

УДК 004.421.6

## ФРЕЙМВОРКИ ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ В РЕЖИМЕ ФОТОГРАФИИ



**В.С. Маркевич**

Студент ГГУ имени Франциска Скорины



**Л.Н. Марченко**

Заведующий кафедрой фундаментальной и  
прикладной математики факультета  
математики и технологий  
программирования

Учреждение образования «Гомельский государственный университет имени Франциска Скорины»,  
Республика Беларусь  
E-mail: [vladismark@gmail.com](mailto:vladismark@gmail.com), [lmarchenko@gsu.by](mailto:lmarchenko@gsu.by)

### **В.С. Маркевич**

Студент 3-го курса специальности «Экономическая кибернетика» факультета математики и технологий программирования

### **Л.Н. Марченко**

Заведующий кафедрой фундаментальной и прикладной математики факультета математики и технологий программирования, кандидат технических наук, доцент,

**Аннотация.** Нейронные сети открыли новые возможности в области распознавания образов. В связи с этим появляется возможность разработки новых методов и алгоритмов для анализа объектов, в частности, фотографий. В работе разработан алгоритм и приложение для распознавания объектов, в котором пользователи могут загружать фотографии и получать заданные готовые изображения с выделенными на нем объектами. Здесь использовался ряд фреймворков языка программирования Python для распознавания образов на фотографии, что в данный момент актуально.

**Ключевые слова:** нейронные сети, фреймворки, распознавание образов, режим фотографии, библиотеки языка программирования Python, искусственный интеллект.

**Введение.** Нейронные сети широко используются для распознавания образов. Существует множество алгоритмов, решающих конкретные задачи в этой сфере. Для распознавания объектов на фотографии обычно используются библиотеки NumPy, SciPy, H5Py, ImageAI, которые используют компьютерное зрение и облегчают создание алгоритмов.

Работа посвящена разработке приложения для распознавания объектов в режиме фотографии. Достижение данной цели предопределило создание алгоритма и кода на языке программирования Python для распознавания объектов, в котором пользователи смогут загружать фотографии и получать заданные готовые изображения с выделенными на нем объектами с использованием библиотек Numpy, SciPy, Pillow, Matplotlib, H5Py, Keras, ImageAI, Tensorflow, OpenCV, которые обладают более мощными возможностями по сравнению с их предшественниками. Приложение должно содержать: подключенные библиотеки и главную функцию их реализации.

*Материалы и методы.* Рассмотрим инструменты, технологии и другие решения, которые позволяют оптимальным образом удовлетворить описанные ранее требования обычно используются при решении таких задач [2].

Приложение для распознавания объектов в режиме фотографии реализовано на языке программирования Python, который широко применяется для разработки алгоритмов нейронных сетей. Для управления программными пакетами использовался менеджер пакетов `pip` [1]. Система управления пакетами – это набор программного обеспечения, который позволяет управлять процессом удаления, установки, настройки и обновления различных компонентов программного обеспечения. Программное обеспечение представлено в виде специальных пакетов, содержащих дистрибутив программного обеспечения, набор определённых метаданных, которые могут включать в себя имя, номер версии, описание пакета, имя разработчика, контрольную сумму, отношения с другими пакетами. Метаданные сохраняются в системной базе данных пакетов. Большинство дистрибутивов Python уже содержат `pip` [3]. Одно из главных преимуществ `pip` – это простота интерфейса командной строки, которая позволяет установить пакеты Python простой командой: `pip install some-package-name`, также легко удалять пакеты: `pip uninstall some-package-name`. `Pip` предоставляет возможность управлять всеми пакетами и их версиями с помощью файла `requirements.txt`.

`NumPy` – библиотека языка Python, предоставляющая поддержку больших многомерных массивов и матриц, и дающая большой спектр высокоуровневых математических функций. В результате любой алгоритм, который может быть выражен в виде последовательности операций над массивами и реализованный с использованием `NumPy`, работает достаточно эффективно. `SciPy` – это библиотека для языка программирования Python, созданная для выполнения научных и инженерных расчетов, которая решает задачи поиска экстремумов функций, вычисления интегралов функций, поддержку специальных функций, обработку сигналов, обработка изображений, работу с генетическими алгоритмами, решение обыкновенных дифференциальных уравнений и др [3]. Библиотека `SciPy` может взаимодействовать с `PyTables` – иерархической базой данных, разработанной для управления большими объемами данных.

