

АРХІТЕКТУРА ТА ПРИНЦИП ДІЇ МОДЕЛІ ГЛИБОКОГО НАВЧАННЯ ДЛЯ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ У BLOSSOM LENS

Малярчук Богдана

ARCHITECTURE AND OPERATING PRINCIPLE OF THE DEEP LEARNING MODEL FOR IMAGE RECOGNITION IN BLOSSOM LENS

Maliarchuk Bohdana

Волинський національний університет імені Лесі Українки, просп. Волі, 13, Луцьк, 43025, Україна

Abstract. Computer vision systems based on deep learning enable a high degree of automation in image classification tasks. Blossom Lens is an example of such implementation, designed for automatic flower recognition from digital images. The model achieves over 90% accuracy and integrates external knowledge sources, transforming the software into an educational tool..

Системи комп'ютерного зору, побудовані на основі глибокого навчання, забезпечують високий рівень автоматизації у завданнях класифікації зображень, виявлення об'єктів, сегментації тощо. Одним із прикладів ефективного впровадження цієї технології є програмне забезпечення Blossom Lens, розроблене з метою автоматичного розпізнавання квітів за цифровими зображеннями. У цій роботі докладно розглянуто архітектуру, алгоритми та етапи обробки зображення, які забезпечують точну ідентифікацію рослин.

На вхід системи подається цифрове зображення у форматі JPEG або PNG, яке представляється як тривимірний масив пікселів. Кожен піксель кодується трьома значеннями відповідно до інтенсивності червоного, зеленого та синього каналів (RGB). Для забезпечення коректної обробки зображення попередньо масштабується до розміру 224×224 пікселі. Наступним етапом є нормалізація піксельних значень — всі значення зводяться до діапазону [0; 1], що забезпечує стабільну роботу моделі під час навчання та класифікації, зменшуючи розмах чисел та ризик перенасичення активацій у шарах нейронної мережі.

Основу системи складає згортоква нейронна мережа (Convolutional Neural Network, CNN), яка реалізує поетапне витягнення візуальних ознак зображення. Перші шари мережі здійснюють виділення низькорівневих характеристик — контурів, меж, напрямків. Це реалізується шляхом згорткових операцій з фільтрами розміром 3×3 або 5×5, які рухаються по зображенню і створюють карти ознак (feature maps), що реагують на локальні патерни.

У середніх шарах нейронна мережа узагальнює локальні ознаки та формує семантичні структури: симетрію, форму пелюсток, характерну текстуру чи орнамент квітки. Саме ці шари дозволяють моделі навчитися диференціювати подібні за формою види, наприклад, хризантему та айстру. Глибші шари узагальнюють усе побачене і створюють високорівневий вектор ознак — стислий числовий опис зображення, що кодує суттєві особливості об'єкта на фото. Цей вектор передається на вхід повнозв'язного шару класифікатора.

Фінальний шар мережі — це шар Softmax, який генерує ймовірнісний розподіл серед усіх визначених класів. У проєкті Blossom Lens модель налаштована для розпізнавання багатьох квітів, серед яких троянда, ромашка, тюльпан, тощо. Клас із найбільшою ймовірністю вважається передбаченим результатом. Разом із назвою класу виводиться рівень впевненості — значення від 0 до 1, яке дозволяє користувачу оцінити достовірність класифікації.

Навчання згортоква нейронної мережі здійснювалось на спеціально сформованому наборі зображень, зібраному з відкритих джерел, з поділом на тренувальну (70%), валідаційну

(20%) і тестову (10%) вибірки. Для запобігання перенавчанню і покращення узагальнюючої здатності моделі застосовувались методи аугментації — штучне розширення даних шляхом випадкових обертань, дзеркального відображення, зміни яскравості та контрасту. Такі перетворення дозволяють моделі бачити різні варіації того самого об'єкта, що значно підвищує її стійкість до зовнішніх впливів, таких як освітлення чи зміна фону.

Процес навчання включав використання функції втрат, що є стандартною для багатокласових задач класифікації. Як оптимізатор було обрано алгоритм Adam з початковим коефіцієнтом навчання 0.0001. Для контролю за точністю використовувались показники асигасу на валідаційній вибірці. Навчання тривало до стабілізації втрат і досягнення максимального значення точності без ознак перенавчання. Результати тестування показали точність понад 90% при класифікації нових зображень, не включених до навчального набору.

Після отримання результату класифікації програма Blossom Lens автоматично звертається до відкритого API Вікіпедії, формує запит на основі передбаченої наукової назви квітки, і отримує короткий опис, який виводиться у графічному інтерфейсі. Інтерфейс програми реалізовано з використанням бібліотеки Tkinter, яка забезпечує доступ до основного функціоналу: вибір зображення, перегляд результатів класифікації, збереження результатів, ведення історії класифікацій, повторне відкриття попередніх ідентифікацій.

У підсумку, Blossom Lens є прикладом повноцінної реалізації архітектури глибокої нейронної мережі для вирішення прикладної задачі класифікації зображень у сфері ботаніки. Завдяки використанню згорткових шарів, методів аугментації, оптимізованих гіперпараметрів і розширеної інтеграції з відкритими джерелами знань, програма забезпечує високу точність розпізнавання та зручність для кінцевого користувача. Результати доводять ефективність обраного підходу та свідчать про високий потенціал подальшого розвитку, зокрема — масштабування на мобільні платформи, додавання нових класів об'єктів і впровадження моделі в інші галузі, пов'язані з аналізом зображень.

Бібліографія

1. TensorFlow. Платформа TensorFlow [Електронний ресурс] // Офіційний сайт TensorFlow. – Режим доступу: <https://www.tensorflow.org/> (дата звернення: 12.05.2025).
2. Keras. Офіційна документація [Електронний ресурс]. – Режим доступу: <https://keras.io/> (дата звернення: 10.05.2025).
3. Учасники проєктів Вікімедіа. Згорткова нейронна мережа [Електронний ресурс] // Wikipedia. – Режим доступу: https://en.wikipedia.org/wiki/Convolutional_neural_network (дата звернення: 11.05.2025).
4. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Електронний ресурс] // arXiv.org. – Режим доступу: <https://arxiv.org/abs/1409.1556> (дата звернення: 08.05.2025).
5. Учасники проєктів Вікімедіа. Аугментація даних [Електронний ресурс] // Wikipedia. – Режим доступу: https://en.wikipedia.org/wiki/Data_augmentation (дата звернення: 09.05.2025).
6. OpenCV. Бібліотека комп'ютерного зору з відкритим кодом [Електронний ресурс]. – Режим доступу: <https://opencv.org/> (дата звернення: 30.04.2025).
7. MediaWiki. Wikipedia API – Головна сторінка [Електронний ресурс] // Розробники MediaWiki. – Режим доступу: https://www.mediawiki.org/wiki/API:Main_page (дата звернення: 29.04.2025).
8. Учасники проєктів Вікімедіа. Softmax-функція [Електронний ресурс] // Wikipedia. – Режим доступу: https://en.wikipedia.org/wiki/Softmax_function (дата звернення: 28.04.2025).
9. Python Software Foundation. Tkinter – офіційна документація [Електронний ресурс] // Python.org. – Режим доступу: <https://docs.python.org/3/library/tkinter.html> (дата звернення: 12.05.2025).
10. DeepAI. Що таке згорткова нейронна мережа (CNN)? [Електронний ресурс] // DeepAI. – Режим доступу: <https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network> (дата звернення: 11.05.2025).