

Аналіз методів побудови моделей нейронних мереж для розпізнавання образів

Вінницький національний технічний університет

Анотація

У даній роботі розглянуто декілька моделей нейронних мереж для подальшого їх використання

Ключові слова: нейронна мережа, алгоритм, класифікації зображень

Abstract

In this article considered the principles of working neural networks with tools for object recognition.

Keyword: neural network, library, algorithm

Вступ

Це зріла, але захоплююча галузь із досить високими темпами розвитку, яка є основою для розвитку суспільних сфер, таких як комп'ютерне бачення, обробка зображень, аналіз тексту та документів, а також нейронні мережі. Розпізнавання образів має застосування в статистичному аналізі даних, обробці сигналів, аналізі зображень, пошуку інформації, біоінформатиці, стисненні даних і комп'ютерній графіці[1]. Він бере свій початок у статистиці та технологіях; завдяки доступності великих даних і збільшенню нових потужностей обробки деякі сучасні методи розпізнавання образів включають використання машинного навчання. Розпізнавання образів можна визначити як класифікацію даних на основі отриманих знань або статистичної інформації, отриманої із зображень або їх представлень.

Аналіз методів побудови моделей нейронних мереж для розпізнавання образів

При підготовці до розробки програми було розглянуто кілька моделей нейронних мереж, у тому числі VGGNet, ResNet та ін.

Архітектура мережі VGGNet була представлена Сімоніаном і Цисерманом у їхній статті 2014 року «Глибокі згорткові мережі для розпізнавання великомасштабних зображень».

Основними типами мереж, що використовуються, є VGG16 і VGG19. «16» і «19» представляють кількість вагових шарів у мережі.

Мережа характеризується своєю простотою, використовуючи лише 3×3 згорткові шари, складені разом на зростаючій глибині. Зменшення розміру тома відбувається за рахунок максимальних агрегатів. Два повністю пов'язаних шару, кожен з яких має 4096 вузлів, за якими йде класифікатор softmax.

Автори виявили, що VGGNet важко навчити (особливо для конвергенції в глибших мережах), тому для полегшення навчання вони спочатку навчили меншу версію VGG з меншою кількістю шарів ваги[2].

Менші мережі об'єднуються, а потім використовуються як ініціалізація для більших, глибших мереж – процес називається попереднім навчанням.

Логічно, що попереднє навчання є дуже трудомістким і виснажливим завданням, яке потребує навчання всієї мережі, перш ніж її можна буде використовувати як ініціалізацію для більш глибоких мереж. Це виявило перший істотний недолік моделі, а саме надзвичайно довгий час навчання мережі. Другий недолік - висока вимогливість до місця на жорсткому диску і пропускної здатності інтернет-каналу.

ResNet – це штучна нейронна мережа, заснована на відомій структурі пірамідних клітин кори головного мозку. Типова модель ResNet реалізована з двома або трьома шарами каналів з нелінійностями (ReLU) і нормалізацією між ними[3]. Однією з мотивів для пропуску шарів є уникнення проблеми зникнення градієнтів шляхом повторного використання активацій із попередніх шарів, доки наступний шар не дізнається свої ваги. Під час тренування вагові коефіцієнти регулюються, щоб придушити верхні шари (що потребують уточнення) і підсилити раніше пропущені шари. У найпростішому випадку регулюються лише ваги з'єднань суміжних шарів без явних вагових коефіцієнтів верхнього рівня. Цей метод найкраще працює, коли окремі нелінійні шари перетинаються або коли всі проміжні шари є лінійними. Якщо ні, слід дослідити явну матрицю ваг для втрачених з'єднань (слід використовувати HighwayNet)[4].

Traversal ефективно спрощує мережу, використовуючи менше рівнів на початкових етапах навчання. Це прискорює навчання, зменшуючи ефект зникнення градієнтів, оскільки існує менше шарів для поширення[5]. Потім мережа поступово відновлює втрачені шари, вивчаючи простір об'єктів. Наприкінці навчання, коли всі шари розгорнуті, він знаходиться ближче до колектора і, таким чином, навчається швидше. Нейронні мережі без залишків досліджують більшу частину простору функцій

Висновки

З аналізу зрозуміло, що не існує універсальної моделі для розпізнавання образів. Нейронні мережі використовують математичні моделі загального призначення для відповіді на конкретні питання за допомогою даних. Протягом багатьох років нейронні мережі використовувались для виявлення спам-листів, створення розумних ракет, інтелектуальних роботів і будинків, виявлення образів за допомогою комп'ютерного зору, розпізнавання мови, а також для створення системи, яка може писати (романи, вірші і т. д.), рекомендувати продукти клієнтам і прогнозувати вартість товарів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Методи розпізнавання образів: Навч. посіб. для студ. / В. М. Заяць, Р. М. Камінський; Нац. ун-т "Львів. політехніка". – Л., 2014. – 173 с. – Бібліогр.: 21 назв.
2. Джордж Стокман, Линда Шапиро. [Computer Vision Компьютерное зрение]. –2010. – 702 с.
3. Su J. One Pixel Attack for Fooling Deep Neural Networks [Електронний ресурс] / Jiawei Su. – 2019. – Режим доступу до ресурсу: <https://arxiv.org/pdf/1710.08864.pdf>
4. Image Classification [Електронний ресурс]. – 2016. – Режим доступу до ресурсу: https://shodhganga.inflibnet.ac.in/bitstream/10603/24460/10/10_chapter5.pdf.
5. He K. Deep Residual Learning for Image Recognition [Електронний ресурс] / Kaiming He. – 2015. – Режим доступу до ресурсу: <https://arxiv.org/pdf/1512.03385.pdf>.

Сідак Дмитрій Анатолійович — студент групи ЗАКІТ-21м, факультет комп'ютерних систем та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: admitriy031@gmail.com

Науковий керівник: **Севастьянов Володимир Миколайович** — к.т.н., доц. каф. АІТ, Вінницький національний технічний університет, м. Вінниця

Dmitriy Anatoliyovych Sidak is a student of the ЗАКІТ-21m group, Faculty of Computer Systems and of automation, Vinnytsia National Technical University, Vinnytsia, e-mail: admitriy031@gmail.com

Academic supervisor: **Sevastyanov Volodymyr Mykolayovych** — Ph.D., Assoc. prof. АІТ, Vinnytsia National Technical University, Vinnytsia