

## СОЗДАНИЕ КАТАЛОГА ТОВАРОВ ЭЛЕКТРОННОЙ КОММЕРЦИИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

**В.В. Сорокина**

*Белорусский государственный университет  
пр. Независимости, 4, 220030, г. Минск, Беларусь  
viktoriasorokina.96@gmail.com*

В данной работе представлена технология для автоматического создания каталога товаров электронной коммерции на основе нейронных сетей. Под каталогом товаров электронной коммерции подразумевается набор обработанных фотографий товаров. Технология включает в себя этапы по определению самого товара (задача object detection) и его основного цвета, сегментации и автоматической обрезке (задача определения частей тела человека). Новизна представленной работы заключается в использовании сверточных нейронных сетей совместно с моделью внимания, которая способна концентрироваться на определенных областях входного изображения.

**Ключевые слова:** Электронная коммерция; каталог товаров; модель внимания; сверточная нейронная сеть; сегментация.

## CREATING A CATALOG OF E-COMMERCE GOODS USING NEURAL NETWORKS

**V. Sorokina**

*Belarussian State University,  
Nezavisimosti av. 4, Minsk, 220050, Belarus, Minsk, Belarus  
viktoriasorokina.96@gmail.com*

In this paper a technology for automatically creating a catalog of e-commerce products based on neural networks is presented. An e-commerce product catalog is a set of processed product photos. The technology includes steps for determining the product itself (the task of object detection) and its main color, segmentation and automatic cropping (the task of determining parts of the human body). The novelty of the presented work lies in the use of convolutional neural networks in conjunction with an attention model that is able to focus on certain areas of the input image.

**Keywords:** E-commerce; catalog of goods; attention model; convolutional neural network; segmentation.

### **Введение**

В электронной коммерции одну из ключевых ролей играет каталог товаров. Каталог товаров – это набор изображений, представляющий товар. Создание каталога – процесс подготовки фотографий от момента съемки

до представления готового набора [1]. С быстрым ростом электронной коммерции и появлением алгоритмов искусственного интеллекта, традиционные системы управления контентом уступают место автоматизированным масштабируемым системам.

При подготовке изображений одежды, например, обычно используется фотография человека в полный рост, представляющего несколько предметов одежды одновременно. Задача состоит в определении положения предмета, его сегментации, обрезки и создания Color swatch – образца цвета товара. В настоящее время процесс подготовки изображений для каталогов производится вручную.

Для автоматизации данного процесса нами разработана технология, позволяющая при помощи решения задачи object detection, сегментации, выделения частей тела человека, а также определения доминантного цвета товара осуществлять автоматическую подготовку каталога товаров электронной коммерции.

## **1. Методология исследования**

В данной работе предлагается технология для автоматической подготовки изображений товаров электронной коммерции и создания каталога. Схема технологии представлена на рисунке 1.

Разработанная на базе данной технологии система представляет собой программу с веб-интерфейсом, реализованным с помощью фреймворка Flask, на вход которой подается RGB-изображение товара/группы товаров. Далее изображение последовательно проходит через алгоритмы, описанные ниже.

В начале определяется тип объекта на изображении. За это отвечает реализованный нами алгоритм на основе нейронной сети YOLACT [2] и стандартизации весов, подробное описание которого представлено в [3].

Результатом предыдущего шага является получение метки класса «человек» / «не человек». Если это человек, то срабатывает алгоритм Smart Cropping [4], где выделяются ключевые точки человеческого тела, вычисляется позиционное соотношение между ними, после чего оно используется для обрезки исходного снимка и создания набора изображений, представляющих товары. Данный алгоритм способен подготавливать изображения плечевой одежды, поясной одежды, головных уборов и обуви. В данной статье для определения частей тела человека используется архитектура OpenPose [5], модифицированная моделью внимания (attention model) [6]. Построенная модель позволяет не только структурировать части тела человека за счет полей сходства частей, но и более детально выделять части человеческого тела благодаря стимулам к

усилению значимых и подавлению незначимых объектов на изображении, что достигается ввиду построения двумерной матрицы оценок для каждой тепловой карты.

