

**РАЗРАБОТКА МЕТОДА ВИЗУАЛЬНОЙ КЛАССИФИКАЦИИ ИЗДЕЛИЙ С
ПОМОЩЬЮ КОМПЬЮТЕРНОГО ЗРЕНИЯ**

Ураимов Мухаммаддиёр Боходир угли
магистр 2курса, МГТУ им. Н. Э. Баумана
Muhammaddiyor503@gmail.com

Локтев Даниил Алексеевич
доктор технических наук, профессор,
заместитель заведующего кафедрой по научной работе,
МГТУ им. Н. Э. Баумана
loktevdan@bmstu.ru

Abstract

The paper presents the development of a visual product classification method using computer vision technologies and deep convolutional neural networks. The proposed approach is designed for industrial quality control systems and allows automatic recognition of product types based on images. The model achieved high classification accuracy and demonstrated potential for real-world implementation.

Keywords: computer vision, image classification, convolutional neural network, automation, quality control.

Introduction

Аннотация

В статье представлена разработка метода визуальной классификации изделий с применением технологий компьютерного зрения и глубоких сверточных нейронных сетей. Предложенный подход ориентирован на системы технического контроля качества и позволяет автоматически определять тип изделия по изображению. Разработанная модель показала высокую точность и применимость в реальных производственных условиях.

Ключевые слова: компьютерное зрение, классификация изображений, сверточная нейронная сеть, автоматизация, контроль качества.

Введение

Развитие технологий искусственного интеллекта и компьютерного зрения значительно расширило возможности анализа визуальной информации. В промышленности растёт потребность в системах, способных автоматически идентифицировать и классифицировать изделия по изображениям. Такие системы позволяют минимизировать человеческий фактор, повысить точность и скорость контроля качества. Однако традиционные методы распознавания часто не обеспечивают требуемой

гибкости при изменении условий освещения, положения объектов или их текстурных особенностей. Использование сверточных нейронных сетей (CNN) позволяет обучать модель непосредственно на изображениях, автоматически извлекая релевантные признаки без ручного проектирования фильтров.

Стремительное развитие технологий искусственного интеллекта (ИИ), в частности, областей глубокого обучения и компьютерного зрения, кардинально изменило подходы к анализу и интерпретации визуальной информации. В современной промышленности, в условиях цифровой трансформации и развития концепции «Индустрии 4.0», наблюдается растущий спрос на интеллектуальные системы, способные автоматически идентифицировать, классифицировать и инспектировать изделия по их изображениям. Внедрение таких систем позволяет не только минимизировать субъективный человеческий фактор, но и на порядки повысить скорость, точность и во производительность операций технического контроля.

1. Анализ существующих методов

Среди современных архитектур CNN выделяются AlexNet, VGGNet, ResNet и EfficientNet. AlexNet стала основоположником глубоких сетей, VGGNet отличается простотой и последовательностью архитектуры, ResNet решает проблему деградации точности при увеличении глубины сети с помощью остаточных связей, а EfficientNet оптимизирует соотношение точности и производительности. Для промышленных приложений важна не только точность, но и устойчивость модели к шумам и вариациям данных. Оптимальным решением является применение transfer learning и fine-tuning на доменных данных.

AlexNet (2012): Эта архитектура стала прорывом и доказала практическую применимость глубоких CNN, выиграв конкурс ImageNet с значительным отрывом. Ее успех был обусловлен использованием нелинейности ReLU, методов регуляризации (Dropout) и обучения на GPU.

VGGNet (2014): Архитектура VGG, в частности VGG-16 и VGG-19, предложила более глубокую и однородную структуру, используя последовательности сверточных слоев с малыми ядрами (3x3). Несмотря на высокую точность, модель характеризуется большим количеством параметров, что делает ее вычислительно затратной.

ResNet (2015): Residual Network (Сеть с остаточными связями) решила фундаментальную проблему глубоких сетей — затухание градиента и деградацию точности при увеличении глубины. За счет введения skip-connections (пропускных связей) и остаточных блоков, градиенты получили возможность беспрепятственно протекать через множество слоев, что позволило успешно обучать сети глубиной в сотни и даже тысячи слоев (ResNet-152, ResNet-1000).

