

ПРОГНОЗИРОВАНИЕ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

© 2024 К. С. Макаров¹, Д. Н. Дронина²

¹*кандидат технических наук, доцент кафедры программного обеспечения
и администрирования информационных систем
e-mail: runaway90@mail.ru*

²*студентка 2 курса магистратуры направления подготовки
«Информатика и вычислительная техника»
e-mail: dn.dronina@yandex.ru*

Курский государственный университет

Нейронные сети широко применяются при решении задач моделирования производственных процессов, в том числе в прогнозировании спроса, оптимизации производства, распознавании дефектов и др. В работе представлены возможности применения нейронных сетей для построения модели обнаружения дефектов на деталях. Описан эксперимент по построению моделей прогнозирования с помощью аппарата нейронных сетей, приведены его результаты. Отмечены достоинства и недостатки обучаемых моделей для решения поставленной задачи.

Ключевые слова: прогнозирование брака продукции, нейронные сети, сверточная нейронная сеть, классификация изображений

FORECASTING PRODUCTION PROCESSES USING NEURAL NETWORKS

© 2024 K. S. Makarov¹, D. N. Dronina²

¹*Candidate of Engineering Sciences,
Associate Professor, Department of Software and Administration Information Systems
e-mail: runaway90@mail.ru*

²*2nd year master's student in the field of study
"Informatics and Computer Science"
e-mail: dn.dronina@yandex.ru*

Kursk State University

Neural networks are widely used in solving problems of modeling production processes, including demand forecasting, production optimization, defect recognition, etc. The paper presents the possibilities of using neural networks to build a model for detecting defects on parts. An experiment on constructing forecasting models using neural networks is described, and its results are presented. The advantages and disadvantages of trained models for solving the problem are noted.

Keywords: product defect prediction, neural networks, convolutional neural network, image classification.

По мере совершенствования технологий производство становится все более эффективным и автоматизированным, а использование нейронных сетей для моделирования производственных процессов приобретает все большее значение.

Нейронные сети представляют собой математические модели, которые способны обучаться на основе имеющихся данных и делать предсказания на основе этого обучения [9].

Нейронные сети в моделировании производственных процессов можно применять в следующих основных направлениях:

1) для прогнозирования результатов производства на основе анализа имеющихся данных;

2) для оптимизации параметров производственных процессов. Например, нейронные сети могут оптимизировать распределение ресурсов или настройку оборудования для достижения максимальной эффективности процесса, анализировать имеющиеся данные, исследовать различные варианты и предлагать наилучшие параметры для оптимизации производства;

3) для обнаружения сбоев в производственных процессах, при анализе данных и выявлении отклонений от нормы. Это позволяет заранее выявлять возможные проблемы и предотвращать их возникновение;

4) для составления оптимального графика работы производственной линии и планирования производственных заданий. Они способны рассчитать оптимальное время начала и окончания каждой задачи, а также распределить ресурсы таким образом, чтобы максимально повысить производительность и сократить время простоев [3].

При прогнозировании производства промышленных деталей на основе машинного обучения значительную часть работы занимает подготовка данных и обучение модели. В данной статье предлагается подход к обнаружению дефектов на деталях и сборочных единицах (ДСЕ) на основе машинного обучения. Необходимо получить нейронную сеть с минимальными ошибками предсказания, валидационные потери которой не должны превышать 15%.

В качестве подхода к решению задачи обнаружения дефектов на деталях можно предложить использование моделей машинного обучения, которые будут обучаться для последующего выявления брака. Результатирующими данными в данном случае будет информация о наличии дефекта на детали.

Для демонстрации применения нейронных сетей для обнаружения деталей с дефектом рассмотрим пример решения задачи обнаружения брака у ДСЕ «Стойка правая ГЖИК.745223.001». Примеры ДСЕ с дефектом и исправной ДСЕ представлены на рисунках 1 и 2 соответственно.