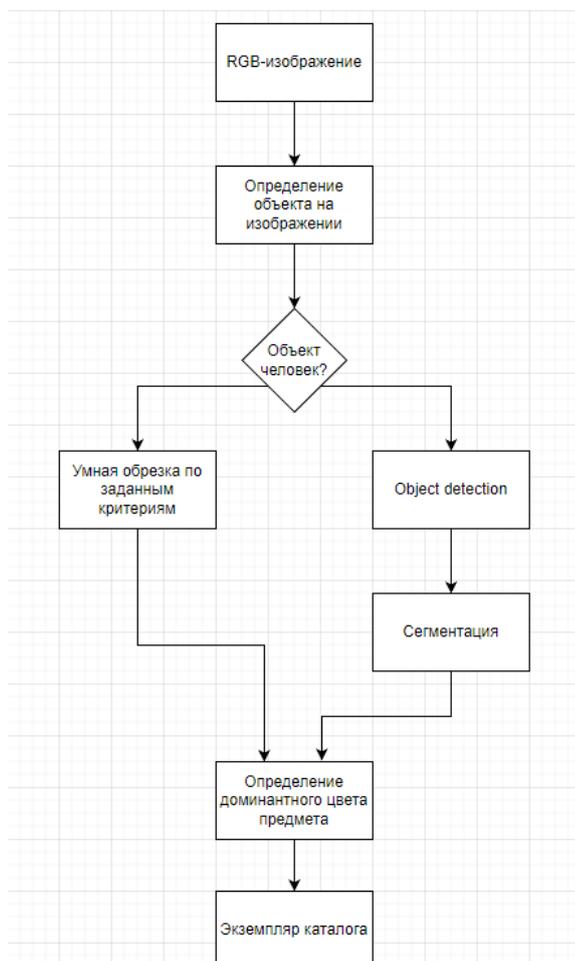


Рис. 1. Схема технологии

Если тип объекта не человек, то происходит классификация и обнаружения положения этого объекта при помощи сверточной нейронной сети YOLACT [2], а также модели внимания, после чего производится сегментация для последующего удаления фона. Модель внимания фиксирует кросс-канальные корреляции признаков, сохраняя при этом независимое представление в метаструктуре. Модуль сети выполняет набор преобразований для вложений низкой размерности и объединяет их выходные данные. Каждое преобразование включает в себя применение модели внимания по каналам, чтобы зафиксировать взаимозависимости карт признаков. Каждое преобразование имеет одну и ту же топологию. Такой подход позволяет ускорить обучение, используя идентичную реализацию, как и у унифицированных операторов CNN.

Последним шагом в технологии создания каталога товаров электронной коммерции является определение доминантного цвета товара и генерация образца – небольшого изображения (обычно 50x50 px), заполненного доминантным цветом. Для этого используется метод k-means [7]. При подборе правильного значения k центр тяжести самого большого кластера будет довольно хорошим представлением доминирующего цвета изображения.

## 2. Результаты и их обсуждение

В ходе исследования были разработаны следующие алгоритмы:

- задача обрезки изображения – алгоритм на основе архитектуры OpenPose с использованием модифицированной модели внимания VGG-19 [8]. Алгоритм показал улучшение точности на 8% и способен распознавать 23 ключевые точки человеческого тела.
- задача обнаружения объекта на изображении – алгоритм с использованием сверточной нейронной сети YOLACT с применением модели внимания. Обучение модели было направлено на распознавание 26 классов объектов электронной коммерции: модели (человека в полный рост), обуви (четырёх классов), одежды (пяти классов), еды (пяти классов), косметики (пяти классов), кухонной техники, аксессуаров и класса заднего фона. Модель внимания использовалась в основе ResNet-101 [9] нейронной сети для выделения наиболее значимых признаков объекта. Она позволила улучшить распознавание объектов в среднем на 3%.
- задача сегментации – алгоритм на основе сверточной нейронной сети YOLACT с применением стандартизации весов. Обучение модели было направлено на распознавание 21 класса объектов электронной коммерции. Стандартизация весов использовалась в сверточных слоях при прямом проходе обучения нейронной сети. Это позволило улучшить классификацию объектов в среднем на 3%, а обнаружение объектов — на 4%.
- задача определения основного цвета товара на изображении – алгоритм на основе k-means, решающий задачу кластеризации цветов, обнаруженных на вырезанном при помощи предыдущего алгоритма изображении товара.