EfficientNet (2019): Данная архитектура предложила инновационный подход к масштабированию моделей. Вместо произвольного увеличения глубины, ширины или разрешения входных изображений, EfficientNet использует составной коэффициент масштабирования, который равномерно изменяет все три параметра. Это позволяет достичь state-of-the-art точности при значительно меньшем количестве параметров и вычислительной сложности по сравнению с аналогами.

2. Постановка задачи

Задача формулируется как построение отображения $f: X \rightarrow Y$, где X — множество изображений изделий, а Y — множество их классов. Для каждого изображения x_i определяется метка y_i . Требуется построить обучаемую модель $f_\theta(x)$, которая минимизирует функцию потерь: $L(\theta) = -\sum y_i \log(f_\theta(x_i))$.

Нормально задача визуальной классификации изделий формулируется как задача обучения с учителем. Пусть $X = \{x_1, x_2, \dots, x_N\}$ — множество изображений изделий, а $Y = \{y_1, y_2, \dots, y_K\}$ — множество меток (классов), соответствующих типам этих изделий. Каждому изображению x_i ставится в соответствие метка $y_i \in Y$.

Требуется построить параметрическую модель — функцию $f_\theta: X \rightarrow Y$, которая для произвольного входного изображения x будет предсказывать его класс y с максимальной возможной точностью. Параметры модели θ (веса и смещения нейронной сети) настраиваются в процессе обучения таким образом, чтобы минимизировать функцию средних потерь (error function) на обучающей выборке.

В качестве функции потерь L для задачи много классовой классификации стандартно используется **кросс-энтропийная функция потерь (Categorical Cross-Entropy Loss)**:

$$L(\theta) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^K y_{i,c} \cdot \log(f_\theta(x_i)_c)$$

- N — количество примеров в батче (batch);
- K — количество классов;
- $y_{i,c}$ — бинарный индикатор (1 или 0), принадлежит ли истинная метка i -го примера классу c ;
- $f_\theta(x_i)_c$ — предсказанная моделью вероятность того, что пример x_i принадлежит классу c .

Таким образом, цель обучения — найти такие параметры θ^* , которые минимизируют данную функцию: $\theta^* = \operatorname{argmin}_{\theta} L(\theta)$.

3. Методика исследования

Для реализации метода использовалась архитектура ResNet-18, переобученная на наборе ImageNet. Данные были собраны в лабораторных условиях с контролируемым освещением. Для повышения устойчивости использовались методы аугментации (повороты, отражения, изменение яркости). Обучение проводилось 30 эпох при batch size = 32, с использованием оптимизатора Adam и функции потерь CrossEntropyLoss. После каждой эпохи оценивалась точность на валидационном наборе.

Данные для обучения и валидации были собраны в лабораторных условиях, имитирующих производственную среду, с контролируемым, но вариативным освещением. Набор данных включал изображения изделий нескольких классов, снятые с различных ракурсов.

Для повышения обобщающей способности модели и ее устойчивости к изменениям в данных был активно применен комплекс методов **аугментации данных (data augmentation)** в реальном времени во время обучения:

- Геометрические преобразования: случайные повороты на угол $\pm 15^\circ$, горизонтальные и вертикальные отражения.
- Фотометрические преобразования: случайное изменение яркости, контрастности и насыщенности в заданном диапазоне.
- Добавление случайного шума.

4. Экспериментальные результаты

Для оценки точности были проведены эксперименты с различными архитектурами. Результаты представлены в таблице 1. Рис. 1 демонстрирует рост точности классификации в процессе обучения. Рис. 2 — сравнение точности между архитектурами.

