

Р. М. Новиков^{1}, Д. Н. Титов¹, Е. В. Рыжкова¹*

Сравнительный анализ эффективности нейросетевых алгоритмов в распознавании дефектов на изображениях

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация
* e-mail: romanov7@live.ru

Аннотация. Современные тенденции к модернизации производства невозможны без применения новых возможностей и технологий, позволяющих повысить скорость распознавания дефектов и принятия решений. Исследование посвящено сравнительному анализу эффективности нейросетевых алгоритмов в задаче распознавания дефектов на изображениях. Эксперимент включает использование различных алгоритмов, позволяющих проанализировать их работу с помощью анализа графиков, процентов успеха. Основным методом обучения - Transfer Learning, позволяет использовать предварительно обученную модель (VGG16) и дообучить ее на собственных данных. Модель эффективно решает задачу обнаружения дефектов в виде трещин на изображениях, достигая высокой точности классификации, до 91,33 %, при распознавании класса изображения.

Ключевые слова: искусственный интеллект, сверточные нейронные сети, машинное зрение, глубокое обучение, повышение качества

R. M. Novikov^{1}, D. N. Titov¹, E. V. Ryzhkova¹*

Comparative analysis of the effectiveness of neural network algorithms in detecting defects in images

¹ Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
* e-mail: romanov7@live.ru

Abstract. Modern tendencies towards modernization of production are impossible without application of new capabilities and technologies that allow to increase the speed of defect detection. capabilities and technologies allowing to increase the speed of defect recognition and decision making. The study is devoted to a comparative analysis of the effectiveness neural network algorithms in the task of recognizing defects in images. Experiment, includes the use of different algorithms, allowing us to analyze their performance by analyzing graphs, success percentages. The main method of training is. Transfer Learning, allows the use of a pre-trained model (VGG16) and further train it on its own data. The model effectively solves the problem of detection of defects in the form of cracks in the images, achieving high accuracy in of defect classification.

Keywords: Artificial intelligence, convolutional neural networks, machine vision, deep learning, quality improvement.

Введение

Обработка изображений становится неотъемлемой частью многих промышленных процессов, особенно в контексте контроля качества продукции. С увеличением сложности производственных процессов и требований к качеству встает

вопрос о эффективных методах распознавания дефектов на изображениях продукции. Нейросетевые алгоритмы представляют собой мощный инструмент для решения таких задач, благодаря их способности к обучению на больших объемах данных и выявлению сложных дефектов [1]. В статье приводятся результаты сравнительного анализа эффективности различных нейросетевых алгоритмов для распознавания дефектов на изображениях. Целью исследования является выбор наиболее эффективного алгоритма распознавания дефектов на изображениях для магистерской работы «Разработка системы размерного контроля на производстве с использованием нейросетевых технологий».

Методы и методики исследования

В качестве основного подхода используется Transfer Learning – этот подход, позволяет использовать предварительно обученную модель (VGG16) и дообучить ее на наших данных. Это особенно полезно при недостаточном количестве данных [2].

Данные для обучения и тестирования модели были собраны из двух директорий, содержащих положительные и отрицательные образцы. Всего было использовано 5200 изображений, из которых 60 % использованы для обучения, а остальные 40 % - для тестирования. Данные были случайно перемешаны перед разделением.

Для оценки эффективности модели использовался тестовый набор данных, состоящий из 100 изображений поверхностей. На этих изображениях были трещины (метка «POSITIVE») или они отсутствовали (метка «NEGATIVE»). Все изображения были предварительно обработаны и уменьшены до размера 120x120 пикселей.

Модель обучалась с использованием оптимизатора Adam и бинарной функции потерь на основе кросс-энтропии, с применением метода ранней остановки для предотвращения переобучения [3].

Алгоритм 1. Применения сверточных нейронных сетей для обнаружения трещин на изображениях

В данном примере исследовано применение сверточных нейронных сетей Convolutional Neural Networks (CNN). CNN являются стандартом в области анализа изображений и показывают высокую эффективность в задачах классификации [4].

