

УДК 004.8

Федосова І. В., Разумний О. В.

РОЗРОБКА ГІБРИДНОЇ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ «TOARIST»

Дана стаття присвячена питанню впровадження згорткових нейронних мереж у туристичні додатки. У статті розглядаються шляхи удосконалення розпізнавання культурних об'єктів з потокового зображення камери смартфона шляхом використання фреймворку Core ML. З того приводу, що вихідний код даної бібліотеки є закритим, були проаналізовані та запропоновані зовнішні шляхи удосконалення. Серед них: препроцесінг, фільтрування, використання систем глобального позиціонування GPS та елементів експертних систем. Більшість вхідних матеріалів для навчання і створення моделі нейромережі були зібрані вручну, шляхом фотографування культурних пам'яток. У результаті роботи був отриманий працездатний прототип мобільного додатку під назвою «ToARist».

Ключові слова: розпізнавання культурних об'єктів, згорткова нейромережа, мобільний додаток.

Постановка проблеми. Мобільні технології з кожним роком набирають обертів і завойовують популярність на ринку туризму. Згідно з результатами дослідження, проведеного в 2015 році в ході щорічної виставки Business Travel Show в Лондоні, одну з провідних позицій в списку найбільш впливових технологій, що впливають на індустрію туризму, займають саме мобільні додатки [1]. Така програма, встановлена на смартфоні зможе уніфікувати і систематизувати наявну інформацію про існуючі культурні об'єкти м. Маріуполь, зручно її скомпонувати і подати кінцевому споживачеві – туристу нашого міста.

Наступним актуальним питанням є спосіб пошуку інформації, про культурний об'єкт, що зацікавив туриста. Далеко не всі пам'ятки і будівлі мають таблички із назвою та інформацією про себе. Більшість туристів не захоче вручну шукати інформацію про них в мережі Інтернет, вводячи зовнішні ознаки, адресу, тощо.

Саме тому, для цього завдання буде ефективно використовувати існуючі рішення в області технологій розпізнавання зображень на мобільних пристроях. У цьому випадку, туристу необхідно буде просто попередньо завантажити спеціальний додаток, а потім вже на місці навести камеру на культурний об'єкт і відсканувати його. Однак, навіть алгоритми Core ML, що використовуються в сучасних мобільних додатках мають проблеми із точністю розпізнавання [2]. У контексті цього питання, буде актуальною постановка проблеми про поліпшення існуючих методів виявлення об'єктів на смартфонах.

Таким чином, актуальність теми дослідження визначається впровадженням систем штучного інтелекту в мобільні туристичні програми та поліпшенням існуючих способів розпізнавання об'єктів з потокового зображення камери смартфона. У доповненні, туристичні програми подібного типу підвищують привабливість і популярність міста.

Аналіз останніх досліджень та публікацій. У ході пошуку вирішення даного питання слід звернути увагу на наступні статті вітчизняних і зарубіжних авторів. Першою слід розглянути статтю О. В. Ліпанова і М. В. Фесенко, в якій були проаналізовані методи розпізнавання об'єктів у системах аналізу візуальної інформації. У ній авторами був запропонований метод для реалізації бібліотеки аналізу відеопотоку, за допомогою якого може бути побудована система розпізнавання об'єктів різних категорій. На жаль, у роботі

представлена лише загальна модель поліпшення, без прив'язки до конкретної області та демонстрації конкретних прикладів роботи програми [3].

Наступні вітчизняні автори, Ю. Тулашвілі та Т. Жовтан, розглядали інформаційне наповнення туристичних інтегрованих систем. У ній вони зібрали базу даних про туристичні об'єкти, запропонували формальну модель і методи інтеграції ресурсів для системи інформаційної підтримки туристів. Додаток має представляти експертну систему, що втілює основну ідею OLAP, що полягає в забезпеченні навігації шляхом надання даних користувачам, щоб вони могли інтерактивно генерувати спеціальні запити без втручання фахівців з інформаційних технологій [4].

