

10. Хамизуллин Ф.Ф. Применение гидроциклонов в химической и нефтехимической промышленности / Ф.Ф. Хамизуллин, С.И. Валеев // В сб. Интенсификация тепло-массообменных процессов, промышленная безопасность и экология. Пятая Всероссийская студенческая научно-техническая конференция. – Казань: КНИТУ, 2018. – С. 212-215.
11. Valeev S.I. Research of equipment for cleaning oil-containing waste-water / Journal of physics: Conference Series, 1399(2019)044021.

NUMERICAL SIMULATION OF HYDRODYNAMIC PROCESSES IN A HYDROCYCLONE

VALEEV Sergey Ildusovich

PhD in Technical Sciences, Associate Professor

SAVCHUK Vladimir Alexandrovich

graduate student

Kazan National Research Technological University
Kazan, Russia

The study of the effective viscosity in a cylindrical hydrocyclone for the separation of emulsions with a low content of light impurities has been carried out. It was found that the effective viscosity in a cylindrical hydrocyclone increases with an increase in the unloading ratio.

Key words: hydrocyclone, swirl, turbulence, pressure, unloading ratio.

АВТОМАТИЗАЦИЯ ОПРЕДЕЛЕНИЯ КОНТРАФАКТНЫХ ЭЛЕКТРОННЫХ КОМПОНЕНТОВ

СЕНИЛОВ Михаил Андреевич

доктор технических наук, профессор

ЗУБАРЕВ Владислав Анатольевич

студент

ФГБОУ ВО «Ижевский государственный технический университет им. М.Т. Калашникова»
г. Ижевск, Россия

Рассматриваются вопросы автоматизации обнаружения контрафактных элементов в электронной компонентной базе (ЭКБ). Для решения данной задачи предлагается использовать технологии машинного обучения и анализа данных, методы бинаризации и распознавания изображений, нейронные сети. Проведен сравнительный анализ известных OCR-систем (optical character recognition – оптическое распознавание символов) и нейронных сетей с целью обоснованного выбора инструментальных средств и их применения в предлагаемой методике. Практическое применение предлагаемого подхода в составе разрабатываемого приложения для обнаружения контрафакта позволит увеличить как точность, так и производительность приложения, обеспечив его работу в реальном времени.

Ключевые слова: контрафактные электронные компоненты, электронная компонентная база, нейронные сети, бинаризация, распознавание изображений.

Контрафактный электронный компонент – компонент, опознавательные или так называемые контрольные точки (например, название и маркировка производителя, маркировка даты и партии) которого были умышленно изменены. Появление на рынке ЭКБ контрафактных компонентов обусловлено несколькими причинами, а именно: устаревание компонентов, время выполнения заказа (например, производитель не может поставить детали в срок), ценообразование (запчасти по более низким ценам у других поставщиков), недобросовестный контроль качества на производстве и в цепочке следования к потребителю, дешевизна инструментов для создания подделок, а также большая стоимость процедуры тестирования. Контрафактные детали, в основном, представляют собой детали с измененной маркировкой (с более новым кодом даты производства), отремонтированные или переупакованные компоненты.

Обнаружение контрафакта в ЭКБ при помощи метода определения характеристик материала и испытаний, направленных на разрушение детали, например, декапсуляции, тре-

буют наличия дорогостоящего инструмента и оборудования. В свою очередь, визуальный осмотр возможно проводить даже невооруженным глазом. Визуальный осмотр может стать первым шагом в процессе обнаружения контрафакта, но далеко не единственным.

Помимо визуального осмотра возможно измерять габаритные размеры и массу образца. Но такие проверки требуют доступ к базе данных, содержащей информацию о размерах и массе компонентов, а также поддержку со стороны производителя, если, к примеру, в базе не окажется информации для какого-либо компонента.

В рамках проводимых исследований создан инспекционный стенд входного контроля, который использует различные датчики для сбора информации о конкретном компоненте: микроскоп, поворотный стол с акселерометром, тензометрический датчик, индуктивный датчик. На основе показаний датчиков и высококачественных снимков проводится сложный анализ данных в подсистеме, состоящей из нескольких блоков обработки информации, см. рисунок 1.



Рисунок 1. Блоки обработки информации

Анализ компонента проводится при помощи нескольких методик, а именно: обработка изображений компонента – бинаризация и анализ при помощи нейронных сетей; сравнение изображения данного компонента с изображением эталона; сравнение физических характеристик эталонного образца,

хранящихся в базе, с характеристиками данной детали; сравнение маркировки.

Обработка изображений компонентов может быть затруднена за счет влияния ряда факторов, таких как неидеальное освещение, «шумы» на фоне или нерегулярное различие цветов. Для того, чтобы добиться максимального

качества распознавания текста и дефектов на изображениях, разработана сложная процедура анализа характеристик изображений электронных компонентов, основанная на использовании алгоритма бинаризации.

