

5. Федорчук О.С., Сторожук О.В. Практика применения инструментов стимулирования прямых иностранных инвестиций зарубежными странами. *Економіка і суспільство*. 2017. № 9. С. 120-126.

6. Чернишова Л.О., Сазонова В.В. Особенности залучення та стимулювання іноземного інвестування в Україні. *Бізнес інформ*. 2013. № 2. С. 87-90.

E-mail: sanazsv@ukr.net

УДК 330.36.012.4

Гнот Т.В.

аспірант,

Національний університет біоресурсів
та природокористування України

АВТОТЕГУВАННЯ ТА ВИКОРИСТАННЯ ВІЗУАЛЬНИХ ХАРАКТЕРИСТИК ЗОБРАЖЕНЬ В ІНТЕРНЕТ-МАРКЕТИНГУ

На сьогоднішній день в світі швидкими темпами розвивається інтернет-маркетинг. Все більше товарів продається через онлайн-ресурси. Це спричиняє продукування великої кількості візуального контенту, який відіграє одну з важливих ролей при виборі товару користувачем. Вміння автоматизовано обробляти весь цей контент дозволить не лише зекономити кошти, але й спричинить до зростання продаж за рахунок ефективних рекомендацій. У статті розглянуто і побудовано моделі для автоматичного тегування товарів з використанням нейронних мереж та "fine-tuning", розроблено підхід до побудови рекомендацій на основі глибинних характеристик зображень та алгоритму найближчих сусідів, а також проведено порівняння глибинних характеристик, отриманих за допомогою різних моделей.

Ключові слова: візуальні рекомендації, глибинні нейронні мережі, інтернет-маркетинг, ResNet50, глибинні характеристики, "fine-tuning" нейронної мережі, DeepFashion, тегування.

АВТОТЕГИРОВАНИЕ И ИСПОЛЬЗОВАНИЕ ВИЗУАЛЬНЫХ ХАРАКТЕРИСТИК ИЗОБРАЖЕНИЙ В ИНТЕРНЕТ-МАРКЕТИНГЕ

Гнот Т.В.

На сегодняшний день в мире быстрыми темпами развивается интернет-маркетинг. Все больше товаров продается через онлайн-ресурсы. Это приводит к продуцированию большого количества визуального контента, который играет одну из важных ролей при выборе товара пользователем. Умение автоматизировано обрабатывать весь этот контент позволит не только сэкономить средства, но и повлечет рост продаж за счет эффективных рекомендаций. В статье рассмотрены и построены модели для автоматического тегирования товаров с использованием нейронных сетей и "fine-tuning", разработан подход к построению рекомендаций на основе глубинных характеристик изображений и алгоритма ближайших соседей, а также проведено сравнение глубинных характеристик, полученных с помощью различных моделей.

Ключевые слова: визуальные рекомендации, глубинные нейронные сети, интернет-маркетинг, ResNet50, глубинные характеристики, "fine-tuning" нейронной сети, DeepFashion, тегирования.

AUTOTAGGING AND USING OF VISUAL CHARACTERISTICS OF IMAGES IN INTERNET-MARKETING

Hnot T.

Internet-marketing is growing fast in today's world. More and more goods are sold through online resources. This causes of a production of the large amount of visual content, which plays one of the important roles when choosing a product by the customer. An ability to process all this content in an automated way will not only save business's money, but will also lead to increase in sales through effective recommendations. In this article, we have built models for images tagging using neural networks and fine-tuning. Also, we have developed an approach for extracting of recommendations based on the deep characteristics of images and nearest neighbors algorithm, as well as have compared deep characteristics calculated with different models.

Keywords: visual recommendations, deep neural networks, Internet-marketing, ResNet50, deep characteristics, fine-tuning of neural network, DeepFashion, tagging.

Постановка проблеми. В сучасному світі є багато методів побудови рекомендацій для користувачів, починаючи з використання атрибутів товарів і закінчуючи аналізом оцінок товарів, проставлених користувачами. В даному дослідженні ми пропонуємо включити в побудову рекомендацій також інформацію, отриману від обробки візуального контенту, для прикладу – з фотографій продуктів.