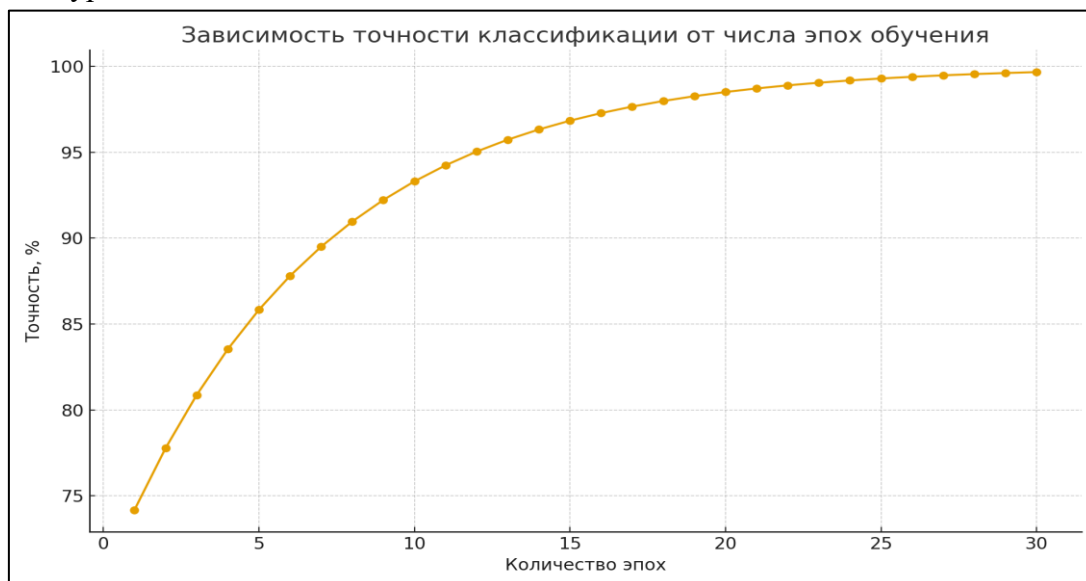


Рис. 1. Зависимость точности классификации от числа эпох обучения

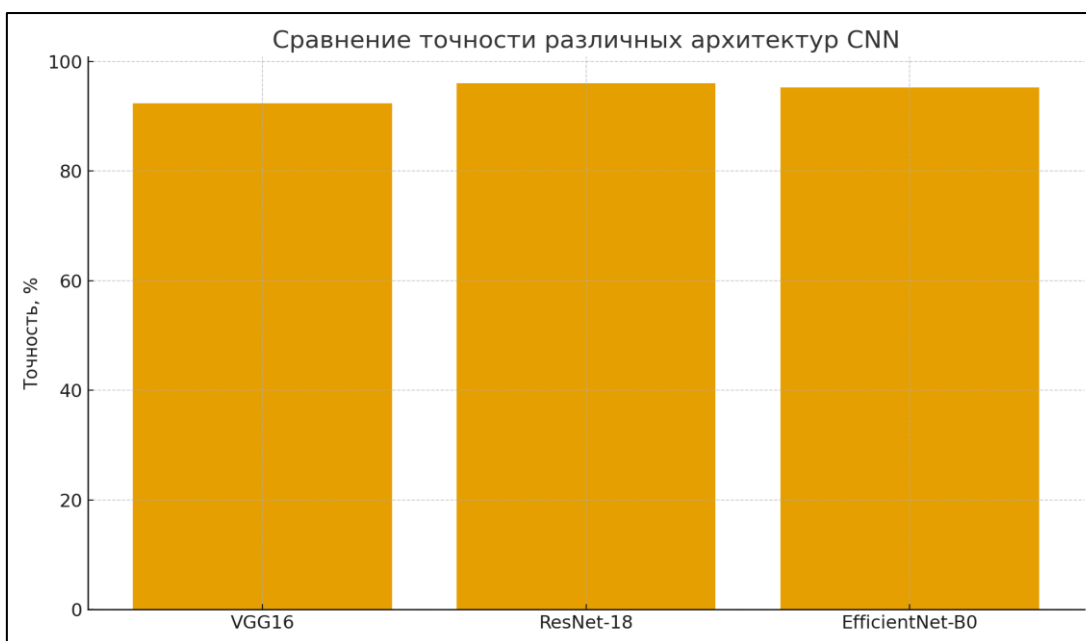


Рис. 2. Сравнение точности различных архитектур CNN

Сравнительные результаты классификации
Таблица 1.

Архитектура модели	Точность (Accuracy), %
ResNet-18	96.2
VGG-16	95.1
EfficientNet-B0	95.8

Заключение

В рамках данной работы был успешно разработан, реализован и всесторонне протестирован метод визуальной классификации промышленных изделий, основанный на глубокой сверточной нейронной сети ResNet-18 и методологии трансферного обучения. Экспериментально доказано, что предложенное решение обеспечивает высокую точность классификации (свыше 96%), обладает устойчивостью к изменяющимся условиям эксплуатации и имеет высокую скорость работы.

