

Технології машинного навчання в системах розпізнавання даних з касового чека

ВИКОНАЛА СТУДЕНТКА ГРУПИ КА-31

КУЗАКА ОЛЬГА

КЕРІВНИК: К.Т.Н., ДОЦЕНТ, ТИМОЩУК О.Л.

Актуальність проблематики

Сьогодні машинне навчання активно використовується для розпізнавання мови і образів, а також на його основі побудовано багато пошукових системи. Машинне навчання настільки міцно закріпилося в нашому повсякденному житті, що щоденне використання його методів залишається для нас, як правило, непоміченим.

В частному випадку:

- необхідність на підприємстві вносити витрати з відрядження з чеків.

Об'єкт, предмет та мета дослідження

Мета дослідження: підібрати тип нейронної мережі, що є найбільш ефективним для розпізнавання тексту з зображення касового чека, реалізувати розпізнавання та інтегрувати з системою управління ресурсами підприємства Ахарта.

Об'єкт дослідження: зображення касових чеків.

Предмет дослідження: нейронні мережі згорткового типу, методи сегментації тексту на зображенні, методи попередньої обробки зображення.

Методи дослідження: для розпізнавання тексту з фотографії використовуються згорткова нейронна мережа, що включає в себе сім шарів різного типу, для сегментації зображення використовується метод зв'язних компонент.

Постановка задачі

- Аналіз предметної області
- Вибір типу нейронної мережі для розпізнавання тексту з зображення
- Створення алгоритму для попередньої обробки зображення.
- Спроекувати і реалізувати програмно модуль, інтегрований з системою управління ресурсами підприємства Ахарта.

Попередня обробка зображень чека

Поворот зображення так, щоб рядки чека розташовувалися максимально горизонтально, знаходження чека на зображенні і бінаризування його

- для пошуку чека була використана функція `adaptive_threshold` з бібліотеки `scikit-image`;
- алгоритм повороту необхідний, щоб максимізувати дисперсію суми яскравості по рядках;
- адаптивна бінаризація з великим порогом, яка залишає пікселі білими в областях з високим градієнтом, а більш однорідні області стають чорними;
- для бінаризації використовується той же `adaptive_threshold`, вікно досить велике, щоб воно містило і текст, і фон

Процедура сегментації тексту

1. Виделення рядків - вихідне зображення тексту необхідно "розрізати" на смуги-рядки потрібної ширини.
2. Сегментація слів - в зображенні текстового рядка виділяємо зображення слів.
3. Сегментація символів - в зображенні слова проводимо кордони символів.