Візуальна інформація зберігається в пікселях зображень. Але репрезентація зображення як піксельної матриці не є найкращим представленням характеристик зображення. Такі значення значною мірою залежать від позиції об'єкта на зображенні, освітленні, тощо. Кращим варіантом представлення зображення є використання «глибших» його характеристик, які можна отримати, використовуючи нейронні мережі.

Таким чином, маючи векторні представлення візуального контенту, система буде спроможна здійснювати пошук «потенційно цікавого» контенту для користувача, базуючись на його минулих вподобаннях чи теперішньому зацікавленню.

Іншою важливою проблемою в онлайн-маркетингу є тегування контенту. Теги відіграють важливу роль при пошуку потрібного продукту. А так як щодня інтернет-магазини поповнюються великою кількістю нових товарів, проставлення тегів потребує багато ручної роботи. З допомогою машинного навчання, цю рутину можна автоматизувати. У дослідженні представлено підхід до побудови класифікаторів зображень для прогнозування тегів, виходячи з умов досить малої розміченої навчальної вибірки.

Аналіз останніх досліджень і публікацій. В статті “Image-based Product Recommendation System with Convolutional Neural Networks” [1] автори розглянули питання пошуку в платформах онлайн-рітейлерів, які зазвичай використовують пошук за ключовими словами. Такий пошук є неефективним і в значній мірі залежить від продавця-покупця. У дослідженні автори представили «розумну» пошукову систему для

онлайн-магазинів. Вона використовує зображення і на їх основі прогнозує інформацію про них у формі тегового описання. Рішення оформлене як дві нейронні мережі: перша прогнозує категорію, друга здійснює пошук подібних зображень. Було досягнуто точності системи близько 50% на даних з Amazon Dataset. На нашу думку, точність системи не є достатньо високою для достовірних прогнозів, а пошук подібних зображень можна було здійснювати більш простішими способами, досягаючи не гірших результатів.

У статті “Deep Learning based Large Scale Visual Recommendation and Search for E-Commerce” [2] автори комплексно підійшли до вирішення проблеми. Вони представили цілісне рішення для візуального пошуку і рекомендацій. Автори запропонували уніфіковане рішення у вигляді глибокої згорточної нейронної мережі VisNet для вивчення характеристик зображень в розрізі декількох семантичних гранулярностей.

В обох статтях автори працювали з досить великими датасетами для побудови класифікатора. У дослідженні ми акцентуємо увагу на тому, що високоточні класифікатори для тегування зображень можуть бути побудовані з використанням не більше кількох сотень зображень на один клас. Більш того, ці класифікатори можуть бути використані і для продукування «глибоких» характеристик для рекомендацій.

Постановка завдання. Метою дослідження є побудова системи для автоматичної генерації тегів зображення, а також пошуку найбільш подібних зображень на основі «глибинних» його характеристик.

Виклад основного матеріалу дослідження. Дані. При побудові алгоритмів було використано частину розмічених зображення з DeepFashion [3] датасету (46985 фотографій, 75% з яких було використано для навчання моделі) для навчання 46-класового класифікатора одягу. DeepFashion датасет містить зображення і розмічені категорії одягу, до яких ці зображення можна віднести. На рис. 1 зображено декілька випадкових зображень з цього датасету.