Бинаризация – это процесс перевода цветного (или в оттенках серого) изображения в двухцветное черно-белое. Существует достаточно большое количество алгоритмов бинаризации, самый распространенный из которых алгоритм Оцу (Otsu's method). Бинаризация может быть как статической, основным параметром которой является порог t – значение, с которым сравнивается яркость каждого пикселя, так и динамической, основанной на зонировании изображения и динамическом подборе такого порога в зависимости от насыщенности того или иного участка изображения [2]. Главная цель бинаризации – радикальное уменьшение количества информации, с которой приходится работать. Так, применив метод бинаризации к обычному изображению, каждый пиксель которого состоит из 256 компонентов цветов красного, синего и зеленого, количество информации сократится более чем в 8 млн. раз, что имеет принципиальное значение для возможности дальнейшей обработки изображения при помощи OCR и нейронных сетей, т. к. существенно сокращает не только объем памяти, необходимой для обработки фотографий, но и время работы программы.

В проводимых исследованиях использован новый метод [11], включающий в себя несколько техник, а именно: нелинейное сглаживание и улучшение контраста; новый метод вычисления порогового значения бинаризации [3]; «вычитание» фона из изображения; проецирование значений цветов пикселя для улучшения контраста.

Определение маркировки компонента – одна из первоначальных задач анализа ЭКБ на наличие контрафакта. Благодаря информации, нанесенной на корпус детали, можно

определить название и тип компонента, код или дату производства, номер партии.

В настоящее время для распознавания текста разработаны и используются несколько программных решений. Однако варианты с открытым исходным кодом являются наиболее привлекательными из-за их низкой стоимости. В рамках проводимого исследования рассмотрены и проанализированы наиболее часто применяемые механизмы OCR: Tesseract [10], Cuineform [4], GOCR [6] и OCRAD [8].

На основе компонентной базы [5] проведено сравнение точности определения текста на компонентах. Для количественной оценки производительности OCR-систем используется такое понятие как счет $F \in [0,1]$ (F -score) [7]. В каждом изображении для эталонной строки S_0 и распознанной S точность вычисляется по следующей формуле:

$$F = 2 * \frac{P * R}{P + R}$$

где P и R , соответственно, точность (*precision*) и полнота (*recall*). Точность и полнота вычисляются, соответственно, по следующим формулам:

$$P = \frac{\max(L_S, L_{S_0}) - d_{lev}(S, S_0)}{L_S}$$

$$R = \frac{\max(L_S, L_{S_0}) - d_{lev}(S, S_0)}{L_{S_0}}$$

где L_{S_0} , L_S обозначают, соответственно, длину эталонной и распознанной строк, d_{lev} – расстояние Левенштейна, описанное в (Расстояние Левенштейна [Электронный ресурс]: Википедия – свободная энциклопедия. – режим доступа: <https://ru.wikipedia.org/>). Общие результаты исследования OCR-систем на точность приведены в табл. 1. Из анализа этих результатов следует, что, по сравнению с GOCR и OCRAD, Tesseract и Cuineform обеспечивают лучшие результаты распознавания текста.

Таблица 1

СРАВНЕНИЕ ТОЧНОСТИ OCR-СИСТЕМ

OCR – система	точность P	полнота R	счет F-score
Tesseract	0,934	0,944	0,939
Cuineform	0,941	0,944	0,943
GOCR	0,711	0,831	0,766
OCRAD	0,851	0,879	0,865

Для обработки изображений целесообразно использовать сверточные нейронные сети (Сверточная нейронная сеть [Электронный ресурс]: Википедия – свободная энциклопедия. – режим доступа: <https://ru.wikipedia.org/>), т. к. именно их архитектура специально нацелена на эффективное распознавание образов. Простейшая архитектура такой сети представлена на рисунке 2. При помощи сверточной

нейронной сети определяются тип маркировки (печать краской, лазерная маркировка или гравировка фрезером), расстояние между соседними ножками, следы шлифования и химического травления. Для исследования были выбраны несколько моделей сверточных нейронных сетей, а именно: YOLO v4, Faster R-CNN ResNet101 V1 640x640 и SSD MobileNet V2 FPNLite 640x640.