Рис. 1. ДСЕ «Стойка правая ГЖИК.745223.001» с дефектом



Рис. 2. Исправная ДСЕ «Стойка правая ГЖИК.745223.001»

Исправные детали от деталей с дефектами отличаются равномерным покрытием, отсутствием гнутых ребер.

Набор данных для обучения состоит из 75 изображений корректных деталей и 19 изображений деталей с браком размерностью 224x224. В ходе предобработки набор данных для обучения был увеличен путем аугментации.

Для решения задач классификации изображений, к которым относится задача обнаружения ДСЕ с браком, могут использоваться различные алгоритмы, однако наибольшую эффективность показываются нейронные сети [1].

Сверточные нейронные сети применяются для классификации изображений и способны распознавать различные объекты. Они используют особенности зрительной системы, включая простые элементы и более сложные элементы, способные распознавать сложные объекты [7]. Архитектура сверточной нейронной сети включает следующие основные элементы.

- Входной слой (Input layer) – это первый слой сети, принимающий входные данные. В данном случае входными данными является изображение детали.
- Сверточный слой (Convolutional layer) – это слой, который выполняет основную функцию сверточной нейронной сети – распознавание образов. Здесь применяются фильтры для свертки изображения и извлечения из него значимых признаков. Каждый фильтр представляет собой матрицу, которая перемещается по изображению и вычисляет скалярное произведение с соответствующими участками изображения. Полученный результат является значением одного канала выходного изображения.
- Функция активации (Activation function) применяется к каждому каналу сверточного слоя для придания нелинейных свойств модели. Обычно для этой цели используется функция ReLU (Rectified Linear Unit) или сигмоида.
- Пулинг (Pooling) уменьшает размерность изображения для уменьшения вычислительной сложности и устранения избыточности. Он отбрасывает часть информации, сохраняя только наиболее значимые признаки.
- Полносвязный слой (Fully connected layer) представляет собой слой нейронов, каждый из которых соединен со всеми нейронами предыдущего слоя. Этот слой используется для классификации входных данных.

Для исследования были выбраны модели ResNet34 и VGG19, которые в последние годы показали себя как одни из наиболее эффективных при решении задач классификации изображений [8].

Для реализации процесса обучения нейронных сетей ResNet34 и VGG19 и оценки его эффективности были использованы фреймворки машинного обучения

PyTorch и TensorFlow соответственно. Данные библиотеки включают в себя множество различных алгоритмов и моделей, позволяя пользователю реализовать глубокие нейронные сети для использования в таких задачах, как обработка естественного языка, распознавание и классификация изображений [6].

При решении задач машинного обучения необходимо обратить внимание на выбор метрики для оценки нейронных сетей. Наиболее часто используются следующие метрики: отзыв, точность, ROC-кривая, F-мера, точность на уровне k и средняя абсолютная ошибка. Выбор метрики должен быть основан на том, какие показатели наиболее важны для данной задачи [5].

С помощью нейросетевых моделей ResNet34 и VGG19 обучим алгоритмы для классификации изображений, а также проведем анализ полученных результатов с помощью метрик [8]. На выходе будут метрики и посчитанное обратное распространение ошибки сверточных нейронных сетей VGG19 и ResNet34 через категориальную кросс-энтропию. Для ResNet34 также будет использоваться метрика F1-score.

Кросс-энтропийная функция обладает рядом важных свойств, которые делают ее привлекательной и эффективной: симметричность, неотрицательность, информационная емкость, чувствительность к вероятностям, расширение на многоклассовую классификацию. Метрика F1-score сочетает в себе точность и полноту, то есть чувствительность модели к положительным классам. Ее преимущество заключается в компромиссе между точностью и полнотой, простоте интерпретации, нечувствительность к дисбалансу классов [7].