Для визуализации результатов расчётов применялась библиотека `Python Imaging Library`, `Matplotlib`, `HippoDraw`, `Chaco`, `Biggles`, `MayaVi` [4].

Библиотека `PIL` (`Python Imaging Library`) предназначена для работы с растровой графикой. В разработанном проекте использовался форк `Pillow`. Возможности библиотеки – это поддержка бинарных, индексированных изображений; разных форматов, таких как: `BMP`, `ESP`, `GIF`, `JPEG`, `PDF`, `PNG`, `TIFF`; преобразование изображений одного формата в другой [5].

Библиотека `MatPlotLib` представляет собой пакет, предназначенный для визуализации данных, который вместе с `NumPy` и `SciPy` представляет возможности для решения задач технических вычислений. Данный пакет поддерживает многие виды графиков и диаграмм. Также с помощью `MatPlotLib` можно создавать анимированные изображения. Набор форматов изображений такой же, как и у библиотеки `PIL`.

`Keras` – это нейросетевая библиотека, предоставляющая настройку над фреймворком `TensorFlow` и направленная на работу с сетями глубокого обучения. `Keras` предоставляет высокоуровневый набор абстракций, благодаря которому можно проще формировать нейронные сети. Эта библиотека содержит многочисленные реализации блоков нейронных сетей.

`TensorFlow` – это программная библиотека для машинного обучения. Она создана для решения задач и построения тренировки нейронной сети с целью автоматического нахождения образов. `TensorFlow` является системой машинного обучения второго поколения. Данная библиотека работает на параллельных процессорах, как `GPU`, так и `CPU`. Вычисления происходят в виде потоков данных через граф состояний. `TensorFlow` хорошо подходит для автоматизированной аннотации изображений [7].

TensorFlow – это продолжение закрытой библиотеки DistBelief. Закрытая система машинного обучения DistBelief разрабатывалась Google Brain для внутренних проектов для работы с нейронными сетями и использовалась во многих проектах группы фирм холдинга Alphabet. После упрощения и оптимизации коды библиотеки увеличили надёжность и удобство и представляют собой TensorFlow [8].

Библиотека ImageAI – это библиотека Python, созданная для создания приложений и систем с автономными возможностями глубокого обучения и компьютерного зрения [9-11].

Библиотека OpenCV – это библиотека алгоритмов компьютерного зрения общего назначения, то есть – это набор данных, классов, функций для обработки изображения. Основные модули библиотеки: ядро, которое содержит основные данные и алгоритмы, базовые операции, матричную алгебру и многое другое; Highgui – модуль для ввода-вывода изображений; Svaux – экспериментальные функции для пространственного зрения; CV – модуль обработки изображения и компьютерного зрения, содержащий множество алгоритмов обработки изображений [12].

#### *Проектирование и реализация приложения «Распознавание объектов».*

Для работы нейронных сетей были подключены рассмотренные библиотеки для реализации алгоритмов распознавания объектов. Для правильной работы приложения алгоритм загружен в корень программы с помощью библиотеки ImageAI достать данный алгоритм и поместить его внутрь (рисунок 1).

