

МОДЕЛЬ НЕЙРОННОЙ СЕТИ ДЛЯ ОПРЕДЕЛЕНИЯ ВИДА ФРАКЦИИ ЩЕБНЯ В КУЗОВЕ ГРУЗОВИКА

Тряпицын Д.Л., Решетников К.И.

ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина»,

Екатеринбург, Россия

try-denis@mail.ru

Аннотация. В строительной сфере применяют различные фракции щебня в качестве засыпных смесей. Учитывая, что стоимость щебня зависит от его вида, требуется автоматизированная система для проверки его типа и исключения человеческих ошибок. В данной работе предлагается метод классификации фракции щебня на изображении с помощью архитектуры EfficientNet-b1 с использованием пространственного внимания (Spatial Attention), совмещенный с функцией потерь LDAM. Для обучения и тестирования модели использовался набор из 635 изображений, разделенный на 7 фракций щебня. Полученная модель показала высокую точность, достигнув уровня 97 %.

Ключевые слова: компьютерное зрение, фракция щебня, EfficientNet, пространственное внимание, LDAM, классификация изображений.

NEURAL NETWORK MODEL FOR DETERMINING GRAVEL FRACTION TYPE IN TRUCK BODY

Tryapitsyn D.L., Reshetnikov K.I.

Ural Federal University, Ekaterinburg, Russia

Abstract. In the construction industry, various fractions of gravel are used as aggregate materials. Considering that the cost of gravel depends on its type, an automated system is required to verify its type and eliminate human errors. This work proposes a method for classifying gravel fractions in images using the EfficientNet-b1 architecture with the addition of Spatial Attention, combined with the

LDAM loss function. A dataset of 635 images, divided into 7 gravel fractions, was used for training and testing the model. The resulting model demonstrated high accuracy, reaching a level of 97 %.

Keywords: computer vision, crushed stone fraction, EfficientNet, spatial attention, LDAM, image classification.

1. Введение

В строительной и дорожной отраслях широко используется разнообразный щебень с различными размерными фракциями для изготовления строительных материалов, включая бетон и асфальт. Точное определение этих фракций щебня играет важную роль в производственном процессе, поскольку от этого зависит качество конечных строительных материалов. Кроме того, цена щебня также зависит от его фракционной характеристики. Тем не менее, существующие методы классификации имеют свои ограничения, включая необходимость высокой квалификации персонала или использование дорогостоящего оборудования. Для преодоления этих ограничений предлагается использовать методы компьютерного зрения и нейронных сетей.

Современные методы прогнозирования состава горных пород, руды или известняка обычно базируются на использовании архитектур сверточных нейронных сетей (CNN), таких как AlexNet, VGG16, ResNet50, InceptionV3 и MobileNet. Среди них одними из наиболее эффективных являются алгоритмы, основанные на переносе обучения (transfer learning), применение которых позволило достичь точности в 94% для MobileNet[1]. В другом исследовании [2] авторы решают схожую задачу: классификацию пяти различных классов изображений щебня в кузове грузовика. Они представляют модель под названием CSDeer, основанную на CNN и текстурных характеристиках, таких как вейвлет Габора, Харалик и Лоуз. CSDeer достигает точности тестирования 89,00%.

Основной проблемой при построении модели компьютерного зрения для классификации фракций щебня, является их взаимное пересечение. В этих классах содержатся фракции с одинаковыми размерами частиц щебня. Кроме того, широко распространена проблема недостаточного количества обучающих данных.

2. Основная часть

В качестве базовой архитектуры была выбрана предобученная CNN EfficientNet-b1. Она обеспечивает высокую производительность при относительно небольшом размере модели [3]. Это важно для уменьшения времени обучения и потребления ресурсов.

