

УДК 004.4:004.032.26

EDN [QCLRZE](#)



Тестирование и анализ нейросетевых моделей для распознавания объектов на изображении в режиме реального времени

М.И. Деев, О.А. Ковалева, С.В. Ковалев

ФГБОУ ВО Тамбовский государственный университет имени Г.Р. Державина,
Тамбов, 392000, Россия

E-mail: deevmih3@yandex.ru

Аннотация. В данной статье проведено сравнение двух нейросетевых моделей глубокого обучения на предмет быстродействия и точности выходных данных при обнаружении объектов на изображении в режиме реального времени. Для проведения тестирования были взяты две нейросетевые модели CenterNet и ThunderNet. Обе модели были разработаны под конкретные задачи по определению объектов на статичном изображении либо при помощи захвата камеры в реальном времени.

Ключевые слова: нейросетевые модели; глубокое обучение; искусственный интеллект; распознавание объектов; компьютерное зрение.

Testing and analysis of neural network models for real-time object recognition in an image

M.I. Deev, O.A. Kovaleva, S.V. Kovalev

Derzhavin Tambov State University, Tambov, 392000, Russia

E-mail: deevmih3@yandex.ru

Abstract. This article compares two deep learning neural network models in terms of speed and accuracy of output data when objects are detected in an image in real time. Two neural network models CenterNet and ThunderNet were taken for testing. Both models were developed for specific tasks of identifying objects in a static image or using real-time camera capture.

Keywords: neural network models; deep learning; artificial intelligence; object recognition; computer vision.

1. Введение

В настоящий момент существует множество нейросетевых моделей под самые различные задачи. Основным направлением развития нейросетей является компьютерное зрение. Данное направление крайне универсально. Оно может применяться и для базовых задач, как например камера смартфона, или же для сложных, таких как определение дорожных знаков беспилотными автомобилями. Чтобы получить такие высокоскоростные и высокоточные данные необходимо пользоваться механизмом параллельных вычислений. Сейчас можно найти множество нейросетевых моделей для проведения глубокого обучения системы, основанных на параллелизме. Для выбора подходящей в конкретном случае модели нужно провести сравнительный анализ их работы в разных условиях.

2. Постановка задачи (Цель исследования)

Цель работы заключается в том, чтобы выявить модель, которая дает преимущественно точные результаты после обучения, располагает высокой производительностью и позволяет выдавать наиболее обширные выходные данные при минимальных входных.

Новизна исследования заключается в том, что оно было проведено с применением сравнительного анализа моделей нейросетей со входными данными, размер которых 32×32 .

3. Методы и материалы исследования

Объектами анализа были выбраны две наиболее распространенные нейросетевые модели: CenterNet и ThunderNet. В ходе изучения используется доступный набор данных CIFAR-10[5], включающий в себя 60000 изображений размером 32×32 разбитых на несколько категорий: животные, автомобили, самолеты. Изображения объединили в две подбазы данных, одна из которых обучающая, а вторая – тестовая. В базе данных для обучения находится 50000 файлов-изображений, в базе для тестирования – 10000. Каждое изображение содержит аннотацию с информацией о четырех координатах знака и его категории.

Первая анализируемая модель CenterNet представляет из себя сочетание уникальной архитектуры и быстродействия. Данная нейросеть определяет объект, как одну точку, которая находится в центре ограничительной рамки. Такие параметры как: размер, ориентация, 3D – форма, направление модель получает в последствии из характеристик изображения около полученной точки. Модель генерирует тепловую карту, пики температуры на ней соответствуют центрам объектов. Характеристики изображения в каждом пике тепловой карты определяют размеры ограничительной рамки вокруг объекта.[1] При помощи CenterNet можно определить как размер 3D – объекта, так и оценить позу человека по двумерному изображению. Архитектура нейросети CenterNet представлена на рисунке 1.