Важно также обратить внимание на эпохи, так как по определению это количество циклов обучения сети на полном наборе данных. Чем больше эпох, тем больше вероятность того, что сеть обучится полностью [4]. Однако слишком большое количество эпох может привести к переобучению сети и снижению ее обобщающей способности. Обычно количество эпох выбирается экспериментально, исходя из специфики задачи и набора данных.

Оптимизаторы – это алгоритмы или методы, используемые для корректировки параметров нейронной сети (таких как веса и величина шага) с целью уменьшения величины потерь. Для модели VGG19 применялся Adam, так как это алгоритм адаптивной оптимизации, который был создан специально для глубокого обучения нейронной сети. Также для модели ResNet34 применен оптимизатор SGD, так как он позволяет минимизировать функцию ошибки модели путем изменения весов связей между нейронами. SGD работает путем вычисления градиента функции ошибки на каждом обучающем примере, а затем обновления весов сети в направлении, обратном к этому градиенту. Для моделей VGG19 оптимальной была скорость обучения 0,001, а для модели ResNet34 – 0,0001.

На рисунке 3 представлен процесс обучения нейронной сети VGG19 длительностью 5 эпох. Анализ графиков показывает, что точность обучения близится к 80%, а валидационная точность также не превышает 80%, то есть данное количество применяемых эпох не подходит для моделирования выявления брака. Для того чтобы улучшить работу данной модели, необходимо увеличить количество обучаемых эпох, чтобы получить необходимые результаты точности.

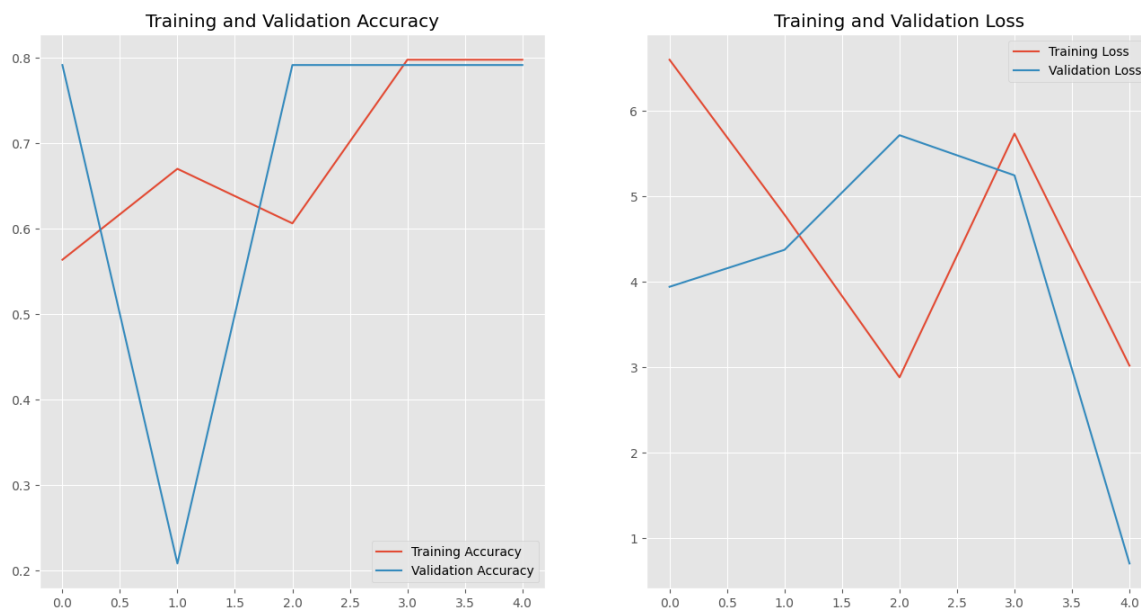


Рис. 3. Процесс обучения нейронной сети VGG19 длительностью 5 эпох

В таблице 1 отображены результаты обучения нейросетевой модели VGG19 на разном количестве эпох.

