

УДК 004.8

АНАЛИЗ ТЕХНОЛОГИЙ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ В МАШИНОСТРОЕНИИ

Киренберг Е.А. , магистрант гр. ИИМ-241, II курс

Научный руководитель: Ванеев О.Н., к.т.н., доцент

Кузбасский государственный технический университет им. Т. Ф. Горбачева,
г. Кемерово

Компьютерное зрение, как область искусственного интеллекта, фокусирующаяся на обработке и анализе визуальной информации, позволяет решать широкий спектр задач, традиционно выполняемых человеком [1]. В этом контексте компьютерное зрение, как ключевая технология искусственного интеллекта, представляющей собой комплекс методов обработки и анализа визуальной информации, получает все более широкое распространение в машиностроении [2].

Современное состояние применения компьютерного зрения в машиностроении демонстрирует как существенные достижения, так и нерешенные проблемы [3]. Уже успешно внедрены системы автоматического контроля качества на основе анализа изображений, системы навигации промышленных роботов, а также решения для мониторинга состояния оборудования. Однако остаются сложности, связанные с обработкой сложных сцен в условиях меняющегося освещения, наличия бликов и теней, а также с необходимостью адаптации систем к новым типам продукции и дефектов.

В зависимости от поставленной задачи в машиностроении используются разные технологические подходы. Классические методы компьютерного зрения базируются на алгоритмической обработке изображений без использования глубокого обучения. К ним относятся: Методы фильтрация (Гауссовский, медианный) [4], преобразования (Фурье, Уолша-Адамара), морфологические операции. SIFT, SURF, ORB для обнаружения уникальных точек на изображении. Пороговая сегментация, watershed, GrabCut для разделения изображения на значимые области [5]. SVM, Random Forest, k-NN для классификации объектов на основе извлеченных признаков [6].

Современные методы компьютерного зрения основаны чаще всего на глубоком обучении, в частности на сверточных нейронных сетях. Их эффективность достигается за счёт комбинации нескольких специализированных технологий [7]: Сверточные слои применяют набор фильтров (ядра) к входному тензору, извлекая локальные признаки. Пулинг (Max Pooling, Average Pooling, Strided Convolution) уменьшает пространственного разрешения. Функции активации (ReLU, Leaky ReLU / ELU / Swish) вводит нелинейность. Нормализация (Batch, Layer, Group

Normalization) нормализует активации внутри слоя, действует как регуляризатор.

Можно выделить несколько классов проблем, решаемых с помощью компьютерного зрения: классификация, детекция, сегментация и оценка позы. Каждый из этих классов опирается на специфические архитектуры нейронных сетей и методы обработки изображений, что определяет их применимость для конкретных технологических операций в машиностроении.

Классификация изображений: Отнесение всего входного изображения или его фрагмента к одному из заданных классов без указания местоположения объекта. Задачи классификации решают модели: ResNet, EfficientNet, ViT-Base и другие. Областью применения может быть, например: определение типа детали, оценка качества поверхности, распознавание дефектов, оценка износа режущего инструмента.

Детекция объектов: Локализация объектов на изображении с помощью ограничивающих рамок и одновременное определение их классов. Детекторы разделяются на одностадийные (YOLO, SSD) и двухстадийные (Faster R-CNN). Одностадийные быстрее и подходят для реального времени, двухстадийные - точнее, но медленнее. Примеры применения: локализация заусенцев на кромках штампованных деталей для автоматической подачи координат удаления фрезером, отслеживать положение работников вблизи опасных зон или движущихся частей оборудования, активируя сигнализацию.

Сегментация: Пиксельная классификация изображения. Различают семантическую (деление на классы без различия экземпляров) и инстанс-сегментацию (выделение каждого отдельного объекта). Сегментация предоставляет более детальную информацию, чем детекция, так как позволяет точно определить границы объектов и их внутреннюю структуру. Примеры: U-Net - архитектура, изначально разработанная для медицинской сегментации, но нашедшая широкое применение в промышленности. Mask R-CNN - Расширение Faster R-CNN, предсказывающее маску для каждого обнаруженного объекта. Примеры применения: получение точного контура для расчёта размеров, контроль геометрии зубчатых передач, сегментация локальных зон перегрева.

