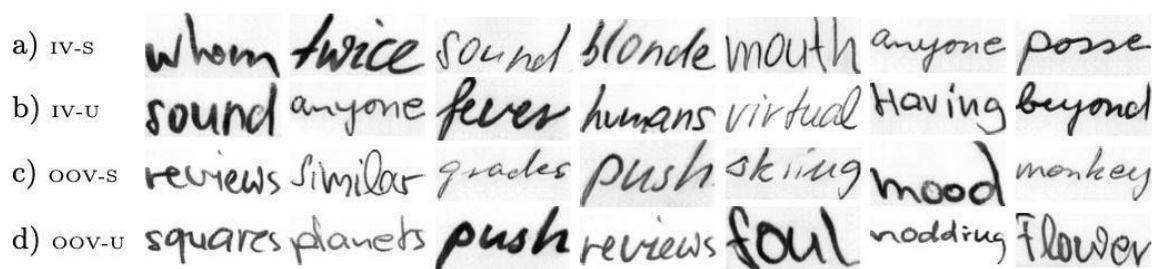


Рисунок 6 – Примеры генерации рукописного текста

Ещё одной интересной демонстрацией GANwriting является смешение стилей разных авторов. Ведь не обязательно на входе использовать 15 входных изображений одного человека. Можно смешивать почерки разных людей, и тем самым разнообразить синтетический датасет. На картинке ниже видно, как модель плавно меняет почерк в зависимости от того, сколько изображений из 15 было от одного человека и сколько от другого.

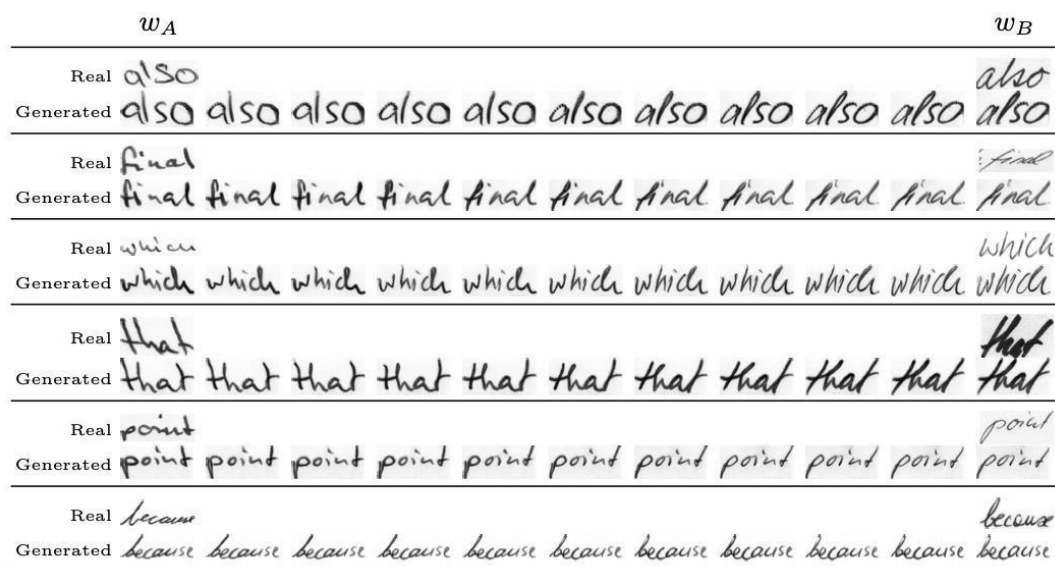


Рисунок 7 – Смешивание стилей/почерков

Исходные коды доступны в репозитории GitHub [4].

Свёрточные нейронные сети

Также в дополнение хотелось бы привести и другие OSR модели помимо GAN, ведь они скорее предназначены для того чтобы создавать данные а не распознавать их. И от GAN хотелось бы перейти к более интересному варианту - свёрточным нейронным сетям (СНС). Звучит как странное сочетание биологии и математики с примесью информатики, но как бы оно не звучало, эти сети — одни из самых влиятельных инноваций в области компьютерного зрения. Впервые нейронные сети привлекли всеобщее внимание в 2012 году, когда Алекс Крижевски благодаря им выиграл конкурс ImageNet, снизив рекорд ошибок классификации с 26% до 15%, что тогда стало прорывом. Сегодня глубинное обучения лежит в основе услуг многих компаний: Facebook использует нейронные сети для алгоритмов автоматического проставления

тегов, Google — для поиска среди фотографий пользователя, Amazon — для генерации рекомендаций товаров, Pinterest — для персонализации домашней страницы пользователя, а Instagram — для поисковой инфраструктуры.

Но классический, и, возможно, самый популярный вариант использования сетей это сортировка изображений (что очень подходит под нашу задачу OSR).

Дообученная нейросеть microsoft/tocr-small-handwritten

Открытая модель **microsoft/tocr-small-handwritten** [5], на основе которой командой энтузиастов из России была обучена нейросеть способная распознавать латиницу и кириллицу. Нейросеть построена на сверточной архитектуре Residual Networks (Resnet-34) [6].

