

ВИКОРИСТАННЯ ТЕХНОЛОГІЙ ГЛИБОКОГО НАВЧАННЯ ТА ЗГОРТАЛЬНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ОЦИФРУВАННЯ ДОКУМЕНТІВ ДЕРЖАВНИХ УСТАНОВ УКРАЇНИ

Денисов М.О., заст.техн.директора

ПП "Лабораторія інформаційних технологій", м.Київ

Науковий керівник – Бобарчук О.А., к.т.н, доцент кафедри КММТ

Зараз в усьому світі практикується відмова від паперового документообігу. Сучасні інформаційні технології дозволяють зберігати дані в більш зручних форматах, в порівнянні зі стандартними паперовими носіями. Документообіг державного сектора

трансформується в цифровий. У зв'язку з цим виникає потреба безпомилково і своєчасно оцифрувати необхідні документи.

Державні оцифровані документи повинні містити в собі індексовані поля для більш зручного пошуку необхідної сторінки в величезному масиві з даних. Для реалізації даного напрямку використовується технологія оптичного розпізнавання символів (*OCR — Optical character recognition*), яка застосовується перед процесом класифікації, так як це необхідно для процедури накладення певних шаблонів, які використовуються для індексування необхідних полів з документа. Однак, на практиці, в аналізованому пакеті даних можуть перебувати зайві документи, які не є необхідними до розпізнавання.

У зв'язку з цим виникає потреба класифікації документів перед їх розпізнаванням, а також їх диференціації залежно від необхідності. Таким чином для досягнення поставленої мети виділяється наступний перелік завдань:

- визначити методи класифікації документа до моменту його розпізнавання на основі їх внутрішнього контенту (матриця чисел);
- знизити вартість процесу оцифрування, а також скоротити навантаження з обчислювальних потужностей підприємств оцифрування;
- оптимізувати процес класифікації документів.

Рішенням поставлених завдань послужило впровадження згортальних нейронних мереж на основі глибоко навчання.

Глибоке навчання входить до складу машинного навчання. Машинне навчання являє собою одну з технологій штучного інтелекту, в якій обчислювальній техніці надається можливість автоматично вчитися і вдосконалюватися на основі пройденого досвіду. У машинному навчанні програмі надаються дані і відповіді, а вона на основі певних алгоритмів проводить аналіз масиву інформації і виявляє певну закономірність, після цього формує свій власний досвід, і становить інструкції, за допомогою яких і дає кінцеву відповідь. Глибоке навчання будується на самостійному моделюванні високорівневих абстракцій.

Штучні нейронні мережі (ШНМ) — це обчислювальні системи, архітектура і логіка яких заснована на принципах біологічних нейронних мереж. ШНМ складається з штучних нейронів, які представляли собою спрощену модель біологічного нейрона. Штучні нейрони в ШНМ зібрані в шари, які пов'язані між собою.

Для створення ШНМ з метою досягнення поставлених завдань була використана одна з найперших відкритих моделей нейромереж — багатошаровий перцептрон по Румельхарту, який має додаткові шари нейронів. В даному перцептроні всі шари навчаються за методом зворотного поширення помилки.

При методі зворотного поширення помилки використовується процес обчислення

градієнта. Обчислення стандартного градієнта при великій кількості штучних нейронів займає тривалий період часу, тому на великих вибірках використовується стохастичний градієнтний спуск для знаходження помилки і нормалізації ваги.

Однак, звичайні повнозв'язні ШНМ недостатньо ефективно справляються з класифікацією зображень, в наслідок чого були розроблені згорткові нейронні мережі (ЗНМ).

Згортальна нейронна мережа є спеціалізованою архітектурою ШНМ. Її основна мета — розпізнавання образів на об'єкті. Основна відмінність згортальних шарів від повнозв'язних в тому, що вони зберігають просторову структуру зображення. Ядро згортки проходить по всьому зображенню, поелементно збільшуючись на кожне його значення, результат підсумовується і заноситься в нову матрицю, яка називається «карта ознак». Так формується один згортальний шар. Цей процес іменується просторовою згорткою (рис 1).

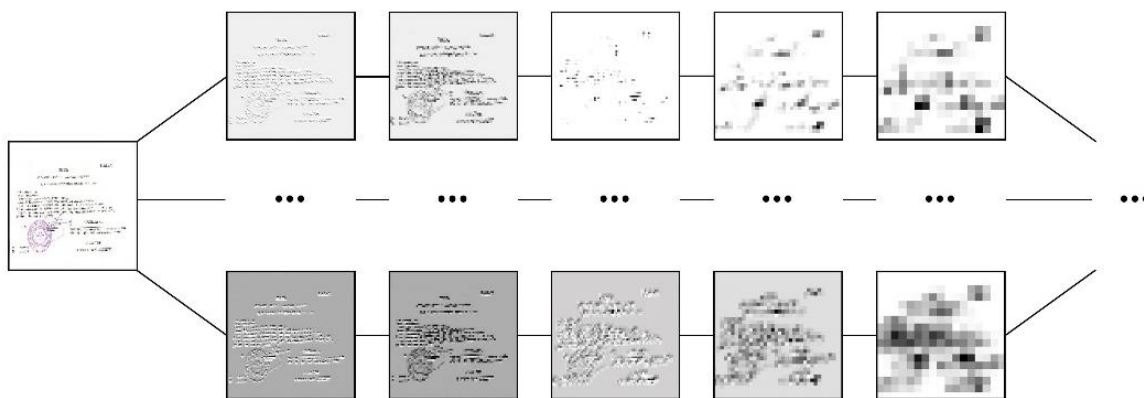


Рис. 1. Операція згортки

Детальніше даних процес можна описати таким чином. ЗНМ приймає на вхід зображення. Для коректної подачі зображення на навчання використовується процес передоброби, далі вихідне зображення групується до значень 175x175 пікселів. Таким чином на виході формується матриця значень. Для прискорення процесу навчання всі значення матриці діляться на 255. У результаті на вхід в перший згортальний шар подається матриця, що складається з 30 625 значень в діапазоні від 0 до 1. Після попередньої обробки зображення проходить через 4 згортальних шари, на кожному з яких виходять карти ознак. На виході з згортальних шарів формується 128 карт ознак, розмірністю 7x7 пікселів, які конвертуються в вектор, що складається з 6 272 елементів. Далі йдуть два повнозв'язних шари, перший вміщує в собі 768 нейронів, а другий кінцеві 4 нейрона, які в кінцевому підсумку визначають тип документа.

Таким чином, така архітектура нейронної мережі видає результат успішності в 93% на 120 тестових зображеннях.

Виходячи з цього можна зробити висновок про те, що використання згортальних нейронних мереж оптимізує і прискорює процес оцифрування документів, а також

знижує вартість даної послуги, внаслідок попередньої класифікації документів в залежності від рівня потреби в процесі оцифрування.

Список використаних джерел

1. Воронцов К.В. Машинне навчання. Курс лекцій / К.В. Воронцов. - 2019. [Електронний ресурс] – Режим доступу до ресурсу: <http://www.machinelearning.ru/wiki/index.php>.