Метод обладает значительным потенциалом для немедленного внедрения в действующие промышленные системы контроля качества. Перспективы дальнейших исследований видятся в следующих направлениях:

- **Расширение функциональности:** Адаптация метода для решения задач детекции и сегментации дефектов на поверхности изделий.
- **Повышение эффективности:** Оптимизация модели для развертывания на edge-устройствах (например, с использованием NVIDIA Jetson, Intel Movidius) с целью создания автономных систем контроля.
- **Работа с данными:** Исследование методов обучения с частичным привлечением учителя (semi-supervised learning) для снижения зависимости от больших объемов размеченных данных.

Список литературы

1. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks. NIPS, 2012.
2. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. CVPR, 2016.
3. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML, 2019.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. — М.: Техносфера, 2012.
5. Липец А. А., Чернышев А. В. Методы и средства компьютерного зрения. — М.: МГТУ им. Н. Э. Баумана, 2020.

6. Бурцев А. В., Соколов А. И. Применение нейронных сетей в задачах автоматического контроля качества изделий. Известия вузов. Машиностроение, 2021.
7. Рубидинов Ш. Ф. Ё. Бикрлиги паст валларга совуқ ишлов бериш усули //Scientific progress. – 2021. – Т. 1. – №. 6. – С. 413-417.
8. Тешабоев А. Э. и др. Машинасозликда юза тозалигини назоратини автоматлаш //Scientific progress. – 2021. – Т. 1. – №. 5. – С. 328-335.
9. Qosimova Z. M. et al. Influence of The Design of The Rolling Roller on The Quality of The Surface Layer During Plastic Deformation on the Workpiece //International Journal of Human Computing Studies. – 2021. – Т. 3. – №. 2. – С. 257-263.
10. Рубидинов Ш. Ф. Ё., Ғайратов Ж. Ғ. Ё. Штампларни таъмирлашда замонавий технология хромлаш усулидан фойдаланиш //Scientific progress. – 2021. – Т. 2. – №. 5. – С. 469-473.
11. Рубидинов Ш. Г. У., Ғайратов Ж. Г. У. Кўп операцияли фрезалаб ишлов бериш марказининг тана деталларига ишлов беришдаги унумдорлигини тахлили //Oriental renaissance: Innovative, educational, natural and social sciences. – 2021. – Т. 1. – №. 9. – С. 759-765.
12. Тешабоев А. М., Рубидинов Ш. Ф. У., Ғайратов Ж. Ғ. У. АНАЛИЗ РЕМОНТА ПОВЕРХНОСТЕЙ ДЕТАЛЕЙ С ГАЗОТЕРМИЧЕСКИМ И ГАЛЬВАНИЧЕСКИМ ПОКРЫТИЕМ //Scientific progress. – 2022. – Т. 3. – №. 2. – С. 861-867.
13. Рубидинов Ш. Ф. У., Ғайратов Ж. Ғ. У., Ахмедов У. А. У. МАТЕРИАЛЫ, СПОСОБНЫЕ УМЕНЬШИТЬ КОЭФФИЦИЕНТ ТРЕНИЯ ДРУГИХ МАТЕРИАЛОВ //Scientific progress. – 2022. – Т. 3. – №. 2. – С. 1043-1048.
14. Mamirov A. R., Rubidinov S. G., Gayratov J. G. Influence and Effectiveness of Lubricants on Friction on the Surface of Materials //CENTRAL ASIAN JOURNAL OF THEORETICAL & APPLIED SCIENCES. – 2022. – Т. 3. – №. 4. – С. 83-89.
15. Шохрух Г. У. Р., Ғайратов Ж. Г. У. Анализ теории разъемов, используемых в процессе подключения радиаторов автомобиля //Science and Education. – 2022. – Т. 3. – №. 9. – С. 162-167.
16. Шохрух Г. У. Р., Ғайратов Ж. Г. У. Анализ технологической системы обработки рабочих поверхностей деталей вала на токарном станках //Science and Education. – 2022. – Т. 3. – №. 8. – С. 23-29.