Таблица 1

Результаты обучения нейросетевой модели VGG19 на разном количестве эпох

Количество эпох обучения	Training Accuracy (%)	Validation Accuracy (%)	Training Loss (%)	Validation Loss (%)
5	80	78.5	30	50
10	99	87	50	70
20	99	88	10	20
25	98	84	10	6

Как видно из приведенных выше результатов, наилучшая точность результатов достигается на 20 эпохах обучения, при этом последующие увеличение эпох приводит к переобучению модели.

Из приведенных исследований модели VGG19 можно сделать вывод, что результаты обучения довольно высокие, но они не являются достаточными для данной задачи, поэтому воспользуемся другой моделью нейронной сети.

Для дальнейшего исследования обратимся к ResNet34 (Residual Network), так как это архитектура сверточной нейронной сети, разработанная специально для решения задач классификации изображений. Данная модель представляет собой усовершенствованную версию классической архитектуры CNN и использует остаточные связи (residual connections) для обучения более глубоких сетей [7].

ResNet34 имеет 34 слоя и состоит из нескольких блоков, каждый из которых содержит несколько сверточных слоев с различными размерами фильтров и остаточных связей. Остаточные связи помогают избежать проблемы исчезающего градиента, которая обычно возникает при обучении глубоких нейронных сетей с большим числом слоев. В результате ResNet34 может достичь более высокой точности классификации по сравнению с классическими архитектурами CNN [8].

Таблица 2

Результаты обучения нейросетевой модели ResNet34 на разном количестве эпох

Количество эпох обучения	Training Accuracy (%)	Validation Accuracy (%)	Training Loss (%)	Validation Loss (%)
5	85	83	30	20
10	87	85	20	20
20	90	87	10	19
22	98	88	9	15
25	99	88	8	13

Из таблицы 2 видно, что при 25 эпохах валидационная точность не превышает 88%, а валидационные потери составляют 13%.

На рисунке 4 отображены графики процессы обучения сети на 25 эпохах.