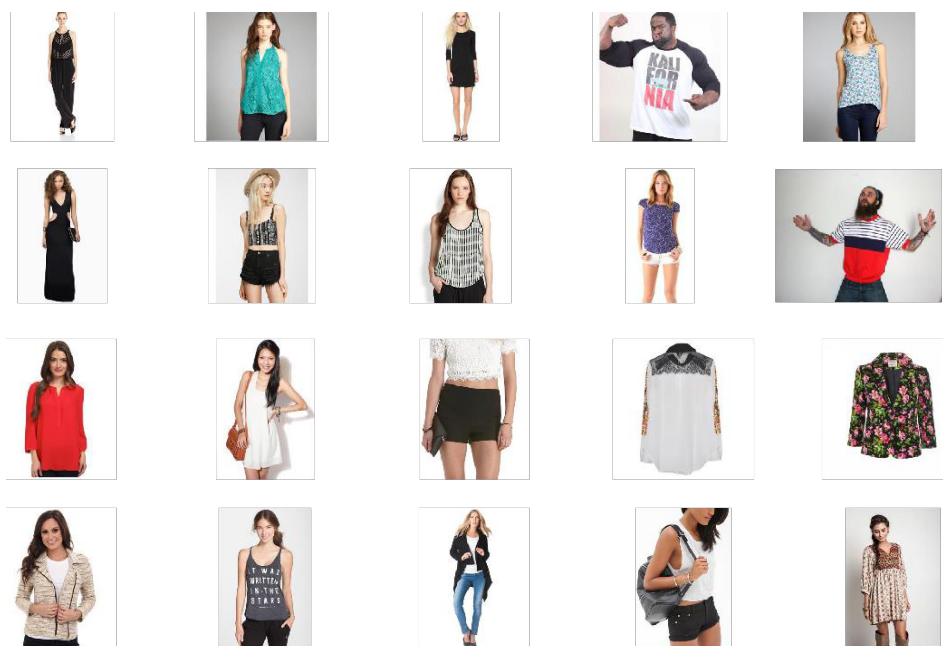


Рис. 1. DeepFashion датасет

Таблиця 1. Список тегів та приклади зображень з навчальної вибірки

Тип тегу	Клас	Приклади			Клас	Приклади		
Патерн (3439)	Квітковий				Однотонний			
	Графічний				В горошок			
	В клітинку				В смужку			
Довжина рукавів (2955)	Без рукавів				Довгі рукави			
	Короткі рукави							
Колір (4456)	Червоний				Зелений			
	Фіолетовий				Білий			
	Жовтий				Бірюзовий			
	Коричневий				Сірий			
	Синій				Чорний			

Також для навчання моделей для прогнозування патернів одягу, довжини рукавів та кольору було завантажено 10850 фотографій з Google (табл. 1).

Для побудови рекомендацій (пошуку найбільш подібного) було використано всі фотографії з DeepFashion (290 тис.).

“Fine-tuning” для побудови класифікаторів. “Fine-tuning” [4] – це підхід, який досить часто використовують для навчання нейронних мереж на картинках при обмеженому наборі навчальної вибірки. Основна ідея такого підходу полягає у використанні натренованої моделі на подібній класифікаційній задачі (в нашому випадку – класифікаторі зображень). В цьому випадку можна зберегти ваги і ядра перших шарів нейронної мережі, які відповідають за розпізнавання базових характеристик, а перенавчити останні, які безпосередньо здійснюють класифікацію. Є певні практики здійснення “fine-tuning”:

1. Поширеним є відрізання останнього шару навченої нейронної мережі і заміна його з softmax-шаром, релевантним до поставленої задачі. Після цього проводять навчання нейронної мережі з новими даними.

2. При навчанні потрібно використовувати досить мале значення кроку (~в 10 разів в порівнянні з базовим навчанням). Так як ми очікуємо, що попередньо навчені ваги є досить непоганими, в порівнянні з випадковими ініціалізованими вагами, і ми не хочемо їх надто відхилити від оптимального значення дуже швидко і забагато.

3. Також часто «заморожують» ваги перших декількох шарів попередньо-натренованої мережі. Це робиться тому, що перші шари вловлюють загальні характеристики зображення, для прикладу криві і кути, які також релевантні до поставленої задачі і ми хочемо зберегти цю взаємодію ваг і ядер. Натомість, ми повинні зосередитися на навчанні характеристик, специфічних до нашого датасету, які вилловлюються останніми шарами.

“Fine-tuning” використовують для навчання, коли немає достатньо даних, обчислювальних ресурсів чи часу. Найбільш точніші нейронні мережі навчаються на мільйонних датасетах протягом декількох днів, і це при умові, що присутні суттєві обчислювальні ресурси. При “fine-tuning” ми можемо використати вже ці навчені моделі, тільки «перенаправити» їх на іншу задачу.