Проте, жоден з вітчизняних авторів не пропонує робочу туристичну систему, яка використовує комп'ютерний зір для розпізнавання об'єкта і надання даних про нього туристу.

У свою чергу, зарубіжні автори часто розглядають хмарні геосервіси, як додаток до комп'ютерного зору для мобільних пристроїв. Там вони стверджують, що комп'ютерний зір може використовуватися для отримання інформації про об'єкт, наприклад ідентифікація будівель, інформаційних знаків, людей, тощо. Це забезпечує індексацію на основі об'єктів, яка в кінцевому підсумку може бути застосована для отримання семантичного опису середовища. У свою чергу, геосервіси обов'язкові для підтримки розпізнавання об'єктів на основі комп'ютерного зору і семантичної індексації декількома способами. Це забезпечує попередню геоприв'язку об'єктів, що зменшує кількість можливих гіпотез об'єкта і, отже, може значно спростити складність розпізнавання. Серед основних недоліків даного підходу слід вважати необхідність використання хмарних технологій для процесу ідентифікації зображення, що потребує наявності швидкісного Інтернет-зв'язку.

З іншого боку, робота таких авторів, як Д. Марімон, Х. Сарасуа і П. Карраско, маючи подібний підхід, спираються лише на комп'ютерний зір і технологи доповненої реальності на локальному пристрої. Їх додаток «MobiAR» – це сервісна платформа Android для туристичної інформації, яка заснована на AR (Argumented Reality). Вона дозволяє користувачам переглядати інформацію і мультимедійний контент про місто через свої смартфони. Платформа не тільки обробляє інформацію про місцезнаходження, але й використовує технології комп'ютерного зору, щоб визначити туристичний ресурс, який цікавить користувача. Подібна система була створена також М. Вільямсом, К. Яо і Д. Нюрсом в роботі «An AR tourism with User Centred Design». Там зроблений упор на інтерв'ювання та опитування користувачів для отримання детального зворотного зв'язку з ними. У даному контексті, автори хотіли підкреслити те, що важливо не просто створити інтелектуальний додаток, здатний розпізнавати туристичні об'єкти, а й дати користувачеві можливість комфортно його використовувати [5].

Таким чином, публікації, що точно характеризує туристичний додаток, який ґрунтується на локальній нейромережі та одночасно використовує допоміжні технології глобального позиціонування виявлено не було.

Постановка завдання. Мета дослідницької роботи – вдосконалення системи пошуку культурних об'єктів м. Маріуполя за допомогою камери смартфона у мобільному додатку на базі згортової нейронної мережі.

У рамках реалізації поставленої мети необхідно вирішити такі завдання:

- розглянути існуючі рішення на ринку туристичних програм. Знайти їх переваги і недоліки;
- проаналізувати алгоритми нейронних мереж, доступні для використання на смартфонах. Вибрати найбільш відповідний для вирішення даної проблеми варіант;
- зібрати базу даних зображень культурних об'єктів м. Маріуполь;
- виявити ефективні підходи по попередній обробці вхідних даних (зображень);

- розробити попередню версію програми, що використовує «чисту» реалізацію обраного алгоритму;
- запропонувати способи збільшення шансу на правильне розпізнавання туристичного об'єкта для обраної нейронної мережі.

Основний матеріал дослідження. У ході вивчення праць на тему використання нейронних мереж у мобільних пристроях було знайдено багато підходів, направлених на покращення існуючих рішень. Однак, більшість із них мають лише концептуальний або математичний опис покращення. Тому для створення мобільного додатку був обраний готовий фреймворк Apple Core ML.

Він заснований на конволюційних (згорнутих) фреймворках BNNS та MPSCNN, які представляють собою сукупності математичних функцій, що повністю користуються швидкими векторними інструкціями процесора. Це дозволяє йому оптимально використовувати наявні ресурси пристрою, адже він може запускати нейромодель на CPU, GPU або на Neural Engine (нейронному процесорі), залежно від того, яке обладнання доступне на пристрої користувача.

