

Рисунок 2 – Число мод для здорового голоса и голоса с патологией

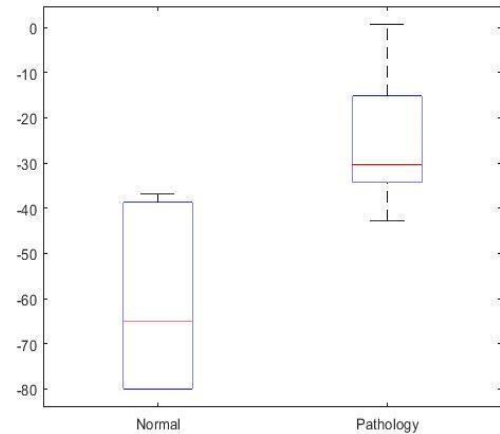


Рисунок 3 – Отношение сигнал/шум (SNR) для здорового голоса и голоса с патологией

По графикам видно, что значения этих двух показателей явно отличаются, что говорит о возможности использования этого метода в области патологий ЛОР-органов, в частности заболеваний гортани, и необходимости дальнейшего развития этого метода.

Список использованных источников:

1. A.-O. Boudraa, J.-Ch. Cexus, "EMD-Based Signal Filtering", IEEE Transactions on instrumentation and measurement, vol. 56, no. 6, pp. 2196-2202, 2007.
2. J. V. Sloten, P. Verdonck, M. Nyssen, J. Haueisen (Eds.): ECIFMBE 2008, IFMBE Proceedings 22, pp. 252–255, 2008.
3. И.А. Воронецкий, Анализ сигналов методом декомпозиции на эмпирические моды и его применение в обработке речевых сигналов / Доклады БГУИР. – 2012. – №7(69). – С. 18 – 24.
4. Juan Rafael Orozco-Arroyave, Elkyn Alexander Belalcazar-Bolaños, Julián David Arias-Londoño, Jesús Francisco Vargas-Bonilla, Sabine Skodda, Jan Ruzs, Khaled Daqrouq, Florian Hönig, and Elmar N'oth, Characterization Methods for the Detection of Multiple Voice Disorders: Neurological, Functional, and Laryngeal Diseases, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 19, NO. 6, 2015
5. Применение комплементарной множественной декомпозиции на эмпирические моды для анализа речевых сигналов / А. К. Алимуратов, Ю. С. Квитка // Измерение. Мониторинг. Управление. Контроль. – 2014. – № 4 (10). – С. 69–75.

ГЛУБОКАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ ДЕТЕКТИРОВАНИЯ ЛИЦ В УСЛОВИЯХ ОГРАНИЧЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ

Суша А.В.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Вашкевич М.И. – к.т.н., доцент

В работе приводятся приёмы проектирования глубоких нейронных сетей, показывающих высокую производительность на встраиваемых и мобильных платформах, на примере решения задачи детектирования лиц. Детектор лиц основан на концепции You Only Look Once, которая заключается в том, что изображение анализируется только один раз при прямом проходе нейронной сети. Приведенные методики проектирования позволили существенно повысить скорость детектирования без серьёзной потери в точности и полноте.

Применение глубоких нейронных сетей для детектирования лиц на изображении позволяет достичь высокого уровня точности и инвариантности к множеству факторов, затрудняющих детектирование: разный масштаб лиц, частичное перекрытие лиц другими объектами, поза людей в момент захвата изображения, выражения лиц, засветка и др. – всё это существенно понижает точность детектирования классическими методами [1]. Несмотря на то, что в настоящее время детекторы на глубоких нейронных сетях показывают высокую скорость детектирования на персональных компьютерах, оснащёнными мощными ЦПУ или ГПУ, всё ещё существуют определенные сложности в достижении высокой производительности в масштабе реального времени

на мобильных и встраиваемых устройствах. Это в первую очередь связано с тем, что в них устанавливается меньшее число вычислителей, при этом обладающими меньшей мощностью.

Последняя работа от авторов YOLO (You Only Look Once) [2] привнесла небольшие улучшения этого подхода к детектированию объектов на изображении. Авторы реализовали многомасштабное детектирование, а также предложили несколько новых архитектур глубоких нейронных сетей, используемых в детекторе. Эти архитектуры (Darknet-53 и Darknet-19) требуют меньших вычислительных затрат, чем нейронные сети, используемые ранее, что позволило авторам достигнуть высокой скорости детектирования с лучшим результатом 22 мс (или 45 FPS) на изображении, размером 320x320 пикселей, и 29 мс (или 34 FPS) на изображении 416x416 пикселей. Измерения проводились на ПК, оснащённом видеокартой Nvidia GeForce GTX Titan X [3].

За последние несколько десятилетий развития глубоких нейронных сетей были предложены множества приёмов организации архитектуры сетей. Анализ научных статей позволил выделить следующие методики повышения производительности глубоких нейронных сетей посредством уменьшения числа операций:

- первая операция свёртки должен иметь шаг окна равный 2, что в 4 раза уменьшает число операций, проводимых в этом слое;

- операции свёртки должны быть факторизованы на несколько меньших последовательных операций. К примеру, свёртка с окном 5x5 выполняется двумя свёртками 3x3 или 5x1 и 1x5, что уменьшает число параметров сети, а также число операций MAC;

- традиционная операция свёртки может быть заменены на последовательные операции поканальной и точечной свёртки, что уменьшает число параметров с $n = c \times h \times w \times k$ на $m = c \times h \times w + c \times k$, где c – число входных каналов, k – число выходных каналов, $h \times w$ – размер фильтра свёртки;

- применение концепции bottleneck (бутылочное горлышко), которая заключается в резком понижении числа каналов операцией точечной свёртки, а затем применением операции обычной свёртки.

Основываясь на приёмах, описанных выше, была спроектирована глубокая нейронная сеть, реализующая концепцию YOLO. Обучение проводилась на базе WIDER [4]. Для тестирования детектора использовался набор Fddb [5]. Время прямого прохода на видеокарте Nvidia GeForce GTX 1070 [6] (сопоставима с Titan X) составило 8-9 мс (110-125 FPS).

Список использованных источников:

1. Dwiyanoro, A. P. J. Performance Showdown of Publicly Available Face Detection Model [Electronic resource] / A. P. J. Dwiyanoro // Medium. – Mode of access: <https://medium.com/nodeflux/performance-showdown-of-publicly-available-face-detection-model-7c725747094a>. – Date of access: 24.03.2019.
2. Redmon, J. YOLOv3: An Incremental Improvement / J. Redmon, A. Farhadi // arXiv preprint, arXiv: 1804.02767. – 2018. – 6 p.
3. GeForce GTX TITAN X [Electronic resource] : Specifications / NVIDIA Corporation. – Mode of access: <https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-titan-x/specifications>. – Date of access: 24.03.2019.
4. Yang, S. WIDER FACE: A Face Detection Benchmark / S. Yang, P. Luo, C. C. Loy, X. Tang // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 5525- 5533.
5. Jain, V. Fddb: A Benchmark for Face Detection in Unconstrained Settings / V. Jain, E. Learned-Miller // Technical Report UM-CS-2010-009, Dept. of Computer Science, University of Massachusetts. – 2010. – 11 p.
6. GeForce GTX 1070 [Electronic resource] : Specifications / NVIDIA Corporation. – Mode of access: <https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-1070/specifications>. – Date of access: 24.03.2019.