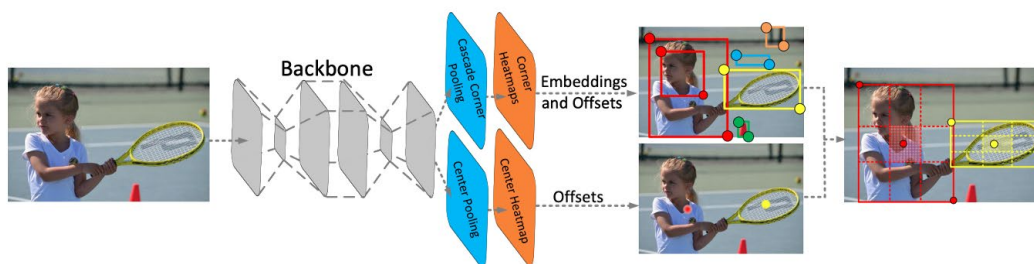


Рисунок 1. Архитектура нейросетевой модели CenterNet.

На ней изображена работа по определению точки в центре объекта и ограничительной рамки по его периметру [2]. Второй анализируемой моделью была выбрана ThunderNet. Нейросеть имеет архитектуру легковесного двухступенчатого детектора. По заявлениям авторов данная модель является первым детектором объектов в реальном времени, который был запущен на платформе ARM. Так как архитектура ARM по последним тенденциям становится повсеместной нейросеть переходит из мобильного класса в кроссплатформенный. При этом нейросеть может обрабатывать изображения со скоростью 24.1 FPS (frames per second, кадров в секунду) что является одним из самых быстрых показателей среди мобильных нейросетей. В то же время по точности выходных данных данная нейросеть сопоставима с десктопными версиями, что является весомым преимуществом. При работе с изображениями нейросеть разбивает их на блоки размером 5x5, а затем анализируя каждый из них в отдельности составляет карту особенностей объекта. [3] В последствии нейросеть сравнивает общий вид объекта и особенности выявленные из процесса классификации по блокам. Конечным результатом работы является как общее описание объекта, так и уникальные черты формы, цвета, контрастности. Высокая детализация при работе нейросети в реальном времени позволяет ее применять даже в таких ресурсно-затратных задачах как распознавание лица, либо построение 3D – модели

интерьеров, по аналогии с технологией Lidar. Архитектура нейросети ThunderNet представлена на рисунке 2.

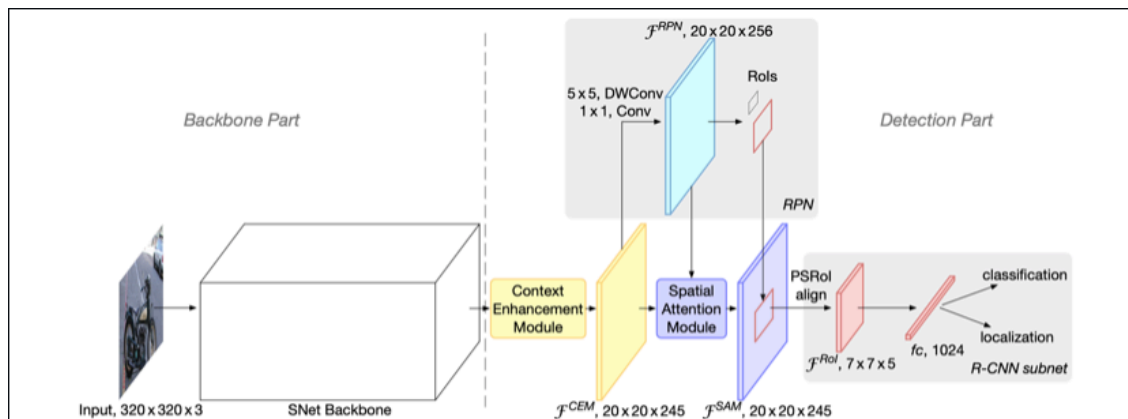


Рисунок 2. Архитектура нейросетевой модели ThunderNet.

На изображении представлен процесс по классификации обнаруженного объекта на изображении по блокам [4]. Исследования проводились на компьютере со следующими техническими характеристиками: графический ускоритель GTX 680 (оперативная память 2 Гб), процессор Intel Core I3 4130 на 3.2 Гц и 16 Гб оперативной памяти. Код написан на языке Python 3.6, применялись стандартные и встраиваемые библиотеки.