Проте використання цієї бібліотеки обмежено браком безпосереднього доступу до вихідного коду [6]. Тому, в даному випадку, доцільно використовувати системний підхід, із використанням стратегії «чорного ящика».

Шляхом варіювання вхідних даних можна протестувати поведінку системи і зрозуміти в яких випадках оброблені дані позитивно позначаються на результаті розпізнавання. Як приклад, перед завантаженням до моделі нейромережі, зображення можна обробити через набір спеціальних фільтрів. Зниження шумів, згладжування, підвищення контрастності й яскравості можуть позитивно вплинути на виділення характерних ознак об'єкта, а як наслідок і на кінцеву якість розпізнавання. Як приклад, можна порівняти роботу нейромережі при використанні моделей з різними ступенями фільтрації та на основі отриманих вихідних даних зробити висновки про доцільність використання різних методів обробки [7].

Деякі керівництва по роботі з Core ML рекомендують використовувати допоміжні інструменти візуалізації для вивчення отриманих моделей. До таких інструментів належить Neutron, який візуально відображає модель і детальну інформацію про її шари. Також він можливість перевірки ваги і зміщення для шарів. Це може стати в нагоді для наочної перевірки отриманих даних, а також оцінки впливу тих чи інших вхідних даних [6].

Модель також можна розбити на окремі складові і подавати в систему у вигляді одиниць, розбитих за певною ознакою. У цьому і полягає суть структурного підходу до розробки систем. При цьому автоматизована система зберігає цілісне представлення, в якому всі складові компоненти взаємопов'язані.

При побудові нейромережевої моделі, призначеної для пошуку і розпізнавання архітектурних об'єктів, вони були розподілені по групах, як масиви зображень, отриманих у ході зйомки культурних пам'яток м. Маріуполь. Вхідні дані для моделі також були зібрані виходячи з доступності інформації про туристичні об'єкти в мережі Інтернет. Приклад серії зображень зроблених за різних сезонних, погодних, добових, людських та інших зовнішніх умов показано на рисунку 1.

Перед завантаженням у модель, у разі, коли зображень було недостатньо, застосовувались такі способи аугментації як: поворот, розмиття, експозиція, шум і фліп (переворот назад). Це дало можливість збільшення вхідної вибірки, однак може збільшити шанс перенавчання [6].



Рисунок 1 – Приклад серії знімків пам'ятника Т.Г. Шевченка

У випадках, коли зображення є низької якості або було знято в умовах недостатньої освітленості, необхідно застосовувати препроцесінг, метою якого є виявлення найбільш важливої частини зображення і видаленні з неї непотрібного шуму.

Фотографії класифікувалися за групами в залежності від їх близькості і місцезнаходження, а також безпосередньо як окремі культурні одиниці. Обсяг навчальної множини повинен становити не більше 10 об'єктів на модель, так як в ході експериментів була виявлена закономірність збільшення помилки прогнозування з перевагою на один з досліджуваних об'єктів. Саме дана закономірність спонукала до досліджень за принципами розбиття моделей на більш дрібні структурні одиниці за принципом глобального позиціонування, яке може бути ідентифіковано за допомогою засобів GPS, яке вбудоване в усі моделі досліджуваних смартфонів. Кожна група повинна мати множину схожих фотографій, знятих з різних ракурсів для перевірки точності одержуваного результату. Проектування, оптимізація, навчання, тестування нейронної мережі і експерименти над нейромережевою моделлю потрібно виконувати за допомогою платформи Create ML.

У процесі навчання і тестування нейромережі були виявлені викиди, виявлені за допомогою методики [8]. Її суть полягає в почерговому виключенні прикладів з навчальної множини і спостереженні за похибкою нейромережі, навченої на цих урізаних множинах. Якщо приклад навчальної вибірки є викидом і випадає з закономірності, характерною для досліджуваної предметної області, то його видалення з навчальної множини призводить до падіння похибки навчання мережі та підвищення її узагальнюючих властивостей, тоді як виняток звичайних прикладів істотного впливу на якість мережі, як правило, не надає.