Для улучшения способности модели к выделению важных пространственных признаков был внедрен метод пространственного внимания (Spatial Attention) [4]. Этот метод позволяет модели сфокусироваться на наиболее значимых областях изображений. Используется один блок пространственного внимания. Входное изображение передается в блок пространственного внимания. Далее результат выхода передается в оригинальную архитектуру Efficientnet-b1. Поскольку предоставленные данные были несбалансированными по классам, был применен метод LDAMLoss [5]. Этот метод позволяет взвешивать функцию ошибки для разных классов во время обучения модели.

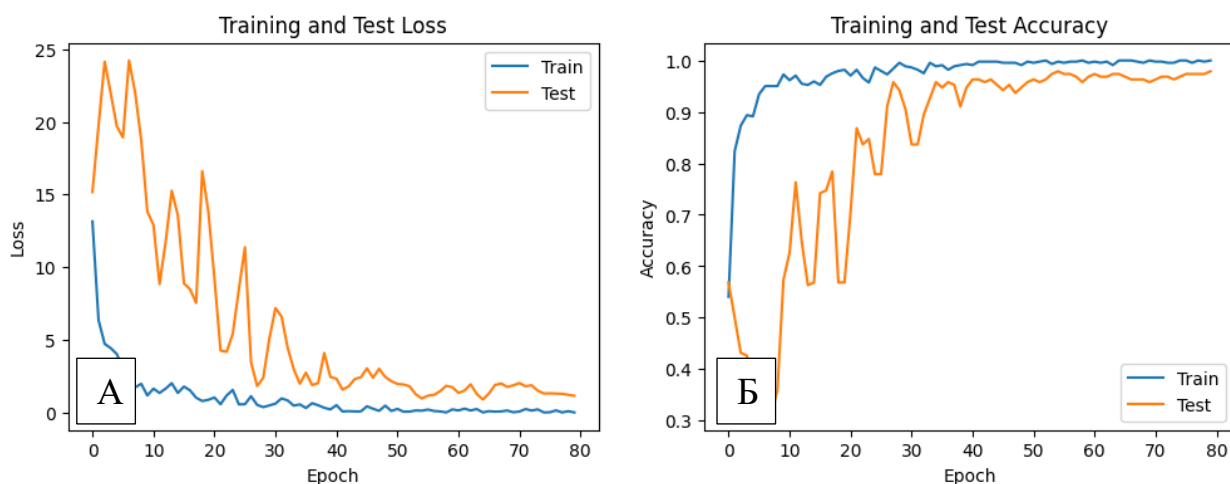


Рисунок 1 – График изменения значения функции потерь (А), график изменения точности классификации во время обучения (Б)

Для обучения и тестирования модели использовался случайно разделенный набор данных, включающий 444 изображений в обучающей выборке и 191 изображений в тестовой выборке, представляющий 7 фракций щебня. На рисунке 1 представлены графики изменения функции ошибки и точности классификации во время обучения для обучающей и тестовой выборок. Результаты классификации модели на тестовых данных достигли 97 %

Увеличение средних потерь объясняется использованием метода LDAM, который способствует более активному обучению модели на менее представленных классах. Повышенное значение функции потерь для редких классов, стимулирует модель уделять большее внимание этим классам. Это приводит к тому, что модель начинает лучше распознавать эти классы, увеличивая общую точность и устойчивость модели.

Для анализа, на какие области изображения обращает внимание модель при принятии решений, был применен метод визуализации Grad-Cam [6]. Интенсивность красного цвета на визуализации, указывает на уровень внимания модели к этой области. На рисунке 2 показана карта внимания моделей, с пространственным вниманием и без.