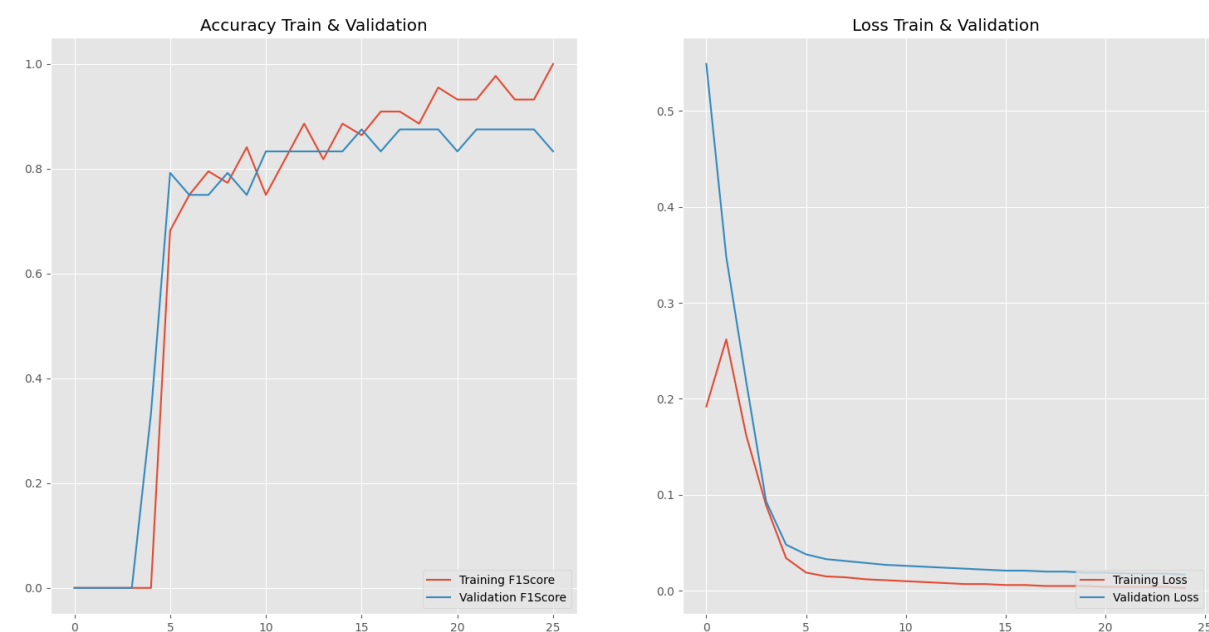


Рис. 4. Процесс обучения нейронной сети ResNet34 длительностью 25 эпох

Исходя из представленных выше результатов обучения двух нейросетевых моделей, можно сделать следующие выводы:

- для получения высоких результатов обучения необходимо в среднем 20–25 эпох для обеих моделей;
- валидационная точность используемых моделей примерно одинакова, но потери при обучении и валидации у ResNet34 меньше, чем у VGG19.

Таким образом, для решения задачи обнаружения дефектов на ДСЕ «Стойка правая ГЖИК.745223.001» предпочтительным является использование нейронной сети ResNet34.

В качестве дальнейшего направления исследований можно отметить построение на основе полученных результатов прогнозной модели производства деталей.

Библиографический список

1. Васкан, В. Д. Обзор архитектур сверточных нейронных сетей для задачи классификации изображений // В. Д. Васкан // Электронный научный журнал «ИТ-Стандарт». – URL: <https://itstd-journal.ru/wp-content/uploads/2021/06/OVERVIEW-OF->

[CONVOLUTIONAL-NEURAL-NETWORK-ARCHITECTURES-FOR-THE-IMAGE-CLASSIFICATION-PROBLEM.pdf](#) (дата обращения: 12.12.2023).

2. Кондаков, А. И. Разработка маршрутных технологических процессов изготовления деталей / А. И. Кондаков, Ю. А. Островский. – Москва: Издательство МГТУ, 2003.

3. Круглов, В. В. Искусственные нейронные сети. Теория и практика / В. В. Круглов, В. В. Борисов. – Москва: Горячая линия – Телеком, 2019.

4. Ларионов, Р. В. Использование изображений с БПЛА для обнаружения незаконных свалок твердых бытовых отходов / Р. В. Ларионов, В. А. Павлов, В. В. Хрящев // Цифровая обработка сигналов и ее применение (DSPA-2021): доклады 23-й международной конференции. – Москва, 2021. – С. 209–213.

5. Маленков, Ю. А. Возможности применения цифровых технологий в стратегическом планировании и прогнозировании устойчивого развития организаций / Ю. А. Маленков, А. Э. Давыдова // Молодой ученый. – 2021. – № 14 (356). – С. 84–89.

6. Нейронные сети в прикладной экономике: [учебное пособие] / Е. А. Трофимова, Вл. Д. Мазуров, Д. В. Гилёв; [под общей редакцией Е. А. Трофимовой]; М-во образования и науки Рос. Федерации, Урал. федер. ун-т. – Екатеринбург: Издательство Урал. ун-та, 2017.

7. Омаров, Т. З. Концепция искусственной нейронной сети / Т. З. Омаров // Современные научные исследования и инновации. Электронный научно-практический журнал. – 2016. – № 5. – URL: <https://web.snauka.ru/issues/2016/05/66203> (дата обращения: 12.12.2023).

8. Хе, Тонг и пр. Мешок трюков для классификации изображений с помощью сверточных нейронных сетей //Материалы конференции IEEE по компьютерному зрению и распознаванию шаблонов. – 2019.

9. Чугреев, В. Л. Использование прогнозной аналитики в информационно-аналитических системах поддержки принятия решений / В. Л. Чугреев, Д. А. Баданин // Молодой ученый. – 2016. – № 6. – С. 49–52.