Після видалення викидів, на отриманих наборах значення оцінки точності коливалося в межах 90-100 %. Таким чином, можна стверджувати, що нейронна мережа засвоїла закономірності моделюється предметної області і тепер її можна використовувати для проведення обчислювальних експериментів.

Отримана модель була завантажена в додаток, де підключені відповідні бібліотеки, необхідні для роботи нейронної мережі (Core ML і Vision). Додатково, в самому додатку, можна використовувати елементи експертної системи у вигляді діалогу з користувачем, коли система буде уточнювати і задавати прямі питання про те, який об'єкт може перед нею перебувати. Наприклад, деякі будівлі або скульптури можуть перебувати в одному місці, при цьому бути дуже схожими між собою. У такому випадку можна задати користувачеві деякі запитання про розташування щодо інших об'єктів, розмір, тип, тощо.

ВИСНОВКИ

Таким чином, у роботі було розглянуто можливості використання нейромережевого підходу у задачах пошуку культурних об'єктів за допомогою смартфона. Були використані шляхи покращення вхідних даних, та їх розбиття за територіальною ознакою, що добре позначилось на якості розпізнавання у тестовому додатку. У майбутньому планується зібрати більше матеріалів про об'єкти для розширення і створення повноцінного туристичного гіда.

Список використаних джерел:

1. Scott, M. Create Your Own Website / M. Scott. – Ed. 4. – Mode of access: <https://www.amazon.com/Create-Your-Own-Website-4th/dp/0672330024>
2. Using Apple Machine Learning Algorithms to Detect and Subclassify Non-Small Cell Lung Cancer / A. A. Borkowski [et al.]. – Mode of access: <https://arxiv.org/ftp/arxiv/papers/1808/1808.08230.pdf>
3. Ліпанов, О. В. Аналіз методів розпізнавання об'єктів в системах аналізу візуальної інформації / О. В. Ліпанов, М. В. Фесенко // Системи обробки інформації. – 2012. – № 3. – С. 78–82.
4. Тулашвілі, Ю. Інтегрована Інформаційна Система для Підтримки Smart-Туризму / Ю. Тулашвілі, Т. Жовтан // Інформаційні технології та комп'ютерне моделювання : матеріали міжнар. науково-практ. конф. (Івано-Франківськ, 14–19 травня 2018 р.). – Івано-Франківськ, 2018. – С. 69–72.
5. Meredydd Williams ToARist: An Augmented Reality Tourism App created through User-Centred Design / Meredydd Williams, Kelvin K. K. Yao, Jason R. C. Nurse // Proceedings of the 31st International BCS Human Computer Interaction Conference (HCI 2017) (HCI) (3–6 July 2017 y.). – Mode of access: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEWjY66OgzIPmAhWqpYsKHR3dC9gQFjABegQIARAB&url=https%3A%2F%2Fwww.scienceopen.com%2Fhosted-document%3Fdoi%3D10.14236%2Ffewic%2FHCI2017.1&usg=AOvVaw2r1K19U0--9u4XZ52O2mLS>
6. Hollemans, M. I. Core ML Survival Guide / M. I. Hollemans. – Mode of access: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEWjt6pm3z4PmAhUVi8MKHWDUAIoQFjABegQIBBAC&url=http%3A%2F%2Fsamples.leanpub.com%2Fcoreml-survival-guide-sample.pdf&usg=AOvVaw1qrbs6nXzcLiLnf66F7Otm>
7. Басюк, Т. Аналіз та класифікація основних методів розпізнавання образів на площині проєкції / Т. Басюк, Я. Пушко // Вісник Нац. ун-ту «Львівська політехніка»: зб. наук. праць. – Львів, 2015. – № 826. – С. 291–298. – (Серія: Комп'ютерні науки та інформаційні технології). – Режим доступу: <http://ena.lp.edu.ua:8080/handle/ntb/31327>
8. Черепанов Ф. М. Нейросетевой фильтр для исключения выбросов в статистической информации / Ф. М. Черепанов, Л. Н. Ясницкий // Вестник Пермского ун-та. – 2018. – № 4. – С. 151–155. – (Серия : Математика. Механика. Информатика)

Федосова И. В., Разумный А. В.