Схема нейронной сети выглядит следующим образом:

- входной слой (Input) размером (120,120,3);
- сверточный слой (Conv2D) с 16 фильтрами и ядром размером (3,3), функция активации - ReLU;
- слой максимального пулинга (MaxPool2D) с размером пула (2,2);
- сверточный слой (Conv2D) с 32 фильтрами и ядром размером (3,3), функция активации - ReLU;
- слой максимального пулинга (MaxPool2D) с размером пула (2,2);
- слой глобального усреднения пулинга (GlobalAveragePooling2D);

– выходной слой Dense с одним нейроном и сигмоидной функцией активации.

График функции потерь алгоритма на валидационных и тестовых данных представлен на рис. 1.

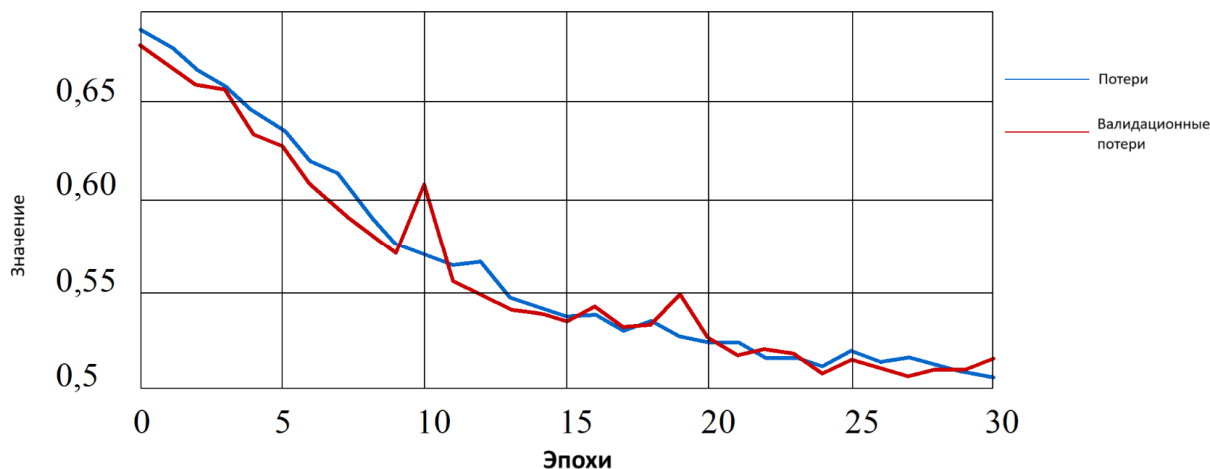


Рис. 1. График обучения алгоритма с CNN

По результатам тестирования модель показала тестовую точность (Test Accuracy) 76,30 %.

Отчет о классификации показал следующие результаты:

- для класса «NEGATIVE»: точность (precision) составила 0,83, полнота (recall) – 0,42, F1-мера (гармоническое среднее точности и полноты) – 0,56;
- для класса «POSITIVE»: точность (precision) составила 0,62, полнота (recall) – 0,92, F1-мера (гармоническое среднее точности и полноты) – 0,74. Что такое F1 –мера.

Результат указывает на то, что модель лучше справляется с определением класса «POSITIVE», но имеет проблемы с определением класса «NEGATIVE».

Результаты уверенности наличия трещины модели (в процентах) на тестовом наборе первых 10 изображений показаны на рис. 2.

```
final_results(old).csv > data
1 ,Filename,Result,Confidance of being Cracked
2 0,10clear.jpg,Negative (Not Crack),0.627637
3 1,1cracked.jpg,Positive(Crack) ,0.813721
4 2,2clear.jpg,Negative (Not Crack),0.543231
5 3,3cracked.jpg,Positive(Crack) ,0.966885
6 4,4cracked.jpg,Positive(Crack) ,0.999512
7 5,5clear.jpg,Negative (Not Crack),0.601615
8 6,6clear.jpg,Negative (Not Crack),0.586648
9 7,7clear.jpg,Positive(Crack) ,0.852246
10 8,8clear.jpg,Positive(Crack) ,0.963136
11 9,9cracked.jpg,Positive(Crack) ,0.964757
12
```

Рис. 2. Результаты первых 10 изображений алгоритма CNN

Можно заметить, что модель часто ошибается, но имеется потенциал для улучшения, особенно в части определения отсутствия трещины на изображении [5-7].

