

АНАЛИЗ СРЕДСТВ РЕАЛИЗАЦИИ ДЛЯ ЗАДАЧИ ОБРАБОТКИ ИЗОБРАЖЕНИЯ

ANALYSIS IMPLEMENTATION FRAMEWORKS FOR THE TASK OF IMAGE PROCESSING

**T. Demenkova
E. Shpieva**

Summary. The paper presents the results of research to solve the problem of image processing based on deep learning methods. The analysis of existing frameworks is carried out and the criteria for the choice of solutions for the task are developed. It is shown that the chosen method using deep convolution neural networks can significantly improve the accuracy of image recognition and achieve results that cannot be obtained by classical methods.

Keywords: image processing, deep learning, neural network, computer vision.

Деменкова Татьяна Александровна

*К.т.н, доцент, МИРЭА — Российский
технологический университет (г. Москва)*

Шпиева Елизавета Викторовна

*Аспирант, МИРЭА — Российский технологический
университет (г. Москва)
elizashpieva@gmail.com*

Аннотация. В работе представлены результаты исследований для решения задачи обработки изображений на основе методов машинного обучения. Проведен анализ существующих фреймворков и разработаны критерии выбора решения для поставленной задачи. Показано, что выбранный метод с использованием глубоких сверточных нейронных сетей позволяет значительно улучшить точность распознавания изображения и достичь результатов, которые невозможно получить классическими методами.

Ключевые слова: обработка изображений, глубокое обучение, нейронная сеть, компьютерное зрение.

В настоящее время цифровые технологии все больше используются в промышленности. Они необходимы для оптимизации процессов и получения большей прибыли. Предметная область данного исследования имеет отношение к цифровизации карьеров по добыче полезных ископаемых. Происходит внедрение автономных карьерных самосвалов и остальной карьерной техники, такой как одноковшовые экскаваторы. Одноковшовые карьерные экскаваторы являются основной горной машиной при ведении погрузочных работ на открытых горных разработках. От их бесперебойной работы в большой степени зависит эффективность работы горного предприятия в целом. При эксплуатации карьерного экскаватора машинист сталкивается с рядом сложностей и проблем, связанных с размерами техники, ее конструктивными особенностями, погодными и климатическими условиями. Достаточно распространенной проблемой является проблема потери зубьев ковша экскаватора. Поскольку порода грубая и твердая, то зубья стесываются об нее и спустя некоторое время могут отломиться. В большинстве погодных условий сильный дождь, снег и яркое солнце мешают водителю экскаватора поймать момент потери зуба. Так как зуб изготавливается из закаленной стали, то его потеря может привести к серьезному повреждению инструментов, конвейерных лент, насосов, экранов, дробилки и другого оборудования при поступлении в систему обработки.

Существуют различные подходы для решения данной проблемы. Хорошие результаты показывают алго-

ритмы, использующие нейронные сети, например, для компьютерного зрения, распознавания речи, обработка естественного языка. К какой бы сфере ни относилась задача при распознавании образов перед построением классификатора зачастую необходимо строить существенные (инвариантные) признаки на основе некоторых исходных признаков, что нередко делается вручную. Задача автоматического выделения признаков является особенно актуальной для систем компьютерного зрения. Последнее связано с тем, что классы объектов, распознаваемых по их изображениям, крайне редко бывают линейно разделимы в пространстве первичных признаков — яркостей пикселей, допустим, при различных погодных условиях на карьере. Построение нелинейных разделяющих поверхностей возможно с помощью таких традиционных методов распознавания образов, как например, метод обобщенных решающих функций или машин опорных векторов. Одним из современных подходов к проблеме автоматического выбора нелинейных признаков является подход на основе сетей глубокого обучения, ставших весьма популярными благодаря способности решать задачи распознавания без использования вручную построенных признаков [1,2]. Глубокое обучение исключает необходимость распознавание характерных признаков в «сырых» входных данных, выполняемых вручную, заменяя их процессом обучения, в результате которого выявляются наиболее полезные структуры. К примеру, сети глубокого обучения были использованы для создания системы распознавания дорожных знаков, впервые в истории искус-

Таблица 1. Анализ инструментов.

Названия ПО	Tensorflow	Keras	Caffe	PyTorch	MXNet	CNTK	Deeplearning4j
Документация	+	+	-	-	-	+	+
Мониторинг	+	+	-	-	-	-	-
Большая среда использования	+	+	-	-	-	-	-
Поддержка распределенного обучения	+	+	-	+	-	-	+
Сложность освоения	+	-	+	+	+	+	-
Низкоуровневый	+	-	+	+	-	+	-
Простое прототипирование	-	+	+	+	+	-	-
Поддержка GPU	+	+	+		+	+	+
Ограниченный функционал	-	+	+	-	-	-	-
Масштабируемость	+	-	-	-	+	+	+
Язык	C++	Python	C++	Python	Python, Java, C++, R	Python, C++, C#, Java	Java, Scala

ственного интеллекта продемонстрировавшей качество решения задачи распознавания выше, чем у человека [3]. Таким образом, методы на основе сетей глубокого обучения являются наилучшими известными методами распознавания образов, по крайней мере, в ряде предметных областей.

В наших исследованиях одной из основных проблем является ограниченная видимость в зоне непосредственной близости экскаватора. Большое количество мертвых зон приводит к тому, что оператор не имеет достаточного обзора для безопасной и безаварийной работы. Также наличие мертвых зон приводит к снижению эффективности работы экскаваторов.

