DOI: 10.58168/MoInSyTe2024_240-245
УДК 004.93'1

IMAGE RECOGNITION BY CONVOLUTIONAL NEURAL NETWORKS

Ahamed Sakib[1], E.A. Anikeev[1]

[1]Voronezh State University of Forestry and Technologies named after G.F. Morozov

Abstract. The capabilities of image recognition technology show how far computers have come. They are having a major impact on many areas such as medical scanning and self-driving cars. Python plays a central role in the development of these technologies because it is easy to use and has many tools, such as TensorFlow and PyTorch, that make complex machine learning models easy to understand. This article details how to create an image recognition system based on Python. It covers every step from collecting datasets and preparing them to selecting neural network models.

Keywords: image recognition, artificial intelligence, COCO data set, neural network, Python

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Ахамед Сакиб[1], Е.А. Аникеев[1]

[1]ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

Аннотация. Возможности технологии распознавания изображений показывают, как далеко продвинулись компьютеры. Они оказывают серьезное влияние на многие области, такие как медицинское сканирование и самоуправляемые автомобили. Python играет центральную роль в разработке этих технологий, поскольку он прост в использовании и имеет множество инструментов, таких как TensorFlow и PyTorch, которые упрощают понимание сложных моделей машинного обучения. В этой статье подробно рассказывается о том, как создать систему распознавания изображений на основе Python. Он охватывает каждый этап - от сбора наборов данных и их подготовки до выбора моделей нейронных сетей.

Ключевые слова: распознавание изображений, искусственный интеллект, набор данных COCO, нейронная сеть, Python

The emergence of digital technology has facilitated a fast and significant increase in the quantity and diversity of digital pictures, hence offering opportunities for the development of complex image recognition algorithms. Artificial intelligence (AI) relies heavily on image recognition, which is the process of identifying and categorizing various features – such as people, objects, symbols, and actions – that may be seen in pictures. There are many different uses for image recognition in many domains. These include applying facial recognition technology to improve security protocols, using medical imaging analysis to diagnose patients more accurately, and incorporating augmented reality technology to improve customer experiences in the retail and entertainment industries. Python's extensive collection of libraries and frameworks, such as TensorFlow, Keras, and PyTorch, has contributed to its growing appeal as a programming language for creating image recognition models. These technologies provide a streamlined depiction of complex algorithms and mathematical procedures, facilitating the creation of sophisticated models with ease. This article offers a comprehensive examination of the procedure involved in the development of an image recognition system utilizing the Python programming language. It encompasses the entire process, from initiation to completion, with a specific focus on practical application and theoretical understanding.