Алгоритм 2. Использование предварительно обученной модели VGG16 для классификации изображений трещин

В алгоритме представлен подход к классификации изображений с использованием предварительно обученной модели VGG16 [8]. Общее количество параметров модели составляет 14715201 (весов и смещений), из которых только 513 обучаемы, в последних слоях. Они могут быть изменены во время обучения. Остальные слои «заморожены» и их параметры не изменяются.

Схема нейронной сети в данном алгоритме выглядит следующим образом:

- входной слой (Input) размером (120,120,3);
- слои предварительно обученной VGG16 (с замороженными весами);
- слой GlobalAveragePooling2D;
- слой Dropout (с вероятностью 0,2);
- выходной слой Dense с одним нейроном и сигмоидной функцией.

График функции потерь при обучении и валидации показан на рис. 3.

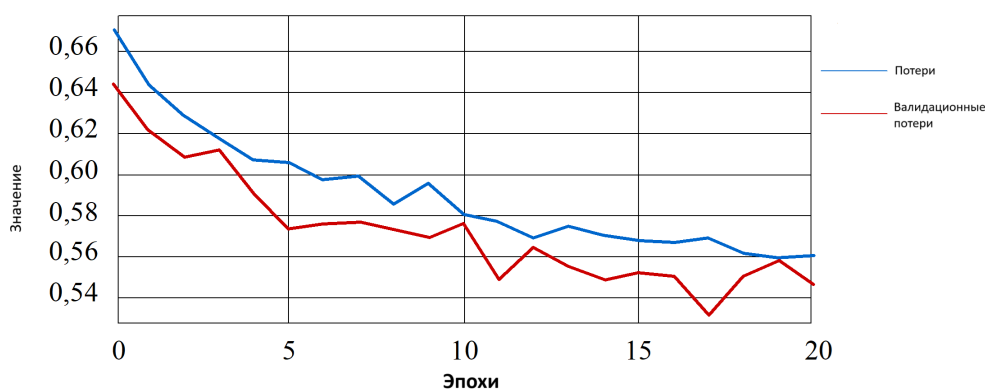


Рис. 3. Функция потерь при обучении и валидации для модели VGG16 с обучаемым слоем

Модель показала значительное улучшение по сравнению с предыдущими результатами, отчет о классификации показал следующие результаты:

- для класса «NEGATIVE»: точность (precision) составила 0,96, полнота (recall) – 0,86, F1-мера (гармоническое среднее точности и полноты) – 0,92;
- для класса «POSITIVE»: точность (precision) составила 0,75, полнота (recall) – 0,96, F1-мера (гармоническое среднее точности и полноты) – 0,86;
- общая точность модели (accuracy) увеличилась до 90 %.

Модель стала значительно лучше определять класс «NEGATIVE», что было проблемой в предыдущем эксперименте. Также стоит отметить улучшение общей точности модели. Результаты (в процентах), на тестовом наборе первых 10 изображений приведены на рис. 4.

```

final_results(VGG16).csv > data
1 ,Filename,Result,Confidance of being Cracked
2 0,10clear.jpg,Negative (Not Crack),0.367081
3 1,1craked.jpg,Positive(Crack) ,0.920216
4 2,2clear.jpg,Negative (Not Crack),0.365792
5 3,3cracked.jpg,Positive(Crack) ,0.872709
6 4,4cracked.jpg,Positive(Crack) ,0.766673
7 5,5clear.jpg,Negative (Not Crack),0.350197
8 6,6clear.jpg,Negative (Not Crack),0.297235
9 7,7clear.jpg,Negative (Not Crack),0.310231
10 8,8clear.jpg,Negative (Not Crack),0.352111
11 9,9cracked.jpg,Positive(Crack) ,0.938156
12

```

Рис. 4. Результаты первых 10 изображений алгоритма VGG16 с обучаемым слоем

Модель показывает высокую уверенность при классификации изображений с трещинами (значения близки к 1), и низкую уверенность при классификации изображений без трещин (значения близки к 0).

Алгоритм 3. Использование предварительно обученной модели VGG16 для классификации изображений трещин с использованием кросс-валидации

В данном алгоритме дополнительно используется 5-кратная кросс-валидация (KFold с n_folds=5). Для каждого из 5 разбиений выполняются все этапы обучения модели, включая оценку на валидационной выборке [9]. График функции потерь и аккуратности при обучении модели с кросс-валидацией для одного фолда показан на рис. 5.