Оценка позы: Определение пространственного положения и ориентации (6 степеней свободы: 3 трансляции + 3 вращения) объекта или его частей. Примеры: PVNet — использует голосование векторов для локализации ключевых точек, устойчив к сильным окклюзиям. GDR-Net - восстанавливает позу через промежуточное представление карт вращения и сдвига. Например, определение положения и ориентацию объектов в рабочей области может позволить роботам-манипуляторам точно захватывать случайно ориентированные объекты.

Рассмотрим более подробно направление по распознаванию дефектов на металлических поверхностях после механической обработки методом классификации. При ручной обработке оператор станка может производить быстрый визуальный контроль качества выполненной операции, тогда как при автоматизированной обработки в условиях ГПС такая возможность

отсутствует. Требуется или перенос заготовки на позицию контроля между операциями или отказ от межоперационного контроля, когда производится только один итоговый контроль после последней операции обработки.

Недостатками первого решения является увеличение затрат времени на транспортировку заготовок, а также необходимость в большем количестве позиций контроля. Недостатками второго решения являются возможные ситуации, когда на одной из ранних операций образуется дефект, который не может быть обнаружен до итогового контроля. В этом случае применение компьютерного зрения для межоперационного контроля может оказаться весьма эффективно, так как почти не требует времени.

Разберем пример использования сверточной нейронной сети для задачи обнаружения и распознавания дефектов. Был использован датасет GC10-DET [8], содержащий изображения металлических поверхностей со складками, включениями, ямками и другими дефектами. Всего 2306 черно-белых изображения, 2280 из которых содержат один или несколько из 10 возможных типов дефектов. Примеры изображений приведены на рис. 1. Рамки с подписями отображены для наглядности. В обучении используются чистые изображения с информацией о наличии каждого из типов дефектов. В результате обучаемая модель классификации сможет определить только наличие или отсутствие каждого из типов дефекта, без указания места.

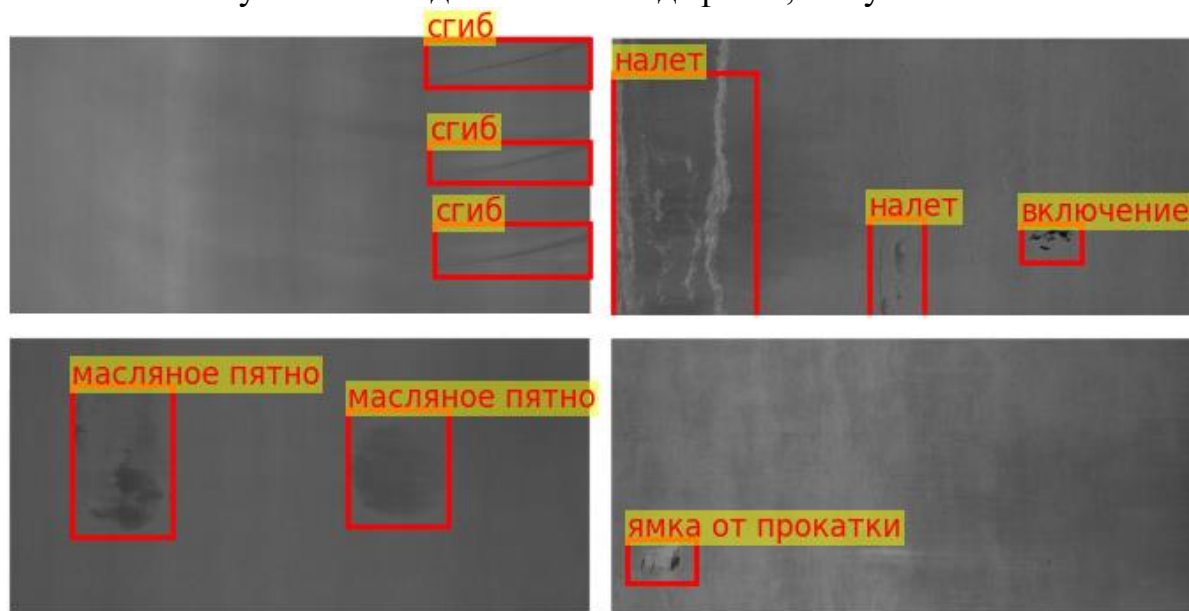


Рис. 1. Примеры из набора данных GC10-DET

Применяется модель VGG-16, предобученная на датасете ImageNet [9]. Архитектурно VGG-16 состоит из двух основных частей: извлекатель признаков и классификатор. Извлекатель признаков состоит из чередующихся сверточных слоев, функций активации, а также слоев субдискретизации. Классификатор состоит из полносвязных слоев, функций активации и функции случайного чередования нейронов. Всего 31 слоев в извлекателе признаков и 7 в классификаторе.

