

Module

# K1879BM8Я

## NM6408

Многоядерный нейропроцессор на базе оригинальной архитектуры NeuroMatrixCore 4 с управляющим RISC-процессором ARM Cortex A5.

Решает задачи цифровой обработки больших массивов данных в режиме реального времени. Работает с 32- и 64-разрядными данными с плавающей точкой одинарной и двойной точности.



Включен в перечни по Постановлениям Правительства РФ № 719 и № 818



### Решает задачи



Реализации обученных нейронных сетей



Цифровой обработки сигналов: изображений, аудио и видео информации

### Технические характеристики

Процессор и топология:	Интерфейсы и память:	ПО и NN:
<p><b>DSP:</b> 16 ядер NMC4 – 1 ГГц</p> <p><b>RISC:</b> 5 ядер Arm Cortex A5 – 800 МГц L2 Cache – 512 Кб</p> <p><b>Производительность:</b> FP32 – 512 GFLOP/s FP64 – 128 GFLOP/s</p> <p><b>Техпроцесс:</b> 28 нм КМОП</p> <p><b>Мощность (типичная/максимальная):</b> 12 Вт / 35 Вт</p> <p><b>Температурный диапазон:</b> -60C .... +85C</p> <p><b>Корпус:</b> 1444 HFCBGA</p> <p><b>Реализация:</b> "ОТК", "5"</p>	<p>5 x DDR3 (до 5 Гб, 32 Гб/с) 4 x PCIe 2.0 1 x Ethernet 10/100 Мб/с (EDCL) 1 x GPIO (8 выводов) 1 x JTAG 1 x SPI</p> <p>Суммарная пропускная способность интерфейсов межпроцессорного обмена: 160 Гбит/с</p> <p><b>Внутренняя память:</b> 76 Мбит</p>	<p><b>NMC SDK:</b> IDE (open VS-code), компиляторы, отладчики ARM и NMC Bare Metal SDK БЗИО, БУПВ (MPI, GAS), OpenCL Набор специализированных библиотек математических функций (BLAS, ЦОС)</p> <p><b>Драйверы:</b> Windows (7,8,10), Linux (включая "AstraLinux" и "Эльбрус ОС")</p> <p><b>Neuro Matrix Deep Learning:</b> Фирменный компилятор ИНС DarkNet и ONNX. Библиотека для применения ИНС. Библиотека трансляции моделей. Библиотека для подготовки изображений. Набор утилит для применения.</p>

# Benchmark NN inference

Ниже приведены значения производительности модуля при реализации (Inference) некоторых современных глубоких нейронных сетей.

**FPS** - число обработанных кадров изображения в секунду.

**Latency** - время обработки одного кадра.

**Batch-mode** - режим одновременной обработки нескольких кадров.

Нейронная сеть (размер изображения в пикселях)	FPS NM Card Mini, NM Mezzo Mini (multi unit mode)	Latency NM Card Mini, NM Mezzo Mini (multi unit mode)	FPS NM Card Mini, NM Mezzo Mini (batch mode)	Latency NM Card Mini, NM Mezzo Mini (batch mode)	FPS NM Quad (multi unit mode)	Latency NM Quad (multi unit mode)	FPS NM Quad (batch mode)	Latency NM Quad (batch mode)
alexnet (227x227)	12,6	79	13	308	50,4	79	52	308
inception v3 (299x299)	12,8	78	20,3	197	51,2	78	81,2	197
inception v3 (512x512)	3,93	254	5,44	735	15,72	254	21,76	735
resnet 18 (224x224)	25	40	47	85	100	40	188	85
resnet 50 (224x224)	12,2	82	20,6	194	48,8	82	82,4	194
squeezenet (224x224)	74,4	13	100	40	297,6	13	400	40
u-net (512x512)*	2	500	2	2000	8	500	8	2000
yolo v2 tiny (416x416)	21	47	30,4	132	84	47	121,6	132
yolo v3 (416x416)	3,7	270	4,5	889	14,8	270	18	889
yolo v3 tiny (416x416)	27,3	36	35,3	113	109,2	36	141,2	113
yolo v5s (640x640)	3,3	303	3,4	1176	13,2	303	13,6	1176

## Модули на базе процессора



**MC 127.05**  
Инструментальный модуль с PCI-e интерфейсом. Функциональный аналог NM Card.




**NM Card Mini**  
Спецвычислитель в форм-факторе PCIe x4.




**NM Mezzo Mini**  
Бортовой спецвычислитель. Подключается к несущей плате по PCIe шине (PCI04).




**NM Quad**  
Серверный вычислитель на базе четырех NM6408 в форм-факторе PCIe x4 (2 слота)



Полное описание модулей доступно по ссылкам



Руководство по эксплуатации  
доступно по ссылке