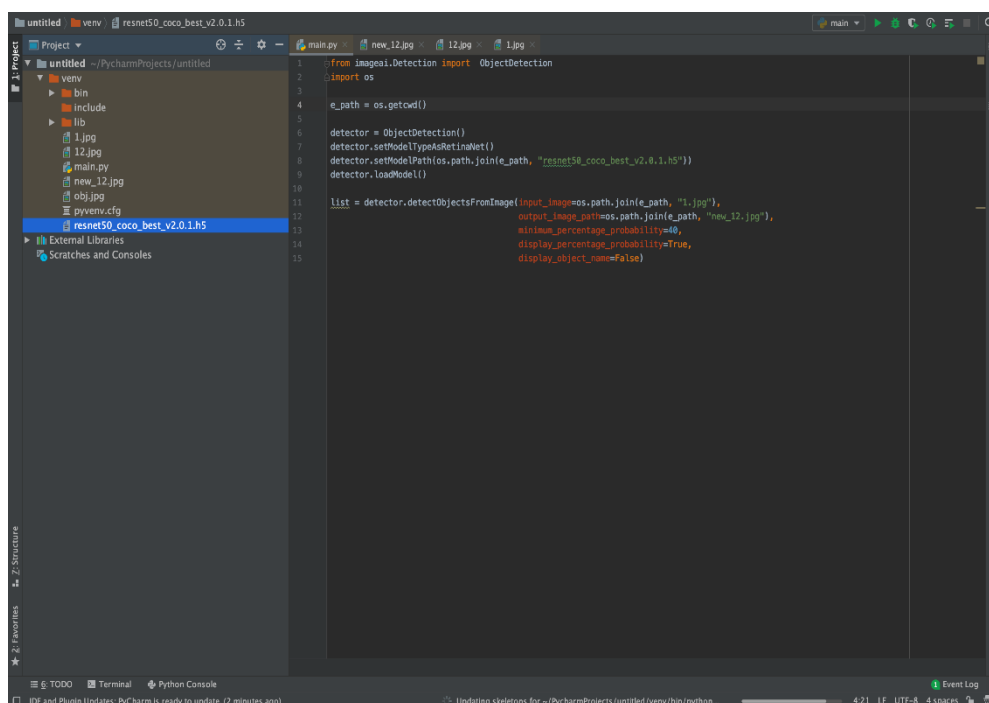


Рисунок 1. – Размещение алгоритма с помощью библиотеки ImageAI

Следует отметить, что алгоритм, который называется resnet50 не способен читаться в текстовых редакторах, он читается на бинарном уровне. Данный алгоритм несет в себе:

- разрезание фотографии на слои через матрицы;
- разбивание точек матрицы на кластеры;
- алгоритм проводит множество прямых через кластеры, тем самым придавая очертание образам, которые желает видеть программист;
- добавление весовых коэффициентов для более точного анализа данных;
- создание функциональных пирамид.

Разрезание фотографий состоит из 106-ти свёрточных слоев и лучше детектирует небольшие объекты. Основная особенность состоит в том, что на выходе есть три слоя, каждый из которых рассчитан на обнаружение объектов разного размера. На рисунке 2 приведен пример такого схематического устройства:

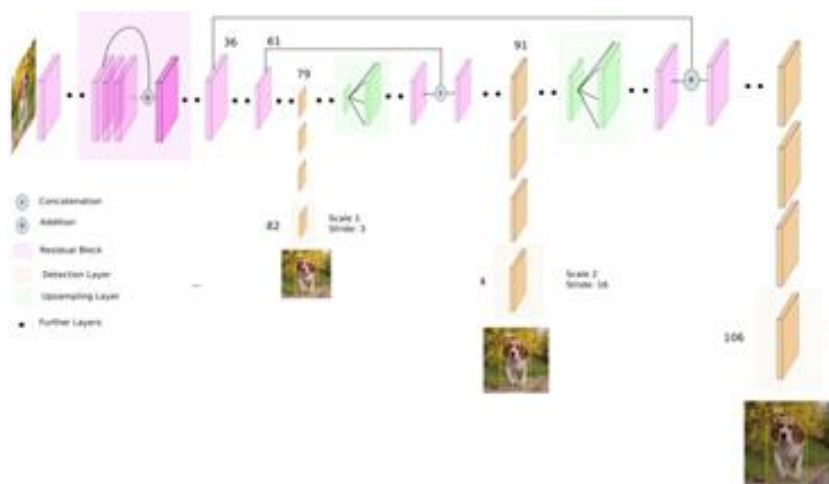


Рисунок 2. – Разрезание фотографий

Рассмотрим этапы работы программы. Загружаем фотографии в корень разработанной программы, называем их «12», «obj» и «1» (рисунок 3).

