

АВТОМАТИЗИРОВАННЫЙ ПОДСЧЕТ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ ГЛУБОКОГО ОБУЧЕНИЯ НА ПРИМЕРЕ СИГАРЕТНЫХ ФИЛЬТРОВ

Егоров А.Т. Email: Egorov629@scientifictext.ru

Егоров Артем Тимурович – магистрант,
кафедра компьютерных интеллектуальных технологий,
Санкт-Петербургский политехнический университет Петра Великого, г. Санкт-Петербург

Аннотация: компьютерные технологии позволяют автоматизировать большое количество задач на производстве, в то же время достижения в увеличении производительности устройств и разработки новых алгоритмов увеличивают круг задач, подверженных автоматизации. В статье рассматривается применение компьютерного зрения и машинного обучения для задачи подсчета объектов на производстве для контроля качества, на примере сигаретных фильтров. Подсчет осуществляется сотрудниками с помощью смартфонов.

Ключевые слова: компьютерное зрение, машинное обучение, нейронные сети, автоматизация.

AUTOMATED CALCULATION OF OBJECTS ON IMAGE WITH THE USE OF DEEP-LEARNING TECHNOLOGY ON THE EXAMPLE OF CIGARETTE FILTERS

Egorov A.T.

Egorov Artem Timurovich – Undergraduate,
COMPUTER INTELLIGENT TECHNOLOGIES,
PETER THE GREAT ST. PETERSBURG POLYTECHNIC UNIVERSITY, ST. PETERSBURG

Abstract: computer technologies allow to automate a large number of tasks in production, at the same time, the achievement in increasing the productivity of devices and the development of new algorithms increase the range of tasks prone to automation. The article deals with the application of computer vision and machine learning for the task of counting objects in the industry for quality control, using the example of cigarette filters. Counting is carried out by employees using smartphones.

Keywords: computer vision, machine learning, neural networks, automation.

УДК 004.93

Дисциплина компьютерного зрения (англ. - computer vision) появилась в середине 1960-х годов. В ее задачи входила разработка методов обработки изображений, позволяющих производить автоматизированное детектирование, отслеживание и классификацию объектов в кадре. На первых этапах системы компьютерного зрения требовали сложного оборудования как для получения изображения, так и для его обработки [1]. На текущем этапе развития технологий персональные устройства со встроенной камерой позволяют получать изображения приемлемого качества для решения большого количества задач компьютерного зрения.

Одновременно, успехи в области машинного обучения и возросшие вычислительные мощности позволили применять нейронные сети для решения задач компьютерного зрения. Одной из главных проблем нейронных сетей являлась сложность их обучения. В 1980 гг. появляются глубокие нейронные сети, состоящие из многих слоев и большого количества нейронов [6]. Применение алгоритмов обучения обычных сетей на глубоких, являлось вычислительно сложной задачей. Однако примерно с 2006 года разными группами исследователей разрабатываются новые эффективные алгоритмы обучения [4], [5], [7].

Одна из возможных сфер применения компьютерного зрения – контроль качества на производстве. Первые системы компьютерного зрения на производстве появились в 1980-х годах [2] и состояли из целого комплекса оборудования. Сегодня вектор развития таких систем идет по пути миниатюризации и упрощения, что выражается, например, в использовании персональных устройств с камерами для решения различных задач контроля.

В статье демонстрируется подход к подсчету объектов на производстве для контроля качества, на примере угольных сигаретных фильтров. На рисунке 1 приведен пример фотографии для обработки.

