

ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ

АДАПТАЦИЯ И ПОРТИРОВАНИЕ НА БОРТОВОЙ МИКРОКОМПЬЮТЕР РОБОТА НЕЙРОСЕТЕВЫХ СРЕДСТВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ ОКРУЖАЮЩЕЙ СРЕДЫ

П. А. Акоюн

Белорусский государственный университет, г. Минск;

pavel.akarjan@gmail.com;

науч. рук. – А. В. Тузиков, д-р физ- мат. наук., проф.;

А. А. Калиновский, науч. сотр. ГНУ «ОИПИ НАН Беларуси»

Была разработана архитектура сверточной нейронной сети (СНС) адаптированная для работы на бортовых микрокомпьютерах (БМК) и метод обучения, позволяющий значительно улучшить качество распознавания. Также в работе предложен метод частичного MC Dropout для повышения качества на этапе прогнозирования. Разработанные СНС были портированы и протестированы на БМК Raspberry Pi 3 в рамках задачи классификации объектов окружающей среды.

Ключевые слова: сверточные нейронные сети; бортовые микрокомпьютеры; распознавание объектов; оптимизация; компьютерное зрение.

ВВЕДЕНИЕ

В настоящее время актуальность роботизированных систем возрастает в связи с развитием потребностей экономики и военной промышленности. Ключевым управляющим компонентом автономных роботов является БМК с установленной на нем системой компьютерного зрения. СНС стали основой передовых систем компьютерного зрения за счет появления возможностей использовать высокопроизводительные сервера с графическими ускорителями (GPU), а также сбора больших объемов обучающей выборки изображений. Проблема в адаптации нейросетевых средств распознавания изображений на БМК заключается в том, что вычислительная мощность процессорных устройств БМК может в десятки, а то и сотни раз быть меньше, чем на серверах с GPU. Это приводит к тому, что стандартные нейросетевые средства не могут работать на роботизированных устройствах в режиме приближенному к реальному времени. Дополнительно при решении прикладных задач системы компьютерного зрения возникают проблемы нехватки размеченной обучающей выборки, которые приводят к сильному ухудшению качества распознавания. В связи с этим возникает вопрос

переиспользованая существующих открытых обучающих выборок изображений, которые создавались для решения других задач.

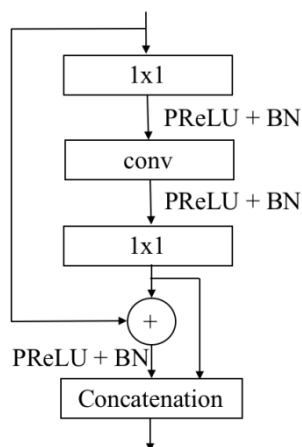
В рамках работы были проведены исследования сверточных нейросетевых архитектур с целью повышения эффективности (соотношение качества распознавания к затраченному времени) алгоритмов, по результатам которых была предложена архитектура, адаптированная для работы на БМК. Также были проведены исследования методов обучения, по итогу которых был предложен комплексный метод обучения, позволяющий значительно улучшить качество распознавания в условиях малых выборок и малых вычислительных ресурсов. В конце работы было произведено портирование на БМК Raspberry Pi 3 разработанной СНС, которая решает задачу классификации объектов окружающей среды.

АДАПТИРОВАННАЯ АРХИТЕКТУРА СНС

В результате проведенных экспериментальных сравнений по эффективности различных современных СНС была разработана следующая архитектура, состоящая из трех последовательных групп.

Группа 1. Сверточное уменьшение изображения

В работе использовалась 8 МП камера с разрешением 1920x1080. Увеличение разрешения входного изображения для СНС существенно повышает время прогнозирования, при этом для многих задач распознавания объектов окружающей среды увеличение может дать лишь небольшое улучшение качества. Сверточное уменьшение размера можно произвести со скоростью сопоставимой с традиционными методами интерполяции, при этом СНС может обучиться делать дополнительные подготовительные преобразования изображения.