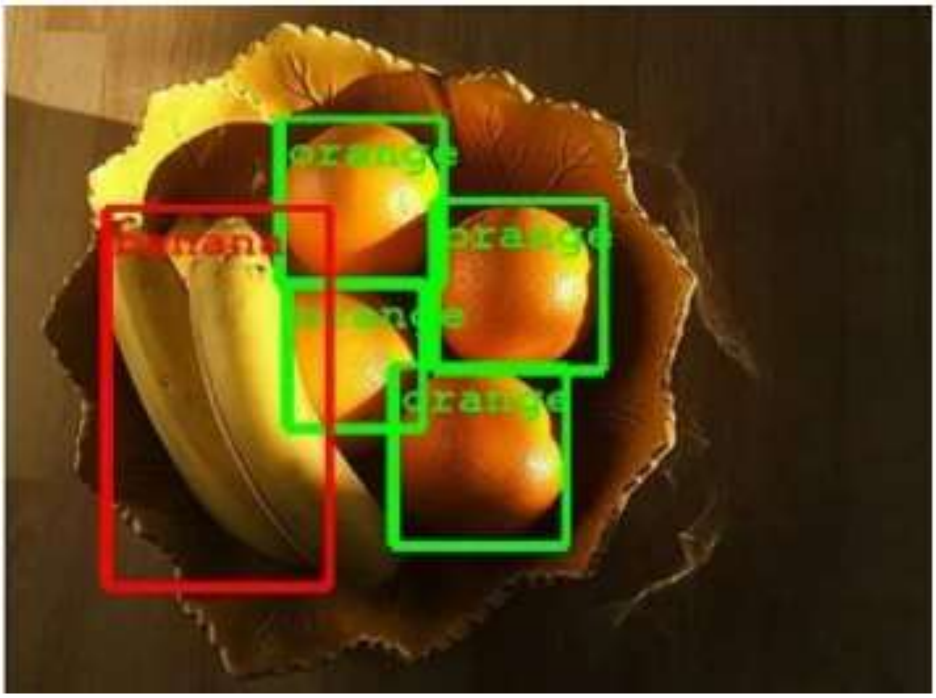

Figure 1 – Image Recognition

Data Collection

It is crucial to have an extensive and all-inclusive dataset. The dataset that is chosen depends on the specific requirements of the image recognition analysis that is being carried out. There are two widely used datasets in the field of image identification, each of which is particularly designed to tackle different goals and complexities. The ImageNet dataset comprises a large collection of over 14 million images that have been accurately labeled, encompassing a diverse array of more than 20,000 distinct categories. The arrangement of the collection is structured according to the WordNet hierarchy, which classifies photographs into many sets of synonyms, ensuring a significant degree of comprehensiveness. ImageNet is particularly beneficial for training and testing deep neural networks in general object recognition and classification applications.

The COCO dataset is a crucial tool for image recognition applications, such as those that deal with segmentation, object identification, and captioning. The dataset consists of an extensive set of 330,000 photos, of which over 200,000 have been carefully annotated in 80 major item categories. Moreover, the collection includes annotations for more than a million data points that correspond to different things. Focusing on item instances in their natural environments, COCO distinguishes itself by offering important contextual information required to build models that comprehend the spatial connections between things in an image.

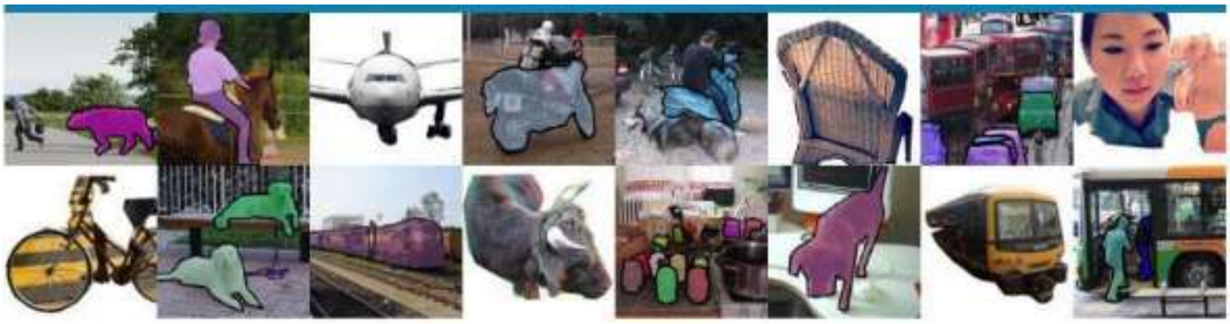

Figure 2 – COCO Dataset

It could be essential in some particular applications to compile a customized dataset. To obtain a thorough representation that includes a wide variety of changes in light, perspective, backdrop, and occlusions, the approach involves establishing clear criteria for image inclusion. Methods for creating a custom dataset include the following components. Web Scraping: Web scraping is the process of automating the retrieval of photos from websites using scripts. Python Beautiful Soup and Selenium packages provide the ability to crawl web pages and retrieve images that meet predefined criteria.

Convolutional Neural Networks (CNNs) can hierarchically handle visual input, they are essential for most photo identification applications. However, the choice of architecture (e.g., AlexNet, ResNet, VGG, etc.) depends on the computing resources available and the difficulty of the task. Transfer learning might save a great deal of time and computer power by using pre-trained models on large datasets and tailoring them to particular applications.