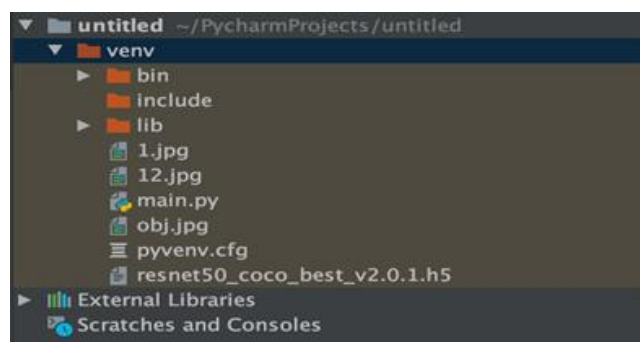


Рисунок 3. – Корень разработанной программы

Все подключенные библиотеки находятся в пакете lib. На рисунке 3 представлены описанные выше библиотеки, которые были задействованы при создании программы.

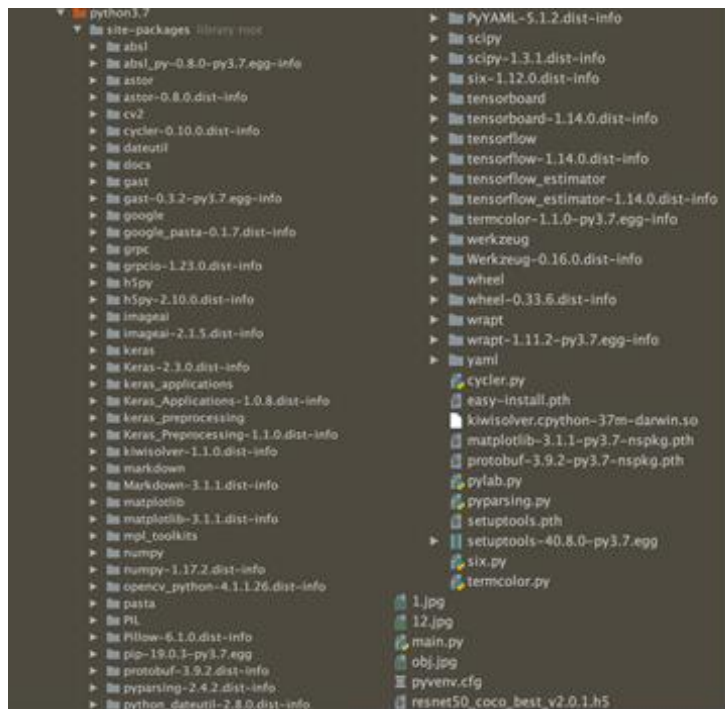


Рисунок 4. – Библиотеки, задействованные при создании программы

Главный класс называется main, который содержит главные методы, загружающие и обрабатывающие фотографии (рисунок 4. )

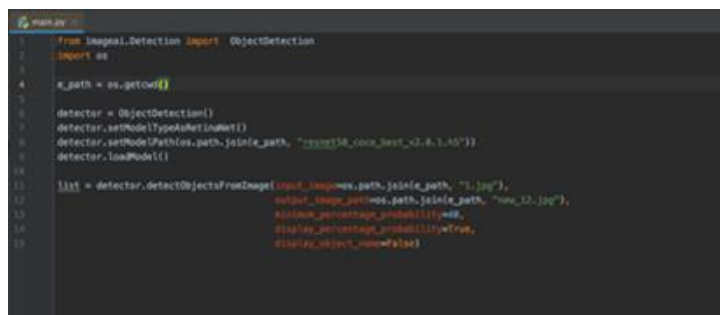


Рисунок 5. – Главный класс main

Когда приложение начинает обработку фотографии, то вначале оно включает все библиотеки, после загружает фотографию и алгоритм начинает «пробегать» по всем слоям (рисунок 5). В результате создается фотография, которая называется «new\_12».

