

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АРХИТЕКТУР НЕЙРОННЫХ СЕТЕЙ ДЛЯ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ ПРЕДПРИЯТИЯ

Нарышкин И.Ю., Печенина Е.Ю., Печенин В.А.
Самарский университет, г. Самара
v.a.pechenin@ssau.ru

Ключевые слова: классификация, stl, сверточная нейронная сеть, техническое зрение.

На промышленных предприятиях стоит задача повышения эффективности учета и движения заготовок, полуфабрикатов и деталей. В настоящее время самое эффективное решение – системы технического зрения. Ключевой частью подобных систем является эффективный алгоритм детектирования интересующих предприятие объектов. В работе делается сравнительный анализ нейронных сетей для классификации объектов (определение типа объекта детектирования) по результатам фото- и видеосъемки.

Для классификации использовалось пять видов деталей, в том числе рабочая лопатка турбины. Основную трудность при внедрении алгоритмов детектирования в системах технического зрения составляет сбор необходимого количества данных для обучения (минимум тысячи и десятки тысяч экземпляров). На основании STL-моделей (рис. 1, а), полученных с использованием результатов сканирования (рис. 1, б), было выполнено моделирование обучающей выборки из 8400 экземпляров, тестовой выборки из 3600 экземпляров и проверочной выборки из 1200 экземпляров.

Все экземпляры сохранялись в изображения. Размер всех изображений был приведен к 112×112 пикселей в оттенках серого. Таким образом, вектором признаков являлись значения интенсивности пикселей в матрицах растровых изображений.

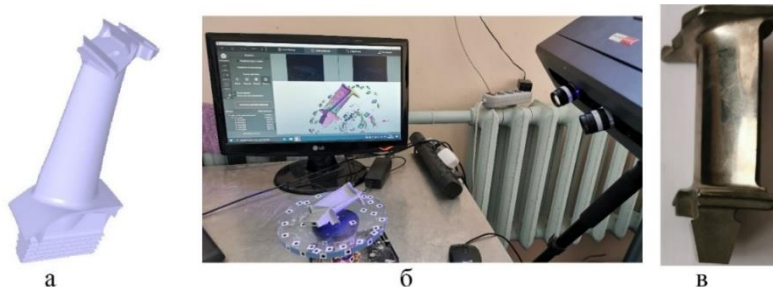


Рис. 1 – Данные для классификации: а – STL-модель; б – сканирование; в – фотография лопатки

Для решения задачи классификации использовалось два варианта архитектуры нейронных сетей: с полносвязными слоями, когда на вход подается вектор значений; и реализация сети из семейства VGG [1], принимающая на вход матрицы изображения.

В качестве функции потерь L была использована категориальная перекрестная энтропия; точность определяется долей $\delta_{\text{клас}}$ правильно классифицированных объектов в общем количестве объектов выборки.

У сетей изменялись параметры для поиска лучшей точности. В первой архитектуре изменяемыми параметрами являлись: количество внутренних слоев (2 и 4); количество эпох обучения N_s ; алгоритм обучения нейронной сети (стохастический градиентный спуск (SGD) [2] и Adam [3]). Для второй архитектуры вместо количества внутренних слоев изменялся размер «пакетов» (маленькие партии данных из общей выборки, одновременно подаваемые для прохождения обучения).

Были выбраны оптимальные, с точки зрения величины функции потерь и точности, параметры сетей, значения величин $\delta_{\text{клас}}$ и L для трех выборок приведены в табл. 1.

Приемлемой точностью при обучении и проверке обладает сеть VGG, при этом в проверочной выборке неверно классифицируются лишь 0,42% данных. Скорость роста

точности и снижения потерь первой нейронной сети крайне низкая; сверточная нейронная сеть быстро увеличивает точность по эпохам и достигает практически 100% точности. Стоит отметить серьезное рассеивание точности в процессе обучения для тестовой выборки, которое к окончанию обучения становится приемлемым (доли процента).

Таблица 1 – Величины $\delta_{\text{клас}}$ и L для всех выборок данных

Первая архитектура нейронной сети								
Число слоев	Число эпох	Алгоритм обучения	Обучающая выборка		Тестовая выборка		Проверочная выборка	
			L	$\delta_{\text{клас}}$	L	$\delta_{\text{клас}}$	L	$\delta_{\text{клас}}$
2	100	SGD	1,1800	0,5384	1,1843	0,5356	1,2192	0,4733
Вторая архитектура нейронной сети (VGG)								
Размер «пакета»	Число эпох	Алгоритм обучения	Обучающая выборка		Тестовая выборка		Проверочная выборка	
			L	$\delta_{\text{клас}}$	L	$\delta_{\text{клас}}$	L	$\delta_{\text{клас}}$
10	50	Adam	0,0004	0,9999	0,0143	0,9978	0,0074	0,9958

Была выполнена классификация деталей по их реальным фотографиям с использованием обученной сети VGG. Всего было сделано 18 фотографий лопатки турбины с разных ракурсов (рис. 1, в). Точность классификации $\delta_{\text{клас}}$ реальных фотографий составила 0,78. Используя предложенную в статье [4] методику подачи данных не в виде одной картинке, а в виде трех картинок-проекций и, обучив таким образом сеть VGG, удалось достигнуть точности классификации $\delta_{\text{клас}}$ фотографий в 0,83. Дальнейшим направлением исследований является предварительная обработка изображений, поступающих на вход в нейронную сеть, для увеличения сходства между обучающими данными и реальными фотографиями, а так же использование нейронных сетей для детектирования разных деталей на одной фотографии.

Список литературы

1. Rosebrock A. Deep Learning for Computer Vision with Python. PyImageSearch.com, 2017, 330 p.
2. Trask A. Grokking Deep Learning. Manning Publications, 2019. 335 p.
3. Kingma D.P., Ba J.L. Adam: A Method for Stochastic Optimization // 3rd International Conference on Learning Representations, ICLR 2015. Conference Track Proceedings. 2015. P. 149801.
4. Pechenin, V.A., Bolotov, M.A., Pechenina, E.Yu. Neural network model of machine parts classification by optical scanning results // Journal of Physics: Conference Series. 2020, V. 1515, no 5. P. 052008.

Сведения об авторах

Нарышкин Илья Юрьевич, магистрант гр. 2121-240405D Самарского университета. Область научных интересов: САЕ-расчеты, технологии сборки.

Печенина Екатерина Юрьевна, аспирант кафедры ТПД Самарского университета. Область научных интересов: координатные измерения, балансировка, математические методы, машинное обучение.

Печенин Вадим Андреевич, канд. техн. наук, доцент кафедры ТПД Самарского университета. Область научных интересов: координатные измерения, процессы сборки, машинное обучение.

COMPARATIVE ANALYSIS OF NEURAL NETWORKS FOR INDUSTRY MACHINE VISION SYSTEM

Naryshkin I.Yu., Pechenina E.Yu., Pechenin V.A.
Samara National Research University, Samara

The paper compares the accuracy parameters of two neural network architectures that can be used in industries machine vision systems. The first architecture is a deep neural network without convolutional layers. The second is a convolutional network of the VGG family. Both networks have the ReLu activation function, the output is using the SoftMax layer. The VGG architecture showed the best applicability for the problem being solved.