Приклад тестової вибірки

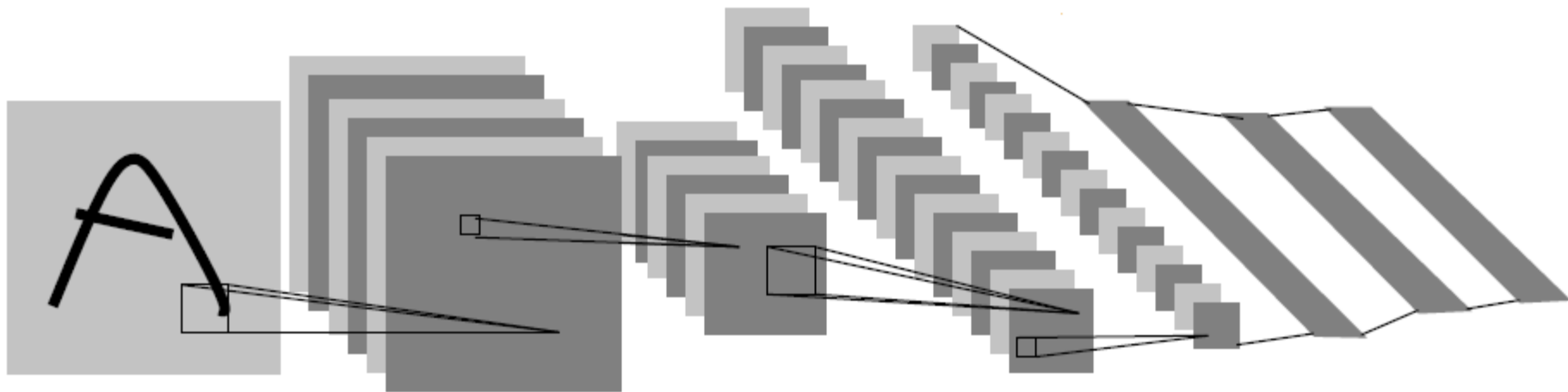
RECIRCLE

98100

Прострелившие

Розпізнавання тексту здійснюється за допомогою згорткової нейромережі, навченої на шрифтах, вирізаних з чеків. На виході з мережі маємо ймовірності для кожної букви і беремо кілька перших варіантів, які в сумі дають близьку до 1 ймовірність (99%). Далі розглядаються всі можливі варіанти складання слів з отриманих букв і перевіряються за словником. Це дозволяє поліпшити точність розпізнавання, виключаючи помилки серед схожих символів (З і Е).

Архітектура використанної нейронної мережі



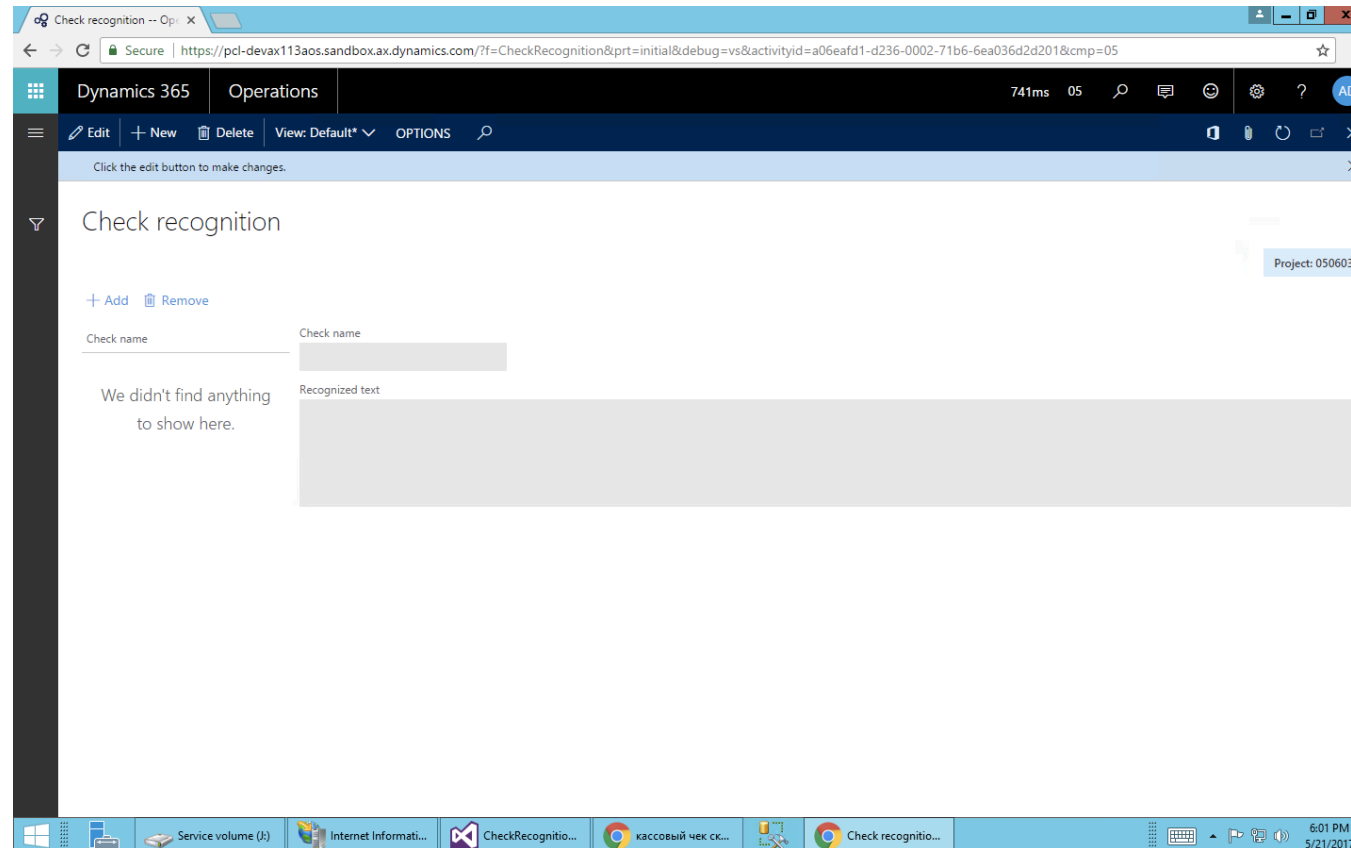
Архітектура використанної нейронної мережі