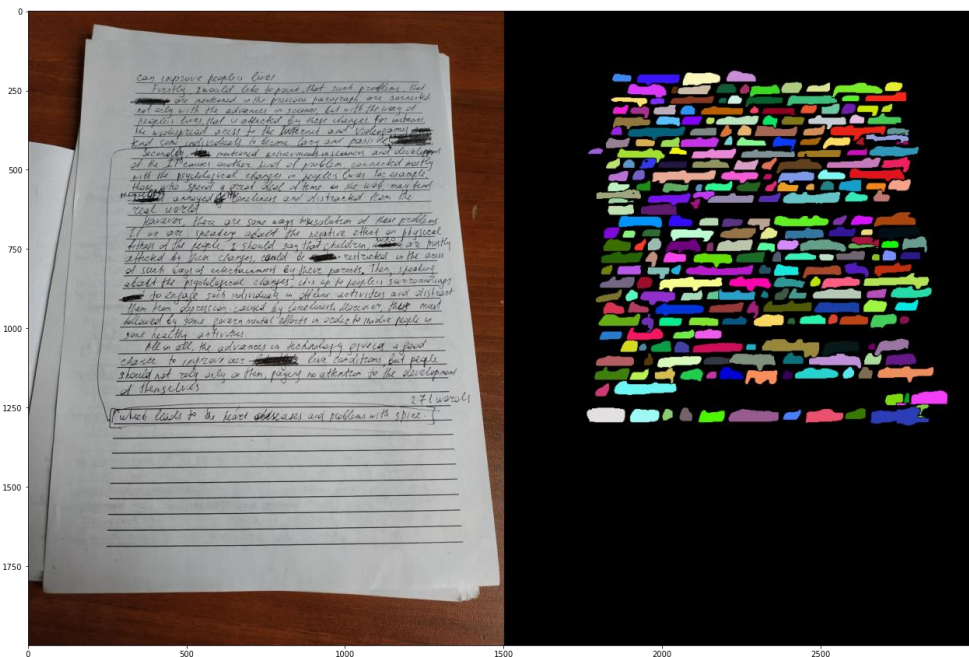


Рисунок 8 – Вычисление границ слов

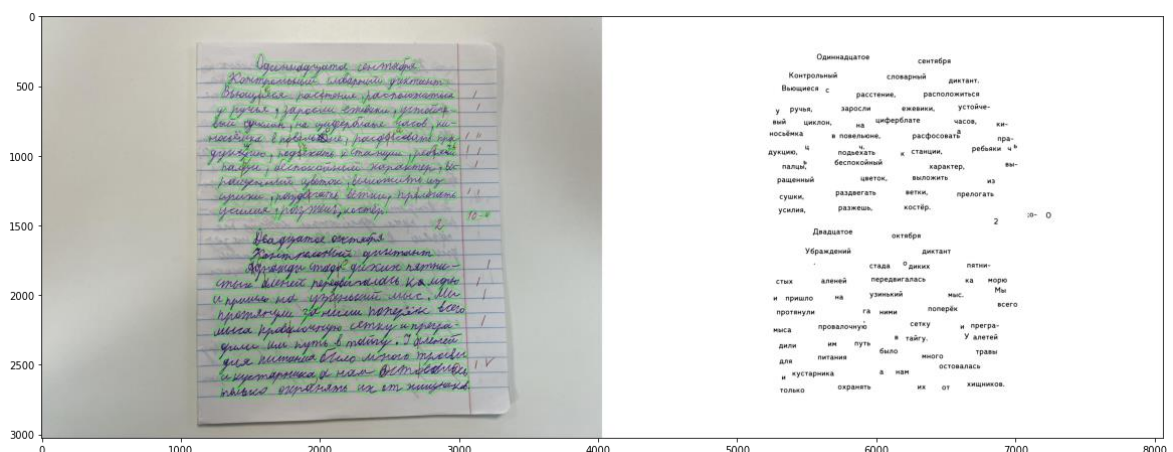


Рисунок 9 – итоги распознавания рукописного текста

Нейросеть HTR

И еще одна модель **HTR** построенная на фреймворке TensorFlow с полностью открытым исходным кодом [7] от sir-timio, обученная на русском и казахском текстах из датасета HKR [8]

