

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.891.2

Ковалев
Сергей Леонидович

Система распределенного интеллектуального анализа изображений

Автореферат
на соискание степени магистра технических наук
по специальности 1-40 80 03 Вычислительные машины и системы

Научный руководитель
Татур Михаил Михайлович
доктор технических наук,
профессор кафедры ЭВМ

Минск 2016

КРАТКОЕ ВВЕДЕНИЕ

На протяжении многих лет перед исследователями стоит задача разработки систем интеллектуального анализа изображений, которые смогли бы покрыть весь спектр проблем существующий в данной области. Но стоит отметить, что на сегодняшний день не существует методов, позволяющих решать задачу семантического анализа изображений в общем виде, а теория в данной области представляет собой сочетание ряда нерешенных к настоящему моменту научных проблем и большого числа подходов не являющихся хорошо проработанной и обобщенной теорией.

На практике, во многих случаях, для решения многих задач распознавания изображений общая проблема семантического анализа может быть приведена к гораздо более простой проблеме обнаружения или распознавания по некоторому заранее известному модельному описанию. Существует целый ряд различных модельных описаний наблюдаемых объектов, которые могут быть использованы для их обнаружения и измерения, начиная с простейших признаковых описаний до высокоспециализированных структурных моделей.

Как было сказано выше, общего алгоритма построения моделей, способных работать с заданной точностью по вышеописанным причинам, на данный момент не существует. Отсюда следует, что разработка моделей, пригодных для эффективного решения задач семантического анализа изображений, в значительной степени требует знания предметной области, отражающего многолетний опыт исследований по решению частных задач определенного типа.

Однако, в последние несколько лет, в том числе и в задачах интеллектуального анализа изображений, получили новый импульс развития алгоритмы глубинного обучения. В частности речь идет о сверточных нейронных сетях, автоэнкодерах и других алгоритмах и их комбинациях. Представляется, что данный набор алгоритмов сможет предоставить обобщенный подход к решению задач семантического анализа изображений в частности и компьютерного зрения в целом.

В связи с вышесказанным, для построения прототипа распределенной системы интеллектуального анализа изображений были выбраны методы глубинного обучения на основе сверточных нейронных сетей. В качестве прикладной задачи взята проблема распознавания злокачественных образований на гистологических снимках лимфы.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель работы - разработка базовых элементов прототипа распределенной системы интеллектуального анализа изображений на основе алгоритмов глубинного обучения. В данном случае под интеллектуальным анализом понимается извлечение значимой информации из изображения. Для каждой конкретной ситуации эта значимая информация может быть разной, будь то распознавание лиц, других объектов на изображении, понимание сцен, сегментация и так далее.

Для достижения поставленной цели исследования необходимо решить следующие задачи:

- определить доступный на сегодня набор фреймворков глубинного обучения;
- провести сравнительный анализ фреймворков глубинного обучения по основным значимым метрикам;
- на основе результатов сравнения определить наиболее подходящий фреймворк для решения поставленной задачи;
- т.к. цель построить распределенную систему анализа изображений, то следовательно необходимо проверить способность фреймворка масштабировать процесс обучения модели на кластере из нескольких машин;
- решить прикладную задачу при помощи выбранного фреймворка.

Объектом исследования и разработки является прототип системы интеллектуального анализа изображений на основе алгоритмов глубинного обучения.

Предметом исследования являются составные части, необходимые для построения системы интеллектуального анализа изображений и их характеристики.

Научными результатами данного исследования являются:

- результаты сравнительного анализа фреймворков глубинного обучения по значимым метрикам;
- проверка характеристик выбранного фреймворка посредством решения прикладной задачи.

Результаты магистерской диссертации представлены в виде публикации и доклада на научной конференции: Вторая Международная Научно-Практическая Конференция BIG DATA и анализ высокого уровня.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулированы цели и задачи исследования, приведено краткое содержание работы.

В первой главе рассмотрены основные принципы функционирования сверточной нейронной сети, ее архитектура и составные элементы. Кроме того, рассмотрены преимущества и недостатки, способы обучения сети и практические области применения.