Для проведения более ускоренного исследования платформу проверяют в стандартных условиях с одной эпохой обучения и использованием моделей CenterNet и ThunderNet, вместе с набором данных CIFAR-10, включающим 60000 изображения. Входные данные в размере 32×32 , пакет данных равен 58. Производительность платформы, использующей как процессор, так и графический ускоритель, увеличивается с помощью оптимизации системы и кода. Часто изображения, представленные в наборе для обучения, имеют отличия в размере или погрешности в освещении. Поэтому, чтобы повысить точность и качество обучения следует провести обработку изображений. Для облегчения этой задачи подходит XnView, программа для обработки фотографий. Для загрузки изображения обрезаются и сводятся к одному размеру в 32×32 .

4. Полученные результаты

Рисунок 3 отображает время обучения двух моделей нейросетей при одинаковом размере входных данных (32×32). График был составлен в сравнении GPU и CPU. На графике можно наблюдать, что производительность намного выше с применением GPU,

нежели чем с CPU. В сравнении двух нейросетевых моделей наиболее производительной себя показала модель ThunderNet. Модель CenterNet имеет примерно одинаковые показатели по времени обучения. Нейросетевая модель ThunderNet изначально проектировалась под ARM процессоры, которые на момент разработки нейросети были малопродуктивны. В дальнейшем процессоры с ARM архитектурой совершенствовали свой тех процесс что позволило ускорить быстродействие приложений, запущенных на ней.

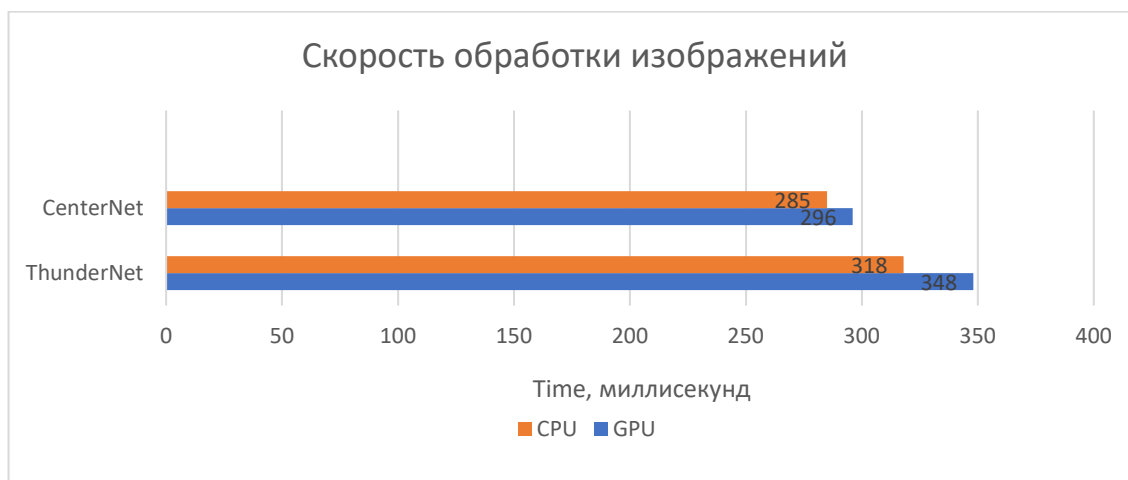


Рисунок 3. График затраченного времени на обучение для каждой нейросетевой модели. По горизонтали размечена временная шкала в секундах, по вертикали отмечены сравниваемые нейросетевые модели.

Модель ThunderNet показала наилучшие результаты по скорости обучения, однако в графике точности выходных данных на рисунке 4 она показала абсолютно противоположные результаты. Полученные итоги тестирования нейросетевых моделей при использовании фиксированных данных будут отличаться на разных процессорах. Это может зависеть от того, что обучение нейросетевых моделей для обнаружения объектов на изображении с использованием графического процессора более перспективно в плане производительности, чем с использованием центрального. Наиболее высокая точность выходных данных CenterNet обуславливается тем, что данная нейросеть работает по принципу тепловой карты и определению центральной точки объекта. В противовес разбиению на блоки нейросетью ThunderNet, этот способ обнаружения объекта на изображении показал себя наиболее точное