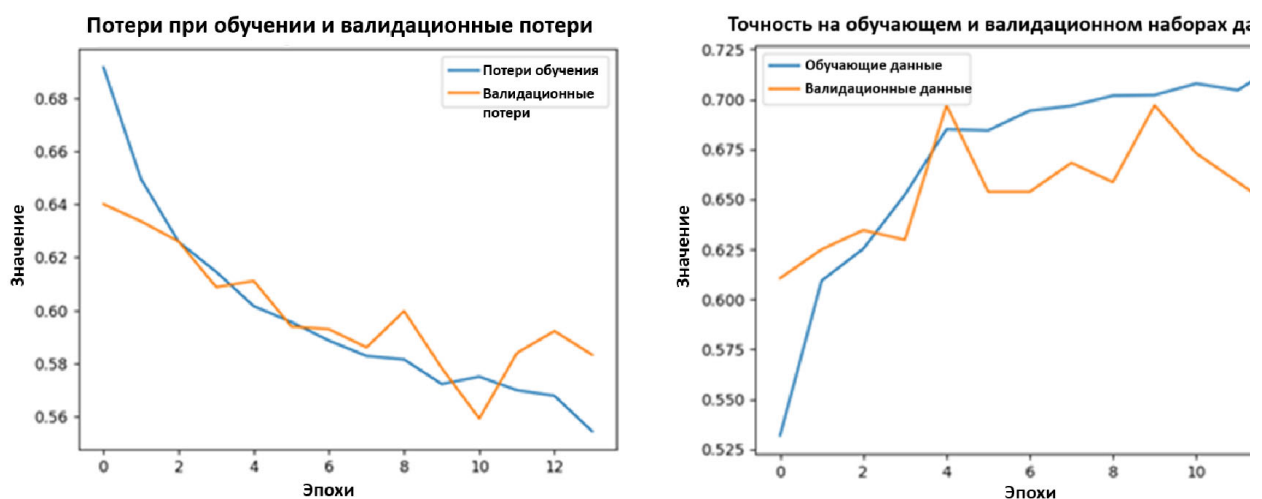


Рис. 5. Функция потерь и точности при обучении кросс-валидацией для модели VGG16 с обучаемым слоем для одного фолда

Результаты (в процентах), на тестовом наборе первых 10 изображений приведен на рис. 6.

```
final_resultsCrossValidation.csv > data
1 ,Filename,Result,Confidance of being Cracked
2 0,10clear.jpg,Negative (Not Crack),0.319881
3 1,1cracked.jpg,Positive(Crack) ,0.982643
4 2,2clear.jpg,Negative (Not Crack),0.378072
5 3,3cracked.jpg,Positive(Crack) ,0.97588
6 4,4cracked.jpg,Positive(Crack) ,0.834913
7 5,5clear.jpg,Negative (Not Crack),0.277851
8 6,6clear.jpg,Negative (Not Crack),0.251782
9 7,7clear.jpg,Negative (Not Crack),0.193464
10 8,8clear.jpg,Negative (Not Crack),0.274406
11 9,9cracked.jpg,Positive(Crack) ,0.99052
12
```

Рис. 6. Результаты первых 10 изображений алгоритма VGG16 с обучаемым слоем и кросс-валидацией

Сравнивая результаты обучения модели с использованием кросс-валидации и без нее, можно отметить следующие особенности:

- уверенность в наличии трещин: например, для изображения «1cracked.jpg» уверенность в наличии трещин увеличивается с 92 % (без кросс-валидации) до 98 % (с кросс-валидацией);

- уверенность в отсутствии трещин: например, в случае изображения «5clear.jpg» с 37 % о наличии трещины (без кросс-валидации) до 26% (с кросс-валидацией).

В отчете по классификации были получены следующие результаты:

- для класса «NEGATIVE» точность составила 0,97, полнота - 0,85, а F1-мера – 0,92. Это означает, что модель идеально предсказывает отсутствие трещин, но немного хуже определяет их наличие;

- для класса «POSITIVE» точность составила 0,76, полнота – 0,97, а F1-мера – 0,85. Это означает, что модель идеально определяет наличие трещин, но немного хуже предсказывает их отсутствие;

- общая точность модели на тестовом наборе (ассурагу) увеличилась до 91,33 %.