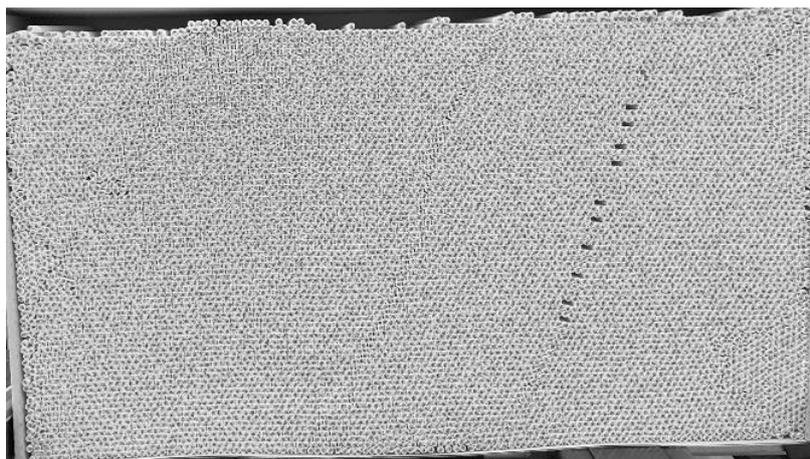


Рис. 1. Пример фотографии для анализа

Одним из способов решения данной задачи является использование алгоритма Виолы-Джонса, позволяющего выделять объекты на изображении, однако специфика объекта подсчета (угольного сигаретного фильтра) не позволяет использовать этот метод, в силу значительных погрешностей, которые демонстрирует алгоритм. Было произведено тестирование данного подхода с использованием библиотеки OpenCV, которая содержит реализацию алгоритма [8], пример изображения, демонстрирующего работу алгоритма, приведен на рисунке 2. Алгоритм находит отдельные фильтры, однако делает множество ошибок и пропускает большую часть, что не позволяет его использовать в полной мере для решения исходной задачи.

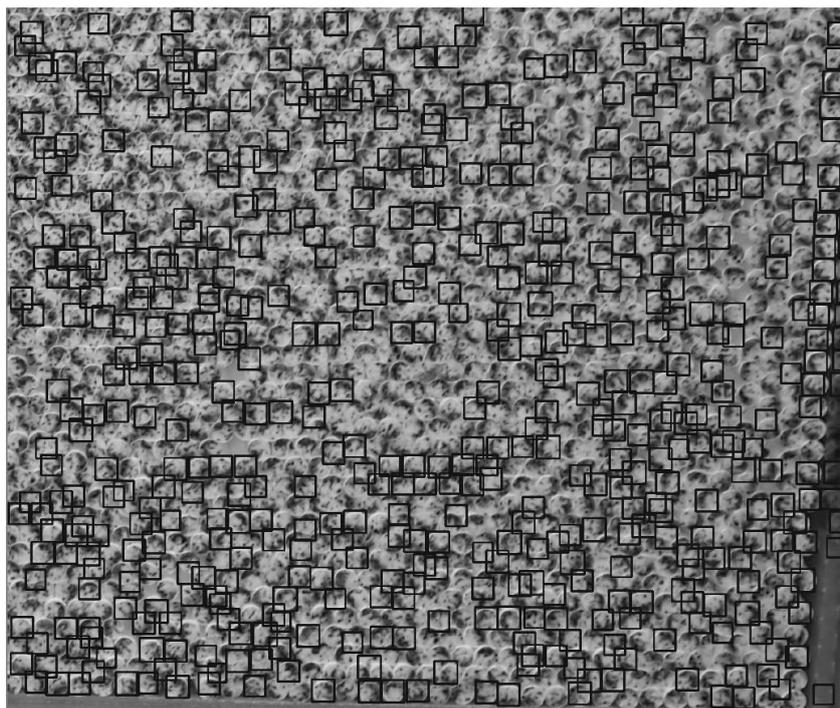


Рис. 2. Результат работы алгоритма Виолы-Джонса

Для подсчета фильтров на фото нет необходимости в поиске точного положения и границ объектов, поэтому можно применить следующий алгоритм:

1. Фильтрация изображения для преобразования к другому виду.
2. Бинаризация изображения для отделения объектов от фона.
3. Нахождение связанных компонент, которые являются искомыми объектами.
4. Подсчет связанных компонент.

Последние три этапа алгоритма могут быть выполнены с использованием уже разработанных библиотек, самым сложным этапом является первый этап.