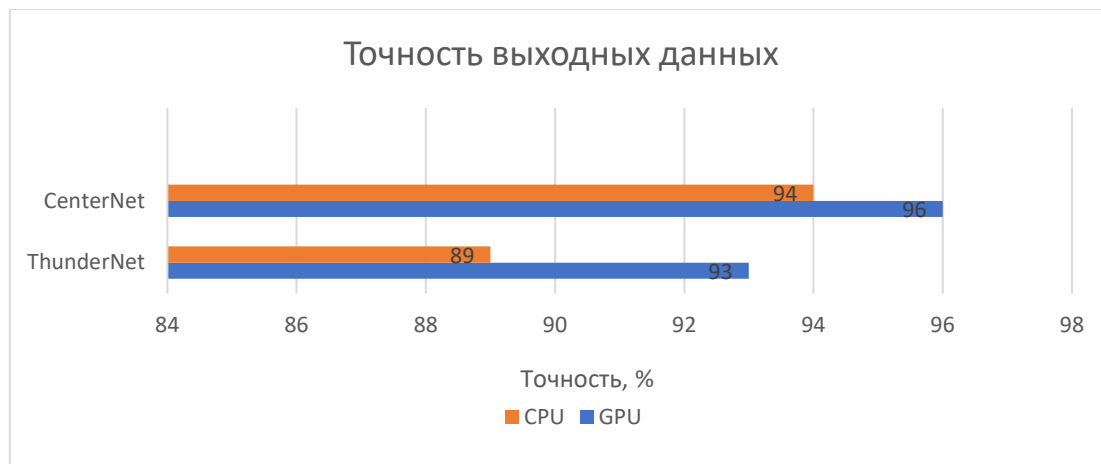


Рисунок 4. График точности выходных данных нейросетевых моделей. По горизонтальной шкале размечена градация точности в процентном соотношении. По вертикальной шкале отмечены сравниваемые нейросетевые модели.

5. Выводы

По итогам исследования можно выделить лидера по точности выходных данных – CenterNet. По скорости обучения наилучшие результаты показала модель ThunderNet. Из-за того, что размеры данных могут не оказывать воздействие на точность результата, при использовании набором данных CIFAR-10 предлагается использовать размеры 32×32 , как размеры входных данных, позволяющие находить объекты на изображении. Это решение экономичнее и не влияет на точность выходных данных. Проведённые в ходе исследования эксперименты показали, что скорость обучения ThunderNet сходна с CenterNet. Это доказывает, что для определения объектов на изображении более глубокие модели, имеющие более длительное время обучения, не в каждом случае предпочтительнее малых. Следующим этапом исследования будет являться использование моделей обработки данных на разных устройствах в условиях реального времени. Анализ полученных данных и использование более обширных баз данных для повышения эффективности обучения поможет улучшить работу моделей нейросетей, но и вызовет новые проблемы и недочёты, которые необходимо будет решать дополнительно по ходу внедрения. Сейчас в большом количестве существуют разнообразные алгоритмы обработки данных, постепенное внедрение которых в будущем поможет анализировать большие объёмы данных распознаванию объектов в реальном времени.

Список литературы

1. Duan Kaiwen. CenterNet: Keypoint Triplets for object Detection. – 2019. [Электронный ресурс] – URL: <https://arxiv.org/abs/1904.08189>

2. CenterNet: object detection, 3D detection, and pose estimation using center point detection. – 2022. [Электронный ресурс] – URL: <https://github.com/xingyizhou/CenterNet>
3. Qin Zheng. ThunderNet: Towards Real-time Generic Object Detection. – 2019. [Электронный ресурс] – URL: <https://arxiv.org/abs/1903.11752>
4. Jian Sun. ThunderNet: Towards Real-time Generic Object Detection. – 2019. // [Электронный ресурс] – URL: <https://paperswithcode.com/method/thundernet>
5. Krizhevsky Alex. The CIFAR-10 dataset. – 2009. [Электронный ресурс] – URL: <https://www.cs.toronto.edu/~kriz/cifar.html>