Результаты и обсуждение

Использование VGG16 с Transfer Learning и совместно с кросс-валидацией приводит к высокой точности классификации, достигающей 91,33 %. Оба подхода демонстрируют высокую точность как для «NEGATIVE» класса (до 0,97 и 0,86), так и для «POSITIVE» класса (до 0,76 и 0,97). Выбор между этими двумя подходами может варьироваться в зависимости от конкретных требований к задаче [10]. Стоит отметить, что при подходе с использованием кросс-валидацией модель может становиться более консервативна.

Результаты всех алгоритмов приведены в табл. 1.

Таблица 1

Сравнение результатов алгоритмов

Название способа	Test Accuracy	Precision «NEGATIVE»	Recall «NEGATIVE»	F1-score «NEGATIVE»	Precision «POSITIVE»	Recall «POSITIVE»	F1-score «POSITIVE»
Convolutional Neural Networks, CNN	76,30 %	0,83	0,42	0,56	0,62	0,92	0,74
VGG16 with Transfer Learning	90,00 %	0,96	0,86	0,92	0,75	0,96	0,86
VGG16 with Transfer Learning and Cross-Validation	91,33 %	0,97	0,85	0,91	0,76	0,97	0,85

Заключение

В данной работе был представлен подход к классификации изображений с использованием предварительно обученной модели VGG16. Полученные результаты показывают, что подход Transfer Learning может быть эффективным для задач классификации изображений. Кроме того, кросс-валидация улучшает способность модели правильно классифицировать изображения с трещинами, увеличивая уверенность в наличии трещин на соответствующих изображениях, но модель может становиться более консервативна. Так же в процессе работы удалось добиться до 91,33 % точности распознавания класса изображения

Эти результаты подтверждают эффективность примененных методов и показывают потенциал для дальнейшего улучшения модели.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Wei. W., Xu. X. Application of crack detection algorithm using convolutional neural network in concrete pavement construction // International Journal of Grid and Utility Computing, 2022. – №13. Т 2-3. – Р. 154–163.
2. Федотов И.С. Трансферное обучение в компьютерном зрении: применение и перспективы // Вестник науки, 2023. – №9. Т. 66. – С. 135–137.
3. Скороход М.В., Гордиенко Е.С., Семенкин Введение в глубинные нейронные сети // Актуальные проблемы авиации и космонавтики, 2020. – №3 – С. 83–85.
4. Маршалко Д.А., Кубанских О.В., Архитектура свёрточных нейронных сетей // Ученые записки Брянского государственного университета, 2019. – №2. Т. 30. – С. 10–13.
5. Бьенвеню М. М., Использование предобученной нейросети (VGG16) для решения задачи переноса стиля изображения // Известия тульского государственного университета. Технические науки 2021. – №18. Т. 2. – С. 243–248.

6. Чернышев Ю. О., Венцов Н. Н., Пшеничный И.С. Использование технологии Transfer Learning для интеллектуального формирования стартовых популяций при решении задач транспортного типа // Известия Южного федерального университета, 2018. – №3. Т. 233. – С. 90–98.
7. Xu Z., Qian S., Ran X., Zhou J. Application of Deep Convolution Neural Network in Crack Identification // Applied Artificial Intelligence 2022. – №36. Т. 1. – Р. 1–22.
8. Камалова Ю. Б., Андриянов Н. А. Распознавание микроскопических изображений пылевых зерен с помощью сверточной нейронной сети VGG-16 // Вестник Южно-Уральского государственного университета, 2022. – №22. Т. 3. – С. 39–46.
9. Ломовцева Н.А., Кувайскова Ю.Е. Использование модуля нейронных сетей в системе Statistica для решения задачи классификации // Известия Южного федерального университета. Вестник Ульяновского государственного технического университета, 2021. – №4(104) – С. 59–61.
10. Pak V., Ugli A. S. Application of transfer learning for medical image classification // Science, technology and education, 2020. – №4. Т. 92. – Р. 39–42.

© Р. М. Новиков, Д. Н. Титов, Е. В. Рыжкова 2024