РАЗРАБОТКА ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «TOARIST»

Данная статья посвящена вопросу внедрения глубоких свёрточных нейронных сетей в туристические приложения. Большинство существующих решений или устарело, или не рассматривает все потенциальные аспекты усовершенствования качества работы нейросети по тем критериям, что предъявлены данной работе. В статье рассматриваются пути совершенствования распознавания культурных объектов с потокового изображения камеры смартфона путем использования фреймворка Core ML. Благодаря тому, что он основан на библиотеках BNNS и MPSCNN это позволяет ему оптимально использовать имеющиеся ресурсы устройства, ведь он может запускать нейромодель на CPU, GPU или на Neural Engine (нейронный процессор), в зависимости от того, какое аппаратное оборудование доступно на устройстве пользователя.

Однако, по причине того, что исходный код Core ML является закрытым, были проанализированы и предложены внешние пути его совершенствования. Среди них: препроцессинг, фильтрация, использование систем глобального позиционирования GPS и элементов экспертных систем. Большинство исходных материалов для обучения и создания модели нейросети были собраны вручную, путем фотографирования культурных памятников, а также поиска в сети Интернет. Фотографии были классифицированы по группам в зависимости от местоположения объекта. Это помогло разбить исходную модель на несколько моделей меньшего размера и увеличить конечную точность обработки.

В процессе создания модели и тестирования модели были обнаружены изображения, которые увеличивают ошибку наблюдения для нейросети. Данные изображения изымались из входного множества, а затем система обучалась, получая обобщенные свойства исследуемых объектов. Модели, точность которых оценивалась фреймворком Create ML на 90-100 % загружались в тестовое приложение. Для него был предложен способ внедрения экспертной системы для повышения точности распознавания путём ведения наводящего диалога с пользователем.

В результате работы был получен работоспособный прототип мобильного приложения под названием «ToARist», который может быть улучшен путём внесения информации о новых культурных объектах.

Ключевые слова: *распознавание культурных объектов, сверточная нейросеть, мобильное приложение.*

Fedosova I. V., Razumnyi O. V.

DEVELOPMENT OF THE TOARIST HYBRID INTELLECTUAL SYSTEM «TOARIST»

This article is devoted to the issue of introducing deep convolutional neural networks into tourism applications. Most of the existing solutions are outdated or do not consider all potential aspects of improving the quality of the neural network according to the criteria presented to this work. The article discusses ways to improve quality of the recognition of cultural objects from the streaming image of the smartphone camera by using the Core ML framework. Proceeding from the fact that it is based on BNNS and MPSCNN libraries, this allows it to optimally use the available resources of the device, because it can run a model on a CPU, GPU or on a Neural Engine, depending on what hardware is available on the user's device.

However, due to the fact that the source code of Core ML is closed, external ways to improve it were analyzed and proposed. Among them: preprocessing, filtering, using of GPS global positioning systems and elements of expert systems. Most of the source materials for training and creating a neural network model were manually collected by photographing cultural sites, as well as searching on the Internet. Photos were classified into groups depending on the location of the object. This helped to split the original model into several smaller ones and increase the final machining accuracy.

In the process of creating the model and testing the model, images were discovered that increase the observation error for the neural network. For this reason, image data was removed from the input set, and then the system was trained, obtaining generalized properties of the objects under study. Models which accuracy was evaluated by the Create ML framework for 90-100 % were loaded into the test application. A method of introducing an expert system was proposed for it to increase recognition accuracy by conducting a suggestive dialogue with the user.

As a result of the work, a prototype of a mobile application called «ToARist» was obtained, and can be improved by adding information of new cultural sites.

Keywords: *cultural object recognition, convolutional neural network, mobile application.*

Рецензент:

Стаття надійшла 11.11.2018 р.