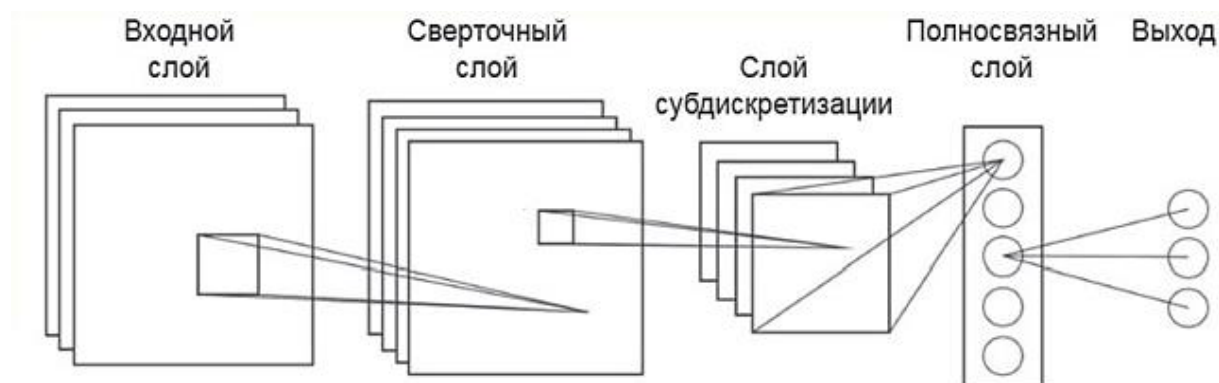


Рисунок 2. Архитектура сверточной нейронной сети

Обучение и тестирование сетей проводилось на компьютере, оборудованном процессором AMD® Ryzen™ 7 2700 3.2 Ghz, оперативной памятью 16 Gb, видеокартой ASUS GeForce® TURBO-GTX1080TI-11, под управлением операционной системы Ubuntu 18.04.5 LTS. Каждая из сетей была обучена при помощи датасета (от англ. *dataset* — набор данных), состоящего из изображений возможных дефектов контрафактных электронных компонентов (рисунок 3). Обучение сетей прово-

дилось итеративно по заданному набору картинок. Всего было проведено 6000 итераций сериями по 64 изображения. Все изображения для обучения и тестирования были отформатированы и приведены к размеру 800x600 пикселей. Для каждого из изображений вручную была создана карта объектов (*label map*) — координаты углов прямоугольников, покрывающих искомые объекты. Продолжительность обучения составила порядка 15-20 часов реального времени.

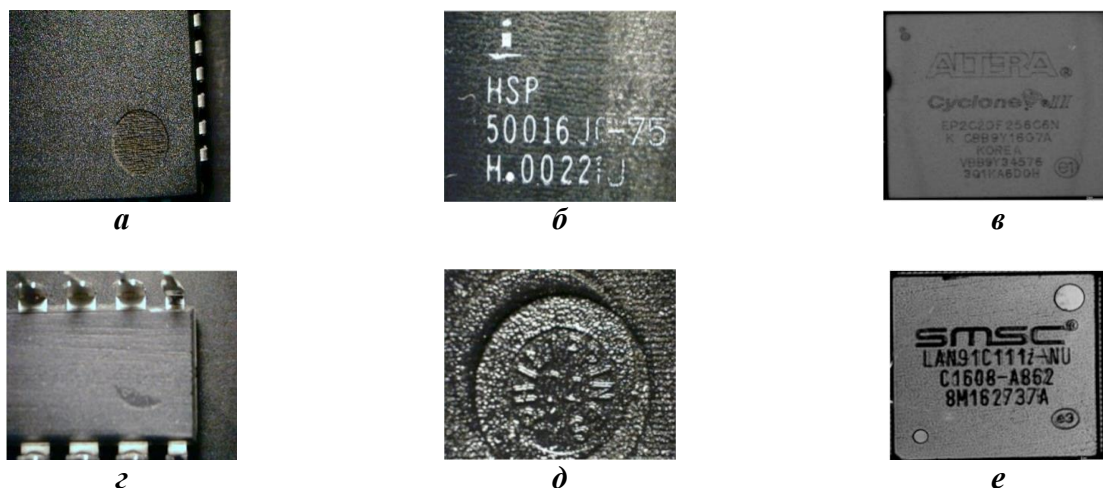


Рисунок 3. Примеры дефектов контрафактных электронных компонентов: а, д – изменение штампа производителя; б, г, е – следы шлифовки; в – следы химического травления

Для проверки точности нейронных сетей, прошедших обучение, был создан вспомогательный датасет, состоящий из таких же примеров дефектов. Используя этот датасет в качестве эталонного, была рассчитана точность определения дефектов при помощи

использованных нейронных сетей. В качестве основного показателя точности использовалась такая характеристика как *mAP* (от англ. *mean average precision*) [9]. Сравнение точности нейронных сетей представлено в таблице 2.

СРАВНЕНИЕ ТОЧНОСТИ НЕЙРОННЫХ СЕТЕЙ

Таблица 2

Нейронная сеть	Точность, mAP	Среднее время обработки кадра, мс
YOLO v4	0,9164	18
Faster R-CNN	0,8944	53
SSD MobileNet	0,8655	39

Как видно из таблицы, нейронная сеть «YOLO v4» добилась наибольших показателей точности, а если учесть, что на обработку одного кадра в среднем уходит 15-20 мс (а это, по сравнению с Faster R-CNN, трехкратное увеличение производительности), также возможно применение данной сети на потоковом видео, например, с микроскопа, с частотой кадров более 60, что позволяет осуществлять распознавание дефектов и определение контрафакта в режиме реального времени.