Выполнять преобразования изображений из одного вида к другому позволяют нейронные сети. Deep Autoencoder – это нейронная сеть, состоящая из двух блоков encoder и decoder, в каждый из которых

входит несколько слоев [9]. Первая часть уменьшает размерность входных данных, позволяя выявить основные существенные для задачи признаки, вторая часть позволяет на основании этих признаков получить требуемые значения зависимых переменных. Такие сети могут осуществлять кодирование данных в другой вид. При работе с изображениями deep autoencoder может убирать шумы или достраивать изображения.

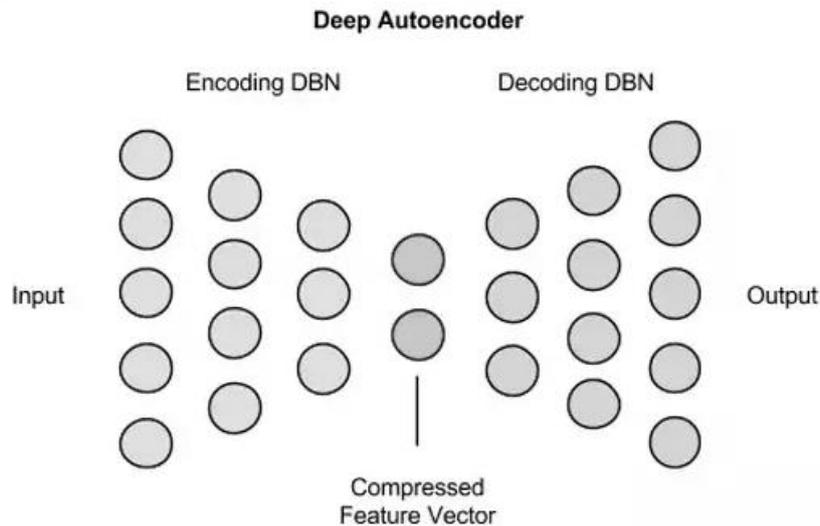


Рис. 3. Схема Deep Autoencoder (<https://deeplearning4j.org/deepautoencoder>)

Для реализации кодирования изображения используется библиотека DeepLearning4j. С ее помощью строится сеть, состоящая из четырех слоев:

- encoder состоит из двух ограниченных машин Больцмана;
- decoder состоит из одной ограниченной машины Больцмана и слоя прямого распространения.

Для обучения сети вручную были размечены центры фильтров на фотографиях, пример разметки приведен на рисунке 4.

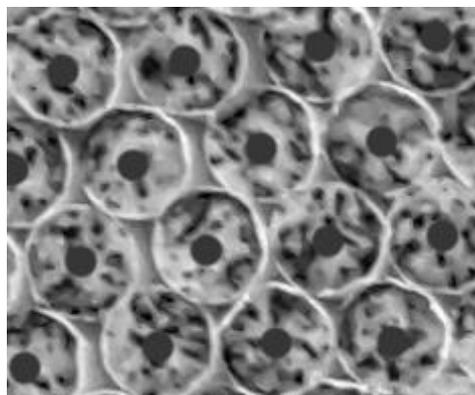


Рис. 4. Размеченная фотография

Для на вход сети подается небольшая часть исходного изображения, для обработки всего изображения, производится его обход с помощью окна заданного размера. Чтобы устранить неравномерность освещения, которая может ухудшать результаты работы сети, применяется эквализация гистограммы и приведение к оттенкам серого.

Для примера рассмотрим обработку фотографии, изображенной на рисунке 1.