- Перший шар - шар згортки з шістьма картами ознак. Кожен нейрон на кожній з карт з'єднаний з областю 5×5 на вхідному зображенні. Розмір однієї карти - 28×28 . Перший шар містить 156 тренуваних ваг і 122304 з'єднань.
- Другий шар - шар субдискретизації з шістьма картами ознак розміром 14×14 . Кожен нейрон на картах з'єднаний з областю 2×2 на відповідній карті з першого шару.
- Третій шар - шар згортки, що складається з 16 карт ознак. Кожен нейрон з'єднаний з областю 5×5 , що знаходиться в однаковому місці на підмножині з карт ознак другого шару.
- Четвертий шар - шар субдискретизації з 16 картами ознак розміром 5×5 . Кожен нейрон в кожній карті ознак з'єднаний з областю 2×2 на відповідній карті третього шару, таким же чином, як і з'єднання між другим і першим шарами

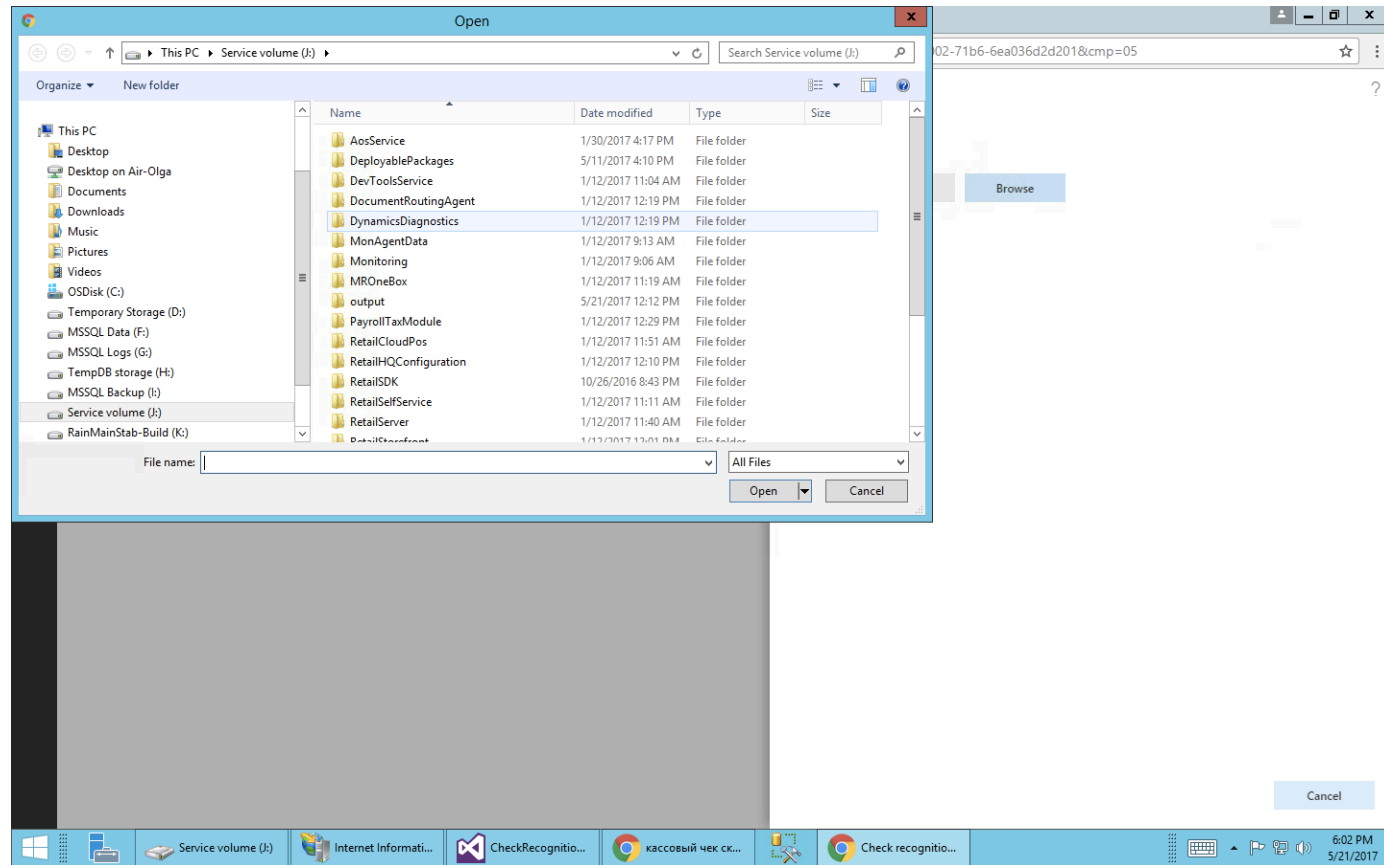
Архітектура використанної нейронної мережі