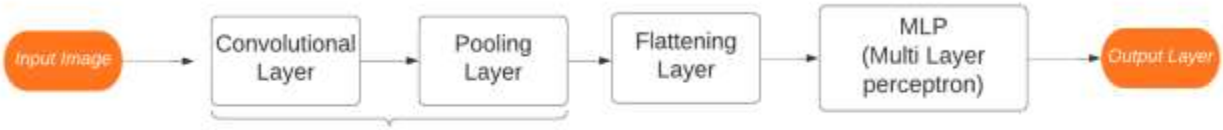

Figure 3 – Image Recogition Model Design

Design

In the design phase, the process involves the configuration of the layers and architecture of the model. Important factors to pay attention to are listed below. Network depth: increasing the number of layers can effectively capture intricate patterns, but it also raises the likelihood of overfitting and computing requirements. Activation functions, such as ReLU (Rectified Linear Unit), provide nonlinearity to the model, allowing it to learn intricate patterns. Pooling layers are a technique used to decrease the complexity and computing burden of a model, while still preserving significant characteristics. Various regularization approaches, like as dropout, L2 regularization, and batch normalization, are employed to mitigate the issue of overfitting and enhance the generalizability of the model.

Implementation

Creating an environment for Python development involves choosing an IDE (Integrated Development Environment) or notebook interface (Jupyter), and getting necessary libraries (PyTorch, TensorFlow, NumPy, and Matplotlib) through package managers (pip or conda). In the realm of image preparation activities, libraries such as OpenCV and PIL (Python Imaging Library) play a crucial role. The implementation of code snippets for resizing, normalization, and augmentation can automate the pretreatment workflow, hence enhancing the efficiency of dataset preparation for training. The architectural design of the model may be established by employing libraries like as TensorFlow or PyTorch. The training procedure starts by compiling the model, which incorporates an optimizer, loss function, and measurements.

As we can see, the Python programming language has wide application possibilities in image recognition, and contributes to the development of this area of artificial intelligence.

References

1. Dilmegani, C. Image Recognition: In-depth Guide for 2024? – URL: https://research.aimultiple.com/image-recognition/ (date of the application: 21.03.2024).

2. Glover, E. What Is Image Recognition? – URL: https://builtin.com/artificial-intelligence/image-recognition (date of the application: 16.03.2024).

3. Javaid, S. Image Data Collection in 2024: What it is and Best Practices? – URL: https://research.aimultiple.com/image-data-collection/ (date of the application: 20.03.2024).

4. Zhaxybayev D.O., Bakiyev M.N. Algorithms for the classification of text documents, taking into account proximity in the attribute space // Modeling of systems and processes. – 2022. – T. 15, № 1. – P. 36-43.

5. Poluektov, A.V. Modeling of oscillatory processes in the MVSTUDIUM package / A.V. Poluektov, K.V. Zolnikov, V.I. Antsiferova // Modeling of systems and processes. – 2021. – T. 14, No. 4. – P. 139-148. – DOI: 10.12737/2219-0767-2021-14-4-139-148.

6. Sazonova, S.A. Features of developing software products using arrays in an object-oriented environment / S.A. Sazonova, A.V. Lemeshkin, V.A. Popov // Modeling of systems and processes. – 2021. – T. 14, No. 4. – P. 90-100. – DOI: 10.12737/2219-0767-2021-14-4-90-100.

Список литературы

1. Дилмегани С. Распознавание изображений: подробное руководство на 2024 год? – URL: https://research.aimultiple.com/image-recognition/ (дата подачи заявки: 21.03.2024).

2. Гловер Э. Что такое распознавание изображений? – URL: https://builtin.com/artificial-intelligence/image-recognition (дата подачи заявки: 16.03.2024).

3. Джавайд С. Сбор данных об изображениях в 2024 году: что это такое и лучшие практики? – URL: https://research.aimultiple.com/image-data-collection/ (дата подачи заявки: 20.03.2024).

4. Жаксыбаев Д.О., Бакиев М.Н. Алгоритмы классификации текстовых документов с учетом близости в признаковом пространстве // Моделирова-ние систем и процессов. – 2022. – Т. 15, № 1. – С. 36-43.

5. Полуэктов, А. В. Моделирование колебательных процессов в пакете MVSTUDIUM / А.В. Полуэктов, К.В. Зольников, В.И. Анциферова // Моделирование систем и процессов. – 2021. – Т. 14, № 4. – С. 139-148. – DOI: 10.12737/2219-0767-2021-14-4-139-148.

6. Сазонова, С.А. Особенности разработки программных продуктов с использованием массивов в объектно-ориентированной среде / С.А. Сазонова, А.В. Лемешкин, В.А. Попов // Моделирование систем и процессов. – 2021. – Т. 14, № 4. – С. 90-100. – DOI: 10.12737/2219-0767-2021-14-4-90-100.