Чтобы сохранить способность переобученной модели выделять общие признаки на изображении, которые будет трудно получить на ограниченном датасете, а также сократить затраты вычислительных ресурсов, перед обучением замораживаются первые 19 слоев извлекателя признаков. Также заменяется классификатор на новый, структура выходных данных которого соответствует детектируемым признакам в датасете.

Пример кода для определения модели и функции потерь, а также запуска цикла обучения:

```
vgg = models.vgg16(pretrained=True)
for i, ch in enumerate(vgg.features.children()):
    if i < 19:
        for p in ch.parameters():
            p.requires_grad = False

vgg.classifier = nn.Sequential(
    nn.Linear(25088, 4096), nn.ReLU(), nn.Dropout(0.5),
    nn.Linear(4096, 512), nn.ReLU(), nn.Dropout(0.5),
    nn.Linear(512, 8)).to(device)

def one_hot_encoder(labels):
    return torch.Tensor([[1 if i == label else 0 for i in range(8)] for label in labels])

def test_loss(model, loss):
    with torch.no_grad():
        images, labels = next(iter(valloader))
        images = images.to(device)
        oh_labels = one_hot_encoder(labels).to(device)
        return loss(model(images), oh_labels)

criterion = nn.CrossEntropyLoss()
optimizer = optim.SGD(vgg_16_network.parameters(), lr = 0.001, momentum=0.9,
weight_decay=0.0005)

def train(model, optimizer, criterion, epochs, load=False, save=False):
    losses, test_losses = [], []
    for epoch in range(epochs):
        td = tqdm(trainloader)
        for i, (images, labels) in enumerate(td):
            images = images.to(device)
            loss = criterion(model(images), one_hot_encoder(labels).to(device))
            optimizer.zero_grad()
            loss.backward()
            td.set_description(f'Loss: {loss.item()}')
```

```

optimizer.step()
if i % 50 == 0:
    tl = test_loss(model, criterion).to("cpu")
    losses.append(loss.item())
    test_losses.append(tl)
print(f'Epoch: {epoch}, Training Loss: {loss}, Val Loss: {tl}')
if (epoch + 1) % 3 == 0:
    for g in optimizer.param_groups: g['lr'] /= 2
    if save: torch.save(model.state_dict(), f'vgg16_benchmark{epoch+1}.pth')
return losses, test_losses

```

```

losses, test_losses = train(vgg_16_network, optimizer, criterion, 10, load=False,
save=True)

```

При проверке обученной модели на тестовой части данных, которая не использовалась в процессе обучения, средняя точность классификации составила 96,14 %. Графики изменения потерь при обучении и валидации в процессе обучения приведены на рис. 2.