- П'ятий шар - шар згортки з 120 картами ознак. Кожен нейрон з'єднаний з областю 5x5 на всіх шістнадцяти карти ознак четвертого шару. Оскільки розмір карт четвертого шару також 5x5, розмір кожної з карт п'ятого шару - 1x1. Досягається повна зв'язність між п'ятим і четвертим шарами.
- Шостий шар - повнозв'язний, складається з 84 нейронів. Містить 10164 параметра.
- Останній, сьомий шар, складається з 10 нейронів, з'єднаних з усіма нейронами попереднього шару. Це і є вихідний шар.

Приклад роботи програмного продукту



Приклад роботи програмного продукту



Приклад роботи програмного продукту

The screenshot displays the Dynamics 365 'Check recognition' interface. The browser address bar shows the URL: <https://pcl-devax113aos.sandbox.ax.dynamics.com/?f=CheckRecognition&prt=initial&debug=vs&activityid=a06eafd1-d236-0002-71b6-6ea036d2d201&cmp=05>. The interface includes a top navigation bar with 'Dynamics 365' and 'Operations' tabs, and a secondary bar with 'Save', '+ New', 'Delete', and 'View: Default*' options. The main content area is titled 'Check recognition' and features a 'Project: 0506033' label. On the left, there is a list of checks with 'check2' selected. The 'Recognized text' field contains the following text:

000"ЛУКОЙЛ-Северо-Западнефтепродукт" АЭС М 98 "КОЛЕВЕННАЯ" СЛБ, Кожевнная линия д.43 лит А Чек н 2771/89466 Оператор Катаргина А.Л. Терн. 2 и п/п 2413 ККМ 00032709 ИНН 007825439514 Л7575 29.02.12 18:24 КАССИР 1 ПрЛАКА и1699 1.Аи-92-4(э) ТРК А 3 н199. 84_л 7.32 К 27.30 ИТОГ 199.84 НАЛИЧНЫМИ 200. 10 СДАЧА -0.26 @П 3603 0870360103 00103176 Л049150 Спасибо за покупку!

Below the text is a scanned image of a receipt with the following details:

000"ЛУКОЙЛ-Северо-Западнефтепродукт"
АЭС М 98 "КОЛЕВЕННАЯ" СЛБ,
Кожевнная линия д.43 лит А
Чек Н 2771/89466
Оператор Катаргина А.Л.
Терн. 2 н п/п 2413
ККМ 00032709 ИНН 007825439514 #7575
29.02.12 18:24 КАССИР 1
ПрЛАКА и1699
1.Аи-92-4(э) ТРК # 3
7.32 X 27.30
ИТОГ ≙199.84
НАЛИЧНЫМИ ≙200.10
СДАЧА ≙0.26
ЗКЛЗ 0870360103
00103176 #049150
Спасибо за покупку!

ВИСНОВКИ ПО РОБОТІ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В даній роботі було проведено:

- огляд нейронних мереж та методів оптимізації роботи нейронних мереж для випадку розпізнавання текстових даних з зображень, що закріплюється обґрунтуванням доцільності використання згорткової нейронної мережі.
- було розглянуто декілька способів сегментації зображення,
- був запропонований алгоритм обробки касового чека перед безпосереднім розпізнаванням, що підвищує результативність роботи нейронної мережі.
- була реалізована система з використанням мови програмування python а також C++ для інтеграції програмного продукту з системою управління ресурсами підприємства Ахарта

Дякую за увагу!