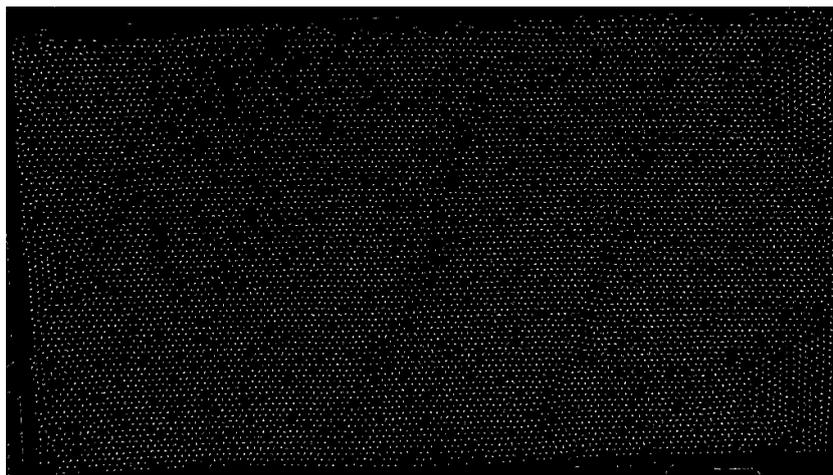


Рис. 5. Обработанное нейронной сетью

На рисунке 5 показан результат обработки нейронной сетью. На следующем этапе происходит бинаризация изображения и с помощью морфологических операций удаляются шумы. Для проверки результата полученные сегменты накладываются на исходное изображение в оттенках серого. Результат работы программы показан на рисунках 6 и 7.



Рис. 6. Результат обработки

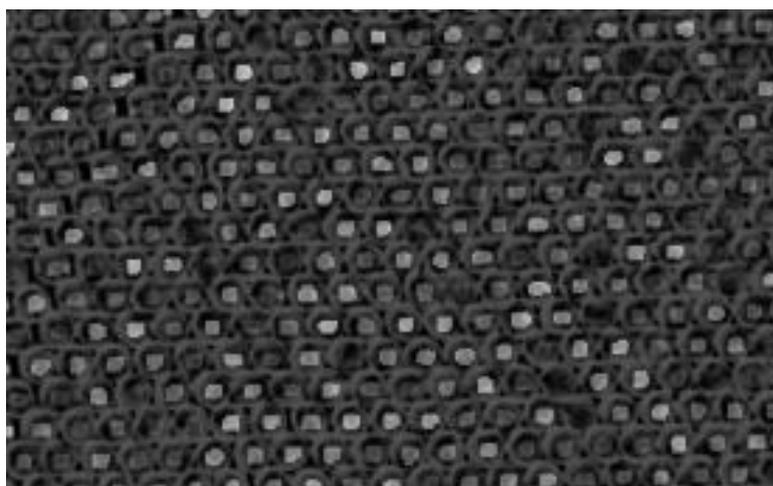


Рис. 7. Результат крупным планом

Для оценки точности работы используется f1-score учитывающая ложноположительные и ложноотрицательные срабатывания. Для данной фотографии точность получилась 98.4%.

Выводы

В ходе работы был разработан и реализован алгоритм подсчета объектов на фотографии сделанной на

смартфон. Следующий этап в обработке изображения – автоматическое определение границ паллеты с сигаретными фильтрами. После завершения этого этапа предстоит внедрение системы на производство. Разработанный алгоритм может быть расширен на поиск других объектов путем обучения сети на других данных.

Список литературы / References

1. *Форсайт Д., Понс Ж.* Компьютерное зрение. Современный подход. М.: ИД Вильямс, 2004.
2. *Braggins D.* Robots sharpen up their vision. *New Scientist*, 1983.
3. Deep Autoencoders [Электронный ресурс]. Режим доступа: <http://deeplearning4j.org/deepautoencoder/> (дата обращения: 05.03.2017).
4. *Hinton G.E.* Learning multiple layers of representation // *Trends in cognitive sciences*, 2007. Т. 11. № 10. С. 428-434.
5. *Hinton G.E., Osindero S., Teh Y.W.* A fast learning algorithm for deep belief nets // *Neural computation*, 2006. Т. 18. № 7. С. 1527-1554.
6. *LeCun Y. et al.* Backpropagation applied to handwritten zip code recognition // *Neural computation*, 1989. Т. 1. – № 4. С. 541-551.
7. *Nielsen M.A.* *Neural Networks and Deep Learning*. Determination Press, 2015.
8. *Viola P., Jones M.* Robust real-time object detection // *International Journal of Computer Vision*. 2001. Т. 4. № 34–47.
9. *Viola P., Jones M.J., Snow D.* Detecting pedestrians using patterns of motion and appearance // *International Journal of Computer Vision*, 2005. Т. 63. № 2. С. 153-161.7.