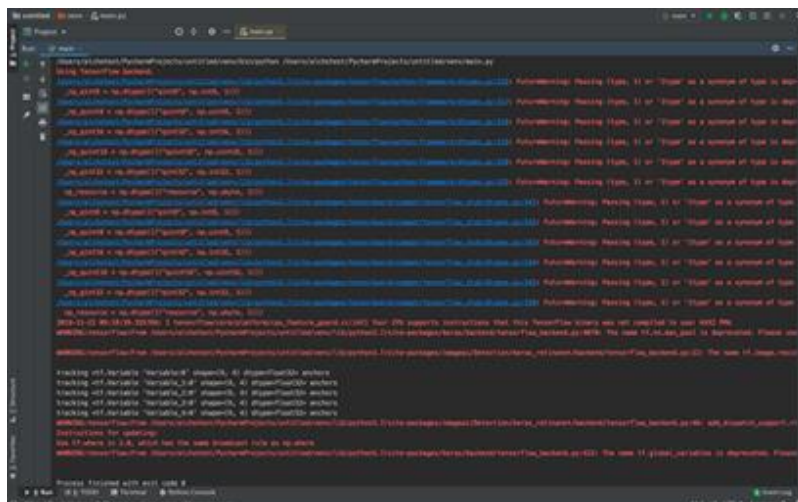


Рисунок 6. – Алгоритм работы программы

Для верификации приложения использовались личные фотографии. На рисунке 6 представлены фотографии до обработки и после обработки: Вероятность распознавания после обучения нейронной сети составила 60 %.



Рисунок 7. – Фотографии до обработки и после

**Заключение.** Распознавание объектов на основе нейронных сетей используется в различных сферах деятельности. Приложение распознавания объектов в режиме фотографии реализовано с помощью библиотек языка программирования Python: Numpy, SciPy, Pillow, Matplotlib, N5Py, Keras, ImageAI, Tensorflow, OpenCV. Приложение удобно и просто в использовании.

#### Список литературы

- [1.] Мюллер, А. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными. Пер с англ. / А. Мюллер, С. Гвидо. – СПб.: ООО «Альфа-книга», 2017. – 480 с.
- [2.] Вандер Плас, Дж. Питон для сложных задач. Наука о данных и машинное обучение / Дж. Вандер Плас. – СПб.: Питер, 2019. – 576 с.
- [3.] Нуньес-Иглесиас, Х. Элегантный SciPy. Научное программирование на Python / Х. Нуньес-Иглесиас, Ш. Уолт, Х. Дэшноу. – М.: ДМК Пресс, 2018. – 266 с.
- [4.] Blanco-Silva, F.J. Learning SciPy for Numerical and Scientific Computing. – Packt Publishing, 2013. – 150 p.
- [5.] Ninad Sathaye. Python Multimedia. – Packt Publishing, 2010. – 292 p.
- [6.] Blanco-Silva, F.J. Mastering SciPy. – Packt Publishing, 2015. – 404 p.
- [7.] Джулли, А. Библиотека Keras / А. Джулли, С. Пал. – М.: ДМК Пресс, 2017. – 294 с.

[8.] Жерон, О. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. Концепции, инструменты и техники для создания интеллектуальных систем / О. Жерон. – М.: Вильямс, 2018. – 688 с.

[9.] Гудфеллоу, Я. Глубокое обучение. Пер с англ. А.А. Слинкина / Я. Гудфеллоу, И. Бенджио, А. Курвилль // 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил.

[10.] Шапиро, Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман. – М., 2013. – 752 с.

[11.] Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. – М.: Вильямс, 2004. – 928 с.

[12.] Bradski, G., Kaehler, A. Learning OpenCV 3: Computer Vision with the OpenCV library. - Packt Publishing: OReilly Media, 2008. – 580 p.

[13.] A - Markevich, V. Application development «Recognition of objects in the photo» on programming language Python. / V. Markevich // 8th Podlasie Conference on Mathematics 8PCM (5-8 December 2019, Bialystok). – Poland, 2019. – P. 100.

## **DEVELOPING THE OBJECT RECOGNITION APPLICATION IN PHOTO MODE**

*Francisk Skorina Gomel State University, Republic of Belarus*

**Abstract.** Neural networks have opened up new possibilities in the field of image recognition. In this regard, it is possible to develop new methods and algorithms for analyzing objects, in particular, photos. The work has developed an algorithm and an application for object recognition, in which users can upload photos and get specified ready-made images with objects highlighted on it. Here we used a number of Python programming language frameworks for image recognition in photos, which is currently relevant.

**Keywords:** neural networks, frameworks, pattern recognition, photography mode, Python programming language libraries, artificial intelligence.