Во всех случаях для обучения использовались видеокарты GPU NVIDIA T4, набор данных COCO [10], размер исходного изображения не

менее 800x800 px. Алгоритмы системы были реализованы на фреймворке PyTorch.

### **Заключение**

Решение таких задач, как обнаружение объекта на изображении, его сегментация, обрезка и определение основного цвета предмета позволили создать технологию для автоматической подготовки изображений для электронного каталога товаров. Данная технология была протестирована на реальных изображениях товаров одного интернет-магазина.

Пример работы представлен на рисунке 2.

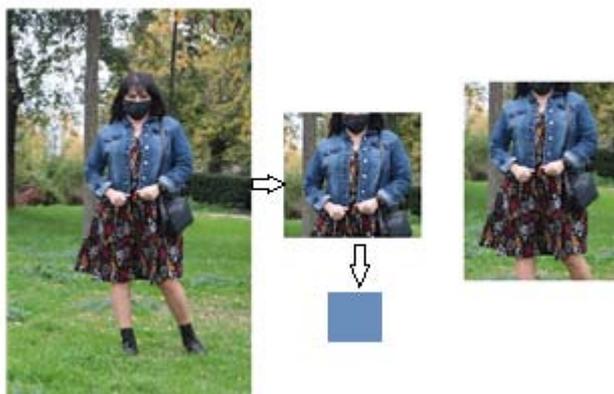


Рис. 2. Пример работы Smart Cropping системы

Результаты, полученные в разработанной программе, могут быть использованы для подготовки изображений других категорий товаров.

По мере того, как данных становится больше и появляются новые технологии, мы постоянно улучшаем построенную систему. Например, планируется оценка релевантности изображения, используя название и описание продукта.

### **Библиографические ссылки**

1. eCommerce Product Image Guide. URL: <https://www.threekit.com/blog/ecommerce-product-image-guide-2020>. (Date of access: 25.03.2021).
2. Bolya D. YOLACT: Real-time Instance Segmentation / D. Bolya [et al.] // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019. P. 9157–9166.
3. Сорокина В.В., Абламейко С.В. Применение метода стандартизации весов в нейросети в задаче сегментации изображений электронной коммерции / В. В. Сорокина // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2020) : доклады XIX Международной конференции, Минск, 19 ноября 2020 г. Минск: ОИПИ НАН Беларуси, 2020 С. 182–187.

4. Viktoria Sorokina and Sergey Ablameyko. Extraction of Human Body Parts from the Image Using Convolutional Neural Network and Attention Mode. Proceedings of 15th International conference “Pattern Recognition and Information Processing”, Minsk, UIIP NASB, 2021. P. 84-88.
5. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Zhe Cao [et al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 21–26 July 2017. Honolulu, 2017. P. 1302–1310.
6. Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural Machine Translation by Jointly Learning to Align and Translate. In 3rd International Conference on Learning Representations.
7. MacQueen J.B. Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Vol. 1, P. 281–297. California: University of California Press.
8. Very Deep Convolutional Networks for Large-Scale Image Recognition / Karen Simonyan and Andrew Zisserman // International Conference on Learning Representations, San Diego, May 7–9, 2015. San Diego, 2015. P. 1137–1149 .
9. He K. Deep Residual Learning for Image Recognition / K. He [et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016. P. 770–778.
10. COCO dataset // COCO 2018 Keypoint Detection Task. URL: <http://cocodataset.org/#overview> (Date of access: 05.04.2019).