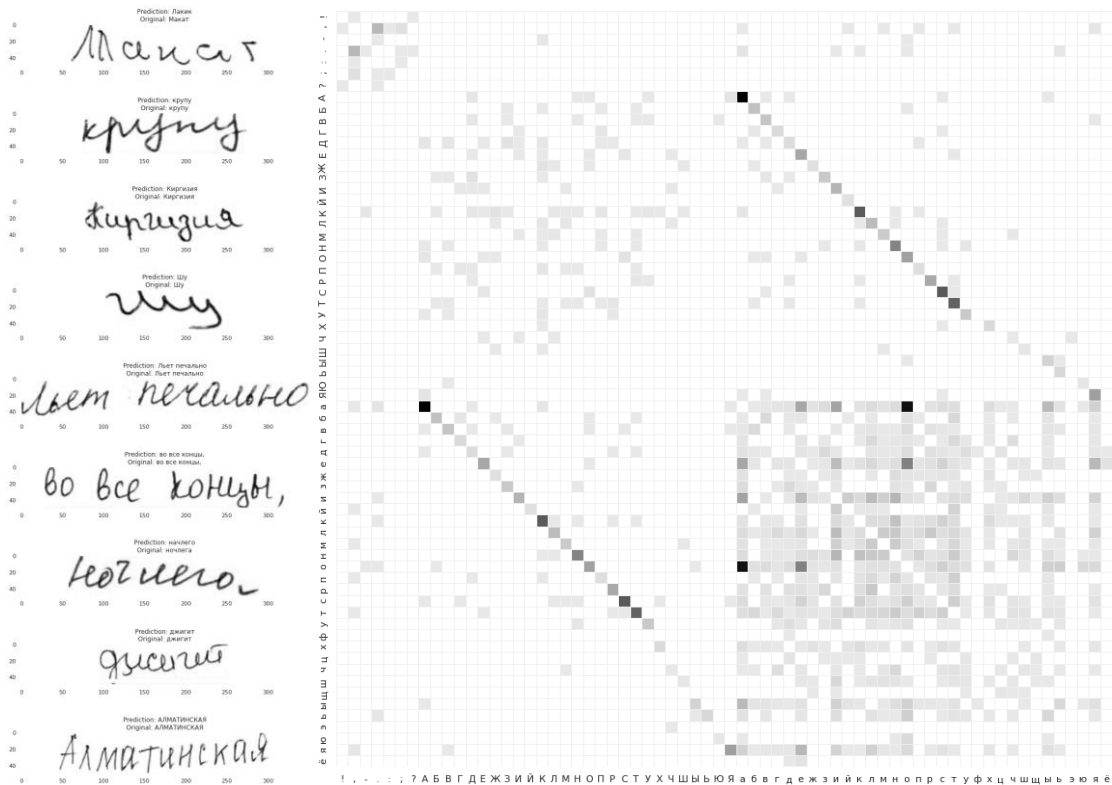


Рисунок 10 – Результат обучения HTR на датасете HKR

Код проекта открытый, и содержится в репозитории github, там же есть ссылки на сайт JUPYTER Notebook, где можно посмотреть итоги обучения, скриншоты оттуда представлены на рисунке 10 слева примеры распознавания слов, справа матрица ошибок (слева исходные символы внизу распознанные, чем чернее квадрат тем чаще эта ошибка совершалась).

Заключение

В процессе прохождения преддипломной практики я расширил теоретические знания и закрепил практические навыки, полученные во время обучения. Ознакомился с содержанием и видами основных работ и исследований, выполняемых на предприятии по месту прохождения практики. Изучил инструментальные средства и технологические цепочки, применяемые при промышленной разработке, отладке и тестировании программных систем на примере предприятия (27 ЦНИИ МО РФ), приобрёл навыки работы в коллективе.

Также мною был проведен анализ существующих методов обработки рукописных текстов, с применением технологии искусственного интеллекта. Был подобран ряд нейросетевых моделей с открытым исходным кодом, реализующих распознавание рукописного текста.

Список используемых источников

1. ScrabbleGAN: Semi-Supervised Varying Length Handwritten Text Generation [Электронный ресурс]. URL: <https://github.com/amzn/convolutional-handwriting-gan> (Дата обращения: 19.02.2023)
2. Обучение рукописной OCR на синте от GAN'ов [Электронный ресурс]. URL: <https://habr.com/ru/company/sberbank/blog/589537/> (Дата обращения: 19.02.2023)
3. GANwriting: Content-Conditioned Generation of Styled Handwritten Word Images [Электронный ресурс]. URL: <https://arxiv.org/abs/2003.02567> (Дата обращения: 19.02.2023)
4. GANwriting: Content-Conditioned Generation of Styled Handwritten Word Images Github [Электронный ресурс]. URL: <https://github.com/omni-us/research-GANwriting> (Дата обращения: 19.02.2023)
5. TrOCR (small-sized model, fine-tuned on IAM) [Электронный ресурс]. URL: <https://huggingface.co/microsoft/trocr-small-handwritten> (Дата обращения: 19.02.2023)
6. Распознавание рукописного текста в школьных тетрадях [Электронный ресурс]. URL: <https://github.com/sergak0/text-recognition> (Дата обращения: 19.02.2023)
7. Распознавание рукописного текста [Электронный ресурс]. URL: <https://github.com/sir-timio/HTR> (Дата обращения: 19.02.2023)
8. Handwritten Kazakh and Russian (HKR) database for text recognition [Электронный ресурс]. URL: https://github.com/abdoelsayed2016/HKR_Dataset (Дата обращения: 19.02.2023)