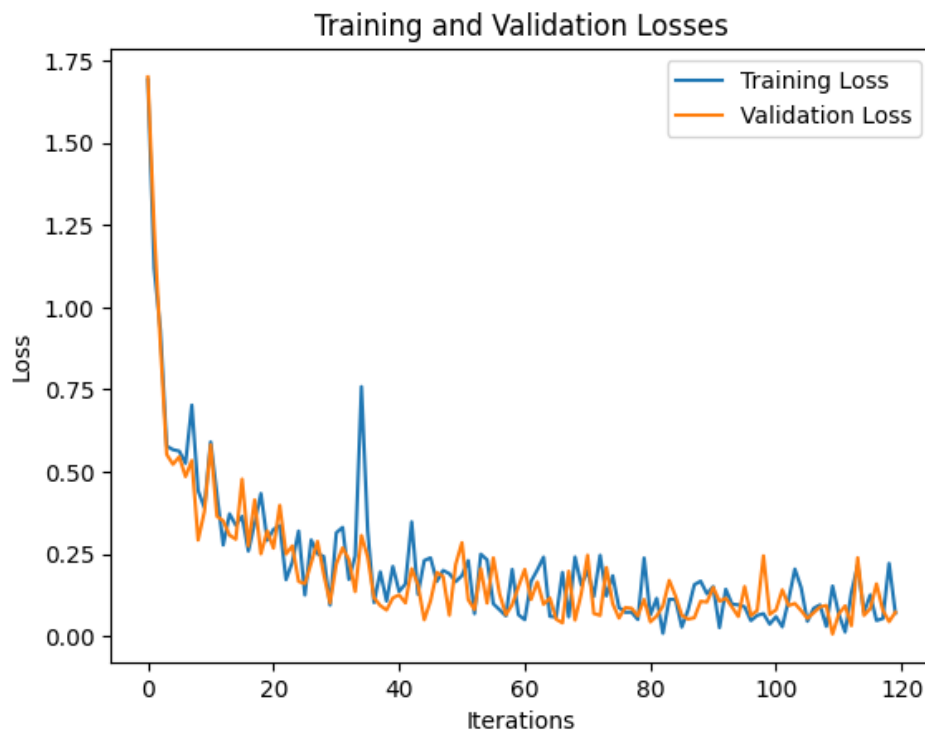


Рис. 2. Визуализация процесса обучения

При условии минимальных затрат времени на проведение контроля, модель компьютерного зрения показывает хорошие результаты определения дефектов поверхности, являясь полезным инструментом в условиях автоматизированных производственных линий. Дальнейшее развитие в области глубокого обучения и нейросетевых архитектур открывает новые перспективы для внедрения компьютерного зрения в промышленные

процессы, что позволит повысить эффективность производства и снизить издержки.

Список источников

1. Сидоров, П. А. Применение компьютерного зрения в роботизированных системах для автоматизации процессов / П. А. Сидоров // Символ науки: международный научный журнал. – 2025. – № 4-2. – С. 62-71. – EDN HJJYBL.
2. Ribeiro J., Lima R., Eckhardt T., Paiva S. Robotic Process Automation and Artificial Intelligence in Industry 4.0 – A Literature review // Procedia Computer Science. 2021. Т. 181. С. 51–58. DOI: 10.1016/j.procs.2021.01.104.
3. Москат, Н. А. Глубокое обучение и компьютерное зрение: достижения, проблемы и перспективы / Н. А. Москат, Е. С. Кручинин // Цифровые инфокоммуникационные технологии : СБОРНИК НАУЧНЫХ ТРУДОВ, Ростов-на-Дону, 27 октября 2023 года. – Ростов-на-Дону: Ростовский государственный университет путей сообщения, 2023. – С. 276-280. – EDN DUDPXW.
4. Петров, Е. П. Метод комбинированной нелинейной фильтрации коррелированных видеоизображений / Е. П. Петров, Е. В. Медведева, А. П. Метелев // Нелинейный мир. – 2010. – Т. 8, № 11. – С. 677-684. – EDN OOUUAV.
5. Об ограничениях уровнево-пороговой сегментации полутоновых растровых изображений / А. В. Григорьев, А. Л. Држевецкий, В. Я. Баннов [и др.] // Труды международного симпозиума "Надежность и качество". – 2014. – Т. 2. – С. 18-21. – EDN SNLWQT.
6. Демидова, Л. А. Использование SVM-алгоритма для уточнения решения задачи классификации объектов с применением алгоритмов кластеризации / Л. А. Демидова, Ю. С. Соколова // Вестник Рязанского государственного радиотехнического университета. – 2015. – № 51. – С. 103-113. – EDN TTWDTR.
7. Kutuzova, M. O. Research On The Architecture Of A Convolutional Neural Network For Image Recognition / M. O. Kutuzova, D. A. Krutko // Молодежь. Общество. Современная наука, техника и инновации. – 2021. – No. 20. – P. 39-42. – EDN OERORH.
8. Deep Metallic Surface Defect Detection: The New Benchmark and Detection Network / X. Lv, F. Duan, J. J. Jiang [et al.] // Sensors. – 2020. – Vol. 20, No. 6. – P. 1562. – DOI 10.3390/s20061562. – EDN LXFMRZ.
9. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv.org. – 2015. – 10 Apr. – DOI:10.48550/arXiv.1409.1556.