В второй главе проведена оценка основных фреймворков глубинного обучения, включая Theano, Torch, Caffe, TensorFlow и Deeplearning4J, по таким характеристикам как: точность классификации при заданном числе эпох обучения, время обучения, время предсказания, а также другие важные параметры. Данная работа ограничена только архитектурой Полносвязной Нейронной Сети, т.к. является первой в серии публикаций по сравнительному анализу фреймворков глубинного обучения.

Результатами второй главы являются выводы приведенные в оценочном стиле. Ниже приведены основные из них:

1. Использование ReLU в качестве нелинейной функции активации вместо гиперболического тангенса ведет к увеличению времени обучения и точности классификации.
2. При тестировании зависимости точности классификации от глубины нейронной сети (количества слоев) Torch и TensorFlow показали неожиданный результат: при увеличении глубины их точность классификации падает. Такой результат может быть объяснен либо некорректным вычислением градиента в процессе обучения, что маловероятно, либо необходимостью более тонкой настройки весовых коэффициентов, что требует более глубоких навыков работы с Torch.
3. В соответствии с результатами сравнения зависимости точности классификации от количества нейронов в скрытом слое, пять фреймворков разделились на 3 группы:
 - i. для фреймворков Caffe и Deeplearning4J точность падает с увеличением количества нейронов;
 - ii. в отличие от предыдущих фреймворков TensorFlow и Torch, точность классификации увеличивается с увеличением числа нейронов в скрытом слое;
 - iii. Theano - Keras показывает стабильную точность классификации даже с относительно малым количеством нейронов в скрытом слое.

4. DeepLearning4J является самым медленным с точки зрения времени тренировки и классификации.
5. В итоге, принимая во внимание результаты тестирования скорости обучения, точности классификации, сложности построения алгоритма (количество строк кода), список фреймворков выглядит следующим образом: Theano - Keras, TensorFlow, Caffe, Torch, DeepLearning4J.

Учитывая, что целью данной работы является разработка базовых элементов прототипа распределенной системы интеллектуального анализа изображений, а на данный момент из лидеров рейтинга только TensorFlow поддерживает распределенное обучение модели, было принято решение взять за основу именно этот фреймворк.

В третьей главе была проведена проверка возможности масштабировать процесс обучения на кластере из нескольких узлов. Для этого решено провести тесты на кластере из 5 машин, каждая из которых содержит несколько GPU устройств.

Как было сказано выше, в качестве прикладной задачи, на которой проверена возможность масштабирования, была выбрана проблема автоматического распознавания злокачественных образований на гистологических снимках лимфы.

Для оценки производительности распределенного обучения модели сверточной нейронной сети были использованы следующие метрики: среднее время обработки одного батча на узле и количество изображений обработанных в секунду всего в системе, т.е. на всем кластере.

Результаты исследования показали способность TensorFlow к масштабированию процесса обучения, однако и выявили ряд недостатков, а именно: производительность распределенного обучения сети очень быстро деградирует с ростом числа узлов. Возможно это связано с использованием в качестве сетевого транспорта технологии Ethernet, который в отличие от технологии Infiniband обладает очень высоким временем отклика сигнала.

В целом же, можно признать опыт использования TensorFlow успешным, а сам фреймворк пригодным для построения распределенной системы анализа изображений.

ЗАКЛЮЧЕНИЕ

В процессе выполнения диссертации получены следующие основные научные и практические результаты:

1. Проведен сравнительный анализ фреймворков глубинного обучения по основным значимым метрикам.
2. На основе результатов сравнения определений наиболее подходящий фреймворк для решения поставленной задачи.
3. Проверена способность фреймворка масштабировать процесс обучения модели на кластере из нескольких машин.
4. Решена прикладная задача.

Библиотека БГУМР

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

Ковалев В. А., Калиновский А. А., Ковалев С. Л. Сравнительный анализ фреймворков глубинного обучения. Вторая Международная Научно-Практическая Конференция BIG DATA и анализ высокого уровня. Минск 2016.

Библиотека БГУИР