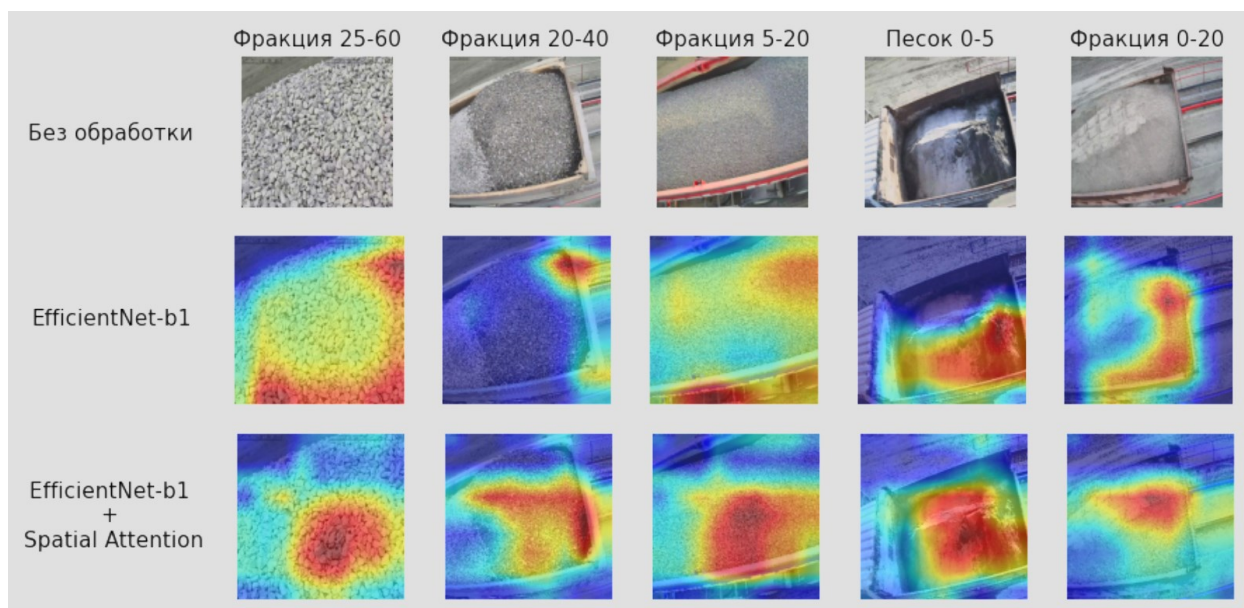


Рисунок 2 – Grad-Cam визуализация карты внимания модели

В случае модели EfficientNet-b1 с использованием Spatial Attention, тепловая карта внимания показывает более высокие значения в областях, соответствующих частицам щебня, по сравнению с моделью без Spatial Attention, когда внимание модели более рассеянное, охватывая как частицы щебня, так и борта кузова. Это указывает на то, что внимание модели в случае использования Spatial Attention более точно направлено на объекты интереса, что улучшает способность к классификации.

3. Заключение

В ходе исследования была разработана модель компьютерного зрения, основанная на архитектуре EfficientNet-b1 с использованием пространственного внимания (Spatial Attention) и метода LDAM. Модель способна классифицировать фракции щебня с высокой точностью, достигающей 97%.

Кроме того, разработанный подход может быть адаптирован и применен в других областях, где требуется точная классификация объектов на основе изображений, таких как анализ геологических образований или распознавание различных материалов.

Библиографический список

1. Ronkin, M. V. Review of deep learning approaches in solving rock fragmentation problems / M.V. Ronkin, E.N. Akimova, V.E. Misilov –// AIMS Mathematics. — 2023. — Т. 8. — № 10. — С. 23900-23940.
2. Charoenkwan, P. CSDeep: A Crushed Stone Image Predictor based on Deep Learning and Intelligently Selected Features / P. Charoenkwan, N. Homkong, 2017.
3. Tan M., Le Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // ArXiv. 2019. Vol. abs/1905.11946. URL: <https://api.semanticscholar.org/CorpusID:167217261>
4. Woo S., Park J., Lee J.-Y., Kweon I.-S. CBAM: Convolutional Block Attention Module // ArXiv. 2018. Vol. abs/1807.06521. URL: <https://api.semanticscholar.org/CorpusID:49867180>

5. Cao, K., Wei, C., Gaidon, A., Arechiga, N., Ma, T. Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss. — 2019. — [arXiv:1906.07413].

6. Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization. — [2019].