Для “fine-tuning” потрібна базова навчена модель. В нашому дослідженні ми використали Res-Net50 [5]

архітектуру, навчену на ImageNet датасеті (1000 класів зображень, 1.28 мільйонів розмічених картинок).

Для прогнозування тегів було побудовано 4 нейронні мережі (категорія одягу, патерн, довжина рукава та колір). Це було спричинено тим, що для різних характеристик були різні набори даних, тому не було можливості будувати спільну модель. Кожна з моделей була натренована подібним шляхом, донавчаючи Res-Net50.

Для класифікатора 46-категорій одягу був використаний наступний підхід до “fine-tuning”:

1. Відрізати останній шар з 1000 неронами.
2. Додати два повносв'язні шари (256, 64 нейрони) з активацією RELU.
3. Додати вихідний шар з 46 можливими виходами, використовуючи SOFTMAX активацію.
4. Заморозити всі ваги, крім тільки що доданих.
5. Навчати нові ваги протягом 10 епох з оптимізатором ADAM.
6. Відморозити всі ваги.
7. “fine-tune” всі ваги протягом 10 епох з дуже малим навчальним кроком, як 0.001.

Точності, отримані для моделей для прогнозування тегів, наведені в табл. 2.

Досить низька точність моделі «Категорії» спричинена досить великим набором категорій, великим перекриттям категорій і тим, що велика кількість фотографій одночасно була розмічена в різні категорії. Щоб покращити точність прогнозу, був введений поріг для кожної з моделей (якщо достовірність прогнозу менша за поріг, тобто прогноз недостатньо надійний, модель пропускає цей кейс і нічого не повертає). Остаточна точність моделей з введеними порогами наведена в табл. 3.

Нижче, на рис. 2 наведено приклади роботи системи.

Система візуальних рекомендацій. Для здійснення рекомендацій на основі візуального контенту, фотографії повинні бути представлені у векторному вираженні. Таким чином, знаючи, що користувачеві сподобався продукт, представлений фотографією a_i , ми зможемо порекомендувати йому продукти, представлені фотографіями $a_{j,...,k}$, якщо $dist(a_i, a_{j,...,k}) \rightarrow \min$, для будь-якого товару $j, ..., k$.

Векторні представлення зображень можна отримати, використавши моделі для класифікацій. Передостанні шари нейронної мережі, перед повносв'язними шарами, які здійснюють класифікацію, містять детальні характеристики зображення, важливі для

Таблиця 2. Точності моделей прогнозування тегів

Модель	Точність	Точність (top3)
Категорії	0.5	0.76
Патерни	0.9	
Довжина рукавів	0.9	
Кольори	0.91	

Таблиця 3. Точність моделей з порогами

Модель	Поріг	Точність
Категорії	0.15	0.65
Патерни	0.81	0.95
Довжина рукавів	0.7	0.93
Кольори	0.81	0.948

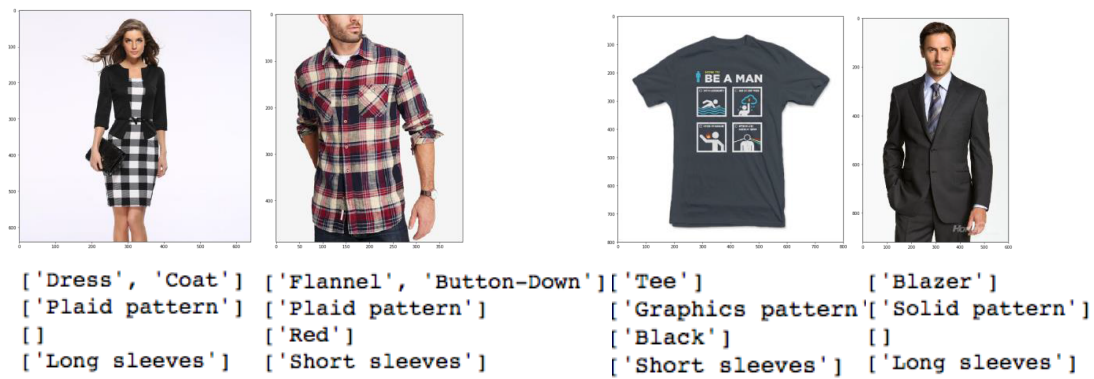


Рис. 2. Приклади роботи системи прогнозування тегів



Рис. 3. Результат рекомендаційної моделі на основі характеристик, отриманих з чотирьох і одного класифікатора