В случае потери зуба приходится останавливать погрузочные работы и проводить его поиски, в противном случае при несвоевременном реагировании потерянный зуб может вывести из строя дорогостоящее оборудование, что в свою очередь приведёт к простоя и незапланированным затратам. Удаление зубьев, попавших в конвейерные ленты или производственные дробилки, тоже весьма опасная и дорогостоящая процедура, для которой требуется привлечение высококвалифицированного (и высокооплачиваемого) персонала. Также стоимость простоя производства очень велика. Для решения необходимо внедрять технологии распознавания образов и машинного обучения, что позволяет отслеживать состояние ковша в автоматическом режиме, предотвращать попадание зубьев в дробилки и конвейеры.

В настоящее время существует множество различных программных средств для обработки и анализа данных. Было проведено сравнение нескольких видов про-

граммного обеспечения, которые можно использовать для данной задачи (таблица 1).

Рассматривая несколько видов фреймворков [4], при выборе одного для поставленной задачи должны учитываться несколько факторов:

- ◆ тип нейронных сетей;
- ◆ язык программирования;
- ◆ количество инструментов и дополнительные параметры;
- ◆ бюджет;
- ◆ характер и общие цели проекта.

Каждая система глубокого обучения или библиотека имеют разные характеристики, предлагая уникальные функции, которые представляют разные решения. При исследовании фреймворков использовались следующие параметры тестирования [5]:

- ◆ продолжительность;
- ◆ потребление памяти;
- ◆ загрузка обрабатывающего и графического процессоров;
- ◆ количество эпох.

В качестве метрик для теста производительности использовались:

- ◆ время обучения нейронной сети;
- ◆ время тестирования нейронной сети;
- ◆ точность обучения / прогнозирования;
- ◆ состоятельность надёжность;

В качестве наборов данных использовались CIFAR-10 [6] и MNIST [7]. По результатам тестирования TensorFlow и Caffe предлагают более высокую точность, исполь-

зую свои собственные настройки CIFAR-10 по умолчанию. В случае Torch была получена более высокая точность, используя настройку по умолчанию CIFAR-10 в TensorFlow, что потребовало огромных затрат времени на обучение. Наконец, когда Caffe использует собственную настройку по умолчанию для CIFAR-10 Обученная модель TensorFlow демонстрирует более высокую надёжность, чем у Caffe. Одной из возможных причин является то, что выпадение в TensorFlow несколько слабее регуляризации, чем снижение веса в Caffe. Для набора данных MNIST модель DNN, обученная TensorFlow, в некоторой степени более устойчива к обоим типам атак, чем модель DNN, обученная Caffe. В качестве метрик тестирования выбрали время обработки и скорость сходимости.

Проведя аналогичные исследования для других вариантов, был выбран фреймворк TensorFlow в связке с Keras. Для него наработана существенная база знаний,

по мнению экспертов он наиболее приспособлен для исследований и обладает большим набором функций для создания своей математической модели. Кроме того, судя по тестам на производительность, скорость его работы при использовании нескольких потоков является наилучшей.

Были проанализированы наиболее распространённые существующие глубокие сверточные нейронные сети: AlexNet; ZF Net; VGG Net; Inception; Inception-v2 и Inception-v3; ResNet; Inception-v4 и Inception-ResNet; G. Densenet and DPN [8–11]. Они позволили значительно улучшить точность распознавания изображений и достичь результатов, которые невозможно получить классическими методами компьютерного зрения. Архитектура сверточной нейронной сети интегрирует выделение элементарных признаков изображения, формирование более сложных признаков на высших уровнях обработки и распознавание поставленных в задаче изображений.

ЛИТЕРАТУРА

1. Schmidhuber J. Deep Learning in neural networks: An overview // *Neural Networks*. 2015. (61). pp. 85–117.
2. Goodfellow I., Bengio Y., Courville A. *Deep Learning* Massachusetts: MIT Press, 2016. pp. 271–290.
3. Ciresan D., Meier U., Masci J., Schmidhuber J. Multi-column deep neural network for traffic sign classification // *Neural Networks*. 2012. (32). pp. 333–338.
4. Bahrampour S., Ramakrishnan N., Schott L., Shah M. Comparative Study of Deep Learning Software Frameworks. 2015. URL: file:///C:/Users/79162/Downloads/Comparative_Study_of_Caffe_Neon_Theano_and_Torch_f.pdf (дата обращения 15.03.2019)
5. Shi S., Wang Q., Xu P., Chu X. Benchmarking state-of-the-art deep learning software tools. 2017. pp. 99–104.
6. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition in *Proceedings of the IEEE*, vol. 86, no. 11, pp.2278–2324, Nov.1998. doi: 10.1109/5.726791
7. Krizhevsky A. Learning multiple layers of features from tiny images. Technical Report TR-2009, Department of Computer Science, University of Toronto, Toronto.
8. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V. Rabinovich A. Going deeper with convolutions. *The IEEE Conference on Computer Vision and Pattern Recognition*. 2015. pp. 1–9.
9. Krizhevsky A., Sutskever I., Hinton G. Imagenet classification with deep convolutional neural networks. In *NIPS*, 2012. pp. 1097–1105.
10. Vanhoucke V., Ioffe S., Shlens J., Zbigniew Wojna Z., Szegedy C. Rethinking the Inception Architecture for Computer Vision. *The IEEE Conference on Computer Vision and Pattern Recognition*. 2016. pp. 2818–2826.
11. Gowda S. N., Yuan C. ColorNet: Investigating the importance of color spaces for image classification. In *CoRR*, 2019. Vol.1. abs/1902.00267.

© Деменкова Татьяна Александровна, Шпиева Елизавета Викторовна (elizashpieva@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»