Базовый блок с шагом 1

Рис. 1. Базовый блок с шагом 1

Для реализации уменьшения размера предлагается использовать два сверточных блока, которые уменьшают разрешение входного изображения в 9 раз. Эти блоки состоят из последовательности поканального сепарабельного сверточного слоя (DWS) [1], PReLU функции активации и слоя батч-нормализации. Для того, чтобы DWS выполнял уменьшение разрешения в 3 раза используется шаг свертки 3 с ядром 6x6. Выбор такой конфигурации мотивирован тем, что ядро 6x6 при шаге 3 не создает закономерных артефактов на выходных изображениях, при этом использование DWS позволяет выполнить свертки с ядром 6x6 без существенных потерь времени. Также PReLU функция активации по сравнению с ReLU функцией активации позволяет повысить эффективность СНС.

Группа 2. Базовые сверточные блоки

В базовых блоках осуществляется основная часть обработки изображения. Общая структура разработанных блоков, представленная на рисунке, существенно отличается существующих.

Во-первых, вместо классических сверточных слоев используются (чередующиеся по блокам) дилационные [2], ассиметричные [3] и пространственные [1] слои. Дилационные и ассиметричные слои позволяют за небольшое количество слоев быстро увеличить рецептивное поле СНС, а пространственный слой выполняется значительно быстрее стандартного. Во-вторых, перед сверточным слоем происходит 8-кратное канальное сжатие свертками 1x1, а после сверточной операции – 8-кратное канальное разжатие. В-третьих, особенностью разработанной архитектуры является использование и суммирующих, и конкатенирующих пробрасывающих связей (skip-connections).

Группа 3. Финальные блоки

Для адаптирования СНС под задачу классификации используются DWS с шагом 2 (взамен использования слоев субдискретизации) и небольшой полносвязный слой с softmax активацией.

МЕТОД ОБУЧЕНИЯ АДАПТИРОВАННОЙ СНС

Для общего повышения качества распознавания адаптированной сети предлагается использовать следующий метод.

Шаг 1. Последовательно переобучаются под целевую задачу слои с конца большой СНС предобученной на данных ImageNet.

Шаг 2. Создается сеть-ансамбль за счет применения метода Монте-Карло исключений [4] (MC Dropout) и аугментаций на тесте.

Шаг 3. Малая СНС с адаптированной архитектурой обучается на размеченных данных из целевой задачи и неразмеченных человеком данных целевой задачи, разметку которых выполнили сетью-ансамблем.

Шаг 4. Дообучение сети с отключением аугментаций и сменой оптимизатора с Adam на SGD с низким шагом обучения.

Шаг 5. Усреднение весов на последних эпохах.

ПОРТИРОВАНИЕ СНС

При портировании на устройство существенно ускоряет и уменьшает размер сети 8-битная квантизация. Для повышения качества на тесте используется частичный MC Dropout ранее не описанный в литературе. Он заключается в применении MC Dropout лишь на последних слоях, чтобы не замедлять значительно выполнение прогнозирования.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СНС

В таблице показано сравнение по точности и времени прогнозирования стандартной VGG архитектурой, мобильной архитектурой MobileNet [1] и предложенной архитектуры до (PiNet) и после (PiNet+) применения комплексного метода обучения, а также с использованием частичного MC Dropout (PiNet++).

Метрика	VGG	MobileNet	PiNet	PiNet+	PiNet++
Точность	0,88	0,84	0,82	0,85	0,86
Время, с	0,19	0,07	0,04	0,04	0,06

ЗАКЛЮЧЕНИЕ

В результате исследований была предложена архитектура СНС, которая превосходит по эффективности существующие адаптированные архитектуры для мобильных устройств. Представленный метод обучения существенно улучшил качество распознавания. Предложенный в работе метод частичного MC Dropout может использоваться для дополнительного улучшения качества распознавания, так как не требует дополнительного обучения и значительного увеличения вычислительных затрат.

Библиографические ссылки

1. Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. Efficient convolutional neural networks for mobile vision applications. / Howard A.G. – arXiv preprint arXiv:1704.04861, 2017.
2. Yu F., Koltun V. Multi-scale context aggregation by dilated convolutions. / Yu F. – arXiv preprint arXiv:1704.04861, 2017.

3. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Wojna. Rethinking the inception architecture for computer vision. / Szegedy C. arXiv preprint arXiv:1512.00567, 2015.
4. Gal Y., Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. / Gal Y. ICML 2016.