класифікації. Їх можна використовувати як числовий опис контенту. В нашому дослідженні ми використали активацію шару, розміщеного перед двома останніми шарами. Це 1000-елементний вектор.

Побудова моделі для рекомендацій була здійснена наступним чином:

1. Обчислено числові описові характеристики всіх зображень з датасету (290 тис.) на основі чотирьох

побудованих моделей та об'єднано їх в один вектор. На виході – кожне зображення описане 4000-елементним вектором.

2. Для пошукового зображення проведено аналогічні розрахунки для отримання 4000-елементний опис.

3. Обчислено Евклідову відстань від пошукового зображення до 290тис. з датасету і знайдено топ найбільш подібних, які і можна вважати хорошими претендентами для рекомендацій.

Звісно, «глибинний» опис зображення можна провести і на основі тільки однієї моделі, але в такому випадку результат буде концентруватися на одну з характеристик як найважливішу, наприклад категорію одягу, частково ігноруючи колір, патерн тощо. На рис. 3 представлено порівняння моделей з використанням тільки категоріальної і моделі з використанням

всіх чотирьох. Перший ряд зображень відповідає за 4000-елементний опис, 2-ий – за опис тільки за допомогою категоріальної моделі. Зображення посортовані в порядку подібності до пошукового зліва направо.

Висновки з проведеного дослідження. Моделі класифікації дозволяють будувати високоефективні механізми автотегування. В дослідженні показано, що ці моделі можна легко і швидко навчати на невеликих вибірках даних, використовуючи таку техніку як “fine-tuning”. За допомогою цих моделей можна також отримати «глибинний» опис візуального контенту, який надалі можна використати в багатьох задачах. Однією з таких задач є рекомендаційна система. Більше того, опис, отриманий на комбінації різних моделей класифікації є більш ефективний, так як дозволяє віднайти подібні продукти за багатьма характеристиками.

1. Luyang Chen, Fan Yang, Heqing Yang. *Image-based Product Recommendation System with Convolutional Neural Networks*, Stanford, 2017.

2. D. Shankar, S. Narumanchi, H. Ananya, P. Kompalli, and K. Chaudhury. *Deep learning based large scale visual recommendation and search for e-commerce*. arXiv preprint arXiv:1703.02344, 2017.

3. Ziwei Liu and Sijie Yan and Ping Luo and Xiaogang Wang and Xiaoou Tang. *Deep-Fashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

4. Ec Lin, Z.; Ji, K.; Kang, M.; Leng, X.; Zou, H. *Deep Convolutional Highway Unit Network for SAR Target Classification with Limited Labeled Training Data*. *IEEE Geosci. Remote Sens. Lett.* 2017, 14, 1091–1095.

5. Kaiming He, Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. *Deep residual learning for image recognition*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

6. Hnot T. *Recommendations Based on Visual Content*. *Proceedings of the 14th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer – Kyiv, Ukraine, May 14-17, 2018*. P. 485-487.

7. Гнот Т.В. Компаративний аналіз ефективності рекомендаційних систем в маркетингу. *Вісник Хмельницького національного університету. Економічні науки*. 2017. № 3. Т.1. С. 135-140.

8. *Технологии цифрового маркетинга: создание рекламной кампании* / В. Н. Григорьева, С. Г. Антонов, П. Ф. Воробьев, А. В. Григорьев, С. В. Федюнин, И. А. Шут / Под ред. В. Н. Григорьевой. СПб.: Издательство «Левша Санкт-Петербург», 2017. 250 с.

E-mail: tarashnot@gmail.com