Целью проводимых исследований в перспективе является полное автоматическое обнаружение контрафактных компонентов, однако за неимением достаточного количества данных, необходимых для обучения нейросети, невозможно гарантировать абсолютную точность представленного метода. Поэтому участие человека на начальном этапе необходимо для дальнейшего обучения нейронных сетей, а

также оптимизации параметров системы. Для этого оператору необходимо вручную выделять нераспознанные объекты и добавлять их в датасет для продолжения обучения нейронной сети. Кроме того, недостающая информация о конкретном электронном компоненте может быть добавлена в базу данных.

Для определения размеров объекта на изображении при известных параметрах линз объектива применяется метод [1]. Полученные в результате применения этого метода приблизительные размеры компонента сравниваются с характеристиками из базы данных. Если полученные данные не находятся в окрестности радиуса $\delta = 10^{-2}$ идеальных данных, то оператору требуется произвести ручное измерение при помощи микрометра. Аналогично определяется и сравнивается масса компонента.

В данной работе представлен метод, поз-

воляющий автоматизировать поиск контрафактных электронных компонентов и повысить его эффективность. В основном рассмотрены алгоритмы нейронных сетей для работы с изображениями: *бинаризация, распознавание текста и изображений*. Благодаря использованию нового метода бинаризации были заметно улучшены контуры и характеристики распознаваемых объектов. Проведенные сравнение и выбор лучших ме-

ханизмов оптического распознавания текста позволили увеличить как точность, так и производительность приложения для обнаружения контрафакта. Проведенные сравнительный анализ нейронных сетей и выбор из них наиболее целесообразных открыли новые возможности в создании автоматизированных систем определения контрафакта, повысив их эффективность и обеспечив возможность работы в реальном времени.

СПИСОК ЛИТЕРАТУРЫ

1. Гейдаров П.И. Алгоритм определения расположения и размеров объектов на основе анализа изображений объектов // Компьютерная оптика, том 35. – № 2. – 2011.
2. A New Local Adaptive Thresholding Technique in Binarization / T. Romen Singh, Sudipta Roy, O. Imocha Singh и др. – IJCSI, vol. 8, problem 6, № 2, 2011.
3. A New Local Adaptive Thresholding Technique in Binarization / T. Romen Singh, Sudipta Roy, O. Imocha Singh и др. – IJCSI, vol. 8, problem 6, № 2, 2011.
4. Cuineform-linux: Cuineform Linux in Launchpad. – URL <https://launchpad.net/cuneiform-linux>.
5. FICS-PCB: A Multi-Modal Image Dataset for Automated Printed Circuit Board Visual Inspection / Hangwei Lu, Dhvani Mehta, Olivia Paradis и др. – Cryptology ePrint Archive, Report 2020/366, 2020.
6. GOCR: GOCR. – URL: <http://jocr.sourceforge.net/>.
7. Jacob J. The Best Metric to Measure Accuracy of Classification Models: CleverTap: Omnichannel Customer Engagement & User Retention Platform. – URL: <https://clevertap.com/>.
8. OCRAD: Ocrad – GNU Project – Free Software Foundation (FSF). – URL: <http://www.gnu.org/software/ocrad/>.
9. Tan R.J. Breaking Down Mean Average Precision (mAP): Towards Data Science. – URL: <https://towardsdatascience.com/>.
10. Tesseract-ocr: Tesseract Open Source OCR Engine repository. – URL: <https://github.com/tesseract-ocr/>.
11. Text Recognition for Information Retrieval in Images of Printed Circuit Boards / W. Li, S. Neullens, M. Breier, M. Bosling и др. – RWTH Aachen University, IECON, 2014.

AUTOMATION OF DETECTION OF COUNTERFEIT ELECTRONIC COMPONENTS

SENILOV Mikhail Andreevich

Grand PhD in Technical Sciences, Professor

ZUBAREV Vladislav Anatolievich

student

M.T. Kalashnikov's Izhevsk State Technical University
Izhevsk, Russia

The issues of automation of detection of counterfeit elements in an electronic component base (EEE) are considered. To solve this problem, it is proposed to use machine learning and data analysis technologies, methods of binarization and image recognition, neural networks. A comparative analysis of well-known OCR systems (optical character recognition) and neural networks is carried out in order to make a reasonable choice of tools and their application in the proposed method. Practical application of the proposed approach as part of the developed application for detecting counterfeit will increase both the accuracy and performance of the application, ensuring its operation in real time.

Key words: counterfeit electronic components, electronic component base, neural networks, binarization, image recognition.